On the Computation of the SVD of Fourier Submatrices

S. Dirckx · D. Huybrechs · R. Ongenae

Received: 30 August 2022 / Revised: 17 February 2023 / Accepted: 1 March 2023 / Published online: 17 April 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
Contiguous submatrices of the Fourier matrix are known to be ill-conditioned. In a recent paper in SIAM review A. Barnett has provided new bounds on the rate of ill-conditioning (Barnett in SIAM Rev 64:105–131, 2022). In this paper we focus on the corresponding singular value decomposition. The singular vectors can be computed from the so-called periodic discrete prolate spheroidal sequences, named in analogy to spheroidal wave functions which are associated with the continuous Fourier transform. Their numerical computation is hampered by the clustering of singular values. We collect and expand known results on the stable numerical computation of the singular value decomposition of Fourier submatrices. The prolate sequences are eigenvectors of a tridiagonal matrix whose spectrum is free of clusters and this enables their computation. We collect these observations in a simple and convenient algorithm. The corresponding singular values can be accurately computed as well, except when they are small. Even then, small singular values can be computed in high-precision arithmetic with modest computational effort, even for large and extremely ill-conditioned submatrices. We illustrate the computations and point out a few applications in which Fourier submatrices arise.

Keywords Discrete Fourier transform (DFT) · Singular value decomposition · Prolate spheroidal wave functions

Mathematics Subject Classification 65T50 · 65F55

1 Introduction

The goal of this paper is an efficient and stable algorithm to compute the singular value decomposition of a contiguous subblock $A$ of the discrete Fourier matrix. The algorithm is
based on the observation that diagonally scaled versions of $A^*A$ and $AA^*$ commute with two analytically known tridiagonal matrices. Whereas the SVD of $A$ contains clusters, which prevent accurate computation of the singular vectors, the spectra of the tridiagonal matrices have well-separated eigenvalues. Hence, the corresponding eigenvectors can be computed stably (see Fig. 1) and, owing to sparsity, efficiently. However, very small singular values must still be computed in higher-precision arithmetic. Our findings will be summarized in the so-called Commuting Tridiagonal Algorithm (CTA), given in Algorithm 1.

The structure of the paper is as follows. We define Fourier submatrices and establish notation in Sect. 2. We describe the computational algorithms in Sect. 3. We end with some illustrations, a few applications and possible links to further research in Sect. 4. First, however, we remark on the context of the results in Sect. 1.1 below. The topic of Fourier submatrices relates to the literature on time-frequency localisation of prolate spheroidal wave functions, as pioneered by Slepian in a series of papers starting with [13]. We do not delve deeper into that topic in the paper, but we sketch the context and origin of our statements and results.

1.1 Historical Perspective of Prolate Sequences

We collect and generalize the formulas for Periodic Discrete Prolate Spheroidal Sequences (P-DPSS) from [6, 7, 15]. The P-DPSS were introduced by Grünbaum [6], and Jain and Ranganath [7]. Earlier studies were devoted to the (continuous) prolate spheroidal wave functions [13] and discrete (but non-periodic) prolate spheroidal sequences [12]. Both of these are eigenfunctions of a bandlimiting operator, but are better computed as eigenfunctions of a related commuting differential operator: see [10] for a comprehensive treatment. To that end, for the periodic discrete setting of P-DPSS, Grünbaum [6] derived a tridiagonal matrix which can be seen as a discrete differential operator, that commutes with a diagonal scaling $C$ of the Gram matrix of the Fourier submatrix $A$. Xu and Chamzas pointed out that for general P-DPSS two singular values of $A$, 0 and $\sqrt{N}$, may have higher multiplicity [15]. This means that the corresponding singular vectors are not uniquely defined. We follow the methodology of [15] and define singular vectors of $A$ using the unique (up to normalization) eigenvectors of the tridiagonal matrix. Asymptotic and non-asymptotic properties of the P-DPSS were studied further in [1, 3, 17].

2 The Discrete Fourier Transform

A common definition of the discrete Fourier transform of a vector \( \{x_k\}_{k=0}^{N-1} \) of length $N$ is

\[
X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i kn/N}, \quad k = 0, \ldots, N - 1.
\]

\[1\] In fact, his construction produces a commuting tridiagonal matrix for both Gramians $C^*C$ and $CC^*$. 

Springer
Its inverse is
\[ x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{2\pi i kn/N}, \quad n = 0, \ldots, N - 1. \]

Other definitions may differ in the position of the $1/N$ term, and in the sign of the exponent in the complex exponential.

For convenience of implementation we prefer to work with vectors $x$ and $X$ in $\mathbb{C}^N$ with indices ranging from 1 to $N$, as is standard in linear algebra, although many of the following formulas are mathematically more elegant with indices starting at 0. With this convention, the $N \times N$ Fourier matrix $F \in \mathbb{C}^{N\times N}$ is given by
\[ F_{j,k} = e^{-2\pi i (j-1)(k-1)/N} = \omega^{-(j-1)(k-1)}, \quad 1 \leq j, k \leq N, \]
in which we have defined the twiddle factor\(^2\)
\[ \omega = e^{2\pi i/N}. \]
The inverse DFT amounts to a matrix–vector product with $F^*/N$, i.e., we have
\[ X = Fx \quad \text{and} \quad x = \frac{1}{N} F^*X \]
where $F^*$ denotes the adjoint of $F$.

### 2.1 Submatrices of the Fourier Matrix

We focus on the contiguous submatrix $A \in \mathbb{C}^{p\times q}$ starting from the $(1, 1)$ entry of $F$,
\[ A_{j,k} = F_{j,k}, \quad 1 \leq j \leq p, \quad 1 \leq k \leq q, \]
with $1 \leq p, q \leq N$. Any other contiguous submatrix of $F$ with the same dimensions as $A$ is easily obtained by diagonal scalings. Shifting the submatrix along rows and columns corresponds to multiplication by diagonal matrices to the right and to the left respectively.

Let us be precise. We define the diagonal matrix $D_n \in \mathbb{C}^n$, with diagonal entries
\[ (D_n)_{j,j} = \omega^{-(j-1)}, \quad j = 1, \ldots, n. \]
Consider the Fourier submatrix $B$ of dimension $p \times q$ but starting from index $(j_0, k_0)$ of $F$, and taking into account periodicity if the submatrix ‘wraps around’,
\[ B_{j,k} = F_{1+\text{mod}(j_0+j-2,N),1+\text{mod}(k_0+k-2,N)}, \quad 1 \leq j \leq p, \quad 1 \leq k \leq q. \]
Then we have
\[ B = \omega^{-(j_0-1)(k_0-1)} D_p^{-1} A D_q^{j_0-1} \]
and, conversely,
\[ A = \omega^{(j_0-1)(k_0-1)} D_p^{-(k_0-1)} B D_q^{-(j_0-1)}. \]
Hence, in the following, without loss of generality we consider only $A$.

---

\(^2\) The name ‘twiddle factor’ originated with [4] and is prevalent in the literature on FFT. It is used here as a synonym for a complex root of unity. In general, powers of $\omega$ are also often called twiddle factors.
Fig. 1 The singular values of Fourier submatrix $A$ cluster towards $\sqrt{N}$ and towards 0 (top panel). The plateau contains approximately $pq/N$ singular values, shown by dashed lines. In contrast, the spectra of the symmetric tridiagonal matrices $J(p, q)$ and $J(q, p)$ (bottom-left and bottom-right) are free of clusters and contain only simple eigenvalues. All eigenvectors can be computed to high accuracy.

2.2 The Singular Value Decomposition of $A$

Our goal is the computation of the singular value decomposition of $A$, 

$$A = U \Sigma V^*. \tag{2}$$

With $r = \min(p, q)$, we aim to compute the reduced SVD with $U \in \mathbb{C}^{p \times r}$, $\Sigma \in \mathbb{C}^{r \times r}$ and $V \in \mathbb{C}^{q \times r}$. A full SVD can also be computed by generating $|p - q|$ additional orthogonal vectors in $U$, if $p > q$, or in $V$, if $q > p$. A typical profile of singular values is illustrated further on in Fig. 1a for $N = 128$ and a few choices of $p$ and $q$.

3 Computation of the SVD

We continue the historical sketch of Sect. 1.1 with some more technical observations to motivate what follows. Grünbaum studied Fourier series of odd length, i.e., with symmetric frequencies, and correspondingly assumed both $p$ and $q$ to be odd. A similar assumption is made in the explicit formulas in [8] in the context of the Fourier extension approximation scheme. Xu and Chamzas adopt the assumption that one of $p$ or $q$ is odd [15]. However, their analysis generalizes straightforwardly to the case of even $p$ and $q$.\(^3\) Jain and Ranganath in [7, §IX.B] do not make any such assumptions in the definition of P-DPSS, but they do not

\(^3\) In particular, their arguments can be used to derive that $A^* A$ has separated eigenvalues if and only if $q \leq p$ and $p + q \leq N$. As a consequence, repeated singular values in the (compact) SVD of $A$ only occur whenever
supply formulas. Edelman et al. in [3] study square submatrices and assume that \( p \) divides \( N \). Zhu et al. [17] and Barnett [1] make no assumptions on \( N \), \( p \) or \( q \), but they do not supply computational formulas. Moreover, due to differences in notation and indexing (starting from 0 or symmetric around 0) in all these references, implementation of the formulas require some attention and care.

The formulas in this section are a generalization of those given in the cited references [6, 8], to accommodate any combination of \( N \), \( p \), and \( q \). They are provided for completeness and ease of implementation. They do not contain new insights and are simply stated without proof. The main complication is the correct treatment of even \( p \) and \( q \).

### 3.1 The Periodic Prolate Matrix

The left and right singular vectors of \( A \) are eigenvectors of \( A^*A \) and \( AA^* \) respectively. These are complex matrices. The computations simplify by considering a more symmetric matrix \( C \) first, which is such that \( C^*C \) and \( CC^* \) are real.

Matrix \( C \in \mathbb{C}^{p \times q} \) is similar to the shifted submatrix \( B \) given by (1), though with possibly fractional shifts if either \( p \) or \( q \) is even:

\[
C_{j,k} = \omega^{-(j-\frac{p+1}{2})(k-\frac{q+1}{2})} D_p^{\frac{(q-1)}{2}} A D_q^{\frac{(p-1)}{2}}.
\]

Equivalently, its entries are

\[
C_{j,k} = \omega^{-(j-\frac{p+1}{2})(k-\frac{q+1}{2})}, \quad 1 \leq j \leq p, 1 \leq k \leq q,
\]

and we have the inverse relation

\[
A = \omega^{-(j-\frac{p+1}{2})(k-\frac{q+1}{2})} D_p^{\frac{(q-1)}{2}} C D_q^{\frac{(p-1)}{2}}.
\]

The real Hermitian matrices \( C^*C \) and \( CC^* \) are both prolate matrices. The periodic prolate matrix \( S(p, q) \in \mathbb{R}^{q \times q} \) is defined by

\[
[S(p, q)]_{j,k} = \frac{\sin \left( \frac{p(j-k)\pi}{N} \right)}{\sin \left( \frac{(j-k)\pi}{N} \right)}, \quad 1 \leq j, k \leq q, j \neq k.
\]

for the off-diagonal entries, and (by Hôpital’s rule)

\[
[S(p, q)]_{j,j} = p
\]

along the diagonal. The matrix \( S(p, q) \) can also be viewed as a Dirichlet kernel matrix.\(^5\) We have

\[
C^*C = S(p, q) \quad \text{and} \quad CC^* = S(q, p).
\]

\(^4\) We have to fix branch cuts in case \( p \) or \( q \) is even. The \((k, k)\) entry of \( D_p^{\frac{(q-1)}{2}} \) here is \( \omega^{(k-1)(q-1)/2} \) and it is best implemented that way explicitly. Numerically raising the diagonal matrix \( D_p \) to the power \( \frac{(q-1)}{2} \) may yield a different outcome, due to choosing a different square root. A similar comment applies to powers of \( D_q \).

\(^5\) For each \( n \) the Dirichlet kernel \( D_n \) is a radial kernel given (up to a factor) by \( D_n(r) = \frac{\sin((2n+1)r/2)}{\sin(r/2)} \).
Hence, possibly up to a normalization factor, the left and right singular vectors of $C$ are given by the eigenvectors of $S(q, p)$ and $S(p, q)$, respectively.

When $q = N$, one can verify that $S(p, N)/N$ is an orthogonal projection. Let $l := \lceil \frac{p}{2} \rceil$, $\tilde{l} := \lfloor \frac{p}{2} \rfloor$ and $\sigma = (1+l-\tilde{l})/2$. Then, with $K := \{0, \ldots, l-1\} \cup \{N-\tilde{l}, \ldots, N-1\}$, $S(p, N)$ projects orthogonally onto $\text{span}\{(\omega^{-k})_{j=0}^{N-1} | k \in K + \sigma\}$. Note the non-integer frequencies for even $p$. In general, $S(p, q)$ is not associated with an orthogonal projection.

### 3.2 Two Commuting Tridiagonal Matrices

The spectrum of the periodic prolate matrix is clustered. This renders the computation of its eigenvectors ill-conditioned, a problem that also appears in the continuous case. The enabling observation for computational purposes, first established in this context by Grünbaum [6], is that $S(p, q)$ commutes with a real and symmetric tridiagonal matrix $J(p, q) \in \mathbb{R}^{q \times q}$.

The diagonal entries of this tridiagonal matrix are

$$[J(p, q)]_{k,k} = \cos \left( \frac{\pi(2k - q - 1)}{N} \right) \cos \left( \frac{p\pi}{N} \right), \quad 1 \leq k \leq q. \quad (7)$$

The first diagonal above the main diagonal is given by

$$[J(p, q)]_{k,k+1} = -\sin \left( \frac{\pi k}{N} \right) \sin \left( \frac{\pi(q - k)}{N} \right), \quad 1 \leq k \leq q - 1. \quad (8)$$

The subdiagonal elements $[J(p, q)]_{k+1,k} = [J(p, q)]_{k,k+1}$ are defined by symmetry. Since these elements are all nonzero, the Sturm sequence property (see [5, Theorem 8.5.1]) ensures that $J(p, q)$ has no eigenvalues of multiplicity greater than one. We have the commutation relation

$$J(p, q)S(p, q) - S(p, q)J(p, q) = 0.$$ 

This, combined with the simple spectrum of $J(p, q)$, implies that the eigenvectors of $J(p, q)$ are also eigenvectors of $S(p, q)$. Indeed, for any vector $v$ such that $J(p, q)v = \lambda v$ we then have that $J(p, q)S(p, q)v = S(p, q)J(p, q)v = \lambda S(p, q)v$, which means that $S(p, q)v$ must lie in the eigenspace of $J(p, q)$ associated with the eigenvalue $\lambda$. Since $J(p, q)$ has simple spectrum, this space is precisely $\text{span}\{v\}$, which implies that $v$ is an eigenvector of $S(p, q)$. Since $S$ may have repeated eigenvalues the converse might not be true. The spectrum of $J(p, q)$ is also free of clusters, and hence the eigenvalue problem of $J(p, q)$ is better conditioned than that of $S(p, q)$ [15]. The above also holds if we flip $p$ and $q$.

The above observation forms the basis for the ‘Commuting Tridiagonal Algorithm’ (CTA, see Algorithm 1 below).\(^6\)

**Remark 1** We illustrate the advantage of the CTA approach in Fig. 2, in which we compare the first singular vector as computed by MatLab’s $\text{svd}$ function (denoted by $\hat{u}_1$) and as computed by CTA (denoted by $u_1$). In exact arithmetic the vectors $u_1$ and $\hat{u}_1$ should coincide, up to a complex scalar multiple. However, due to the clustering of singular values of $A$, $\hat{u}_1$ is ‘contaminated’ by singular vectors with singular values close to $\sqrt{N}$. Similarly, the vectors corresponding to the smallest singular values are contaminated. In the plunge region, the

\(^6\) A complete and user-friendly implementation of the algorithm is available at the time of writing in a Julia package called ‘FourierSubmatrices.jl’, available from https://github.com/daanhb/FourierSubmatrices.jl. The experiments of this paper have been computed in Matlab, except when high precision was required.
contamination is modest to negligible, as illustrated by studying the vectors $u_m$ and $\hat{u}_m$ at $m = \frac{pq}{N}$.

We illustrate and quantify this contamination further in Fig. 3. We use the notation $U$ for the matrix containing the left CTA vectors, i.e. $U := [u_1, \ldots, u_p]$. To study how contamination depends on the subblock size $(p, q)$, we introduce a novel contamination measure. Given a triple $(p, q, N)$, let $P_k$ denote the orthogonal projection onto the $k$th CTA vector i.e. $P_k = u_k u_k^\ast$. We then set

$$
\mu_k(p, q; N) := \|P_k^\perp \hat{u}_1\|_2.
$$

Of particular interest are $\mu_1$ and $\mu_p$, since they measure how contaminated the extremal singular vectors are. Similar measures can be introduced to study the right singular vectors. The dependence of $\mu_1$ and $\mu_p$ on $(p, q)$ is reported on in Fig. 4.

### 3.3 The Singular Vectors of $A$

Let $\tilde{v}_k \in \mathbb{R}^q$ be an eigenvector of $J(p, q)$ with eigenvalue $\tilde{\lambda}_k \in \mathbb{R}$,

$$
J(p, q)\tilde{v}_k = \tilde{\lambda}_k \tilde{v}_k.
$$

All quantities are real-valued owing to the symmetry of $J(p, q)$. Since $J(p, q)$ and $S(p, q)$ commute, we also have

$$
S(p, q)\tilde{v}_k = \lambda_k \tilde{v}_k
$$

for some $\lambda_k \in \mathbb{R}$ which in general differs from $\tilde{\lambda}_k$, but which is still real. Consequently, by (6),

$$
C^*C\tilde{v}_k = \lambda_k \tilde{v}_k.
$$
Fig. 3 Top: Inner products $U^* \hat{u}_m$ of the CTA vectors associated with $J(p, q)$ with the singular vectors $\hat{u}_1, \hat{u}_m$ of $A$ as computed by MatLab’s $\text{svd}$. Here $m = \frac{pq}{N}$ and $(N, p, q) = (2500, 800, 351)$. Bottom: corresponding singular values of $A$. The dashed line delineates $\sigma_k < 1 - \sqrt{\epsilon}$, with $\epsilon$ the machine precision. The vector $\hat{u}_1$ is contaminated by all singular vectors to the left of this line.

Fig. 4 Contamination measures $\mu_1(p, q; N)$ (left) and $\mu_p(p, q; N)$ (right) for $N = 1000$ as $(p, q)$ vary in $\{1, \ldots, N\}$. These two plots together show that significant contamination occurs for any subblock $A = F_N(1 : p, 1 : q)$ except for the very thinnest.

We can conclude that the corresponding singular value of $C$ is $\sigma_k = \sqrt{\lambda_k}$. However, $\sigma_k$ is not computed this way, since we solve the eigenvalue problem (9) for the tridiagonal matrix rather than (10). There is no known relation between $\tilde{\lambda}_k$ and $\lambda_k$.

Moving to $A$, from (3) we find that

$$A^s AD_q^{-\frac{(p-1)}{2}} \tilde{v}_k = \lambda_k D_q^{-\frac{(p-1)}{2}} \tilde{v}_k.$$  

Hence, up to normalization, we find that

$$v_k = D_q^{-\frac{(p-1)}{2}} \tilde{v}_k$$  

(11)
is a right singular vector of $A$.

For the left singular vectors we proceed similarly. Thus, $\tilde{u}_k \in \mathbb{R}^p$ is an eigenvector of $J(q, p)$ with eigenvalue $\tilde{\lambda}_k \in \mathbb{R}$ (different from the eigenvalue in (9) above),

$$J(q, p)\tilde{u}_k = \tilde{\lambda}_k \tilde{u}_k.$$  \hspace{1cm} (12)

The corresponding eigenvalue for $S(q, p)$ is $\lambda_k \in \mathbb{R}$. We have

$$CC^*\tilde{u}_k = \lambda_k \tilde{u}_k.$$  \hspace{1cm} (13)

The corresponding singular value is $\sigma_k = \sqrt{\lambda_k}$. From (3) we find that

$$AA^*D_p^{(q-1)/2}u_k = \lambda_k D_p^{(q-1)/2}u_k.$$  \hspace{1cm} (14)

Hence, again up to normalization, the left singular vector $u_k$ of $A$ is

$$u_k = D_p^{(q-1)/2} \tilde{u}_k.$$  \hspace{1cm} (15)

### 3.4 The Singular Values of $A$

It remains to determine the singular value and a suitable normalization for $u_k$ and $v_k$. This is achieved most easily by computing the projections

$$\hat{\sigma}_k = u_k^*A v_k, \quad 1 \leq k \leq \min(p, q) \hspace{1cm} (14)$$

where we have ordered the eigenvectors $v_k, u_k$ by decreasing corresponding eigenvalues. The computed value $\hat{\sigma}_k$ may be complex, yet $\sigma_k$ should be real by the definition of the SVD. Thus, the singular value is

$$\sigma_k = |\hat{\sigma}_k|.$$  

We can incorporate the phase difference into one of the singular vectors, and in our implementation we arbitrarily choose to do so in $u_k$:

$$u_k \rightarrow \frac{|\hat{\sigma}_k|}{\hat{\sigma}_k} u_k, \quad 1 \leq k \leq \min(p, q).$$  \hspace{1cm} (15)

**Remark 2** Small singular values can not be computed to high accuracy in this way. Indeed, since all entries of $A$ have modulus 1, small singular values result from large cancellation effects in (14). Hence, in the absence of alternative methods, the computation of small singular values to high accuracy currently requires high-precision arithmetic. We emphasize that such is not the case for the singular vectors.

### 3.5 Efficient Matrix–Vector Product Using an FFT of Length $N$

The computation of the singular values via (14) requires $\min(p, q)$ matrix vector products with $A$, which is a costly affair. However, since $A$ is a submatrix of the DFT matrix of dimension $N$, an efficient matrix–vector product is also available using the FFT of length $N$ [8]. The multiplication $Av$ can be computed by zero-padding $v$ to a vector of length $N$ (note that $v$ has length $q$), followed by the FFT, followed by restricting to the first $p$ entries. We write that as

$$A = R_p F E_q,$$  \hspace{1cm} (16)
with $E_q \in \mathbb{R}^{N \times q}$, $F$ the DFT matrix of dimension $N$, and $R_p \in \mathbb{R}^{p \times N}$.

Moreover, we have the relation

$$A v_k = \sigma_k u_k, \quad 1 \leq k \leq \min(p, q).$$

Hence, once $v_k$ is determined, both $\sigma_k \in \mathbb{R}$ and $u_k$ can be determined from (17) using the FFT.

Whether the FFT-based approach is more efficient than computing the matrix–vector product with $A$ directly depends on the relative sizes of $p$ and $q$ versus $N$: the matrix vector product has cost $O(pq)$, the FFT has cost $O(N \log N)$. In some applications $N$ may be much larger than $p$ and $q$, for example in the parallel FFT computation of [3].

Algorithm 1: Commuting Tridiagonal Algorithm (CTA)

| input : Triple $(p, q, N)$ representing leading block $A = F_N(1 : p, 1 : q)$ |
| output: Approximate SVD $A \approx USV^*$ |
| /* Part 1: compute singular vectors $(U, V)$ */ |
| 1 Construct $J_{pq}$ and $J_{qp}$ by Eq. 7 and 8; |
| 2 $r \leftarrow \min\{p, q\}$; |
| 3 $V \leftarrow \text{eig}(J_{pq})$; // sorted eigenvectors of (symmetric positive definite) $J_{pq}$ |
| 4 $U \leftarrow \text{eig}(J_{qp})$; // sorted eigenvectors of (symmetric positive definite) $J_{qp}$ |
| 5 $V \leftarrow D_q^{-(p-1)/2} V(:, 1 : r)$; // see Eq. 11 |
| 6 $U \leftarrow D_p^{-(q-1)/2} U(:, 1 : r)$; // see Eq. 13 |
| /* Part 2: compute singular values $S$ */ |
| 7 $S \leftarrow U^* R_p \text{fft}(E_q V)$; // column-wise fft, see Eq. 16 |
| 8 $F \leftarrow \exp(-i \cdot \arg(S))$; // $S$ is diagonal |
| 9 $S \leftarrow FS$; |
| 10 $U \leftarrow U/F$; |

3.6 Computational Complexity

Summarizing, we solve (9) for $\tilde{v}_k$, and use (11) to find the right singular vectors of $A$. At this point we have a choice. We can solve (12) for $\tilde{u}_k$, use (13) to find the left singular vectors of $A$, compute the singular value from (14) and finally normalize $u_k$ following (15). The alternative to (12)–(15) is to use (17) to find both $\sigma_k$ and $u_k$ simultaneously.

The computational complexities of these steps are as follows:

- Computing the eigenvalues in (9) for the symmetric tridiagonal matrix $J(p, q)$, using for example the symmetric QR algorithm, requires $O(q^2)$ floating point operations (FLOP), as shown in [5, Ch. 8], after which computing each eigenvector requires an additional $O(q)$ FLOP (see [14]), for a total of $O(q^2)$ FLOP. Similarly, the eigendecomposition problem in (12) requires $O(p^2)$ FLOP.
- The diagonal scaling in (11) requires just $O(q)$ FLOP, and that of (13) requires $O(p)$ FLOP.
- Determining the singular values from (14) or (17) requires $\min(p, q)$ matrix–vector products with $A$. The cost in FLOP is:
Fig. 5 Condition number of all Fourier submatrices of the DFT matrix of dimension 512, shown in base-10 logarithmic scale, with $p$ and $q$ varying between 1 and 512. The contour plot is after Barnett [1]. The surface plot shows a peak along the $p = q$ anti-diagonal (right panel).

- $O(\min(p, q) pq)$ when multiplying with the dense matrix $A$, and
- $O(\min(p, q) N \log N)$ when using the FFT for the matrix–vector products with $A$.

The computation of the singular values is the most expensive step in all cases and it determines the overall computational complexity.

**Remark 3** These formulas have been stated for the reduced SVD. The full SVD of $A$ can also be computed by computing $|p - q|$ additional orthogonal vectors from the full eigenvalue decomposition of $J(p, q)$ if $q > p$ or from that of $J(q, p)$ if $p > q$. This does not change the complexity, as there remain at most $\min(p, q)$ nonzero singular values.

**Remark 4** The methodology can also be adapted to compute a part of the spectrum, for example the singular values and vectors associated with the plunge region. We refer to [8, §4.1] for the details of this computation and for an estimate of the index range that corresponds to the plunge region. It is a topic of considerable interest that the size of the plunge region scales logarithmically with $N$ [3, Theorem 1]. For example, it determines the computational complexity of computing Fourier extension approximations using the AZ algorithm [2].

## 4 Illustrations and Applications

For the specific application in the Fourier extension approximation scheme we refer to [8, 9], and for a generic algorithm for solving linear systems involving DFT submatrices to [2]. Here, we illustrate the properties and review a few applications which are more directly related to the SVD of $A$.

### 4.1 The Spectra of the Jacobi Matrices $J(p, q)$ and $J(q, p)$

Some examples of the spectra of the matrices involved are shown in Fig. 1 for the case $N = 128$. The singular values of $A$ in panel (a) exhibit the familiar plateau and plunge regions, varying with $p$ and $q$. The eigenvalues of the tridiagonal matrices $J(p, q)$ and $J(q, p)$ are shown in panels (b) and (c), respectively. The values themselves are not relevant.
here as they are not used: what matters is that they are simple and, in comparison to the singular values of $A$, well-separated.

4.2 Exponential Ill-Conditioning of $A$

For the computation of the condition number of $A$, we have only to compute the smallest and largest singular values. In turn, we compute the smallest and largest eigenvalues of $J(p, q)$, making use of the fact that it is a tridiagonal matrix. This can be achieved with a computational cost proportional to $\max(p, q)$. This enables the computation of the condition number in linear time. This compares favourably to a computation based on the full SVD of $A$, which requires $O(pq^2)$ or $O(p^2q)$ operations.

We do note that efficiency is not the main computational problem here. Instead it is the exponential ill-conditioning of $A$. Following Remark 2, in order to compute small singular values we have to resort to high-precision arithmetic. This is the case for submatrices whose condition number is comparable to or exceeds $1/\varepsilon$, where $\varepsilon$ is floating-point precision.

The condition numbers of all submatrices of the DFT matrix of length $N = 512$ are shown in Fig. 5. Similar figures were shown in [1] for $N = 8$, $N = 16$ and $N = 32$. In our experiment we first ran all computations in standard double precision. All condition numbers estimated to be greater than $10^{13}$ have been recomputed using higher precision arithmetic. For $N = 512$ the largest condition number thus encountered is $1.2 \times 10^{127}$, for $p = q = 256$, necessitating arithmetic with about 150 digits. Calculating the full matrix is time-consuming mainly because it has many elements. Each individual matrix entry takes only a fraction of a second, even in high-precision.7

4.3 Growth Rate of the Condition Number of $A$

The growth rate of the condition number is analyzed in [1] as a function of the parameters $\alpha = p/N$ and $\beta = q/N$, in the square $(\alpha, \beta) \in (0, 1)^2$. These parameters are kept fixed as $N$ increases. The main results are lower bounds on the condition numbers of Fourier

7 Using the Julia package ‘FourierSubmatrices.jl’, the command \texttt{cond(DFTBlock}BigFloat\texttt{(512,256,256))} takes 0.05 seconds when using 512 bit precision on a contemporary laptop.
Comparison of the numerically estimated growth rate $\hat{\rho}(\alpha, \beta)$ with the known lower bound $\rho_1(\alpha, \beta)$, along the diagonal (top-left panel) and anti-diagonal (top-right panel) of the previous figure. The former also shows the second lower bound $\rho_2(\alpha, \beta)$ and the cross-over point $\alpha^* \approx 0.117$ of the bounds. While the cross-over point is relevant to determine the sharpest lower bound, it does not seem to play a role in the condition number itself. The bottom panels zoom in on the diagonal for small $\alpha$, showing growth rate and also the condition number for $N = 1500$ submatrices, which also gives rise to lower bounds on the exponential rate $\rho(\alpha, \beta)$. The main such bound results from [1, Theorem 2] and takes the form

$$ \kappa(A) = \Omega \left( N^{-3/2} e^{\rho_1(\alpha, \beta) N} \right), \quad N \to \infty, \quad (18) $$

with $\rho_1(\alpha, \beta) = \frac{\pi}{2} [\min(\alpha, \beta) - \alpha \beta]$. A second bound is shown to be sharper in the small-parameter corner $(\alpha, \beta) \in (0, \alpha^*)^2$, with $\alpha^* \approx 0.117$ the cross-over point. That bound, which is also valid in a larger square $(\alpha, \beta) \in (0, \frac{4}{e\pi})^2 \approx (0, 0.468)^2$ but not in all of $(0, 1)^2$, is

$$ \kappa(A) = \Omega \left( N^{-1/2} e^{\rho_2(\alpha, \beta) N} \right), \quad N \to \infty, \quad (19) $$

with $\rho_2(\alpha, \beta) = \min(\alpha, \beta) \log \frac{4}{e\pi \max(\alpha, \beta)}$. This bound follows from [1, Theorem 3].
The sharpness of these bounds was illustrated with limited numerical experiments, since only matrices had been considered with a condition number smaller than 10^{13}. With that restriction, estimates of \( \rho \) are practical only for rational numbers \( \alpha \) and \( \beta \) with small denominators. The methods of this paper allow a much finer resolution.

We numerically estimate the rate from a linear fit to the logarithms of the condition numbers, exactly as in [1], but with \( N \) chosen adaptively without constraint on the condition number. The result is shown in Fig. 8, in which samples of \( \alpha \) and \( \beta \) have been taken from a 60 \times 60 grid. Although Figs. 5 and 8 show different things, they are visually very similar. Both illustrate the dependence of the condition number on the size of the submatrix.

We show the sharpness of the lower bounds \( \rho_1 \) and \( \rho_2 \) in Fig. 7. Here, we first consider the diagonal \( \alpha = \beta \) and anti-diagonal \( \alpha = 1 - \beta \) using smaller steps of size 0.01. The figure confirms that the regime with \( \alpha \) and \( \beta \) fixed as \( N \) increases corresponds to exponential growth of the condition number. Yet, it also shows that there is some room to improve the lower bounds on the exponential growth factor, at least on these diagonals. In the bottom row we zoom in on the diagonal for small \( \alpha \) at a higher resolution with step 0.002. In the bottom right panel we show the corresponding condition numbers for \( N = 1500 \). The latter figure serves as a reminder that any gap between the actual growth rate and its lower bound results in an exponentially large gap between the condition number itself and its bound. The two bounds in the lower-right panel refer to the bounds as stated in Theorem 2 and Theorem 3, respectively, in [1].

These computational results suggest that a closer fit to the actual rate can be obtained in the form \( \rho(\alpha, \beta) \approx F(\min(\alpha, \beta), \max(\alpha, \beta)) \) with a first-degree polynomial

\[
F_1(a, b) = 2(a - ab),
\]
i.e., with \( \pi/2 \) replaced by 2 in \( \rho_1 \), or with the second degree polynomial

\[
F_2(a, b) = \pi a - 2\pi ab + \pi a^2b + \pi ab^2 - \pi a^2b^2.
\]

However, neither of these are lower bounds on \((0, 1)^2\). The fits were computed numerically and do not provide novel insights: they are merely better fits.
4.4 Low-Rank Relation Between DFTs of Different Length

The efficient evaluation of a DFT is usually based on recursive computation of DFTs of shorter length, the celebrated example being the FFT itself. The recursive relations are typically based on number-theoretic properties of $N$ in combination with algebraic properties of the transform.

Using low-rank Fourier submatrices, we can also establish an approximate numerical link between DFTs of different length. This is based on a Hadamard product, the element-wise product of matrices, which was used for the computation of non-uniform FFTs in [11]. We consider just one example, namely between the DFT of length $N$ and of length $N + 1$.

Consider the $N \times N$ submatrix $G$ of the DFT matrix of dimension $N + 1$, and compare to the DFT matrix $F$ of dimension $N$. Thus,

$$F_{j,k} = e^{-2\pi i \frac{(j-1)(k-1)}{N}} \quad \text{and} \quad G_{j,k} = e^{-2\pi i \frac{(j-1)(k-1)}{N+1}}.$$  

Using the Hadamard product notation $\circ$, we write the link as

$$F = G \circ H. \quad (20)$$

Elementwise, this corresponds to $F_{j,k} = G_{j,k} H_{j,k}$. From the explicit formulas of the matrices $G$ and $F$ above, we deduce that

$$H_{j,k} = e^{-2\pi i \frac{(j-1)(k-1)}{N(N+1)}}.$$  

This is essentially due to the identity

$$\frac{1}{N} = \frac{1}{N+1} + \frac{1}{N(N+1)}.$$  

Hence, $H$ is precisely the $N \times N$ submatrix of the much larger DFT matrix of dimension $N(N + 1)$.

The singular values of $H$ exhibit no plateau, since the general formula $\frac{pq}{\dim(F)}$ here becomes, with $p = q = N$ and $\dim(F) = N(N + 1)$,

$$\frac{N^2}{N(N + 1)} = \frac{N}{N + 1} < 1.$$  

There is only a plunge region with $\mathcal{O}(\log N)$ values larger than any given fixed threshold. From (20) we can also write

\[123\]

\[\text{Fig. 9} \quad \text{The singular values of } H \text{ in Eq. (20) for different values of } N. \text{ The matrices have no plateau, only a plunge region, which scales logarithmically with } N: H \text{ is low rank. These singular values where computed using the CTA algorithm.} \]
Fig. 10 Illustration of the separation property of the prolate sequence matrices $U$ and $V$ in the frequency domain. The matrices $U$ and $V$ are computed using the CTA algorithm (Alg. 1). Plotted are magnitudes $\log |F_p U(i, j)|$, and $\log |F_q V(i, j)|$ over the indices in the matrices $F_p U$ and $F_q V$ respectively. The figure shows that $F_p U$ and $F_q V$ each consist of three regions: purely low frequency, purely high-frequency and a transitory regime in-between. Here $N = 2000$, $p = 1000$, $q = 500$.

This makes $G$ a Butterfly Matrix (see [16]). The main computational technique of [11] is to expedite matrix–vector products of the form $(A \circ B)x$. This is shown to be possible at least if $A$ has a fast matrix–vector product and $B$ has low rank. If $B$ has rank $r$, then the matrix–vector product with $A \circ B$ requires $r$ matrix–vector products with $A$. That is exactly the setting of this subsection. However, this observation is mainly of mathematical interest. It is unlikely to result in any computational advantage compared to existing FFT methods, because even with $r = \mathcal{O} (\log N)$, the Hadamard matrix–vector product remains expensive in comparison.

4.5 Localization Properties of P-DPSS

Finally, we illustrate the time-frequency properties of the singular vectors, i.e., of the periodic discrete prolate spheroidal sequences. There are three regimes, associated with the three regimes of the singular values shown in Fig. 1a. The singular vectors corresponding to the larger singular values are approximately localized in the lower part of the spectrum, while those corresponding to small singular values have high-frequency content. The plunge region is associated with a transitory regime covering the full spectrum.

These properties are conveniently visualized simply by taking the column-wise FFT of the matrices $U$ and $V$ of length $p$ and $q$ respectively. This is shown in Fig. 10. The separation between low and high frequency is the ratio $p/N = 1/2$ in the left panel, and $q/N = 1/4$ in the right panel. In both panels, the left part of the figure corresponds to large singular values and has the exact opposite localization properties as the right part of the figure, which corresponds to small singular values. The plunge region appears as a small band in the middle covering the full frequency spectrum.
Acknowledgements The authors are grateful to Alex Barnett, Nick Trefethen and Thijjs Steel for their suggestions and discussions on the topic, to Emiel Dhondt for his contributions in the implementation, and to the anonymous reviewers for their constructive remarks.

Funding This work was supported in part by the FWO-Flanders (Project G088622N).

Data availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Barnett, A.H.: How exponentially ill-conditioned are contiguous submatrices of the Fourier matrix? SIAM Rev. 64, 105–131 (2022)
2. Coppé, V., Huybrechs, D., Matthysen, R., Webb, M.: The AZ algorithm for least squares problems with a known incomplete generalized inverse. SIAM J. Mat. Anal. Appl. 46, 1237–1259 (2020)
3. Edelman, A., McCorquodale, P., Toledo, S.: The future Fast Fourier transform? SIAM J. Sci. Comput. 20(3), 1094–1114 (1999)
4. Gentleman, W.M., Sande, G.: Fast Fourier transforms: for fun and profit. In: Proceedings of the November 7–10, 1966, Fall Joint Computer Conference, AFIPS ’66 (Fall), pp. 563–578. Association for Computing Machinery, New York (1966). https://doi.org/10.1145/1464291.1464352
5. Golub, G.H., van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press, Baltimore (1996)
6. Grünbaum, F.A.: Eigenvectors of a Toeplitz matrix: discrete version of the prolate spheroidal wave functions. SIAM J. Alg. Discrete Methods 2(2), 136–141 (1981)
7. Jain, A., Ranganath, S.: Extrapolation algorithms for discrete signals with application in spectral estimation. IEEE Trans. Acoust. Speech Signal Process. 29(4), 830–845 (1981)
8. Matthysen, R., Huybrechs, D.: Fast algorithms for the computation of Fourier extensions of arbitrary length. SIAM J. Sci. Comput. 38(2), A899–A922 (2016)
9. Matthysen, R., Huybrechs, D.: Function approximation on arbitrary domains using Fourier frames. SIAM J. Numer. Anal. 56, 1360–1385 (2018)
10. Osipov, A., Rokhlin, V., Xiao, H.: Prolate Spheroidal Wave Functions of Order Zero. Springer, New York (2013)
11. Ruiz-Antolín, D., Townsend, A.: A nonuniform fast Fourier transform based on low rank approximation. SIAM J. Sci. Comput. 40(1), A529–A547 (2018)
12. Slepian, D.: Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: the discrete case. Bell Syst. Tech. J. 57, 1371–1430 (1978)
13. Slepian, D., Pollak, H.: Prolate spheroidal wave functions, Fourier analysis, and uncertainty-I. Bell Syst. Tech. J. 40(1), 43–63 (1961)
14. Van Dooren, P., Laudadio, T., Mastronardi, N.: Computing the eigenvectors of nonsymmetric tridiagonal matrices. Comput. Math. Math. Phys. 61, 733–749 (2021)
15. Xu, W.Y., Chanzas, C.: On the periodic discrete prolate spheroidal sequences. SIAM J. Appl. Math. 44(6), 1210–1217 (1984)
16. Yang, H.: A unified framework for oscillatory integral transforms: when to use nufft or butterfly factorization? J. Comput. Phys. 388, 103–122 (2019)
17. Zhu, Z., Karnik, S., Davenport, M.A., Romberg, J., Wakin, M.B.: The eigenvalue distribution of discrete periodic time-frequency limiting operators. IEEE Signal Process. Lett. 25(1), 95–99 (2018). https://doi.org/10.1109/LSP.2017.2751578

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.