Improved characterization of the eigenvalue behavior of discrete prolate spheroidal sequences

Santhosh Karnik, Justin Romberg, Mark A. Davenport

June 2, 2020

Abstract

The discrete prolate spheroidal sequences (DPSSs) are a set of orthonormal sequences in $\ell_2(\mathbb{Z})$ which are strictly bandlimited to a frequency band $[-W, W]$ and maximally concentrated in a time interval $\{0, \ldots, N - 1\}$. The timelimited DPSSs (sometimes referred to as the Slepian basis) are an orthonormal set of vectors in $\mathbb{C}^N$ whose discrete time Fourier transform (DTFT) is maximally concentrated in a frequency band $[-W, W]$. Due to these properties, DPSSs have a wide variety of signal processing applications. The DPSSs are the eigensequences of a timelimit-then-bandlimit operator and the Slepian basis vectors are the eigenvectors of the so-called prolate matrix. The eigenvalues in both cases are the same, and they exhibit a particular clustering behavior – slightly fewer than $2NW$ eigenvalues are very close to $1$, slightly fewer than $N - 2NW$ eigenvalues are very close to $0$, and very few eigenvalues are not near $1$ or $0$. This eigenvalue behavior is critical in many of the applications in which DPSSs are used. There are many asymptotic characterizations of the number of eigenvalues not near $0$ or $1$. In contrast, there are very few non-asymptotic results, and these don’t fully characterize the clustering behavior of the DPSS eigenvalues. In this work, we establish two novel non-asymptotic bounds on the number of DPSS eigenvalues between $\epsilon$ and $1 - \epsilon$. Furthermore, we obtain bounds detailing how close the first $\approx 2NW$ eigenvalues are to $1$, how close the last $\approx N - 2NW$ eigenvalues are to $0$, as well as how small certain summations of these eigenvalues are. We also present simulations demonstrating the quality of these non-asymptotic bounds on the number of DPSS eigenvalues between $\epsilon$ and $1 - \epsilon$.

1 Introduction

A fundamental fact of Fourier analysis is that no non-zero signal can be simultaneously bandlimited and timelimited. Thus, a compactly supported non-zero function cannot have a compactly supported Fourier transform, and a non-zero function whose Fourier transform is compactly supported cannot itself be compactly supported. Between 1960 and 1978, Landau, Pollak, and Slepian published a series of seminal papers [1–5] exploring the problem of finding bandlimited signals which are maximally concentrated in a given time interval. They formulate this as an eigenproblem whose solutions are the prolate spheroidal wave functions (PSWFs) in the continuous case and the discrete prolate spheroidal sequences (DPSSs) in the discrete case.

By truncating the DPSSs, one can form the so-called Slepian basis vectors, which are an efficient basis for representing a window of samples from bandlimited signals [6–8]. As such, Slepian basis vectors can be used in a variety of applications. Some classic applications include prediction of bandlimited signals based on past samples [5] and Thomson’s multitaper method for spectral analysis [9]. More recent applications include time-variant channel estimation [10, 11], wideband compressive radio receivers [12], compressed sensing of analog signals [6], target detection [13, 14], and a fast method [15] for computing Fourier extension series coefficients [16, 17].
The eigenvalues associated with the DPSSs (which we will refer to in this paper as the DPSS eigenvalues) exhibit a particular clustering behavior. Most of the DPSS eigenvalues are very close to 1 or 0, and very few eigenvalues are not near 1 or 0. This clustering behavior plays a critical role in many applications. In [18], it is shown that the bias of Thomson’s multitaper spectral estimator depends on the sum of the leading DPSS eigenvalues. In [6], it is shown that the number of DPSS eigenvalues not near 1 determines the effective dimension of a vector of samples from a bandlimited signal. Also, the sum of the trailing DPSS eigenvalues bounds the error in approximating a vector of samples from a bandlimited signal by a linear combination of the leading DPSSs. In [8, 15, 19], the fact that only a small number of DPSS eigenvalues are not near 1 or 0 is exploited to perform fast computations. As such, the behavior of the DPSS eigenvalues is of considerable interest. There are many results (both asymptotic and non-asymptotic) regarding the eigenvalues associated with the PSWFs [1, 3, 20–23]. However, there are far fewer results regarding the DPSS eigenvalues, and the existing non-asymptotic results don’t fully capture the behavior of these eigenvalues.

The main contribution of this work is establishing novel non-asymptotic bounds on the number of DPSS eigenvalues which are not close to 1 or 0, as well as non-asymptotic bounds on the DPSS eigenvalues themselves. This work is organized as follows. In Section 2.1 we define DPSSs and Slepian basis vectors as well as review their basic properties. In Section 2.2, we will outline existing results on the DPSS eigenvalues. In Section 2.3, we state our new results. In Section 3, we prove the two novel non-asymptotic bounds on the number of DPSS eigenvalues that are not within \( \epsilon \) of 1 or 0. In Section 4, we use the two new bounds to derive non-asymptotic bounds on the DPSS eigenvalues themselves, as well as the sum of the leading and trailing DPSS eigenvalues. Finally, we conclude the paper in Section 5 with some numerical results to demonstrate the quality of our non-asymptotic bounds.

2 Overview of Main Results

2.1 Discrete prolate spheroidal sequences and Slepian basis vectors

We begin by defining the discrete prolate spheroidal sequences (DPSSs) and Slepian basis vectors and reviewing some of their key properties. For a discrete signal \( x \in \ell_2(\mathbb{Z}) \), we define its discrete time Fourier transform \( \hat{x} \in L_2([−\frac{1}{2}, \frac{1}{2}]) \) by

\[
\hat{x}(f) = \sum_{n=-\infty}^{\infty} x[n] e^{-j2\pi fn} \quad \text{for} \quad f \in [−\frac{1}{2}, \frac{1}{2}],
\]

where we use the notation \( i = \sqrt{-1} \). The inverse DTFT is given by

\[
x[n] = \int_{-1/2}^{1/2} \hat{x}(f) e^{j2\pi fn} \, df \quad \text{for} \quad n \in \mathbb{Z}.
\]

With these definitions, any \( x, y \in \ell_2(\mathbb{Z}) \) satisfy the Parseval-Plancherel identity \( \langle x, y \rangle_{\ell_2(\mathbb{Z})} = \langle \hat{x}, \hat{y} \rangle_{L_2([−1/2, 1/2])} \). For any \( N \in \mathbb{N} \), we say that \( x \in \ell_2(\mathbb{Z}) \) is timelimited to \( n \in \{0, \ldots, N−1\} \) if \( x[n] = 0 \) for \( n \in \mathbb{Z} \setminus \{0, \ldots, N−1\} \). Also, for any \( W \in (0, \frac{1}{2}) \), we say that \( x \in \ell_2(\mathbb{Z}) \) is bandlimited to \( |f| \leq W \) if \( \hat{x}(f) = 0 \) for \( |f| > W \).

We can now ask the question “what signal bandlimited to \( |f| \leq W \) has a maximum concentration of energy over the time indices \( n \in \{0, \ldots, N−1\} ? \)”, i.e.,

\[
\maximize_{x \in \ell_2(\mathbb{Z})} \sum_{n=0}^{N−1} |x[n]|^2 \quad \text{subject to} \quad ||x||_{\ell_2(\mathbb{Z})}^2 = 1 \quad \text{and} \quad \hat{x}(f) = 0 \text{ for } |f| > W.
\]

To help answer this question, we define two self-adjoint operators. For a given \( N \in \mathbb{N} \) we define a timelimiting operator \( T_N : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z}) \) by

\[
(T_N x)[n] = \begin{cases} 
  x[n] & \text{if } n \in \{0, \ldots, N−1\} \\
  0 & \text{for } n \in \mathbb{Z} \setminus \{0, \ldots, N−1\} 
\end{cases}
\]
For a given bandwidth parameter $W \in (0, \frac{1}{2})$, we define a bandlimiting operator $B_W : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ by

$$(B_W x)[n] = \sum_{m=-\infty}^{\infty} \frac{\sin[2\pi W(m-n)]}{\pi(m-n)} x[m] \quad \text{for} \quad n \in \mathbb{Z}.$$  

Note that the DTFT of $B_W x$ satisfies $\hat{B_W x}(f) = \hat{x}(f)$ for $|f| \leq W$ and $\hat{B_W x}(f) = 0$ for $|f| > W$.

For bandlimited signals $x \in \ell_2(\mathbb{Z})$, we can write

$$\sum_{n=0}^{N-1} |x[n]|^2 = \langle x, T_N x \rangle_{\ell_2(\mathbb{Z})} = \langle B_W x, T_N B_W x \rangle_{\ell_2(\mathbb{Z})} = \langle x, B_W T_N B_W x \rangle_{\ell_2(\mathbb{Z})}.$$  

Subject to the constraint $\|x\|^2_{\ell_2(\mathbb{Z})} = 1$, this is maximized by the eigensequence of $B_W T_N B_W$ corresponding to the largest eigenvalue. Slepian defined the discrete prolates spheroidal sequences (DPSSs) $s_0, \ldots, s_{N-1} \in \ell_2(\mathbb{Z})$ as the $N$ orthonormal eigensequences of $B_W T_N B_W$ corresponding to non-zero eigenvalues. The corresponding eigenvalues $1 > \lambda_0 > \lambda_1 > \cdots > \lambda_{N-1} > 0$ are referred to as the DPSS eigenvalues and are sorted in descending order. Slepian [5] showed that these eigenvalues are all distinct and strictly between 0 and 1. In addition to $s_0$ being the bandlimited sequence in $\ell_2(\mathbb{Z})$ with a maximal concentration of energy in $\{0, \ldots, N-1\}$, it is also true that for each $k = 1, \ldots, N-1$, $s_k$ is the bandlimited sequence in $\ell_2(\mathbb{Z})$ with a maximal concentration of energy in $\{0, \ldots, N-1\}$ subject to the additional constraint of being orthogonal to $s_0, \ldots, s_{k-1}$. Furthermore $\lambda_k$ is equal to the amount of energy $s_k$ has in the time interval $\{0, \ldots, N-1\}$.

We can also ask the question “what signal timelimited to $n \in \{0, \ldots, N-1\}$ has a maximum concentration of energy in the frequency band $|f| \leq W$?”, i.e.

$$\text{maximize} \quad \int_{-W}^{W} |\hat{x}(f)|^2 \, df \quad \text{subject to} \quad \|x\|^2_{\ell_2(\mathbb{Z})} = 1 \quad \text{and} \quad x[n] = 0 \text{ for } n \in \mathbb{Z} \setminus \{0, \ldots, N-1\}.$$  

For timelimited signals $x \in \ell_2(\mathbb{Z})$, we can write

$$\int_{-W}^{W} |\hat{x}(f)|^2 \, df = \langle \hat{x}, \hat{B_W x} \rangle_{\ell_2([-1/2,1/2])} = \langle x, B_W x \rangle_{\ell_2(\mathbb{Z})} = \langle T_N x, B_W T_N x \rangle_{\ell_2(\mathbb{Z})} = \langle x, T_N B_W T_N x \rangle_{\ell_2(\mathbb{Z})}.$$  

Since $T_N B_W T_N$ is self-adjoint, the sequence $x \in \ell_2(\mathbb{Z})$ which solves the above maximization problem is the eigensequence of the operator $T_N B_W T_N$ corresponding to the largest eigenvalue.

Clearly, the range of $T_N B_W T_N$ and the orthogonal complement of the kernel of $T_N B_W T_N$ is the $N$-dimensional space of timelimited signals. Hence, we can reduce this eigenproblem on $\ell_2(\mathbb{Z})$ to an eigenproblem on $\mathbb{R}^N$. With respect to the Euclidean basis for the space of timelimited signals, the matrix representation of $T_N B_W T_N$ is given by

$$B[m, n] = \frac{\sin[2\pi W(m-n)]}{\pi(m-n)} \quad \text{for} \quad m, n = 0, \ldots, N-1. \quad (1)$$

This matrix $B \in \mathbb{R}^{N \times N}$ is known in the literature as the prolate matrix [24, 25]. The Slepian basis vectors $s_0, \ldots, s_{N-1} \in \mathbb{R}^N$ are the orthonormal eigenvectors of $B$, where again the eigenvalues $1 > \lambda_0 > \lambda_1 > \cdots > \lambda_{N-1} > 0$ are sorted in descending order. Note that the eigenvalues of $B$ are the same as the eigenvalues of $T_N B_W T_N$, which are the same as the eigenvalues of $B_W T_N B_W$. Hence, we can reuse the notation $\lambda_k$ for $k = 0, \ldots, N-1$ to denote the eigenvalues of $B$. The eigensequences $\tilde{s}_0, \ldots, \tilde{s}_{N-1} \in \ell_2(\mathbb{Z})$ of $T_N B_W T_N$ are then given by $\tilde{s}_k[n] = s_k[n]$ for $n \in \{0, \ldots, N-1\}$ and $\tilde{s}_k[n] = 0$ for $n \in \mathbb{Z} \setminus \{0, \ldots, N-1\}$. Note that in addition to $\tilde{s}_0$ being the timelimited sequence in $\ell_2(\mathbb{Z})$ whose DTFT has a maximal concentration of energy in $[-W, W]$, it is also true that for each $k = 1, \ldots, N-1$, $\tilde{s}_k$ is a timelimited sequence in $\ell_2(\mathbb{Z})$ whose DTFT has a maximal concentration of energy in $[-W, W]$ subject to the additional constraint of being orthogonal to $\tilde{s}_0, \ldots, \tilde{s}_{k-1}$. Furthermore, the eigenvalue $\lambda_k$ is equal to the amount of energy that the DTFT of $\tilde{s}_k$ has in the frequency band $[-W, W]$.
Figure 1: A plot of the DPSS Eigenvalues for $N = 1000$ and $W = \frac{1}{5}$. These eigenvalues satisfy $\lambda_{243} \approx 0.9997$ and $\lambda_{256} \approx 0.0003$. Only 14 of the 1000 DPSS eigenvalues lie in $(0.001, 0.999)$.

### 2.2 DPSS eigenvalue concentration

Showing that the DPSS eigenvalues are strictly between 0 and 1 is a trivial consequence of the facts that $\lambda_k = \left(\int_{-W}^{W} |\hat{s}_k(f)|^2 \, df \right) / \left(\int_{-1/2}^{1/2} |\hat{s}_k(f)|^2 \, df \right)$ and that $\hat{s}_k(f)$ is a non-zero analytic function. It is also easy to check that the sum of all the DPSS eigenvalues is $\sum_{k=0}^{N-1} \lambda_k = \text{trace}(B) = 2NW$. What is perhaps more interesting is that the DPSS eigenvalues obey a particular clustering behavior. For any $\epsilon \in (0, \frac{1}{2})$, slightly fewer than $2NW$ eigenvalues lie in $[1 - \epsilon, 1)$, slightly fewer than $N - 2NW$ eigenvalues lie in $(0, \epsilon)$, and very few eigenvalues lie in the transition region $(\epsilon, 1 - \epsilon)$. In Figure 1, we demonstrate this phenomenon by plotting the DPSS eigenvalues for $N = 1000$ and $W = \frac{1}{5}$ (so $2NW = 250$). The first 243 eigenvalues lie in $[0, 0.999]$ and the last 743 eigenvalues lie in $(0, 0.001]$. Only 14 eigenvalues lie between $0.001$ and $0.999$.

Experimentally, we can see that the width of this transition region behaves like $\# \{k : \epsilon < \lambda_k < 1 - \epsilon \} = O(\log(NW) \log(\frac{1}{\epsilon}))$. This can be seen in Figures 2 and 3 in Section 5. Our main contribution will be to demonstrate this analytically, but before we do, we will briefly review some of the prior results along these lines.

We begin with the original results from Slepian [5]. For any fixed $W \in (0, \frac{1}{2})$ and $b \in \mathbb{R}$,

$$\lambda_{[2NW+(b/\pi) \log N]} \sim \frac{1}{1 + e^{b\pi}} \quad \text{as} \quad N \to \infty.$$  

From this result, it is easy to show that for any fixed $W \in (0, \frac{1}{2})$ and $\epsilon \in (0, \frac{1}{2})$,

$$\# \{k : \epsilon < \lambda_k < 1 - \epsilon \} \sim \frac{2}{\pi^2 \log N} \log \left(\frac{1}{\epsilon} - 1\right) \quad \text{as} \quad N \to \infty.$$  

This asymptotic bound on the width of the transition region correctly captures the logarithmic dependence on both $N$ and $\epsilon$, but not the dependence on $W$. Slepian also stated that if $0.2 < \lambda_k < 0.8$, then

$$\lambda_k \approx \left[1 + \exp \left(-\frac{\pi^2 (2NW - k - \frac{1}{2})}{\log[8N \sin(2\pi W)] + \gamma}\right)\right]^{-1}$$

is a good approximation to $\lambda_k$ where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant. This would suggest that

$$\# \{k : \epsilon < \lambda_k < 1 - \epsilon \} \approx \frac{2}{\pi^2} \log [8\epsilon N \sin(2\pi W)] \log \left(\frac{1}{\epsilon} - 1\right)$$

for $\epsilon \in (0.2, 0.5)$. This correctly captures the logarithmic dependence on $N$, $W$, and $\epsilon$, but only holds for large values of $\epsilon$.  

4
Very few papers provide non-asymptotic bounds regarding the width of the transition region \( \# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \). Zhu and Wakin [14] showed that

\[
\# \{ k : \epsilon \leq \lambda_k \leq 1 - \epsilon \} \leq \frac{2}{\epsilon} \log(N - 1) + \frac{2}{\epsilon} \frac{2N-1}{N-1} \frac{\log(NW)}{\epsilon(1-\epsilon)}
\]

for all integers \( N \geq 2 \), \( W \in (0, \frac{1}{2}) \), and \( \epsilon \in (0, \frac{1}{2}) \). This non-asymptotic bound correctly highlights the logarithmic dependence on \( N \), but fails to capture the dependence on \( W \). Also, the dependence on \( \epsilon \) is \( O\left(\frac{1}{\epsilon}\right) \) as opposed to \( O\left(\log\frac{1}{\epsilon}\right) \). When \( \epsilon \) is small, this bound is considerably worse than a \( O\left(\frac{1}{\epsilon}\right) \) bound. Furthermore, when \( \epsilon < \frac{2 \log(\frac{N-1}{2N})}{\pi^{2} NW} \), this bound is worse than the trivial bound of \( \# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \leq N \).

Recently, Boulsane, Bourguiba, and Karoui [26] improved this bound to

\[
\# \{ k : \epsilon \leq \lambda_k \leq 1 - \epsilon \} \leq \frac{1}{\pi^{2}} \log(2NW) + 0.45 - \frac{3}{4} W^{2} + \frac{\sin^{2}(2\pi NW)}{e^{\pi NW}} \epsilon(1-\epsilon)
\]

for all integers \( N \geq 1 \), \( W \in (0, \frac{1}{2}) \), and \( \epsilon \in (0, \frac{1}{2}) \). For a fixed \( W \in (0, \frac{1}{2}) \) and large \( N \), this bound is roughly half of (2). Also, this bound correctly captures the logarithmic dependence on \( 2NW \) as opposed to just \( N \). However, this bound still has a dependence on \( \epsilon \) that is \( O\left(\frac{1}{\epsilon}\right) \). Boulsane et. al. also proved that

\[
\lambda_k \leq 2 \exp\left[-\frac{k-2NW}{\log(\frac{\pi NW}{2})} + 5\right]
\]

for \( 2NW + \log(\pi NW) + 6 \leq k \leq \pi NW \).

where \( \eta = 0.069 \) is the constant specified in [21], and that

\[
\lambda_k \leq 2 \exp\left[-(2k+1) \log\left(\frac{2k+2}{\epsilon \pi NW}\right)\right]
\]

for \( 2 \leq \frac{\epsilon \pi}{NW} \leq k \leq N - 1 \).

However, with no similar lower bounds on the DPSS eigenvalues \( \lambda_k \) for \( k < 2NW \), they were unable to obtain a bound on \( \# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \) which has a logarithmic dependence on \( \epsilon \).

In [8], the authors of this paper along with Zhu and Wakin proved that

\[
\# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \leq \left(\frac{8}{\pi^{2}} \log(8N) + 12\right) \log\left(\frac{15}{\epsilon}\right)
\]

for all \( N \in \mathbb{N} \), \( W \in (0, \frac{1}{2}) \), and \( \epsilon \in (0, \frac{1}{2}) \). This bound correctly captures the logarithmic dependence on both \( N \) and \( \epsilon \), but not the dependence on \( W \). Also, the leading constant \( \frac{8}{\pi^{2}} \) is four times larger than that of the asymptotic results by Slepian.

### 2.3 Main results

In this paper, we will establish the following two non-asymptotic bounds on the number of DPSS eigenvalues in the transition region \((\epsilon, 1 - \epsilon)\).

**Theorem 1.** For any \( N \in \mathbb{N} \), \( W \in (0, \frac{1}{2}) \), and \( \epsilon \in (0, \frac{1}{2}) \),

\[
\# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \leq 2 \left[ \frac{1}{\pi^{2}} \log(4N) \log\left(\frac{4}{\epsilon(1-\epsilon)}\right) \right].
\]

**Theorem 2.** For any \( N \in \mathbb{N} \), \( W \in (0, \frac{1}{2}) \), and \( \epsilon \in (0, \frac{1}{2}) \),

\[
\# \{ k : \epsilon < \lambda_k < 1 - \epsilon \} \leq 2 \frac{2}{\pi^{2}} \log(100NW + 25) \log\left(\frac{5}{\epsilon(1-\epsilon)}\right) + 7.
\]

Both Theorem 1 and Theorem 2 capture the logarithmic dependence of the width of the transition region on \( N \) and \( \epsilon \). Also, both bounds have a leading constant of \( \frac{2}{\pi^{2}} \), which is consistent with the asymptotic result by Slepian. Furthermore, Theorem 2 also captures the logarithmic dependence on \( W \). We choose to include Theorem 1 since the proof is much simpler and since the bound in Theorem 1 is better than the bound in Theorem 2 when \( W \geq \frac{1}{25} \).

With the non-asymptotic bounds in Theorem 1 and Theorem 2, the following bounds on the eigenvalues themselves are almost immediate.
Corollary 1. For any $N \in \mathbb{N}$, $W \in (0, \frac{1}{2})$, we have

$$\lambda_k \geq 1 - \min \left\{ 8 \exp \left[ -\frac{1}{2} \left( 2NW \right) - k - 2 \right], 10 \exp \left[ -\frac{1}{2} \right] \sum_{k=0}^{N-1} \log(2NW + 25) \right\} \text{ for } 0 \leq k \leq \left[ 2NW \right] - 1$$

and

$$\lambda_k \leq \min \left\{ 8 \exp \left[ -\frac{1}{2} \left( 2NW \right) - k - 1 \right], 10 \exp \left[ -\frac{1}{2} \right] \sum_{k=0}^{N-1} \log(2NW + 25) \right\} \text{ for } \left[ 2NW \right] \leq k \leq N - 1.$$
3 Proof of Bounds on the Width of the Transition Region

3.1 Proof overview

For any rectangular matrix $X \in \mathbb{C}^{M \times N}$, we use the notation $\sigma_k(X)$ to denote the $k$th largest singular value of $X$. If $k > \min\{M, N\}$, we define $\sigma_k(X) = 0$. Also, for a Hermitian matrix $A \in \mathbb{C}^{N \times N}$, we use the notation $\mu_k(A)$ to denote $k$th largest eigenvalue of a symmetric matrix. Again, if $k > N$, we define $\mu_k(A) = 0$. To be consistent with standard notation, we define $\lambda_k = \mu_{k+1}(B)$ for $k \in \{0, \ldots, N - 1\}$, i.e. $\lambda_k$ is the $(k + 1)$th largest eigenvalue of the $N \times N$ prolate matrix $B$ with bandwidth parameter $W$, which is defined in (1). Although both $B$ and $\lambda_k$ depend on $N$ and $W$, our notation will omit this dependence for convenience.

Intuitively, we aim to show that $B - B^2 = X^* X$ for a “matrix” $X$ with $N$ columns and infinitely many rows, and then show that $X$ has a low numerical rank. Hence, $B - B^2$ also has a low numerical rank. Therefore, very few of the eigenvalues of $B - B^2$ are not near 0, and thus, very few of the eigenvalues of $B$ are not near 1 or 0. The following lemma allows us to start a rigorous version of this argument.

**Lemma 1.** Suppose for some $r \in \{0, \ldots, N - 1\}$ and $L_0 \in \mathbb{N}$, there exists a sequence of matrices $X_L \in \mathbb{R}^{2L \times N}$ for $L = L_0, L_0 + 1, \ldots$, such that:

- $\lim_{L \to \infty} \|B - B^2 - X_L^* X_L\|_F = 0$,
- $\sigma_{r+1}(X_L) \leq \sqrt{\epsilon(1 - \epsilon)}$ for all $L \geq L_0$.

Then, $\#\{k : \epsilon < \lambda_k < 1 - \epsilon\} \leq r$.

**Proof.** By the first property, $\lim_{L \to \infty} \mu_{r+1}(X_L^* X_L)$ exists and is equal to $\mu_{r+1}(B - B^2)$. Then by using the fact that $\mu_{r+1}(X_L^* X_L) = \sigma_{r+1}(X_L)^2$ for all $L \geq L_0$ along with the second property, we have

$$\mu_{r+1}(B - B^2) = \lim_{L \to \infty} \mu_{r+1}(X_L^* X_L) = \lim_{L \to \infty} \sigma_{r+1}(X_L)^2 \leq \epsilon(1 - \epsilon).$$

The eigenvalues of $B - B^2$ are $\{\lambda_k(1 - \lambda_k)\}_{k=0}^{N-1}$. Also, the function $\lambda \mapsto \lambda(1 - \lambda)$ is increasing for $\lambda < \frac{1}{2}$ and decreasing for $\lambda > \frac{1}{2}$ and symmetric about $\lambda = \frac{1}{2}$. As a result, $\epsilon < \lambda < 1 - \epsilon$ if and only if $\lambda(1 - \lambda) > \epsilon(1 - \epsilon)$. Therefore,

$$\#\{k : \epsilon < \lambda_k < 1 - \epsilon\} = \#\{k : \lambda_k(1 - \lambda_k) > \epsilon(1 - \epsilon)\} = \#\{k : \mu_k(B - B^2) > \epsilon(1 - \epsilon)\} \leq r.$$

We will prove both Theorem 1 and Theorem 2 by using Lemma 1. In Section 3.2, we construct a sequence of matrices $X_L \in \mathbb{R}^{2L \times N}$ which satisfies the first property above. In Section 3.3, we show that the singular values of each matrix $X_L$ decays exponentially, which allows us to obtain the bound in Theorem 1. In Section 3.4, we refine the rate at which the singular values of each $X_L$ decay, which allows us to obtain the bound in Theorem 2.

3.2 Constructing the sequence of matrices $X_L$

First, we begin by proving an identity involving the sinc function.

**Lemma 2.** For any $W \in (0, \frac{1}{2})$ and any $m, n \in \mathbb{Z}$,

$$\sum_{\ell=-\infty}^{\infty} \frac{\sin[2\pi W(\ell - m)]}{\pi(\ell - m)} \frac{\sin[2\pi W(\ell - n)]}{\pi(\ell - n)} = \frac{\sin[2\pi W(m - n)]}{\pi(m - n)}.$$

**Proof.** For each $m \in \mathbb{Z}$, define the shifted sinc sequence $x_m \in \ell_2(\mathbb{Z})$ by

$$x_m[\ell] = \frac{\sin[2\pi W(\ell - m)]}{\pi(\ell - m)} \text{ for } \ell \in \mathbb{Z},$$
The DTFT of the unshifted sinc sequence \( x_0(f) = 1_{[-W,W]}(f) \). So for any \( m \in \mathbb{Z} \), the DTFT of the shifted sinc sequence \( x_m \) is \( \hat{x}_m(f) = e^{-2\pi fm}1_{[-W,W]}(f) \). By using the Parseval-Plancherel identity, we obtain

\[
\sum_{\ell=-\infty}^{\infty} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)} = \sum_{\ell=-\infty}^{\infty} \frac{x_m[\ell]}{\pi} \frac{x_n[\ell]}{\pi} = \langle x_m, x_n \rangle_{L^2((-1/2,1/2])}
\]

By using the identity in Lemma 2 we can write the entries of \( B - B^2 \) as:

\[
(B^2)(m,n) = B[m,n] - \sum_{\ell=0}^{N-1} B[m,\ell] B[\ell,n] = \sum_{\ell=-\infty}^{\infty} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)}
\]

\[
= \sum_{\ell=-\infty}^{\infty} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)} - \sum_{\ell=0}^{N-1} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)}
\]

\[
= \sum_{\ell=-\infty}^{\infty} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)} + \sum_{\ell=N}^{\infty} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)}
\]

where the rearranging of terms is valid since the summands decay like \( O(|\ell|^{-2}) \) as \( \ell \to \pm\infty \), and thus, all the sums are absolutely convergent.

For each integer \( L \geq 1 \), we define an index set

\[
\mathcal{I}_L = \{-L, -L+1, \ldots, -2, -1\} \cup \{N, N+1, \ldots, N+L-2, N+L-1\}
\]

and we define \( X_L \in \mathbb{R}^{2L \times N} \) by

\[
X_L[\ell,n] = \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)} \text{ for } \ell \in \mathcal{I}_L \text{ and } n \in \{0, \ldots, N-1\}.
\]

Note that we index the rows of \( X_L \) by \( \mathcal{I}_L \) instead of the usual 0, 1, \ldots, \( 2L-1 \) for convenience. We will also index the rows and/or columns of other matrices with dimension \( 2L \) by \( \mathcal{I}_L \). With this definition, the entries of \( X_L^* X_L \) are

\[
(X_L^* X_L)[m,n] = \sum_{\ell \in \mathcal{I}_L} X_L[\ell,m] X_L[\ell,n]
\]

\[
= \sum_{\ell=-L}^{-1} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)} + \sum_{\ell=N}^{N+L-1} \frac{\sin[2\pi W(\ell-m)]}{\pi(\ell-m)} \frac{\sin[2\pi W(\ell-n)]}{\pi(\ell-n)}
\]

From Equations 7 and 8 above, \( \lim_{L \to \infty} (X_L^* X_L)[m,n] = (B - B^2)[m,n] \) for each of the \( N^2 \) entries. Therefore,

\[
\lim_{L \to \infty} \| (B - B^2) - X_L^* X_L \|_F^2 = \lim_{L \to \infty} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left| (B - B^2)[m,n] - (X_L^* X_L)[m,n] \right|^2 = 0.
\]

This shows that the sequence of matrices \( X_L \in \mathbb{R}^{2L \times N} \) satisfies the first property of Lemma 1. We will now focus on bounding the singular values of each \( X_L \) in order to prove that these matrices \( X_L \) satisfy the second property of Lemma 1.
3.3 Proof of Theorem 1

In [27], Beckermann and Townsend showed that matrices with a low-rank displacement have rapidly decaying singular values provided the spectra of the left and right displacement matrices are well separated. More specifically, suppose a matrix $X \in \mathbb{C}^{M \times N}$ satisfies the displacement equation

$$CX - XD = UV^*$$

where $C \in \mathbb{C}^{M \times M}$ and $D \in \mathbb{C}^{N \times N}$ are normal matrices and $U \in \mathbb{C}^{M \times \nu}$ and $V \in \mathbb{C}^{N \times \nu}$. If there are closed, disjoint subsets $E$, $F$ of $\mathbb{C}$ such that $\text{Spec}(C) \subset E$ and $\text{Spec}(D) \subset F$ then the singular values of $X$ satisfy

$$\sigma_{\nu k+1}(X) \leq \sigma_1(X) Z_k(E, F)$$

for all integers $k \geq 0$, where $Z_k(E, F)$ are the Zolotarev numbers [28] for the sets $E$ and $F$. As a rule of thumb, when $E$ and $F$ are “well-separated”, $Z_k(E, F)$ decays exponentially with $k$. For more details about Zolotarev numbers, see Appendix A.

In Appendix B, we build on the work of Beckermann and Townsend to prove the following theorem.

**Theorem 3.** Suppose $X \in \mathbb{C}^{M \times N}$ satisfies the displacement equation

$$CX - XD = UV^*$$

where $C \in \mathbb{C}^{M \times M}$ and $D \in \mathbb{C}^{N \times N}$ are normal matrices and $U \in \mathbb{C}^{M \times \nu}$ and $V \in \mathbb{C}^{N \times \nu}$. If $\text{Spec}(C) \subset (-\infty, c_1] \cup [c_2, \infty)$ and $\text{Spec}(D) \subset [d_1, d_2]$ where $c_1 < d_1 < d_2 < c_2$, then for any integer $k \geq 0$,

$$\sigma_{\nu k+1}(X) \leq 4\|X\| \exp \left[ -\frac{\pi^2 k}{\log(16\gamma)} \right] \text{ where } \gamma = \frac{(c_2 - d_1)(d_2 - c_1)}{(c_2 - d_2)(d_1 - c_1)}.$$

We now show that the matrices $X_L$ defined in Section 3.2 satisfy a low-rank displacement equation, and use Theorem 3 to bound their singular values. Define a diagonal matrix $D \in \mathbb{R}^{N \times N}$ by $D[n, n] = n$ for $n \in \{0, \ldots, N-1\}$. For each integer $L \geq 1$, define a diagonal matrix $C_L \in \mathbb{R}^{2L \times 2L}$ by $C_L[\ell, \ell] = \ell$ for $\ell \in \mathcal{I}_L$ (again, we index $C_L$ by $\ell \in \mathcal{I}_L$ for convenience). With this definition, we have

$$(C_LX_L - X_LD)[\ell, n] = C_L[\ell, \ell]X_L[\ell, n] - X_L[\ell, n]D[n, n]$$

$$= \ell \cdot \frac{\sin[2\pi W(\ell - n)]}{\pi(\ell - n)} - \frac{\sin[2\pi W(\ell - n)]}{\pi(\ell - n)} \cdot n$$

$$= \frac{1}{\pi} \sin[2\pi W(\ell - n)]$$

$$= \frac{1}{\pi} \left[ \sin(2\pi W\ell) \cos(2\pi Wn) - \cos(2\pi W\ell) \sin(2\pi Wn) \right].$$

From this, it is clear that we can factor

$$C_LX_L - X_LD = U_LV^*$$

where $U_L \in \mathbb{R}^{2L \times 2}$ is defined by

$$U_L[\ell, 0] = \frac{1}{\sqrt{\pi}} \sin(2\pi W\ell) \quad \text{and} \quad U_L[\ell, 1] = \frac{1}{\sqrt{\pi}} \cos(2\pi W\ell) \quad \text{for} \quad \ell \in \mathcal{I}_L,$$

and $V \in \mathbb{R}^{N \times 2}$ is defined by

$$V[n, 0] = \frac{1}{\sqrt{\pi}} \cos(2\pi Wn) \quad \text{and} \quad V[n, 1] = -\frac{1}{\sqrt{\pi}} \sin(2\pi Wn) \quad \text{for} \quad n \in \{0, \ldots, N-1\}.$$

In other words, $X_L$ has a rank-2 displacement with respect to the matrices $C_L$ and $D$.  

9
Since \( \text{Spec}(C_L) = \mathcal{I}_L \subset (-\infty, -1) \cup [N, \infty) \) and \( \text{Spec}(D) = \{0, \ldots, N - 1\} \subset [0, N - 1] \), we can apply Theorem 3 with parameters \( c_1 = -1, d_1 = 0, d_2 = N - 1, c_2 = N, \) and \( \nu = 2 \). The theorem tells us that for every integer \( k \geq 0 \),

\[
\sigma_{2k+1}(X_L) \leq 4\|X_L\| \exp \left[ -\frac{\pi^2 k}{\log(16\gamma)} \right],
\]

where \( \gamma = \frac{(c_2 - d_1)(d_2 - c_1)}{(c_2 - d_2)(d_1 - c_1)} = N^2 \).

For any \( L \geq 1 \), \( X_L \) is a submatrix of \( X_{L+1} \), and so, \( X_L^*X_L \preceq X_{L+1}^*X_{L+1} \). Hence, \( X_L^*X_L \preceq \lim_{L \to \infty} X_L^*X_L = B - B^2 \). Therefore,

\[
\|X_L\|^2 = \|X_L^*X_L\| \leq \|B - B^2\| = \max_k [\lambda_k - \lambda_k^2] \leq \max_{0 \leq \lambda \leq 1} [\lambda - \lambda^2] = \frac{1}{4},
\]

and thus, \( \|X_L\| \leq \frac{1}{2} \) for all \( L \geq 1 \). Substituting \( \gamma = N^2 \) and \( \|X_L\| \leq \frac{1}{2} \) into the above bound yields

\[
\sigma_{2k+1}(X_L) \leq 2 \exp \left[ -\frac{\pi^2 k}{\log(16N^2)} \right],
\]

for all integers \( k \geq 0 \). So, if we set

\[
k = \left[ \frac{1}{\pi^2} \log(16N^2) \log \left( \frac{2}{\sqrt{\epsilon(1 - \epsilon)}} \right) \right] = \left[ \frac{1}{\pi^2} \log(4N) \log \left( \frac{4}{\epsilon(1 - \epsilon)} \right) \right],
\]

we obtain \( \sigma_{2k+1}(X_L) \leq \sqrt{\epsilon(1 - \epsilon)} \) for all \( L \geq 1 \).

This proves the second property in Lemma 1 for \( r = 2k \) and \( L_0 = 1 \). Therefore, we have proved that

\[
\#\{k : \epsilon < \lambda_k < 1 - \epsilon\} \leq 2k = 2 \left[ \frac{1}{\pi^2} \log(4N) \log \left( \frac{4}{\epsilon(1 - \epsilon)} \right) \right],
\]

which is exactly the content of Theorem 1.

### 3.4 Proof of Theorem 2

First, note that if \( W \in [\frac{1}{4}, \frac{1}{2}] \), the bound in Theorem 2 is greater than the bound in Theorem 1, which has already been established. So we will henceforth assume that \( W \in (0, \frac{1}{4}) \).

Now, set \( L_1 = \left\lfloor \frac{1}{W} \right\rfloor \) (clearly, \( L_1 \geq 1 \)). For each integer \( L \geq L_1 + 1 \), we partition the index set

\[\mathcal{I}_L = \{-L, -L + 1, \ldots, -2, -1\} \cup \{N, N + 1, \ldots, N + L - 2, N + L - 1\}\]

into three sets

\[\mathcal{I}_L^{(0)} = \{-L, -L + 1, \ldots, -L - 2, -L - 1\} \cup \{N + L_1, N + L_1 + 1, \ldots, N + L - 2, N + L - 1\}\]

\[\mathcal{I}_L^{(1)} = \{-L_1, -L_1 + 1, \ldots, -2, -1\}\]

\[\mathcal{I}_L^{(2)} = \{N, N + 1, \ldots, N + L_1 - 2, N + L_1 - 1\}\]

and then accordingly partition \( X_L \) into three submatrices \( X_L^{(0)} \in \mathbb{R}^{2(L-L_1) \times N}, X_L^{(1)} \in \mathbb{R}^{L_1 \times N}, \) and \( X_L^{(2)} \in \mathbb{R}^{L_1 \times N} \) defined by

\[X_L^{(i)}[\ell, n] = X_L[\ell, n] \quad \text{for} \quad \ell \in \mathcal{I}_L^{(i)} \text{ and } n \in \{0, \ldots, N - 1\}.
\]

Once again, we index the rows of each \( X_L^{(i)} \) by \( \ell \in \mathcal{I}_L^{(i)} \) for convenience. We proceed to bound the singular values of \( X_L^{(0)}, X_L^{(1)}, X_L^{(2)} \), and then use these bounds to bound the singular values of \( X_L \).
Singular values of $X_L^{(0)}$

The submatrix $X_L^{(0)}$, has the same low-rank displacement structure as $X_L$. Specifically, we can write

$$C_L^{(0)}X_L^{(0)} - X_L^{(0)}D = U_L^{(0)}V^*$$

where $C_L^{(0)} \in \mathbb{R}^{2(L-L_1) \times 2(L-L_1)}$ is the diagonal submatrix of $C_L$ defined by $C_L^{(0)}[\ell, \ell] = \ell$ for $\ell \in \mathcal{I}_L^{(0)}$, $U_L^{(0)} \in \mathbb{R}^{2(L-L_1) \times 2}$ is the submatrix of $U_L$ defined by $U_L^{(0)}[\ell, q] = U_L[\ell, q]$ for $\ell \in \mathcal{I}_L^{(0)}$ and $q \in \{0,1\}$, and $D$ and $V$ are the same as defined in Section 3.3.

Since $\text{Spec}(C_L^{(0)}) = \mathcal{I}_L^{(0)} \subset (-\infty, -L_1 - 1] \cup [N + L_1, \infty)$ and $\text{Spec}(D) = \{0, \ldots, N-1\} \subset [0, N-1]$, we can once again apply Theorem 3, but with the parameters $c_1 = -L_1 - 1$, $d_1 = 0$, $d_2 = N - 1$, $c_2 = N + L_1$, and $\nu = 2$. Then, the theorem tells us that for any integer $k_0 \geq 0$,

$$\sigma_{2k_0+1}(X_L^{(0)}) \leq 4\|X_L^{(0)}\| \exp \left[ -\frac{\pi^2 k_0}{\log(16\gamma)} \right] \quad \text{where} \quad \gamma = \frac{(c_2 - d_1)(d_2 - c_1)}{(c_2 - d_2)(d_1 - c_1)} = \left( \frac{N + L_1}{L_1 + 1} \right)^2.$$  

Since $X_L^{(0)}$ is a submatrix of $X_L$, we have $\|X_L^{(0)}\| \leq \|X_L\| \leq \frac{1}{2}$. Also, since $L_1 = \left\lfloor \frac{1}{3W} \right\rfloor \geq \frac{1}{3W} - 1 > 0$ and $\frac{N+\frac{1}{2}}{x+1}$ is a non-increasing function of $x > 0$, we can bound

$$\gamma = \left( \frac{N + L_1}{L_1 + 1} \right)^2 \leq \left( \frac{N + \left( \frac{1}{3W} - 1 \right)}{\left( \frac{1}{3W} - 1 \right) + 1} \right)^2 = (4NW + 1 - 4W)^2 \leq (4NW + 1)^2.$$  

Substituting $\gamma \leq (4NW + 1)^2$ and $\|X_L^{(0)}\| \leq \frac{1}{2}$ into the above bound yields

$$\sigma_{2k_0+1}(X_L^{(0)}) \leq 2 \exp \left[ -\frac{\pi^2 k_0}{\log(16(4NW + 1)^2)} \right] = 2 \exp \left[ -\frac{\pi^2 k_0}{2\log(16NW + 4)} \right]$$

for all integers $k_0 \geq 0$.

Singular values of $X_L^{(1)}$

To bound the singular values of $X_L^{(1)}$, we exploit the fact that its entries $X_L^{(1)}[\ell, n] = \frac{\sin[2\pi W(\ell - n)]}{\pi(\ell - n)}$ are a smooth function of $\ell$ and $n$ to construct a tunable low-rank approximation of $X_L^{(1)}$.

Define the sinc function $g(t) = \frac{\sin(2\pi Wt)}{\pi t}$. For each $n = 0, \ldots, N - 1$, define

$$g_n(t) = g(t - n) = \frac{\sin[2\pi W(t - n)]}{\pi(t - n)},$$

and let

$$P_{k,n}(t) = \sum_{m=0}^{k-1} p_{m,n} t^m$$

be the degree $k-1$ Chebyshev interpolating polynomial for $g_n(t)$ on the interval $[-L_1, -1]$.

We now define the low rank approximation $\widetilde{X}_L^{(1)} \in \mathbb{R}^{L_1 \times N}$ by

$$\widetilde{X}_L^{(1)}[\ell, n] = P_{k,n}(\ell) = \sum_{m=0}^{k-1} p_{m,n} \ell^m \quad \text{for} \quad \ell \in \{-L_1, \ldots, -1\} \text{ and } n \in \{0, \ldots, N - 1\}.$$  

We can factor $\widetilde{X}_L^{(1)} = WP$ where $W \in \mathbb{R}^{L_1 \times k}$ and $P \in \mathbb{R}^{k \times N}$ are defined by $W[\ell, m] = \ell^m$ and $P[m, n] = p_{m,n}$. Hence, rank($\widetilde{X}_L^{(1)}$) $\leq k$.  

11
By Theorem 7 (in Appendix C), the Chebyshev interpolating polynomial satisfies
\[ |g_n(t) - P_{k,n}(t)| \leq \frac{(L_1 - 1)^k}{2^{2k-1}k!} \max_{\xi \in [-L_1, -1]} |g_n^{(k)}(\xi)| \quad \text{for all } t \in [-L_1, -1]. \]

Also, by Lemma 5 (in Appendix C), the derivatives of the unshifted sinc-function \( g(t) \) can be bounded by
\[ |g^{(k)}(t)| \leq (2\pi W)^k \min \left\{ \frac{2W}{k+1}, \frac{2}{\pi|t|} \right\} \quad \text{for all } t \in \mathbb{R}. \]

Hence, for any \( \ell \in \mathcal{I}_L^{(1)} \) and \( n \in \{0, \ldots, N - 1\} \), we have
\[
\left| X_L^{(1)}[\ell, n] - \tilde{X}_L^{(1)}[\ell, n] \right| = |g_n(\ell) - P_{k,n}(\ell)| \\
\leq \frac{(L_1 - 1)^k}{2^{2k-1}k!} \max_{\xi \in [-L_1, -1]} |g_n^{(k)}(\xi)| \\
= \frac{(L_1 - 1)^k}{2^{2k-1}k!} \max_{\xi \in [-L_1, -1]} |g^{(k)}(\xi - n)| \\
= \frac{(L_1 - 1)^k}{2^{2k-1}k!} \max_{t \in [-L_1-n, -n-1]} |g^{(k)}(t)| \\
\leq \frac{(L_1 - 1)^k}{2^{2k-1}k!} \max_{t \in [-L_1-n, -n-1]} (2\pi W)^k \min \left\{ \frac{2W}{k+1}, \frac{2}{\pi|t|} \right\} \\
= \frac{(L_1 - 1)^k}{2^{2k-1}k!} (2\pi W)^k \min \left\{ \frac{2W}{k+1}, \frac{2}{\pi(n+1)} \right\} \\
= \frac{4(W(L_1 - 1))^k}{k!} \min \left\{ \frac{W}{k+1}, \frac{1}{\pi(n+1)} \right\}.
\]

We proceed to bound the Frobenius norm of \( X_L^{(1)} - \tilde{X}_L^{(1)} \). Set \( N_1 = \left\lfloor \frac{k+1}{\pi W} \right\rfloor \). Then,
\[
\left\| X_L^{(1)} - \tilde{X}_L^{(1)} \right\|_F^2 = \sum_{n=0}^{N_1-1} \sum_{\ell = -L_1}^{-1} \left| X_L^{(1)}[\ell, n] - \tilde{X}_L^{(1)}[\ell, n] \right|^2 \\
\leq \sum_{n=0}^{N_1-1} \sum_{\ell = -L_1}^{-1} \frac{16(W(L_1 - 1))^{2k}}{(k!)^2} \min \left\{ \frac{W^2}{(k+1)^2}, \frac{1}{\pi^2(n+1)^2} \right\} \\
= \sum_{n=0}^{N_1-1} 16L_1(W(L_1 - 1))^{2k} \min \left\{ \frac{W^2}{(k+1)^2}, \frac{1}{\pi^2(n+1)^2} \right\} \\
\leq \sum_{n=0}^{N_1} 16L_1(W(L_1 - 1))^{2k} \min \left\{ \frac{W^2}{(k+1)^2}, \frac{1}{\pi^2(n+1)^2} \right\} \\
\leq \frac{16L_1(W(L_1 - 1))^{2k}}{(k!)^2} \left[ \sum_{n=0}^{N_1-1} \frac{W^2}{(k+1)^2} + \sum_{n=N_1}^{\infty} \frac{1}{\pi^2(n+1)^2} \right] \\
\leq \frac{16L_1(W(L_1 - 1))^{2k}}{(k!)^2} \left[ \frac{W^2N_1}{(k+1)^2} + \frac{1}{\pi^2N_1} \right],
\]
where the last line follows from the bound \( \sum_{n=1}^{\infty} \frac{1}{(n+1)^2} \leq \frac{1}{N_1} \).
We proceed to weaken this result to obtain a more usable upper bound as follows:

\[
\left\| X^{(1)}_L - \widetilde{X}^{(1)}_L \right\|_F^2 \leq \frac{16L_1(\pi W(L_1 - 1))^{2k}}{(k!)^2} \left[ \frac{W^2 N_1}{(k+1)^2} + \frac{1}{\pi^2 N_1} \right] \\
\leq \frac{16(L_1 - 1)(\pi WL_1)^{2k}}{(k!)^2} \left[ \frac{W^2 N_1}{(k+1)^2} + \frac{1}{\pi^2 N_1} \right] \\
= \frac{16(\pi W L_1)^{2k}}{(k!)^2} \left[ \frac{W^2 N_1 L_1}{(k+1)^2} + \frac{L_1}{\pi^2 (N_1 + 1)} \right] \\
\leq \frac{16(\pi W \cdot \frac{1}{2\pi})^{2k}}{(k!)^2} \left[ \frac{W^2 \cdot \frac{k+1}{\pi} \frac{1}{2\pi}}{(k+1)^2} + \frac{1}{\pi^2 (N_1 + 1)} \right] \\
\leq \frac{8}{\pi(k+1)(k!)^2} \left( \frac{\pi}{8} \right)^{2k} \leq \frac{5600}{\pi} \left( \frac{\pi}{48} \right)^{2k} .
\]

The 2nd line follows from the inequalities \((L_1 - 1)^{2k} \leq L_1^{2k}\) and \((L_1 - 1)^{2k} \leq (L_1 - 1)L_1^{2k}\). The 4th line holds since \(\frac{L_1 - 1}{N_1} \leq \frac{L_1}{N_1 + 1}\) is equivalent to \(L_1 \leq N_1 + 1\), which is true since \(L_1 \leq \frac{1}{\pi W} \leq \frac{1}{4W} \leq \left\lfloor \frac{k+1}{\pi W} \right\rfloor + 1 = N_1 + 1\). The last line holds due to the fact that \((k+1)(k!)^2 \geq \frac{1}{1600} 6^{2k}\) for all integers \(k \geq 0\).

Since \(\text{rank}(\widetilde{X}^{(1)}_L) \leq k\), we have \(\sigma_{k+1}(\widetilde{X}^{(1)}_L) = 0\). Hence, we can bound

\[
\sigma_{k+1}(X^{(1)}_L) = \left| \sigma_{k+1}(X^{(1)}_L) - \sigma_{k+1}(\widetilde{X}^{(1)}_L) \right| \leq \left| X^{(1)}_L - \widetilde{X}^{(1)}_L \right| \leq \left\| X^{(1)}_L - \widetilde{X}^{(1)}_L \right\|_F \leq \sqrt{\frac{5600}{\pi} \left( \frac{\pi}{48} \right)^k}.
\]

**Singular values of** \(X^{(2)}_L\)

We can exploit the symmetry between \(X^{(2)}_L\) and \(X^{(1)}_L\) to show that the singular values of \(\widetilde{X}^{(2)}_L\) are the same as those of \(\widetilde{X}^{(1)}_L\). Specifically, for any indices \(\ell \in \mathcal{I}^{(2)}_L\) and \(n \in \{0, \ldots, N - 1\}\), we have that \(N - 1 - \ell \in \mathcal{I}^{(1)}_L\) and \(N - 1 - n \in \{0, \ldots, N - 1\}\), and that

\[
X^{(2)}_L[\ell, n] = \frac{\sin[2\pi W(\ell - n)\pi]}{\pi(\ell - n)} = \frac{\sin[2\pi W((N - 1 - \ell) - (N - 1 - n))]}{\pi((N - 1 - \ell) - (N - 1 - n))} = X^{(1)}_L[N - 1 - \ell, N - 1 - n].
\]

Since the singular values of a matrix are invariant under permutations of rows/columns, we have

\[
\sigma_{k+1}(X^{(2)}_L) = \sigma_{k+1}(X^{(1)}_L) \leq \sqrt{\frac{5600}{\pi} \left( \frac{\pi}{48} \right)^k}
\]

for all integers \(k \geq 0\).

**Singular values of** \(X_L\)

Due to the way we partitioned \(X_L\) into three submatrices, we have

\[
X_L^* X_L = X_L^{(0)*} X_L^{(0)} + X_L^{(1)*} X_L^{(1)} + X_L^{(2)*} X_L^{(2)}.
\]

By the Weyl eigenvalue inequalities, we have

\[
\mu_{2k_0 + 2k+1}(X_L^* X_L) \leq \mu_{2k_0 + 1}(X_L^{(0)*} X_L^{(0)}) + \mu_{k+1}(X_L^{(1)*} X_L^{(1)}) + \mu_{k+1}(X_L^{(2)*} X_L^{(2)}).
\]
Hence, we can bound
\[
\sigma_{2k_0+2k+1}(X_L)^2 = \mu_{2k_0+2k+1}(X_L X_L)
\]
\[
\leq \mu_{2k_0+1}(X_L^0 X_L^0) + \mu_{k+1}(X_L^{(1)} X_L^{(1)}) + \mu_{k+1}(X_L^{(2)} X_L^{(2)})
\]
\[
= \sigma_{2k_0+1}(X_L^0)^2 + \sigma_{k+1}(X_L^{(1)})^2 + \sigma_{k+1}(X_L^{(2)})^2
\]
\[
\leq 4 \exp \left[ -\frac{\pi^2 k_0}{\log(16NW + 4)} \right] + \frac{5600}{\pi} \left( \frac{\pi}{48} \right)^{2k} + \frac{5600}{\pi} \left( \frac{\pi}{48} \right)^{2k}
\]
\[
= 4 \exp \left[ -\frac{\pi^2 k_0}{\log(16NW + 4)} \right] + \frac{11200}{\pi} \left( \frac{\pi}{48} \right)^{2k}
\]
for any integers \(k_0 \geq 0\) and \(k \geq 1\).

If we set
\[
k_0 = \left\lfloor \frac{1}{\pi^2} \log(16NW + 4) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) \right\rfloor
\]
and
\[
k = \left\lfloor \frac{1}{2 \log(\frac{48}{\pi})} \log \left( \frac{56000}{\pi} \frac{\epsilon}{\epsilon(1-\epsilon)} \right) \right\rfloor
\]
then we obtain
\[
\sigma_{2k_0+2k+1}(X_L)^2 \leq 4 \exp \left[ -\frac{\pi^2 k_0}{\log(16NW + 4)} \right] + \frac{11200}{\pi} \left( \frac{\pi}{48} \right)^{2k}
\]
\[
\leq \frac{4\epsilon(1-\epsilon)}{5} + \frac{\epsilon(1-\epsilon)}{5}
\]
\[
= \epsilon(1-\epsilon),
\]
i.e., \(\sigma_{2k_0+2k+1}(X_L) \leq \sqrt{\epsilon(1-\epsilon)}\). Our steps hold for all \(L \geq L_1 + 1\).

This proves the second property of Lemma 1 for \(r = 2k_0 + 2k\) and \(L_0 = L_1 + 1\). Therefore,
\[
\# \{k : \epsilon < \lambda_k < 1 - \epsilon\} \leq 2k_0 + 2k = 2 \left\lfloor \frac{1}{\pi^2} \log(16NW + 4) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) \right\rfloor + 2 \left\lfloor \frac{1}{2 \log(\frac{48}{\pi})} \log \left( \frac{56000}{\pi} \frac{\epsilon}{\epsilon(1-\epsilon)} \right) \right\rfloor.
\]

We can loosen this bound to make it more "user friendly" as follows:
\[
\# \{k : \epsilon < \lambda_k < 1 - \epsilon\} \leq 2 \left\lfloor \frac{1}{\pi^2} \log(16NW + 4) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) \right\rfloor + 2 \left\lfloor \frac{1}{2 \log(\frac{48}{\pi})} \log \left( \frac{56000}{\pi} \frac{\epsilon}{\epsilon(1-\epsilon)} \right) \right\rfloor
\]
\[
\leq \frac{2}{\pi^2} \log(16NW + 4) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + \frac{1}{\log(\frac{48}{\pi})} \log \left( \frac{56000}{\pi} \frac{\epsilon}{\epsilon(1-\epsilon)} \right) + 4
\]
\[
= \frac{2}{\pi^2} \log(16NW + 4) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + \frac{1}{\log(\frac{48}{\pi})} \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + \frac{\log(\frac{11200}{\pi} \log(\frac{48}{\pi}))}{\log(\frac{48}{\pi})} + 4
\]
\[
= \left( \frac{2}{\pi^2} \log(16NW + 4) + \frac{1}{\log(\frac{48}{\pi})} \right) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + \frac{\log(\frac{11200}{\pi} \log(\frac{48}{\pi}))}{\log(\frac{48}{\pi})} + 4
\]
\[
= \frac{2}{\pi^2} \log \left( \exp \left( \frac{\pi^2}{2 \log(\frac{48}{\pi})} \right) (16NW + 4) \right) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + \frac{\log(\frac{11200}{\pi} \log(\frac{48}{\pi}))}{\log(\frac{48}{\pi})} + 4
\]
\[
\leq \frac{2}{\pi^2} \log(100NW + 25) \log \left( \frac{5}{\epsilon(1-\epsilon)} \right) + 7,
\]
which establishes theorem 2.
4 Proof of Eigenvalue Bounds (Corollaries 1 and 2)

First, we state a result from [7] which bounds \( \lambda_k \) for two values of \( k \) near \( 2NW \).

**Lemma 3.** For any \( N \in \mathbb{N} \) and \( W \in (0, \frac{1}{2}) \),

\[
\lambda_{[2NW]-1} \geq \frac{1}{2} \geq \lambda_{[2NW]}.
\]

To derive bounds on \( \lambda_k \), we will set \( \epsilon \) such that the transition region is too narrow to contain \( k \), and thus conclude either \( \lambda_k \geq 1 - \epsilon \) (if \( k \leq [2NW] - 1 \)) or \( \lambda_k \leq \epsilon \) (if \( k \geq [2NW] \)). To derive bounds on \( \sum_{k=0}^{K-1} (1 - \lambda_k) \) and \( \sum_{k=K}^{N-1} \lambda_k \), we will simply apply the bounds on \( \lambda_k \) and the formula for the sum of a geometric series.

4.1 Lower bounds on \( \lambda_k \) for \( k \leq [2NW] - 1 \)

For any integer \( k \) such that \( 0 \leq k \leq [2NW] - 1 \), set

\[
\epsilon = 8 \exp \left( -\frac{[2NW] - k - 2}{2 \pi^2 \log(4N)} \right),
\]

and suppose for sake of contradiction that \( \lambda_k < 1 - \epsilon \).

By using the assumption \( k \leq [2NW] - 1 \) and Lemma 3, we have \( \frac{1}{2} \leq \lambda_{[2NW]-1} \leq \lambda_k < 1 - \epsilon \), i.e., \( \epsilon < \frac{1}{2} \). Therefore, \( \epsilon < \frac{1}{2} \leq \lambda_{[2NW]-1} \leq \lambda_k < 1 - \epsilon \), i.e. both \( k \) and \( [2NW] - 1 \) are in the transition region \( \{k' : \epsilon < \lambda_{k'} < 1 - \epsilon\} \), and thus, so are all the indices \( k' \) between \( k \) and \( [2NW] - 1 \). Hence,

\[
\# \{k' : \epsilon < \lambda_{k'} < 1 - \epsilon\} \geq \# \{k : k \leq [2NW] - 1\} = [2NW] - k.
\]

However, since \( \epsilon < \frac{1}{2} \), by Theorem 1 we have

\[
\# \{k' : \epsilon < \lambda_{k'} < 1 - \epsilon\} \leq 2 \left[ \frac{1}{\pi^2} \log(4N) \log \left( \frac{4}{\epsilon(1 - \epsilon)} \right) \right] < \frac{2}{\pi^2} \log(4N) \log \left( \frac{8}{\epsilon} \right) + 2 = [2NW] - k.
\]

This is a contradiction. Therefore,

\[
\lambda_k \geq 1 - \epsilon = 1 - 8 \exp \left( -\frac{[2NW] - k - 2}{2 \pi^2 \log(4N)} \right) \quad \text{for} \quad 0 \leq k \leq [2NW] - 1.
\]

In a similar manner, we can assume \( \lambda_k < 1 - \epsilon \), where

\[
\epsilon = 10 \exp \left( -\frac{[2NW] - k - 7}{2 \pi^2 \log(100NW + 25)} \right),
\]

and then invoke theorem 2 to obtain a contradiction. Therefore,

\[
\lambda_k \geq 1 - \epsilon = 1 - 10 \exp \left( -\frac{[2NW] - k - 7}{2 \pi^2 \log(100NW + 25)} \right) \quad \text{for} \quad 0 \leq k \leq [2NW] - 1.
\]

Combining these two bounds establishes the first part of Corollary 1.

4.2 Upper bounds on \( \lambda_k \) for \( k \geq [2NW] \)

For any integer \( k \) such that \( [2NW] \leq k \leq N - 1 \), set

\[
\epsilon = 8 \exp \left( -\frac{k - [2NW] - 1}{2 \pi^2 \log(4N)} \right),
\]

and suppose for sake of contradiction that \( \lambda_k > \epsilon \).
By using the assumption $k \geq \lceil 2NW \rceil$ and Lemma 3, we have $\epsilon < \lambda_k \leq \lambda_{\lceil 2NW \rceil} \leq \frac{1}{2}$, i.e., $\epsilon < \frac{1}{2}$.
Therefore, $\epsilon < \lambda_k \leq \lambda_{\lceil 2NW \rceil} \leq \frac{1}{2} < 1 - \epsilon$, i.e., both $k$ and $\lceil 2NW \rceil$ are in the transition region $\{ k' : \epsilon < \lambda_{k'} < 1 - \epsilon \}$, and thus, so are all the indices $k'$ between $k$ and $\lceil 2NW \rceil$. Hence,

$$\# \{ k' : \epsilon < \lambda_{k'} < 1 - \epsilon \} \geq \# \{ k' : \lceil 2NW \rceil \leq k' \leq k \} = k - \lceil 2NW \rceil + 1.$$ 

However, since $\epsilon < \frac{1}{2}$, by Theorem 1 we have

$$\# \{ k' : \epsilon < \lambda_{k'} < 1 - \epsilon \} \leq 2 \left[ \frac{1}{\pi^2} \log(4N) \log \left( \frac{4}{\epsilon(1 - \epsilon)} \right) \right] < 2 \frac{1}{\pi^2} \log(4N) \log \left( \frac{8}{\epsilon} \right) + 2 = k - \lceil 2NW \rceil + 1.$$ 

This is a contradiction. Therefore,

$$\lambda_k \leq \epsilon = 8 \exp \left[ - \frac{k - \lceil 2NW \rceil - 1}{2 \pi^2 \log(4N)} \right] \quad \text{for} \quad \lceil 2NW \rceil \leq k \leq N - 1.$$ 

In a similar manner, we can assume $\lambda_k > \epsilon$, where

$$\epsilon = 10 \exp \left[ - \frac{k - \lceil 2NW \rceil - 6}{2 \pi^2 \log(100NW + 25)} \right],$$

and then invoke Theorem 2 to obtain a contradiction. Therefore,

$$\lambda_k \leq \epsilon = 10 \exp \left[ - \frac{k - \lceil 2NW \rceil - 6}{2 \pi^2 \log(100NW + 25)} \right] \quad \text{for} \quad \lceil 2NW \rceil \leq k \leq N - 1.$$ 

Combining these two bounds establishes the second part of Corollary 1.

### 4.3 Bounds on $\sum_{k=0}^{K-1} (1 - \lambda_k)$ for $K \leq \lceil 2NW \rceil$

For any integer $K$ such that $1 \leq K \leq \lceil 2NW \rceil$, we can apply the first part of the lower bound for $\lambda_k$ in Corollary 1 along with the inequality $\frac{e^{-x}}{1 - e^{-x}} \leq \frac{1}{x}$ for $x > 0$ to obtain

$$\sum_{k=0}^{K-1} (1 - \lambda_k) \leq \sum_{k=0}^{K-1} 8 \exp \left[ - \frac{\lceil 2NW \rceil - k - 2}{2 \pi^2 \log(4N)} \right] \leq \sum_{k=0}^{K-1} 8 \exp \left[ - \frac{\lceil 2NW \rceil - k - 2}{2 \pi^2 \log(4N)} \right] \leq 16 \log(4N) \exp \left[ - \frac{\lceil 2NW \rceil - k - 2}{2 \pi^2 \log(4N)} \right].$$

In a similar manner, we can apply the second part of the lower bound for $\lambda_k$ in Corollary 1 instead of
the first part of the lower bound to obtain

\[
\sum_{k=0}^{K-1} (1 - \lambda_k) \leq \sum_{k=0}^{K-1} 10 \exp \left[-\frac{\lfloor 2NW \rfloor - k - 7}{2\pi \log(100NW + 25)}\right]
\]

\[
\leq \sum_{k=-\infty}^{K-1} 8 \exp \left[-\frac{\lfloor 2NW \rfloor - K - 6}{2\pi \log(100NW + 25)}\right]
\]

\[
= \frac{1 - \exp \left[-\frac{1}{2\pi \log(100NW + 25)}\right]}{1 - \exp \left[-\frac{2\pi \log(100NW + 25)}{2\pi \log(100NW + 25)}\right]}
\]

\[
\leq \frac{20}{\pi^2} \log(100NW + 25) \exp \left[-\frac{\lfloor 2NW \rfloor - K - 6}{2\pi \log(100NW + 25)}\right].
\]

Combining these two bounds establishes the first part of Corollary 2.

### 4.4 Bounds on \(\sum_{k=K}^{N-1} \lambda_k\) for \(K \geq \lceil 2NW \rceil\)

For any integer \(K\) such that \(\lceil 2NW \rceil \leq K \leq N - 1\), we can apply the first part of the upper bound for \(\lambda_k\) in Corollary 1 along with the inequality \(\frac{e^{1-x}}{1-e^{-x}} \leq \frac{1}{x}\) for \(x > 0\) to obtain

\[
\sum_{k=K}^{N-1} \lambda_k \leq \sum_{k=K}^{N-1} 8 \exp \left[-\frac{k - \lfloor 2NW \rfloor - 1}{2\pi \log(4N)}\right]
\]

\[
\leq \sum_{k=K}^{\infty} 8 \exp \left[-\frac{k - \lfloor 2NW \rfloor - 1}{2\pi \log(4N)}\right]
\]

\[
= \frac{1 - \exp \left[-\frac{1}{2\pi \log(4N)}\right]}{1 - \exp \left[-\frac{2\pi \log(4N)}{2\pi \log(4N)}\right]}
\]

\[
\leq \frac{16}{\pi^2} \log(4N) \exp \left[-\frac{K - \lfloor 2NW \rfloor - 2}{2\pi \log(4N)}\right].
\]

In a similar manner, we can apply the second part of the upper bound for \(\lambda_k\) in Corollary 1 instead of the first part of the upper bound to obtain

\[
\sum_{k=K}^{N-1} \lambda_k \leq \sum_{k=K}^{N-1} 10 \exp \left[-\frac{k - \lfloor 2NW \rfloor - 6}{2\pi \log(100NW + 25)}\right]
\]

\[
\leq \sum_{k=K}^{\infty} 10 \exp \left[-\frac{k - \lfloor 2NW \rfloor - 6}{2\pi \log(100NW + 25)}\right]
\]

\[
= \frac{10 \exp \left[-\frac{K - \lfloor 2NW \rfloor - 6}{2\pi \log(100NW + 25)}\right]}{1 - \exp \left[-\frac{2\pi \log(100NW + 25)}{2\pi \log(100NW + 25)}\right]}
\]

\[
\leq \frac{20}{\pi^2} \log(100NW + 25) \exp \left[-\frac{K - \lfloor 2NW \rfloor - 7}{2\pi \log(100NW + 25)}\right].
\]

Combining these two bounds establishes the second part of Corollary 2.
5 Numerical Results

We demonstrate the quality of our bounds on the width of the transition region \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) with some numerical computations. First, we fix \(W = \frac{1}{4}\) (a large value of \(W\)), and for each integer \(2^4 \leq N \leq 2^{16}\) we use the method described in [29] to compute \(\lambda_k\) for a range \(k_{\text{min}} \leq k \leq k_{\text{max}}\) such that \(\lambda_{k_{\text{min}}} > 1 - 10^{-13}\) and \(\lambda_{k_{\text{max}}} < 10^{-13}\). From this, we can determine \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) as well as the upper bound on \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) from Theorem 1 in Figure 2. We note that over this range of parameters, the difference between the bound in Theorem 1 and the true width of the transition region \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) is between 1 and 14.

![Figure 2: Plots of the width of the transition region \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) vs. \(N\) where \(W = \frac{1}{4}\) and \(\epsilon = 10^{-3}\) (blue), \(\epsilon = 10^{-8}\) (green), and \(10^{-13}\) (red) are fixed. The dashed lines indicate the upper bound from Theorem 1.](image)

Next, we fix \(N = 2^{16}\) and for 10001 logarithmically spaced values of \(W\) between \(2^{-14}\) and \(2^{-2}\), we use the method described in [29] to compute \(\lambda_k\) for a range \(k_{\text{min}} \leq k \leq k_{\text{max}}\) such that \(\lambda_{k_{\text{min}}} > 1 - 10^{-13}\) and \(\lambda_{k_{\text{max}}} < 10^{-13}\). From this, we can determine \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) as well as the upper bound on \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) from Theorem 2 in Figure 3. We note that over this range of parameters, the difference between the bound in Theorem 2 and the true width of the transition region \(#\{k : \epsilon < \lambda_k < 1 - \epsilon\}\) is between \(\approx 9.8\) and \(\approx 30.7\).

In Figure 2, we see that the plots of both \(\frac{1}{2} \log(4N) \log\left(\frac{1}{\epsilon(1-\epsilon)}\right)\) (the bound in Theorem 1) and the actual width of the transition region increase roughly linearly with \(\log N\) and at roughly the same rate. However, the difference between the bound in Theorem 1 and the actual width of the transition region is noticeably larger for smaller values of \(\epsilon\) than for larger values of \(\epsilon\). This provides numerical evidence that for a large bandwidth \(W\), the bound’s dependence on \(N\) is close to correct, but the dependence on \(\epsilon\) has some room for improvement.

In Figure 3, we see that the plots of both \(\frac{1}{2} \log(100NW + 25) \log\left(\frac{5}{\epsilon(1-\epsilon)}\right) + 7\) and the actual width of the transition region increase roughly linearly with \(\log(NW)\) and at roughly the same rate. However, the
difference between the bound in Theorem 2 and the actual width of the transition region is quite noticeable. This provides numerical evidence that the leading constant of $\frac{8}{\pi^2}$ is indeed correct, but that the other constants leave significant room for improvement.

Finally, we note that for the range of parameters in both plots, the non-asymptotic bounds on the width of the transition region given by (2), (3), and (6) (in Section 2.2) would all be well above the range of the plots in Figures 2 and 3. The bounds in (2) and (3) are proportional to $\frac{1}{\epsilon(1-\epsilon)}$. Thus, they are only useful when $\epsilon$ is not too small. Also, the bound in (6) is rather large since the leading constant $\frac{8}{\pi^2}$ being 4 times larger than that in Theorems 1 and 2, and the trailing constant $12 \log(8N)$ when $N$ isn’t too large. In particular, for $\epsilon = 10^{-3}$ and any $N \in \mathbb{N}$, if we impose the mild constraint that $NW \geq \frac{1}{2}$, then the bound in (2) is at least $\frac{4/\pi^2}{\epsilon(1-\epsilon)} \approx 405$, the bound in (3) is at least $\frac{0.45-1/6}{\epsilon(1-\epsilon)} \approx 283$, and the bound in (6) is at least $(\frac{8}{\pi^2} \log(8) + 12) \log(\frac{15}{\epsilon}) \approx 131$.

References

[1] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. I. *Bell Systems Tech. J.*, 40(1):43–64, 1961.

[2] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. II. *Bell Systems Tech. J.*, 40(1):65–84, 1961.

[3] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. III. *Bell Systems Tech. J.*, 41(4):1295–1336, 1962.

[4] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. IV. *Bell Systems Tech. J.*, 43(6):3009–3058, 1964.

[5] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V – The discrete case. *Bell Systems Tech. J.*, 57(5):1371–1430, 1978.

[6] M. A. Davenport and M. B. Wakin. Compressive sensing of analog signals using discrete prolate spheroidal sequences. *Applied and Computational Harmonic Analysis*, 33(3):438–472, 2012.
[7] Z. Zhu and M. B. Wakin. Approximating sampled sinusoids and multiband signals using multiband modulated dpss dictionaries. *Journal of Fourier Analysis and Applications*, 23(6):1263–1310, 2017.

[8] S. Karnik, Z. Zhu, M. B. Wakin, J. Romberg, and M. A. Davenport. The fast slepine transform. *Applied and Computational Harmonic Analysis*, 46(3):624–652, 2019.

[9] D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, 1982.

[10] T. Zemen and C. Mecklenbräuker. Time-variant channel estimation using Discrete Prolate Spheroidal Sequences. *IEEE Trans. Signal Processing*, 53(9):3597–3607, 2005.

[11] T. Zemen, C. F. Mecklenbrauker, F. Kaltenberger, and B. H. Fleury. Minimum-energy band-limited predictor with dynamic subspace selection for time-variant flat-fading channels. *IEEE Transactions on Signal Processing*, 55(9):4534–4548, 2007.

[12] M. A. Davenport, S. R. Schnelle, J. P. Slavinsky, R. G. Baraniuk, M. B. Wakin, and P. T. Boufounos. A wideband compressive radio receiver. In *2010-MILCOM 2010 Military Communications Conference*, pages 1193–1198. IEEE, 2010.

[13] F. Yin, C. Debes, and A. M. Zoubir. Parametric waveform design using discrete prolate spheroidal sequences for enhanced detection of extended targets. *IEEE Transactions on Signal Processing*, 60(9):4525–4536, 2012.

[14] Z. Zhu and M. B. Wakin. Wall clutter mitigation and target detection using Discrete Prolate Spheroidal Sequences. In *Proc. Int. Work. on Compressed Sensing Theory Appl. Radar, Sonar and Remote Sens. (CoSeRa)*, Pisa, Italy, Jun. 2015.

[15] R. Matthysen and D. Huybrechs. Fast algorithms for the computation of fourier extensions of arbitrary length. *SIAM Journal on Scientific Computing*, 38(2):A899–A922, 2016.

[16] D. Huybrechs. On the Fourier extension of nonperiodic functions. *SIAM J. Numer. Anal.*, 47(6):4326–4355, 2010.

[17] B. Adcock, D. Huybrechs, and J. Martín-Vaquero. On the numerical stability of fourier extensions. *Foundations of Computational Mathematics*, 14(4):635–687, 2014.

[18] L. D. Abreu and J. L. Romero. Mse estimates for multitaper spectral estimation and off-grid compressive sensing. *IEEE Transactions on Information Theory*, 63(12):7770–7776, 2017.

[19] S. Karnik, J. Romberg, and M. A. Davenport. Fast multitaper spectral estimation. In 2019 *International Conference on Sampling Theory and Applications (SampTA)*. IEEE, 2019.

[20] A. Bonami and A. Karoui. Spectral decay of time and frequency limiting operator. *Applied and Computational Harmonic Analysis*, 42(1):1–20, 2017.

[21] A. Bonami, P. Jaming, and A. Karoui. Non-asymptotic behaviour of the spectrum of the sinc kernel operator and related applications. *arXiv preprint arXiv:1804.01257*, 2018.

[22] J. A. Hogan and J. D. Lakey. Duration and bandwidth limiting: prolate functions, sampling, and applications. Springer Science & Business Media, 2011.

[23] A. Israel. The eigenvalue distribution of time-frequency localization operators. *arXiv preprint arXiv:1502.04404*, 2015.

[24] J. M. Varah. The prolate matrix. *Linear Algebra and its Applications*, 187:269 – 278, 1993.

[25] A. W. Bojanczyk, R. P. Brent, F. R. De Hoog, and D. R. Sweet. On the stability of the bareiss and related toeplitz factorization algorithms. *SIAM J. Matrix Analysis and Applications*, 16:40–57, 1995.
A Zolotarev numbers

In this section, we review some properties of Zolotarev numbers, which will be useful in our analysis in Appendix B. With the exception of Corollary 4, all the results here have been proven elsewhere. However, we state these results and outline the proofs for sake of completeness.

For any integer \( k \geq 0 \), we let \( \mathcal{R}_{k,k} \) denote the set of rational functions \( \varphi(z) = \frac{p(z)}{q(z)} \) such that \( p(z) \) and \( q(z) \) are polynomials with degree at most \( k \). For any two disjoint, closed subsets of the Riemann sphere \( E, F \subset \mathbb{C} \cup \{\infty\} \), the Zolotarev number \( Z_k(E, F) \) is defined as

\[
Z_k(E, F) = \inf_{\varphi \in \mathcal{R}_{k,k}} \sup_{z \in E} |\varphi(z)| \quad \inf_{z \in F} |\varphi(z)|.
\]

Note that any rational function \( \varphi(z) = \frac{p(z)}{q(z)} \) can be extended to a continuous function on the Riemann sphere \( \mathbb{C} \cup \{\infty\} \) by defining \( \varphi(\infty) = \lim_{|z| \to \infty} \varphi(z) \) and \( \varphi(z) = \infty \) for any \( z \) such that \( q(z) = 0 \).

Beckermann and Townsend [27] proved the following bound on the Zolotarev numbers for the intervals \( E = [-b, -a] \) and \( F = [a, b] \).

**Theorem 4.** [27] For any reals \( b > a > 0 \), and any integer \( k \geq 0 \),

\[
Z_k([-b, -a], [a, b]) \leq 4 \exp \left[ -\frac{\pi^2 k}{\log(4b/a)} \right].
\]

The proof of Theorem 4 involves using theory of elliptic functions to construct a rational function \( \varphi \in \mathcal{R}_{k,k} \) for which

\[
\sup_{z \in [-b, -a]} |\varphi(z)| \quad \inf_{z \in [a, b]} |\varphi(z)| \leq 4 \exp \left[ -\frac{\pi^2 k}{\log(4b/a)} \right].
\]

A fact about Zolotarev numbers is that they are invariant under invertible Möbius transforms [30].

**Lemma 4.** For any two disjoint, closed subsets of the Riemann sphere \( E, F \subset \mathbb{C} \cup \{\infty\} \) and any Möbius transform \( \phi(z) = \frac{\beta_1 z + \beta_2}{\beta_3 z + \beta_4} \) such that \( \beta_1 \beta_4 \neq \beta_2 \beta_3 \), we have \( Z_k(\phi(E), \phi(F)) = Z_k(E, F) \) for all integers \( k \geq 0 \).

This fact is easily proved by noting that \( \varphi^* \in \mathcal{R}_{k,k} \) is the extremal rational function for \( (\phi(E), \phi(F)) \) if and only if \( \varphi^* \circ \phi \in \mathcal{R}_{k,k} \) is the extremal rational function for \( (E, F) \).

Using this fact, Beckermann and Townsend proved the following bound on the Zolotarev numbers for two non-overlapping intervals.
Corollary 3. [27] For any two intervals $[c_1, c_2]$ and $[d_1, d_2]$ that are nonoverlapping, and any integer $k \geq 0$,
\[
Z_k([c_1, c_2], [d_1, d_2]) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(16\gamma)} \right) \quad \text{where} \quad \gamma = \frac{(d_1 - c_1)(d_2 - c_2)}{(d_2 - c_1)(d_1 - c_2)}.
\]

Proof. It is trivial to check that $\gamma > 1$ when $[c_1, c_2]$ and $[d_1, d_2]$ do not overlap. Now, set $\alpha = 2\gamma - 1 + 2\sqrt{\gamma^2 - \gamma}$ and define the Möbius transforms
\[
\phi_1(z) = \frac{(d_2 - d_1)(z - c_2)}{(d_2 - c_2)(z - d_1)} \quad \text{and} \quad \phi_2(z) = \frac{(\alpha - 1)(z + 1)}{(\alpha + 1)(z - 1)}.
\]

One can check that $\phi_1([c_1, c_2]) = [0, \frac{\gamma - 1}{\gamma}] = [0, \frac{(\alpha - 1)^2}{(\alpha + 1)^2}] = \phi_2([-\alpha, -1])$ and $\phi_1([d_1, d_2]) = [1, \infty] = \phi_2([1, \alpha])$, and that both $\phi_1$ and $\phi_2$ are bijections. Thus, the Möbius transform $\phi = \phi_2^{-1} \circ \phi_1$ satisfies $\phi([c_1, c_2]) = [-\alpha, -1]$ and $\phi([d_1, d_2]) = [1, \alpha]$. So by applying Theorem 4, Lemma 4, and the bound $\alpha = 2\gamma - 1 + 2\sqrt{\gamma^2 - \gamma} \leq 4\gamma$, we have
\[
Z_k([c_1, c_2], [d_1, d_2]) = Z_k([-\alpha, -1], [1, \alpha]) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(4\alpha)} \right) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(16\gamma)} \right).
\]

\[
\square
\]

In a nearly identical manner, we can also prove the following bound.

Corollary 4. For any real numbers $c_1 < d_1 < d_2 < c_2$, and any integer $k \geq 0$,
\[
Z_k([-\infty, c_1] \cup [c_2, \infty], [d_1, d_2]) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(16\gamma)} \right) \quad \text{where} \quad \gamma = \frac{(c_2 - d_1)(d_2 - c_1)}{(c_2 - d_2)(d_1 - c_1)}.
\]

Proof. Again, since $c_1 < d_1 < d_2 < c_2$, we have $\gamma > 1$. Now, set $\alpha = 2\gamma - 1 + 2\sqrt{\gamma^2 - \gamma}$ and define the Möbius transforms
\[
\phi_1(z) = \frac{(d_2 - d_1)(z - c_1)}{(d_2 - c_1)(z - d_1)} \quad \text{and} \quad \phi_2(z) = \frac{(\alpha - 1)(z + 1)}{(\alpha + 1)(z - 1)}.
\]

One can check that $\phi_1([-\infty, c_1] \cup [c_2, \infty]) = [0, \frac{\gamma - 1}{\gamma}] = [0, \frac{(\alpha - 1)^2}{(\alpha + 1)^2}] = \phi_2([-\alpha, -1])$ and $\phi_1([d_1, d_2]) = [1, \infty] = \phi_2([1, \alpha])$, and that both $\phi_1$ and $\phi_2$ are bijections. Thus, the Möbius transform $\phi = \phi_2^{-1} \circ \phi_1$ satisfies $\phi([-\infty, c_1] \cup [c_2, \infty]) = [-\alpha, -1]$ and $\phi([d_1, d_2]) = [1, \alpha]$. So by applying Theorem 4, Lemma 4, and the bound $\alpha = 2\gamma - 1 + 2\sqrt{\gamma^2 - \gamma} \leq 4\gamma$, we have
\[
Z_k([-\infty, c_1] \cup [c_2, \infty], [d_1, d_2]) = Z_k([-\alpha, -1], [1, \alpha]) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(4\alpha)} \right) \leq 4 \exp \left( -\frac{\pi^2 k}{\log(16\gamma)} \right).
\]

\[
\square
\]

B  Singular values of matrices with low rank displacement

With the exception of Theorem 3, the results in this section have all been proven elsewhere. Furthermore, the proof of Theorem 3 is very similar to that of Theorem 6. However, we state these results and the proof of Theorem 3 for sake of completeness.

Throughout this section, we suppose that $X \in \mathbb{C}^{M \times N}$ satisfies the displacement equation
\[
CX - XD = UV^*;
\]
where $C \in \mathbb{C}^{M \times M}$ and $D \in \mathbb{C}^{N \times N}$ are Hermitian matrices, and $U \in \mathbb{C}^{M \times \nu}$ and $V \in \mathbb{C}^{N \times \nu}$ (where it is understood that $\nu \ll \min\{M, N\}$ for the results in this section to be useful). Our goal is to show that $X$ is approximately low-rank under certain assumptions on $\text{Spec}(C)$ and $\text{Spec}(D)$.

Beckermann and Townsend [27] showed that the numerical rank of $X$ can be bounded in terms of Zolotarev numbers.
Theorem 5. [27] If Spec(C) ⊂ E and Spec(D) ⊂ F, then the singular values of X satisfy
\[ \sigma_{v,k+j}(X) \leq \sigma_j(X)Z_k(E,F) \]
for any integers \( j \geq 1, k \geq 0 \).

The proof involves showing that for any rational function \( \varphi \in \mathcal{R}_{k,k} \), we can construct a rank-(\( \nu k + j - 1 \)) matrix \( Y \) such that
\[ X - Y = \varphi(C)(X - X_{j-1})\varphi(D)^{-1} \]
where \( X_{j-1} \) is the best rank-(\( j - 1 \)) approximation to \( X \). Then, by applying the facts that
\[ \| \varphi(C) \| \leq \sup_{z \in E} | \varphi(z) |, \quad \| \varphi(D)^{-1} \| \leq \sup_{z \in F} | \varphi(z) |^{-1} = \left( \inf_{z \in F} | \varphi(z) | \right)^{-1}, \quad \text{and} \quad \| X - X_{j-1} \| = \sigma_j(X) \]
along with the submultiplicativity of the matrix norm, we obtain
\[ \sigma_{\nu k+j}(X) \leq \| X - Y \| \leq \| \varphi(C) \| \cdot \| X - X_{j-1} \| \cdot \| \varphi(D)^{-1} \| \leq \sigma_j(X) \cdot \sup_{z \in E} | \varphi(z) | \cdot \inf_{z \in F} | \varphi(z) |. \]

This bound holds for any \( \varphi \in \mathcal{R}_{k,k} \). Taking the infimum over all \( \varphi \in \mathcal{R}_{k,k} \) yields \( \sigma_{\nu k+j}(X) \leq \sigma_j(X)Z_k(E,F) \).

By combining Theorem 5 (with \( j = 1 \)) along with Corollary 3, Beckermann and Townsend established the following result.

Theorem 6. [27] If Spec(C) ⊂ [\( c_1, c_2 \)] and Spec(D) ⊂ [\( d_1, d_2 \)] where \( [c_1, c_2] \) and \( [d_1, d_2] \) are nonoverlapping, then for any integer \( k \geq 0 \),
\[ \sigma_{\nu k+1}(X) \leq 4\| X \| \exp \left[ -\frac{\pi^2 k}{\log(16 \gamma)} \right] \quad \text{where} \quad \gamma = \frac{(d_1 - c_1)(d_2 - c_2)}{(d_2 - c_1)(d_1 - c_2)}. \]

Finally, by combining Theorem 5 (with \( j = 1 \)) along with Corollary 4, we obtain Theorem 3 (stated in Section 3.3).

C Polynomial approximations of the sinc function

For a bandwidth parameter \( W > 0 \), we define the sinc function
\[ g(t) = \frac{\sin(2\pi W t)}{\pi t} \quad \text{for} \quad t \in \mathbb{R}. \]

First, we prove the following bound on the derivatives of \( g(t) \).

Lemma 5. For any non-negative integer \( k \),
\[ |g^{(k)}(t)| \leq (2\pi W)^k \min \left\{ \frac{2W}{k+1}, \frac{1}{\pi |t|} \right\} \quad \text{for all} \quad t \in \mathbb{R}. \]

Proof. For \( k = 0 \), we can apply the inequality \( | \sin \theta | \leq \min\{ |\theta|, 2 \} \) to obtain \( |g(t)| \leq \min\{2W, \frac{2}{\pi |t|} \} \). Hence, we can proceed with the case where \( k \geq 1 \). Note that we can write the sinc function as
\[ g(t) = \frac{\sin(2\pi W t)}{\pi t} = \int_{-W}^{W} e^{j2\pi ft} df. \]

By differentiating under the integral sign \( k \) times, we obtain
\[ g^{(k)}(t) = \frac{d^k}{dt^k} \left[ \int_{-W}^{W} e^{j2\pi ft} df \right] = \int_{-W}^{W} \frac{d^k}{dt^k} \left[ e^{j2\pi ft} \right] df = \int_{-W}^{W} (j2\pi f)^k e^{j2\pi ft} df. \]
Applying the triangle inequality yields the bound
\[
\left| g^{(k)}(t) \right| = \left| \int_{-W}^{W} (j2\pi f)^k e^{j2\pi ft} \, df \right| \leq \int_{-W}^{W} \left| (j2\pi f)^k e^{j2\pi ft} \right| \, df = \int_{-W}^{W} (2\pi f)^k \, df = \frac{(2\pi W)^{k+1}}{\pi(k+1)}.
\]
Alternatively, we can use integration by parts before applying the triangle inequality to obtain
\[
\left| g^{(k)}(t) \right| = \left| \int_{-W}^{W} (j2\pi f)^k e^{j2\pi ft} \, df \right|
\]
\[
= \left| \left[ (j2\pi f)^k e^{j2\pi ft} \right]_{f=-W}^{f=W} - \int_{-W}^{W} j2\pi k (j2\pi f)^{k-1} e^{j2\pi ft} \, df \right|
\]
\[
= \left| \frac{(2\pi W)^k}{j2\pi} + \frac{(2\pi W)^k}{j2\pi} + \frac{\int_{-W}^{W} k^2 (2\pi f)^{k-1} \, df}{j2\pi} \right|
\]
\[
= \left| \frac{(2\pi W)^k}{j2\pi} + \frac{\int_{-W}^{W} k^2 (2\pi f)^{k-1} \, df}{j2\pi} \right|
\]
Combining the two bounds yields
\[
\left| g^{(k)}(t) \right| \leq \min \left\{ \frac{(2\pi W)^{k+1}}{\pi(k+1)}, \frac{2(2\pi W)^k}{\pi|t|} \right\} = (2\pi W)^k \min \left\{ \frac{2W}{k+1}, \frac{2}{\pi|t|} \right\}.
\]

We finish this section by noting a well-known theorem on Chebyshev interpolation.

**Theorem 7.** [31] Suppose \( g \in C^k[a, b] \) for some positive integer \( k \). Define the Chebyshev interpolating polynomial of degree \( k-1 \) by
\[
P_k(t) = \sum_{m=1}^{k} g(t_m) \prod_{m' = 1, \ldots, k}^{m \neq m'} \frac{t - t_{m'}}{t_m - t_{m'}}
\]
where
\[
t_m = \frac{b + a}{2} + \frac{b - a}{2} \cos \left( \frac{2m - 1}{2} \pi \right) \quad \text{for} \quad m = 1, \ldots, k
\]
are the Chebyshev nodes used for interpolation. Then, for any \( t \in [a, b] \), we have
\[
\left| g(t) - P_k(t) \right| \leq \frac{(b-a)^k}{2^{2k-1}k!} \max_{\xi \in [a, b]} \left| g^{(k)}(\xi) \right|.
\]
We will use both Lemma 5 and Theorem 7 in Section 3.4 to prove Theorem 2.