

1 **REVIVING CIRCULANT PRECONDITIONERS FOR ADAPTIVE
2 MESH REFINEMENT ***

3 K. WALL †

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

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

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

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

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

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

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

29 Toeplitz matrices commonly arise in PDE discretization, signal processing, and
30 control theory. Often the Toeplitz matrices are also symmetric positive definite (SPD).
31 Given an SPD Toeplitz system $Tx = b$, the idea introduced by Strang and Chan is
32 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/>).

33 iterations [?].

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

45 **TODO:** citations for facts about FEM and Toeplitz matrices

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

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

57 We emphasize that our unique contributions are:

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

61 **2. Background.** To understand the need for our preconditioner, we must give
 62 a bit more detail about AFEM. We restrict our attention to problems on a one-
 63 dimensional domain, $[a, b]$ with the discretization $a = x_0, x_1, \dots, x_n = b$. Often FEM
 64 is done on a uniform mesh, that is each element $[x_i, x_{i+1}]$ is size $x_{i+1} - x_i$ for all
 65 $0 \leq i \leq n - 1$.

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

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

82 of the matrix, and smaller blocks—or indeed 1×1 blocks—near the boundary.

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

$$89 \quad T = \begin{pmatrix} t_0 & t_1 & t_2 & & \\ t_1 & t_0 & t_1 & \ddots & \\ t_2 & t_1 & t_0 & \ddots & \\ \ddots & \ddots & \ddots & \ddots & \end{pmatrix}.$$

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

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

$$95 \quad T_m = \begin{pmatrix} t_0 & t_1 & \cdots & t_{m-1} \\ t_1 & t_0 & \cdots & t_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m-1} & t_{m-2} & \cdots & t_0 \end{pmatrix}.$$

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

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

$$101 \quad T = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & t_{n-2} & t_{n-1} \\ t_1 & t_0 & t_1 & \cdots & t_{n-3} & t_{n-2} \\ t_2 & t_1 & t_0 & \cdots & t_{n-4} & t_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ t_{n-2} & t_{n-3} & t_{n-4} & \cdots & t_0 & t_1 \\ t_{n-1} & t_{n-2} & t_{n-3} & \cdots & t_1 & t_0 \end{pmatrix} H = \begin{pmatrix} t_2 & t_3 & t_4 & \cdots & t_{n-1} & 0 & 0 \\ t_3 & t_4 & t_5 & \cdots & 0 & 0 & 0 \\ t_4 & t_5 & t_6 & \cdots & 0 & 0 & t_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & t_5 & t_4 & t_3 \\ 0 & 0 & t_{n-1} & \cdots & t_4 & t_3 & t_2 \end{pmatrix}.$$

102 Then Bini and Benedetto's preconditioner is defined as $\tau := T - H$. H is a Hankel
 103 matrix, that is it is constant on the antidiagonals. Both T and H are symmetric, so
 104 they can be completely represented by at most n unique values: the diagonals of
 105 $T, t_0, t_1, \dots, t_{n-1}$, and the antidiagonals of $H, t_2, t_3, \dots, t_{n-1}, 0, 0$.

106 We summarize the important properties of τ (see [?] for details):

- 107 • τ is a circulant matrix.
- 108 • τ can be applied in $n \log n$ time .
- 109 • τ is diagonalized by the type-I discrete sine transform matrix, S . We write
 110 $\tau = SAS^{-1} = S\Lambda S$.
- 111 • Each eigenvalue of τ can be written as the truncated function t_m evaluated
 112 somewhere on the unit circle, IE $\lambda_k(\tau) = t_m(z_k)$ where $|z_k| = 1$.

113 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 [?].
 114
 115

116 **3. Our Preconditioner.** In this section we set forth the properties we require
 117 from a preconditioner, how we use τ in building our preconditioner, and how to build
 118 and apply our preconditioner efficiently.
 119

120 A good preconditioner for an iterative method must in general decrease the total
 121 number of iterations, without increasing the cost of a single iteration. We borrow
 122 Kailath's [?] criteria for preconditioners, though similar criteria has been established
 123 in literature (for example [?]).
 124

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

128 How tightly the eigenvalues cluster around 1 will determine the speed of con-
 129 vergence. In proving results about the clustering we will refer to the infinite matrix
 130 framework established previously. First we summarize results established previously
 131 about the spectral clustering of τ applied to a single Toeplitz block T . This result
 132 will then be used as a lemma in proving the spectral clustering for our stiffness matrix
 133 from the adaptive mesh. We will show that for m large enough
 134

135 We make this last point more precise:

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

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

144 We can now construct a preconditioner, \mathcal{T} for the adaptive system. To explicitly
 145 construct \mathcal{T} we can calculate the Hankel correction H_i for every Toeplitz block, T_i .
 Then we define $\tau_i = T_i - H_i$, resulting in k matrices $\tau_1, \tau_2, \dots, \tau_k$. Finally we assemble
 \mathcal{T} as the block diagonal matrix with τ_i as the i -th block.

$$\mathcal{T} = \begin{bmatrix} \tau_1 & & & \\ & \tau_2 & & \\ & & \ddots & \\ & & & \tau_m \end{bmatrix}.$$

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

148 To apply it $m \log m$ instead, (implicit) Get sizes of Toeplitz blocks n_j , calculate
 $\lambda_i = c_i[S\mathbf{t}]_i$ for $1 \leq i \leq n_j$, apply $S\Lambda^{-1}S$.

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

152 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

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

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

151 **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,

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

152 LEMMA 4.2. Let $a_m(z)$ be the truncated generating function with $2m - 1$ terms,
153 $\sum_{k=-(m-1)}^{m-1} a_k z^k$. Then for each $\lambda_k(S_Q(A))$ there exists z_k on the unit circle such
154 that $\lambda_k = a_m(z_k)$.

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

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

$$158 \quad a(z) = a_m(z) + \sum_{-\infty}^{-m} a_k z^k + \sum_m^{\infty} a_k z^k$$

$$159 \quad \implies a_m(z) = a(z) - \sum_{-\infty}^{-m} a_k z^k - \sum_m^{\infty} a_k z^k$$

$$160 \quad \implies a(z) - 2\varepsilon < a_m(z)$$

$$161 \quad 0 < a(z) - 2\varepsilon < a_m(z) \quad \square$$

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

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

166 Lemma statement

$$167 \quad (4.1) \quad S_Q(A) = A + H + B$$

$$168 \quad (4.2) \quad A = S_Q(A) - (H + B)$$

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

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

170 So it suffices to show the spectrum of $S_Q(A)^{-1}(H + B)$ is clustered around zero.

171 Let $\varepsilon > 0$ and choose N such that $\sum_{N+1}^{\infty} |a_k| < \varepsilon$. We can then split $H + B$ into
172 the sum of a low-rank matrix A_{lr} and a small norm matrix A_{sn} . Here A_{lr} contains

173 the diagonals with entries a_0, \dots, a_N . Let $s := \text{rank}(A_{lr}) \ll m$. Now $A_{sn} :=$
 174 $(H + B) - A_{lr}$ is a hermitian $m \times m$ matrix with at most two copies of a_{N+1}, \dots, a_m
 175 in each row/column. Thus $\|A_{sn}\|_2 = \sqrt{\|A_{sn}\|_1 \|A_{sn}\|_\infty} = \|A_{sn}\|_1 < 2\varepsilon$. Hence by
 176 Weyl's Inequality at least $m - s$ of the eigenvalues of $H + B$ are clustered within 2ε
 177 of zero.

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

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

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

188 4.2. Full Matrix Proof.

189 **4.2.1. setup.** A single block preconditioner is τ the block diagonal precondi-
 190 tioner 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

$$A = A_D + A_E + B$$

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

$$193 (4.3) \quad \mathcal{T}^{-1}A = \mathcal{T}^{-1}(\mathcal{T} + H + A_E + B) = I + \mathcal{T}^{-1}H + \mathcal{T}^{-1}A_E + \mathcal{T}^{-1}B$$

194 **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
 195 clustered around zero. First notice that $\mathcal{T}^{-1}H$ is block diagonal and the spectrum of
 196 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

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

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

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

$$\begin{aligned} 201 \quad B &= \sum_{C \in \text{off-diag}} \left[\left(\sum_{i=1}^{r_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)*} \right) + \left(\sum_{i=r_C+1}^{n_C} \sigma_i^{(C)} \mathbf{u}_i^{(C)} \mathbf{v}_i^{(C)*} \right) \right] \\ 202 \quad &= \left(\sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^* \right) + \left(\sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^* \right) \\ 203 \end{aligned}$$

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

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

$$\begin{aligned} 207 \quad B &= \left(\sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^* \right) + \left(\sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^* \right) \\ 208 \quad H &= H|_{a_0, \dots, a_N} + H|_{a_{N+1}, \dots, a_m}. \end{aligned}$$

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

212 **Bound on number of off diagonal blocks**

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

$$\begin{aligned} 214 \quad A_{SN} &= H|_{a_{N+1}, \dots, a_m} + \sum_{i=r_B+1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^* + A_E \\ 215 \quad A_{LR} &= H|_{a_0, \dots, a_N} + \sum_{i=1}^{r_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^*. \\ 216 \end{aligned}$$

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

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

219 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
 220 $\tilde{H} = H|_{a_{N+1}, \dots, a_m}$, so that $A_{SN} = \tilde{H} + \tilde{B} + A_E$.

$$222 \quad \|\mathcal{T}^{-1} A_{SN}\|_2 \leq \|\mathcal{T}^{-1} \tilde{H}\|_2 + \|\mathcal{T}^{-1} \tilde{B}\|_2 + \|\mathcal{T}^{-1} A_E\|_2$$

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

$$224 \quad \|\mathcal{T}^{-1} A_E\|_2 \leq \|\mathcal{T}^{-1}\|_2 \|A_E\|_2 = \sigma_{\max}(\mathcal{T}^{-1}) \sigma_{\max}(A_E) = \frac{\sigma_{\max}(A_E)}{\lambda_{\min}(\mathcal{T})}.$$

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

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

233 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 ?

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

240 5. Numerical Results.

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

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