SIAM J. MATRIX ANAL. APPL.
Vol. 46, No. 4, pp. 2614–2634
© 2025 Society for Industrial and Applied Mathematics

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

MATRIX PERTURBATION ANALYSIS OF METHODS FOR EXTRACTING SINGULAR VALUES FROM APPROXIMATE SINGULAR SUBSPACES*

LORENZO LAZZARINO†, HUSSAM AL DAAS‡, AND YUJI NAKATSUKASA†

Abstract. Given (orthonormal) approximations $\tilde{U}$ and $\tilde{V}$ to the left and right subspaces spanned by the leading singular vectors of a matrix A, we discuss methods to approximate the leading singular values of A and study their accuracy. In particular, we focus our analysis on the generalized Nyström approximation, as surprisingly, it is able to obtain significantly better accuracy than classical methods, namely Rayleigh–Ritz and (one-sided) projected SVD. A key idea of the analysis is to view the methods as finding the exact singular values of a perturbation of A. In this context, we derive a matrix perturbation result that exploits the structure of such 2×2 block matrix perturbation. This leads to sharp bounds that predict well the approximation error trends and explain the difference in the behavior of these methods. Finally, we present an approach to derive an a posteriori version of those bounds, which are more amenable to computation in practice.

Key words. singular values, generalized Nyström approximation, Rayleigh–Ritz approximation, randomized SVD, matrix perturbation

MSC codes. 15A18, 15A42, 65F15, 68W20

DOI. 10.1137/24M169343X

1. Introduction.

Approximating the singular values of a matrix A is important for several reasons across various fields of mathematics, engineering, and data science. In principal component analysis, the singular values of the data matrix provide information on the variance captured by each principal component. In linear inverse problems, access to singular values is important to build effective regularization techniques [14, 15, 27]. Additionally, estimating leading singular values is essential for determining where to truncate to achieve a meaningful low-rank approximation. This is beneficial in any context where low-rank approximations are applied.

Subspace iteration, Golub–Kahan bidiagonalization, and randomized techniques are widely used methods to approximate the left and right dominant singular subspaces. Once an approximation to the left/right singular subspaces $\tilde{U}$ and/or $\tilde{V}$ is available, a number of strategies exist to extract the singular values. The Rayleigh–Ritz technique [6, 25, 35] is the standard method used to extract the singular values once the singular subspaces are approximated. A significant motivation behind this work is the efficient randomized methods for low-rank approximation, such as the Halko–Martinsson–Tropp (HMT) [13] and the generalized Nyström [3, 22, 34], which are based on first finding approximations to the left and/or right dominant singular subspaces of A.

*Received by the editors September 13, 2024; accepted for publication (in revised form) August 19, 2025; published electronically December 5, 2025.
https://doi.org/10.1137/24M169343X
Funding: This work is supported by EPSRC grants EP/Y008200/1, EP/Y010086/1, and EP/Y030990/1, and by the Ada Lovelace Centre Programme at the Scientific Computing Department, STFC.
†Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK (lazzarino@maths.ox.ac.uk, nakatsukasa@maths.ox.ac.uk).
‡Scientific Computing Department, STFC, Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire, OX11 0QX, UK (hussam.al-daas@stfc.ac.uk).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Motivated by streaming data analysis, Rayleigh–Ritz, HMT, generalized Nyström, and other methods to extract an approximation of the singular values can be categorized into two families, one-pass and multipass, based on the number of accesses to the matrix of interest that are required to obtain the approximation. It can be observed numerically that these methods have different approximation behavior. The goal of this paper is to understand and characterize the reasons for such difference. This is accomplished by interpreting the methods as perturbations of the original matrix and studying their accuracy through a certain matrix perturbation result, to obtain error bounds for individual singular values.

By design, generalized Nyström and HMT are (randomized) algorithms for finding a low-rank approximation $A \approx \tilde{A}$. The approximation quality of genralized Nyström has been extensively analyzed in [32, 22], showing roughly that the error is optimal up to a modest factor and a slight oversampling. Analogous results are available for HMT, which comes with a slightly better error [13, sect. 10]. These results hold with very high quantifiable probability. It is possible to use such matrix approximation bounds to derive error bounds for the singular values. For example, a low-rank approximation error bound in the 2-norm $\|A - \tilde{A}\|_2$ would give an error bound for all singular values. However, as we will see, the errors are often not uniform, with the leading singular values typically having significantly higher accuracy. Such a property cannot be captured by a global, low-rank approximation error. For HMT (and more generally subspace iteration), convergence analysis of singular values is provided in [12, 26].

We review in section 2 several methods to approximate the leading singular values of a matrix given the approximate singular subspaces. Motivated by numerical observation showing its better accuracy compared to other single-pass strategies, we analyze the generalized Nyström approximation and generalize the analysis to the Rayleigh–Ritz and (randomized) SVD approximations. Specifically, in section 3, we present a formal interpretation of the generalized Nyström approximation as a structured matrix perturbation and detail its underlying structure. Classic results from perturbation analysis of singular values such as Weyl's inequality provide pessimistic bounds that cannot provide insights into the expected accuracy of the approximated singular values by each of these methods. Therefore, in section 4 we develop results that exploit the structure of the perturbation. In particular, we study perturbations of 2×2 block matrices. This result is a generalization of the result for eigenvalues based on studying perturbations of symmetric matrices in [21]. Then, in section 5 we apply this formally derived and general result to the problem above, obtaining bounds for each of the mentioned strategies. Finally, in section 6, we illustrate how these bounds are often sharper than Weyl's inequality and how they can be used to compare methods, and we present an a posteriori alternative to them.

From a computational perspective, our results suggest the following strategy for approximating singular values. When only one subspace $\tilde{U}$ or $\tilde{V}$ is available, the SVD approach is recommended. When $\tilde{U}, \tilde{V}$ are both available, generalized Nyström tends to give significantly higher accuracy than other single-pass methods, as our bounds indicate and Figure 1 illustrates; we regard this as a key contribution of this paper. When multiple passes with A are allowed, we recommend HMT, or better yet, combined with several iterations of subspace iteration as much as the computational budget permits.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

2. Background and motivation. Let $A \in \mathbb{R}^{m\times n}$ and its economic SVD $A = U\Sigma V^*$ with singular values $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. The first singular values and their corresponding singular vectors are often referred to as *leading,* and they play a crucial role in applications involving, e.g., low-rank approximation. Throughout the paper, we say that a matrix M is orthonormal if it is a tall matrix with orthonormal columns, i.e., $M^*M = I$. For simplicity, we assume the matrices are real, although everything carries over verbatim to the complex case. Thus, the superscript $*$ denoting the conjugate transpose can be interpreted as the matrix transpose.

Suppose we have access to orthonormal approximations $\tilde{U}$ and $\tilde{V}$ to the leading singular subspaces of the matrix A. Sometimes we have approximations to different dimensions of left and right singular subspaces, that is, $\tilde{V} \in \mathbb{R}^{n\times r}$ and $\tilde{U} \in \mathbb{R}^{m\times(r+\ell)}$, orthonormal approximations to the leading r-dimensional left and $(r+\ell)$-dimensional right singular subspaces, where $\ell \geq 0$ is the so-called oversampling parameter. In other cases, we may have just one of them. In either case, a classical numerical linear algebra problem is to approximately find the leading r singular values of A. There are multiple strategies to accomplish such a task. The most classical one is the Rayleigh–Ritz method, a strategy usually presented for the extraction of eigenvalues [25, sect. 4–5], but that can be easily generalized to singular values [6, 35]. It extracts the leading r singular values by considering the singular values of the matrix $A_{RR,\tilde{V},\tilde{U}} := \tilde{U}^*A\tilde{V}$, i.e., for $i = 1,\dots,r$,

$$\sigma_i \approx \sigma_i(\tilde{U}^*A\tilde{V}) =: \sigma_i^{RR}.$$

If we have access to only one of the two approximations, say $\tilde{V}$, then we can extract the leading singular values by looking at the matrix $A_{SVD,\tilde{V}} := A\tilde{V}$, i.e., for $i = 1,\dots,r$,

$$\sigma_i \approx \sigma_i(A\tilde{V}) =: \sigma_i^{SVD}.$$

We refer to this as the *SVD approximation,* as it is based on a (one-sided) projected SVD. Less intuitive methods can also be employed. The general idea is to use the available approximations in the expression of low-rank approximations of A. For example, we can consider the generalized Nyström low-rank approximation [3, 22, 34]. That is, for some $X \in \mathbb{R}^{n\times r}$ and $Y \in \mathbb{R}^{m\times(r+\ell)}$, we construct an approximation of A with column space AX and row space Y^*A. In particular,

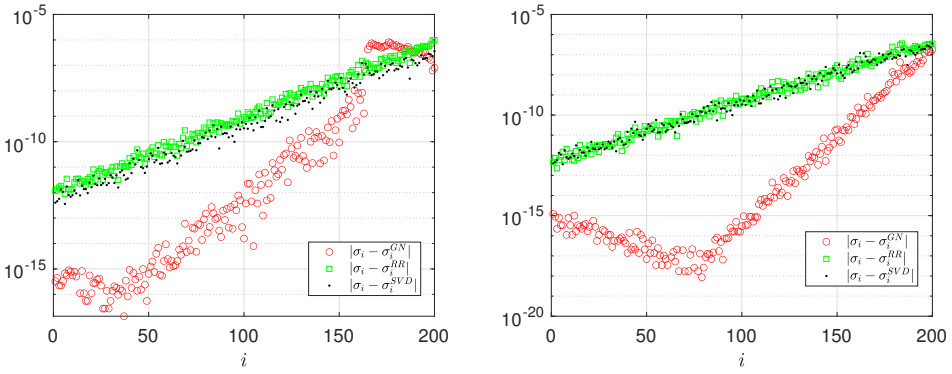
$$A_{GN,X,Y} := AX(Y^*AX)^\dagger Y^*A, \tag{2.1}$$

where the superscript † denotes the Moore–Penrose pseudoinverse as defined in [11, sect. 5.5.2]. This expression has been introduced in [33] and then used and analyzed in the randomized low-rank approximation context in [3, 22, 34]. In its classic definition, X and Y are taken to be random matrices, e.g., Gaussian matrices or more structured matrices such as SRTT [2, 30]. Thus, we can interpret this approximation as a two-sided projection process. The idea is to use the generalized Nyström expression to obtain an approximation to singular values. That is, we use $\tilde{U}$ and $\tilde{V}$ in (2.1) (instead of X and Y) and then compute the singular values, i.e., for $i = 1,\dots,r$,

$$\sigma_i \approx \sigma_i(A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger\tilde{U}^*A) =: \sigma_i^{GN}.$$

We will refer to this as the *generalized Nyström approximation*; see Algorithm 6.1 for an implementation. It is worth noting that, unlike the other single-pass methods, $\tilde{U}$ and $\tilde{V}$ need not be orthonormal in generalized Nyström. We will, however, assume they are, as this simplifies the analysis without losing generality.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Without Oversampling ($\ell = 0$). (b) With Oversampling ($\ell > 0$).

FIG. 1. ***Accuracy of single-pass extraction methods.*** *We compare the approximation quality of the studied single-pass methods: generalized Nyström (red circles), Rayleigh–Ritz (green squares), and SVD (black dots). The matrix in this example is constructed such that its singular values decay exponentially. The V-shaped errors with generalized Nyström are likely an artifact of roundoff errors; in exact arithmetic, we expect the error to be roughly proportional to* $1/\sigma_i$. *Further details on the experiment are given in section* 6.

A similar approach can be used for the case where just $\tilde{V}$ is available. We consider the low-rank approximation computed by the HMT method, also known as randomized SVD [4, 13, 24]. Given a random matrix $X \in \mathbb{R}^{n\times r}$, we compute the product AX and perform a QR factorization $AX = QR$. Then, we define the low-rank approximation by

$$A_{HMT,X} = Q(Q^*A). \tag{2.2}$$

The idea is to use the HMT process to extract singular values. Thus, we start the HMT process with $\tilde{V}$ (instead of X) and then compute the singular values of Q^*A. We have, for $i = 1, \ldots, r$,

$$\sigma_i \approx \sigma_i(Q^*A) =: \sigma_i^{HMT}.$$

Note that due to the orthogonality of Q, there is no need to perform the left multiplication by it in (2.2) for the purpose of extracting singular values. We will refer to this as the *HMT approximation;* see Algorithm 6.4 for an implementation.

Although both the HMT and the generalized Nyström approximations need two multiplications by A, only the generalized Nyström approximation is single-pass, that is, can be computed by accessing A once (as are the SVD approximation and the Rayleigh–Ritz method). Moreover, all these methods return exact singular values if $\tilde{U}$ and $\tilde{V}$ are exact leading singular subspaces.

Figure 1 shows how the generalized Nyström approximation typically tends to give the best approximation among single-pass methods, providing significantly better approximations of the singular values, especially for the leading ones. This serves as a motivation to investigate further how to characterize the accuracies of the various strategies to understand why the generalized Nyström approximation performance is often significantly better. For this reason, we focus our analysis on the generalized Nyström approximation, and then, in subsection 5.2, generalize it to the other methods.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

All these strategies can be interpreted as perturbations of the original matrix A [31] that, after particular orthogonal transformations, have peculiar structures that we will exploit. A standard result in this field is *Weyl's inequality* [16, Cor. 7.3.5], [28, Cor. I.4.31]: for any matrix M and for all i,

$$|\sigma_i(M) - \sigma_i(M+E)| \leq \|E\|_2. \tag{2.3}$$

An analogous result holds for eigenvalues of symmetric matrices [16, Cor. 6.3.4], [23, Thm. 10.3.1], [29, Cor. III.4.10]. Both results are usually referred to as Weyl's inequality, bound, or theorem. As is clear from (2.3), this bound neither captures nor exploits the structure of the perturbation E. Moreover, it does not depend on i so it does not predict the difference between the change of the leading and smallest eigen/singular values. We note that for some structured cases, better bounds than (2.3) are available, i.e., those that give relative perturbation bounds [7, 18]. However, typically relative perturbation bounds are helpful more for the smallest singular values, and not for the leading ones. Notable exceptions are the bounds for the extraction of (leading) singular values by the HMT method in [12, 26]. For this reason, in subsection 6.4, we compare our results for HMT with the one in [26].

3. Generalized Nyström and matrix perturbation. The main results of this paper are based on a matrix perturbation interpretation of the aforementioned methods. The following illustrates how the generalized Nyström approximation can be interpreted as a perturbation of the original matrix A [31].

For any orthonormal $\tilde{U}$ and $\tilde{V}$, considering their orthogonal complements $\tilde{V}_\perp$ and $\tilde{U}_\perp$, we can define the orthogonal matrices

$$Q_1 = \begin{bmatrix}\tilde{U} & \tilde{U}_\perp\end{bmatrix} \in \mathbb{R}^{m\times m}, \quad Q_2 = \begin{bmatrix}\tilde{V} & \tilde{V}_\perp\end{bmatrix} \in \mathbb{R}^{n\times n},$$

for which

$$\begin{aligned}(Q_1^*AQ_2)_{GN,\left[\begin{smallmatrix}I_r\\0\end{smallmatrix}\right],\left[\begin{smallmatrix}I_{r+\ell}\\0\end{smallmatrix}\right]} &= Q_1^*AQ_2\begin{bmatrix}I_r\\0\end{bmatrix}\left(\begin{bmatrix}I_{r+\ell} & 0\end{bmatrix}Q_1^*AQ_2\begin{bmatrix}I_r\\0\end{bmatrix}\right)^\dagger\begin{bmatrix}I_{r+\ell} & 0\end{bmatrix}Q_1^*AQ_2\\ &= Q_1^*\left(A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger\tilde{U}^*A\right)Q_2 = Q_1^*\left(A_{GN,\tilde{V},\tilde{U}}\right)Q_2.\end{aligned}$$

Due to the orthogonality of the transformations Q_1 and Q_2, this trivially implies that

$$\begin{aligned}|\sigma_i(A) - \sigma_i(A_{GN,\tilde{V},\tilde{U}})| &= |\sigma_i(Q_1^*AQ_2) - \sigma_i(Q_1^*A_{GN,\tilde{V},\tilde{U}}Q_2)|\\ &= \left|\sigma_i(Q_1^*AQ_2) - \sigma_i\left((Q_1^*AQ_2)_{GN,\left[\begin{smallmatrix}I_r\\0\end{smallmatrix}\right],\left[\begin{smallmatrix}I_{r+\ell}\\0\end{smallmatrix}\right]}\right)\right|.\end{aligned} \tag{3.1}$$

This means that for any orthonormal $\tilde{U}$ and $\tilde{V}$ we can find a set of coordinates in which we can study the generalized Nyström approximation with sampling matrices

$$\begin{array}{c} \\ r \\ n-r \end{array}\!\!\begin{array}{c} \begin{array}{c} r \end{array} \\ \begin{bmatrix} I_r \\ 0 \end{bmatrix} \end{array}, \qquad \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{c} \begin{array}{c} r+\ell \end{array} \\ \begin{bmatrix} I_{r+\ell} \\ 0 \end{bmatrix} \end{array},$$

applied to the matrix $\bar{A} := Q_1^*AQ_2$. Dividing $\bar{A}$ into blocks of appropriate sizes, i.e.,

$$\bar{A} := \begin{array}{c} \\ r+\ell \\ m-(r+\ell) \end{array}\!\!\begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \end{array}, \tag{3.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

we can analyze the effect the generalized Nyström approximation has on each block of $\bar{A}$. We have

$$\bar{A}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} = \begin{bmatrix} \bar{A}_{11} \\ \bar{A}_{21} \end{bmatrix} (\bar{A}_{11})^{\dagger} \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \end{bmatrix}.$$

That is,

$$\bar{A}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} = \begin{bmatrix} \bar{A}_{11}\bar{A}_{11}^{\dagger}\bar{A}_{11} & \bar{A}_{11}\bar{A}_{11}^{\dagger}\bar{A}_{12} \\ \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{11} & \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{12} \end{bmatrix}.$$

A reasonable and generic assumption to make is that the submatrix $\bar{A}_{11}$ is full-rank, i.e., $\operatorname{rank}(\bar{A}_{11}) = r$. In this case, $\bar{A}_{11}$ has full column rank and, thus, the pseudoinverse acts as a left inverse, i.e., $\bar{A}_{11}^{\dagger}\bar{A}_{11} = I_r$ [29, Thm. III.1.2]. For general ℓ, the same cannot be said for $\bar{A}_{11}\bar{A}_{11}^{\dagger}$. This implies that, under the above assumption,

$$\bar{A}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{11}\bar{A}_{11}^{\dagger}\bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{12} \end{bmatrix}. \tag{3.3}$$

Equation (3.3) leads to a clear relation between the matrix A and $A_{GN,\tilde{V},\tilde{U}}$, i.e.,

$$Q_1^*\left(A - A_{GN,\tilde{V},\tilde{U}}\right)Q_2 = \begin{bmatrix} 0 & \bar{A}_{12} - \bar{A}_{11}\bar{A}_{11}^{\dagger}\bar{A}_{12} \\ 0 & \bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{12} \end{bmatrix} =: E_{GN}. \tag{3.4}$$

Therefore, the generalized Nyström approximation of a matrix A can be seen as a perturbation of A that, up to orthogonal transformations, has the particular structure as in (3.4), where the nonzero perturbations are, respectively, the difference between $\bar{A}_{12}$ and its orthogonal projection into the column space of $\bar{A}_{11}$ [29, Thm. III.1.3] ($(1,2)$ block in (3.4)), and the (generalized) Schur complement [36] of $\bar{A}$ ($(2,2)$ block in (3.4)). In the case where $\tilde{U}$ and $\tilde{V}$ are given with the same column size (i.e., without oversampling, $\ell = 0$), the pseudoinverse acts also as a right inverse, as, in this case, A_{11} has full row rank, and, therefore, the top-right block is also 0. At first sight, this seems an advantageous situation; however, it is important to notice that the considered blocks change size depending on the size of $\tilde{U}$ and $\tilde{V}$, resulting in a different, and potentially larger, bottom-right perturbation.

It is also worth noting that by transforming A into $\bar{A}$ we typically (assuming $\tilde{U}, \tilde{V}$ are high quality) construct a matrix with small off-diagonal blocks and leading singular values far from the spectrum of the $\bar{A}_{22}$ block. More precisely, the magnitude of the off-diagonal blocks will depend on the accuracy of the approximate subspaces. In section 6, we see that, for the generalized Nyström approximation, when A is low rank, little information about A can be sufficient to give good enough approximate singular subspaces. In the next section, we aim to derive a bound that exploits both the structure of $\bar{A}$ and the perturbation in (3.4).

4. Matrix perturbation results. In this section, we step away from the problem of extracting singular values and derive a general matrix perturbation result. We derive a bound on the change of singular values when a perturbation is applied that captures and exploits the structure of the perturbation and that predicts the difference in change between different singular values. Our analysis starts from a result for eigenvalues of symmetric block tridiagonal matrices presented in [21]. We

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

start by considering the result proven in [21] and, using a technique based on the Jordan–Wielandt theorem presented in [17], we derive the result for the general case.

Theorem 3.2 in [21] gives us a bound on the change in the eigenvalues when a symmetric perturbation is applied to a symmetric 2×2 block matrix. In more detail, considering the symmetric 2×2 block $n \times n$ matrix

$$A := \begin{bmatrix} A_{11} & A_{21}^* \\ A_{21} & A_{22} \end{bmatrix}$$

with eigenvalues $\{\lambda_i\}_{i=1}^n$ and the perturbed matrix

$$\hat{A} := A + \begin{bmatrix} E_{11} & E_{21}^* \\ E_{21} & E_{22} \end{bmatrix} =: A + E$$

with eigenvalues $\{\hat{\lambda}_i\}_{i=1}^n$, we can define, for $i = 1, \dots, n$,

$$\tau_i = \left(\frac{\|A_{21}\|_2 + \|E_{21}\|_2}{\min_k |\lambda_i - \lambda_k(A_{22})| - 2\|E\|_2} \right).$$

Then, for all i for which $\tau_i > 0$, we have

$$|\lambda_i - \hat{\lambda}_i| \le \|E_{11}\|_2 + 2\|E_{21}\|_2 \tau_i + \|E_{22}\|_2 \tau_i^2. \tag{4.1}$$

This bound is derived by considering the eigendecompositions of $A + tE$ for $t \in [0, 1]$ and observing that small components of the t-dependent eigenvector in positions corresponding to the dominant elements of E imply the eigenvalue has a small derivative with respect to t. Hence, bound (4.1) is found by bounding those components (a special case of the Davis–Kahan generalized $\sin\Theta$ theorem [5, Thm. 6.1] is used) and integrating. The condition $\tau_i > 0$ comes from the well-definiteness of the intermediate bound on the components of the t-dependent eigenvector and it means that the best situation for this bound is when the analyzed eigenvector is well-separated from the spectrum of A_{22} and the perturbation is small in norm. For (4.1) to be better than Weyl's inequality for eigenvalues, it is necessary that $\tau_i < 1$. If $\|E_{11}\| \ll \|E\|$ and λ_i is far from the spectrum of A_{22}, then $\tau_i \ll 1$.

In the following, we use a technique presented in [17] to show that an analogous result holds for the singular values of a nonsymmetric rectangular 2×2 block $m \times n$ matrix

$$H := \begin{bmatrix} G_1 & B \\ C & G_2 \end{bmatrix}.$$

This is based on the *Jordan–Wielandt theorem* [29, Thm. I.4.2], [16, Thm. 7.3.3]: Let $\{\sigma_i\}_{i=1}^n$ be the singular values of a matrix $M \in \mathbb{R}^{m \times n}$ with $m \ge n$. Then, the symmetric matrix

$$\begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}$$

has eigenvalues (in descending order)

$$\{\lambda_j\}_{j=1}^{m+n} = \left\{ \sigma_1, \dots, \sigma_n, \underbrace{0, \dots, 0}_{m-n \text{ times}}, -\sigma_n, \dots, -\sigma_1 \right\}. \tag{4.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

This can be used to study the eigenvalues of

$$H_{JW} := \left[\begin{array}{c|cc} 0 & H \\ \hline H^* & 0 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ 0 & 0 & C & G_2 \\ \hline G_1^* & C^* & 0 & 0 \\ B^* & G_2^* & 0 & 0 \end{array}\right]$$

to obtain results on the singular values of H. However, we still want to obtain results that are translatable back to the original blocks G_1, G_2, B, and C. Hence, we consider a permutation of H_{JW} with the same eigenvalues but different structure. In other words, we first exchange the second and third block rows (red) and then the second and third block columns (blue) of H_{JW}, to obtain the permuted matrix H_p:

$$\left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ 0 & 0 & C & G_2 \\ \hline G_1^* & C^* & 0 & 0 \\ B^* & G_2^* & 0 & 0 \end{array}\right] \overset{P_{\text{red}}}{\hookrightarrow} \left[\begin{array}{cc|cc} 0 & 0 & G_1 & B \\ G_1^* & C^* & 0 & 0 \\ \hline 0 & 0 & C & G_2 \\ B^* & G_2^* & 0 & 0 \end{array}\right] \overset{P_{\text{blue}}}{\hookrightarrow} \left[\begin{array}{cc|cc} 0 & G_1 & 0 & B \\ G_1^* & 0 & C^* & 0 \\ \hline 0 & C & 0 & G_2 \\ B^* & 0 & G_2^* & 0 \end{array}\right] =: H_p. \tag{4.3}$$

In this way, the diagonal blocks are written in terms of only the corresponding diagonal blocks of H and, at the same time, the off-diagonal blocks do not depend on them. Moreover, since the permutations P_{red} and P_{blue} are the inverse of each other ($P_{\text{red}} = P_{\text{blue}}^{-1}$), the resulting matrix H_p is similar (i.e., it has the same eigenvalues) to H_{JW} [16, Cor. 1.3.4]:

$$H_p = P_{\text{red}} H_{JW} P_{\text{red}}^{-1}.$$

By the same process that led us from H to H_p, we can go from a perturbed version

$$\hat{H} := H + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} =: H + F$$

to the symmetric perturbed matrix

$$\hat{H}_p = H_p + \left[\begin{array}{cc|cc} 0 & F_{11} & 0 & F_{12} \\ F_{11}^* & 0 & F_{21}^* & 0 \\ \hline 0 & F_{21} & 0 & F_{22} \\ F_{12}^* & 0 & F_{22}^* & 0 \end{array}\right] =: H_p + F_p. \tag{4.4}$$

We can now apply the result (4.1) on the matrices H_p with eigenvalues $\{\lambda_j\}_{j=1}^{m+n}$ and $\hat{H}_p$ with eigenvalues $\{\hat{\lambda}_j\}_{j=1}^{m+n}$. Thus, for $j = 1, \ldots, m+n$, we can define

$$\tau_j = \left(\frac{\left\| \begin{bmatrix} 0 & C \\ B^* & 0 \end{bmatrix} \right\|_2 + \left\| \begin{bmatrix} 0 & F_{21} \\ F_{12}^* & 0 \end{bmatrix} \right\|_2}{\min_k |\lambda_j - \lambda_k \left(\begin{bmatrix} 0 & G_2 \\ G_2^* & 0 \end{bmatrix} \right)| - 2 \|F_p\|_2} \right),$$

and conclude that, for all j for which $\tau_j > 0$, we have

$$|\lambda_j - \hat{\lambda}_j| \leq \left\| \begin{bmatrix} 0 & F_{11} \\ F_{11}^* & 0 \end{bmatrix} \right\|_2 + 2 \left\| \begin{bmatrix} 0 & F_{21} \\ F_{12}^* & 0 \end{bmatrix} \right\|_2 \tau_j + \left\| \begin{bmatrix} 0 & F_{22} \\ F_{22}^* & 0 \end{bmatrix} \right\|_2 \tau_j^2.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

The next step is to write this bound in terms of the original blocks to find a bound for the original problem. By the Jordan–Wielandt theorem and, in particular, by (4.2), the first n eigenvalues of H_p and $\hat{H}_p$ are exactly the singular values $\{\sigma_i\}_{i=1}^n$ and $\{\hat{\sigma}_i\}_{i=1}^n$ of H and $\hat{H}$, respectively, and, analogously, $\|F_p\|_2 = \|F\|_2$. Then, since

$$\left\| \begin{bmatrix} 0 & M_1 \\ M_2 & 0 \end{bmatrix} \right\|_2 = \max\{\|M_1\|_2, \|M_2\|_2\},$$

we obtain the desired generalization of [21, Thm. 3.2].

THEOREM 4.1. *Consider a* 2×2 *block matrix*

$$H := \begin{bmatrix} G_1 & B \\ C & G_2 \end{bmatrix} \tag{4.5}$$

with singular values $\{\sigma_i\}_{i=1}^n$ *and the perturbed matrix*

$$\hat{H} := H + \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} =: H + F$$

with singular values $\{\hat{\sigma}_i\}_{i=1}^n$. *Define, for* $i = 1, \ldots, n$,

$$\tau_i = \left(\frac{\max\{\|B\|_2, \|C\|_2\} + \max\{\|F_{12}\|_2, \|F_{21}\|_2\}}{\min_k |\sigma_i - \sigma_k(G_2)| - 2\|F\|_2} \right).$$

Then, for all i *for which* $\tau_i > 0$, *it holds that*

$$|\sigma_i - \hat{\sigma}_i| \le \|F_{11}\|_2 + 2\max\{\|F_{12}\|_2, \|F_{21}\|_2\}\tau_i + \|F_{22}\|_2 \tau_i^2. \tag{4.6}$$

By construction, this result inherits all the properties of the symmetric case studied in [21]. In particular, we need $\tau_i < 1$ for the bound to be better than Weyl's inequality (2.3), and if $\|F_{11}\| \ll \|F\|$ and σ_i is far from the spectrum of G_2, then $\tau_i \ll 1$.

COROLLARY 4.2 (perturbation "on the right"). *Consider a matrix as in* (4.5) *and assume that just the two right blocks are perturbed, i.e.,*

$$F = \begin{bmatrix} 0 & F_1 \\ 0 & F_2 \end{bmatrix}.$$

Define, for $i = 1, \ldots, n$,

$$\tau_i = \left(\frac{\max\{\|B\|_2, \|C\|_2\} + \|F_1\|_2}{\min_k |\sigma_i - \sigma_k(G_2)| - 2\|F\|_2} \right).$$

Then, for all i *for which* $\tau_i > 0$, (4.6) *trivially becomes*

$$|\sigma_i - \hat{\sigma}_i| \le 2\|F_1\|_2 \tau_i + \|F_2\|_2 \tau_i^2. \tag{4.7}$$

As shown in section 3, the generalized Nyström approximation can be seen, up to orthogonal transformations, as a perturbation "on the right" of the correspondent original matrix. While in the general matrix perturbation theory framework, $\|F\|_2$ (i.e., Weyl's theorem) is not guaranteed to be small, in the extraction of singular values framework, if good approximate subspaces are available, F will be in general

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

quite small. Indeed, with A low rank, this will follow from the fact that generalized Nyström approximation performs a low-rank approximation. However, it is the (nonzero) structure of F and the spectral distribution of the (transformed) matrix $\bar{A}$ that will determine the quality of the approximation. The denominator of τ_i will be big since the structure of the transformed matrix $\bar{A}$ typically implies that the leading singular values are well-separated from the spectrum of its $(2,2)$ block $\bar{A}_{22}$. Thus, in subsection 5.1, we use Corollary 4.2 to obtain a bound on the generalized Nyström approximation, and, in section 6, we numerically illustrate its behavior.

5. Application to methods for extracting singular values. The discussion of section 4 can be used to analyze methods for extracting singular values through a matrix perturbation approach. In the following, we show how this can be done for the generalized Nyström approximation and then generalize the derivation to other extracting methods.

5.1. Application to generalized Nyström. In section 3 we saw that, with the appropriate orthogonal transformations, the generalized Nyström approximation can be interpreted as a perturbation of the original matrix in the top and bottom-right blocks, i.e., as a perturbation "on the right." Thus, we can apply Corollary 4.2 to the transformed matrix $\bar{A}$ and its generalized Nyström approximation with sampling matrices (3.3), obtaining, by (3.1), a bound on the difference between the extracted singular values $\sigma_i^{GN} := \sigma_i(A_{GN,\tilde{V},\tilde{U}})$ and the exact singular values $\sigma_i := \sigma_i(A)$, for $i = 1, \ldots, n$.

COROLLARY 5.1. *Consider an $m \times n$ matrix A with singular values $\{\sigma_i\}_{i=1}^n$ and approximations $\tilde{U}$ and $\tilde{V}$ to its leading singular subspaces. Consider the generalized Nyström approximation $A_{GN,\tilde{V},\tilde{U}} = A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger\tilde{U}^*A$ and its singular values $\{\sigma_i^{GN}\}_{i=1}^n$. Define $\bar{A} = Q_1^*AQ_2$, with $Q_1 = \begin{bmatrix}\tilde{U} & \tilde{U}_\perp\end{bmatrix}$ and $Q_2 = \begin{bmatrix}\tilde{V} & \tilde{V}_\perp\end{bmatrix}$, and its blocks as in* (3.2). *Let, for $i = 1, \ldots, n$,*

$$\tau_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\} + \left\|\bar{A}_{12} - \bar{A}_{11}\bar{A}_{11}^\dagger\bar{A}_{12}\right\|_2}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{GN}\|_2}.$$

Then, for all i for which $\tau_i > 0$, it holds that

$$|\sigma_i - \sigma_i^{GN}| \le 2\left\|\bar{A}_{12} - \bar{A}_{11}\bar{A}_{11}^\dagger\bar{A}_{12}\right\|_2\tau_i + \left\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger\bar{A}_{12}\right\|_2\tau_i^2. \tag{5.1}$$

As already mentioned, in the case where $\tilde{U}$ and $\tilde{V}$ are given with the same column size (i.e., without oversampling, $\ell = 0$), we have $\|\bar{A}_{12} - \bar{A}_{11}\bar{A}_{11}^\dagger\bar{A}_{12}\|_2 = 0$. Thus, the bound can be slightly simplified. However, we note once again that, since the block sizes change compared to the oversampled case, this does not mean that $\ell = 0$ implies better approximations.

5.2. Application to other extraction methods. The generalized Nyström approximation is not the only method that can be analyzed through a matrix perturbation approach. Indeed, all methods mentioned in the introduction can be interpreted as a perturbation of the original matrix. Moreover, also in these other cases, we can achieve, by orthogonal transformations, a special structure of the perturbation that can be exploited, using Theorem 4.1, to derive a bound on the singular values extraction error. Therefore, as explored in subsection 6.3, our analysis can be seen as a tool for potential comparisons between the different strategies. For example, let us consider the Rayleigh–Ritz method

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\sigma_i := \sigma_i(A) \approx \sigma_i(\tilde{U}^* A \tilde{V}) =: \sigma_i^{RR}, \quad i = 1, \dots, r.$$

Then, considering the orthogonal transformations

$$Q_1 = \begin{bmatrix} \tilde{U} & \tilde{U}_\perp \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \tilde{V} & \tilde{V}_\perp \end{bmatrix},$$

we have

$$(Q_1^* A Q_2)_{RR, \left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} = \begin{bmatrix} I_{r+\ell} & 0 \end{bmatrix} Q_1^* A Q_2 \begin{bmatrix} I_r \\ 0 \end{bmatrix} = \tilde{U}^* A \tilde{V} = A_{RR, \tilde{V}, \tilde{U}}.$$

This means that instead of studying the original extraction error, we can study the error of the Rayleigh–Ritz method with sampling matrices applied to $\bar{A} = Q_1^* A Q_2$, i.e., for $i = 1, \dots, r$,

$$\begin{aligned} |\sigma_i - \sigma_i^{RR}| &= \left| \sigma_i(\bar{A}) - \sigma_i \left(\bar{A}_{RR, \left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} \right) \right| = |\sigma_i(\bar{A}) - \sigma_i(\bar{A}_{11})| \\ &= \left| \sigma_i(\bar{A}) - \sigma_i \left(\begin{bmatrix} \bar{A}_{11} & 0 \\ 0 & 0 \end{bmatrix} \right) \right| = \left| \sigma_i(\bar{A}) - \sigma_i \left(\bar{A} - \underbrace{\begin{bmatrix} 0 & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}}_{=: E_{RR}} \right) \right|. \end{aligned}$$

Therefore, it is possible to apply Theorem 4.1 and, as a consequence, obtain that, for all i for which $\tau_i^{RR} := \frac{2 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2)} > 0$,

$$\begin{aligned} |\sigma_i - \sigma_i^{RR}| \le\ & 4 \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2} \\ & + \|\bar{A}_{22}\|_2 \frac{4 \max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}^2}{\left(\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2 \right)^2}. \end{aligned} \tag{5.2}$$

Note that the Rayleigh–Ritz bound is largely influenced by the nonzero off-diagonal perturbations and that the significant difference in accuracy between generalized Nyström and Rayleigh–Ritz may arise from the fact that, unlike Rayleigh–Ritz, off-diagonal perturbations in generalized Nyström can be much smaller (or zero if oversampling is not used).

To obtain a bound for the SVD approximation we note that

$$\begin{aligned} |\sigma_i - \sigma_i^{SVD}| = |\sigma_i - \sigma_i(A\tilde{V})| &= \left| \sigma_i(AQ_2) - \sigma_i \left((AQ_2)_{SVD, \left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right]} \right) \right| \\ \overset{\tilde{A} := AQ_2}{=} & \left| \sigma_i(\tilde{A}) - \sigma_i(\tilde{A} - \underbrace{\begin{bmatrix} 0 & \tilde{A}_2 \end{bmatrix}}_{=: E_{SVD}}) \right|, \end{aligned}$$

where $\tilde{A}_2$ is the $m \times (n-r)$ right block of $\tilde{A}$. Then, using Theorem 4.1, we obtain, for all i for which $\tau_i^{SVD} := \frac{2\|\tilde{A}_2\|_2}{\sigma_i - 2\|E_{SVD}\|_2} > 0$,

$$|\sigma_i - \sigma_i^{SVD}| \le 4 \frac{\|\tilde{A}_2\|_2^2}{\sigma_i - 2\|E_{SVD}\|_2}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Algorithm 6.1 The generalized Nyström approximation (MATLAB notation): given $A \in \mathbb{R}^{m\times n}, \tilde{V} \in \mathbb{R}^{n\times r}$, and $\tilde{U} \in \mathbb{R}^{m\times(r+\ell)}$, approximate the first r singular values of A.

```
AtV = A*tV;
tUA = tU' * A;
[~, R1] = qr(AtV, 0);          ▷ QR factorization of external factor
[~, R2] = qr(tUA', 0);         ▷ QR factorization of external factor
[Q3, R3] = qr(tU' * AtV, 0);          ▷ QR factorization to compute
                                                          pseudoinverse
P = Q3' * R2';              ▷ Note:  The order of operations matters!
AGN = (R1/R3) * P;       ▷ Compute generalized Nyström approximation
svGN = svd(AGN); ▷ Use build-in function to compute singular values
```

Finally, to obtain a bound for the HMT approximation, we notice that, as shown in [22, Table 1], if Q is the orthogonal matrix from the QR factorization of $A\tilde{V}$, then

$$A_{HMT} := QQ^*A = A_{GN,\tilde{V},A\tilde{V}}. \tag{5.3}$$

Thus, the HMT approximation can be seen as a specific generalized Nyström approximation enabling us to use Corollary 5.1. Gu [12] and Saibaba [26] provide bounds for the extraction of singular values by the HMT method. In subsection 6.4, we show that, in contrast with bound (5.1), while these bounds are particularly sharp when $\tilde{V}$ is a random Gaussian matrix, they are less precise when $\tilde{V}$ has information of A.

6. Numerical illustrations and comparisons. Having derived results that respect the original goal of adaptability depending on each singular value and the use of the perturbation structure, we now numerically illustrate the sharpness of the derived bounds, compare different methods, and explore the possibility of making the derived bounds computable in practice.

We construct a 1000×1000 matrix A starting from its SVD $A = U\Sigma V^*$, with U and V Haar matrices [20]. We construct the singular subspace approximations from Gaussian matrices subjected to one power iteration. That is, taking Gaussian matrices Ω_1 and Ω_2, we define $\tilde{V} \in \mathbb{R}^{1000\times 200}$ and $\tilde{U} \in \mathbb{R}^{1000\times(200+\ell)}$ as the orthonormal matrices of the QR factorization of $A^*\Omega_1$ and $A\Omega_2$, respectively. There are various alternative strategies to generate approximate singular subspaces, including subspace iteration and Krylov methods. This means, e.g., generating $\tilde{V}$ and $\tilde{U}$ as the orthogonal basis computed by the Lanczos algorithm on A^*A and AA^* after r and $r+\ell$ iterations, respectivetly. We have also experimented with the latter, but ultimately considered the case described above, which is of more interest. Indeed, when generating subspaces by Krylov methods, A is applied repetitively, thus AA^*v is contained in the span of V for all but the last vector in $[v_0, AA^*v_0, \ldots, (AA^*)^r v]$. Thus, the second multiplication by A, typical of generalized Nyström, has less impact than in other situations, leading to rather small improvements on other single-pass methods.

We consider two singular value distributions. First, the singular values will follow an exponential decay:

$$\sigma_i = e^{-(i-1)\frac{30}{999}\ln 10}, \tag{6.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

where the largest singular value is 1, the smallest 10^{-30}, and $\sigma_r = \sigma_{200}$ is of the order of 10^{-6}. This can be implemented by the MATLAB built-in function `logspace(-30,0,n)'`.[1] Second, the singular values will follow an algebraic decay:

$$\sigma_i = \left(\frac{1}{i}\right)^4,$$

where the largest singular value is 1, the smallest 10^{-12}, and $\sigma_r = \sigma_{200} = 6.25 \cdot 10^{-10}$.

We compute the generalized Nyström approximation and the extracted singular values by Algorithm 6.1, in which we exploit the structure of the approximation instead of forming the $m \times n$ approximate matrix. Namely, first, we compute the QR (LQ) factorization of the external factors (lines 3 and 4)

$$\begin{aligned} &A\tilde{V} = Q_1 R_1, \\ &\tilde{U}A^* = Q_2 R_2 \implies \tilde{U}^* A = R_2^* Q_2^*. \end{aligned}$$

To compute the pseudoinverse in (2.1), we use the QR factorization of the matrix to invert $\tilde{U}^* A \tilde{V} = Q_3 R_3$ (line 5). Then, since singular values do not change under orthogonal transformations, we extract the positive ones by looking only at the small matrix formed by the R-factors and the pseudoinverse of the core matrix (lines 6 and 7)

$$\begin{aligned} \sigma^{GN} &= \sigma(Q_1 R_1 (Q_3 R_3)^\dagger R_2^* Q_2^*) \\ &= \sigma(Q_1 R_1 R_3^\dagger Q_3^* R_2^* Q_2^*) \\ &= \sigma(R_1 R_3^\dagger Q_3^* R_2^*). \end{aligned}$$

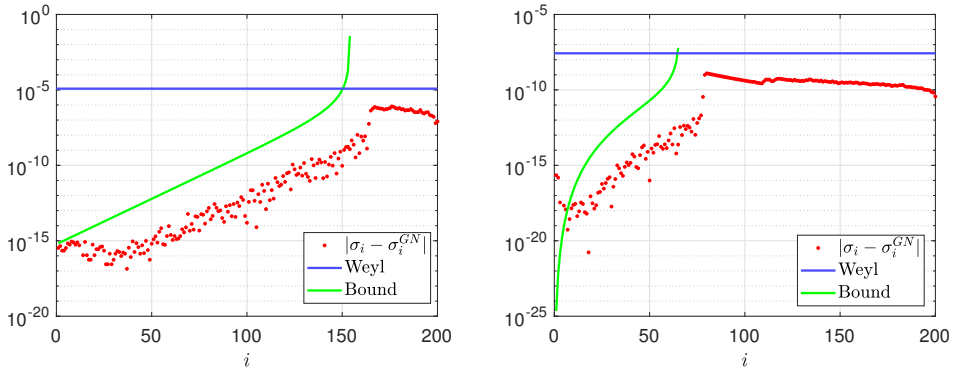
The singular values of the small matrix are then computed by the MATLAB built-in function `svd`. Alternatively, it is possible to use the Jacobi SVD from [8, 9], which can compute singular values with high relative accuracy for a certain class of matrices. In the whole procedure, we need to perform two matrix-matrix multiplications by A (lines 1 and 2) but, since they are not sequential (and hence parallelizable), we need to access A only once.

To understand the complexity of this procedure, we start by considering the cost N_r and $N_{r+\ell}$ of computing $A\tilde{V}$ or $\tilde{U}^* A$, respectively. This varies depending on the properties of A, e.g., sparsity, and those of the approximate singular subspaces. The QR factorization (and consequently computing the SVD) of an $m \times n$ ($m > n$) matrix costs $\mathcal{O}(mn^2)$. Therefore, from Algorithm 6.1, we see that the cost of the generalized Nyström procedure is dominated by the matrix multiplications with A of cost $N_{2r+\ell} := N_r + N_{r+\ell}$, and by the QR factorizations needed to extract singular values, particularly the one in line 3. Hence, we obtain an overall cost of $N_{2r+\ell} + \mathcal{O}((m+n)r^2)$.

6.1. Equal column size singular approximate subspace ($\ell = 0$). Let us consider the case without oversampling, i.e., $\ell = 0$, and, where $\tilde{U}$ and $\tilde{V}$ belong to $\mathbb{R}^{1000 \times 200}$. In Figure 2, we observe that, for both exponential and algebraic decay, the bound in Corollary 5.1 is significantly better than Weyl's inequality and that it can better predict the slope of the error. It is important to note that the apparent misbehavior of the bound for the very first singular value errors is simply due to the bounds falling below machine precision, as our analysis does not account for roundoff errors.

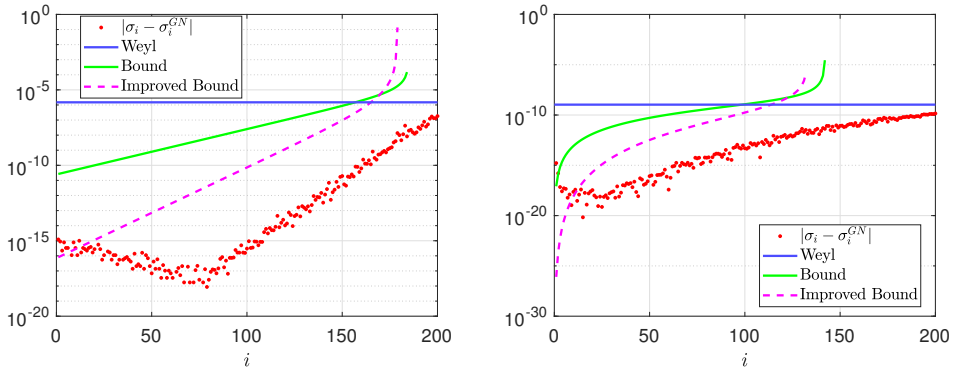
[1] All the implementations are available in the supplementary material (SupplementaryMaterials.zip [local/web 12.2KB]).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Exponential Decay. (b) Algebraic Decay.

FIG. 2. ***Singular value approximation errors of generalized Nyström without oversampling.*** *We show the error in extracting the first* $r = 200$ *singular values by generalized Nyström approximation with* $\ell = 0$ *(red dots), i.e.,* $|\sigma_i - \sigma_i^{GN}|$ *for* $i = 1, \ldots, 200$*, Weyl's inequality* (2.3) *(blue), and bound* (5.1) *(green). Both exponential* (a) *and algebraic* (b) *singular values decays are considered.*



(a) Exponential Decay. (b) Algebraic Decay.

FIG. 3. ***Singular value approximation errors of generalized Nyström with oversampling.*** *We show the error in extracting the first* $r = 200$ *singular values by generalized Nyström approximation with* $\ell = 1.5r$ *(red dots), i.e.,* $|\sigma_i - \sigma_i^{GN}|$ *for* $i = 1, \ldots, 200$*, Weyl's inequality* (2.3) *(blue), bound* (5.1) *(green), and the improved bound (magenta). Both exponential* (a) *and algebraic* (b) *singular values decays are considered.*

6.2. Different column size approximate singular subspace ($\ell = 0.5r$). Let us consider the case in which the approximations are given with different column sizes. In particular, we set $r + \ell = 1.5r$ following the recommendation in [22], so that, for $r = 200$, we take $\tilde{V} \in \mathbb{R}^{1000 \times 200}$ and $\tilde{U} \in \mathbb{R}^{1000 \times 300}$. Figure 3 shows that for both algebraic and exponential singular value decays, bound (5.1) is still significantly better than Weyl's. However, the sharpness observed in the previous experiment seems partially lost. Qualitatively speaking, this could be caused by the need to compute the pseudoinverse of a rectangular matrix ($\bar{A}_{11}^{\dagger}$) and by not achieving the typical structure discussed in section 3 when constructing $\bar{A}$. In view of this, we propose a heuristics-based technique that, after extensive numerical tests, seems to improve this bound and usually estimate the error reasonably accurately.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Heuristic-based improvement. The goal is to reduce ourselves to a situation similar to the nonoversampling framework. This means that we aim at modifying the sizes of the blocks used to compute bound (5.1), so as to have a square $(1,1)$ block. Moreover, we rely on the fact that oversampling generally improves generalized Nyström approximations. Naively modifying the sizes of $\bar{A}$ blocks (as in (3.2)) would probably ruin even more the desired structure described in section 3. Thus, we first apply orthogonal transformations to the (already transformed) matrix $\bar{A}$ to diagonalize its $(1,1)$ block, $\bar{A}_{11} = X\Sigma Y^*$, and obtain the matrix

$$\bar{\bar{A}} := \begin{bmatrix} X^* & 0 \\ 0 & I \end{bmatrix} \bar{A} \begin{bmatrix} Y & 0 \\ 0 & I \end{bmatrix}.$$

Since $\bar{\bar{A}}$ is an orthogonal transformation of $\bar{A}$, for the same reasons illustrated at the start of section 3, we have

$$\left| \sigma_i(\bar{A}) - \sigma_i\left(\bar{A}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]} \right) \right| = \left| \sigma_i(\bar{\bar{A}}) - \sigma_i\left(\bar{\bar{A}}_{GN,\left[\begin{smallmatrix} Y^* \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} X^* \\ 0 \end{smallmatrix}\right]} \right) \right|.$$

Moreover, due to the orthogonal invariance of the 2-norm, bound (5.1) reduces to the same expression when computed for $|\sigma_i(\bar{\bar{A}}) - \sigma_i(\bar{\bar{A}}_{GN,\left[\begin{smallmatrix} Y^* \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} X^* \\ 0 \end{smallmatrix}\right]})|$ and for $|\sigma_i(\bar{\bar{A}}) - \sigma_i(\bar{\bar{A}}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_{r+\ell} \\ 0 \end{smallmatrix}\right]})|$. Finally, note that the latter is, in general, smaller than the error obtained if no oversampling is employed, i.e., $|\sigma_i(\bar{\bar{A}}) - \sigma_i(\bar{\bar{A}}_{GN,\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right],\left[\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right]})|$. Thus, we can expect that bound (5.1) computed with blocks of $\bar{\bar{A}}$ whose sizes are chosen so that $\bar{\bar{A}}_{11}$ is $r \times r$ (in the specific example 200×200) is going to bound the original accuracy error.

In Figure 4, we illustrate the effects that this process has on the matrix with exponentially decaying singular values taken into account above. Figure 3 shows the behavior of bound (5.1) and the heuristically improved version for both exponential and algebraic singular value decays. Note that, as shown by both Figures 2 and 3, in general, there exists an $i < r$ for which the presented bounds became worse than Weyl's one. This is because, when approaching the rth singular value, the gap becomes smaller, making τ_i greater than 1. Thus, we suggest considering the bound given by the minimum, for each i, between (5.1) and Weyl's inequality.

6.3. Methods comparison. We implement and compare the various methods. We compute the Rayleigh–Ritz and the SVD approximations by Algorithms 6.2 and 6.3, respectively. In both cases, we first construct the approximate matrix A_{RR} or A_{SVD} and then use the MATLAB built-in function `svd` to compute the singular values. We stress that, in both cases, we need to access the matrix A only once, i.e., these are single-pass algorithms, and we need to perform only one matrix-matrix multiplication by A. The cost of the Rayleigh–Ritz procedure is given by $N_r + O(mr(r+\ell))$ operations plus the cost of computing the singular values of an $(r+\ell) \times r$ matrix that, assuming $\ell = \mathcal{O}(r)$, is $\mathcal{O}(r^3)$, while to obtain the SVD approximation we need to compute $A\tilde{V}$, of cost N_r, and compute singular values of an $m \times r$ matrix, of cost $\mathcal{O}(mr^2)$.

To implement the HMT method, Algorithm 6.4, we need to access and multiply by A twice; see lines 1 and 3. Moreover, the multiplication of computing Q^*A almost invariably involves multiplying an unstructured, dense matrix Q^* to A, in contrast to computing $A\tilde{V}$ or $\tilde{U}^*A$, which can be faster for structured (e.g. sparse) $\tilde{V}, \tilde{U}$. We thus distinguish the cost of computing $\tilde{U}^*A$ (which is N_r or $N_{r+\ell}$) from that of Q^*A, which we denote by $\tilde{N}_r$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

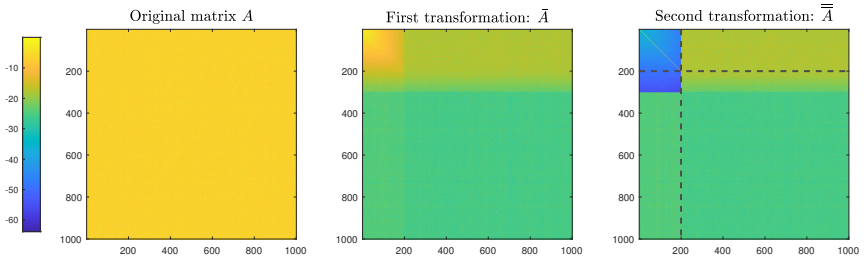


FIG. 4. ***Graphic representation of the process to obtain an heuristic-based bound improvement.*** *The figures show the magnitude of the entries in logaritmic scale of the matrices $A, \bar{A}$, and $\bar{\bar{A}}$, with A generated with exponentially decaying singular values. The dotted lines in the figure corresponding to $\bar{\bar{A}}$ indicate the choice of the blocks sizes used in computing the heuristic-based bound.*

Algorithm 6.2 The Rayleigh–Ritz approximation (MATLAB notation): given $A \in \mathbb{R}^{m\times n}, \tilde{V} \in \mathbb{R}^{n\times r}$, and $\tilde{U} \in \mathbb{R}^{m\times(r+\ell)}$, approximate the first r singular values of A.

```
AtV = A*tV;
ARR = tU' * AtV;                          ▷ Compute Rayleigh--Ritz approximation
svRR = svd(ARR); ▷ Use build-in function to compute singular values
```

Algorithm 6.3 The SVD approximation (MATLAB notation): given $A \in \mathbb{R}^{m\times n}$ and $\tilde{V} \in \mathbb{R}^{n\times r}$, approximate the first r singular values of A.

```
ASVD = A*tV;                                  ▷ Compute SVD approximation
svSVD = svd(ASVD);       ▷ Use build-in function to compute singular
values
```

Algorithm 6.4 The HMT method (MATLAB notation): given $A \in \mathbb{R}^{m\times n}$, and $\tilde{V} \in \mathbb{R}^{n\times r}$, approximate the first r singular values of A.

```
AtV = A*tV;
[Q,~] = qr(AtV, 0);
AHMT = Q' * A;                                ▷ Compute HMT approximation
svHMT = svd(AHMT);       ▷ Use build-in function to compute singular
values
```

In Table 1 we summarize the main features of the presented algorithms. In Figure 5, we compare the methods, their extraction errors, and the corresponding bounds (derived in subsection 5.2) applied to the case where A has exponentially decaying exact singular values as in (6.1). Analogous experiments with similar results for algebraically decaying exact singular values are in the supplementary material (Supplementary-Materials.zip [local/web 12.2KB]). Since the derived bounds seem to be good predictors of extraction errors, we can use them to explain the behavior of the methods. In particular, we consider the generalized Nyström approximation and the Rayleigh–Ritz method without oversampling. First of all, we note that the bounds of both methods are computed with the blocks of the same transformed matrix $\bar{A}$. Moreover, Figure 5 confirms that Weyl's inequality is not descriptive of the difference between these methods and that the perturbation norms are close to each other, i.e., $\|E_{GN}\| \approx \|E_{RR}\|$. Therefore, with $\ell = 0$, we have

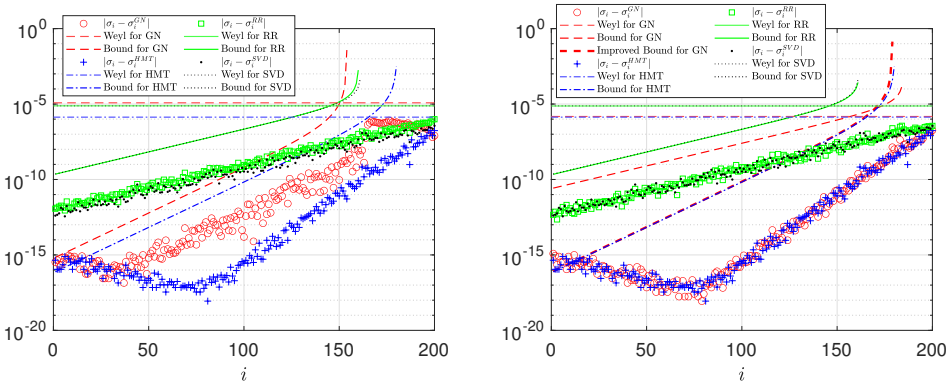
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

TABLE 1

Comparison of algorithms for the extraction of singular values. *N_r and $N_{r+\ell}$ are the cost of forming the product $A\tilde{V}$ or $\tilde{U}^*A$, respectively, and $N_{2r+\ell}$ is their sum. $\tilde{N}_r$ is the cost of computing Q^*A in the HMT method (Algorithm 6.4, line 3). Here we assume $\ell = \mathcal{O}(r)$. Recall that for generalized Nyström, we need not assume $\tilde{U}, \tilde{V}$ are orthonormal.*

	Pass	MatMul by A	Complexity
Rayleigh–Ritz	Single	1	$N_r + O(mr^2)$
SVD approximation	Single	1	$N_r + \mathcal{O}(mr^2)$
Generalized Nyström	Single	2	$N_{2r+\ell} + \mathcal{O}((m+n)r^2)$
HMT	Double	2	$N_r + \tilde{N}_r + \mathcal{O}(mr^2)$



(a) Without oversampling. (b) With oversampling.

FIG. 5. ***Method comparison for exponentially decaying exact singular values.*** *We show the approximation errors and relative bounds for generalized Nyström (red), Rayleigh–Ritz (green), HMT (blue), and SVD (black). We consider both the case without* (a) *and with* (b) *oversampling.*

$$\tau_i^{GN} = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{GN}\|_2} \approx \frac{1}{2}\frac{2\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{\min_k |\sigma_i - \sigma_k(\bar{A}_{22})| - 2\|E_{RR}\|_2} = \frac{1}{2}\tau_i^{RR}. \tag{6.2}$$

Now, we note that in bound (5.2), the first term is dominant, i.e.,

$$|\sigma_i - \sigma_i^{RR}| \lesssim 2\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}\tau_i^{RR}, \tag{6.3}$$

while, by using (6.2) in (5.1) with $\ell = 0$,

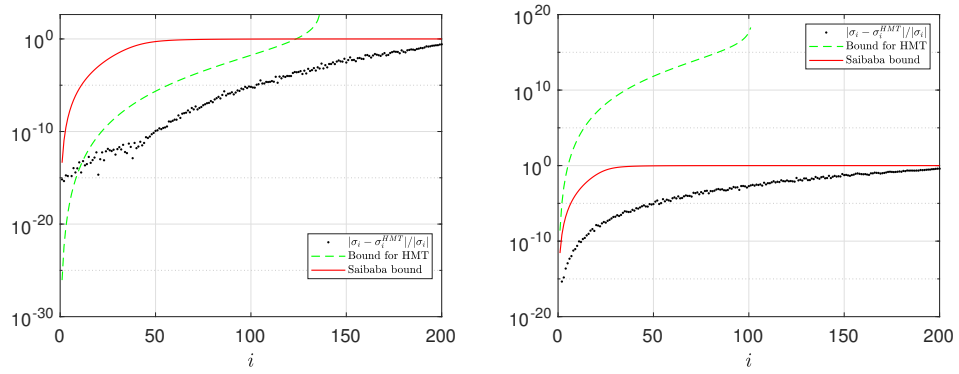
$$|\sigma_i - \sigma_i^{GN}| \lesssim \frac{1}{4}\left\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{12}\right\|_2 (\tau_i^{RR})^2. \tag{6.4}$$

Recall that the derived bounds intersect the corresponding Weyl bound when $\tau_i = 1$. Therefore, $\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}$ and $\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^{\dagger}\bar{A}_{12}\|_2$ are close to the respective perturbation norms and, consequently, close to each other. Thus, the presence of $(\tau_i^{RR})^2$ in (6.4) (as opposed to τ_i^{RR} in (6.3)) gives us a possible explanation of the significantly better behavior of the generalized Nyström approximation.

6.4. Comparison with bound in [26]. As previously mentioned, Saibaba [26] provides a bound on the accuracy of the HMT method in extracting singular values, that, given $\gamma_j = \frac{\sigma_{r+1}}{\sigma_j}$ and

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) $\tilde{V}$ from Gaussian subject to one power iteration.

(b) $\tilde{V}$ randomly generated.

FIG. 6. ***Comparison between relative bound from Corollary*** 5.1 ***and bound in*** [26]. *We show the relative error in extracting the first* $r = 200$ *singular values by the HMT method with* $\ell = 0$ *(black dots), Saibaba's bound in* [26] *(red), and the relative bound derived by Corollary* 5.1 *(dashed green). We considered* A *with algebraically decaying singular values, and approximate subspaces generated* (a) *starting from a Gaussian matrix subject to one power iteration and* (b) *randomly.*

$$\begin{bmatrix} \Omega_1 \\ \Omega_2 \end{bmatrix} := V^* \tilde{V},$$

can be written as

$$\frac{|\sigma_j - \sigma_j^{HMT}|}{|\sigma_j|} \leq 1 - \frac{1}{\sqrt{1 + \left(\frac{\sigma_{r+1}}{\sigma_j}\right)^2 \|\Omega_2 \Omega_1^{\dagger}\|_2^2}}.$$

Since this is a bound on the relative error, as opposed to our bound that is on the absolute error, we divide Corollary 5.1 by $|\sigma_i|$, to obtain a comparable bound. We perform experiments with the matrix with algebraically decaying singular values as described above and no oversampling, i.e., $\ell = 0$. We consider the approximate subspace $\tilde{V}$ to be, at first, as in the previous experiments and then randomly generated following the Haar distribution, i.e., as the orthogonal factor of a Gaussian matrix.

Figure 6 shows a comparison of the bound in [26] and the bound proposed here. We note that for approximate subspaces that are randomly generated, Saibaba's bound is very sharp and describes better the error compared to the relative bound derived from Corollary 5.1. However, when the approximate subspace contains information of the matrix A, we see that the proposed bound became sharper, while Saibaba's bound does not clearly depict the relative error.

6.5. Computability. A bound is usually derived to be used as a theoretical tool to study the behavior of the bounded quantity and to be used in practical computations. In the previous sections, we have shown how the derived bounds can serve the first purpose. However, as expressed so far, these bounds are impractical to use. Indeed, the presence of the exact singular values makes the bounds useless in practical computations. In this section, we discuss an approach that can overcome this issue. Once again we analyze the generalized Nyström approximation, and, for the sake of simplicity, we restrict ourselves to the case where no oversampling is employed ($\ell = 0$). As seen before, the theory can be easily generalized to the other strategies and to the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

case with oversampling; see the supplementary material (Supplementary-Materials.zip [local/web 12.2KB]) for the implementation of each case.

To obtain bound (5.1), we used (3.4) to interpret the generalized Nyström approximation as a perturbation of $\bar{A}$, i.e., $\bar{A}_{GN} = \bar{A} - E_{GN}$. This results in the final expression of the bound including quantities that depend on $\bar{A}$, and consequently on A, which are typically unknown in practice. To address this, another less intuitive interpretation is trivially available. Specifically, the roles can be reversed, viewing the matrix $\bar{A}$ as a perturbation of the corresponding generalized Nyström approximation, i.e., $\bar{A} = \bar{A}_{GN} + E_{GN}$, where the perturbation E_{GN} remains the same. We can then apply Theorem 4.1 to this interpretation.

COROLLARY 6.1. *Consider an $m \times n$ matrix A with singular values $\{\sigma_i\}_{i=1}^n$ and approximations $\tilde{V} \in \mathbb{R}^{n\times r}$ and $\tilde{U} \in \mathbb{R}^{m\times r}$ to its leading singular subspaces. Consider the generalized Nyström approximation $A_{GN,\tilde{V},\tilde{U}} = A\tilde{V}(\tilde{U}^*A\tilde{V})^\dagger\tilde{U}^*A$ and its singular values $\{\sigma_i^{GN}\}_{i=1}^n$. Define $\bar{A} = Q_1^*AQ_2$, with $Q_1 = \begin{bmatrix}\tilde{U} & \tilde{U}_\perp\end{bmatrix}$ and $Q_2 = \begin{bmatrix}\tilde{V} & \tilde{V}_\perp\end{bmatrix}$, and its blocks as in* (3.2)*. Let, for $i = 1, \dots, n$,*

$$\bar{\tau}_i = \frac{\max\{\|\bar{A}_{12}\|_2, \|\bar{A}_{21}\|_2\}}{\min_k |\sigma_i(A_{GN}) - \sigma_k(\bar{A}_{21}\bar{A}_{11}^\dagger\bar{A}_{12})| - 2\|E_{GN}\|_2}.$$

Then, for all i for which $\bar{\tau}_i > 0$, it holds that

$$|\sigma_i - \sigma_i^{GN}| \le \left\|\bar{A}_{22} - \bar{A}_{21}\bar{A}_{11}^\dagger\bar{A}_{12}\right\|_2 \bar{\tau}_i^2. \tag{6.5}$$

We refer to bound (6.5) as the backward bound. Notice that instead of the exact singular values, as in (5.1), we now have the computed singular values. However, the gap in the denominator of $\bar{\tau}_i$ is more complicated to compute. For this reason, we can use the approximation

$$\min_k |\sigma_i(A_{GN}) - \sigma_k(\bar{A}_{21}\bar{A}_{11}^\dagger\bar{A}_{12})| \approx \sigma_i(A_{GN}) - \|\bar{A}_{21}\bar{A}_{11}^\dagger\bar{A}_{12}\|_2$$

to obtain an approximated form of the backward bound. Figure 7 shows that, for both exponentially and algebraically decaying exact singular values, both the backward bound and the approximated backward bound are not far from the forward bound (5.1). In practice, the backward bound can be further approximated by, e.g., estimating norms by random sampling [19], approximating the norm of the Schur complement [1], [10, sect. 8.2], and, for the generalized Nyström approximation with oversampling, making use of properties of the orthogonal projection $\bar{A}_{11}^\dagger\bar{A}_{11}$.

7. Discussion. We conclude with a discussion of the possible extensions and open problems.

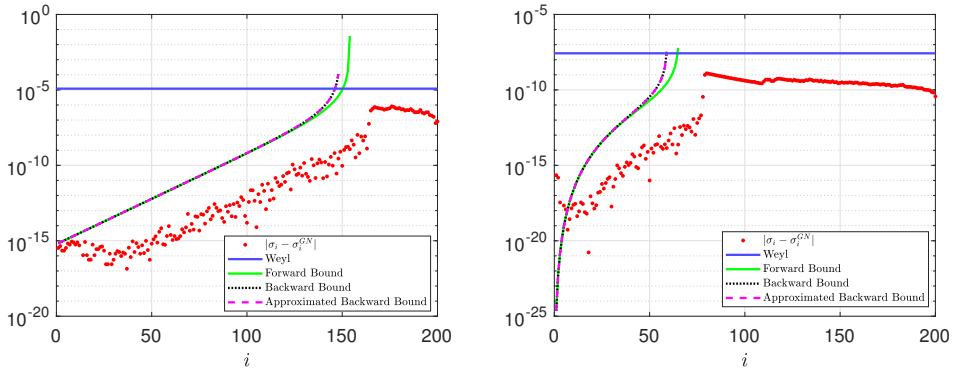
We believe there is room to further explore how Theorem 4.1 can be used as a tool to formally compare strategies to extract singular values. For example, a study of the dominant term in each bound could lead to further understanding of a possible accuracy hierarchy.

As presented, the technique to improve the bound for the generalized Nyström approximation when $\ell > 0$ is based on heuristics. Thus, there is space to investigate the formal validity of such strategy and to devise alternative ones.

It is worth noting that the behavior of the presented bounds differs from what is illustrated in section 6 when, instead of good approximations of the leading singular subspaces, we take $\tilde{U}$ and $\tilde{V}$ to be random matrices. Our experiments show that,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy



(a) Exponential Decay. (b) Algebraic Decay.

FIG. 7. ***Computable bound for generalized Nyström without oversampling.*** *We show the error in extracting the first $r = 200$ singular values by generalized Nyström approximation with $\ell = 0$ (red dots), Weyl's inequality* (2.3) *(blue), the forward bound* (5.1) *(green), the backward bound* (6.5) *(black), and the approximated backward bound (magenta). Both exponential* (a) *and algebraic* (b) *singular values decays are considered.*

in this case, the change of the errors in extracting the leading singular values of the generalized Nyström approximation depends less on i, giving very similar accuracy errors for the first r singular values. In fact, in this case Weyl's inequality is more descriptive than in the previous experiments, indicating that the need to find better bounds is diminished.

Finally, the presented analysis could be used to characterize the difference in the behavior of the generalized Nyström approximation with and without oversampling. This would have an impact also on the analysis of the generalized Nyström low-rank approximation.

REFERENCES

[1] M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numer., 14 (2005), pp. 1–137, https://doi.org/10.1017/S0962492904000212.

[2] C. Boutsidis and A. Gittens, *Improved matrix algorithms via the subsampled randomized Hadamard transform*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1301–1340, https://doi.org/10.1137/120874540.

[3] K. L. Clarkson and D. P. Woodruff, *Numerical linear algebra in the streaming model*, in Proceedings of the 41st Annual ACM Symposium on the Theorey of Computing, 2009, pp. 205–214, https://doi.org/10.1145/1536414.

[4] K. L. Clarkson and D. P. Woodruff, *Low-rank approximation and regression in input sparsity time*, J. ACM, 63 (2017), pp. 1–45, https://doi.org/10.1145/3019134.

[5] C. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation.* III, SIAM J. Numer. Anal., 7 (1970), pp. 1–46, https://doi.org/10.1137/0707001.

[6] A. Dax, *On extremum properties of orthogonal quotients matrices*, Linear Algebra Appl., 432 (2010), pp. 1234–1257, https://doi.org/10.1016/j.laa.2009.10.034.

[7] Z. Drmač, *Numerical methods for accurate computation of the eigenvalues of Hermitian matrices and the singular values of general matrices*, SeMA J., 78 (2021), pp. 53–92, https://doi.org/10.1007/s40324-020-00229-8.

[8] Z. Drmač and K. Veselić, *New fast and accurate Jacobi SVD algorithm.* I, SIAM J. Matrix Anal. Appl., 29 (2008), pp. 1322–1342, https://doi.org/10.1137/050639193.

[9] Z. Drmač and K. Veselić, *New fast and accurate Jacobi SVD algorithm.* II, SIAM J. Matrix Anal. Appl., 29 (2008), pp. 1343–1362, https://doi.org/10.1137/05063920X.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

[10] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Numer. Mathe. Sci. Comput., Oxford University Press, Oxford, UK, 2005.
[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, 2013.
[12] M. Gu, *Subspace iteration randomization and singular value problems*, SIAM J. Sci. Comput., 37 (2015), pp. A1139–A1173, https://doi.org/10.1137/130938700.
[13] N. Halko, P. G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288, https://doi.org/10.1137/090771806.
[14] P. C. Hansen, *Discrete Inverse Problems*, SIAM, Philadelphia, 2010, https://doi.org/10.1137/1.9780898718836.
[15] P. C. Hansen, V. Pereyra, and G. Scherer, *Least Squares Data Fitting with Applications*, Johns Hopkins University Press, Baltimore, 2013.
[16] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2012.
[17] C.-K. Li and R.-C. Li, *A note on eigenvalues of perturbed Hermitian matrices*, Linear Algebra Appl., 395 (2005), pp. 183–190, https://doi.org/10.1016/j.laa.2004.08.026.
[18] R.-C. Li, *Relative perturbation theory:* I. *Eigenvalue and singular value variations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 956–982, https://doi.org/10.1137/S089547989629849X.
[19] P.-G. Martinsson and J. A. Tropp, *Randomized numerical linear algebra: Foundations and algorithms*, Acta Numer., 29 (2020), pp. 403–572, https://doi.org/10.1017/S0962492920000021.
[20] E. S. Meckes, *The Random Matrix Theory of the Classical Compact Groups*, Cambridge Tracts in Math. 218, Cambridge University Press, Cambridge, UK, 2019.
[21] Y. Nakatsukasa, *Eigenvalue perturbation bounds for Hermitian block tridiagonal matrices*, Appl. Numer. Math., 62 (2012), pp. 67–78, https://doi.org/10.1016/j.apnum.2011.09.010.
[22] Y. Nakatsukasa, *Fast and Stable Randomized Low-Rank Matrix Approximation*, https://arxiv.org/abs/2009.11392v1, 2020.
[23] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Classics Appl. Math. 20, SIAM, Philadelphia, 1998, https://doi.org/10.1137/1.9781611971163.
[24] V. Rokhlin, A. Szlam, and M. Tygert, *A randomized algorithm for principal component analysis*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1100–1124, https://doi.org/10.1137/080736417.
[25] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, 2nd ed., Classics Appl. Math. 66, SIAM, Philadelphia, 2011, https://doi.org/10.1137/1.9781611970739.
[26] A. K. Saibaba, *Randomized subspace iteration: Analysis of canonical angles and unitarily invariant norms*, SIAM J. Matrix Anal. Appl., 40 (2019), pp. 23–48, https://doi.org/10.1137/18M1179432.
[27] G. W. Stewart, *On the early history of the singular value decomposition*, SIAM Rev., 35 (1993), pp. 551–566, https://doi.org/10.1137/1035134.
[28] G. W. Stewart, *Matrix Algorithms. Volume* 1*: Basic Decompositions*, SIAM, Philadelphia, 1998.
[29] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Computer Science and Scientific Computing, Academic Press, New York, 1990.
[30] J. A. Tropp, *Improved analysis of the subsampled randomized Hadamard transform*, Adv. Adapt. Data Anal., 3 (2011), pp. 115–126, https://doi.org/10.1142/S1793536911000787.
[31] J. A. Tropp and R. J. Webber, *Randomized Algorithms for Low-Rank Matrix Approximation: Design, Analysis, and Applications*, https://arxiv.org/abs/2306.12418, 2023.
[32] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher, *Practical sketching algorithms for low-rank matrix approximation*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 1454–1485, https://doi.org/10.1137/17M1111590.
[33] J. Wedderburn, *Lectures on Matrices*, Amer. Math. Soc. Colloq. Publ. 17, AMS, Providence, RI, 1934.
[34] F. Woolfe, E. Liberty, V. Rokhlin, and M. Tygert, *A fast randomized algorithm for the approximation of matrices*, Appl. Comput. Harmon. Anal., 25 (2008), pp. 335–366, https://doi.org/10.1016/j.acha.2007.12.002.
[35] L. Xin-guo, *On Rayleigh quotient theory for the eigenproblem and the singular value problem*, J. Comput. Math., 10 (1992), pp. 216–224, https://www.jstor.org/stable/45340282.
[36] F. Zhang, ed., *The Schur Complement and Its Applications*, Numer. Methods Algorithms 4, Springer, Berlin, 2005.

Downloaded 12/15/25 to 188.64.37.199 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.