REVIVING CIRCULANT PRECONDITIONERS FOR ADAPTIVE MESH REFINEMENT *

K. WALL †

In collaboration with: James Adler, Xiaozhe Hu, Misha Kilmer

Abstract. We present a preconditioner for solving fractional partial differential equations (PDEs) on an adaptive mesh. Adaptive refinement of the problem domain results in a stiffness matrix with Toeplitz blocks along the main diagonal, while the fractional PDE yields a dense stiffness matrix, where off-diagonal blocks are stored as low-rank approximations. Our preconditioner utilizes ideas from the circulant preconditioner of Chan and Strang [SIAM Journal on Scientific Computing, 1989], which takes advantage of the Toeplitz blocks on the diagonal and also accounts for the low-rank nature of the off-diagonal blocks. We demonstrate its effectiveness at accelerating convergence for our systems and emphasize its efficient application. This work presents theoretical results about the spectral clustering of the preconditioned system. In order to prove these results, special consideration is taken on how the low-rank blocks perturb the eigenvalues of the Toeplitz block-diagonal system. Numerical tests for various fractional orders are used to inspect any assumptions and validate our results.

Key words. Preconditioner, Adaptive Refinement, Toeplitz, Circulant, DST, DCT

1. Introduction.

There has long been interest in solving Toeplitz linear systems efficiently. A matrix A is called Toeplitz if $a_{ij} = a_{i-j}$, in other words, A has constant diagonals. Arbitrary $n \times n$ matrices have up to n^2 unique entries and are solved directly by traditional techniques in $\mathcal{O}(n^3)$ time. Since a Toeplitz matrix has just at most $2n-1$ unique entries, we may expect to be able to solve it in $\mathcal{O}(n^2)$ time. This is indeed the case via techniques such as Levinson's algorithm [**?**]. Even this improvement, however, is infeasible for large systems. Instead we turn to iterative Krylov and multigrid methods. For these methods we can still take advantage of Toeplitz structure by using circulant preconditioners.

A circulant matrix is a Toeplitz matrix, that additionally has the "wrap-around" property where the last entry each row is the first entry of the subsequent row.

$$T = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{-(n-1)} \\ t_1 & t_0 & \cdots & t_{-(n-2)} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \cdots & t_0 \end{pmatrix} \qquad C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{pmatrix} \tag{1.1}$$

Equation 1: T is a Toeplitz matrix and C is circulant.

Toeplitz matrices commonly arise in PDE discretization, signal processing, and control theory. Often the Toeplitz matrices are also symmetric positive definite (SPD). Given an SPD Toeplitz system $Tx = b$, the idea introduced by Strang and Chan is to use certain circulant preconditioners C so that $C^{-1}Tx = C^{-1}b$ is solved in fewer

*This work was funded by NSF.

†Tufts University, (kate.wall@tufts.edu, https://katejeanw.github.io/).

iterations [?].

We leverage this idea to build a preconditioner for stiffness matrices generated from the adaptive finite element method (AFEM) for fractional PDEs. In this setting the problem is discretized on a nonuniform mesh, and the resulting stiffness matrix is dense. In the usual finite element method (FEM) setting an uniform mesh results in a Toeplitz stiffness matrix. In the adaptive setting, after an initial solve on a uniform grid, the error on each element is estimated and the elements with largest error are refined via bisection. It is often the case that neighboring elements are refined the same number of times. So although the mesh is not globally uniform, there are areas of local uniformity. To build an effective preconditioner, we will take advantage of these locally uniform areas and their corresponding Toeplitz blocks in the stiffness matrix.

TODO: citations for facts about FEM and Toeplitz matrices

Although dense, the stiffness matrix can be effectively stored as a hierarchical matrix ($\mathcal{H}$-matrix). Due to weaker interaction between elements that are further apart in the domain, off-diagonal blocks are well-suited for low rank approximation. (See [?] for more stiffness matrix details.) The low rank representation makes for fast computations, but complicates both the implementation of the preconditioner and the spectral clustering of the preconditioned system.

In this paper we investigate how to precondition such systems using circulant matrices. Our investigation is focused on $\mathcal{H}$-matrices as in [?], but the same methods could be used on any matrix with Toeplitz blocks on the diagonal. We prove the preconditioned system has eigenvalues clustered around 1 and demonstrate numerical results with superlinear convergence.

We emphasize that our unique contributions are:

- building circulant preconditioners for adaptive meshes
- proving the preconditioned system has eigenvalues clustered around 1
- something else? numerical results? dealing with low-rank blocks?

2. Background.

To understand the need for our preconditioner, we must give a bit more detail about AFEM. We restrict our attention to problems on a one-dimensional domain, $[a, b]$ with the discretization $a = x_0, x_1, \ldots, x_n = b$. Often FEM is done on a uniform mesh, that is each element $[x_i, x_{i+1}]$ is size $x_{i+1} - x_i$ for all $0 \leq i \leq n-1$.

If the mesh is not fine enough to give the desired accuracy, one approach is to increase the number of elements, n. While this approach preserves uniformity, it usually requires recomputing the stiffness matrix entirely. Alternatively, if the level of refinement gives sufficiently small error for some parts of the domain, we can leave those unchanged and only refine in areas of larger error. This approach allows us to take advantage of computations that have already been performed, but the mesh is no longer uniform.

In practice we find that adjacent elements are often refined to the same level, so a group of elements forms a locally uniform mesh. Since uniform meshes give rise to SPD Toeplitz systems, we can see that if we formed the stiffness matrix for just a locally uniform subdomain we would have an SPD Toeplitz matrix. So wherever there are adjacent elements of the same size we can find a corresponding Toeplitz block on the main diagonal of our stiffness matrix, A. The size of this block depends on how many adjacent elements are the same size. We have also observed that the boundary of the domain almost always requires the greatest level of refinement. In general we have larger Toeplitz blocks from locally uniform subdomains in the middle of the matrix, and smaller blocks—or indeed 1×1 blocks—near the boundary.

To build an effective preconditioner and investigate its properites, we have to take full advantage of these SPD Toeplitz blocks. Toeplitz matrices have many unique properties that give rise to efficient algorithms (see for example [?]). For our purposes we focus on their connection to functions in the Wiener class. This will allow us to take our problem from matrix operator theory to function theory. Suppose we have a singly infinite, symmetric Toeplitz matrix

$$T = \begin{pmatrix} t_0 & t_1 & t_2 & \\ t_1 & t_0 & t_1 & \ddots \\ t_2 & t_1 & t_0 & \ddots \\ & \ddots & \ddots & \ddots \end{pmatrix}.$$

Assume $\sum_{k=-\infty}^{\infty} |t_k| < \infty$. Then the function $t(z) = \sum_{k=-\infty}^{\infty} t_k z^k$ is real, positive, and in the Wiener class for $|z| = 1$. It will be convenient to define the corresponding truncated function for a finite subsection of the infinite matrix:

Definition 2.1. *The $m \times m$ finite subsection of the singly infinite matrix T is denoted T_m and defined as*

$$T_m = \begin{pmatrix} t_0 & t_1 & \cdots & t_{m-1} \\ t_1 & t_0 & \cdots & t_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m-1} & t_{m-2} & \cdots & t_0 \end{pmatrix}.$$

Similarly this matrix induces a function from the truncated series, $t_m(z) = \sum_{k=-(m-1)}^{m-1} t_k z^k$.

Previous work on Toeplitz systems offers a few options for circulant preconditioners. Although their construction differs, the spectrum of the preconditioned systems are asymptotically the same [?]. We use the construction given by Bini and Benedetto [?]. Given a symmetric Toeplitz matrix T we build a Hankel correction, H where

$$T = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & t_{n-2} & t_{n-1} \\ t_1 & t_0 & t_1 & \cdots & t_{n-3} & t_{n-2} \\ t_2 & t_1 & t_0 & \cdots & t_{n-4} & t_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ t_{n-2} & t_{n-3} & t_{n-4} & \cdots & t_0 & t_1 \\ t_{n-1} & t_{n-2} & t_{n-3} & \cdots & t_1 & t_0 \end{pmatrix} H = \begin{pmatrix} t_2 & t_3 & t_4 & \cdots & t_{n-1} & 0 & 0 \\ t_3 & t_4 & t_5 & \cdots & 0 & 0 & 0 \\ t_4 & t_5 & t_6 & \cdots & 0 & 0 & t_{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & t_5 & t_4 & t_3 \\ 0 & 0 & t_{n-1} & \cdots & t_4 & t_3 & t_2 \end{pmatrix}.$$

Then Bini and Benedetto's preconditioner is defined as $\boldsymbol{\tau} := T - H$. H is a Hankel matrix, that is it is constant on the antidiagonals. Both T and H are symmetric, so they can be completely represented by the at most n unique values: the diagonals of $T, t_0, t_1, \ldots, t_{n-1}$, and the antidiagonals of $H, t_2, t_3, \ldots, t_{n-1}, 0, 0$.

We summarize the important properties of $\boldsymbol{\tau}$ (see [?] for details):

- $\boldsymbol{\tau}$ is a circulant matrix.
- $\boldsymbol{\tau}$ can be applied in $n \log n$ time .
- $\boldsymbol{\tau}$ is diagonalized by the type-I discrete sine transform (DST) matrix, S. We write $\boldsymbol{\tau} = S\Lambda S^{-1} = S\Lambda S$.
- The k-th eigenvalue of $\boldsymbol{\tau}$ is proportional to the k-th entry of St_1 where t_1 is

the first column of T. Specifically, define $c_k := \sqrt{\frac{n+1}{2}}\frac{1}{\sin(\frac{\pi k}{n+1})}$, then

$$\lambda_k(\boldsymbol{\tau}) = c_k[St_1]_k \tag{2.1}$$

.

- Each eigenvalue of $\boldsymbol{\tau}$ can be written as the truncated function t_m evaluated somewhere on the unit circle, IE $\lambda_k(\boldsymbol{\tau}) = t_m(z_k)$ where $|z_k| = 1$.

These preconditioner can also be thought of as coming from the kernel of a displacement operator. This framework is useful for generating yet other circulant preconditioners, see [**?**].

3. Our Preconditioner.

In this section we set forth the properties we require from a preconditioner, how we use $\boldsymbol{\tau}$ in building our preconditioner, and how to build and apply our preconditioner efficiently.

A good preconditioner for an iterative method must in general decrease the total number of iterations, without increasing the cost of a single iteration. We borrow Kailath's [**?**] criteria for preconditioners, though similar criteria has been established in literature (for example [**?**]).

1. Complexity of constructing applying $\boldsymbol{\tau}$ should be $\mathcal{O}(m \log m)$.
2. A linear system with $\boldsymbol{\tau}$ should be solved in $\mathcal{O}(m \log m)$ operations.
3. The spectrum of $\boldsymbol{\tau}^{-1}A$ should be clustered around 1

How tightly the eigenvalues cluster around 1 will determine the speed of convergence. In proving results about the clustering we will refer to the infinite matrix framework established previously. First we summarize results established previously about the spectral clustering of $\boldsymbol{\tau}$ applied to a single Toeplitz block T. This result will then be used as a lemma in proving the spectral clustering for our stiffness matrix from the adaptive mesh. We will show that for m large enough

We make this last point more precise:

Definition 3.1 (Eigenvalue Clustering). *For any $\varepsilon > 0$ we say the eigenvalues of a matrix $C^{-1}T_m$ are clustered around a real number ρ if there exists N_1 and N_2 such that for all $m > N_1$ there are at most N_2 eigenvalues of $C^{-1}T_m$ that do not lie within $[\rho - \varepsilon, \rho + \varepsilon]$.*

Assume our stiffness matrix has k Toeplitz blocks $T_1, T_2, \dots, T_k$ of respective sizes $m_1, m_2, \dots, m_k$. Assume additionally these are ordered as they appear along the main diagonal. Each block can be thought of as a singly infinite matrix with a corresponding generating function, $t_1(z), t_2(z), \dots t_k(z)$. Assume that for all $1 \le i \le k, t_i(z) \sum_{j=-\infty}^{\infty} |t_j| < \infty$ and $t_i(z) > 0$ for z on the unit circle.

We can now construct a preconditioner, $\mathcal{T}$ for the adaptive system. To explicitly construct $\mathcal{T}$ we can calculate the Hankel correction H_i for every Toeplitz block, T_i. Then we define $\boldsymbol{\tau}_i = T_i - H_i$, resulting in k matrices $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_k$. Finally we assemble $\mathcal{T}$ as the block diagonal matrix with $\boldsymbol{\tau}_i$ as the i-th block.

$$\mathcal{T} = \begin{pmatrix} \boldsymbol{\tau}_1 & & & \\ & \boldsymbol{\tau}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{\tau}_m \end{pmatrix}.$$

This explicit construction is not the most efficient, instead we can apply $\mathcal{T}^{-1}$ implicitly in $m \log m$ time where $m = \sum_{j=1}^{k} m_j$.

Let S_{m_i} denote the type-I DST matrix of size m_i. Using equation 2.1 to calculate the eigenvalues of $\boldsymbol{\tau}_i$, we can write $\boldsymbol{\tau}_i = S_{m_i}\Lambda_i S_{m_i}$.

$$\mathcal{T}^{-1} = \begin{pmatrix} S_{m_1} & & & \\ & S_{m_2} & & \\ & & \ddots & \\ & & & S_{m_k} \end{pmatrix} \begin{pmatrix} \Lambda_1^{-1} & & & \\ & \Lambda_2^{-1} & & \\ & & \ddots & \\ & & & \Lambda_m^{-1} \end{pmatrix} \begin{pmatrix} S_{m_1} & & & \\ & S_{m_2} & & \\ & & \ddots & \\ & & & S_{m_k} \end{pmatrix}.$$

Multiplication of the $m \times m$ matrices that consist of the DST blocks can be done via the fast Fourier transform in $m \log m$ time. The eigenvalue matrix is diagonal and can be inverted in m operations. The scaling done by this diagonal matrix could in the worst case cost $\mathcal{O}(m^2)$, but as we will see in the next section, we need not scale our entire stiffness matrix.

is this true?

4. Theoretical Results.

Story: we know how this works on one block, what about block diagonal, what about with things happening outside of block diagonal? First, we recall some classic eigenvalue results that will be of use to us later.

4.1. Eigenvalue Bounds.

LEMMA 4.1 (Weyl's Inequality).
Let M and E be Hermitian $n \times n$ matrices. Then for $A := M + E$ we have

$$|\lambda_k(A) - \lambda_k(M)| \leq ||E||_2 \ , 1 \leq k \leq n.$$

That is the eigenvalues of A are at most $||E||_2$ away from the eigenvalues of M.

4.1.1. Kailath-Olshevsky Proof rewritten.

Fix $\varepsilon > 0$. Assumptions:

1. Generating function is in the Wiener Class,

$$a(z) = \sum_{k=-\infty}^{\infty} a_k z^k, \quad \sum_{k=-\infty}^{\infty} |a_k| < \infty.$$

2. Generating function is bounded away from zero on unit circle,

$$a(z) > 2\varepsilon, \ |z| = 1.$$

LEMMA 4.2. *Let $a_m(z)$ be the truncated generating function with $2m-1$ terms, $\sum_{k=-(m-1)}^{m-1} a_k z^k$. Then for each $\lambda_k(S_Q(A))$ there exists z_k on the unit circle such that $\lambda_k = a_m(z_k)$.*

LEMMA 4.3. *Choose m and $N < m$ big enough so that $\sum_{N+1}^{\infty} |a_k| < \varepsilon$. Then assumption 2 implies that $a_m(z)$ is positive on the unit circle.*

Proof. First notice $\varepsilon > \sum_{N+1}^{\infty} |a_k| > \sum_{m}^{\infty} |a_k|$. Now

$$a(z) = a_m(z) + \sum_{-\infty}^{-m} a_k z^k + \sum_{m}^{\infty} a_k z^k$$

$$\implies a_m(z) = a(z) - \sum_{-\infty}^{-m} a_k z^k - \sum_{m}^{\infty} a_k z^k$$

$$\implies a(z) - 2\varepsilon < a_m(z)$$

$$0 < a(z) - 2\varepsilon < a_m(z)$$

∎

COROLLARY 4.4. *The matrices $S_Q(A_m)$ and $S_Q(A_m)^{-1}$ are positive definite.*

First we present the proof for a single Toeplitz matrix/block, adapted from [**?**]. This will be used in the proof of the full matrix spectral clustering.

Lemma statement

$$S_Q(A) = A + H + B \tag{4.1}$$

$$A = S_Q(A) - (H + B) \tag{4.2}$$

Where A is Toeplitz (given), H is Hankel, and B is 'border' matrix, at most nonzero in exterior rows and columns. Thus

$$S_Q(A)^{-1}A = I - S_Q(A)^{-1}(H + B).$$

So it suffices to show the spectrum of $S_Q(A)^{-1}(H + B)$ is clustered around zero. Let $\varepsilon > 0$ and choose N such that $\sum_{N+1}^{\infty} |a_k| < \varepsilon$. We can then split $H + B$ into the sum of a low-rank matrix A_{lr} and a small norm matrix A_{sn}. Here A_{lr} contains the diagonals with entries $a_0, \ldots, a_N$. Let $s := \operatorname{rank}(A_{lr}) << m$. Now $A_{sn} := (H + B) - A_{lr}$ is a hermitian $m \times m$ matrix with at most two copies of $a_{N+1}, \ldots, a_m$ in each row/column. Thus $||A_{sn}||_2 = \sqrt{||A_{sn}||_1 ||A_{sn}||_\infty} = ||A_{sn}||_1 < 2\varepsilon$. Hence by Weyl's Inequality at least $m - s$ of the eigenvalues of $H + B$ are clustered within 2ε of zero.

Now we use the min-max theorem to bound the eigenvalues of $S_Q(A)^{-1}(H + B)$.

$$\begin{aligned}
\lambda_k(S_Q(A)^{-1}(H + B)) &= \min_{\dim V = k} \max_{x \in V} \left(\frac{((H + B)x, x)}{(S_Q(A)x, x)} \right) \\
&\leq \min_{\dim V = k} \left[\max_{x \in V} \left(\frac{((H + B)x, x)}{(x, x)} \right) \max_{x \in V} \left(\frac{(x, x)}{(S_Q(A)x, x)} \right) \right] \\
&\leq \left[\min_{\dim V = k} \max_{x \in V} \left(\frac{((H + B)x, x)}{(x, x)} \right) \right] \max_{x \in \mathbb{R}^n} \left(\frac{(x, x)}{(S_Q(A)x, x)} \right) \\
&= \lambda_k(H + B) \max_{x \in \mathbb{R}^n} \left(\frac{(x, x)}{(S_Q(A)x, x)} \right) \\
&\leq \lambda_k(H + B) \frac{1}{\lambda_{\min}(S_Q(A))} \\
&= \lambda_k(H + B) \frac{1}{a_m(z_{\min})} \\
&\leq \lambda_k(H + B) \frac{1}{\min_{|z|=1} a_m(z)}
\end{aligned}$$

Can be simplified with $B = \mathbf{0}$. Actual condition: $a(z) > 2\varepsilon$.

4.2. Full Matrix Proof.

4.2.1. setup. A single block preconditioner is τ the block diagonal preconditioner is $\mathcal{T}$.

On a single block we write $\tau = A - H$, but for the full adaptive matrix A includes off diagonal blocks. Denote the diagonal (Toeplitz blocks) as A_D and everything else A_E so that

$$A = A_D + A_E + B$$

. And thus the splitting as in [?] is expressed $A_D = \mathcal{T} + H$ and $A = A_E + B + \mathcal{T} + H$. So

$$\mathcal{T}^{-1}A = \mathcal{T}^{-1}(\mathcal{T} + H + A_E + B) = I + \mathcal{T}^{-1}H + \mathcal{T}^{-1}A_E + \mathcal{T}^{-1}B \tag{4.3}$$

4.2.2. Proof. It suffices to show that $\mathcal{T}^{-1}H$, $\mathcal{T}^{-1}B$ and $\mathcal{T}^{-1}A_E$ have spectra clustered around zero. First notice that $\mathcal{T}^{-1}H$ is block diagonal and the spectrum of each block can be characterized using the former proof on each block.

since we don't really choose block size in practice the actual block size dictates the size of ε. Over all the blocks we can take the max ε for a uniform bound, but many will be clustered tighter than that. Supports argument that bigger Toeplitz blocks = better clustering

Assume the off-diagonal-by-one blocks are low-rank. Let C be such a block with dimensions $n_C \times n_C$ and rank $r_C << n_C$. Using the SVD we can split C as

$$C = \left(\sum_{i=1}^{r_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)} *\right) + \left(\sum_{i=r_C+1}^{n_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)} *\right).$$

With a slight abuse of notation, we can embed this decomposition in the appropriate "off-diagonal" position of an $m \times m$ matrix. Doing this for all such off-diagonal blocks we write

$$\begin{aligned} B &= \sum_{C \in \text{off-diag}} \left[\left(\sum_{i=1}^{r_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)} *\right) + \left(\sum_{i=r_C+1}^{n_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)} *\right)\right] \\ &= \left(\sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i *\right) + \left(\sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i *\right) \end{aligned}$$

where $r_B = \max_{C \in \text{off-diag}} r_C$.

We additionally split H by separating the anti-diagonals with coefficients $a_0, \ldots, a_N$ and the anti-diagonals comprising of $a_{N+1}, \ldots a_m$. So we have two splittings,

$$\begin{aligned} B &= \left(\sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i *\right) + \left(\sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i *\right) \\ H &= H|_{a_0,\ldots,a_N} + H|_{a_{N+1},\ldots,a_m}. \end{aligned}$$

The first term in each sum can be thought of as our 'low-rank' equivalent from before and similarly the second term is our 'small-norm' summand.

Bound on number of off diagonal blocks

Finally we can make the splitting $A = A_{SN} + A_{LR}$ where

$$\begin{aligned} A_{SN} &= H|_{a_{N+1},\ldots,a_m} + \sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i * + A_E \\ A_{LR} &= H|_{a_0,\ldots,a_N} + \sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i *. \end{aligned}$$

A_{LR} represent outliers, IE $s := \text{rank}(A_{LR}) \le N + r_B$ bounds the number of outliers.

Is the N part of this bound true? 2N?

So the work is showing $||\mathcal{T}^{-1}A_{SN}||_2 \leq \varepsilon$. Define $\tilde{B} = \sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i*$ and $\tilde{H} = H|_{a_{N+1},\ldots,a_m}$, so that $A_{SN} = \tilde{H} + \tilde{B} + A_E$.

$$||\mathcal{T}^{-1}A_{SN}||_2 \leq ||\mathcal{T}^{-1}\tilde{H}||_2 + ||\mathcal{T}^{-1}\tilde{B}||_2 + ||\mathcal{T}^{-1}A_E||_2$$

We can bound $||\mathcal{T}^{-1}\tilde{H}||_2$ as in [**?**]. We can bound $||\mathcal{T}^{-1}A_E||_2$ with Weyl's inequality:

$$||\mathcal{T}^{-1}A_E||_2 \leq ||\mathcal{T}^{-1}||_2||A_E||_2 = \sigma_{\max}(\mathcal{T}^{-1})\sigma_{\max}(A_E) = \frac{\sigma_{\max}(A_E)}{\lambda_{\min}(\mathcal{T})}.$$

Finally we bound $||\mathcal{T}^{-1}\tilde{B}||_2$.

$$\begin{aligned}
\lambda_k(\mathcal{T}^{-1}\tilde{B}) &= \min_{\dim V = k} \max_{x\in V} \left(\frac{(\tilde{B}x, x)}{(\mathcal{T}x, x)} \right) \\
&\leq \min_{\dim V = k} \left[\max_{x\in V} \left(\frac{(\tilde{B}x, x)}{(x, x)} \right) \max_{x\in V} \left(\frac{(x, x)}{(\mathcal{T}x, x)} \right) \right] \\
&\leq \left[\min_{\dim V = k} \max_{x\in V} \left(\frac{(\tilde{B}x, x)}{(x, x)} \right) \right] \max_{x\in \mathbb{R}^n} \left(\frac{(x, x)}{(\mathcal{T}x, x)} \right) \\
&= \lambda_k(\tilde{B}) \max_{x\in \mathbb{R}^n} \left(\frac{(x, x)}{(\mathcal{T}x, x)} \right) \\
&\leq \lambda_k(\tilde{B}) \frac{1}{\lambda_{\min}(\mathcal{T})} \\
&= \lambda_k(\tilde{B}) \min_{n\in n_k} \min_{1\leq i\leq n} \frac{\sin(\frac{\pi i}{n+1})}{\sum_{j=1}^{n} t_j \sin(\frac{\pi i j}{n+1})}
\end{aligned}$$

Since $\tilde{B}$ made of blocks that have form $\sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i*$ what can we say about λ_k?

- numerical test confirming off-diag low rank
- Explanation and tests showing off-off-diag are small norm
- technically lots of 1×1 blocks at boundaries, these get jacobi inverse treatment so are clustered around 1
- Comment - all problems come from boundaries
- Extend proof to different kinds of circulant preconditioner

5. Numerical Results.

- enough info to reproduce
- Single block clustering
- Adaptive clustering (what happens to smallest eigenvalue?)
- behavior for different α
- Verify assumptions from proof
- convergence of solving with PCG (superlinear convergence)

6. Conclusion. Future work: how to build adaptive mesh to increase block size, other circulant preconditioners, tensor preconditioners, higher dimension domain, mixed precision