Università degli Studi dell'Insubria

Ph.D. in Computer Science and Mathematics of Computation – XXXIX Cycle

Asymptotic Linear Algebra and Approximation Theory for Structured Problems

Valerio Loi



Advisors: Luisa Fermo
Stefano Serra Capizzano
Academic year: 2025/2026

Abstract

This thesis studies structure-preserving approximations and representations in asymptotic linear algebra, numerical linear algebra, and approximation theory. The central question is how a tractable transformation preserves the asymptotic and algebraic/analytic properties relevant to the original problem. The analysis then determines when the preserved information supports effective preconditioning, stable quadrature, or convergent interpolation, and which additional hypotheses are required for these numerical conclusions.

1) For block matrix-sequences with irregularly sized Toeplitz blocks and rational limiting proportions, the thesis aligns the physical sequence with a canonical ambient Toeplitz model by asymptotically full extradimensional compression. It then constructs a fast structured surrogate S_n through block replication and Toeplitz approximation in fast matrix algebras. The difference sequence $\{S_n - A_n\}_n$ is shown to be zero-distributed under the basic approximation assumptions. Additional continuity and block-level clustering assumptions yield strong clustering of that difference at zero, while a.e. full-rank symbol assumptions yield weak singular-value clustering of the appropriate preconditioned products at one. In the square Hermitian case with a Hermitian positive-definite preconditioner, the corresponding weak eigenvalue clustering is proved. Use of PCG additionally requires the system matrix to be Hermitian positive definite. The numerical experiments assess these conclusions and the observed Krylov behavior.

2) For single-axis quaternion block multilevel Toeplitz sequences, the complex symplectic representation transfers the problem exactly to structured complex matrix-sequences. Within this setting, the thesis derives asymptotic singular-value distributions for integrable matrix-valued symbols and canonical right-eigenvalue distributions under the stated Hermitian or non-Hermitian assumptions. It also establishes Hermitian localization and positivity criteria, a block-size-uniform Schatten-norm estimate, and right-circulant approximating classes with the required doubled normalization.

3) For the circular Hilbert transform, the thesis constructs paired Szegő and anti-Szegő rules whose para-orthogonal parameter depends on the evaluation point. The construction preserves the exact opposite-error identity on the stated Laurent-polynomial space while separating both node sets from the kernel singularity. Under the chapter's regularity and approximation hypotheses, this geometry yields logarithmic pointwise operator and quadrature error bounds and the stated periodic Sobolev upper rate.

4) For interpolation at anti-Gauss Jacobi nodes, the thesis derives modified Christoffel–Darboux, barycentric, and orthogonal-expansion formulas that represent the same interpolant exactly. Their finite-precision behavior is compared experimentally. Under the stated internality and endpoint-weight conditions, local node, polynomial, derivative, and Christoffel estimates give an $O(\log n)$ upper bound for the weighted Lebesgue constants and the resulting weighted convergence estimates.

The results share a method across asymptotic linear algebra, approximation theory, and numerical linear algebra. We construct a tractable structured transformation, prove which asymptotic or exact invariants it preserves, and identify the structure-specific estimates needed for rigorous numerical conclusions.

Publications and Manuscripts Related to the Thesis

This thesis incorporates five works as substantial parts of four core research chapters. Three additional contributions to structured matrix-sequence analysis are listed separately below. The material incorporated into the thesis has been revised, reorganized, and adapted to its structure.

Works substantially incorporated into the thesis

1. N. Barakitis, M. Donatelli, S. Ferri, V. Loi, S. Serra-Capizzano, and R. L. Sormani, *Blocking structures, approximation, and preconditioning.* Corresponds to Chapter 3 [9]. Published online on 8 July 2025 in *Numerical Algorithms* 100(4) (2025), 1893–1927, DOI: 10.1007/s11075-025-02157-y; arXiv:2501.14874v1 [math.NA].

2. A. Lailoune, V. Loi, and S. Serra-Capizzano, *Weyl distributions, spectral properties, and circulant approximation results for quaternion block multilevel Toeplitz matrix sequences.* Corresponds to Chapter 4 [112]. Accepted for publication in *Boletín de la Sociedad Matemática Mexicana.* Preprint available as arXiv:2511.22422v3 [math.NA], first posted on 27 November 2025 and last revised on 7 August 2026.

3. L. Fermo and V. Loi, *Approximation of the Hilbert transform on the unit circle* [68]. Substantially incorporated into Chapter 5. Published on 10 September 2025 in *Journal of Scientific Computing* 105(1), article 23 (2025), DOI: 10.1007/s10915-025-03044-3; arXiv:2409.07810v1 [math.NA].

4. P. Díaz de Alba, L. Fermo, and V. Loi, *Anti-Gauss Lagrange interpolation: Christoffel-Darboux form, barycentric representation, and orthogonal expansion* [46]. Substantially incorporated into Chapter 6. Public preprint arXiv:2606.18018v1 [math.NA], posted on 16 June 2026.

5. P. Díaz de Alba, L. Fermo, V. Loi, and D. Occorsio, *Lagrange interpolation processes based on the zeros of anti-Gauss Jacobi polynomials* [47]. Substantially incorporated into Chapter 6. Public preprint arXiv:2607.21714v1 [math.NA], posted on 23 July 2026.

Additional related works

1. N. Barakitis, V. Loi, and S. Serra-Capizzano, *A note on eigenvalues and singular values of variable Toeplitz matrices and matrix-sequences, with application to variable two-step BDF approximations to parabolic equations* [12]. Published as a conference paper in *Recent Developments in Spectral and Approximation Theory* (ICSAT 2023), Trends in Mathematics, Birkhäuser, Cham, 2025, pp. 37–67; first published online on 27 July 2025, DOI: 10.1007/978-3-031-90240-6_4; arXiv:2407.00792v3 [math.NA].

2. S. Ferri, C. Giraudo, V. Loi, M. Kuchta, and S. Serra-Capizzano, *Spectral analysis of the stiffness matrix sequence in the approximated Stokes equation* [72]. Published online on 30 July 2026 in *Boletín de la Sociedad Matemática Mexicana* 32(3), article 102 (2026), DOI: 10.1007/s40590-026-00935-9; arXiv:2510.25252v2 [math.NA].

3. A. Ilyas, M. F. Khan, V. Loi, and S. Serra-Capizzano, *A maximal distribution result and numerical tests for geometric means of HPD GLT matrix-sequences with degenerate commuting/non-commuting GLT symbols* [100]. Available online in *Linear Algebra and its Applications* since 8 August 2025 as an in-press corrected proof, DOI: 10.1016/j.laa.2025.08.004; arXiv:2505.03256v1 [math.NA].

Contents

Abstract **2**

Publications and Manuscripts Related to the Thesis **3**

List of Tables **7**

List of Figures **8**

Notation and Conventions **10**

1 Introduction **12**
1.1 Motivation, scope, and research questions 12
1.2 Background and related work 12
1.3 Main contributions and scope 16
1.4 Organization of the thesis 16

2 Matrix Sequences and Approximation Theory **18**
2.1 Analytical framework and conventions 18
2.2 Spectral distributions and matrix-sequence approximation 18
2.3 Toeplitz, Circulant, and Hankel sequences 28
2.4 Generalized Locally Toeplitz Sequences 35
2.5 Exactness, best approximation, and stability 41
2.6 Gauss–Christoffel and Szegő quadrature 47
2.7 Anti-Gauss and anti-Szegő quadrature 50
2.8 Connections with the subsequent chapters 56

I Structured Matrix Sequences **57**

3 Blocking Structures, Approximation, and Preconditioning **58**
3.1 From block distributions to preconditioning 58
3.2 Problem setting and asymptotic block geometry 58
3.3 Block Toeplitz structures and limiting distributions 60
3.4 Approximation and preconditioning 62
3.5 Numerical experiments 66
3.6 Conclusions 77

4 Quaternion Block Multilevel Toeplitz Sequences **78**
4.1 Scope and main results 78
4.2 Definitions and preliminary results 79

4.3 Quaternion spectral distributions and approximation 84
4.4 Single-axis quaternion Toeplitz sequences 87
4.5 Numerical results 97
4.6 Conclusions 103
4.7 Supplementary calculations and proofs 104

II Structured Approximation Theory 118

5 Approximation of the Hilbert Transform on the Unit Circle 119
5.1 Hilbert transform and numerical obstruction 119
5.2 Function spaces and approximation 120
5.3 Paired quadrature schemes 121
5.4 Pointwise approximation of the Hilbert transform 124
5.5 Numerical tests 130
5.6 Conclusions 136

6 Anti-Gauss Lagrange and Jacobi Interpolation 137
6.1 Representations of the anti-Gauss Lagrange interpolant 137
6.2 Weighted interpolation at anti-Gauss Jacobi nodes 143
6.3 Conclusions 164

7 Conclusions and Perspectives 165
7.1 Summary of the main results 165
7.2 Scope and limitations 166
7.3 Directions for future research 166

List of Tables

1.1 Main theoretical results, numerical consequences, and scope. 17

3.1 Published outlier data for the cluster at 1. 71
3.2 Published Krylov iteration counts. 72
3.3 Example 3.11: eigenvalue outlier counts and CG/PCG iterations. 73
3.4 Example 3.12: reported complex-eigenvalue counts and GMRES/PGMRES iterations. 74
3.5 Example 3.13: exploratory two-level outlier counts and CG/PCG iterations. . 77

4.1 Eigenvalues for the Hermitian embedded symbols and singular values for the non-Hermitian embedded symbols at $\boldsymbol{\theta}_0 = (\pi/3, \pi/4)$. 105

5.1 Direct Hilbert-transform approximations for Example 5.2. 125
5.2 Prescribed Hilbert-transform approximations for Example 5.8. 131
5.3 Errors of the prescribed formulas for Example 5.8. 131
5.4 Paired quadrature and Hilbert-transform errors for Example 5.9. 132
5.5 Paired errors and computable indicators for Example 5.10. 133
5.6 Paired errors for Example 5.11. 134
5.7 Paired errors for Example 5.12. 135

6.1 Sampled interpolation errors for f_1 and f_2. 142
6.2 Gauss and anti-Gauss interpolation errors for Example 6.18. 163

List of Figures

3.1 Singular-value clustering of $\{A_n - S_n\}_n$ at 0, Group 1. 69
3.2 Singular-value clustering of $\{A_n - S_n\}_n$ at 0, Group 2. 69
3.3 Singular-value clustering of $\{A_n - S_n\}_n$ at 0, Group 3. 70
3.4 Singular-value clustering of $\{S_n^{-1}A_n\}_n$ at 1, Group 1. 70
3.5 Singular-value clustering of $\{A_nS_n^{\dagger}\}_n$ at 1, Group 2. 70
3.6 Singular-value clustering of $\{S_n^{-1}A_n\}_n$ at 1, Group 3. 71
3.7 Example 3.11: Hermitian eigenvalues before and after preconditioning. 74
3.8 Example 3.12: non-Hermitian eigenvalues before and after preconditioning. . . 74
3.9 Example 3.13: exploratory two-level eigenvalues before and after preconditioning. 77

4.1 Continuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$. 99
4.2 Continuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$. 99
4.3 Continuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$. 100
4.4 Continuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$. 100
4.5 Discontinuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$. 101
4.6 Discontinuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$. 102

4.7 Discontinuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$. 103
4.8 Discontinuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$. 103

5.1 Errors for Example 5.9 at $n = 4$. 133

Notation and Conventions

This list collects the conventions used across the thesis and the principal symbols of the research chapters. Local notation is defined where it is introduced; the references below give the full definitions.

Symbol	Meaning and scope	Ref.
General conventions		
$\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$; $\Gamma = e^{\mathrm{i}\mathbb{T}}$	Torus, represented by $[-\pi, \pi)$, and unit circle	Sec. 2.6.2
$\mathbb{P}_n$, $\Lambda_{-n,n}$	Algebraic polynomials of degree at most n; Laurent polynomials with powers $-n, \dots, n$	Secs. 2.6.1, 2.6.2
$\mu_\ell(D) = \lvert D\rvert$	Lebesgue measure of D; distribution formulas use $\lvert D\rvert^{-1}\,\mathrm{d}x$	Def. 2.3
A^*, $A^\dagger$	Conjugate transpose; Moore–Penrose pseudoinverse	Ch. 2
$\lVert A\rVert$, $\lVert A\rVert_{S_p}$	Spectral and Schatten p-norms; also used for quaternion matrices in Chapter 4	Defs. 2.1, 4.7
$\mathrm{spec}(A)$; $\lambda_j(A)$, $\sigma_j(A)$	Matrix spectrum, eigenvalues, and singular values	Def. 2.3
$a_n \asymp b_n$; $a_n \sim b_n$	Two-sided comparability; ratio asymptotics $a_n/b_n \to 1$	Sec. 6.2.2
Matrix sequences		
p_n, q_n, r_n; d_n	Row and column dimensions, active size $r_n = \min\{p_n, q_n\}$; full square matrix order d_n	Def. 2.3
$\boldsymbol{n} = (n_1, \dots, n_d)$, $N_{\boldsymbol{n}}$	Multilevel index and product size $N_{\boldsymbol{n}} = \prod_j n_j$; limits have $\min_j n_j \to \infty$	Def. 2.22
$\{A_n\}_n \sim_\sigma f$	Singular-value distribution, normalized by r_n	Def. 2.3
$\{A_n\}_n \sim_\lambda f$	Eigenvalue distribution, normalized by d_n	Def. 2.3
$\{Z_n\}_n \sim_\sigma 0$	Zero-distributed sequence	Def. 2.10
$\xrightarrow{\text{a.c.s.}}$; a.c.s.; g.a.c.s.	Approximating class of sequences in the rank–norm sense; its generalized version	Def. 2.16; Ch. 3
$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$	Generalized Locally Toeplitz relation with symbol κ	Def. 2.41
$\mathrm{ER}_\sigma(F)$, $\mathrm{ER}_\lambda(G)$	Singular-value and spectral essential ranges of matrix-valued symbols	Thm. 2.8
$\widehat{f}_{\boldsymbol{k}}$; $T_{\boldsymbol{n},\boldsymbol{m}}(f)$	Fourier coefficient and rectangular multilevel Toeplitz matrix	Def. 2.22

Continued on the next page

Symbol	Meaning and scope	Ref.
$D_{\boldsymbol{n}}(a)$, $D_n(a;\mathcal{X}_n)$	Block-diagonal sampling on a canonical grid or an empirical node array	Def. 2.28; Prop. 2.30
Quadrature and approximation		
π_n, p_n (Chapter 2)	Monic and orthonormal interval orthogonal polynomials, respectively	Sec. 2.6.1
$p_n(x)$, $p_n(w,x)$ (Chapter 6)	Monic and orthonormal polynomials, respectively; the orthonormal notation includes the weight w	Sec. 6.1.2
$I_\mu(f)$, $Q_n(f)$	Target integral and quadrature rule	Def. 2.51
$\widetilde{Q}_n$, Q_n^{avg}	Second rule in an opposite-error pair and its weighted average with Q_n	Prop. 2.70
G_n, S_n	Gauss and Szegő quadrature rules	Secs. 2.6.1, 2.6.2
L_nf, $\ell_{k,n}$	Lagrange interpolant and fundamental polynomials	Def. 2.55
$\Lambda_n(x)$, Λ_n	Lebesgue function and Lebesgue constant	Def. 2.55
E_m^{alg}, E_m^{Γ}	Best uniform algebraic and symmetric Laurent approximation errors; Chapter 5 abbreviates $E_n^{\Gamma}(f)$ to $E_n(f)$	Eq. (2.23); Sec. 5.2
Principal chapter-specific objects		
$\{A_n\}_n$, $\{S_n\}_n$ (Chapter 3)	Given matrix sequence and structured surrogate sequence	Eqs. (3.2), (3.12)
$\mathbb{H}$, $\mathbf{i},\mathbf{j},\mathbf{k}$, $\mathbb{C}_\mu$	Hamilton quaternions, Cartesian units, and the complex slice generated by the imaginary unit μ	Sec. 4.2.1
Φ_{grp}, Φ_{ent}, $\Phi_{\mathrm{blk}}^{(s,t)}$	Grouped, entrywise, and block-ordered complex symplectic realizations	Sec. 4.2.2; Eq. (4.25)
$\lambda_j^+(A)$; $\sim_{\lambda^+}$	Canonical upper-half-plane right-eigenvalue representatives; the Chapter 4 distribution relation	Thm. 4.5; Def. 4.20
$T_{\boldsymbol{n}}^{(\tau)}(F)$, G_τ	Quaternion Toeplitz matrix and its complex embedded symbol; τ specifies the kernel convention	Thm. 4.32
$(Hf)(\phi)$	Circular Hilbert transform	Eq. (5.3)
H_n^{pr}, $\widetilde{H}_n^{\mathrm{pr}}$, $\widehat{H}_n^{\mathrm{pr}}$	Prescribed Szegő, anti-Szegő, and averaged discrete Hilbert operators	Eqs. (5.20)–(5.22)
$\widetilde{L}_{n+1}(f,x)$	Anti-Gauss Lagrange interpolant	Eq. (6.2)
$v^{\rho,\sigma}$, u; C_u, $E_n(f)_u$	Jacobi weights; weighted uniform space and best polynomial approximation error	Sec. 6.2.2
$\widetilde{\Lambda}_{n+1}(x)$, $\widetilde{\Lambda}_{n+1}(v^{\alpha,\beta})$	Weighted anti-Gauss Lebesgue function and Lebesgue constant	Eqs. (6.77), (6.76)

Chapter 1

Introduction

1.1 Motivation, scope, and research questions

Many problems in numerical linear algebra and approximation theory lead to growing matrix-sequences, quadrature rules, and interpolation operators whose exact analysis becomes impractical as the dimension or polynomial degree increases. Their large-size behavior can often be described by simpler quantities. This approach belongs to asymptotic linear algebra, in which matrix-valued symbols encode limiting singular-value or eigenvalue distributions. The corresponding descriptions in approximation theory involve quadrature exactness relations, empirical node distributions, and weighted operator or polynomial estimates.

This thesis studies which asymptotic or algebraic properties are retained by structured approximations and representations, and which additional estimates yield preconditioning, stability, or convergence results for each problem. Part I concerns irregular-block Toeplitz matrix-sequences and quaternion block multilevel Toeplitz matrix-sequences generated by single-axis Fourier kernels, while Part II concerns prescribed-node quadrature for the Hilbert transform on the unit circle and interpolation at anti-Gauss Jacobi nodes for weighted approximation on compact intervals.

The two parts use different analytical notions of convergence. Both distinguish global information from the local or operator-level estimates required by the problem. This distinction leads to two principal questions:

(i) How can a structured approximation, or an algebraic representation, retain the relevant singular-value or eigenvalue distributions of the matrix-sequences studied in Part I? When can approximations with these properties serve as practical, fast preconditioners for the associated linear systems?

(ii) In weighted approximation problems, which geometric and weighted estimates must supplement the asymptotic distribution of interpolation nodes in Part II to yield optimal and stable quadrature rules and convergent interpolation processes?

1.2 Background and related work

1.2.1 Structured matrix-sequences and computational approximations

Toeplitz matrices provide a classical model in which a generating function describes a growing matrix sequence. The classical Szegő theory and the Avram-Parter singular-value results are

the starting point of this theory [93, 138, 6]. In modern scalar and block formulations, Toeplitz sequences generated by the Fourier coefficients of an integrable symbol f are distributed in singular values as $|f|$ in the scalar case and according to the singular values of the matrix-valued symbol in the block case. Under the corresponding Hermitian, normal, or topological assumptions, an eigenvalue-distribution result also holds. The development of this result can be traced through, for example, [169, 168, 167, 27]. Approximating classes of sequences (a.c.s.) and the Generalized Locally Toeplitz (GLT) theory extend this principle to non-stationary structures through algebraic operations, functional calculus, perturbations that are small in relative rank or norm and are encoded by zero-distributed sequences, and an algebraic treatment of sampling matrices of Riemann integrable functions [156, 157, 76, 77, 79, 78]. GLT theory also treats multilevel, matrix-valued, and rectangular symbols [17, 16, 15]. Generalized a.c.s. (g.a.c.s.) permit approximating matrices of different dimensions and exhaustion arguments for moving or unbounded domains in PDE problems [5]. A GLT-type algebra has also been introduced for problems on unbounded domains of finite measure [3].

Matrix-sequences assembled from rectangular Toeplitz blocks of irregular sizes do not fit the standard square-block GLT framework, and their distributions cannot be obtained by a direct application of the a.c.s. machinery and the usual GLT functional calculus alone. An analysis based on the extradimensional approach was first published for a 2×2 block setting; that account also traces the idea of the extradimensional technique to a 1997 private discussion between S. Serra-Capizzano and E. E. Tyrtyshnikov [11]. Using similar principles, subsequent results identify the eigenvalue and singular-value distributions for asymptotically rational block-size ratios and then for asymptotic ratios that are not necessarily rational, the latter through g.a.c.s. and a functional analytic description of the limiting distribution [1, 2, 4]. Thus, for unilevel Toeplitz blocks, the asymptotic distributions are available for both rational and irrational limiting block-size ratios. Irregular block structures arise in scientific computing. Unequal and rectangular block geometries occur, for example, in coupled or interface PDE discretizations and in incomplete-data and image-restoration models [22, 54, 58, 61, 33, 34, 43].

In Chapter 3 we address part of the computational problem left by these distributional descriptions. The chapter constructs an implementable approximation that retains the relevant asymptotic distribution and belongs to a structured algebra that is sparser and suitable for fast application and inversion. Classical work on Toeplitz-generated algebras, circulant and multilevel approximations, spectral localization, and non-Hermitian preconditioners supplies the basic building blocks for the unequal-block construction and explains how such approximants can support preconditioning, see [152, 160, 57, 36, 35, 150, 149, 45]. The asymptotic distribution and clustering of matrix sequences provide diagnostics and heuristics for Krylov preconditioning, but they do not by themselves imply a dimension-independent Krylov convergence bound. Choosing a preconditioner for a specific matrix structure requires further analytical tools, particularly for localizing the spectrum and estimating the number and asymptotic behavior of outliers (see, e.g., the analysis of their influence on preconditioned conjugate gradient convergence in [7]). For the localization results, techniques based on linear positive operators and dyadic decomposition can be used [159]. In this thesis, the rank–norm decomposition in Theorem 2.11 and the bounded-rank refinement in Theorem 3.7 are used to study zero distribution and strong singular-value clustering of the approximation error, respectively. In particular, it is known that multilevel fast matrix-algebra preconditioners for multilevel structures stemming from multidimensional problems need not yield superlinear or optimal convergence rates in general [154, 133]. Recent related work includes GLT analyses of block Jacobi and Gauss-Seidel methods and Schwarz preconditioners, which begin with

sequences already carrying a GLT symbol. The notion of generalized locally Hankel sequences also studies related objects that appear implicitly in Chapter 3 [80, 143].

1.2.2 Quaternion Toeplitz structures

In finite-dimensional quaternion linear algebra, the eigenvalue problem differs from its complex counterpart. Left and right eigenvalues are distinct, and quaternion right eigenvalues, the ones most used in practice, occur in similarity classes rather than as finite isolated scalars. An asymptotic spectral analysis of quaternion matrix sequences therefore calls for a canonical choice of spectral representatives. This thesis uses canonical representatives in the closed upper half-plane of a fixed complex slice, selected from the spectrum of the standard complex symplectic representation described in [115, 175, 144]. Nonreal representatives are selected with positive imaginary part; each real eigenvalue is retained with half its even algebraic multiplicity in the doubled complex representation. The representation is compatible with matrix products and adjoints and relates quaternion singular values and canonical right eigenvalues to the corresponding spectral quantities of the embedded complex matrix. Quaternion linear algebra provides compact, geometry-preserving models for coupled multichannel data in signal and image processing; see the overview [127] and representative matrix applications [104, 38, 136, 119].

Quaternion spectral analysis is also motivated by potential applications to preconditioning. Recent results on structure-preserving quaternion Krylov methods and circulant block-diagonalization provide finite-dimensional computational tools [103, 116, 117, 137]. QCGNR methods with quaternion circulant or block-circulant preconditioners have been developed for quaternion Toeplitz and block-Toeplitz linear systems [98], while preconditioned QGMRES has been studied for broader structured quaternion coefficient matrices [97]. Recent spectral work also treats scalar, unilevel, Hermitian quaternion Toeplitz sequences through a 2×2 complex representation and develops a Strang-circulant PCG method [118]. Other studies concern finite-dimensional quaternion Toeplitz algebras [107] or the classification of normal Toeplitz matrices over commutative quaternions [106]. So far, the relevant literature does not cover block multilevel Weyl distributions for matrix-valued integrable symbols or an a.c.s. theory for quaternion matrix-sequences. Within the single-axis setting, Chapter 4 identifies the exact complex matrix-valued symbols for the left, right, and sandwich Toeplitz constructions. We then derive singular-value distributions, canonical-eigenvalue distributions under the stated additional assumptions, Hermitian localization results, Schatten estimates, and right-circulant approximations with the required doubled normalization.

1.2.3 Hilbert transforms and prescribed-node quadrature

Numerical methods for Cauchy principal-value integrals and Hilbert transforms include singularity subtraction, specialized quadrature formulas, and spectral or orthogonal-polynomial constructions. Representative methods for periodic, circular, and related integral transforms are given in [37, 60, 59, 126, 134, 165]. On the unit circle, Szegő rules are generated by zeros of para-orthogonal polynomials; their nodes can be computed from unitary Hessenberg matrices, and the free para-orthogonal parameter can be used to prescribe a node [105, 92, 32, 30]. Szegő-Lobatto rules extend prescribed-node constructions while retaining maximal Laurent-polynomial exactness [101, 31]. Unlike the analogous Gaussian rules on the real line, they have a free para-orthogonal parameter, which makes the analysis of node behavior and quadrature weights more delicate.

Anti-Szegő rules form a second construction that can be paired with Szegő rules. Under the

assumptions of the original construction, the two errors satisfy an opposite-error relation on specified Laurent spaces. This relation motivates averaged rules and numerical approximation of quadrature errors [108, 114]. However, in strongly singular integral approximation, exactness and opposite-error relations control quadrature functionals rather than the geometry of the nodes relative to the kernel singularity at the evaluation point. Likewise, prescribed-node theory for an individual rule does not by itself separate both interlaced node sets from that singularity.

The previous results leave open a singularity-dependent prescription for the paired Szegő-anti-Szegő construction and its pointwise stability and quadrature-error analysis. In Chapter 5 we choose the free para-orthogonal parameters as functions of the evaluation point so that both node sets remain separated from the kernel singularity at the natural n^{-1} scale, providing the node separation needed for numerical stability. Under the stated assumptions, this separation yields logarithmic pointwise stability and quadrature-error estimates.

1.2.4 Anti-Gauss quadrature and weighted interpolation

Gaussian quadrature links orthogonal polynomials, Jacobi matrices, and maximal polynomial exactness for quadrature rules on the real line [83, 84, 166]. Laurie introduced the anti-Gauss companion, whose error has the same magnitude and opposite sign as the Gauss error on a prescribed polynomial space [114]. Generalized anti-Gauss and averaged Gauss constructions extend this principle to other degrees, measures, and applications [139, 163, 164, 140, 141]; the survey [53] reviews this literature.

Interpolation introduces requirements not contained in quadrature exactness. Faber's theorem and later work on optimal or near-optimal node systems explain the role of Lebesgue constants. On bounded intervals, logarithmic growth is the optimal order, while weighted uniform-space theory gives Jacobi-based constructions and convergence results [65, 173, 123, 124, 122]. Christoffel-Darboux kernels and barycentric formulas provide exact structured representations of classical interpolants [130, 20, 95]. A directly related anti-Gauss interpolation result concerns Laguerre polynomials and truncated weighted processes on the positive half-line [69], rather than the bounded Jacobi setting of Chapter 6.

Exact representations of the interpolant. The anti-Gauss quadrature literature determines the nodes, weights, and quadrature functionals, but it does not provide anti-Gauss-specific Christoffel-Darboux, barycentric, or orthogonal-expansion representations of the associated Lagrange interpolant, nor does it establish their numerical stability. Chapter 6 derives these three exact forms and compares their finite-precision behavior [46]. The three formulas represent the same interpolation polynomial in exact arithmetic. In the reported tests, these forms return finite errors at degrees for which the direct nodal product returns `NaN`.

Weighted convergence at anti-Gauss Jacobi nodes. Interlacing and the asymptotic nodal distribution are likewise insufficient for convergence of weighted anti-Gauss interpolation. The analysis also requires internality, extreme-node and local-spacing estimates, asymptotic bounds for the relevant polynomials and derivatives in uniform norm, and control of the weighted Lebesgue function. For anti-Gauss Jacobi interpolation schemes, Chapter 6 establishes these properties, proves sufficient conditions on the Jacobi parameters for obtaining an $O(\log n)$ upper bound for the weighted Lebesgue constants, and derives weighted convergence estimates [47].

1.3 Main contributions and scope

The four core chapters cover five studies. Table 1.1 summarizes the main theoretical result, numerical consequence, and scope of each work. The corresponding core chapters give the complete hypotheses and statements.

The five studies are based on the published work on blocking approximations [9], the quaternion Toeplitz manuscript [112], the published work on prescribed Hilbert-transform quadrature [68], and the two anti-Gauss studies [46, 47].

1.4 Organization of the thesis

In Chapter 2 we introduce the notation for spectral distributions and clustering, a.c.s. convergence, the GLT calculus, orthogonal-polynomial and quadrature tools, and the weighted approximation spaces used in the subsequent chapters.

Part I contains the structured-matrix chapters. In Chapter 3 we develop blocking approximations, preconditioning results, and numerical experiments. Chapter 4 contains the quaternion Toeplitz theory, its numerical tests, and the right-circulant approximation construction.

Part II concerns quadrature and interpolation. Chapter 5 treats paired and prescribed-node quadrature for the circular Hilbert transform. In Chapter 6 we first derive exact structured representations of the anti-Gauss interpolant and then develop weighted interpolation at anti-Gauss Jacobi zeros. Chapter 7 summarizes the principal results, scope, limitations, and directions for future research.

Table 1.1: Main theoretical results, numerical consequences, and scope.

Problem	**Main theoretical result**	**Numerical consequence**	**Scope**
irregular-block Toeplitz approximation and preconditioning	For a fixed block count and rational limiting ratios, the stated blockwise approximations give a zero-distributed error sequence $\{S_n - A_n\}_n$. Continuity and strong blockwise error clustering yield strong error clustering. Rank assumptions give weak preconditioned singular-value clustering at 1; the eigenvalue conclusion needs the stated Hermitian setting and an HPD preconditioner.	Ordered singular-value and eigenvalue plots illustrate the proved clustering, and the reported Krylov comparisons for HPD, Hermitian, non-Hermitian, and rectangular cases show that the preconditioners are effective in practice.	We assume rational limiting proportions, a strong hypothesis, and restrict the Toeplitz blocks to the unilevel case. Spectral localization results are not proved. The two-level test is only exploratory.
Quaternion block multilevel Toeplitz embeddings	The complex symplectic representation identifies matrix-valued symbols for the left, right, and sandwich single-axis Fourier kernels. Under the stated hypotheses, the analysis establishes singular-value and canonical-eigenvalue distributions, Hermitian localization results, Schatten estimates, and right-circulant approximations with doubled normalization.	Qualitative comparisons of rearranged symbol values with eigenvalues or singular values cover continuous and L^1 symbols and all three constructions.	The analysis is exhaustive but proves no extremal asymptotics. A guide to constructing practical optimal preconditioners and a complete quaternion GLT calculus are under development.
Circular Hilbert-transform quadrature	Opposite Szegő and anti-Szegő errors cancel under averaging on the stated Laurent-polynomial space. A quarter-step prescription separates both node sets from the kernel singularity by distances comparable to $1/n$. Under the stated measure, Dini-continuity, prescribed-node, and approximation assumptions, the discrete operators satisfy $O(\log n)$ pointwise bounds and quadrature-error bounds. The stated Sobolev rate is an upper estimate.	The examples illustrate the loss of accuracy caused by an unconstrained node approaching the singularity, the effect of the prescribed-node sets, the improvement from averaging, and the paired-rule difference as a numerical error indicator. Some reported errors use a high-order numerical reference rather than a closed-form exact value.	More complex measures, the application to Cauchy integral equations, and analysis of the structured matrices generated by this scheme lie outside the chapter.
Anti-Gauss interpolant representations	The modified Christoffel-Darboux, barycentric, and orthogonal-expansion formulas are exact representations of the same anti-Gauss Lagrange interpolant.	The reported double-precision experiment compares the finite-precision behavior of the three structured formulas with that of the direct nodal product.	The chapter does not establish a general forward- or backward-stability theorem.
Anti-Gauss Jacobi interpolation	A rank-two Jacobi-matrix comparison gives weak convergence of the empirical zero-counting measures to the arc-sine measure. Local spacing, endpoint, polynomial, derivative, and Christoffel estimates yield an $O(\log n)$ upper bound for the weighted Lebesgue constants and weighted convergence estimates in the stated shifted parameter window.	The interpolation-error table illustrates convergence and the effect of function smoothness in the tested cases.	The endpoint-weight conditions are sufficient for the stated logarithmic upper bound. Necessity, matching lower bounds, and a classification of the admissible parameter region remain open.

Chapter 2

Matrix Sequences and Approximation Theory

2.1 Analytical framework and conventions

The analysis of structured numerical methods in this thesis relies on two forms of approximation. For matrix sequences, the central question is whether an approximation preserves the limiting singular-value or eigenvalue distribution. For quadrature and interpolation, it is whether exactness on a finite-dimensional space, together with suitable stability bounds, yields convergence.

Sections 2.2–2.4 develop the first framework: spectral distributions, approximating classes of sequences, Toeplitz and related matrix models, and the generalized locally Toeplitz (GLT) calculus. These results are stated for complex matrix sequences; Chapter 4 establishes the embeddings and spectral conventions needed to apply them to quaternion sequences. Sections 2.5–2.7 develop best-approximation estimates and the Gauss, Szegó, and anti-quadrature constructions used later.

The two frameworks use different notions of convergence. Both identify approximation errors whose effect can be controlled. Section 2.8 explains how each research chapter uses these tools.

Throughout the chapter, we use the following conventions. Given a rectangular matrix

$$A_n \in \mathbb{C}^{p_n \times q_n}, \qquad r_n = \min\{p_n, q_n\},$$

we call r_n the active singular-value dimension. Multilevel limits are understood with $\min_j n_j \to \infty$. Symbol domains have finite positive measure, and the measurable symbols are identified up to equality almost everywhere. The measure used for distribution formulas is the normalized Lebesgue measure unless stated otherwise, while the matrix-valued L^p norms in Definition 2.1 instead use unnormalized integrals.

2.2 Spectral distributions and matrix-sequence approximation

Spectral distributions describe the limiting behavior of normalized singular-value and eigenvalue averages. After introducing these notions, we identify perturbations that preserve them and approximation processes that allow distributions to be obtained from simpler matrix sequences.

2.2.1 Matrix norms and measurable symbols

We first fix the matrix norms and measurability conventions used in the spectral distribution framework. Let $D \subseteq \mathbb{R}^\ell$ be measurable, with $0 < \mu_\ell(D) < \infty$. A complex matrix-valued function is measurable when its entries are measurable. We write $L^0(D;\mathbb{C}^{a\times b})$ for these functions modulo equality almost everywhere.

Definition 2.1 (Schatten norms and matrix-valued L^p spaces). For $A \in \mathbb{C}^{a\times b}$, put $r = \min\{a,b\}$ and order its singular values as $\sigma_1(A) \geq \cdots \geq \sigma_r(A) \geq 0$. For $1 \leq p < \infty$, define

$$\|A\|_{S_p} = \left(\sum_{j=1}^{r} \sigma_j(A)^p\right)^{1/p}, \qquad \|A\|_{S_\infty} = \|A\| = \sigma_1(A).$$

The defined norms are unitarily invariant; S_1, S_2, and S_∞ give the trace, Frobenius, and operator norms, respectively. For a measurable $f : D \to \mathbb{C}^{a\times b}$, set

$$\|f\|_{L^p(D;\mathbb{C}^{a\times b})} = \left(\int_D \|f(x)\|_{S_p}^p \, \mathrm{d}x\right)^{1/p}, \quad 1 \leq p < \infty,$$
$$\|f\|_{L^\infty(D;\mathbb{C}^{a\times b})} = \operatorname*{ess\,sup}_{x\in D} \|f(x)\|.$$

The space $L^p(D;\mathbb{C}^{a\times b})$ consists of the corresponding equivalence classes modulo equality almost everywhere with finite norm. The integrals are not normalized by $\mu_\ell(D)$.

Lemma 2.2 (Spectral measurability). For every $a \in \mathbb{N}$, there exist Borel maps

$$\mu_1, \ldots, \mu_a : \mathbb{C}^{a\times a} \longrightarrow \mathbb{C}$$

whose values enumerate the eigenvalues counted with algebraic multiplicity. Ordered singular values and Schatten norms are continuous functions of the matrix values. Consequently, for a measurable matrix field, its ordered singular values, its Schatten norms, and every continuous symmetric function of its eigenvalues are measurable.

Proof. By [8, Corollary 2], a unitary Schur factor can be chosen as a Borel function $U(A)$ of a complex square matrix A. The diagonal elements of $U(A)^* A U(A)$ give the required maps μ_j. Composition with a measurable matrix field, followed by a continuous symmetric function, gives the eigenvalue assertion independently of the enumeration. Continuity of ordered singular values gives the remaining claims; see also [16, Lemma 2.5 and Remark 2.6]. □

2.2.2 Spectral distributions and clustering

We describe a spectral distribution by testing normalized averages against continuous functions of compact support. In this Weyl-type formulation, the limit is expressed as an integral involving a measurable symbol.

Definition 2.3. Let $D \subseteq \mathbb{R}^\ell$ be measurable, with $0 < \mu_\ell(D) < \infty$. For singular values, let $A_n \in \mathbb{C}^{p_n\times q_n}$, set $r_n = \min\{p_n, q_n\} \to \infty$, and let $f : D \to \mathbb{C}^{s\times t}$ be measurable. Set $r = \min\{s,t\}$. We write $\{A_n\}_n \sim_\sigma f$ if, for every $\psi \in C_c([0,\infty))$,

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_{j=1}^{r_n} \psi(\sigma_j(A_n)) = \frac{1}{\mu_\ell(D)} \int_D \frac{1}{r} \sum_{j=1}^{r} \psi(\sigma_j(f(x))) \, \mathrm{d}x.$$

Pertaining to eigenvalues, let $B_n \in \mathbb{C}^{d_n\times d_n}$, with $d_n \to \infty$, and let $g : D \to \mathbb{C}^{s\times s}$ be measurable. We write $\{B_n\}_n \sim_\lambda g$ if, for every $\psi \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{j=1}^{d_n} \psi(\lambda_j(B_n)) = \frac{1}{\mu_\ell(D)} \int_D \frac{1}{s} \sum_{j=1}^{s} \psi(\lambda_j(g(x))) \, \mathrm{d}x.$$

The domain is suppressed when it is clear from the context, and measurable symbols are identified up to equality almost everywhere.

Remark 2.4. If $\varphi : D \to D$ is measurable and measure-preserving, then replacing a symbol f by $f \circ \varphi$ leaves its distribution unchanged, in either the singular-value or eigenvalue sense. Indeed, the integrals against all the test functions in Definition 2.3 are unchanged. This observation also applies when the admissible test functions are restricted by a symmetry condition.

A singular-value distribution does not in general determine an eigenvalue distribution. The latter requires a separate argument; Hermitian structure will provide a sufficient hypothesis in several results below.

Eigenvalues are treated as multisets since they are counted with algebraic multiplicities. Singular values are ordered nonincreasingly in the theoretical analysis and nondecreasingly in the numerical computations. In numerical tests, we also order real eigenvalues nondecreasingly. In the standard fixed-block setting, $p_n = sN_n$ and $q_n = tN_n$.

Definition 2.5. Given a closed set S and $\varepsilon > 0$, let

$$B(S, \varepsilon) = \{z : \operatorname{dist}(z, S) < \varepsilon\}.$$

Let $A_n \in \mathbb{C}^{p_n \times q_n}$, set $r_n = \min\{p_n, q_n\} \to \infty$, and let $S \subseteq [0, \infty)$ be closed. The sequence $\{A_n\}_n$ is weakly clustered at S in the singular-value sense if, for every $\varepsilon > 0$,

$$\#\{1 \le j \le r_n : \sigma_j(A_n) \notin B(S, \varepsilon)\} = o(r_n).$$

It is instead strongly clustered at S in the singular-value sense if the same count is $O(1)$.

If $p_n = q_n = d_n$, with $d_n \to \infty$, and $S \subset \mathbb{C}$ is closed, then the relation below defines weak and strong eigenvalue clustering

$$\#\{1 \le j \le d_n : \lambda_j(A_n) \notin B(S, \varepsilon)\} = o(d_n)$$

and $O(1)$, respectively, for every $\varepsilon > 0$.

A constant scalar eigenvalue symbol c gives weak eigenvalue clustering at c. For a constant scalar singular-value symbol c, the corresponding cluster point is $|c|$. Similarly, a constant matrix-valued symbol C gives weak clustering at $\{\sigma_j(C)\}$ in the singular-value sense and at $\{\lambda_j(C)\}$ in the eigenvalue sense. These conclusions allow $o(r_n)$ outliers in the rectangular singular-value case and $o(d_n)$ in the square eigenvalue case. Strong clustering requires a bound on the number of outliers that is independent of the matrix size.

Definition 2.6 (Spectral essential ranges). Given a measurable matrix-valued function $f : D \to \mathbb{C}^{s \times t}$, with $r = s \wedge t$, we define its singular-value essential range by

$$\mathrm{ER}_\sigma(f) := \left\{ \rho \in [0, \infty) : \mu_\ell\left(\left\{ x \in D : \min_{1 \le j \le r} |\sigma_j(f(x)) - \rho| < \varepsilon \right\} \right) > 0 \text{ for every } \varepsilon > 0 \right\}.$$

For a measurable square symbol $g : D \to \mathbb{C}^{s \times s}$, define its spectral essential range by

$$\mathrm{ER}_\lambda(g) := \left\{ z \in \mathbb{C} : \mu_\ell\left(\left\{ x \in D : \min_{1 \le j \le s} |\lambda_j(g(x)) - z| < \varepsilon \right\} \right) > 0 \text{ for every } \varepsilon > 0 \right\}.$$

The spectral definition agrees with the labeling-free matrix-valued essential range in [88, Definition 4.1].

Remark 2.7. For any measurable enumeration $\lambda_1, \dots, \lambda_s$ of the eigenvalues of g, supplied by Lemma 2.2,

$$\mathrm{ER}_\lambda(g) = \bigcup_{j=1}^{s} \operatorname{ess\,range} \lambda_j.$$

By definition, each set on the right is contained in the left. Conversely, if z is outside every essential range of the eigenvalues, choose $\varepsilon_j > 0$ for which $\{x : |\lambda_j(x) - z| < \varepsilon_j\}$ has measure zero. With $\varepsilon = \min_j \varepsilon_j$, the spectral-distance set is contained in a finite union of these null sets. Thus z is also outside $\mathrm{ER}_\lambda(g)$. In particular, the definition does not depend on the specific enumeration of the spectrum.

A point is said to strongly attract the eigenvalues or singular values with infinite order if, after ordering them by their distance from that point, the j-th distance tends to zero for every fixed j. The next theorem connects this property with the essential range of the symbol. Earlier results on distribution and clustering appear in [169]; the scalar and matrix-valued infinite-order attraction statements are [88, Theorems 2.4 and 4.2]. The singular-value assertion follows by applying the scalar statement to $(x, j) \mapsto \sigma_j(f(x))$ on the disjoint union of $s \wedge t$ copies of D; see also the singular-value distribution framework in [158, Theorem 2.1].

Theorem 2.8. Let $D \subseteq \mathbb{R}^\ell$ have finite positive measure.

(i) If $\{A_n\}_n \sim_\sigma f$, then $\{A_n\}_n$ is weakly clustered at $\mathrm{ER}_\sigma(f)$ in the singular-value sense, and every point of $\mathrm{ER}_\sigma(f)$ strongly attracts the singular values with infinite order.

(ii) If $\{B_n\}_n \sim_\lambda g$, then $\{B_n\}_n$ is weakly clustered at $\mathrm{ER}_\lambda(g)$, and every point of $\mathrm{ER}_\lambda(g)$ strongly attracts the eigenvalues with infinite order.

The localization consequence below is immediate from Theorem 2.8: every point of the symbol's essential range is approached by eigenvalues, all of which belong to S.

Corollary 2.9. If $\{A_n\}_n \sim_\lambda f$ for a scalar symbol and $\mathrm{spec}(A_n) \subseteq S$ for every n, then

$$\operatorname{ess\,range}(f) \subseteq \overline{S}.$$

2.2.3 Zero-distributed sequences

Boundary modifications, finite-section corrections, and dimension adjustments often produce perturbations whose rank is negligible relative to the matrix size or whose norm approaches zero. Zero-distributed sequences provide a common description of these perturbations.

Definition 2.10. Let $Z_n \in \mathbb{C}^{p_n \times q_n}$, with $r_n = \min\{p_n, q_n\} \to \infty$. The sequence $\{Z_n\}_n$ is zero-distributed if

$$\{Z_n\}_n \sim_\sigma 0.$$

The following characterization separates a zero-distributed perturbation into a term of negligible relative rank and a term of vanishing norm. For the square rank–norm characterization, see [16, Theorem 2.18]; see also [153] for the Schatten sufficient criterion. For completeness, we give the proof for rectangular matrices.

Theorem 2.11. The sequence $\{Z_n\}_n$ is zero-distributed if and only if it admits a decomposition

$$Z_n = R_n + N_n, \qquad \frac{\operatorname{rank} R_n}{r_n} \to 0, \qquad \|N_n\| \to 0.$$

Moreover, for every $1 \le p < \infty$, the condition

$$\|Z_n\|_{S_p}^p = o(r_n)$$

is sufficient for the zero distribution property. For $p = \infty$, the corresponding sufficient condition is $\|Z_n\| = o(1)$.

Proof. Write the singular values in nonincreasing order. The defining convergence to zero is equivalent to

$$\frac{1}{r_n}\#\{j : \sigma_j(Z_n) > \varepsilon\} \longrightarrow 0 \qquad \text{for every } \varepsilon > 0.$$

For the forward implication, choose $\chi_\varepsilon \in C_c([0,\infty))$ with $0 \le \chi_\varepsilon \le 1$, $\chi_\varepsilon(0) = 1$, and $\chi_\varepsilon = 0$ on $[\varepsilon, \infty)$. Then

$$\frac{1}{r_n}\#\{j : \sigma_j(Z_n) \ge \varepsilon\} \le 1 - \frac{1}{r_n}\sum_{j=1}^{r_n} \chi_\varepsilon(\sigma_j(Z_n)) \longrightarrow 0.$$

Conversely, for $\varphi \in C_c([0,\infty))$, split the average over $\sigma_j(Z_n) \le \varepsilon$ and its complement, and then let $n \to \infty$ followed by $\varepsilon \downarrow 0$. Uniform continuity of φ gives convergence to $\varphi(0)$.

Assume first that $\{Z_n\}_n$ is zero-distributed. A diagonal choice gives numbers $\varepsilon_n \downarrow 0$ such that

$$k_n := \#\{j : \sigma_j(Z_n) > \varepsilon_n\} = o(r_n).$$

In a singular value decomposition of Z_n, retain in R_n precisely the singular values above ε_n, and place the remaining terms in N_n. Then $\operatorname{rank} R_n = k_n$ and $\|N_n\| \le \varepsilon_n$.

Conversely, suppose that $Z_n = R_n + N_n$, with $\operatorname{rank} R_n = o(r_n)$ and $\|N_n\| \to 0$. If $\|N_n\| < \varepsilon$, the singular-value inequality for a rank $\operatorname{rank} R_n$ perturbation gives

$$\#\{j : \sigma_j(Z_n) > \varepsilon\} \le \operatorname{rank} R_n.$$

The threshold criterion above therefore yields zero distribution. Finally, for $1 \le p < \infty$,

$$\#\{j : \sigma_j(Z_n) > \varepsilon\} \le \varepsilon^{-p}\|Z_n\|_{S_p}^p,$$

which proves the Schatten sufficient condition; the case $p = \infty$ is immediate. □

If $Z_n = Z_n^*$, the rank-norm decomposition may be replaced by its Hermitian parts. The resulting zero-distributed perturbation principle has the square block formulations given in [16, Theorems 2.34–2.35 and Remark 2.36].

Corollary 2.12. Let $A_n, Z_n \in \mathbb{C}^{p_n \times q_n}$, with $r_n = \min\{p_n, q_n\} \to \infty$. If $\{A_n\}_n \sim_\sigma f$ and $\{Z_n\}_n \sim_\sigma 0$, then

$$\{A_n + Z_n\}_n \sim_\sigma f.$$

If $A_n, Z_n \in \mathbb{C}^{d_n \times d_n}$ are Hermitian, $\{A_n\}_n \sim_\lambda g$, and $\{Z_n\}_n \sim_\sigma 0$, then

$$\{A_n + Z_n\}_n \sim_\lambda g.$$

Throughout the thesis, *zero-distributed* refers to the singular-value relation $\sim_\sigma 0$. The threshold criterion in the proof identifies this with weak singular-value clustering at zero; strong clustering imposes the additional requirement of a uniformly bounded outlier count.

2.2.4 Sparse singular-value tails

To control products, we need to limit the proportion of large singular values. For pseudoinverses, the corresponding quantity to control is the proportion of singular values near zero. Sparse unboundedness and sparse vanishing make these two conditions precise, without requiring uniform bounds on every singular value.

Definition 2.13. Let $A_n \in \mathbb{C}^{p_n \times q_n}$, with $r_n = \min\{p_n, q_n\} \to \infty$. The sequence $\{A_n\}_n$ is *sparsely unbounded* (s.u.) if

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{1}{r_n}\#\{1 \le j \le r_n : \sigma_j(A_n) > M\} = 0. \tag{2.1}$$

It is *sparsely vanishing* (s.v.) if

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{1}{r_n} \#\{1 \le j \le r_n : \sigma_j(A_n) < M^{-1}\} = 0. \tag{2.2}$$

For a multi-indexed family $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}}$, the same conditions are required along every scalar sequence $\boldsymbol{n} = \boldsymbol{n}(n)$ such that $\min_{1\le\ell\le d} n_\ell \to \infty$.

Sparse unboundedness for matrix sequences is introduced in [152]. The relation between these tail conditions and a square matrix-valued symbol is given in [17, Propositions 2.20 and 2.25]. In the rectangular formulation below, maximal rank replaces invertibility. The argument uses the singular-value branches in the defining C_c-relations, normalized by the active dimension.

Proposition 2.14. Let $D \subseteq \mathbb{R}^\ell$ be measurable, with $0 < \mu_\ell(D) < \infty$, and let $f : D \to \mathbb{C}^{s\times t}$ be measurable. Set $r = \min\{s,t\}$, and suppose that $\{A_n\}_n \sim_\sigma f$. Then $\{A_n\}_n$ is sparsely unbounded. Moreover, $\{A_n\}_n$ is sparsely vanishing if and only if

$$\operatorname{rank} f(x) = r \qquad \text{for almost every } x \in D.$$

In particular, if $\{A_n\}_n$ is sparsely vanishing, then $\{A_n^\dagger\}_n$ is sparsely unbounded.

Rectangular products require the active dimensions of the factors and product to remain comparable.

Definition 2.15. For positive-integer sequences $a = (a_n)$, $b = (b_n)$, and $c = (c_n)$, write

$$\mathrm{PC}(a,b,c) \quad\Longleftrightarrow\quad (a_n \wedge b_n) + (b_n \wedge c_n) = O(a_n \wedge c_n). \tag{2.3}$$

This condition compares the active dimensions of two composable factors with that of their product. It is automatic for square sequences and for fixed block sizes over a common grid.

The product-compatibility condition above and the corresponding varying-aspect-ratio product rule are stated explicitly in [112, Section 3.1 and Theorem A.9].

2.2.5 Approximating classes of sequences

An approximating class of sequences is defined through a doubly indexed family. For square sequences, the low-rank-plus-small-norm definition is given in [152, Definition 2.1]. The active-dimensional rectangular extension is introduced in [15, Definition 2.6].

Definition 2.16. Let $A_n \in \mathbb{C}^{p_n\times q_n}$, set $r_n = \min\{p_n, q_n\} \to \infty$, and let $\{\{B_{n,m}\}_n\}_m$ have compatible dimensions. We write

$$\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$$

if, for every fixed m, all sufficiently large n admit

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \tag{2.4}$$

$$\operatorname{rank} R_{n,m} \le c(m) r_n, \qquad \|N_{n,m}\| \le \omega(m), \tag{2.5}$$

where $c(m) \to 0$ and $\omega(m) \to 0$ as $m \to \infty$.

For each fixed m, the rank and norm bounds hold for all sufficiently large matrix sizes. The limit $n \to \infty$ is taken first, and then $m \to \infty$ removes the remaining defects. This order of limits is part of the definition.

To transfer spectral distributions through an approximating class, we also need convergence of the associated symbols. The appropriate notion is convergence in measure.

Definition 2.17. Let $D \subseteq \mathbb{R}^\ell$ be measurable, with $0 < \mu_\ell(D) < \infty$, and let $g_m, f : D \to \mathbb{C}^{s\times t}$ be measurable, where s and t are fixed. We say that $g_m \to f$ in measure on D if, for every $\varepsilon > 0$,

$$\mu_\ell\big(\{x \in D : \|g_m(x) - f(x)\| > \varepsilon\}\big) \to 0.$$

For scalar functions, the matrix norm is replaced by the absolute value. Since the matrix sizes are fixed, the definition is independent of the chosen matrix norm.

The following rules transfer distributions from approximants to their limits and describe how approximants may be combined. The transfer, Hermitian realization, block assembly, and Schatten criteria are part of the complex a.c.s. calculus in [152] and [77, Chapter 5], with the active-dimensional rectangular normalization of [15, Definition 2.6].

Products require particular care because an unbounded factor can amplify a small-norm error. For the common-size argument, see [77, Proposition 5.5, pp. 87–88]; fixed-block rectangular multiplication is treated in [15, Theorem 4.5]. The product rule under PC, allowing varying aspect ratios, uses the extension in [112, Section 3.1 and Theorem A.9]. Its proof below separates bounded factors from terms of negligible relative rank.

ACS 1. Let $D \subseteq \mathbb{R}^\ell$ have finite positive measure. Then $\{A_n\}_n \sim_\sigma f$ if and only if there exist symbols f_m and matrix-sequences $\{B_{n,m}\}_n$ such that

$$\{B_{n,m}\}_n \sim_\sigma f_m, \qquad \{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n, \qquad f_m \to f \quad \text{in measure on } D.$$

ACS 2. Suppose that every A_n is Hermitian and that $f : D \to \mathbb{C}^{s\times s}$ is Hermitian almost everywhere. Then $\{A_n\}_n \sim_\lambda f$ if and only if there exist Hermitian matrix-sequences $\{B_{n,m}\}_n$ and Hermitian symbols f_m such that

$$\{B_{n,m}\}_n \sim_\lambda f_m, \qquad \{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n, \qquad f_m \to f \quad \text{in measure on } D.$$

ACS 3. Let

$$A_n, B_{n,m} \in \mathbb{C}^{d_n\times e_n}, \qquad A'_n, B'_{n,m} \in \mathbb{C}^{d'_n\times e'_n},$$

and assume

$$\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n, \qquad \{B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A'_n\}_n.$$

Then

$$\{B^*_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A^*_n\}_n.$$

If $(d'_n, e'_n) = (d_n, e_n)$, then, for every $\alpha, \beta \in \mathbb{C}$,

$$\{\alpha B_{n,m} + \beta B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{\alpha A_n + \beta A'_n\}_n.$$

If $d'_n = e_n$, the target sequences are s.u., and $\mathrm{PC}(d, e, e')$, then

$$\{B_{n,m}B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_nA'_n\}_n.$$

The same conclusion holds for multiplication on either side by a fixed compatible s.u. sequence under the corresponding product-compatibility condition.

Here a fixed multiplier may depend on the matrix size n, but not on the approximation index m.

Product assertions in ACS 3. We make the varying-dimension rank estimate explicit. Write $C_n = A'_n$, $D_{n,m} = B'_{n,m}$, and $f_n = e'_n$, and set

$$r_n = d_n \wedge e_n, \qquad s_n = e_n \wedge f_n, \qquad t_n = d_n \wedge f_n.$$

Product compatibility gives $r_n + s_n \le \Gamma t_n$ eventually, for some $\Gamma > 0$. Choose common a.c.s. controls $c(m), \omega(m) \to 0$ and write

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \qquad C_n = D_{n,m} + R'_{n,m} + N'_{n,m},$$

with ranks at most $c(m)r_n$, $c(m)s_n$, respectively, and both small-norm terms bounded by $\omega(m)$.

Sparse unboundedness and singular-value truncation give, for every fixed $M > 0$ and all sufficiently large n, decompositions

$$A_n = A^{\mathrm{bd}}_{n,M} + E_{n,M}, \qquad C_n = C^{\mathrm{bd}}_{n,M} + E'_{n,M},$$

where both bounded parts have norm at most M, and

$$\operatorname{rank} E_{n,M} \le \rho(M) r_n, \qquad \operatorname{rank} E'_{n,M} \le \rho(M) s_n, \qquad \rho(M) \longrightarrow 0.$$

This is the truncation used in [77, Propositions 5.3 and 5.5]. Suppressing the indices in the next identity, define

$$AC - BD = \underbrace{RD + (A-N)R' + EN' + NE'}_{\mathcal{R}} + \underbrace{A^{\mathrm{bd}}N' + NC^{\mathrm{bd}} - NN'}_{\mathcal{N}}.$$

Consequently,

$$\frac{\operatorname{rank}\mathcal{R}}{t_n} \le \Gamma\big(c(m) + \rho(M)\big), \qquad \|\mathcal{N}\| \le 2M\omega(m) + \omega(m)^2.$$

Choosing $M = M_m = (\omega(m) + m^{-2})^{-1/2}$ makes both bounds tend to zero. The threshold in n may depend on m and M_m, as allowed by the definition of an a.c.s. This proves the two-approximant product assertion without a fixed aspect ratio.

For fixed right multiplication, only the multiplier sequence $\{C_n\}_n$ must be sparsely unbounded. The same truncation gives

$$(A_n - B_{n,m})C_n = \underbrace{R_{n,m}C_n + N_{n,m}E'_{n,M}}_{\mathcal{R}} + \underbrace{N_{n,m}C^{\mathrm{bd}}_{n,M}}_{\mathcal{N}},$$

with $\operatorname{rank}\mathcal{R}/t_n \le \Gamma(c(m) + \rho(M))$ and $\|\mathcal{N}\| \le M\omega(m)$. Take $M = M_m$ again. Fixed left multiplication follows by taking adjoints. These arguments supply the product clauses of the rule; the remaining clauses are the standard a.c.s. calculus cited above. □

ACS 4. Assume a fixed compatible block partition. If each block family satisfies

$$\{B^{(m)}_{n,ij}\}_n \xrightarrow{\text{a.c.s.}} \{A_{n,ij}\}_n,$$

then the corresponding assembled block matrices satisfy

$$\{[B^{(m)}_{n,ij}]_{i,j}\}_n \xrightarrow{\text{a.c.s.}} \{[A_{n,ij}]_{i,j}\}_n.$$

This includes the fixed-block rectangular arrangements over a common grid used later in the thesis.

ACS 5. Let $A_n, B_{n,m} \in \mathbb{C}^{d_n \times e_n}$, set $r_n = d_n \wedge e_n \to \infty$, and let $p \in [1, \infty]$. Assume that, for every m, all sufficiently large n satisfy

$$\|A_n - B_{n,m}\|_{S_p} \le \varepsilon(m,n)\, r_n^{1/p},$$

with $r_n^{1/\infty} := 1$, and that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \varepsilon(m,n) = 0.$$

Then

$$\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n.$$

As an immediate consequence of ACS 1, if $\{A_n - B_{n,m}\}_n \sim_\sigma g_m$ on a fixed finite-measure domain and $g_m \to 0$ in measure, then

$$\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n. \tag{2.6}$$

2.2.6 Extradimensional compression and dimension transfer

Extradimensional compression compares a matrix sequence with an auxiliary sequence whose dimension differences are negligible relative to the active singular-value dimension. The auxiliary matrices may have a simpler structure, making their distribution easier to identify. The method is attributed to a 1997 discussion between Tyrtyshnikov and Serra-Capizzano in [11]. Spectral invariance upon compression is established in [125, Theorem 4.3 and Corollary 4.4]; rectangular extensions appear in [73, Theorem 2.4] and [1, Theorem 3.8].

We distinguish compression, which returns to the original matrix space, from ambient realization, which embeds that space into the larger one. This distinction is needed when applying algebraic operations to the auxiliary sequences.

Definition 2.18 (Extradimensional compression and ambient realization). Let

$$X_n \in \mathbb{C}^{P_n \times Q_n}, \qquad E_n \in \mathbb{C}^{P_n \times p_n}, \qquad F_n \in \mathbb{C}^{Q_n \times q_n},$$

where

$$E_n^* E_n = I_{p_n}, \qquad F_n^* F_n = I_{q_n}.$$

The isometries E_n and F_n are called *compression matrices*. The physical matrix

$$A_n := E_n^* X_n F_n \in \mathbb{C}^{p_n \times q_n} \tag{2.7}$$

is the (E_n, F_n)-*extradimensional compression* of X_n. Conversely, for any physical matrix $A_n \in \mathbb{C}^{p_n \times q_n}$, its (E_n, F_n)-*ambient realization* is

$$\mathcal{A}_n := E_n A_n F_n^* \in \mathbb{C}^{P_n \times Q_n}. \tag{2.8}$$

Write

$$\Pi_n^E := E_n E_n^*, \qquad \Pi_n^F := F_n F_n^*, \qquad \rho_n := \min\{P_n, Q_n\}.$$

The sequence of pairs $\{(E_n, F_n)\}_n$, and the associated compression or ambient realization, is called *asymptotically full* if

$$\rho_n \to \infty, \qquad (P_n - p_n) + (Q_n - q_n) = o(\rho_n). \tag{2.9}$$

If each compression matrix consists of distinct columns of the identity matrix, we call the operation *coordinate compression*.

Compression and ambient realization satisfy

$$E_n^*(E_n A_n F_n^*) F_n = A_n, \tag{2.10}$$

whereas performing the operations in the opposite order gives

$$E_n(E_n^* X_n F_n) F_n^* = \Pi_n^E X_n \Pi_n^F. \tag{2.11}$$

An asymptotically full compression preserves the limiting distribution, as stated in [1, Theorems 3.7 and 3.8].

Proposition 2.19 (Extradimensional principle). Let $A_n = E_n^* X_n F_n$, with the sequence of pairs $\{(E_n, F_n)\}_n$ asymptotically full. Then, for every measurable matrix-valued symbol f,

$$\{X_n\}_n \sim_\sigma f \quad \Longleftrightarrow \quad \{A_n\}_n \sim_\sigma f.$$

If, in addition, $P_n = Q_n$, $F_n = E_n$, and every X_n is Hermitian, then, for every measurable square matrix-valued symbol g,

$$\{X_n\}_n \sim_\lambda g \quad \Longleftrightarrow \quad \{A_n\}_n \sim_\lambda g.$$

Two sequences of different sizes can therefore be compared by realizing them in a common ambient space. If the two realizations differ by a zero-distributed sequence, they have the same distribution. The next corollary combines [1, Theorems 3.6–3.8] with Corollary 2.12.

Corollary 2.20 (Aligned extradimensional transfer). Let $A_n \in \mathbb{C}^{p_n \times q_n}$ and $\widehat{A}_n \in \mathbb{C}^{\widehat{p}_n \times \widehat{q}_n}$ admit asymptotically full ambient realizations

$$\mathcal{A}_n := E_n A_n F_n^*, \qquad \widehat{\mathcal{A}}_n := \widehat{E}_n \widehat{A}_n \widehat{F}_n^*,$$

both belonging to $\mathbb{C}^{P_n \times Q_n}$. If

$$\{\mathcal{A}_n - \widehat{\mathcal{A}}_n\}_n \sim_\sigma 0,$$

then, for every measurable matrix-valued symbol f,

$$\{A_n\}_n \sim_\sigma f \quad \Longleftrightarrow \quad \{\widehat{A}_n\}_n \sim_\sigma f.$$

If, in addition, $P_n = Q_n$, the physical matrices A_n and $\widehat{A}_n$ are Hermitian, and

$$F_n = E_n, \qquad \widehat{F}_n = \widehat{E}_n,$$

then, for every measurable square matrix-valued symbol g,

$$\{A_n\}_n \sim_\lambda g \quad \Longleftrightarrow \quad \{\widehat{A}_n\}_n \sim_\lambda g.$$

The two realizations need not use the same compression matrices.

The following finite-dimensional identities describe the algebra of ambient realizations. In particular, the product identity requires the two factors to use compatible compressions.

Lemma 2.21 (Algebraic identities for ambient realizations). Let

$$A, C \in \mathbb{C}^{p \times q}, \qquad B \in \mathbb{C}^{q \times r},$$

and let

$$E \in \mathbb{C}^{P \times p}, \qquad F \in \mathbb{C}^{Q \times q}, \qquad H \in \mathbb{C}^{R \times r},$$

with $E^*E = I_p$, $F^*F = I_q$, and $H^*H = I_r$. Define

$$\mathcal{A} := EAF^*, \qquad \mathcal{C} := ECF^*, \qquad \mathcal{B} := FBH^*.$$

Then, for all $\alpha, \beta \in \mathbb{C}$,

$$FA^*E^* = \mathcal{A}^*, \tag{2.12}$$

$$E(\alpha A + \beta C)F^* = \alpha\mathcal{A} + \beta\mathcal{C}, \tag{2.13}$$

$$E(AB)H^* = \mathcal{A}\mathcal{B}, \tag{2.14}$$

$$FA^\dagger E^* = \mathcal{A}^\dagger, \tag{2.15}$$

where $\dagger$ denotes the Moore–Penrose pseudoinverse.

If, in addition, $p = q$, $A = A^*$, and $F = E$, so that $\mathcal{A} = EAE^* \in \mathbb{C}^{P \times P}$, then every function $g : \mathrm{spec}(A) \cup \{0\} \to \mathbb{C}$, interpreted through spectral calculus, satisfies

$$g(\mathcal{A}) = Eg(A)E^* + g(0)(I_P - EE^*). \tag{2.16}$$

In particular, this identity applies to every continuous function $g : \mathbb{R} \to \mathbb{C}$.

Proof. The adjoint and linearity identities follow directly from the definitions. For the product, the common intermediate isometry F gives

$$\mathcal{AB} = EA(F^*F)BH^* = E(AB)H^*.$$

For the pseudoinverse, set $Z := FA^\dagger E^*$. Then

$$\mathcal{A}Z = EAA^\dagger E^*, \qquad Z\mathcal{A} = FA^\dagger AF^*.$$

Both matrices are Hermitian, since $AA^\dagger$ and $A^\dagger A$ are Hermitian. Moreover,

$$\mathcal{A}Z\mathcal{A} = E(AA^\dagger A)F^* = \mathcal{A}, \qquad Z\mathcal{A}Z = F(A^\dagger AA^\dagger)E^* = Z.$$

Thus Z satisfies the four Penrose equations for $\mathcal{A}$, and uniqueness of the Moore–Penrose pseudoinverse gives $Z = \mathcal{A}^\dagger$.

Finally, in the Hermitian case, complete E to a unitary matrix $\mathcal{U} = [\, E \; E_\perp \,] \in \mathbb{C}^{P\times P}$. Then

$$\mathcal{U}^*\mathcal{A}\mathcal{U} = A \oplus 0_{P-p}.$$

The finite-dimensional spectral calculus therefore yields

$$\mathcal{U}^* g(\mathcal{A})\mathcal{U} = g(A) \oplus g(0)I_{P-p}.$$

Conjugating by $\mathcal{U}$ and using $E_\perp E_\perp^* = I_P - EE^*$ proves (2.16). When $P = p$, the complementary blocks are absent. □

2.3 Toeplitz, Circulant, and Hankel sequences

The matrix models in this section provide the building blocks for the GLT calculus. Toeplitz matrices describe interactions determined by index differences, while diagonal sampling matrices describe spatial variation. Circulants replace nonperiodic shifts by periodic ones and admit exact Fourier diagonalization. Hankel matrices describe another type of boundary correction.

We first establish the distributions and norm estimates of these models. We then compare Toeplitz matrices with circulants through their boundary terms and identify the fixed-symbol Hankel corrections that are zero-distributed.

2.3.1 Toeplitz sequences and spectral distributions

In a Toeplitz matrix, each block depends only on the difference between its row and column indices. Fourier coefficients of a generating function supply the blocks. The following notation accommodates both multilevel indexing and rectangular blocks.

Definition 2.22. Let $d \ge 1$, let $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathbb{N}^d$ and $\boldsymbol{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$, and set

$$\Lambda_{\boldsymbol{n}} = \prod_{\ell=1}^{d} \{0, \dots, n_\ell - 1\}, \qquad N_{\boldsymbol{n}} = \prod_{\ell=1}^{d} n_\ell,$$

with the analogous definitions for $\Lambda_{\boldsymbol{m}}$ and $N_{\boldsymbol{m}}$. Let $f \in L^1([-\pi,\pi]^d; \mathbb{C}^{s\times t})$, with Fourier coefficients

$$\widehat{f}_{\boldsymbol{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} f(\boldsymbol{\theta}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{k},\boldsymbol{\theta})} \, \mathrm{d}\boldsymbol{\theta}, \qquad \boldsymbol{k} \in \mathbb{Z}^d. \tag{2.17}$$

The rectangular d-level block Toeplitz matrix generated by f is

$$T_{\boldsymbol{n},\boldsymbol{m}}(f) = [\, \widehat{f}_{\boldsymbol{i}-\boldsymbol{j}} \,]_{\boldsymbol{i}\in\Lambda_{\boldsymbol{n}},\, \boldsymbol{j}\in\Lambda_{\boldsymbol{m}}} \in \mathbb{C}^{sN_{\boldsymbol{n}}\times tN_{\boldsymbol{m}}}, \tag{2.18}$$

where the multi-indices are ordered lexicographically. When $\boldsymbol{n} = \boldsymbol{m}$, we write $T_{\boldsymbol{n}}(f) = T_{\boldsymbol{n},\boldsymbol{n}}(f)$. The function f is the Toeplitz generating function or symbol.

In the unilevel case, with zero-based indices,

$$T_{n,m}(f) = [\,\widehat{f}_{i-j}\,]_{\substack{0\le i\le n-1\\ 0\le j\le m-1}} \in \mathbb{C}^{sn\times tm}.$$

The notation $T_n(f)$ means $T_{n,n}(f)$. Fixed matrix blocks affect only the inner dimensions s and t.

The generating function determines the limiting singular-value distribution and, in the Hermitian case, the eigenvalue distribution. This is the multilevel block Szegő–Tilli theorem of [168], recalled in [16, Theorem 2.43]. The nonsquare-block statement is also covered by [15, Section 4.3.1 and Theorem 4.4]. Scalar one-level antecedents include [138, 6], with related multilevel developments in [170].

Theorem 2.23. If $f \in L^1([-\pi,\pi]^d;\mathbb{C}^{s\times t})$, then, as $\min_\ell n_\ell \to \infty$,

$$\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}} \sim_\sigma f.$$

If $s = t$ and f is Hermitian almost everywhere, then also

$$\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}} \sim_\lambda f.$$

For bounded non-Hermitian symbols, the block multilevel Tilli theorem in [57] gives an eigenvalue distribution under geometric conditions on the spectral essential range. We use Definition 2.6, representing the torus by $[-\pi,\pi)^d$.

Theorem 2.24 (Block multilevel Tilli theorem). Let $G \in L^\infty(\mathbb{T}^d;\mathbb{C}^{a\times a})$. If

$$\operatorname{int}\mathrm{ER}_\lambda(G) = \varnothing, \qquad \mathbb{C}\setminus \mathrm{ER}_\lambda(G) \text{ is connected},$$

then, as $\min_j n_j \to \infty$,

$$\{T_{\boldsymbol{n}}(G)\}_{\boldsymbol{n}} \sim_\lambda (G,\mathbb{T}^d).$$

To pass from approximation of a symbol to approximation of the Toeplitz matrices it generates, we use a Schatten norm estimate. The matrix-valued L^p norms are those of Definition 2.1, with unnormalized Lebesgue integrals.

The estimate is proved for scalar symbols in [160, Corollary 4.2] and extended to multilevel blocks in [155, Corollary 3.5]. Rectangular blocks are treated in Section 3.3, particularly Eqs. (34)–(35), of the latter paper; see also [16, Theorem 2.44].

Theorem 2.25.
Let $f \in L^p([-\pi,\pi]^d;\mathbb{C}^{s\times t})$, with $1 \le p \le \infty$, and let $N_{\boldsymbol{n}} = \prod_{\ell=1}^d n_\ell$. Then

$$\|T_{\boldsymbol{n}}(f)\|_{S_p} \le \left(\frac{N_{\boldsymbol{n}}}{(2\pi)^d}\right)^{1/p} \|f\|_{L^p([-\pi,\pi]^d;\mathbb{C}^{s\times t})}. \tag{2.19}$$

Here S_∞ denotes the spectral norm and $1/\infty = 0$. The constant 1 is optimal.

For the constant symbol $f(\boldsymbol{\theta}) = C$, one has $T_{\boldsymbol{n}}(f) = I_{N_{\boldsymbol{n}}} \otimes C$, and equality holds in (2.19); hence the constant 1 is optimal.

For Hermitian symbols, essential bounds also locate every eigenvalue of every Toeplitz matrix. This finite-dimensional conclusion complements the limiting distribution. For the broader theory and the asymptotics of extreme eigenvalues, see [147].

Theorem 2.26.
Let $f \in L^1([-\pi,\pi]^d;\mathbb{C}^{s\times s})$ be Hermitian almost everywhere, and set

$$m_f = \operatorname*{ess\,inf}_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\min}(f(\boldsymbol{\theta})), \qquad M_f = \operatorname*{ess\,sup}_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\max}(f(\boldsymbol{\theta})).$$

The endpoints may be infinite. For every $\boldsymbol{n} \in \mathbb{N}^d$, $T_{\boldsymbol{n}}(f)$ is Hermitian and

$$m_f \le \lambda_j(T_{\boldsymbol{n}}(f)) \le M_f, \qquad j = 1, \dots, sN_{\boldsymbol{n}},$$

where the inequalities are interpreted in the extended real line.

The strict scalar refinement is classical; see [93, Chapter 5, Section 2].

Theorem 2.27. Let $f \in L^1([-\pi,\pi])$ be real-valued, and set $m_f = \operatorname{ess\,inf} f$ and $M_f = \operatorname{ess\,sup} f$. Every eigenvalue of $T_n(f)$ lies in $[m_f, M_f]$, and the inclusion is strict when f is not almost everywhere constant. In particular, $m_f > 0$ implies that every $T_n(f)$ is positive definite.

2.3.2 Canonical and empirical sampling sequences

A diagonal sampling matrix places function values on its diagonal, or uses them as diagonal blocks when the function is matrix-valued. The canonical grid on the unit cube is the Cartesian product of the points i_ℓ/n_ℓ, $1 \le i_\ell \le n_\ell$, in each coordinate direction. Thus it consists of the right endpoints of the uniform subintervals in each direction. We first consider this grid and then allow node arrays with a prescribed limiting distribution.

Definition 2.28. Let $d \ge 1$, let $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathbb{N}^d$, and set $N_{\boldsymbol{n}} = \prod_{\ell=1}^d n_\ell$. Let $a : [0,1]^d \to \mathbb{C}^{s\times t}$ be finite at the canonical grid points. Define

$$D_{\boldsymbol{n}}(a) = \operatorname*{diag}_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} a(\boldsymbol{i}/\boldsymbol{n}),$$

where the quotient is coordinatewise, and the multi-indices are ordered lexicographically. Then $D_{\boldsymbol{n}}(a) \in \mathbb{C}^{sN_{\boldsymbol{n}}\times tN_{\boldsymbol{n}}}$. The definition alone does not assert an asymptotic distribution.

The canonical sampling rule below is recorded in [15, Section 4.3.2 and GLT 3 in Section 5]; for square blocks, see also [16, Theorems 4.11 and 4.23].

Proposition 2.29. Let $a : [0,1]^d \to \mathbb{C}^{s\times t}$ be a finite-valued measurable function that is entrywise continuous almost everywhere. Then, as $\min_\ell n_\ell \to \infty$,

$$\{D_{\boldsymbol{n}}(a)\}_{\boldsymbol{n}} \sim_\sigma a.$$

If $s = t$ and $a(x) = a(x)^*$ for every $x \in [0,1]^d$, then also

$$\{D_{\boldsymbol{n}}(a)\}_{\boldsymbol{n}} \sim_\lambda a.$$

In particular, these conclusions hold for entrywise Riemann-integrable functions, with the same Hermitian assumption for eigenvalues.

Proof of the distributional assertions. Set $r = \min\{s,t\}$. For $\varphi \in C_c([0,\infty))$, define

$$h_\varphi(x) = \frac{1}{r}\sum_{j=1}^{r} \varphi\big(\sigma_j(a(x))\big).$$

This function is bounded and continuous at every continuity point of a. It is therefore Riemann integrable, even when a itself is unbounded. The singular values of a block-diagonal matrix are those of its blocks, counted with multiplicity, so

$$\frac{1}{rN_{\boldsymbol{n}}}\sum_{j=1}^{rN_{\boldsymbol{n}}} \varphi\big(\sigma_j(D_{\boldsymbol{n}}(a))\big) = \frac{1}{N_{\boldsymbol{n}}}\sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} h_\varphi(\boldsymbol{i}/\boldsymbol{n}) \longrightarrow \int_{[0,1]^d} h_\varphi(x)\,\mathrm{d}x.$$

This gives the singular-value distribution. For Hermitian blocks, replace singular values by eigenvalues and take $\varphi \in C_c(\mathbb{R})$. Continuity of the ordered eigenvalues on Hermitian matrices gives the same conclusion. □

The almost-everywhere continuity requirement cannot be replaced by measurability alone. For example, $a = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ vanishes almost everywhere but gives $D_n(a) = I_n$. Thus arbitrary changes of a representative on a null set can change the sampling sequence and even its limiting distribution.

For noncanonical nodes, the limiting node measure takes the place of the uniform grid. The proposition below transports this measure through the sampled function and, for matrix-valued functions, through its singular values or Hermitian eigenvalues. It follows from the continuous mapping theorem [21, Theorem 2.7] and continuity of the relevant ordered spectral maps.

On a compact rectangle or torus, an array is *asymptotically uniform* in the empirical-measure sense when its averages of every continuous test function converge to integration against normalized Lebesgue measure. This describes the distribution of the nodes, not their ordering or their individual distances from canonical grid points.

Proposition 2.30. Let D be a compact metric space, let $\mathcal{X}_n = (x_{k,n})_{k=1}^{N_n} \in D^{N_n}$, with $N_n \to \infty$, and let ν be a Borel probability measure on D. Assume that

$$\frac{1}{N_n}\sum_{k=1}^{N_n} \psi(x_{k,n}) \longrightarrow \int_D \psi(x)\,\mathrm{d}\nu(x), \qquad \psi \in C(D).$$

If $a : D \to \mathbb{C}$ is finite-valued, Borel measurable, and continuous ν-almost everywhere, then

$$\frac{1}{N_n}\sum_{k=1}^{N_n} \varphi\big(a(x_{k,n})\big) \longrightarrow \int_D \varphi\big(a(x)\big)\,\mathrm{d}\nu(x), \qquad \varphi \in C_c(\mathbb{C}).$$

More generally, let $a : D \to \mathbb{C}^{s\times t}$ be finite-valued and Borel measurable, with entries continuous ν-almost everywhere, and set $r = \min\{s,t\}$. Then

$$\frac{1}{rN_n}\sum_{k=1}^{N_n}\sum_{j=1}^{r} \varphi\big(\sigma_j(a(x_{k,n}))\big) \longrightarrow \frac{1}{r}\int_D \sum_{j=1}^{r} \varphi\big(\sigma_j(a(x))\big)\,\mathrm{d}\nu(x), \qquad \varphi \in C_c([0,\infty)).$$

If $s = t$ and $a(x)$ is Hermitian for every $x \in D$, then

$$\frac{1}{sN_n}\sum_{k=1}^{N_n}\sum_{j=1}^{s} \varphi\big(\lambda_j(a(x_{k,n}))\big) \longrightarrow \frac{1}{s}\int_D \sum_{j=1}^{s} \varphi\big(\lambda_j(a(x))\big)\,\mathrm{d}\nu(x), \qquad \varphi \in C_c(\mathbb{R}).$$

Proof. Set $\nu_n = N_n^{-1}\sum_{k=1}^{N_n} \delta_{x_{k,n}}$. The node hypothesis is the weak convergence $\nu_n \Rightarrow \nu$. By the almost-everywhere continuous mapping theorem [21, Theorem 2.7], integration against ν_n converges to integration against ν for every bounded Borel function that is continuous ν-almost everywhere.

For the scalar statement, apply this fact to $h = \varphi \circ a$. For singular values, apply it to

$$h(x) = \frac{1}{r}\sum_{j=1}^{r} \varphi\big(\sigma_j(a(x))\big).$$

The ordered singular values depend continuously on the matrix entries, so h is continuous at every continuity point of a. Moreover, h is bounded because φ is bounded; no boundedness assumption on a is needed. The block-diagonal matrix has exactly the singular values of its diagonal blocks, counted with multiplicity, which gives the stated normalization. The Hermitian statement follows in the same way from continuity of the ordered eigenvalues on the Hermitian matrix space. □

When D is a compact rectangle, or a torus represented by a compact coordinate cell, and ν is normalized Lebesgue measure, entrywise Riemann integrability of a is sufficient.

Equivalently, define the noncanonical block-diagonal sequence $\{D_n(a;\mathcal{X}_n)\}_n$ by

$$D_n(a;\mathcal{X}_n) = \operatorname{diag}\big(a(x_{1,n}),\dots,a(x_{N_n,n})\big).$$

It has the transported singular-value distribution above and, in the square Hermitian case, the corresponding eigenvalue distribution. Empirical distributions depend only on the sampled values, whereas a GLT symbol also records their spatial ordering. For example, with $a(x) = x$ and $x_{k,n} = 1 - k/n$, the empirical node measures converge to uniform measure on $[0,1]$, but

$$D_n(a; \mathcal{X}_n) = \operatorname{diag}(1 - 1/n, \dots, 0) = D_n(b), \qquad b(x) = 1 - x.$$

The canonical GLT symbol is therefore $b(x)$, not $a(x)$; these functions have the same value distribution but different spatial dependence [15, Section 4.3.2].

2.3.3 Circulant sequences and Fourier diagonalization

Replacing the nonperiodic shifts of a Toeplitz matrix by periodic shifts produces a circulant. Its exact discrete Fourier diagonalization makes circulants useful both as approximants and as computational models for Toeplitz matrices.

Definition 2.31. For $n \in \mathbb{N}$ and $q \in \mathbb{Z}$, define the periodic shift $P_n^{(q)} \in \{0,1\}^{n \times n}$ by

$$[P_n^{(q)}]_{ij} = \mathbf{1}_{\{i \equiv j+q \pmod n\}}, \qquad i, j = 0, \dots, n-1.$$

For $\boldsymbol{n} = (n_1, \dots, n_d)$ and $\boldsymbol{\rho} = (\rho_1, \dots, \rho_d) \in \mathbb{Z}^d$, set

$$P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} = P_{n_1}^{(\rho_1)} \otimes \cdots \otimes P_{n_d}^{(\rho_d)}.$$

Let $H : \Lambda_{\boldsymbol{n}} \to \mathbb{C}^{s \times t}$. The associated d-level $s \times t$ block circulant is

$$\mathcal{C}_{\boldsymbol{n}}(H) = \sum_{\boldsymbol{\rho} \in \Lambda_{\boldsymbol{n}}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes H(\boldsymbol{\rho}) \in \mathbb{C}^{sN_{\boldsymbol{n}} \times tN_{\boldsymbol{n}}}.$$

The discrete Fourier transform diagonalizes the periodic shifts simultaneously. For block circulants, this yields one matrix block at each frequency [131, 89].

Proposition 2.32. Let

$$[F_n]_{kj} = \frac{1}{\sqrt{n}} \mathrm{e}^{-2\pi \mathrm{i} kj/n}, \qquad k, j = 0, \dots, n-1,$$

let $F_{\boldsymbol{n}} = F_{n_1} \otimes \cdots \otimes F_{n_d}$, and set

$$U_L = F_{\boldsymbol{n}} \otimes I_s, \qquad U_R = F_{\boldsymbol{n}} \otimes I_t.$$

Then

$$U_L \mathcal{C}_{\boldsymbol{n}}(H) U_R^* = \operatorname{Bdiag}\big(\widehat{H}_{\boldsymbol{n}}(\boldsymbol{k})\big)_{\boldsymbol{k} \in \Lambda_{\boldsymbol{n}}},$$

where

$$\widehat{H}_{\boldsymbol{n}}(\boldsymbol{k}) = \sum_{\boldsymbol{\rho} \in \Lambda_{\boldsymbol{n}}} H(\boldsymbol{\rho}) \mathrm{e}^{-2\pi \mathrm{i} \sum_{\ell=1}^{d} k_\ell \rho_\ell / n_\ell}.$$

Thus a rectangular block circulant is two-sided unitarily equivalent to its frequency blocks; in the square-block case, the factorization is a unitary similarity.

Lemma 2.33 (DFT reversal identities). Let F_n and $F_{\boldsymbol{n}}$ be the discrete Fourier matrices of Proposition 2.32. Define the modular-reversal permutation by

$$A_n e_k = e_{(-k) \bmod n}, \qquad 0 \le k < n, \qquad A_{\boldsymbol{n}} = A_{n_1} \otimes \cdots \otimes A_{n_d}.$$

Then

$$F_n^* F_n = I_n, \qquad F_n^T = F_n, \qquad F_n^2 = A_n.$$

The same identities hold for $F_{\boldsymbol{n}}$, with $I_{N_{\boldsymbol{n}}}$ and $A_{\boldsymbol{n}}$ in place of I_n and A_n.

Proof. The entries of F_n are symmetric in their indices, and finite Fourier orthogonality gives both unitarity and

$$[F_n^2]_{jk} = \frac{1}{n}\sum_{\ell=0}^{n-1} \mathrm{e}^{-2\pi \mathrm{i}\ell(j+k)/n} = \mathbf{1}_{\{j+k\equiv 0 \pmod n\}}.$$

Tensor products give the multilevel identities. □

The singular values of a block circulant are therefore the singular values of its frequency blocks, counted together. In the square case, the same statement holds for eigenvalues. The factorization also supports matrix-vector products through discrete Fourier transforms and, when the square frequency blocks are invertible, linear solvers.

Definition 2.34. Let

$$p(\boldsymbol{\theta}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\boldsymbol{\rho}} \mathrm{e}^{\mathrm{i}(\boldsymbol{\rho},\boldsymbol{\theta})}, \qquad p_{\boldsymbol{\rho}} \in \mathbb{C}^{s\times t},$$

where $\mathcal{F} \subset \mathbb{Z}^d$ is finite. The associated multilevel block circulant is

$$C_{\boldsymbol{n}}(p) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\boldsymbol{\rho}}.$$

For $\boldsymbol{k} \in \Lambda_{\boldsymbol{n}}$, let $\boldsymbol{\theta}_{\boldsymbol{k}}^{(\boldsymbol{n})}$ denote the torus point satisfying

$$\boldsymbol{\theta}_{\boldsymbol{k}}^{(\boldsymbol{n})} \equiv -2\pi\left(\frac{k_1}{n_1},\dots,\frac{k_d}{n_d}\right) \pmod{2\pi},$$

where the congruence is componentwise and the representative is chosen in $[-\pi,\pi)^d$. The DFT factorization gives the exact identity

$$(F_{\boldsymbol{n}} \otimes I_s) C_{\boldsymbol{n}}(p) (F_{\boldsymbol{n}}^* \otimes I_t) = \mathrm{Bdiag}\big(p(\boldsymbol{\theta}_{\boldsymbol{k}}^{(\boldsymbol{n})})\big)_{\boldsymbol{k}\in\Lambda_{\boldsymbol{n}}}.$$

The sign of the Fourier grid reflects the chosen DFT convention and has no effect on the limiting distribution.

Corollary 2.35. Let $p : [-\pi,\pi]^d \to \mathbb{C}^{s\times t}$ be a trigonometric polynomial. Then, as $\min_\ell n_\ell \to \infty$,

$$\{C_{\boldsymbol{n}}(p)\}_{\boldsymbol{n}} \sim_\sigma p.$$

If $s = t$ and p is Hermitian, then also

$$\{C_{\boldsymbol{n}}(p)\}_{\boldsymbol{n}} \sim_\lambda p.$$

This distribution follows directly from the block diagonalization and Riemann-sum convergence. To compare the circulant with its Toeplitz counterpart, we next count the columns affected by replacing nonperiodic shifts with periodic ones.

Lemma 2.36. For $q \in \mathbb{Z}$, define the nonperiodic shift

$$[J_n^{(q)}]_{ij} = \mathbf{1}_{\{i=j+q\}}, \qquad i,j = 0,\dots,n-1,$$

and, for $\boldsymbol{\rho} \in \mathbb{Z}^d$, set

$$J_{\boldsymbol{n}}^{(\boldsymbol{\rho})} = J_{n_1}^{(\rho_1)} \otimes \cdots \otimes J_{n_d}^{(\rho_d)}.$$

With $x_+ = \max\{x, 0\}$, one has

$$\mathrm{rank}\big(J_{\boldsymbol{n}}^{(\boldsymbol{\rho})} - P_{\boldsymbol{n}}^{(\boldsymbol{\rho})}\big) = N_{\boldsymbol{n}} - \prod_{\ell=1}^{d} (n_\ell - |\rho_\ell|)_+ \le \sum_{\ell=1}^{d} |\rho_\ell| \frac{N_{\boldsymbol{n}}}{n_\ell}.$$

Proof. Let D be the diagonal projection indexed by $\Lambda_{\boldsymbol{n}}$ whose $\boldsymbol{j}$-th diagonal entry equals one precisely when $\boldsymbol{j}+\boldsymbol{\rho}\in\Lambda_{\boldsymbol{n}}$. On these columns, the periodic and nonperiodic shifts agree; all other columns of the nonperiodic shift are zero. Hence

$$J_{\boldsymbol{n}}^{(\boldsymbol{\rho})} = P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} D.$$

Since $P_{\boldsymbol{n}}^{(\boldsymbol{\rho})}$ is unitary, the rank of the difference is $N_{\boldsymbol{n}} - \operatorname{rank} D$. There are exactly $\prod_{\ell=1}^{d}(n_\ell - |\rho_\ell|)_+$ retained columns, which proves the equality. The number of excluded columns is at most the sum, over the coordinate directions, of the numbers excluded in that direction:

$$N_{\boldsymbol{n}} - \operatorname{rank} D \le \sum_{\ell=1}^{d} \min\{n_\ell, |\rho_\ell|\}\frac{N_{\boldsymbol{n}}}{n_\ell} \le \sum_{\ell=1}^{d} |\rho_\ell|\frac{N_{\boldsymbol{n}}}{n_\ell}.$$

□

The boundary-rank estimate makes the Toeplitz–circulant difference zero-distributed by Theorem 2.11; see also [16, Theorem 2.18] and [15, Section 4.3.3].

Corollary 2.37. Let p be a fixed $s\times t$ trigonometric polynomial. Then

$$\{T_{\boldsymbol{n}}(p) - C_{\boldsymbol{n}}(p)\}_{\boldsymbol{n}}$$

is zero-distributed as $\min_\ell n_\ell \to \infty$.

Proof. Writing

$$T_{\boldsymbol{n}}(p) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} J_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\boldsymbol{\rho}}, \qquad C_{\boldsymbol{n}}(p) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\boldsymbol{\rho}},$$

rank subadditivity and the Kronecker-product rank identity give

$$\operatorname{rank}\big(T_{\boldsymbol{n}}(p) - C_{\boldsymbol{n}}(p)\big) \le \sum_{\boldsymbol{\rho}\in\mathcal{F}} \operatorname{rank}(p_{\boldsymbol{\rho}}) \sum_{\ell=1}^{d} |\rho_\ell|\frac{N_{\boldsymbol{n}}}{n_\ell} = o(N_{\boldsymbol{n}}).$$

The fixed block sizes do not change the vanishing relative-rank conclusion, so the difference is zero-distributed. □

A fixed polynomial Toeplitz sequence therefore differs from its periodic counterpart by a correction of vanishing relative rank. In one level the rank is uniformly bounded; in several levels it can grow like the number of grid points on the boundary. Thus the multilevel argument establishes zero distribution, not a uniformly bounded number of outliers. For example, if $d=2$, $\boldsymbol{n}=(n,n)$, and $p(\theta_1,\theta_2)=\mathrm{e}^{\mathrm{i}\theta_1}$, then the difference has rank n and exactly n singular values equal to one.

2.3.4 Hankel sequences

Hankel matrices depend on the sum, rather than the difference, of their indices. They arise in truncation and nonperiodic block assembly. Here we consider Hankel matrices formed from the Fourier coefficients of a fixed integrable symbol.

Definition 2.38. For $f\in L^1([-\pi,\pi];\mathbb{C}^{s\times t})$, define the rectangular block Hankel matrix by

$$H_{n,n'}(f) = [\,\widehat{f}_{i+j}\,]_{i=0,\dots,n-1;\,j=0,\dots,n'-1} \in \mathbb{C}^{sn\times tn'}.$$

We write $H_n(f)=H_{n,n}(f)$. For a fixed integer q, the version with entries $\widehat{f}_{i+j+q}$ is generated by $\mathrm{e}^{-\mathrm{i}q\theta}f(\theta)$.

Block-diagonal approximation arguments also involve rectangular Toeplitz and Hankel matrices and their products with reversal matrices. For bounded symbols, the uniform norm estimates

$$\|T_{n,n'}(f)\| \le \|f\|_{L^\infty}, \qquad \|H_{n,n'}(f)\| \le \|f\|_{L^\infty}$$

hold for all admissible dimensions; see [155, Corollaries 3.5–3.6 and Section 3.3].

The following is the unilevel specialization of the Fasino–Tilli zero-distribution theorem [66, Theorem 1].

Theorem 2.39. Let $f \in L^1([-\pi,\pi];\mathbb{C}^{s\times t})$. Then

$$\{H_n(f)\}_n \sim_\sigma 0,$$

with active singular-value normalization $\min\{sn, tn\}$.

The same zero-distribution conclusion holds for the family of rectangular Hankel matrices $\{H_{n,n'}(f)\}_{n,n'}$ when $\min\{n,n'\} \to \infty$ and the ratio n/n' stays bounded above and away from zero, with active normalization $r_{n,n'} = \min\{sn, tn'\}$.

Justification of the rectangular case. Set $m = \max\{n,n'\}$. The matrix $H_{n,n'}(f)$ is obtained from $H_m(f)$ by row and column compression, so, for every $\varepsilon > 0$,

$$\#\{j \le r_{n,n'} : \sigma_j(H_{n,n'}(f)) > \varepsilon\} \le \#\{j \le m\min\{s,t\} : \sigma_j(H_m(f)) > \varepsilon\}.$$

Dividing the displayed inequality by $r_{n,n'}$ and applying [66, Theorem 1] prove the claim. □

This argument uses comparability of the active dimensions, rather than the asymptotically full compression required by the extradimensional principle. Multilevel versions with fixed block sizes and equal row and column outer index sets are included in [66].

Zero distribution is preserved by fixed Fourier-index shifts, by multiplication of these Hankel matrices on either side by reversal matrices, and by coordinate compressions with comparable active dimensions. It is not implied by reversing an arbitrary Toeplitz matrix. For example, if Y_n is the reversal matrix, then $T_n(1)Y_n = Y_n$, whose singular values are all one. Although Y_n has Hankel structure, its nonzero anti-diagonal moves with n, so it is not covered by the fixed-symbol hypothesis above.

Remark 2.40. For unilevel symbols continuous on the torus $\mathbb{T}$, as assumed in Chapter 3, the Hankel corrections satisfy a stronger bounded-rank plus arbitrarily small-norm decomposition. Because this refinement is needed only for the strong-clustering conclusion in that chapter, it is established directly in the proof of Theorem 3.7 rather than recorded here as a separate theorem.

Under the hypotheses stated above, Toeplitz, zero-distributed, and canonical sampling sequences generate the GLT class introduced next. Polynomial circulants inherit the GLT symbols of their Toeplitz counterparts through the zero-distributed difference just established.

2.4 Generalized Locally Toeplitz Sequences

The GLT calculus combines the matrix models of the preceding section and transfers algebraic operations to their symbols. It extends and modifies the locally Toeplitz framework of Tilli [167], with an early systematic development by Serra-Capizzano [156]. We use the square multilevel block theory of [16] and its rectangular extension in [15].

Fix $d \ge 1$, and set

$$\boldsymbol{n} = (n_1,\dots,n_d), \qquad N_{\boldsymbol{n}} = \prod_{\ell=1}^{d} n_\ell, \qquad D_d = [0,1]^d \times [-\pi,\pi]^d.$$

Block sizes are fixed, all sequences involved in an algebraic operation use the same multilevel grid, and limits are taken as $\min_\ell n_\ell \to \infty$. The notation also applies along any prescribed sequence of such multi-indices.

2.4.1 Definition and fundamental rules

A constructive description of a GLT sequence uses finite sums of products of sampling and Toeplitz matrices, followed by an a.c.s. limit. This characterization follows from [15, Theorems 4.7 and 4.9] and is equivalent to the block-extension definition in [15, Definition 4.1].

Definition 2.41 (GLT sequence and symbol). Let $s, t \geq 1$ be fixed, let $A_{\boldsymbol{n}} \in \mathbb{C}^{sN_{\boldsymbol{n}} \times tN_{\boldsymbol{n}}}$, and let $\kappa : D_d \to \mathbb{C}^{s\times t}$ be measurable. We write

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$$

and call $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ a d-level (s,t)-block GLT sequence with symbol κ if there exist finite sums

$$\kappa_m(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{i=1}^{M_m} a_{i,m}(\boldsymbol{x}) f_{i,m}(\boldsymbol{\theta}),$$
$$B_{\boldsymbol{n},m} = \sum_{i=1}^{M_m} D_{\boldsymbol{n}}(a_{i,m} I_s) T_{\boldsymbol{n}}(f_{i,m})$$

such that

$$\kappa_m \to \kappa \quad \text{in measure on } D_d, \qquad \{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \xrightarrow{\text{a.c.s.}} \{A_{\boldsymbol{n}}\}_{\boldsymbol{n}}.$$

Here $a_{i,m} \in C^\infty([0,1]^d;\mathbb{C})$, and each $f_{i,m}$ is a rectangular trigonometric monomial of the form

$$f_{i,m}(\boldsymbol{\theta}) = \mathrm{e}^{\mathrm{i}(\boldsymbol{\rho}_{i,m},\boldsymbol{\theta})} e^{(s)}_{\alpha_{i,m}} (e^{(t)}_{\beta_{i,m}})^*,$$

where $\boldsymbol{\rho}_{i,m} \in \mathbb{Z}^d$, $1 \leq \alpha_{i,m} \leq s$, $1 \leq \beta_{i,m} \leq t$, and $e^{(r)}_j$ is the j-th standard basis vector of $\mathbb{C}^r$.

The approximation in this definition is tied to the prescribed spatial and block ordering, so a GLT symbol contains more information than a singular-value distribution alone. We retain the rule numbers of [15, Section 5]. The Hermitian spectral and functional-calculus results are stated together in Proposition 2.42, rather than as GLT 2 and GLT 5.

The first two rules identify the symbol uniquely and connect it to the singular-value distribution [15, Theorems 4.3, 4.13, and 4.4].

GLT 0. If $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$, then, for every measurable $\xi : D_d \to \mathbb{C}^{s\times t}$,

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi \quad \Longleftrightarrow \quad \kappa = \xi \quad \text{almost everywhere.}$$

Conversely, every measurable $\kappa : D_d \to \mathbb{C}^{s\times t}$ is the symbol of an (s,t)-block GLT sequence along every prescribed multilevel index sequence with $\min_\ell n_\ell \to \infty$.

GLT 1. If $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$, then

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_\sigma \kappa,$$

with normalization by $N_{\boldsymbol{n}} \min\{s,t\}$. For square non-Hermitian sequences this rule alone gives no eigenvalue-distribution conclusion.

The three generator families have the following symbols [15, Section 4.3].

GLT 3. For $f \in L^1([-\pi,\pi]^d;\mathbb{C}^{s\times t})$,

$$\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} f(\boldsymbol{\theta}).$$

If $a : [0,1]^d \to \mathbb{C}^{s\times t}$ is finite-valued, measurable, and entrywise continuous almost everywhere, then

$$\{D_{\boldsymbol{n}}(a)\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} a(\boldsymbol{x}).$$

The regularity assumption applies to the pointwise representative that is actually sampled. Finally, for $Z_{\boldsymbol{n}} \in \mathbb{C}^{sN_{\boldsymbol{n}}\times tN_{\boldsymbol{n}}}$,

$$\{Z_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} O_{s,t} \quad \Longleftrightarrow \quad \{Z_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\sigma} 0,$$

where $O_{s,t}$ denotes the $s \times t$ zero matrix.

Compatible algebraic operations on GLT sequences correspond to the same operations on their symbols. Pseudoinversion is included when the symbol has maximal rank almost everywhere [15, Theorems 4.5 and 4.6].

GLT 4. Suppose that $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$ and $\{B_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi$. Then

$$\{A_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa^*.$$

If κ and ξ have the same dimensions, then

$$\{\alpha A_{\boldsymbol{n}} + \beta B_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \alpha\kappa + \beta\xi, \qquad \alpha, \beta \in \mathbb{C}.$$

If κ has size $s \times t$ and ξ has size $t \times r$, then

$$\{A_{\boldsymbol{n}} B_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa\xi.$$

If $\operatorname{rank}\kappa = \min\{s,t\}$ almost everywhere, then

$$\{A_{\boldsymbol{n}}^\dagger\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa^\dagger.$$

In the square case, if each $A_{\boldsymbol{n}}$ is invertible and κ is invertible almost everywhere, this gives $\{A_{\boldsymbol{n}}^{-1}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa^{-1}$.

The maximal-rank hypothesis concerns the symbol, not every finite matrix. In fact, it is equivalent to sparse vanishing of the GLT sequence by Proposition 2.14.

Block restriction and assembly also pass to the symbols, provided that the matrices use the corresponding grid-first ordering [15, Theorems 4.10–4.12].

GLT 6. Let $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$, where κ has size $s \times t$. Let $R \in \mathbb{C}^{s\times\widetilde{s}}$ and $S \in \mathbb{C}^{t\times\widetilde{t}}$ consist of fixed selections of distinct standard basis vectors. Then

$$\{(I_{N_{\boldsymbol{n}}} \otimes R^*) A_{\boldsymbol{n}} (I_{N_{\boldsymbol{n}}} \otimes S)\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} R^*\kappa S.$$

Thus the same internal rows and columns are selected in every block; this is not an arbitrary selection of global matrix rows and columns.

More generally, let $u, v \geq 1$ and $s_i, t_j \geq 1$ be fixed, and let

$$A_{\boldsymbol{n}}^{(ij)} = [A_{\boldsymbol{n};\boldsymbol{k},\boldsymbol{l}}^{(ij)}]_{\boldsymbol{k},\boldsymbol{l}\in\Lambda_{\boldsymbol{n}}}, \qquad A_{\boldsymbol{n};\boldsymbol{k},\boldsymbol{l}}^{(ij)} \in \mathbb{C}^{s_i\times t_j}.$$

Set $s = \sum_i s_i$, $t = \sum_j t_j$, and define the matrix whose outer indexing is by grid points by

$$\widetilde{A}_{\boldsymbol{n}} = \Big[[A_{\boldsymbol{n};\boldsymbol{k},\boldsymbol{l}}^{(ij)}]_{i=1,\dots,u}^{j=1,\dots,v} \Big]_{\boldsymbol{k},\boldsymbol{l}\in\Lambda_{\boldsymbol{n}}}.$$

For measurable $\kappa_{ij} : D_d \to \mathbb{C}^{s_i\times t_j}$,

$$\{\widetilde{A}_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} [\kappa_{ij}]_{i=1,\dots,u}^{j=1,\dots,v} \quad \Longleftrightarrow \quad \{A_{\boldsymbol{n}}^{(ij)}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa_{ij} \quad \text{for all } i,j.$$

The usual block array $[A_{\boldsymbol{n}}^{(ij)}]_{i,j}$ must first be permuted from the ordering $(i,\boldsymbol{k},\alpha)$ to $(\boldsymbol{k},i,\alpha)$ in its rows, and analogously in its columns, to obtain $\widetilde{A}_{\boldsymbol{n}}$.

Within the GLT class, a.c.s. approximation of matrices is equivalent to convergence of their symbols in measure [15, Theorems 4.7–4.9]. All sequences in an a.c.s. relation have the same dimensions at each $\boldsymbol{n}$.

GLT 7. One has $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa$ if and only if there are GLT sequences $\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa_m$ such that

$$\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \xrightarrow{\text{a.c.s.}} \{A_{\boldsymbol{n}}\}_{\boldsymbol{n}}, \qquad \kappa_m \to \kappa \quad \text{in measure on } D_d.$$

GLT 8. Suppose that $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa$ and $\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa_m$. Then

$$\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \xrightarrow{\text{a.c.s.}} \{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \quad \Longleftrightarrow \quad \kappa_m \to \kappa \quad \text{in measure on } D_d.$$

In particular, two GLT sequences have the same symbol almost everywhere if and only if their difference is zero-distributed.

GLT 9. If $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa$, the approximants in Definition 2.41 can be chosen so that

$$\kappa_m(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{i=1}^{M_m} a_{i,m}(\boldsymbol{x}) f_{i,m}(\boldsymbol{\theta}) \longrightarrow \kappa(\boldsymbol{x},\boldsymbol{\theta}) \quad \text{almost everywhere,}$$

and

$$\left\{ \sum_{i=1}^{M_m} D_{\boldsymbol{n}}(a_{i,m} I_s) T_{\boldsymbol{n}}(f_{i,m}) \right\}_{\boldsymbol{n}} \xrightarrow{\text{a.c.s.}} \{A_{\boldsymbol{n}}\}_{\boldsymbol{n}}.$$

For Hermitian sequences, the calculus also describes eigenvalue distributions and continuous functions of the matrices [16, Proposition 4.10 and Theorems 4.11 and 4.18].

Proposition 2.42 (Hermitian GLT calculus). Let $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \kappa$, with every $A_{\boldsymbol{n}} \in \mathbb{C}^{sN_{\boldsymbol{n}} \times sN_{\boldsymbol{n}}}$ Hermitian. Then κ is Hermitian almost everywhere,

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\lambda} \kappa, \qquad \{g(A_{\boldsymbol{n}})\}_{\boldsymbol{n}} \sim_{\text{GLT}} g(\kappa) \quad \text{for every } g \in C(\mathbb{R};\mathbb{C}).$$

Moreover, suppose that $\{P_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \xi$, and every $P_{\boldsymbol{n}}$ is Hermitian positive definite. Then

$$\{P_{\boldsymbol{n}}^{-1/2}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \xi^{-1/2}.$$

The last assertion follows from the proofs in [77, Solution of Exercise 8.4, and Corollary 3.1], with the GLT axioms interpreted in the block setting.

Related eigenvalue results are available beyond the Hermitian case. For normal matrices in the scalar unilevel GLT setting, see [14, Theorems 3.3 and 3.4].

2.4.2 Extradimensional compression in the GLT setting

A matrix sequence may have the structure of a GLT model without having exactly the required dimensions. The ambient realizations of Subsection 2.2.6 allow us to apply the GLT calculus in a common larger space and then return the distributional conclusions to the original dimensions. Let

$$E_{\boldsymbol{n}} \in \mathbb{C}^{aN_{\boldsymbol{n}} \times p_{\boldsymbol{n}}}, \qquad F_{\boldsymbol{n}} \in \mathbb{C}^{bN_{\boldsymbol{n}} \times q_{\boldsymbol{n}}}$$

satisfy

$$\begin{aligned} E_{\boldsymbol{n}}^* E_{\boldsymbol{n}} &= I_{p_{\boldsymbol{n}}}, & F_{\boldsymbol{n}}^* F_{\boldsymbol{n}} &= I_{q_{\boldsymbol{n}}}, \\ aN_{\boldsymbol{n}} - p_{\boldsymbol{n}} &= o(N_{\boldsymbol{n}}), & bN_{\boldsymbol{n}} - q_{\boldsymbol{n}} &= o(N_{\boldsymbol{n}}). \end{aligned} \tag{2.20}$$

These pairs are asymptotically full because $\min\{aN_{\boldsymbol{n}}, bN_{\boldsymbol{n}}\} = N_{\boldsymbol{n}} \min\{a,b\}$. The isometries themselves need not be GLT sequences.

The next proposition follows from the zero-distributed generator and the GLT algebra [15, Section 4.3.3 and Theorem 4.5].

Proposition 2.43. Assume (2.20), and let

$$\mathcal{X}_{\boldsymbol{n}} \in \mathbb{C}^{aN_{\boldsymbol{n}} \times bN_{\boldsymbol{n}}}, \qquad \{\mathcal{X}_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa, \qquad A_{\boldsymbol{n}} = E_{\boldsymbol{n}}^* \mathcal{X}_{\boldsymbol{n}} F_{\boldsymbol{n}}.$$

Then the ambient realizations

$$\mathcal{A}_{\boldsymbol{n}} := E_{\boldsymbol{n}} A_{\boldsymbol{n}} F_{\boldsymbol{n}}^* \tag{2.21}$$

satisfy $\{\mathcal{A}_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa$.

Proof. Set $\Pi_{\boldsymbol{n}}^E = E_{\boldsymbol{n}} E_{\boldsymbol{n}}^*$ and $\Pi_{\boldsymbol{n}}^F = F_{\boldsymbol{n}} F_{\boldsymbol{n}}^*$. Since $\mathcal{A}_{\boldsymbol{n}} = \Pi_{\boldsymbol{n}}^E \mathcal{X}_{\boldsymbol{n}} \Pi_{\boldsymbol{n}}^F$,

$$\mathcal{X}_{\boldsymbol{n}} - \mathcal{A}_{\boldsymbol{n}} = (I_{aN_{\boldsymbol{n}}} - \Pi_{\boldsymbol{n}}^E)\mathcal{X}_{\boldsymbol{n}} + \Pi_{\boldsymbol{n}}^E \mathcal{X}_{\boldsymbol{n}} (I_{bN_{\boldsymbol{n}}} - \Pi_{\boldsymbol{n}}^F).$$

Consequently,

$$\operatorname{rank}(\mathcal{X}_{\boldsymbol{n}} - \mathcal{A}_{\boldsymbol{n}}) \le (aN_{\boldsymbol{n}} - p_{\boldsymbol{n}}) + (bN_{\boldsymbol{n}} - q_{\boldsymbol{n}}) = o(N_{\boldsymbol{n}}).$$

The difference is zero-distributed, hence has GLT symbol zero by GLT 3. GLT 4 then gives the asserted symbol of $\{\mathcal{A}_{\boldsymbol{n}}\}_{\boldsymbol{n}}$. □

Corollary 2.44 (Return from ambient to physical dimensions). Assume (2.20), and let $A_{\boldsymbol{n}} \in \mathbb{C}^{p_{\boldsymbol{n}} \times q_{\boldsymbol{n}}}$. If

$$\{E_{\boldsymbol{n}} A_{\boldsymbol{n}} F_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa,$$

then $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_\sigma \kappa$, with active normalization $\min\{p_{\boldsymbol{n}}, q_{\boldsymbol{n}}\}$. If, additionally, $a = b$, $p_{\boldsymbol{n}} = q_{\boldsymbol{n}}$, $F_{\boldsymbol{n}} = E_{\boldsymbol{n}}$, and every $A_{\boldsymbol{n}}$ is Hermitian, then $\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_\lambda \kappa$.

Proof. Apply GLT 1 to the ambient realizations and then Proposition 2.19 to their exact compressions. In the Hermitian case the ambient realizations are Hermitian, so Proposition 2.42 supplies the eigenvalue distribution before the same dimension-transfer step. □

The return to physical dimensions preserves the distribution, but the physical sequence need not have the same GLT symbol as its ambient realization. The coordinate ordering still matters. For example, if Y_n is the reversal matrix and $a(x) = x$, $b(x) = 1 - x$, then

$$Y_n^* D_n(a) Y_n = D_n(b) + n^{-1} I_n.$$

The sequence $\{Y_n^* D_n(a) Y_n\}_n$ has GLT symbol $b(x)$, whereas the ambient sequence through $E_n = F_n = Y_n$ is $\{D_n(a)\}_n$, with symbol $a(x)$.

For algebraic operations, compatible ambient realizations are essential. In addition to $E_{\boldsymbol{n}}, F_{\boldsymbol{n}}$, let $H_{\boldsymbol{n}} \in \mathbb{C}^{cN_{\boldsymbol{n}} \times r_{\boldsymbol{n}}}$ satisfy $H_{\boldsymbol{n}}^* H_{\boldsymbol{n}} = I_{r_{\boldsymbol{n}}}$ and $cN_{\boldsymbol{n}} - r_{\boldsymbol{n}} = o(N_{\boldsymbol{n}})$, with c fixed. Let

$$A_{\boldsymbol{n}}, C_{\boldsymbol{n}} \in \mathbb{C}^{p_{\boldsymbol{n}} \times q_{\boldsymbol{n}}}, \qquad B_{\boldsymbol{n}} \in \mathbb{C}^{q_{\boldsymbol{n}} \times r_{\boldsymbol{n}}},$$

and suppose that

$$\{\mathcal{A}_{\boldsymbol{n}} := E_{\boldsymbol{n}} A_{\boldsymbol{n}} F_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa,$$
$$\{\mathcal{C}_{\boldsymbol{n}} := E_{\boldsymbol{n}} C_{\boldsymbol{n}} F_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \zeta,$$
$$\{\mathcal{B}_{\boldsymbol{n}} := F_{\boldsymbol{n}} B_{\boldsymbol{n}} H_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi.$$

Lemma 2.21 and GLT 4 give

$$\{E_{\boldsymbol{n}}(\alpha A_{\boldsymbol{n}} + \beta C_{\boldsymbol{n}}) F_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \alpha\kappa + \beta\zeta,$$
$$\{E_{\boldsymbol{n}}(A_{\boldsymbol{n}} B_{\boldsymbol{n}}) H_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa\xi.$$

The product uses the same intermediate isometry $F_{\boldsymbol{n}}$. If κ has maximal rank almost everywhere, then

$$\{F_{\boldsymbol{n}} A_{\boldsymbol{n}}^\dagger E_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} = \{\mathcal{A}_{\boldsymbol{n}}^\dagger\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa^\dagger.$$

In the square Hermitian case, with the same isometry on both sides, Lemma 2.21 gives

$$E_{\boldsymbol{n}} g(A_{\boldsymbol{n}}) E_{\boldsymbol{n}}^* = g(\mathcal{A}_{\boldsymbol{n}}) - g(0)(I_{aN_{\boldsymbol{n}}} - E_{\boldsymbol{n}} E_{\boldsymbol{n}}^*).$$

For fixed $g \in C(\mathbb{R};\mathbb{C})$, the correction has rank $o(N_{\boldsymbol{n}})$. Hence

$$\{E_{\boldsymbol{n}} g(A_{\boldsymbol{n}}) E_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} g(\kappa).$$

If every physical $A_{\boldsymbol{n}}$ is positive definite and κ is positive definite almost everywhere, then

$$\{E_{\boldsymbol{n}} A_{\boldsymbol{n}}^{-1/2} E_{\boldsymbol{n}}^*\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa^{-1/2}.$$

Indeed, the pseudoinverse rule applies to the possibly singular ambient matrix, and

$$h(\mathcal{A}_{\boldsymbol{n}}^\dagger) = E_{\boldsymbol{n}} A_{\boldsymbol{n}}^{-1/2} E_{\boldsymbol{n}}^*, \qquad h(t) = \sqrt{\max\{t, 0\}}.$$

2.4.3 Symbol matching and preconditioning

When a matrix sequence and its preconditioner share a maximal-rank symbol, the appropriate preconditioned product has identity symbol. For rectangular matrices, the side of preconditioning depends on whether the symbol has full column or full row rank. The next theorem applies the product and pseudoinverse rules of [15, Theorems 4.5 and 4.6]; the Hermitian corollary also uses [16, Theorems 4.11 and 4.18].

Theorem 2.45 (Pseudoinverse symbol matching). Let

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa, \qquad \{P_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi,$$

where $A_{\boldsymbol{n}}, P_{\boldsymbol{n}} \in \mathbb{C}^{sN_{\boldsymbol{n}} \times tN_{\boldsymbol{n}}}$, and assume that $\operatorname{rank} \xi = \min\{s,t\}$ almost everywhere. Equivalently, $\{P_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ is sparsely vanishing. Then, when dimensions are compatible,

$$\{P_{\boldsymbol{n}}^\dagger A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi^\dagger \kappa, \qquad \{A_{\boldsymbol{n}} P_{\boldsymbol{n}}^\dagger\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa \xi^\dagger.$$

Under symbol matching $\xi = \kappa$ almost everywhere:

(a) If $s \geq t$, then $\xi^\dagger \kappa = I_t$ almost everywhere, and the sequence $\{P_{\boldsymbol{n}}^\dagger A_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ is weakly clustered at one in the singular-value sense.

(b) If $s \leq t$, then $\kappa \xi^\dagger = I_s$ almost everywhere, and the sequence $\{A_{\boldsymbol{n}} P_{\boldsymbol{n}}^\dagger\}_{\boldsymbol{n}}$ is weakly clustered at one in the singular-value sense.

Both conclusions apply when $s = t$. In this square case, if every $P_{\boldsymbol{n}}$ is invertible, its pseudoinverse can be replaced by its inverse.

Proof. The full-rank symbol permits the pseudoinverse rule, and the product rule gives both displayed GLT symbols. In the full-column-rank case, $\kappa^\dagger \kappa = I_t$; in the full-row-rank case, $\kappa \kappa^\dagger = I_s$. GLT 1 therefore gives singular-value distribution at one for the indicated products, which is equivalent to weak singular-value clustering at one. □

For a Hermitian sequence and a positive-definite preconditioner, symmetric preconditioning gives a Hermitian matrix. Similarity then transfers its eigenvalue distribution to the usual left-preconditioned product.

Corollary 2.46 (Hermitian positive-definite preconditioning). Let

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \kappa, \qquad \{P_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi,$$

where every $A_{\boldsymbol{n}} \in \mathbb{C}^{sN_{\boldsymbol{n}} \times sN_{\boldsymbol{n}}}$ is Hermitian, every $P_{\boldsymbol{n}}$ is Hermitian positive definite, and ξ is positive definite almost everywhere. Set

$$H_{\boldsymbol{n}} = P_{\boldsymbol{n}}^{-1/2} A_{\boldsymbol{n}} P_{\boldsymbol{n}}^{-1/2}, \qquad \eta = \xi^{-1/2} \kappa \xi^{-1/2}.$$

Then $H_{\boldsymbol{n}}$ is Hermitian and

$$\{H_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \eta, \qquad \{H_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\lambda} \eta.$$

Moreover,

$$P_{\boldsymbol{n}}^{-1} A_{\boldsymbol{n}} = P_{\boldsymbol{n}}^{-1/2} H_{\boldsymbol{n}} P_{\boldsymbol{n}}^{1/2}, \qquad \{P_{\boldsymbol{n}}^{-1} A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\lambda} \eta.$$

If $\kappa = \xi$ almost everywhere, equivalently $\eta = I_s$ almost everywhere, the eigenvalues of both sequences are weakly clustered at one.

Proof. Proposition 2.42 gives $\{P_{\boldsymbol{n}}^{-1/2}\}_{\boldsymbol{n}} \sim_{\mathrm{GLT}} \xi^{-1/2}$. The product rule gives the symbol η of $\{H_{\boldsymbol{n}}\}_{\boldsymbol{n}}$, and the Hermitian part of the same proposition gives its eigenvalue distribution. The displayed similarity transfers the eigenvalues, with their multiplicities, to $P_{\boldsymbol{n}}^{-1} A_{\boldsymbol{n}}$. If $\kappa = \xi$, then $\eta = I_s$, yielding weak clustering at one. Conversely, multiplication of $\eta = I_s$ on both sides by $\xi^{1/2}$ gives $\kappa = \xi$. □

The symmetric product and similarity argument require no commutativity between κ and ξ.

2.5 Exactness, best approximation, and stability

We now turn from asymptotic spectral distributions to approximation in norm. If an approximation operator and its target agree on a subspace, subtracting any element of that subspace leaves their error unchanged. Best approximation measures how small the remaining residual can be; the norm of the error operator controls its amplification. This argument underlies the quadrature and interpolation estimates below.

For a sequence of bounded linear operators $T_n : X \to Y$, we call $\|T_n\|_{X \to Y}$ its *stability factor*. The sequence $\{T_n\}_n$ is *uniformly stable* if

$$\sup_n \|T_n\|_{X \to Y} < \infty.$$

Boundedness of every individual operator does not imply uniform stability.

2.5.1 Best approximation and the exactness–stability principle

Definition 2.47. Let X be a normed linear space over $\mathbb{R}$ or $\mathbb{C}$, and let $V \subseteq X$ be a linear subspace. For $f \in X$, the error of best approximation from V is

$$E_X(f; V) = \inf_{v \in V} \|f - v\|_X. \tag{2.22}$$

An element $v^* \in V$ is a best approximant if $\|f - v^*\|_X = E_X(f; V)$. The infimum need not be attained for an arbitrary subspace.

For a compact Hausdorff space K, equip $C(K)$ with the uniform norm and write

$$E(f; V; K) = \inf_{v \in V} \|f - v\|_{\infty, K}, \qquad \|f\|_{\infty, K} = \max_{x \in K} |f(x)|. \tag{2.23}$$

If $\mathbb{P}_m$ denotes the algebraic polynomials of degree at most m, over the same scalar field as $C([a, b])$, set

$$E_m^{\mathrm{alg}}(f; [a, b]) = \inf_{p \in \mathbb{P}_m} \|f - p\|_{\infty, [a,b]}. \tag{2.24}$$

The superscript distinguishes algebraic approximation from the Laurent, trigonometric, and weighted approximation errors used later.

Finite-dimensionality guarantees existence of a best approximant. For polynomial spaces on an interval, the Weierstrass theorem also ensures that the error tends to zero as the degree increases; see [122, Section 1.1] and [39].

Proposition 2.48. Every finite-dimensional subspace V of a normed linear space X contains a best approximant to each $f \in X$. Moreover, for every $f \in C([a,b])$,

$$E_{m+1}^{\mathrm{alg}}(f;[a,b]) \le E_m^{\mathrm{alg}}(f;[a,b]), \qquad E_m^{\mathrm{alg}}(f;[a,b]) \longrightarrow 0.$$

Remark 2.49. For real-valued $f \in C([a,b])$, best approximation from the real polynomial space $\mathbb{P}_m$ is unique. If $f \notin \mathbb{P}_m$, a polynomial $p^* \in \mathbb{P}_m$ is the best approximant if and only if there are $a \le t_0 < \cdots < t_{m+1} \le b$ and $\varepsilon \in \{-1,1\}$ such that

$$f(t_j) - p^*(t_j) = \varepsilon(-1)^j \|f - p^*\|_{\infty,[a,b]}, \qquad j = 0,\dots,m+1.$$

This is the Chebyshev alternation theorem; see [122, Section 1.1]. The real-valued hypothesis is essential for this formulation.

Exactness and stability enter the error estimate through the following bounded-operator argument. Its quadrature forms are developed in [42, 29]; its interpolation form is Lebesgue's inequality [86, Section 4.2, Eq. (12)].

Theorem 2.50 (Exactness and stability). Let X, Y be normed linear spaces over the same scalar field, let $V_n \subseteq X$ be a linear subspace, and let $S, T_n : X \to Y$ be bounded linear operators satisfying

$$T_n v = Sv, \qquad v \in V_n. \tag{2.25}$$

Then, for every $f \in X$,

$$\begin{aligned} \|Sf - T_n f\|_Y &\le \|S - T_n\|_{X\to Y}\, E_X(f;V_n) \\ &\le \big(\|S\|_{X\to Y} + \|T_n\|_{X\to Y}\big) E_X(f;V_n). \end{aligned} \tag{2.26}$$

In particular,

$$\big(\|S\|_{X\to Y} + \|T_n\|_{X\to Y}\big) E_X(f;V_n) \longrightarrow 0 \quad \Longrightarrow \quad T_n f \longrightarrow Sf.$$

Proof. For every $v \in V_n$, exactness gives

$$Sf - T_n f = (S - T_n)(f - v).$$

Taking norms and then the infimum over $v \in V_n$ proves the first inequality. The second follows from $\|S - T_n\| \le \|S\| + \|T_n\|$. □

2.5.2 Quadrature as a bounded functional

A quadrature rule approximates an integral by a finite weighted sum of function values. Viewed as a bounded functional on a space of continuous functions, it fits the exactness–stability principle directly. Its norm is determined by the total variation of the associated discrete measure; for positive rules exact on constants, this norm equals the mass of the integration measure.

Definition 2.51. Let K be a compact Hausdorff space and μ a finite positive regular Borel measure on K, with $\mu(K) > 0$. In this subsection, $C(K) = C(K;\mathbb{C})$, and

$$I_\mu(f) = \int_K f \, \mathrm{d}\mu.$$

For nodes $x_{k,n} \in K$ and weights $w_{k,n} \in \mathbb{C}$, define

$$Q_n(f) = \sum_{k=1}^{N_n} w_{k,n} f(x_{k,n}), \qquad e_n(f) = I_\mu(f) - Q_n(f).$$

The quadrature rule is *exact* on a finite-dimensional linear space $V_n \subseteq C(K)$ if $Q_n(v) = I_\mu(v)$ for every $v \in V_n$. If $K = [a,b]$ and the rule is exact on constants, its *algebraic degree of exactness* is

$$\deg_{\mathrm{alg}}(Q_n) = \sup\bigl\{m \in \mathbb{N} \cup \{0\} : Q_n(p) = I_\mu(p) \text{ for every } p \in \mathbb{P}_m\bigr\},$$

where $+\infty$ is allowed. The rule is *positive* if all its weights are real and nonnegative. For real weights, the same definitions and estimates apply to real-valued functions.

The discrete measure associated with the rule is

$$\nu_n = \sum_{k=1}^{N_n} w_{k,n} \delta_{x_{k,n}}. \tag{2.27}$$

The norm identity in the Riesz representation theorem gives

$$\|I_\mu\| = \mu(K), \qquad \|Q_n\| = |\nu_n|(K), \qquad \|I_\mu - Q_n\| = |\mu - \nu_n|(K),$$

where the norms are functional norms on $C(K)$ and $|\nu|$ denotes the total-variation measure. More explicitly,

$$\|Q_n\| = \sum_{x \in \{x_{1,n},\dots,x_{N_n,n}\}} \left| \sum_{k:\, x_{k,n} = x} w_{k,n} \right| \le \sum_{k=1}^{N_n} |w_{k,n}|.$$

Thus, when the nodes are distinct, or coincident nodes have been consolidated by adding their weights,

$$\|Q_n\| = \sum_{k=1}^{N_n} |w_{k,n}|. \tag{2.28}$$

Without consolidation, cancellation at repeated nodes can make the last sum strictly larger than the functional norm.

Combining these norm identities with exactness gives the usual best-approximation estimate [42, 29]. The measure formulation covers complex weights on a general compact space K.

Proposition 2.52. If Q_n is exact on V_n, then, for every $f \in C(K)$,

$$|I_\mu(f) - Q_n(f)| \le \|I_\mu - Q_n\|\, E(f; V_n; K) \tag{2.29}$$

$$\le \bigl(\mu(K) + \|Q_n\|\bigr) E(f; V_n; K) \tag{2.30}$$

$$\le \left(\mu(K) + \sum_{k=1}^{N_n} |w_{k,n}|\right) E(f; V_n; K). \tag{2.31}$$

Proof. Apply Theorem 2.50 with $X = C(K)$, $Y = \mathbb{C}$, $S = I_\mu$, and $T_n = Q_n$, and use $\|I_\mu\| = \mu(K)$ and $\|Q_n\| \le \sum_k |w_{k,n}|$. □

For positive rules, exactness on constants fixes both the total weight and the functional norm. The stability factor is therefore independent of the number and positions of the nodes.

Corollary 2.53. Suppose that Q_n is positive, exact on V_n, and $1 \in V_n$. Then

$$\sum_{k=1}^{N_n} w_{k,n} = \mu(K), \qquad \|Q_n\| = \mu(K),$$

and, for every $f \in C(K)$,

$$|I_\mu(f) - Q_n(f)| \le 2\mu(K) E(f; V_n; K). \tag{2.32}$$

For a probability measure, the coefficient is 2.

Proof. Exactness on constants gives $\sum_k w_{k,n} = Q_n(1) = I_\mu(1) = \mu(K)$. Positivity yields $\|Q_n\| = Q_n(1) = \mu(K)$. Substitute in (2.30). □

Uniformly bounded functional norms turn approximation by the exactness spaces into convergence of the quadrature rules [42, 29].

Corollary 2.54. Let Q_n be exact on V_n. For a fixed $f \in C(K)$,

$$E(f;V_n;K) \longrightarrow 0, \qquad \sup_n \|Q_n\| < \infty \quad \Longrightarrow \quad Q_n(f) \longrightarrow I_\mu(f).$$

For positive rules exact on constants, it is enough that $E(f;V_n;K) \to 0$.

If $E(f;V_n;K) \to 0$ for every $f \in C(K)$, convergence of the rules for every continuous function is equivalent to $\sup_n \|Q_n\| < \infty$: sufficiency follows from the corollary, and necessity from the uniform boundedness principle.

The spaces are *nested* when $V_n \subseteq V_{n+1}$ for every index. If their union is dense, any chosen approximant belongs to all sufficiently large spaces, which gives the approximation hypothesis. In particular, on an interval it suffices that $V_n \supseteq \mathbb{P}_{m_n}$, with $m_n \to \infty$. For nonnested spaces, density of the union alone does not ensure approximation along the full sequence.

Convergence for each fixed integrand is weaker than convergence in functional norm. For positive rules exact on constants, the following identity shows why the latter can fail. If $\mu(\{x_{k,n}\}) = 0$ at every quadrature node, then μ and ν_n are mutually singular and

$$\|I_\mu - Q_n\| = |\mu - \nu_n|(K) = \mu(K) + \nu_n(K) = 2\mu(K).$$

This applies, in particular, to nonatomic measures. The error norm then remains constant even when $Q_n(f) \to I_\mu(f)$ for every fixed $f \in C(K)$.

2.5.3 Interpolation projectors and Lebesgue constants

An *interpolation array* is a sequence of finite node sets, arranged as rows indexed by the approximation order. Within each row the nodes are distinct; no inclusion relation between successive rows is required. Polynomial interpolation maps the values prescribed at one row of nodes to the unique polynomial of degree at most one less than the number of nodes. Its stability is measured by the Lebesgue constant.

Definition 2.55. Let

$$\mathcal{X} = (x_{k,n})_{1\le k\le N_n,\ n\ge 1}, \qquad x_{k,n} \in [a,b], \qquad N_n \to \infty,$$

be an interpolation array with distinct nodes in each row. Define the nodal polynomial

$$\omega_n(x) = \prod_{j=1}^{N_n} (x - x_{j,n}) \tag{2.33}$$

and the fundamental Lagrange polynomials

$$\ell_{k,n}(x) = \prod_{\substack{1\le j\le N_n \\ j\ne k}} \frac{x - x_{j,n}}{x_{k,n} - x_{j,n}} = \frac{\omega_n(x)}{\omega_n'(x_{k,n})(x - x_{k,n})}. \tag{2.34}$$

The last quotient is interpreted by continuity at $x = x_{k,n}$. For prescribed nodal values $f(x_{k,n})$, the unique interpolating polynomial in $\mathbb{P}_{N_n-1}$ is

$$L_n f(x) = \sum_{k=1}^{N_n} f(x_{k,n})\ell_{k,n}(x).$$

The Lebesgue function and Lebesgue constant are

$$\Lambda_n(x) = \sum_{k=1}^{N_n} |\ell_{k,n}(x)|, \qquad \Lambda_n = \max_{x\in[a,b]} \Lambda_n(x).$$

The barycentric form represents the same polynomial as a quotient of two weighted sums. We use the classical second barycentric formula [20, Eqs. (3.2) and (4.2)]; earlier formulations and historical discussion appear in [86, Section 4.1].

Proposition 2.56. Set

$$\omega_{k,n} = \frac{1}{\omega_n'(x_{k,n})}, \qquad k = 1, \dots, N_n. \tag{2.35}$$

For $x \notin \{x_{1,n}, \dots, x_{N_n,n}\}$,

$$L_n f(x) = \frac{\displaystyle\sum_{k=1}^{N_n} \frac{\omega_{k,n}}{x - x_{k,n}} f(x_{k,n})}{\displaystyle\sum_{k=1}^{N_n} \frac{\omega_{k,n}}{x - x_{k,n}}}. \tag{2.36}$$

The denominator is nonzero away from the nodes. At each node the quotient has a removable singularity and extends to the prescribed value. A common nonzero rescaling of all weights leaves the quotient unchanged.

Proof. The cardinal-polynomial formula gives, away from the nodes,

$$L_n f(x) = \omega_n(x) \sum_{k=1}^{N_n} \frac{\omega_{k,n}}{x - x_{k,n}} f(x_{k,n}).$$

Since $\sum_k \ell_{k,n}(x) = 1$, the same identity for the constant polynomial gives

$$\sum_{k=1}^{N_n} \frac{\omega_{k,n}}{x - x_{k,n}} = \frac{1}{\omega_n(x)} \neq 0.$$

Division proves the formula. Its extension at the nodes equals the polynomial $L_n f$, and a common scaling cancels in the quotient. □

Remark 2.57. Lagrange, Newton, barycentric, and suitable orthogonal-expansion forms can represent the same polynomial while behaving differently in finite precision. The Lebesgue constant measures the norm of the exact interpolation operator, not the rounding error of its implementation; see [20, 95]. The polynomial barycentric formula uses the weights (2.35), up to a common factor. Arbitrary nonproportional weights generally define a different, rational interpolant rather than another representation of L_n.

The interpolation operator reproduces its polynomial range:

$$L_n p = p, \qquad p \in \mathbb{P}_{N_n - 1}. \tag{2.37}$$

The norm identity and the resulting Lebesgue inequality are classical; see [122, Section 1.4], [20, Section 9.1], and [86, Section 4.2, Eq. (12)]. The proof also identifies the exact norm of the error operator on a nondegenerate interval.

Proposition 2.58. As linear operators on $C([a,b])$, with the uniform norm,

$$\|L_n\| = \Lambda_n, \qquad \|I - L_n\| = 1 + \Lambda_n. \tag{2.38}$$

Consequently, for every $f \in C([a,b])$,

$$\|f - L_n f\|_{\infty,[a,b]} \leq (1 + \Lambda_n) E^{\mathrm{alg}}_{N_n - 1}(f; [a,b]). \tag{2.39}$$

Proof. For fixed $x \in [a,b]$, the functional $f \mapsto L_n f(x)$ is represented by $\sum_k \ell_{k,n}(x) \delta_{x_{k,n}}$. Its norm is the sum of the absolute values of the coefficients, namely $\Lambda_n(x)$, because the nodes are distinct. Taking the supremum over x gives $\|L_n\| = \Lambda_n$.

If x is not a node, the error-evaluation functional is represented by

$$\delta_x - \sum_{k=1}^{N_n} \ell_{k,n}(x)\delta_{x_{k,n}},$$

whose norm is $1 + \Lambda_n(x)$. The complement of the finite node set is dense in $[a,b]$, and $\Lambda_n(x)$ is continuous. Therefore $\|I - L_n\| = 1 + \Lambda_n$. Finally, apply Theorem 2.50 with $S = I$, $T_n = L_n$, and $V_n = \mathbb{P}_{N_n-1}$, using (2.37). □

Unlike the norms of positive quadrature rules, the norms of polynomial interpolation operators cannot remain uniformly bounded as the number of nodes increases. Faber's theorem [65] gives the resulting obstruction to convergence for all continuous functions; see also [122, Section 1.4] and [86, Section 4.2].

Theorem 2.59 (Faber's obstruction). There is an absolute constant $c > 0$, independent of the interval, the number of nodes, and their positions, such that every interpolation array as above satisfies

$$\Lambda_n \geq c \log N_n. \tag{2.40}$$

Consequently, $\{L_n\}_n$ is not uniformly stable on $C([a,b])$, and for every such array there exists $f \in C([a,b])$ for which

$$\limsup_{n\to\infty} \|L_n f\|_{\infty,[a,b]} = \infty. \tag{2.41}$$

Deduction of the divergence assertion. The logarithmic lower bound is the cited classical result; affine changes of interval preserve Lebesgue constants. Since $N_n \to \infty$, $\|L_n\| = \Lambda_n$ is unbounded. The uniform boundedness principle on the Banach space $C([a,b])$ yields an f such that $\sup_n \|L_n f\|_{\infty,[a,b]} = \infty$. Every individual norm is finite, so the limsup is infinite as well. □

Lebesgue's inequality nevertheless gives the following sufficient criterion for a fixed function.

Corollary 2.60. For $f \in C([a,b])$,

$$(1 + \Lambda_n) E^{\mathrm{alg}}_{N_n-1}(f;[a,b]) \longrightarrow 0 \quad \Longrightarrow \quad \|L_n f - f\|_{\infty,[a,b]} \longrightarrow 0.$$

This condition is not necessary. For example, on $[0,1]$, the usual interpolation remainder formula [86, Eq. (11)] gives

$$\|\mathrm{e}^x - L_n(\mathrm{e}^x)\|_{\infty,[0,1]} \leq \frac{\mathrm{e}}{N_n!}$$

for every choice of distinct nodes. For each fixed $N_n \geq 2$, however, $E^{\mathrm{alg}}_{N_n-1}(\mathrm{e}^x;[0,1]) > 0$, while Λ_n can be made arbitrarily large by letting two nodes approach each other with the remaining nodes fixed. Thus an array can be chosen for which interpolation of e^x converges but the product in the corollary does not tend to zero. Faber's theorem concerns failure for some continuous function on each fixed array.

The later analysis of weighted interpolation uses the same residual argument in a specified weighted norm. Its application requires bounded point evaluations, a bounded interpolation operator, and reproduction of the approximation space in that setting.

Positive quadrature exact on constants has a uniform stability factor equal to $\mu(K)$, so approximation by its exactness spaces suffices for convergence. Polynomial interpolation has stability factor Λ_n, which is necessarily unbounded along every node array. A decay estimate for $(1 + \Lambda_n) E^{\mathrm{alg}}_{N_n-1}(f;[a,b])$ is one way to prove interpolation convergence for a particular function, but it is not the only one.

2.6 Gauss–Christoffel and Szegő quadrature

Orthogonality provides quadrature rules with high exactness, while finite-dimensional spectral problems provide their nodes and weights. The exactness–stability principle of Section 2.5 then gives convergence. On an interval, the nodes are zeros of orthogonal polynomials. On the unit circle, those zeros lie in the open disk, so quadrature instead uses para-orthogonal polynomials whose zeros lie on the circle. Throughout, $n \geq 1$.

2.6.1 Gauss–Christoffel quadrature on an interval

Let μ be a finite positive Borel measure with infinite support contained in a compact interval $[a, b]$, and set

$$M = \mu([a, b]), \qquad I_\mu(f) = \int_a^b f(x)\, \mathrm{d}\mu(x).$$

Let π_j be the monic orthogonal polynomial of degree j, with $\pi_0 = 1$, and let p_j be its orthonormal counterpart with positive leading coefficient. With $p_{-1} = 0$, $p_0 = M^{-1/2}$, the recurrence is

$$xp_j(x) = \beta_{j+1}p_{j+1}(x) + \alpha_j p_j(x) + \beta_j p_{j-1}(x), \qquad j \geq 0, \tag{2.42}$$

where $\alpha_j \in \mathbb{R}$, $\beta_0 = 0$, and $\beta_j > 0$ for $j \geq 1$. These are the standard three-term recurrence coefficients [85, 122]. Define

$$J_n = \begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}, \qquad K_m(x, y) = \sum_{j=0}^{m} p_j(x)p_j(y),$$

with $J_1 = [\alpha_0]$.

The recurrence identifies Gaussian nodes with eigenvalues of the Jacobi matrix and recovers their weights from its eigenvectors. The quadrature construction originates with Gauss [83]; the zero-location, interlacing, and Christoffel-number results are developed in [85, 122]. The Jacobi eigenvalue formulation leads to the Golub–Welsch algorithm [90].

Theorem 2.61 (Gauss–Christoffel nodes and weights). The zeros $x_{1,n} < \cdots < x_{n,n}$ of π_n are simple, lie in the interior of $\operatorname{conv}(\operatorname{supp}\mu)$, and are the eigenvalues of J_n. Consecutive orthogonal polynomial zeros strictly interlace. The positive Christoffel numbers

$$\lambda_{k,n} = K_{n-1}(x_{k,n}, x_{k,n})^{-1} \tag{2.43}$$

define $G_n(f) = \sum_{k=1}^{n} \lambda_{k,n} f(x_{k,n})$. For a normalized eigenvector $v_{k,n}$ of J_n associated with $x_{k,n}$,

$$\lambda_{k,n} = M|(v_{k,n})_1|^2. \tag{2.44}$$

The convex hull in this statement allows measures with gaps in their support, since nodes may lie in those gaps. The resulting rule has the maximal algebraic degree of exactness among rules with the same number of function evaluations [85, Section 1.4.2].

Theorem 2.62 (Gaussian exactness and optimality). The rule G_n has algebraic degree of exactness $2n - 1$:

$$G_n(p) = I_\mu(p), \qquad p \in \mathbb{P}_{2n-1}. \tag{2.45}$$

No function-value rule with at most n nodes in $[a, b]$ is exact on all of $\mathbb{P}_{2n}$.

Deduction of optimality. The exactness assertion is the cited Gaussian quadrature theorem. For any nodes $y_1, \ldots, y_n \in [a, b]$, the polynomial

$$q(x) = \prod_{k=1}^{n} (x - y_k)^2$$

has zero quadrature sum, whereas $I_\mu(q) > 0$ because μ has infinite support. A rule with fewer nodes is covered by repeating nodes in this product. Thus exactness on $\mathbb{P}_{2n}$ is impossible. □

Corollary 2.63. For every $f \in C([a,b])$,

$$|I_\mu(f) - G_n(f)| \le 2M E^{\mathrm{alg}}_{2n-1}(f;[a,b]). \tag{2.46}$$

Consequently, $G_n(f) \to I_\mu(f)$ for every continuous f.

Proof. Positivity and exactness on constants give $\|G_n\| = M$. Corollary 2.53 gives the estimate; Weierstrass approximation gives convergence. □

2.6.2 Szegő quadrature on the unit circle

Let $\Gamma = \{z \in \mathbb{C} : |z| = 1\}$, let ν be a finite positive Borel measure on Γ with infinite support, and set

$$M_\Gamma = \nu(\Gamma), \qquad I_\nu(f) = \int_\Gamma f \, \mathrm{d}\nu, \qquad \langle f, g\rangle_\nu = \int_\Gamma f\overline{g} \, \mathrm{d}\nu.$$

The inner product is linear in its first argument. Let Φ_j be the monic orthogonal polynomial of degree j, with $\Phi_0 = 1$. For a polynomial q of degree j, write

$$q^*(z) = z^j \overline{q(1/\overline{z})},$$

interpreted as a polynomial also at zero. We use the plus-sign recurrence

$$\Phi_j = z\Phi_{j-1} + \gamma_j \Phi^*_{j-1}, \qquad \Phi^*_j = \overline{\gamma_j} z\Phi_{j-1} + \Phi^*_{j-1}, \qquad j \ge 1, \tag{2.47}$$

where $\gamma_j = \Phi_j(0)$ and $|\gamma_j| < 1$. Define

$$\delta_j = \|\Phi_j\|^2_{L^2(\nu)}, \qquad \varphi_j = \Phi_j/\sqrt{\delta_j}, \qquad \delta_0 = M_\Gamma, \qquad \delta_j = \delta_{j-1}(1 - |\gamma_j|^2).$$

This is $\gamma_j = -\overline{\alpha_{j-1}}$ in the standard Verblunsky-coefficient convention [161, Section 1.5]; the plus-sign notation is also used in [102, Eqs. (1.4)–(1.7)].

For $\tau \in \Gamma$, define the para-orthogonal polynomial

$$\Psi_n(z;\tau) = z\Phi_{n-1}(z) + \tau \Phi^*_{n-1}(z). \tag{2.48}$$

Here τ is the *terminal recurrence parameter*. In the alternative parametrization $\Phi_n + \eta\Phi^*_n$, $\eta \in \Gamma$, the recurrence gives

$$\Phi_n + \eta\Phi^*_n = (1 + \eta\overline{\gamma_n})\Psi_n(\cdot;\tau), \qquad \tau = \frac{\gamma_n + \eta}{1 + \overline{\gamma_n}\eta}.$$

Since $|\gamma_n| < 1$, the prefactor is nonzero and the two polynomials have the same zeros. Their parameters are related by the displayed fractional-linear map: a prescribed η is converted through this map rather than by setting $\tau = \eta$. In the convention $z\Phi_{n-1} - \overline{\beta}\Phi^*_{n-1}$, one instead has $\tau = -\overline{\beta}$ [162].

For integers $r \le s$, set $\Lambda_{r,s} = \operatorname{span}\{z^j : r \le j \le s\}$, and define

$$E^\Gamma_m(f) = \inf_{L \in \Lambda_{-m,m}} \|f - L\|_{\infty,\Gamma}. \tag{2.49}$$

Laurent polynomials form a self-adjoint algebra containing the constants and separating points. Stone-Weierstrass therefore gives $E^\Gamma_m(f) \to 0$ for every $f \in C(\Gamma)$.

The unit-circle counterpart of the Jacobi matrix is a unitary Hessenberg matrix. We use Gragg's construction [92], in the convention of [102, Eqs. (2.2)–(2.3)]. The associated moment and quadrature theory is developed in [105].

Definition 2.64 (Szegő Hessenberg matrix). Set $\gamma_0 = 1$ and $\mathcal{D}_n = \operatorname{diag}(\delta_0, \dots, \delta_{n-1})$. For $1 \le i, j \le n$, define

$$[\widehat{H}_n(\tau)]_{ij} = \begin{cases} -\overline{\gamma_{i-1}}\gamma_j, & i \le j < n, \\ -\overline{\gamma_{i-1}}\tau, & j = n, \\ 1 - |\gamma_j|^2, & i = j+1, \\ 0, & i > j+1, \end{cases}$$

and set

$$H_n(\tau) = \mathcal{D}_n^{-1/2} \widehat{H}_n(\tau) \mathcal{D}_n^{1/2}. \tag{2.50}$$

In particular, $H_1(\tau) = [-\tau]$.

The paraorthogonal-node and reproducing-kernel weight theory is developed in [105]. For the Hessenberg realization above, the eigenvector expression for the weights is obtained by the recurrence and normalization argument below. The probability-normalized formula in [102, Proposition 2.1] is written below with the total mass retained explicitly.

Theorem 2.65 (Szegő nodes and weights). For every $\tau \in \Gamma$, $H_n(\tau)$ is unitary and

$$\det(zI_n - H_n(\tau)) = \Psi_n(z;\tau).$$

Its eigenvalues $z_{1,n}(\tau), \dots, z_{n,n}(\tau)$ are simple and lie on Γ. Define

$$\mathcal{K}_m(z,w) = \sum_{j=0}^{m} \varphi_j(z)\overline{\varphi_j(w)}.$$

The positive weights

$$\lambda_{k,n}(\tau) = \mathcal{K}_{n-1}(z_{k,n}(\tau), z_{k,n}(\tau))^{-1} \tag{2.51}$$

define $S_{n,\tau}(f) = \sum_{k=1}^{n} \lambda_{k,n}(\tau) f(z_{k,n}(\tau))$. For every normalized right eigenvector $u_{k,n}(\tau)$ at this node,

$$\lambda_{k,n}(\tau) = M_\Gamma |(u_{k,n}(\tau))_1|^2. \tag{2.52}$$

Verification of conventions and weight normalization. The cited Hessenberg construction gives unitarity and the characteristic polynomial with terminal parameter τ. Its subdiagonal entries are $\sqrt{1-|\gamma_j|^2} > 0$, so the matrix is unreduced Hessenberg and has a cyclic vector. A cyclic normal matrix has simple eigenvalues; unitarity puts them on Γ.

At a zero z of $\Psi_n(\cdot;\tau)$, set $q(z) = (\varphi_0(z), \dots, \varphi_{n-1}(z))^T$. The recurrence gives

$$H_n(\tau)^T q(z) = z q(z).$$

Conjugation and unitarity imply $H_n(\tau)\overline{q(z)} = z\overline{q(z)}$. Hence a normalized right eigenvector is

$$u(z) = \frac{\overline{q(z)}}{\sqrt{\mathcal{K}_{n-1}(z,z)}}.$$

Since $|\varphi_0(z)|^2 = M_\Gamma^{-1}$, the two weight formulas agree. The probability-normalized spectral weight formula scales by M_Γ when the measure is rescaled; the Hessenberg matrix itself is unchanged. This also gives positivity and total weight M_Γ. □

Laurent exactness is the classical Szegő quadrature theorem [105, 92]; see also [102, Eq. (1.10)].

Theorem 2.66 (Laurent exactness and optimality). For every $\tau \in \Gamma$,

$$S_{n,\tau}(L) = I_\nu(L), \qquad L \in \Lambda_{-(n-1),n-1}. \tag{2.53}$$

No function-value rule with at most n nodes on Γ is exact on all of $\Lambda_{-n,n}$.

The optimality assertion follows by applying any proposed rule to $|\prod_{k=1}^{n}(z-z_k)|^2$, viewed on Γ as an element of $\Lambda_{-n,n}$. Its quadrature sum is zero and its integral is positive.

Corollary 2.67. For every $f \in C(\Gamma)$,

$$\sup_{\tau\in\Gamma}|I_\nu(f) - S_{n,\tau}(f)| \le 2M_\Gamma E_{n-1}^{\Gamma}(f) \longrightarrow 0. \tag{2.54}$$

In particular, convergence holds for every parameter sequence $\tau_n \in \Gamma$.

Proof. Apply Corollary 2.53 to $K = \Gamma$ and $V_n = \Lambda_{-(n-1),n-1}$, then use Laurent polynomial density. The estimate is independent of τ. □

The terminal parameter can be chosen to include a prescribed point among the nodes. When the same point is prescribed at consecutive orders, the remaining nodes satisfy a strict cyclic interlacing property [162, Theorem 1.4(ii)].

Proposition 2.68. For every $z_p \in \Gamma$, the unique parameter making z_p a zero of $\Psi_n(\cdot;\tau)$ is

$$\tau_n(z_p) = -z_p\frac{\Phi_{n-1}(z_p)}{\Phi_{n-1}^*(z_p)} \in \Gamma. \tag{2.55}$$

When the same node is prescribed at orders n and $n+1$, it is their only common node. The n nodes at order n strictly interlace in cyclic order with the n nonprescribed nodes at order $n+1$.

Proof. The zeros of Φ_{n-1} lie in the open disk [161, Section 1.7]. Both polynomial values in the quotient are therefore nonzero on Γ, and their moduli agree there. Solving $\Psi_n(z_p;\tau) = 0$ gives the unique displayed unimodular parameter. The common-node and interlacing assertions are exactly the cited consecutive-degree theorem. □

For normalized Lebesgue measure, the orthogonal polynomials are monomials and the nodes form a rotated uniform grid [102, Examples 1.1 and 2.1].

Corollary 2.69. For $\mathrm{d}\nu = \mathrm{d}\theta/(2\pi)$, $z = \mathrm{e}^{\mathrm{i}\theta}$,

$$\Phi_j(z) = \varphi_j(z) = z^j, \qquad \gamma_j = 0 \ (j \ge 1), \qquad \Psi_n(z;\tau) = z^n + \tau.$$

The nodes form an equally spaced rotated grid, every weight is $1/n$, and $\tau_n(z_p) = -z_p^n$.

2.7 Anti-Gauss and anti-Szegő quadrature

An anti-quadrature rule is designed to cancel the error of a base rule on a specified space. Laurie's anti-Gauss construction [114] and the anti-Szegő construction of Kim–Reichel [108] produce positive companions whose errors are negative multiples of the base errors. A suitable weighted average therefore has a larger exactness space, while the difference of the two rules estimates the base error. The following proposition gives the common algebraic argument.

Proposition 2.70 (Opposite-error pairing). Let X be a linear space of functions, let $W_n \subseteq X$ be a linear subspace, and let $I, Q_n, \widetilde{Q}_n : X \to \mathbb{C}$ be linear functionals. Suppose that $\rho_n > 0$ and

$$\widetilde{Q}_n(v) = (1+\rho_n)I(v) - \rho_n Q_n(v), \qquad v \in W_n. \tag{2.56}$$

Then

$$I(v) - \widetilde{Q}_n(v) = -\rho_n\big(I(v) - Q_n(v)\big), \qquad v \in W_n, \tag{2.57}$$

and

$$I(v) - Q_n(v) = \frac{\widetilde{Q}_n(v) - Q_n(v)}{1 + \rho_n}, \qquad v \in W_n. \tag{2.58}$$

The convex average

$$Q_n^{\mathrm{avg}} = \frac{\rho_n Q_n + \widetilde{Q}_n}{1 + \rho_n} \tag{2.59}$$

is exact on W_n. If Q_n is exact on $V_n \subseteq W_n$, so is $\widetilde{Q}_n$.

Proof. Subtract the assumed identity from $I(v)$ and rearrange. Furthermore,

$$I - Q_n^{\mathrm{avg}} = \frac{\rho_n(I - Q_n) + (I - \widetilde{Q}_n)}{1 + \rho_n}$$

vanishes on W_n. Substituting $Q_n(v) = I(v)$ gives the last assertion. □

For a general function, the estimator error is exactly the error of the averaged rule:

$$(I - Q_n)(f) - \frac{(\widetilde{Q}_n - Q_n)(f)}{1 + \rho_n} = (I - Q_n^{\mathrm{avg}})(f), \qquad f \in X.$$

If X is normed and these functionals are bounded, its modulus is at most $\|I - Q_n^{\mathrm{avg}}\| E_X(f; W_n)$, by Theorem 2.50. This is an absolute error bound; relative accuracy also depends on the size of $(I - Q_n)(f)$. For real-valued functionals on real $v \in W_n$, the opposite signs of the errors place $I(v)$ between the two rule values for every $\rho_n > 0$. Outside W_n, bracketing requires further information about the integrand.

2.7.1 Anti-Gauss quadrature

Retain the interval measure μ, its mass M, and the Jacobi coefficients from Section 2.6.1. With $\mathbf{e}_n = (0, \dots, 0, 1)^T \in \mathbb{R}^n$, define

$$\widetilde{J}_{n+1} = \begin{bmatrix} J_n & \sqrt{2}\,\beta_n \mathbf{e}_n \\ \sqrt{2}\,\beta_n \mathbf{e}_n^T & \alpha_n \end{bmatrix}. \tag{2.60}$$

Only the final off-diagonal coupling of J_{n+1} is changed. The next theorem is Laurie's anti-Gauss construction [114]; see also [48, Section 2.3.2] for this Jacobi-matrix formulation.

Theorem 2.71 (Anti-Gauss quadrature). The matrix $\widetilde{J}_{n+1}$ has simple eigenvalues $\widetilde{x}_{1,n+1} < \cdots < \widetilde{x}_{n+1,n+1}$. For normalized eigenvectors $\widetilde{v}_{k,n+1}$, set

$$\widetilde{\lambda}_{k,n+1} = M|(\widetilde{v}_{k,n+1})_1|^2 > 0, \qquad \widetilde{G}_{n+1}(f) = \sum_{k=1}^{n+1} \widetilde{\lambda}_{k,n+1} f(\widetilde{x}_{k,n+1}).$$

The weights sum to M, and the rule is defined whenever f is defined at all these nodes. Its monic nodal polynomial is

$$\widetilde{\pi}_{n+1}(x) = \pi_{n+1}(x) - \beta_n^2 \pi_{n-1}(x), \tag{2.61}$$

or, equivalently,

$$\widetilde{\pi}_{n+1}(x) = 2\pi_{n+1}(x) + (\alpha_n - x)\pi_n(x). \tag{2.62}$$

The nodes strictly interlace as

$$\widetilde{x}_{1,n+1} < x_{1,n} < \widetilde{x}_{2,n+1} < \cdots < x_{n,n} < \widetilde{x}_{n+1,n+1}. \tag{2.63}$$

Only the two extreme anti-Gauss nodes can lie outside $[a, b]$. Moreover,

$$I_\mu(p) - \widetilde{G}_{n+1}(p) = -\big(I_\mu(p) - G_n(p)\big), \qquad p \in \mathbb{P}_{2n+1}. \tag{2.64}$$

The algebraic degree of exactness of $\widetilde{G}_{n+1}$ is $2n - 1$.

The monic recurrence is $\pi_{n+1} = (x-\alpha_n)\pi_n - \beta_n^2\pi_{n-1}$; doubling its last coefficient β_n^2 gives both displayed nodal-polynomial formulas. Strict interlacing follows from the irreducible symmetric tridiagonal matrix and its leading principal submatrix J_n. Exactness on $\mathbb{P}_{2n-1}$ follows from the cited opposite-error identity; it fails on π_n^2, since $G_n(\pi_n^2) = 0$ and $\widetilde{G}_{n+1}(\pi_n^2) = 2I_\mu(\pi_n^2) > I_\mu(\pi_n^2)$.

Corollary 2.72. The positive averaged rule

$$G_n^{\mathrm{avg}} = \tfrac{1}{2}(G_n + \widetilde{G}_{n+1}) \tag{2.65}$$

is exact on $\mathbb{P}_{2n+1}$, and

$$I_\mu(p) - G_n(p) = \tfrac{1}{2}\big(\widetilde{G}_{n+1}(p) - G_n(p)\big), \qquad p \in \mathbb{P}_{2n+1}. \tag{2.66}$$

This is Proposition 2.70 with $\rho_n = 1$. The average can have higher exactness for special measures; $2n+1$ is a guaranteed lower bound, not always its exact degree. If all anti-Gauss nodes lie in $[a,b]$, positivity also gives

$$|I_\mu(f) - G_n^{\mathrm{avg}}(f)| \le 2ME_{2n+1}^{\mathrm{alg}}(f;[a,b]), \qquad f \in C([a,b]).$$

If there are exterior nodes, a function given only on $[a,b]$ must first be extended to them. The resulting operator on $C([a,b])$ then depends on the chosen extension.

2.7.2 Anti-Szegő quadrature

Retain ν, M_Γ, and the plus-sign terminal convention of Subsection 2.6.2. The anti-Szegő companion is obtained by changing the terminal parameter of a Szegő rule for the same measure. The parameter is chosen so that the errors on the first nonexact moments are negative multiples of one another. For $c > 0$, this requirement leads to the auxiliary functional on $\Lambda_{-n,n}$,

$$I_{n,\tau,c}^{\mathrm{AS}} = (1+c)I_\nu - cS_{n,\tau}. \tag{2.67}$$

It agrees with I_ν on $\Lambda_{-(n-1),n-1}$, so the lower-order data are unchanged. The degree-n orthogonality condition then determines the formal terminal coefficient

$$\widetilde{\tau}_{n,\tau,c} = (1+c)\gamma_n - c\tau = \gamma_n + c(\gamma_n - \tau). \tag{2.68}$$

To obtain a Szegő rule, this coefficient must lie on the unit circle. That condition selects the error ratio and the companion parameter in the theorem below. At this stage the auxiliary functional is used only on the stated Laurent space; the positive companion will represent it on that space.

This is the Kim–Reichel construction [108], also described in [102, Section 5]. The proof first determines the parameter, then verifies the error identity on the two moments not already covered by the base exactness.

Theorem 2.73 (Anti-Szegő quadrature). For every $\tau \in \Gamma$, the unique positive c for which $|\widetilde{\tau}_{n,\tau,c}| = 1$ is

$$c_{n,\tau} = \frac{1-|\gamma_n|^2}{|\tau-\gamma_n|^2} > 0. \tag{2.69}$$

The resulting parameter is

$$\widetilde{\tau}_{n,\tau} = \gamma_n + c_{n,\tau}(\gamma_n - \tau) = \frac{\gamma_n - \tau}{1 - \overline{\gamma_n}\tau} \in \Gamma. \tag{2.70}$$

Define the companion as another Szegő rule for the *same* measure:

$$\widetilde{S}_{n,\tau} := S_{n,\widetilde{\tau}_{n,\tau}}.$$

Its nodes $\widetilde{z}_{k,n}(\tau)$ are the zeros of $\Psi_n(\cdot;\widetilde{\tau}_{n,\tau})$, equivalently the eigenvalues of $H_n(\widetilde{\tau}_{n,\tau})$, and its weights are

$$\widetilde{\lambda}_{k,n}(\tau) = \mathcal{K}_{n-1}(\widetilde{z}_{k,n}(\tau), \widetilde{z}_{k,n}(\tau))^{-1} > 0.$$

They sum to M_Γ. The two node sets strictly interlace in cyclic order, and

$$I_\nu(L) - \widetilde{S}_{n,\tau}(L) = -c_{n,\tau}\big(I_\nu(L) - S_{n,\tau}(L)\big), \qquad L \in \Lambda_{-n,n}. \tag{2.71}$$

In particular,

$$\widetilde{S}_{n,\tau}(L) = I_\nu(L), \qquad L \in \Lambda_{-(n-1),n-1}. \tag{2.72}$$

Equivalently, the companion represents $I^{\mathrm{AS}}_{n,\tau,c_{n,\tau}}$ on $\Lambda_{-n,n}$.

Proof. Write $\gamma = \gamma_n$. Direct expansion gives

$$|\gamma + c(\gamma - \tau)|^2 - 1 = |\gamma - \tau|^2(c+1)(c - c_{n,\tau}).$$

Since $|\gamma| < 1$, the displayed $c_{n,\tau}$ is positive and is the unique positive root. Simplification gives the fractional formula for $\widetilde{\tau}_{n,\tau}$. Also $\widetilde{\tau}_{n,\tau} - \tau = (1 + c_{n,\tau})(\gamma - \tau) \neq 0$. Positivity and base exactness follow from Szegő quadrature, while strict cyclic interlacing follows from [162, Theorem 1.3].

To verify the enlarged-space error identity, note that

$$I_\nu(\Phi^*_{n-1}) = \langle z^{n-1}, \Phi_{n-1}\rangle_\nu = \delta_{n-1}, \qquad I_\nu(\Phi_n) = 0.$$

Since $\Psi_n(\cdot;\tau) = \Phi_n + (\tau - \gamma_n)\Phi^*_{n-1}$ is monic and vanishes at every node of $S_{n,\tau}$, lower-degree exactness gives

$$(I_\nu - S_{n,\tau})(z^n) = (\tau - \gamma_n)\delta_{n-1}.$$

Replacing τ by $\widetilde{\tau}_{n,\tau}$ multiplies this error by $-c_{n,\tau}$. The corresponding negative moment is its complex conjugate because both rules and the measure are positive. All intervening moments are exact, proving the identity on $\Lambda_{-n,n}$. □

The representation by the companion explains the scope of the auxiliary functional. At the selected coefficient it is nonnegative on $|q|^2$ for $q \in \mathbb{P}_n$, since these values are computed by a positive atomic rule. It is only semidefinite on $\mathbb{P}_n$: the nodal polynomial $\Psi_n(\cdot;\widetilde{\tau}_{n,\tau})$ has zero squared norm. Thus the construction terminates at this degree rather than producing an infinite orthogonal-polynomial sequence.

Corollary 2.74. The positive weighted average

$$S^{\mathrm{avg}}_{n,\tau} = \frac{c_{n,\tau}S_{n,\tau} + \widetilde{S}_{n,\tau}}{1 + c_{n,\tau}} \tag{2.73}$$

is exact on $\Lambda_{-n,n}$, and

$$I_\nu(L) - S_{n,\tau}(L) = \frac{\widetilde{S}_{n,\tau}(L) - S_{n,\tau}(L)}{1 + c_{n,\tau}}, \qquad L \in \Lambda_{-n,n}. \tag{2.74}$$

Moreover, for every $f \in C(\Gamma)$,

$$\sup_{\tau\in\Gamma}|I_\nu(f) - S^{\mathrm{avg}}_{n,\tau}(f)| \le 2M_\Gamma E^\Gamma_n(f) \longrightarrow 0.$$

The opposite-error pairing gives exactness of the average, and the positive-rule estimate gives convergence. Even when $c_{n,\tau}$ varies, the average remains convex and its total weight is M_Γ, so its norm is uniformly bounded.

Corollary 2.75. As functionals on $\Lambda_{-n,n}$, the two errors are negatives of each other if and only if $c_{n,\tau} = 1$. Equivalently,

$$|2\gamma_n - \tau| = 1 \quad \Longleftrightarrow \quad \mathrm{Re}(\gamma_n \overline{\tau}) = |\gamma_n|^2. \tag{2.75}$$

In this case $\widetilde{\tau}_{n,\tau} = 2\gamma_n - \tau$, and the exact average is the arithmetic mean $S^{\mathrm{avg}}_{n,\tau} = \frac{1}{2}(S_{n,\tau} + \widetilde{S}_{n,\tau})$.

For the converse, compare the two nonzero errors using $(\tau - \gamma_n)\delta_{n-1}$ on z^n. For real-valued $L \in \Lambda_{-n,n}$, every $c_{n,\tau} > 0$ gives bracketing; the balanced choice additionally makes the error magnitudes equal and the arithmetic mean exact.

Corollary 2.76. For normalized Lebesgue measure on Γ,

$$\gamma_n = 0, \qquad c_{n,\tau} = 1, \qquad \widetilde{\tau}_{n,\tau} = -\tau.$$

The two nodal polynomials are $z^n + \tau$ and $z^n - \tau$, and both rules have weights $1/n$. Writing $-\tau = \mathrm{e}^{\mathrm{i}\vartheta}$, their node arguments are respectively

$$\frac{\vartheta + 2k\pi}{n}, \qquad \frac{\vartheta + (2k+1)\pi}{n}, \qquad k = 0, \dots, n-1.$$

Their arithmetic mean is the equally spaced $2n$-point periodic trapezoidal rule and is exact on $\Lambda_{-(2n-1),2n-1}$.

The stronger exactness follows by summing Fourier modes on the union grid. A mode outside that space also gives a simple limit to the error-estimation argument. At a fixed order n, choose $\tau = 1$ and $f(z) = \mathrm{Re}(z^{2n})$. Both rules return one, whereas $I_\nu(f) = 0$. Their difference vanishes although each error is -1. Thus smoothness alone does not extend the bracketing property beyond the space on which the opposite-error identity holds.

2.7.3 Chebyshev anti-Gauss/anti-Szegő correspondence

For the first-kind Chebyshev weight, retain the scaled functional

$$I(g) = \frac{1}{2\pi} \int_{-1}^{1} \frac{g(x)}{\sqrt{1-x^2}}\, \mathrm{d}x, \tag{2.76}$$

whose total mass is $I(1) = 1/2$. Its Gaussian rule is

$$\mathcal{G}_n(g) = \frac{1}{2n} \sum_{j=1}^{n} g\left(\cos \frac{(2j-1)\pi}{2n}\right).$$

These are the classical Gauss–Chebyshev data [85, 122], multiplied by the chosen mass factor. The corresponding anti-Gauss rule is

$$\widetilde{\mathcal{G}}_{n+1}(g) = \sum_{j=0}^{n} \widetilde{\omega}_{j,n}\, g(\widetilde{x}_{j,n}), \tag{2.77}$$

with

$$\widetilde{x}_{j,n} = \cos(j\pi/n), \qquad \widetilde{\omega}_{j,n} = \begin{cases} 1/(4n), & j = 0, n, \\ 1/(2n), & 1 \le j \le n-1. \end{cases} \tag{2.78}$$

The endpoint weights are half the interior weights. Laurie's identity gives

$$\widetilde{\mathcal{G}}_{n+1}(p) = 2I(p) - \mathcal{G}_n(p), \qquad p \in \mathbb{P}_{2n+1}. \tag{2.79}$$

For the unscaled anti-Gauss data see [48, Theorem 2].

The change from circle to interval identifies conjugate node pairs, while the endpoints remain unpaired. The following correspondence specializes the interval–circle relations of [30, Theorems 3.2–3.3 and 4.2–4.3] to these grids, retaining the mass normalization and the endpoint weights.

Theorem 2.77 (Chebyshev correspondence). Let $n \geq 1$, and define

$$\mathcal{I}(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta.$$

For $g \in C([-1,1])$, its symmetric lift is

$$(\mathcal{U}g)(z) = \frac{1}{2} g\left(\frac{z + z^{-1}}{2}\right), \qquad z \in \Gamma. \tag{2.80}$$

Let $S_{2n} = S_{2n,1}$ and $\widetilde{S}_{2n} = S_{2n,-1}$ for normalized Lebesgue measure. Their nodes are respectively

$$z_k = \exp\left(\frac{(2k-1)\pi\mathrm{i}}{2n}\right), \qquad \widetilde{z}_k = \exp\left(\frac{(k-1)\pi\mathrm{i}}{n}\right), \qquad k = 1, \dots, 2n,$$

and every weight is $1/(2n)$. Then

$$\begin{aligned} \mathcal{I}(\mathcal{U}g) &= I(g), \\ S_{2n}(\mathcal{U}g) &= \mathcal{G}_n(g), \\ \widetilde{S}_{2n}(\mathcal{U}g) &= \widetilde{\mathcal{G}}_{n+1}(g). \end{aligned} \tag{2.81}$$

In particular, for $p \in \mathbb{P}_{2n+1}$,

$$\begin{aligned} \widetilde{S}_{2n}(\mathcal{U}p) &= 2\mathcal{I}(\mathcal{U}p) - S_{2n}(\mathcal{U}p), \\ \mathcal{I}(\mathcal{U}p) - \widetilde{S}_{2n}(\mathcal{U}p) &= -\big(\mathcal{I}(\mathcal{U}p) - S_{2n}(\mathcal{U}p)\big). \end{aligned} \tag{2.82}$$

The anti-Szegő rule has $2n$ distinct nodes, not $2(n+1)$.

Proof. Symmetry and $x = \cos\theta$ give

$$\mathcal{I}(\mathcal{U}g) = \frac{1}{4\pi} \int_{-\pi}^{\pi} g(\cos\theta)\, \mathrm{d}\theta = \frac{1}{2\pi} \int_{0}^{\pi} g(\cos\theta)\, \mathrm{d}\theta = I(g).$$

Each conjugate pair of Szegő nodes has the same image under $(z + z^{-1})/2$. Its two weights $1/(2n)$, together with the factor $1/2$ in the lift, give the Gaussian weight $1/(2n)$. For the anti-Szegő grid, 1 and -1 occur only once, while the other nodes occur in conjugate pairs. Hence

$$\widetilde{S}_{2n}(\mathcal{U}g) = \frac{g(1) + g(-1)}{4n} + \frac{1}{2n} \sum_{j=1}^{n-1} g\left(\cos\frac{j\pi}{n}\right) = \widetilde{\mathcal{G}}_{n+1}(g).$$

This proves the correspondence and node count. Substitution into (2.79) proves the error identities. □

In this special case the arithmetic average has *exact degree* $4n - 1$, sharper than the general guarantee $2n + 1$. Indeed,

$$\tfrac{1}{2}(\mathcal{G}_n + \widetilde{\mathcal{G}}_{n+1})(g) = \frac{g(1) + g(-1)}{8n} + \frac{1}{4n} \sum_{j=1}^{2n-1} g\left(\cos\frac{j\pi}{2n}\right).$$

It is the scaled $(2n+1)$-node Chebyshev–Lobatto rule, obtained by folding the equally spaced $4n$-point circle average. That average is exact on $\Lambda_{-(4n-1),4n-1}$, and $\mathcal{U}p \in \Lambda_{-m,m}$ for $p \in \mathbb{P}_m$. Exactness therefore extends to $\mathbb{P}_{4n-1}$, and so do the two opposite-error identities above. It fails at T_{4n}, where $T_j(\cos\theta) = \cos(j\theta)$: the integral is zero but every nodal value is one, giving quadrature value $1/2$.

These correspondences carry nodes, weights, and exact error identities between the interval and the circle. For general integrands, the error of the paired estimator is the error of the averaged rule, as described at the beginning of this section. Best approximation therefore gives an absolute bound beyond the exactness space.

2.8 Connections with the subsequent chapters

The results of this chapter serve two complementary purposes. Spectral distributions and the GLT calculus describe the asymptotic behavior of matrix sequences; exactness spaces and operator norms control errors in quadrature and interpolation. The subsequent chapters combine these foundations with the structure of their particular problems.

Chapter 3 studies approximation and pseudoinverse preconditioning for rectangular block Toeplitz matrices. Zero-distributed perturbations and the Fasino–Tilli theorem control the blockwise approximation errors. Extradimensional alignment then places the matrices and their approximants in a common space, where GLT symbol matching gives the preconditioning results. The chapter develops the rational block geometry and the stronger estimates needed for strong clustering under continuous-symbol assumptions.

Chapter 4 extends Toeplitz spectral analysis to quaternion matrices. Canonical right eigenvalues and singular-value duplication connect its quaternion conventions to the complex distributions and a.c.s. calculus of Section 2.2. Exact single-axis embeddings identify the complex symbols to which Subsection 2.3.1 applies. For circulant approximation, the chapter combines quaternion frequency pairing with the complex Fourier identities of Subsection 2.3.3.

Chapter 5 applies quadrature to the singular Hilbert kernel. Laurent exactness and the Szegő/anti-Szegő pairing provide the base rules and error identities; the Chebyshev lift relates them to interval quadrature. The terminal parameter is chosen as a function of the evaluation point. Separation of the resulting node sets and regularization of the kernel then allow pointwise error bounds at a moving singularity.

Chapter 6 studies interpolation at anti-Gauss nodes. The quadrature constructions determine the nodes, while the Hermitian perturbation principle helps describe their limiting distribution. Exactness–stability and Lebesgue bounds support the interpolation analysis. The chapter derives modified-kernel, barycentric, and modal representations and establishes node and Christoffel estimates for Jacobi weights, together with sufficient conditions for weighted stability.

In each case, the common results identify which approximation errors can be controlled. The research chapters supply the further geometric, spectral, or weighted estimates needed for their specific conclusions.

Part I

Structured Matrix Sequences

Chapter 3

Blocking Structures, Approximation, and Preconditioning

3.1 From block distributions to preconditioning

This chapter studies structured approximation and preconditioning for block Toeplitz matrix-sequences whose outer blocks have unequal dimensions. Their limiting singular-value distributions, and their eigenvalue distributions under the corresponding Hermitian assumptions, are known from earlier work [11, 73, 4]. Following [9], we use these results to construct a finite structured approximant S_n and analyze its approximation and preconditioning properties. The definitions of distribution, clustering, and zero-distributed perturbations are those of Chapter 2.

Such block structures arise in coupled discretizations of ordinary and partial differential equations and in inverse problems from signal and image reconstruction; see [22, 16, 17, 54, 58, 61] and [33, 34, 43], respectively. The limiting block-size ratios determine the assembled symbol, but computational use also requires an approximant that can be applied or inverted efficiently.

This unilevel construction uses rational limiting block-size ratios. We first establish approximation properties of S_n and then analyze clustering of the corresponding preconditioned sequences. The numerical experiments examine approximation errors, preconditioned spectra, and Krylov iteration counts, including an exploratory two-level example.

Section 3.2 fixes the block geometry. Section 3.3 recalls the distribution results and constructs the blocking surrogates. Section 3.4 develops approximation, extradimensional alignment, and preconditioning. Sections 3.5 and 3.6 present the experiments and conclusions.

3.2 Problem setting and asymptotic block geometry

3.2.1 Unequal block dimensions and limiting ratios

Let ν and $n_1, n_2, \ldots, n_\nu$ be positive integers, and let $n = \sum_{i=1}^{\nu} n_i$. We consider the block matrix $A_n = [A_{ij}]_{i,j=1}^{\nu}$ given by

$$A_n = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1\nu} \\ A_{21} & A_{22} & \cdots & A_{2\nu} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\nu 1} & A_{\nu 2} & \cdots & A_{\nu\nu} \end{bmatrix}, \tag{3.1}$$

in which, for $i = 1, \dots, \nu$, each diagonal block A_{ii} is a matrix of size $sn_i \times tn_i$, where s and t are fixed positive integers, and the off-diagonal blocks A_{ij}, $i \neq j$, are general $sn_i \times tn_j$ rectangular matrices. When $s = t$, the diagonal blocks A_{ii}, $i = 1, \dots, \nu$, are square matrices. Overall, A_n has size $sn \times tn$ and is composed of ν^2 blocks.

We study the matrix-sequence $\{A_n\}_n$ as $n \to \infty$. The block sizes may differ at each finite dimension, but their proportions are required to converge. Following [11, 73], we impose the following assumptions throughout the chapter.

Assumption 3.1. The parameters ν, s, t are fixed independently of n, and $\nu \geq 2$, $s, t \geq 1$. The case $\nu = 1$ reduces to the standard single-block setting.

Assumption 3.2. For each $i = 1, \dots, \nu$, the block size $n_i = n_i(n)$ is a positive integer and, as $n \to \infty$,

$$\frac{n_i(n)}{n} \longrightarrow c_i, \qquad c_i \in (0,1), \qquad \sum_{i=1}^{\nu} c_i = 1.$$

Because $c_i > 0$, this condition implies $n_i(n) \to \infty$ for every $i = 1, \dots, \nu$.

For the finite blocking construction used in Sections 3.3 and 3.4, we additionally assume that $c_i \in \mathbb{Q}^+$. Write

$$c_i = \frac{\alpha_i}{\beta_i} = \frac{m_i}{m}, \qquad m = \operatorname{lcm}(\beta_1, \dots, \beta_\nu), \qquad \sum_{i=1}^{\nu} m_i = m.$$

Here α_i and β_i are relatively prime positive integers, and $m_i = m\alpha_i/\beta_i \in \mathbb{N}$. The common denominator permits a fixed finite collection of replicated blocks of a common base size. This is the geometry used to construct $\widehat{A}_n$ and S_n below.

Rationality is required for this finite common-denominator construction, not for every distributional treatment of unequal block structures. The generalized a.c.s. framework of [5], applied in [4], extends the Weyl distribution analysis to irrational limiting ratios. That extension does not, without an additional argument, provide the finite preconditioner or the clustering results considered here.

Coupled PDEs, including the linear-elasticity problem [28] and the EMI model, give rise to block structures. In the equal-volume EMI example of [18], the limiting ratio is 1/2; rationality is a feature of that geometric setup, not a general consequence of coupling PDEs. Irrational ratios are also relevant, for example for a circular domain embedded in the square $[0,1]^2$ or for mixed discretizations [22].

3.2.2 Inherited spectral tools

Chapter 2 supplies the rank–norm characterization of zero distribution (Theorem 2.11) and the matrix norm bounds used in the approximation proofs. The bounded-rank refinement for strong clustering is established below in Theorem 3.7.

For preconditioning, Lemma 3.8 constructs a common ambient realization of the block model. The argument then uses pseudoinverse symbol matching (Theorem 2.45), the Hermitian GLT calculus, and the exact identities in Lemma 2.21. Corollary 2.44 returns the resulting distributions to the original dimensions.

3.3 Block Toeplitz structures and limiting distributions

We now specialize the block model to unilevel Toeplitz matrices. The recalled distribution results lead to two auxiliary sequences, $\{\widetilde{A}_n\}_n$ and $\{\widehat{A}_n\}_n$, whose approximation properties underlie the construction in Section 3.4.

3.3.1 Toeplitz block model

We specialize the blocks of (3.1) as follows:

- the diagonal blocks A_{ii}, for $i = 1, \dots, \nu$, are $sn_i \times tn_i$ Toeplitz matrices $T_{n_i}(f_{i,i})$, following Definition 2.22;
- the off-diagonal blocks A_{ij}, for $i \neq j$, are $sn_i \times tn_j$ rectangular matrices $T_{n_i,n_j}(f_{i,j})$, according to Definition 2.22;

where $f_{i,j}$, $i, j = 1, \dots, \nu$, are given functions belonging to $L^1([-\pi,\pi], \mathbb{C}^{s\times t})$. More explicitly, the matrix A_n has the form

$$A_n = \begin{bmatrix} T_{n_1}(f_{1,1}) & T_{n_1,n_2}(f_{1,2}) & \cdots & T_{n_1,n_\nu}(f_{1,\nu}) \\ T_{n_2,n_1}(f_{2,1}) & T_{n_2}(f_{2,2}) & \cdots & T_{n_2,n_\nu}(f_{2,\nu}) \\ \vdots & \vdots & \ddots & \vdots \\ T_{n_\nu,n_1}(f_{\nu,1}) & T_{n_\nu,n_2}(f_{\nu,2}) & \cdots & T_{n_\nu}(f_{\nu,\nu}) \end{bmatrix}. \tag{3.2}$$

We next recall the distribution results for equal block sizes, asymptotically equal block sizes, and general rational limiting ratios.

3.3.2 From equal blocks to rational limiting ratios

The following results are recalled from [73]; their symbols and auxiliary constructions are used in the preconditioning analysis. Equal block sizes give the basic fixed block symbol, asymptotically equal sizes allow negligible dimension discrepancies, and rational limiting ratios produce the replicated block symbol used below.

Theorem 3.3 (Equal block sizes). Assume that a matrix A_n structured as in (3.2) is given, with $n_1 = n_2 = \cdots = n_\nu = \frac{n}{\nu}$, and set $n(\nu) = \frac{n}{\nu}$. For any positive integer μ, define the

following permutation matrix

$$\Pi_\mu = \Pi_{n,\nu,\mu} = \begin{bmatrix} I_\nu \otimes \mathbf{e}_1^\mathsf{T} \\ I_\nu \otimes \mathbf{e}_2^\mathsf{T} \\ \vdots \\ I_\nu \otimes \mathbf{e}_{n(\nu)}^\mathsf{T} \end{bmatrix} \otimes I_\mu,$$

such that

$$T_{n(\nu)}(F) = \Pi_s A_n \Pi_t^\mathsf{T},$$

where $F = (f_{i,j})_{i,j=1}^{\nu}$. Under these assumptions,

$$\{A_n\}_n \sim_\sigma \ F.$$

If, in addition, F is Hermitian-valued (which implies $s = t$), then

$$\{A_n\}_n \sim_\lambda \ F.$$

Theorem 3.4 (Asymptotically equal block sizes). Assume that A_n is a matrix structured as in (3.2), with $n_j = n(\nu) + o(n)$, $j = 1, \dots, \nu$, and set $n(\nu) = \frac{n}{\nu}$, $F = (f_{i,j})_{i,j=1}^{\nu}$. Under these assumptions,

$$\{A_n\}_n \sim_\sigma \ F.$$

If, in addition, F is Hermitian-valued (which implies $s = t$), then

$$\{A_n\}_n \sim_\lambda \ F.$$

For positive integers a, b, let $I_{a,b} \in \mathbb{R}^{a\times b}$ be the rectangular identity, $(I_{a,b})_{uv} = \delta_{uv}$. With the integers m_j, m fixed above, define the canonical symbol by

$$F(\theta) := \left[I_{m_j,m_k} \otimes f_{j,k}(\theta) \right]_{j,k=1}^{\nu} \in \mathbb{C}^{sm\times tm}. \tag{3.3}$$

Thus its (j,k) block contains $\min\{m_j, m_k\}$ diagonal copies of $f_{j,k}$ and zeros in the remaining replica positions.

Theorem 3.5 (Rational asymptotic block ratios). Let A_n be given by (3.2), and suppose that Assumptions 3.1–3.2 hold with $c_j = m_j/m \in \mathbb{Q}^+$. For the symbol (3.3),

$$\{A_n\}_n \sim_\sigma F.$$

If every A_n is Hermitian, then also $\{A_n\}_n \sim_\lambda F$.

3.3.3 Distribution-preserving blocking surrogates

Following [73], we construct two auxiliary sequences in the original dimensions. The truncation $\widetilde{A}_n$ retains the leading equal-order Toeplitz core of each block; the surrogate $\widehat{A}_n$ splits that core into replicated Toeplitz matrices and one residual Toeplitz matrix. Here the order counts block rows and columns, so an equal-order core need not be square when $s \neq t$.

We use the same zero-padding rule throughout. Put $h_{n,j,k} := \min\{n_j, n_k\}$. For a matrix $C \in \mathbb{C}^{sh_{n,j,k}\times th_{n,j,k}}$, let $\mathcal{P}_{n,j,k}(C)$ be the $sn_j \times tn_k$ matrix obtained by placing C in the leading rows and columns and filling any remaining entries with zeros. Then

$$\widetilde{A}_n := \left[\mathcal{P}_{n,j,k}(T_{h_{n,j,k}}(f_{j,k})) \right]_{j,k=1}^{\nu}. \tag{3.4}$$

Set $\ell_n := \lfloor n/m \rfloor$. The rational-ratio assumptions imply $n_j = m_j \ell_n + o(\ell_n)$. For every block pair, define

$$p_{j,k} := \min\{m_j, m_k\}, \qquad r_{n,j,k} := h_{n,j,k} - (p_{j,k} - 1)\ell_n.$$

Then $r_{n,j,k} = \ell_n + o(\ell_n) > 0$ for all sufficiently large n; we henceforth discard the finitely many preceding indices. Define

$$C_{n,j,k} := \left[I_{p_{j,k}-1} \otimes T_{\ell_n}(f_{j,k})\right] \oplus T_{r_{n,j,k}}(f_{j,k}), \tag{3.5}$$

where a direct-sum term with zero replicas is absent, and set

$$\widehat{A}_n := \left[\mathcal{P}_{n,j,k}(C_{n,j,k})\right]_{j,k=1}^{\nu}. \tag{3.6}$$

The residual Toeplitz matrix has its exact order $r_{n,j,k}$, whether this order is smaller or larger than ℓ_n.

The truncation and replication arguments in [73] give

$$\{A_n - \widetilde{A}_n\}_n \sim_\sigma 0, \tag{3.7}$$

$$\{\widetilde{A}_n - \widehat{A}_n\}_n \sim_\sigma 0. \tag{3.8}$$

Consequently,

$$\{A_n - \widehat{A}_n\}_n \sim_\sigma 0. \tag{3.9}$$

This is the comparison needed for the approximation theorem. Under continuity, the strong-clustering proof below refines both construction errors through a single polynomial approximation.

3.4 Approximation and preconditioning

We replace the Toeplitz matrices in $\widehat{A}_n$ by structured approximants, following [9]. The approximation results are proved in the original dimensions. Only afterwards do we introduce the common ambient space needed to analyze the preconditioned products.

3.4.1 Construction of the structured approximant

For every j,k, choose a principal approximant $S_{\ell_n}(f_{j,k}) \in \mathbb{C}^{s\ell_n \times t\ell_n}$ and, independently, a residual approximant $R_{\ell_n,j,k} \in \mathbb{C}^{sr_{n,j,k} \times tr_{n,j,k}}$. Assume

$$\begin{aligned} \{S_{\ell_n}(f_{j,k}) - T_{\ell_n}(f_{j,k})\}_n &\sim_\sigma 0, \\ \{R_{\ell_n,j,k} - T_{r_{n,j,k}}(f_{j,k})\}_n &\sim_\sigma 0. \end{aligned} \tag{3.10}$$

Define the approximant core and the assembled matrix by

$$D_{n,j,k} := \left[I_{p_{j,k}-1} \otimes S_{\ell_n}(f_{j,k})\right] \oplus R_{\ell_n,j,k}, \tag{3.11}$$

$$S_n := \left[\mathcal{P}_{n,j,k}(D_{n,j,k})\right]_{j,k=1}^{\nu}. \tag{3.12}$$

Thus S_n and $\widehat{A}_n$ have the same constituent block sizes and zero padding.

The independent residual choice is an adaptation of [9, Section 4, equations (9)–(10)], where the residual is described by adding or deleting trailing rows and columns of the principal approximant. Here its approximation property is explicitly required in (3.10). If a chosen structured algebra supplies approximants at every order, one may take $R_{\ell_n,j,k} = S_{r_{n,j,k}}(f_{j,k})$, provided the same approximation property holds along $r_{n,j,k} \to \infty$.

The hypotheses above control approximation, not computational cost. Efficient application or solution additionally depends on the structured algebra used for the principal and residual approximants.

3.4.2 Approximation error and clustering at zero

The following two results are adapted from [9, Theorems 11–12].

Theorem 3.6 (Zero-distributed block approximation). Assume the hypotheses of Theorem 3.5, and let S_n be constructed by (3.12) with the principal and residual errors satisfying (3.10). Then

$$\{S_n - A_n\}_n \sim_\sigma 0.$$

If S_n and A_n are Hermitian for every n, then also $\{S_n - A_n\}_n \sim_\lambda 0$.

Proof. By (3.9), it suffices to treat $S_n - \widehat{A}_n$. Its (j,k) block is the zero padding of $D_{n,j,k} - C_{n,j,k}$, a fixed direct sum of the errors in (3.10). Since $r_{n,j,k} \sim \ell_n \sim n/m$, the rank–norm characterization in Theorem 2.11 gives, for each block,

$$(S_n - \widehat{A}_n)^{(j,k)} = R_n^{(j,k)} + N_n^{(j,k)}, \qquad \operatorname{rank} R_n^{(j,k)} = o(n), \quad \|N_n^{(j,k)}\| \longrightarrow 0.$$

Assemble these decompositions into $S_n - \widehat{A}_n = R_n + N_n$. Since ν is fixed,

$$\operatorname{rank} R_n \le \sum_{j,k=1}^{\nu} \operatorname{rank} R_n^{(j,k)} = o(n), \qquad \|N_n\| \le \sum_{j,k=1}^{\nu} \|N_n^{(j,k)}\| \longrightarrow 0.$$

Therefore

$$\{S_n - \widehat{A}_n\}_n \sim_\sigma 0. \tag{3.13}$$

Adding (3.9) proves the singular-value claim. In the Hermitian case, the singular values of $S_n - A_n$ are the absolute values of its eigenvalues, which gives the eigenvalue conclusion. □

Zero distribution allows a sublinear number of outliers. The next result bounds their number independently of n, outside each fixed neighborhood of zero.

Theorem 3.7 (Strong zero clustering under continuity). Under the hypotheses of Theorem 3.6, suppose that every $f_{j,k}$ is continuous on $\mathbb{T}$ and that both error sequences in (3.10) are strongly clustered at zero in the singular-value sense. Then $\{S_n - A_n\}_n$ is strongly clustered at zero in the singular-value sense.

Proof. Write $\mathbf{f} = (f_{j,k})_{j,k=1}^{\nu}$ and make the linear dependence of the construction on the symbols explicit by setting

$$K_n(\mathbf{f}) := A_n(\mathbf{f}) - \widehat{A}_n(\mathbf{f}).$$

For $\varepsilon > 0$, choose trigonometric polynomials $\mathbf{g} = (g_{j,k})_{j,k=1}^{\nu}$ such that $\max_{j,k} \|f_{j,k} - g_{j,k}\|_{L^\infty(\mathbb{T})} < \varepsilon$. Linearity gives

$$K_n(\mathbf{f}) = K_n(\mathbf{g}) + K_n(\mathbf{f} - \mathbf{g}).$$

For the polynomial term, use the two construction stages:

$$K_n(\mathbf{g}) = [A_n(\mathbf{g}) - \widetilde{A}_n(\mathbf{g})] + [\widetilde{A}_n(\mathbf{g}) - \widehat{A}_n(\mathbf{g})].$$

The first difference contains only couplings across the cutoff at $h_{n,j,k}$ in each rectangular block. The second contains only couplings across the $p_{j,k} - 1$ interfaces between its retained blocks. Toeplitz matrices generated by fixed trigonometric polynomials are banded, so these couplings are confined to a fixed number of block rows or columns near the interfaces. Consequently, for the chosen polynomials,

$$\sup_n \operatorname{rank} K_n(\mathbf{g}) < \infty.$$

For the remainder, the norm bounds for Toeplitz matrices apply to both the rectangular blocks of A_n and the direct-sum cores of $\widehat{A}_n$. Zero padding does not increase the norm, and there are ν^2 block pairs; hence

$$\sup_n \|K_n(\mathbf{f} - \mathbf{g})\| \le 2\nu^2 \varepsilon.$$

The bounded-rank plus small-norm criterion proves strong zero clustering of $\{A_n - \widehat{A}_n\}_n$.

The same criterion and the block assembly used in the preceding proof show that the assumed strong clustering of the principal and residual errors passes to $\{S_n - \widehat{A}_n\}_n$. Adding the two differences proves the result. □

3.4.3 Extradimensional alignment with the canonical Toeplitz model

The approximation results were proved in the original, or *physical*, dimensions $sn \times tn$. To analyze preconditioning, we now construct one common embedding into an *ambient* space containing complete replicas of the canonical Toeplitz model. This makes explicit the extradimensional comparison used in [9, proofs of Theorems 11 and 13].

Lemma 3.8 (Extradimensional realization of the block model). Under the hypotheses of Theorem 3.5, let $\widehat{A}_n$ be defined by (3.6). There exist integers $L_n \sim n/m$ and coordinate isometries

$$E_n^{(s)} \in \mathbb{C}^{smL_n \times sn}, \qquad E_n^{(t)} \in \mathbb{C}^{tmL_n \times tn},$$

such that

$$\operatorname{rank}\big(E_n^{(s)} \widehat{A}_n (E_n^{(t)})^* - T_{L_n}(F)\big) = o(n). \tag{3.14}$$

Moreover, for every sequence $\{X_n\}_n$ with $X_n \in \mathbb{C}^{sn \times tn}$,

$$\{X_n - \widehat{A}_n\}_n \sim_\sigma 0 \quad \Longrightarrow \quad \{E_n^{(s)} X_n (E_n^{(t)})^*\}_n \sim_{\mathrm{GLT}} F. \tag{3.15}$$

When $s = t$, the two isometries can be chosen equal.

Proof. **Embedding and ordering.** Set

$$q_{n,j} := n_j - (m_j - 1)\ell_n, \qquad L_n := \max\{\ell_n, q_{n,1}, \dots, q_{n,\nu}\}.$$

Since $q_{n,j} = \ell_n + o(\ell_n)$, we have $L_n \sim \ell_n \sim n/m$. Partition the scalar coordinates of block j into $m_j - 1$ intervals of length ℓ_n and a final interval of length $q_{n,j}$. Embed each interval into the leading coordinates of an interval of length L_n by the coordinate isometry

$$J_n := \bigoplus_{j=1}^{\nu} \big[(I_{m_j - 1} \otimes I_{L_n, \ell_n}) \oplus I_{L_n, q_{n,j}} \big].$$

As before, a term with zero replicas is absent. Let $\Pi_{n,s}$ be the perfect shuffle taking a block/replica-first row index (j, a, α, μ) to the Toeplitz ordering (α, j, a, μ), where $1 \le a \le m_j$, $1 \le \alpha \le L_n$, and $1 \le \mu \le s$. Define $\Pi_{n,t}$ analogously for columns, and set

$$E_n^{(s)} := \Pi_{n,s}(J_n \otimes I_s), \qquad E_n^{(t)} := \Pi_{n,t}(J_n \otimes I_t).$$

Their columns are distinct canonical basis vectors, so $(E_n^{(s)})^* E_n^{(s)} = I_{sn}$ and $(E_n^{(t)})^* E_n^{(t)} = I_{tn}$. Since $mL_n - n = o(n) = o(L_n)$, the isometries satisfy (2.20) and are asymptotically full.

A common boundary estimate. Put $\delta_n := \max_j |q_{n,j} - \ell_n| = o(\ell_n)$. For all sufficiently large n, $2\delta_n < \ell_n$ and $|r_{n,j,k} - \ell_n| \le \delta_n$ for every pair: if $m_j < m_k$, then $n_j < n_k$ and $r_{n,j,k} = q_{n,j}$; equal multiplicities give $r_{n,j,k} = \min\{q_{n,j}, q_{n,k}\}$, and the remaining case is symmetric. In each ambient replica retain only the internal coordinates

$$H_n := \{\delta_n + 1, \dots, \ell_n - \delta_n\}.$$

Every such coordinate is in the image of the embedding. For a fixed pair (j, k), the retained coordinates in replicas $1, \dots, p_{j,k}$ lie in the corresponding diagonal blocks of $C_{n,j,k}$; all retained coordinates in later replicas lie outside its core. In particular, shortening a residual removes

only excluded trailing coordinates, and any extension into a later replica is confined to its first δ_n coordinates.

Before the perfect shuffle, a retained row (j,a,α,μ) and column (k,b,β,η) therefore give the same entry in the embedded surrogate and in $\Pi_{n,s}^{*}T_{L_n}(F)\Pi_{n,t}$. If $a=b\le p_{j,k}$, this entry is $(\widehat{f_{j,k}}_{\alpha-\beta})_{\mu,\eta}$; otherwise it is zero. Thus the difference is supported in the excluded rows or columns. Their total number, unchanged by the shuffles, gives

$$\operatorname{rank}\bigl(E_n^{(s)}\widehat{A}_n(E_n^{(t)})^{*}-T_{L_n}(F)\bigr)\le (s+t)m\bigl(L_n-\ell_n+2\delta_n\bigr)=o(n).$$

Transfer of the symbol. Embedding $X_n-\widehat{A}_n$ preserves its nonzero singular values and adds only $o(n)$ zeros. The extradimensional principle (Proposition 2.19) therefore preserves zero distribution. Combining this fact with (3.14) shows that $\{E_n^{(s)}X_n(E_n^{(t)})^{*}-T_{L_n}(F)\}_n$ is zero-distributed. Since $\{T_{L_n}(F)\}_n\sim_{\mathrm{GLT}}F$, the GLT zero-generator and algebra rules give (3.15). When $s=t$, the shuffles and embeddings on the two sides coincide. □

3.4.4 Clustering of the preconditioned sequence

The alignment lemma places both A_n and S_n in the same ambient GLT setting. We can now apply the pseudoinverse calculus under a maximal-rank hypothesis on F. The following result is adapted from [9, Theorem 13].

Theorem 3.9 (Weak clustering of the preconditioned sequence). Assume the hypotheses of Theorem 3.6, and suppose that the canonical symbol satisfies

$$\operatorname{rank}F(\theta)=m\min\{s,t\}\qquad\text{for almost every }\theta.$$

Then

$$\begin{aligned} s\ge t \quad&\Longrightarrow\quad \{S_n^{\dagger}A_n\}_n\sim_{\sigma}1,\\ s\le t \quad&\Longrightarrow\quad \{A_nS_n^{\dagger}\}_n\sim_{\sigma}1.\end{aligned}$$

If, additionally, $s=t$, every A_n is Hermitian, and every S_n is Hermitian positive definite, then

$$\{S_n^{-1}A_n\}_n\sim_{\lambda}1.$$

Proof. **Common ambient symbol.** Use the isometries of Lemma 3.8 to define

$$\mathcal{A}_n:=E_n^{(s)}A_n(E_n^{(t)})^{*},\qquad \mathcal{S}_n:=E_n^{(s)}S_n(E_n^{(t)})^{*}.$$

Relations (3.9) and (3.13), followed by (3.15), give directly

$$\{\mathcal{A}_n\}_n\sim_{\mathrm{GLT}}F,\qquad \{\mathcal{S}_n\}_n\sim_{\mathrm{GLT}}F.$$

Singular-value conclusions. The maximal-rank hypothesis permits pseudoinverse symbol matching (Theorem 2.45):

$$\begin{aligned}\{\mathcal{S}_n^{\dagger}\mathcal{A}_n\}_n&\sim_{\mathrm{GLT}}F^{\dagger}F=I_{tm},\quad s\ge t,\\ \{\mathcal{A}_n\mathcal{S}_n^{\dagger}\}_n&\sim_{\mathrm{GLT}}FF^{\dagger}=I_{sm},\quad s\le t.\end{aligned}$$

By Lemma 2.21, these products are the exact ambient realizations of the physical products:

$$\begin{aligned}\mathcal{S}_n^{\dagger}\mathcal{A}_n&=E_n^{(t)}(S_n^{\dagger}A_n)(E_n^{(t)})^{*},\\ \mathcal{A}_n\mathcal{S}_n^{\dagger}&=E_n^{(s)}(A_nS_n^{\dagger})(E_n^{(s)})^{*}.\end{aligned}$$

Corollary 2.44 therefore gives the asserted singular-value distributions. Both orientations apply when $s=t$.

Hermitian eigenvalue conclusion. Choose the same isometry E_n on both sides. Then $\mathcal{A}_n$ is Hermitian and $\mathcal{S}_n$ is Hermitian positive semidefinite. The Hermitian GLT calculus and the nonnegative spectra of $\mathcal{S}_n$ show that F is positive semidefinite almost everywhere; maximal rank makes it positive definite almost everywhere.

The ambient matrix $\mathcal{S}_n$ may have padding zeros, even though S_n is positive definite. Accordingly, define

$$\mathcal{R}_n := (\mathcal{S}_n^\dagger)^{1/2} = E_n S_n^{-1/2} E_n^*.$$

Apply continuous functional calculus to $h(t) = \sqrt{\max\{t, 0\}}$ after pseudoinversion, as in Section 2.4.2. This gives $\{\mathcal{R}_n\}_n \sim_{\mathrm{GLT}} F^{-1/2}$, and hence

$$\mathcal{R}_n \mathcal{A}_n \mathcal{R}_n = E_n (S_n^{-1/2} A_n S_n^{-1/2}) E_n^*,$$

$$\{\mathcal{R}_n \mathcal{A}_n \mathcal{R}_n\}_n \sim_{\mathrm{GLT}} F^{-1/2} F F^{-1/2} = I_{sm}.$$

The physical matrices in this realization are Hermitian, so the same transfer corollary yields $\{S_n^{-1/2} A_n S_n^{-1/2}\}_n \sim_\lambda 1$. Similarity with $S_n^{-1} A_n$ proves the final assertion. □

3.5 Numerical experiments

We first consider three groups with polynomial generating functions and banded Toeplitz blocks. For these groups, we examine approximation-error plots, preconditioned singular values, and Krylov iteration counts. We then consider dense fractional-diffusion examples, including an exploratory two-level problem. The figures and tables reproduce the experiments reported in [9, Section 5].

3.5.1 Matrix sequences and example groups

Each sequence $\{A_n\}_n$ has the Toeplitz block structure in (3.2). Group 1 uses scalar-valued generators, Group 2 uses row-vector-valued generators and right preconditioning, and Group 3 uses square matrix-valued generators. For Groups 1 and 2, the common block template is

$$A_n = \begin{bmatrix} T_{n_1}(f_{1,1}) & T_{n_1,n_2}(f_{1,2}) \\ T_{n_2,n_1}(f_{2,1}) & T_{n_2}(f_{2,2}) \end{bmatrix}.$$

Group 1: 2×2 blocks with scalar-valued generators. We set $\nu = 2$ and $s = t = 1$. For the size parameter $\eta \in \mathbb{N}$, the two configurations are

(a) $n_1 = 3\eta$, $n_2 = 2\eta$;

(b) $n_1 = 3\eta$, $n_2 = 2\eta + 20$.

The scalar-valued trigonometric polynomials are

$$\begin{aligned} f_{1,1}(t) &= 2 - 2\cos(t); \quad & f_{2,2}(t) &= 2 - 2\cos(t) - 6\cos(2t); \\ f_{1,2}(t) &= 1 - 2\cos(t); \quad & f_{2,1}(t) &= 1 - 2\cos(t). \end{aligned} \tag{3.16}$$

Group 2: 2×2 blocks with rectangular matrix-valued generators. We use $\nu = 2$, $s = 1$, and $t = 2$. Thus each generator is row-vector-valued and $A_n \in \mathbb{C}^{(n_1+n_2)\times 2(n_1+n_2)}$ is rectangular and underdetermined. The two block-size configurations are

(a) $n_1 = \eta$, $n_2 = 2\eta$;

(b) $n_1 = \eta$, $n_2 = 2\eta + \lceil \sqrt{\eta} \rceil$.

The generators are

$$\begin{aligned} f_{1,1}(t) &= \big(2 - 2\cos(t), \quad 4 + 6\cos(2t)\big); \\ f_{2,2}(t) &= \big(3 + 2\cos(t), \quad 4 + 6\cos(t) - 2\cos(2t)\big); \\ f_{1,2}(t) &= \big(1 + \mathrm{e}^{\mathrm{i}t}, \quad 1 - \mathrm{e}^{-\mathrm{i}t}\big); \quad f_{2,1}(t) = f_{1,2}(t). \end{aligned} \tag{3.17}$$

The published article [9, Section 5.1.2] interchanges s and t in the opening sentence of its Group 2 description. The row-vector generators, the underdetermined formulation, and the right-preconditioned product $A_n S_n^\dagger$ determine the dimensionally consistent convention $s = 1$, $t = 2$ used here.

Group 3: 3×3 blocks with square matrix-valued generators. We set $\nu = 3$ and $s = t = 2$, and let η range over even positive integers so that all block sizes are integral. The block sizes are

$$n_1 = \eta, \quad n_2 = \frac{\eta}{2}, \quad n_3 = 2\eta - 2.$$

The generators combine symbols from discretizations and multigrid transfer operators; the assembled block matrix is used as a test problem built from these components. The discretization symbols are introduced through the one-dimensional, constant-coefficient problem

$$\begin{cases} -u''(x) = \psi(x) & x \in (0,1), \\ u(0) = u(1) = 0, \end{cases} \tag{3.18}$$

where $\psi \in L^2(0,1)$.

Equation (16) of [9] is printed with $u'' = \psi$; the minus-sign convention in (3.18) is retained here.

We select the following four building blocks:

1. **Finite-difference (FD) discretization:** The FD discretization of (3.18) leads to the Toeplitz matrix generated by $f(\theta) = 2 - 2\cos\theta$. However, as discussed in [58, Remark 1.3], the symbol is not unique and, by viewing the matrix in 2×2 blocks, the corresponding 2×2 matrix-valued symbol is given by [99]:

   $$f^{[2]}(\theta) = \hat{f}_0^{[2]} + \hat{f}_{-1}^{[2]}\mathrm{e}^{-\mathrm{i}\theta} + \hat{f}_1^{[2]}\mathrm{e}^{\mathrm{i}\theta},$$

   where

   $$\hat{f}_0^{[2]} = T_2(2 - 2\cos\theta), \quad \hat{f}_{-1}^{[2]} = -\mathbf{e}_2\mathbf{e}_1^{\mathsf{T}}, \quad \hat{f}_1^{[2]} = (\hat{f}_{-1}^{[2]})^{\mathsf{T}} = -\mathbf{e}_1\mathbf{e}_2^{\mathsf{T}}.$$

2. $\mathbb{Q}_2$ **Lagrange finite-element (FEM) approximation:** The scaled stiffness matrix is a Toeplitz matrix generated by $f_{\mathbb{Q}_2}(\theta)$ [81]:

   $$f_{\mathbb{Q}_2}(\theta) = \frac{1}{3}\left(\begin{bmatrix} 16 & -8 \\ -8 & 14 \end{bmatrix} + \begin{bmatrix} 0 & -8 \\ 0 & 1 \end{bmatrix}\mathrm{e}^{\mathrm{i}\theta} + \begin{bmatrix} 0 & 0 \\ -8 & 1 \end{bmatrix}\mathrm{e}^{-\mathrm{i}\theta}\right).$$

3. **B-spline discretization:** We consider B-spline discretizations for different degrees p and regularities k. Specifically, we examine the pairs $(p,k) = (2,0)$ and $(3,1)$. The resulting structures $A_{p,k}$ are low-rank corrections of Toeplitz matrices generated by the following functions [82]:

   $$f^{(2,0)}(\theta) = \frac{1}{3}\left(\begin{bmatrix} 4 & -2 \\ -2 & 8 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ 0 & -2 \end{bmatrix}\mathrm{e}^{\mathrm{i}\theta} + \begin{bmatrix} 0 & 0 \\ -2 & -2 \end{bmatrix}\mathrm{e}^{-\mathrm{i}\theta}\right),$$

   $$f^{(3,1)}(\theta) = \frac{1}{40}\left(\begin{bmatrix} 48 & 0 \\ 0 & 48 \end{bmatrix} + \begin{bmatrix} -15 & -15 \\ -3 & -15 \end{bmatrix}\mathrm{e}^{\mathrm{i}\theta} + \begin{bmatrix} -15 & -3 \\ -15 & -15 \end{bmatrix}\mathrm{e}^{-\mathrm{i}\theta}\right).$$

The published version [9, Section 5.1.3] prints -1 in the lower-right entry of the coefficient multiplying $e^{i\theta}$. The value -15 agrees with the independent B-spline symbol in [82, Section 2.4.2, p. 1654]; with this choice, the positive- and negative-frequency coefficient matrices are transposes of one another. The blocking article does not state which coefficient was used in its computations, so the numerical tables below are retained as reported and are not a verified reproduction for the coefficient displayed here.

4. **Transfer-polynomial distinction:** Classical geometric multigrid transfer operators provide the background [145, 25, 28, 55]. The Group 3 construction below uses $p_{\mathbb{Q}_2}$ as displayed. It must be distinguished from the finite-element prolongation polynomial in [71, Section 4.2, Eq. (17), p. 9]:

$$p_{\mathbb{Q}_2}(\theta) = \begin{bmatrix} \frac{3}{4} & \frac{3}{8} \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \frac{3}{8} \\ 0 & 0 \end{bmatrix} e^{i\theta} + \begin{bmatrix} \frac{3}{4} & -\frac{1}{8} \\ 1 & 0 \end{bmatrix} e^{-i\theta} + \begin{bmatrix} 0 & -\frac{1}{8} \\ 0 & 0 \end{bmatrix} e^{-2i\theta}.$$

In the cited finite-element construction the last nonzero coefficient multiplies $e^{2i\theta}$, not $e^{-2i\theta}$. Writing $C = \left[\begin{smallmatrix} 0 & -1/8 \\ 0 & 0 \end{smallmatrix}\right]$, the two definitions are related by

$$p^{\mathrm{FEM}}_{\mathbb{Q}_2}(\theta) = p_{\mathbb{Q}_2}(\theta) + C\big(e^{2i\theta} - e^{-2i\theta}\big).$$

This is a substantive difference. With $z = e^{i\theta}$, direct calculation gives

$$\det p^{\mathrm{FEM}}_{\mathbb{Q}_2}(\theta) = \frac{(z+1)^3}{8z^2}, \qquad \det p_{\mathbb{Q}_2}(\theta) = \frac{(z+1)(3z^2+1)}{8z^3}.$$

Their determinants therefore have different zero orders at $\theta = \pi$. A global reversal of the Fourier convention would preserve that order and cannot remove this difference. Accordingly, no classical two-grid optimality claim for $p^{\mathrm{FEM}}_{\mathbb{Q}_2}$ is transferred here to the displayed $p_{\mathbb{Q}_2}$ or to the full Group 3 matrix. The latter continues to use the displayed polynomial; the numerical tables and figures are retained as results reported in [9], not as verified reproductions for either variant. Establishing which variant generated those results requires the implementation or a controlled rerun; it cannot be inferred from the displayed formulas alone.

Combining these four building blocks gives the Group 3 matrix-sequence $\{A_n\}_n$:

$$A_n = \begin{bmatrix} T_{n_1}(f_{1,1}) & T_{n_1,n_2}(f_{1,2}) & T_{n_1,n_3}(f_{1,3}) \\ T_{n_2,n_1}(f_{2,1}) & T_{n_2}(f_{2,2}) & T_{n_2,n_3}(f_{2,3}) \\ T_{n_3,n_1}(f_{3,1}) & T_{n_3,n_2}(f_{3,2}) & T_{n_3}(f_{3,3}) \end{bmatrix},$$

where

$$\begin{aligned} &f_{1,1}(\theta) = f_{\mathbb{Q}_2}(\theta), \quad f_{1,2}(\theta) = f_{2,1}(\theta) = f^{(2,0)}(\theta), \quad f_{1,3}(\theta) = f_{3,1}(\theta) = f^{(3,1)}(\theta); \\ &f_{2,2}(\theta) = p^*_{\mathbb{Q}_2}(\theta)p_{\mathbb{Q}_2}(\theta) + p^*_{\mathbb{Q}_2}(\theta+\pi)p_{\mathbb{Q}_2}(\theta+\pi); \\ &f_{2,3}(\theta) = f_{3,2}(\theta) = p_{\mathbb{Q}_2}(\theta); \quad f_{3,3}(\theta) = f^{[2]}(\theta). \end{aligned} \tag{3.19}$$

The limiting ratios are $(c_1, c_2) = (3/5, 2/5)$ for Group 1, $(c_1, c_2) = (1/3, 2/3)$ for Group 2, and $(c_1, c_2, c_3) = (2/7, 1/7, 4/7)$ for Group 3. Thus all three groups remain within the rational-ratio setting used by the finite blocking construction.

3.5.2 Distribution, clustering, and Krylov experiments

The experiments in this subsection are those of [9, Section 5.2]. The tables use $\eta \in \{100, 500\}$, and the clustering figures use $\eta = 100$.

For the banded Toeplitz blocks, the published study considers Strang-type circulants and uses Frobenius-optimal circulants when the Strang choice is singular; see [35]. Its Krylov discussion reports the use of the Frobenius-optimal choice. The experiments address three questions:

1. **Zero clustering**: the singular values of $A_n - S_n$ are plotted to illustrate Theorem 3.6 and, under the additional continuity hypotheses, Theorem 3.7.
2. **Clustering at** 1: the singular values of $S_n^{-1}A_n$ in the square cases and of $A_n S_n^{\dagger}$ in the rectangular case are compared with those of A_n.
3. **Krylov behavior**: the published iteration counts for the selected preconditioned methods are compared with the corresponding unpreconditioned counts when the latter are reported.

Zero clustering analysis

Figures 3.1–3.3 show the reported singular values of $A_n - S_n$ for $\eta = 100$. Most values lie near zero, with a small visible set outside the central cluster. The plots illustrate the approximation behavior; the strong-clustering conclusion follows from Theorem 3.7 under its stated hypotheses.

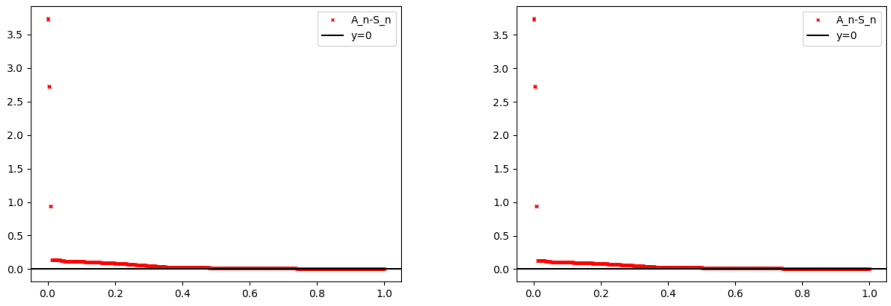


Figure 3.1: Singular values of $A_n - S_n$ for Group 1 with $\eta = 100$. The left and right panels show cases (a) and (b), respectively.

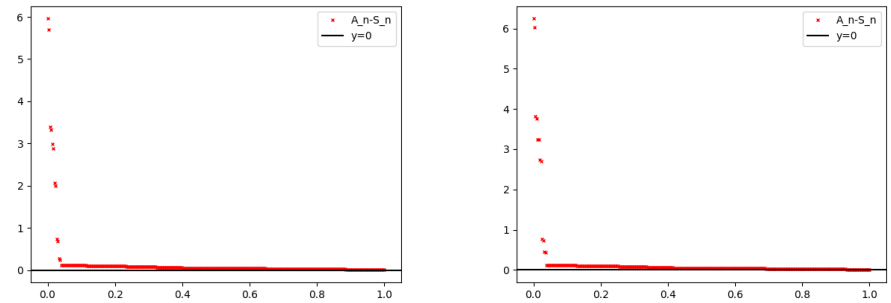


Figure 3.2: Singular values of $A_n - S_n$ for Group 2 with $\eta = 100$. The left and right panels show cases (a) and (b), respectively.

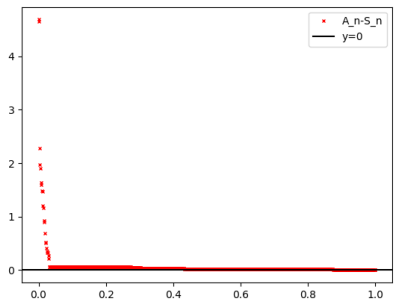


Figure 3.3: Singular values of $A_n - S_n$ for Group 3 with $\eta = 100$.

Clustering at 1 of the preconditioned matrix-sequence

Figures 3.4–3.6 show the reported singular values of $S_n^{-1}A_n$ in the square cases and $A_nS_n^\dagger$ in the rectangular case, for $\eta = 100$. The Group 1 and Group 2 plots also show the singular values of A_n for comparison; the Group 3 plot shows only the preconditioned values.

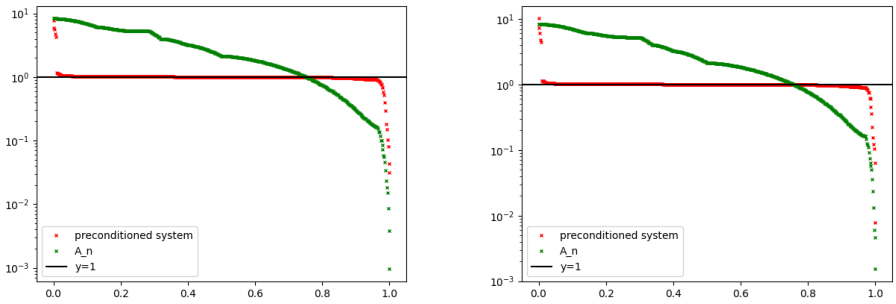


Figure 3.4: Singular values of A_n and $S_n^{-1}A_n$ for Group 1 with $\eta = 100$. Green markers show the unpreconditioned values, red markers show the preconditioned values, and the horizontal line marks 1. The left and right panels show cases (a) and (b), respectively.

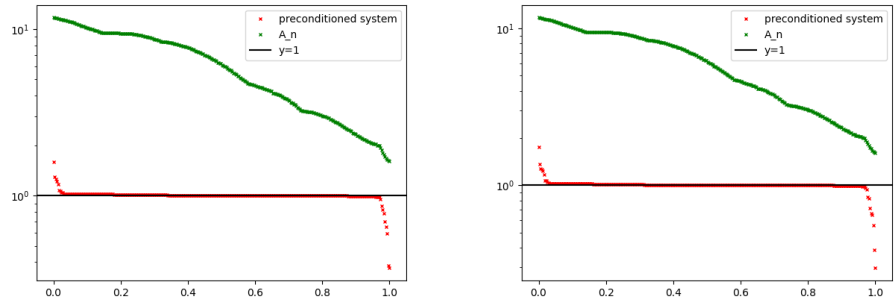


Figure 3.5: Singular values of A_n and $A_nS_n^\dagger$ for Group 2 with $\eta = 100$. Green markers show the unpreconditioned values, red markers show the preconditioned values, and the horizontal line marks 1. The left and right panels show cases (a) and (b), respectively.

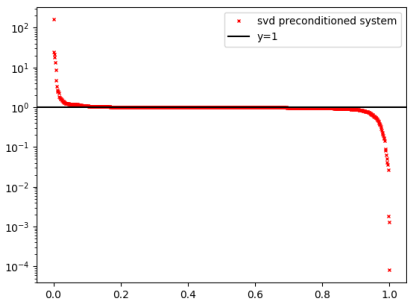


Figure 3.6: Singular values of $S_n^{-1}A_n$ for Group 3 with $\eta = 100$.

Table 3.1 reproduces the outlier counts and ratios from [9, Table 1], retaining the printed numerical precision. The article does not specify the counting neighborhood of 1 alongside that table.

Table 3.1: Outliers of the preconditioned matrix with respect to the cluster at 1, reproduced from [9, Table 1].

Group	Case	η	Outliers below	Ratio below	Outliers above	Ratio above
Group 1	(a)	100	8	0.0160	5	0.010
		500	14	0.0056	5	0.002
	(b)	100	8	0.0150	5	0.009
		500	14	0.0055	5	0.002
Group 2	(a)	100	9	0.030	8	0.027
		500	9	0.006	7	0.005
	(b)	100	9	0.029	8	0.026
		500	9	0.006	7	0.005
Group 3	–	100	27	0.038	18	0.026
		500	53	0.015	23	0.006

The reported ratios decrease between $\eta = 100$ and $\eta = 500$ in every row pair, whereas some absolute counts increase. This is consistent with weak clustering at 1; the two reported sizes do not establish a uniform bound on the outlier counts or strong clustering.

Convergence of the preconditioned Krylov methods

The published experiments use a zero initial guess and, for restarted GMRES, a restart length of 100. Convergence is declared when the residual norm is below 10^{-8} before 1000 iterations. The article does not specify the right-hand side or whether the stopping residual is absolute, relative, or preconditioned.

Group 1 is tested with PGMRES in both configurations, and Group 2 case (b) with CGNE and PCGNE. For the CG and PCG tests, Group 3 is modified: the matrix is symmetrized, the diagonal block at position $(2, 2)$ is scaled by 1.78, and every off-diagonal block is scaled by 0.7. These tests therefore use the modified positive-definite system, not the unmodified Group 3 matrix used in the earlier clustering plots.

Table 3.2 gives the iteration counts reported in [9, Table 2]; terminal residuals and condition numbers are not reported there.

The unpreconditioned GMRES counts for Group 1 are absent because the method does not reach the tolerance in fewer than 1000 iterations. For Group 2, preconditioning reduces

Table 3.2: Numbers of iterations required by PGMRES, CGNE, PCGNE, CG, and PCG to attain the published tolerance 10^{-8}, reproduced from [9, Table 2]. The Group 3 columns refer to the modified system described in the text.

η	PGMRES Group 1 (a)	PGMRES Group 1 (b)	CGNE Group 2 (b)	PCGNE Group 2 (b)	CG Group 3	PCG Group 3
100	23	21	57	18	226	23
500	33	33	56	18	496	28

the reported count from 57 to 18 at $\eta = 100$ and from 56 to 18 at $\eta = 500$. For the modified Group 3 system, it reduces the count from 226 to 23 and from 496 to 28, respectively.

Remark 3.10 (Clustered spectra and PCG). For Hermitian positive definite problems, classical PCG theory shows that if all but a controlled number of eigenvalues of the symmetrically preconditioned operator lie in a fixed interval $[a, b] \subset (0, \infty)$, then the exceptional eigenvalues can be separated from the Chebyshev estimate for the clustered bulk; see [7, 172]. We use this result only as qualitative motivation. It applies to the Hermitian positive-definite PCG setting, not to GMRES, nonnormal preconditioned matrices, or conclusions drawn solely from singular-value clustering.

3.5.3 Numerical results from fractional diffusion equations

The remaining experiments, reported in [9, Section 5.3], include dense Toeplitz blocks generated by fractional operators. The unilevel examples use Frobenius-optimal circulant approximants, whereas the two-level example uses the τ algebra. For these choices, application of the preconditioner inverse costs $O(n^d \log n)$ arithmetic operations, with $d = 1$ in the unilevel case and $d = 2$ in the two-level case. For the role of τ approximants with even generating functions, see [148] and the fractional-diffusion studies [56, 10].

In this subsection, n denotes the base discretization size rather than the sum of the outer block sizes. We write d_n for the total matrix order and P_n for the chosen preconditioner. The unilevel examples remain within the rational-ratio geometry considered above; the status of the two-level example is discussed separately below.

In Examples 3.11 and 3.12, we take $s = t = 1$, $\nu = 2$, $n_2 = 2n$, and $n_1 = n$, so that $d_n = 3n$. The matrix and its preconditioner are

$$A_n = \begin{bmatrix} T_{n_1}(f_{1,1}) & T_{n_1,n_2}(f_{1,2}) \\ T_{n_2,n_1}(f_{2,1}) & T_{n_2}(f_{2,2}) \end{bmatrix}$$

and the preconditioner

$$P_n = \begin{bmatrix} C_{n_1}(T_{n_1}(f_{1,1})) & C_{n_1}(T_{n_1}(f_{1,2})) & \mathbf{0}_{n_1 \times n_1} \\ C_{n_1}(T_{n_1}(f_{2,1})) & C_{n_1}(T_{n_1}(f_{2,2})) & \mathbf{0}_{n_1 \times n_1} \\ \mathbf{0}_{n_1 \times n_1} & \mathbf{0}_{n_1 \times n_1} & C_{n_1}(T_{n_1}(f_{2,2})) \end{bmatrix}$$

where $C_n(T_n(f))$ denotes the Frobenius-optimal circulant approximant to $T_n(f)$ [36, 150, 45]. The circulant blocks permit application and solution at $O(n \log n)$ cost. Their placement follows the replicated structure in (3.6); the discarded couplings contribute a zero-distributed error, as in Theorem 3.5.

The two-level example uses the same outer-block proportions, with tensor-product Toeplitz blocks and total order $d_n = 3n^2$. For each example, we give the generating functions and discuss the reported spectra and iteration counts.

Unilevel fractional-diffusion examples

Table 3.3: Example 3.11: absolute and relative numbers of eigenvalues of $P_n^{-1}A_n$ outside $[0.95, 1.05]$, together with the CG and PCG iteration counts, for $d_n \in \{300, 600, 1200\}$. The relative count is normalized by d_n.

n_1	d_n	N	N/d_n	Iter. CG	Iter. PCG
100	300	22	0.073	> 100	15
200	600	28	0.046	> 100	16
400	1200	37	0.030	> 100	18

The first example uses a Hermitian positive-definite matrix and preconditioner and is tested with CG and PCG. The second uses an indefinite preconditioner and is tested with GMRES and PGMRES. Its eigenvalue observations are not an application of the Hermitian positive-definite preconditioner clause of Theorem 3.9.

Example 3.11. We use the generating functions

$$f_{1,1}(\theta) = -e^{-i\theta}(1-e^{i\theta})^{\alpha} - e^{i\theta}(1-e^{-i\theta})^{\alpha}, \qquad f_{1,2}(\theta) = -1 + e^{-i\theta},$$
$$f_{2,1}(\theta) = f_{1,2}(-\theta), \qquad f_{2,2}(\theta) = 4 - 2\cos(\theta),$$

The function $f_{1,1}$ comes from a first-order finite-difference discretization of a fractional diffusion equation with constant coefficients equal to 1 and fractional exponent $\alpha = 1.5$. Its Fourier expansion is [56]

$$f_{1,1}(\theta) = -2g_1^{(\alpha)} - (g_0^{(\alpha)} + g_2^{(\alpha)})(e^{-i\theta} + e^{i\theta}) - \sum_{k=2}^{\infty} g_{k+1}^{(\alpha)}(e^{-ki\theta} + e^{ik\theta}),$$

with

$$g_k^{(\alpha)} = (-1)^k \binom{\alpha}{k} = \frac{(-1)^k}{k!}\alpha(\alpha-1)\cdots(\alpha-k+1), \quad k = 0, 1, \ldots. \tag{3.20}$$

The function is real and positive apart from its zero of order α at 0. The Toeplitz matrix $T_{n_1}(f_{1,1})$ is dense, real symmetric, positive definite, and ill-conditioned. Its smallest eigenvalue satisfies $\lambda_{\min}(T_{n_1}(f_{1,1})) = O(n_1^{-\alpha})$; see [147]. Both A_n and P_n are real symmetric and positive definite, so A_n and $P_n^{-1}A_n$ have positive eigenvalues.

In Figure 3.7, the unpreconditioned eigenvalues span a broad range, whereas the preconditioned eigenvalues are concentrated near 1.

Table 3.3 shows the relative outlier count decreasing from 0.073 to 0.030 as the matrix order grows from 300 to 1200, while PCG requires 15, 16, and 18 iterations. These values are consistent with weak clustering at 1 and mild growth of the iteration count over the tested sizes. The entries > 100 report lower bounds on the CG counts. The article does not specify the right-hand side or stopping residual for these tests, or the reporting procedure behind those lower bounds.

Example 3.12. We use the generating functions

$$f_{1,1}(\theta) = 2 - 2\cos(\theta), \qquad f_{1,2}(\theta) = -e^{-i\theta}(1-e^{i\theta})^{\alpha} - e^{i\theta}(1-e^{-i\theta})^{\alpha},$$
$$f_{2,1}(\theta) = f_{1,2}(-\theta), \qquad f_{2,2}(\theta) = 4 - 2\cos(\theta),$$

with $\alpha = 1.7$. Unlike the first example, both A_n and the preconditioner have negative eigenvalues. In this example, $P_n^{-1}A_n$ has complex eigenvalues, whose reported concentration near 1 is shown in Figure 3.8.

Table 3.4 reports relative counts decreasing from 0.103 to 0.041 and PGMRES counts of 18, 21, and 26. The article describes the counts using the real interval $[0.95, 1.05]$ but does not specify a corresponding neighborhood in the complex plane. The reported counts therefore cannot be treated as a verified complex-neighborhood clustering test. The entries > 100 give lower bounds on the unpreconditioned GMRES counts.

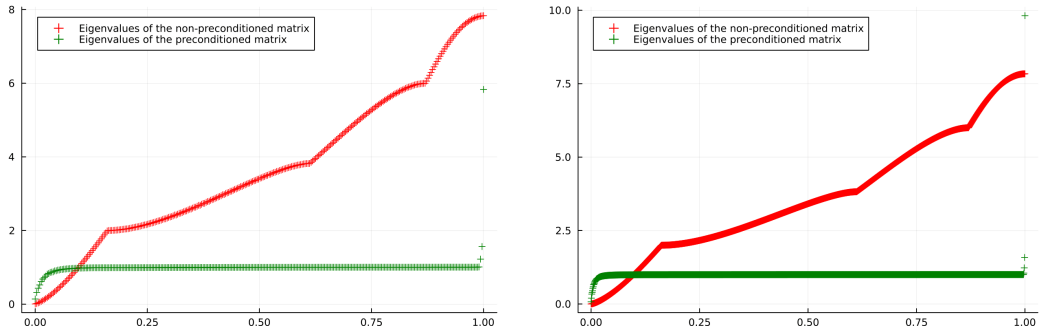


Figure 3.7: Example 3.11: eigenvalues of the unpreconditioned and preconditioned matrices, arranged in nondecreasing order, for $n_1 = 100$ (left) and $n_1 = 400$ (right), with $n_2 = 2n_1$.

Table 3.4: Example 3.12: reported absolute and relative counts for the complex eigenvalues of $P_n^{-1}A_n$, together with GMRES and PGMRES iterations, for $d_n \in \{300, 600, 1200\}$.

n_1	d_n	N	N/d_n	Iter. GMRES	Iter. PGMRES
100	300	31	0.103	> 100	18
200	600	38	0.063	> 100	21
400	1200	50	0.041	> 100	26

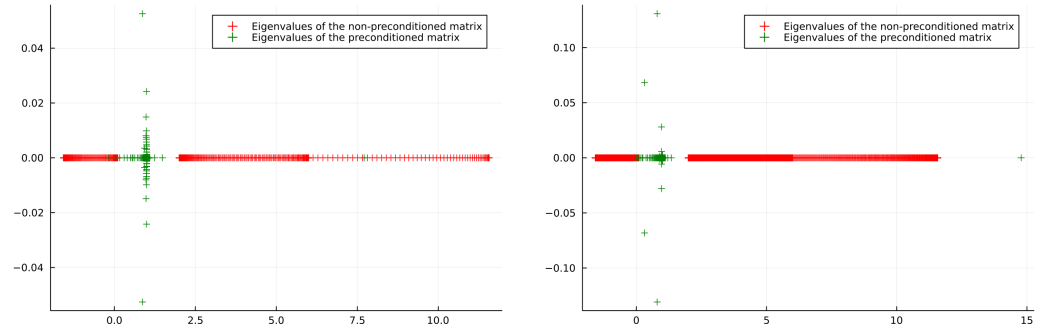


Figure 3.8: Example 3.12: eigenvalues of the unpreconditioned and preconditioned matrices for $n_1 = 100$ (left) and $n_1 = 400$ (right), with $n_2 = 2n_1$. The red and green markers show eigenvalues in the complex plane; the horizontal and vertical coordinates are their real and imaginary parts, respectively.

Exploratory two-level example

The final experiment tests the preconditioning construction on a two-level problem. It is exploratory: Theorems 3.6–3.9 provide the unilevel analysis, not a corresponding multilevel approximation or preconditioning result.

Example 3.13. We consider a two-level block matrix of the form

$$A_n = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix},$$

The diagonal blocks $A_{1,1}$ and $A_{2,2}$ arise from a second-order finite-difference discretization of a two-dimensional fractional differential equation with constant coefficients:

$$\begin{aligned} A_{1,1} &= 4(I_n \otimes T_n(q_\alpha(\theta_1)) + T_n(q_\beta(\theta_2)) \otimes I_n), \\ A_{2,2} &= I_{2n} \otimes T_n(q_\alpha(\theta_1)) + T_{2n}(q_\beta(\theta_2)) \otimes I_n, \end{aligned}$$

with $q_\gamma(\theta) = w_\gamma(\theta) + w_\gamma(-\theta)$ and $w_\gamma(\theta) = -\left(\frac{2-\gamma(1-e^{-i\theta})}{2}\right)\left(1-e^{i\theta}\right)^\gamma$. The function $q_\gamma(\theta)$ has the Fourier expansion

$$\begin{aligned} q_\gamma(\theta) &= -2w_1^{(\gamma)} - (w_0^{(\gamma)} + w_2^{(\gamma)})(e^{-i\theta} + e^{i\theta}) - \sum_{k=2}^{\infty} w_{k+1}^{(\gamma)}(e^{-ki\theta} + e^{ki\theta}), \\ w_0^{(\gamma)} &= \frac{\gamma}{2} g_0^{(\gamma)}, \quad w_k^{(\gamma)} = \frac{\gamma}{2} g_k^{(\gamma)} + \frac{2-\gamma}{2} g_{k-1}^{(\gamma)}, \quad k \geq 1, \end{aligned}$$

where $g_k^{(\gamma)}$ is defined in (3.20). The diagonal blocks are two-level Toeplitz matrices generated by $f_{i,i}(\theta_1, \theta_2) = \gamma_i(q_\alpha(\theta_1) + q_\beta(\theta_2))$ for $i = 1, 2$, with $\gamma_1 = 4$ and $\gamma_2 = 1$. Their discretization orders are $\boldsymbol{m}_1 = (n, n)$ and $\boldsymbol{m}_2 = (n, 2n)$, respectively. Throughout this example, the first coordinate varies fastest; the selector E below uses this ordering, which is a permutation of the lexicographic ordering used in the common toolbox.

Both blocks use the fixed parameters $\alpha = 1.5$ and $\beta = 1.8$. For fixed $1 < \gamma < 2$, the real, even symbol q_γ is positive away from zero and satisfies $q_\gamma(\theta) \asymp |\theta|^\gamma$; see [128] for the fractional-diffusion symbol. Garoni's two-sided estimate and Toeplitz eigenvalue comparison [75, Theorem 2.1 and Remark 4.1 in the author manuscript] therefore give $\lambda_{\min}(T_n(q_\gamma)) \asymp n^{-\gamma}$. Applying this estimate to the displayed Kronecker sums yields

$$\lambda_{\min}(A_{i,i}) \asymp n^{-\alpha} + n^{-\beta} \asymp n^{-\min\{\alpha,\beta\}}, \qquad i = 1, 2.$$

The comparison constants are independent of n but may depend on the fixed orders. The two-variable symbols have anisotropic behavior $f_{i,i}(\theta_1, \theta_2) \asymp |\theta_1|^\alpha + |\theta_2|^\beta$, rather than a two-sided radial zero of order $\min\{\alpha, \beta\}$ when $\alpha \neq \beta$. The diagonal blocks are positive definite and ill-conditioned. These estimates concern only $A_{1,1}$ and $A_{2,2}$, not the smallest eigenvalue or iteration count of the full coupled system.

To specify the rectangular coupling, write

$$\ell(\theta) = 2 - 2\cos\theta, \qquad L_{n,m} = I_m \otimes T_n(\ell) + T_m(\ell) \otimes I_n, \qquad E = \begin{bmatrix} I_{n^2} \\ 0_{n^2 \times n^2} \end{bmatrix}.$$

Here the rectangular upper part means the first n^2 rows of the *unscaled* Dirichlet Laplacian $L_{n,2n}$: $A_{1,2} = E^\mathsf{T} L_{n,2n}$ and $A_{2,1} = L_{n,2n} E$. The positivity assertion refers to this explicit normalization and ordering. Put $Q_{n,m} = I_m \otimes T_n(q_\alpha) + T_m(q_\beta) \otimes I_n$. For the separate coupling comparison, let $u = (\pi - \theta)/2$ for $0 \leq \theta \leq \pi$. The defining symbol gives

$$q_\gamma(\theta) = \left(2\sin(\theta/2)\right)^\gamma \left[\gamma\cos((2-\gamma)u) - (2-\gamma)\cos(\gamma u)\right].$$

For fixed $1 < \gamma < 2$, the bracket is bounded below by $2(\gamma-1)\cos((2-\gamma)\pi/2) > 0$ and above by 2. For the parameters used here, $\alpha = 3/2$ and $\beta = 9/5$, this yields, for $3/2 \le \gamma < 2$ and $0 < |\theta| \le \pi$,

$$\frac{q_\gamma(\theta)}{\ell(\theta)} > 2^{\gamma-2}\,2(\gamma-1)\cos\!\left(\frac{(2-\gamma)\pi}{2}\right) \ge \frac{1}{2}.$$

The last expression is increasing in γ and equals $1/2$ at $\gamma = 3/2$. Strict positivity of the difference symbol almost everywhere and the Toeplitz quadratic-form identity give $Q_{n,m} - \frac{1}{2}L_{n,m} \succ 0$. The identity $E^{\mathsf{T}} L_{n,2n} E = L_{n,n}$ yields a decomposition of A_n into a positive definite matrix and a positive semidefinite matrix:

$$A_n = \begin{bmatrix} 4Q_{n,n} - 2L_{n,n} & 0 \\ 0 & Q_{n,2n} - \frac{1}{2}L_{n,2n} \end{bmatrix} + \begin{bmatrix} \sqrt{2}E^{\mathsf{T}} \\ I_{2n^2}/\sqrt{2} \end{bmatrix} L_{n,2n} \begin{bmatrix} \sqrt{2}E & I_{2n^2}/\sqrt{2} \end{bmatrix}$$

Thus the explicitly specified A_n is real symmetric positive definite for every $n \ge 1$. This is a proof for the stated construction, not a verification of the implementation used to generate the reported data.

The preconditioner is chosen as

$$P_n = \begin{bmatrix} \tau_n(f_{1,1}) & \tau_n(f_{2,1}) & \mathbf{0}_{n^2\times n^2} \\ \tau_n(f_{2,1}) & \tau_n(f_{2,2}) & \mathbf{0}_{n^2\times n^2} \\ \mathbf{0}_{n^2\times n^2} & \mathbf{0}_{n^2\times n^2} & \tau_n(f_{2,2}) \end{bmatrix},$$

with nonzero blocks from the τ algebra. The analysis in [10] treats the specified positive constant-coefficient fractional Kronecker-sum class. It motivates the choice of the individual τ blocks, but does not establish the corresponding preconditioning guarantees for the assembled unequal-block matrix used in this exploratory example. The blocks have the representation

$$\tau_n(f(\theta_1,\theta_2)) = (\mathbb{S}_n \otimes \mathbb{S}_n)\, D\, (\mathbb{S}_n \otimes \mathbb{S}_n),$$

where D is diagonal and contains samples of f on an $n \times n$ grid, and $\mathbb{S}_n$ is the discrete sine transform matrix defined by

$$[\mathbb{S}_n]_{i,j} = \sqrt{\frac{2}{n+1}}\sin(i\theta_j), \qquad i,j = 1,\dots,n.$$

Since the Laplacian belongs to the τ algebra, $\tau_n(f_{2,1})$ is the Laplacian itself. The two-dimensional transforms cost $O(n^2 \log n)$, equivalently $O(d_n \log d_n)$ up to fixed factors because $d_n = 3n^2$. Positive definiteness follows directly from the sine-grid blocks. At each sine-grid point (θ_j, θ_k), where $\theta_j = j\pi/(n+1)$, put $q = q_\alpha(\theta_j) + q_\beta(\theta_k)$ and $l = \ell(\theta_j) + \ell(\theta_k)$. Here $f_{2,1}$ is the Laplacian symbol $\ell(\theta_1) + \ell(\theta_2)$ of the specified coupling. After the orthogonal sine transform and a permutation, P_n is a direct sum of matrices

$$\begin{bmatrix} 4q & l & 0 \\ l & q & 0 \\ 0 & 0 & q \end{bmatrix}.$$

The inequality $q > l/2 > 0$ proved above gives $4q^2 - l^2 > 0$, so all these matrices, and hence P_n, are positive definite. Consequently $P_n^{-1}A_n$ is similar to the symmetric positive definite matrix $P_n^{-1/2}A_nP_n^{-1/2}$ and has real positive eigenvalues.

Figure 3.9 compares the eigenvalues before and after preconditioning. Table 3.5 shows the relative outlier count decreasing from 0.088 to 0.045 as the total matrix order grows from 1200 to 4800, while PCG requires 14, 16, and 18 iterations. The entries > 120, > 169, and > 215 give lower bounds on the CG counts; the stopping or reporting procedure behind these bounds is not specified in the article. The results support the use of the preconditioner at the tested sizes and motivate a multilevel extension.

Table 3.5: Example 3.13: absolute and relative numbers of eigenvalues of $P_n^{-1}A_n$ outside $[0.95, 1.05]$, together with the CG and PCG iteration counts, for $d_n \in \{1200, 2700, 4800\}$. The relative count is normalized by d_n.

n	d_n	N	N/d_n	Iterations for CG	Iterations for PCG
20	1200	106	0.088	> 120	14
30	2700	162	0.060	> 169	16
40	4800	217	0.045	> 215	18

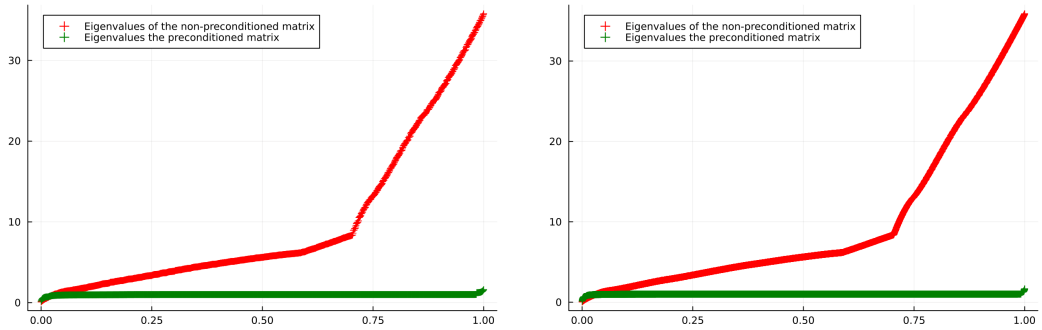


Figure 3.9: Example 3.13: eigenvalues of the unpreconditioned and preconditioned matrices in the two-dimensional case, for global matrix order $d_n = 1200$ (left) and $d_n = 4800$ (right), arranged in non-decreasing order.

3.6 Conclusions

Following [9], this chapter constructed a structured approximant S_n for unequal-block unilevel Toeplitz sequences with rational limiting ratios. Truncation and replication preserve the limiting distributions, and the blockwise replacement yields a zero-distributed error. Continuity and strong blockwise clustering strengthen the error result, while the stated rank and definiteness hypotheses give weak clustering at one for the appropriate preconditioned sequences.

The reported experiments cover square and rectangular systems, scalar- and matrix-valued symbols, and dense fractional-diffusion examples. They show the approximation behavior, spectral concentration, and reduction in iteration counts at the tested sizes.

The two-level experiment motivates a multilevel extension; the approximation and preconditioning results proved here remain unilevel.

For the matrices studied here, the limiting distributions were known before the structured approximant was constructed. The next chapter begins instead by identifying the complex symbols associated with quaternion Toeplitz sequences and explaining how their spectra encode canonical quaternion right eigenvalues.

Chapter 4

Quaternion Block Multilevel Toeplitz Sequences

4.1 Scope and main results

Quaternion matrices arise in multichannel signal and image processing, polarization models, and low-rank quaternion approximation; see, for example, [104, 38, 136, 119]. Toeplitz and circulant structures are particularly relevant when the underlying operations are discrete convolutions on one- or higher-dimensional tori. Their large-scale analysis requires asymptotic information on singular values and, after an appropriate spectral convention has been fixed, on eigenvalues. Two features distinguish the quaternion setting from the complex theory. Right eigenvalues occur in similarity classes, so eigenvalue distributions require a canonical representative of each class. Noncommutativity also makes the placement of the Fourier kernels consequential: left, right, and sandwich constructions generally produce different finite matrices and different complex symbols.

Based on the manuscript [112], this chapter studies quaternion block multilevel Toeplitz sequences with a fixed Fourier axis. Here *single-axis* means that all Fourier kernels use the same imaginary unit $\mathbf{i}$; the generating function may depend on several variables and take values in the full quaternion algebra. The central step is an exact representation of each kernel construction by a complex block Toeplitz matrix with an explicit $2s \times 2t$ symbol.

The symplectic representation provides singular-value distributions for rectangular sequences and canonical-eigenvalue distributions for square sequences, with normalizations accounting for the doubled dimensions and paired spectral data. It also connects quaternion and complex approximating classes of sequences. The product rule includes a compatibility condition for varying aspect ratios. For L^1 Toeplitz generators, the embeddings yield singular-value distributions and, for Hermitian sequences, eigenvalue distributions. The non-Hermitian eigenvalue result requires boundedness and the stated Tilli-class hypotheses. Further results concern Schatten norm estimates, Hermitian spectral localization and positivity, and circulant approximating classes for right Toeplitz sequences.

The chapter follows the passage from quaternion matrices to complex symbols. Section 4.2 fixes the algebraic and spectral conventions; Section 4.3 connects them to the distributions and a.c.s. calculus of Chapter 2. Section 4.4 gives the exact Toeplitz embeddings, their spectral consequences, and the circulant construction. Section 4.5 presents the numerical comparisons, followed by the conclusions in Section 4.6. Worked evaluations and supporting quaternion proofs are collected in Section 4.7.

4.2 Definitions and preliminary results

The common matrix norms and measurable-symbol conventions are those of Subsection 2.2.1. Here we fix the additional quaternion choices: a complex slice, a symplectic representation, canonical right eigenvalues, and the placement of the Fourier kernels. We follow the standard conventions of quaternion linear algebra; see [144] for a systematic treatment. The connection with complex linear algebra is based primarily on [175].

4.2.1 Quaternion algebra and complex slices

The classical Hamilton quaternion algebra is the skew field

$$\mathbb{H} := \operatorname{span}_{\mathbb{R}}\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}, \qquad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1,$$

where, by construction,

$$\mathbf{ij} = \mathbf{k}, \qquad \mathbf{jk} = \mathbf{i}, \qquad \mathbf{ki} = \mathbf{j}.$$

Throughout the chapter, we use the representation $\mathbf{k} = \mathbf{ij}$. Every $q \in \mathbb{H}$ has the Cartesian form

$$q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}, \qquad q_j \in \mathbb{R}.$$

Its real and imaginary parts are $\Re q := q_0$ and $\Im q := q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, respectively. Quaternion conjugation and modulus are defined by

$$\overline{q} = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}, \qquad |q| := \sqrt{q\,\overline{q}},$$

and satisfy $|pq| = |p|\,|q|$. For a quaternion or complex matrix $A = (a_{k\ell})$, we write $\overline{A} = (\overline{a_{k\ell}})$ for its entrywise conjugate.

A quaternion $\mu \in \mathbb{H}$ is an imaginary unit if $\Re\mu = 0$ and $\mu^2 = -1$. It determines the complex slice

$$\mathbb{C}_\mu := \operatorname{span}_{\mathbb{R}}\{1, \mu\} \cong \mathbb{C}.$$

If ν is a unit imaginary quaternion orthogonal to μ, equivalently $\nu\mu = -\mu\nu$, then [175, Thm. 2.1] gives the orthogonal decomposition

$$\mathbb{H} \;=\; \mathbb{C}_\mu \oplus \mathbb{C}_\mu\nu, \qquad \text{i.e.} \quad q = z + w\,\nu \text{ with } z, w \in \mathbb{C}_\mu, \tag{4.1}$$

and this representation is unique. Unless stated otherwise, we fix $\mu = \mathbf{i}$ and $\nu = \mathbf{j}$. The decomposition (4.1) provides the complex coordinates used by the symplectic representation below. The following identities will be used repeatedly:

$$\mathbf{j}\alpha = \overline{\alpha}\,\mathbf{j} \quad \text{for all} \quad \alpha \in \mathbb{C}_{\mathbf{i}}, \tag{4.2}$$

and the associated matrix version

$$\mathbf{j}A = \overline{A}\,\mathbf{j} \quad \text{for all} \quad A \in \mathbb{C}_{\mathbf{i}}^{m\times n}. \tag{4.3}$$

These identities determine how the complex Fourier factors interact with the two slice components.

4.2.2 Symplectic embedding

Let $x = z + w\mathbf{j}$, with $z, w \in \mathbb{C}_{\mathbf{i}}$, be the decomposition in (4.1). The classical complex symplectic representation [115] is

$$\Phi(x) := \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \;\in\; \mathbb{C}^{2\times 2}. \tag{4.4}$$

For $A = Z + W\mathbf{j}$, with $Z, W \in \mathbb{C}_{\mathbf{i}}^{m\times n}$, define the grouped realization

$$\Phi_{\mathrm{grp}}(A) \coloneqq \begin{bmatrix} Z & W \\ -\overline{W} & \overline{Z} \end{bmatrix} \in \mathbb{C}^{2m\times 2n}.$$

The entrywise extension is denoted separately: for $A = (a_{k\ell}) \in \mathbb{H}^{m\times n}$, set

$$\Phi_{\mathrm{ent}}(A) \coloneqq \big[\Phi(a_{k\ell})\big]_{k,\ell} \in \mathbb{C}^{2m\times 2n}.$$

We use Φ_{ent} for the spectral distributions and a.c.s. calculus, and Φ_{grp} to display the slice components Z, W. In Subsection 4.4.3, Φ_{blk} will apply the grouped map within each Toeplitz block while retaining the outer indices. These are coordinate orderings of the same representation, not different spectral models. For the two orderings defined here,

$$P_m^*\, \Phi_{\mathrm{ent}}(A)\, P_n \;=\; \Phi_{\mathrm{grp}}(A), \tag{4.5}$$

where $P_n \in \{0,1\}^{2n\times 2n}$ is the permutation

$$P_n \;\coloneqq\; \sum_{i=1}^{n} \big(e_{2i-1}e_i^* + e_{2i}e_{n+i}^*\big),$$

and $\{e_j\}_{j=1}^{2n}$ is the standard basis of $\mathbb{C}^{2n}$. Thus the two realizations have the same singular values; when $m = n$, they also have the same eigenvalues. The entrywise map Φ_{ent} is the injective real-linear $*$-algebra representation used for products and inverses [115, 175]. Here and below, X^* denotes the quaternion or complex conjugate transpose, according to the entries of X.

Lemma 4.1 ($*$-algebra property). Φ_{ent} is an injective real-linear $*$-algebra map. In particular,

$$\Phi_{\mathrm{ent}}(AB) = \Phi_{\mathrm{ent}}(A)\Phi_{\mathrm{ent}}(B), \qquad \Phi_{\mathrm{ent}}(A^*) = \Phi_{\mathrm{ent}}(A)^*.$$

4.2.3 Spectral tools for quaternion matrices

Because quaternion multiplication is noncommutative, the left and right eigenvalue problems are distinct. This chapter uses right eigenvalues.

Canonical right eigenvalues

Definition 4.2 (Right eigenpair). A nonzero $x \in \mathbb{H}^m$ and $\lambda \in \mathbb{H}$ form a right eigenpair of $A \in \mathbb{H}^{m\times m}$ if $Ax = x\lambda$.

A quaternion matrix may have infinitely many right eigenvalues, but they are organized into similarity classes.

Lemma 4.3 (Conjugacy classes). If $Ax = x\lambda$, then $A(xq) = (xq)(q^{-1}\lambda q)$ for every nonzero q. Thus right eigenvalues occur in conjugacy classes

$$[\lambda] \coloneqq \{\, q^{-1}\lambda q : q \neq 0 \,\}.$$

The number of right-eigenvalue classes is finite.

Lemma 4.4 ([175, Thm. 5.4]). Every $A \in \mathbb{H}^{n\times n}$ has exactly n right-eigenvalue classes, counted with multiplicity.

If a right eigenvalue is real, its similarity class is a singleton. This observation is particularly useful for Hermitian matrices.

A standard convention selects one complex representative with nonnegative imaginary part from each right-eigenvalue class.

Theorem 4.5 ([175, Thm. 5.4]). Every $A \in \mathbb{H}^{n\times n}$ has exactly n right eigenvalues in the closed upper half-plane, counted with multiplicity.

The canonical eigenvalues can be recovered from $\Phi_{\text{ent}}(A) \in \mathbb{C}^{2n\times 2n}$. The eigenvalues of the embedded matrix occur in conjugate pairs $\{z, \overline{z}\}$, and every real eigenvalue has even algebraic multiplicity. Retaining the eigenvalues in the open upper half-plane and half the algebraic multiplicity of each real eigenvalue gives the n canonical representatives. The measurability needed for distribution formulas is discussed in Subsection 4.2.4.

Singular values and Schatten norms

The singular value decomposition (SVD) over $\mathbb{H}$ mirrors the complex case.

Theorem 4.6 (Quaternion SVD, [175, Thm. 7.2]). For $A \in \mathbb{H}^{m\times n}$ there exist quaternion unitary matrices U, V of appropriate dimensions, $\rho = \operatorname{rank}_{\mathbb{H}}(A)$, and $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_\rho)$ with $\sigma_j > 0$, $j = 1, \dots, \rho$, such that

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^*.$$

The σ_j are the positive square roots of the positive eigenvalues of A^*A; the remaining eigenvalues are zero.

Definition 4.7 (Quaternion Schatten norms). For $A \in \mathbb{H}^{m\times n}$, use the Schatten-norm formulas of Definition 2.1 with the quaternion singular values from Theorem 4.6, counted once rather than twice. We write $\|A\|_{S_p}$, with $\|A\| = \|A\|_{S_\infty}$. The quaternion SVD gives unitary invariance and the same trace, Frobenius, and spectral norm conventions.

Proposition 4.8 (Symplectic embedding duplicates singular values). Let $A \in \mathbb{H}^{m\times n}$. If $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the nonzero singular values of A, then the singular values of $\Phi_{\text{ent}}(A) \in \mathbb{C}^{2m\times 2n}$ are

$$\sigma_1, \sigma_1,\ \sigma_2, \sigma_2,\ \dots,\ \sigma_r, \sigma_r,$$

followed by zeros to reach size $\min\{2m, 2n\}$.

Proof of Proposition 4.8. By Theorem 4.6 and Lemma 4.1, $\Phi_{\text{ent}}(A)$ is unitarily equivalent to $\Phi_{\text{ent}}(D)$, where

$$D = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}.$$

Since D is real, $\Phi_{\text{ent}}(D)$ is, up to fixed row and column permutations, equal to $D \oplus D$. Hence every nonzero singular value of A occurs twice among the singular values of $\Phi_{\text{ent}}(A)$, while all the remaining singular values are zero. □

Corollary 4.9 (Rank and Schatten norms under Φ_{ent}). For $A \in \mathbb{H}^{m\times n}$ and $1 \leq p \leq \infty$,

$$\operatorname{rank}_{\mathbb{C}}\big(\Phi_{\text{ent}}(A)\big) = 2\operatorname{rank}_{\mathbb{H}}(A), \qquad \big\|\Phi_{\text{ent}}(A)\big\| = \|A\|, \qquad \big\|\Phi_{\text{ent}}(A)\big\|_{S_p} = 2^{1/p}\|A\|_{S_p}.$$

Equivalently,

$$\|A\|_{S_p} = 2^{-1/p}\,\big\|\Phi_{\text{ent}}(A)\big\|_{S_p}.$$

4.2.4 Measurability and L^p spaces

Let $D \subset \mathbb{R}^k$ be measurable, with $0 < |D| < \infty$. Through $\mathbb{H}^{m\times n} \cong \mathbb{R}^{4mn}$, measurability means measurability of the real coordinate functions. It is equivalent to measurability of $\Phi_{\mathrm{ent}} \circ f$, since the embedding and its inverse on its image are continuous. As in Subsection 2.2.1, L^0 denotes measurable fields modulo equality almost everywhere.

Lemma 2.2 handles the complex spectral data. The local lemma records the quaternion consequences, including the canonical eigenvalue selection. Here a symmetric function is invariant under permutations of its arguments.

Lemma 4.10 (Measurability of quaternion spectral data). Let $f : D \to \mathbb{H}^{r\times r}$ be measurable and let $p \in [1, \infty]$. Then:

(a) the map $x \mapsto \|f(x)\|_{S_p}$ is measurable;

(b) the ordered singular-value map

$$x \longmapsto \big(\sigma_1(f(x)), \dots, \sigma_r(f(x))\big)$$

is measurable;

(c) if $\lambda_1^+, \dots, \lambda_r^+$ are the Borel maps provided by Lemma 4.48, then, for every continuous symmetric function $g : \mathbb{C}^r \to \mathbb{C}$, the map

$$x \longmapsto g\big(\lambda_1^+(f(x)), \dots, \lambda_r^+(f(x))\big)$$

is measurable.

A proof is given in Section 4.7.2.

Definition 4.11 (Quaternion matrix-valued L^p spaces). For measurable $f : D \to \mathbb{H}^{m\times n}$, define $L^p(D; \mathbb{H}^{m\times n})$ by Definition 2.1, using the quaternion Schatten norms of Definition 4.7. The pointwise norm is the spectral norm when $p = \infty$, and functions are identified up to equality almost everywhere. As in the complex case, the L^p integrals are unnormalized; the factor $|D|^{-1}$ appears in distribution averages, not in these norms.

Remark 4.12 (Compatibility with Φ_{ent}). By Corollary 4.9, for $1 \le p \le \infty$,

$$\|f\|_{L^p(D;\mathbb{H}^{m\times n})} = 2^{-1/p}\, \|\Phi_{\mathrm{ent}} \circ f\|_{L^p(D;\mathbb{C}^{2m\times 2n})},$$

with the usual convention $2^{1/\infty} = 1$.

4.2.5 Left, right, and sandwich trigonometric polynomials

We first distinguish the three kernel placements, then define their Fourier coefficients and describe their complex slice representations. Throughout, $\mathbf{i}$ is the fixed Hamilton unit introduced in Subsection 4.2.1, and all exponential kernels belong to the slice $\mathbb{C}_{\mathbf{i}}$. We use $[-\pi, \pi)^d$ as the standard representative cell of $\mathbb{T}^d$, equipped with Lebesgue measure $\mathrm{d}\boldsymbol{\theta}$, and set

$$\langle \boldsymbol{m}, \boldsymbol{\theta} \rangle := \sum_{j=1}^{d} m_j \theta_j$$

for $\boldsymbol{m} \in \mathbb{Z}^d$ and $\boldsymbol{\theta} \in \mathbb{T}^d$. Because quaternion multiplication is noncommutative, the exponential factors may act on the left, on the right, or on both sides of the coefficients. The last case is called a sandwich kernel. For $d \in \mathbb{N}$, write $[d] := \{1, \dots, d\}$. For an ordered disjoint partition $S_L \dot\cup S_R = [d]$, define

$$\langle \boldsymbol{m}, \boldsymbol{\theta} \rangle_{S_L} := \sum_{j\in S_L} m_j \theta_j, \qquad \langle \boldsymbol{m}, \boldsymbol{\theta} \rangle_{S_R} := \sum_{j\in S_R} m_j \theta_j,$$

and define the corresponding left and right exponential kernels by

$$L_{S_L}(\boldsymbol{\theta},\boldsymbol{m}) \coloneqq \exp\bigl(-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle_{S_L}\bigr), \qquad R_{S_R}(\boldsymbol{\theta},\boldsymbol{m}) \coloneqq \exp\bigl(-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle_{S_R}\bigr).$$

The purely left choice is $S_L = [d]$, $S_R = \varnothing$, whereas the purely right choice is $S_L = \varnothing$, $S_R = [d]$.

Definition 4.13 (Left, right, and sandwich trigonometric polynomials). Let $A_{\boldsymbol{m}} \in \mathbb{H}^{m\times n}$ and $\mathcal{F} \subset \mathbb{Z}^d$ be finite. Define

$$p_{\mathrm{L}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}} e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle} A_{\boldsymbol{m}}, \qquad p_{\mathrm{R}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}} A_{\boldsymbol{m}} e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle},$$

and, for a fixed partition $S_L \dot\cup S_R = [d]$,

$$p^{(S_L,S_R)}(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}} L_{S_L}(\boldsymbol{\theta},\boldsymbol{m})\, A_{\boldsymbol{m}}\, R_{S_R}(\boldsymbol{\theta},\boldsymbol{m}).$$

We call $p^{(S_L,S_R)}$ a sandwich trigonometric polynomial. For this representation, $\deg p$ denotes the degree bound $\deg p \coloneqq \max_{\boldsymbol{m}\in\mathcal{F}} \|\boldsymbol{m}\|_\infty$, with $\deg p \coloneqq 0$ when $\mathcal{F} = \varnothing$.

The Fourier coefficients and partial sums depend on the position of the kernel factors.

Definition 4.14 (Sandwich Fourier coefficients). For $F \in L^1(\mathbb{T}^d;\mathbb{H}^{m\times n})$ and $\boldsymbol{m} \in \mathbb{Z}^d$, define

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{m}) \coloneqq (2\pi)^{-d} \int_{\mathbb{T}^d} L_{S_L}(\boldsymbol{\theta},\boldsymbol{m})\, F(\boldsymbol{\theta})\, R_{S_R}(\boldsymbol{\theta},\boldsymbol{m})\, \mathrm{d}\boldsymbol{\theta} \in \mathbb{H}^{m\times n}. \tag{4.6}$$

For $N \in \mathbb{N}$, the rectangular partial sums are

$$S_N^{(S_L,S_R)} F(\boldsymbol{\theta}) \coloneqq \sum_{\|\boldsymbol{m}\|_\infty \le N} L_{S_L}(\boldsymbol{\theta},-\boldsymbol{m})\, \widehat{F}^{(S_L,S_R)}(\boldsymbol{m})\, R_{S_R}(\boldsymbol{\theta},-\boldsymbol{m}).$$

The left and right cases are obtained by $S_L = [d], S_R = \varnothing$ and $S_L = \varnothing, S_R = [d]$, respectively.

If

$$p^{(S_L,S_R)}(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}} L_{S_L}(\boldsymbol{\theta},\boldsymbol{m})\, A_{\boldsymbol{m}}\, R_{S_R}(\boldsymbol{\theta},\boldsymbol{m}),$$

then

$$\widehat{p}^{(S_L,S_R)}(\boldsymbol{k}) = A_{-\boldsymbol{k}},$$

where coefficients outside $\mathcal{F}$ are understood to be zero. The negative index is a consequence of the negative-exponential convention in both the polynomial representation and the coefficient definition.

Lemma 4.15 (Transport under the symplectic embedding). Write $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ with $Z, W : \mathbb{T}^d \to \mathbb{C}_{\mathbf{i}}^{m\times n}$. Define the sign-flip $\tilde{\boldsymbol{m}} \in \mathbb{Z}^d$ by $\tilde m_j = m_j$ if $j \in S_L$ and $\tilde m_j = -m_j$ if $j \in S_R$. Then, for every $\boldsymbol{m} \in \mathbb{Z}^d$,

$$\Phi_{\mathrm{grp}}\bigl(\widehat{F}^{(S_L,S_R)}(\boldsymbol{m})\bigr) = \begin{bmatrix} \widehat{Z}(\boldsymbol{m}) & \widehat{W}(\tilde{\boldsymbol{m}}) \\ -\widehat{\overline{W}}(-\tilde{\boldsymbol{m}}) & \widehat{\overline{Z}}(-\boldsymbol{m}) \end{bmatrix},$$

where the hats on the right denote ordinary complex matrix Fourier coefficients on $\mathbb{T}^d$.

A proof is given in Section 4.7.2.

The same representation also gives the following density result.

Proposition 4.16 (Density in L^p). For $1 \le p < \infty$, the linear span of sandwich trigonometric polynomials $\{p^{(S_L,S_R)}\}$ is dense in $L^p(\mathbb{T}^d;\mathbb{H}^{m\times n})$.

A proof is given in Section 4.7.2.

The commutation identity (4.2) reduces every single-axis sandwich kernel to a right-kernel representation. This is a special case of the splitting techniques used for quaternion Fourier transforms with two axes; see [96]. For matrix-valued polynomials and Fourier coefficients, the reduction is as follows.

Proposition 4.17 (Reduction to the right case). Let $S_L \dot\cup S_R = [d]$ and define

$$\tilde m_j = \begin{cases} m_j, & j \in S_L, \\ -m_j, & j \in S_R. \end{cases}$$

For finite $\mathcal{F} \subset \mathbb{Z}^d$ and $A_{\boldsymbol m} \in \mathbb{H}^{m\times n}$, write $A_{\boldsymbol m} = Z_{\boldsymbol m} + W_{\boldsymbol m}\mathbf{j}$ with $Z_{\boldsymbol m}, W_{\boldsymbol m} \in \mathbb{C}_{\mathbf i}^{m\times n}$. With L_{S_L}, R_{S_R} as in Definitions 4.13 and 4.14, the sandwich polynomial

$$p^{(S_L,S_R)}(\boldsymbol\theta) = \sum_{\boldsymbol m\in\mathcal{F}} L_{S_L}(\boldsymbol\theta,\boldsymbol m)\, A_{\boldsymbol m}\, R_{S_R}(\boldsymbol\theta,\boldsymbol m)$$

splits into two complex slice-valued trigonometric polynomials, with the second multiplied by $\mathbf j$ on the right:

$$p^{(S_L,S_R)}(\boldsymbol\theta) = \sum_{\boldsymbol m\in\mathcal{F}} \Big(Z_{\boldsymbol m} e^{-\mathbf i\langle \boldsymbol m,\boldsymbol\theta\rangle} + W_{\boldsymbol m} e^{-\mathbf i\langle \tilde{\boldsymbol m},\boldsymbol\theta\rangle}\mathbf j \Big). \tag{4.7}$$

Moreover, for $F = Z + W\mathbf j \in L^1(\mathbb{T}^d;\mathbb{H}^{m\times n})$ with $Z, W : \mathbb{T}^d \to \mathbb{C}_{\mathbf i}^{m\times n}$, the sandwich Fourier coefficients

$$\widehat F^{(S_L,S_R)}(\boldsymbol m) = (2\pi)^{-d} \int_{\mathbb{T}^d} L_{S_L}(\boldsymbol\theta,\boldsymbol m)\, F(\boldsymbol\theta)\, R_{S_R}(\boldsymbol\theta,\boldsymbol m)\, \mathrm{d}\boldsymbol\theta$$

satisfy

$$\widehat F^{(S_L,S_R)}(\boldsymbol m) = \widehat Z(\boldsymbol m) + \widehat W(\tilde{\boldsymbol m})\,\mathbf j, \tag{4.8}$$

where the hats on the right are the usual complex (right) Fourier coefficients on $\mathbb{T}^d$ in the slice $\mathbb{C}_{\mathbf i}$. In particular,

$$\widehat F^{(\mathrm L)}(\boldsymbol m) = \overline{\widehat{\overline F}^{(\mathrm R)}(-\boldsymbol m)}\,. \tag{4.9}$$

A proof is given in Section 4.7.2.

The transformation $\boldsymbol m \mapsto \tilde{\boldsymbol m}$ acts on frequency indices. Later, the Toeplitz reduction will express the corresponding change through a reflection of the generating function's spatial argument.

Remark 4.18. Applied to the purely left kernel, the proposition gives $\widehat F^{(\mathrm R)}(\boldsymbol m) = \overline{\widehat{\overline F}^{(\mathrm L)}(-\boldsymbol m)}$.

4.3 Quaternion spectral distributions and approximation

We apply Definition 2.3 to the symplectic embeddings. Matrix dimensions and singular-value lists double, while eigenvalues occur in conjugate pairs; the formulas below record the resulting normalization in the standard Weyl framework [16, 26, 79, 160, 170]. Eigenvalues are counted with algebraic multiplicity and singular values in nonincreasing order. Lemma 4.10 ensures measurability of the spectral integrands.

Eigenvalue distributions via the symplectic embedding

Definition 4.19 (Embedding-based eigenvalue symbol). Let $A_n \in \mathbb{H}^{d_n\times d_n}$ be square quaternion matrices with $d_n \to \infty$. Let $D \subset \mathbb{R}^k$ be measurable with $0 < |D| < \infty$, and let

$G : D \to \mathbb{C}^{q\times q}$ be a measurable matrix-valued function. We say that the quaternion sequence $\{A_n\}_n$ is distributed in eigenvalues as G over D, and we write

$$\{A_n\}_n \sim_\lambda (G, D),$$

if and only if the complex embeddings satisfy $\{\Phi_{\mathrm{ent}}(A_n)\}_n \sim_\lambda (G, D)$ in the sense of Definition 2.3. Concretely, for every $F \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{2d_n} \sum_{j=1}^{2d_n} F\big(\lambda_j(\Phi_{\mathrm{ent}}(A_n))\big) = \frac{1}{|D|} \int_D \frac{1}{q} \sum_{i=1}^{q} F\big(\lambda_i(G(x))\big)\,\mathrm{d}x, \tag{4.10}$$

where the eigenvalues $\lambda_j(\Phi_{\mathrm{ent}}(A_n))$ and $\lambda_i(G(x))$ are listed with algebraic multiplicity.

Let $\lambda_1^+(A_n), \dots, \lambda_{d_n}^+(A_n)$ denote the canonical eigenvalues of A_n, as defined in Theorem 4.5. Then the eigenvalues of $\Phi_{\mathrm{ent}}(A_n)$ occur in conjugate pairs $\{z, \overline{z}\}$; real eigenvalues appear as pairs $\{z, z\}$. Therefore, for every $F \in C_c(\mathbb{C})$,

$$\frac{1}{2d_n} \sum_{j=1}^{2d_n} F\big(\lambda_j(\Phi_{\mathrm{ent}}(A_n))\big) = \frac{1}{d_n} \sum_{j=1}^{d_n} \widetilde{F}\big(\lambda_j^+(A_n)\big), \qquad \widetilde{F}(z) := \tfrac{1}{2}\big(F(z) + F(\overline{z})\big). \tag{4.11}$$

Equation (4.11) averages each conjugate pair $\{z, \overline{z}\}$ via $\widetilde{F}$. For conjugation-symmetric test functions, $\widetilde{F} = F$, so the normalized average can be written directly in terms of the canonical eigenvalues.

Canonical eigenvalues and conjugate pairing

The canonical formulation uses conjugation-symmetric test functions, which do not distinguish a complex value from its conjugate. Remark 4.21 gives the implication from the embedding-based notion and a sufficient symmetry condition for the converse.

Definition 4.20 (Canonical-eigenvalue distribution). Let $A_n \in \mathbb{H}^{d_n\times d_n}$ with $d_n \to \infty$, and let $\lambda_1^+(A_n), \dots, \lambda_{d_n}^+(A_n)$ be the canonical eigenvalues as in Theorem 4.5. Let $D \subset \mathbb{R}^k$ be measurable with $0 < |D| < \infty$, and let $G \in L^0(D; \mathbb{C}^{q\times q})$. We write $\{A_n\}_n \sim_{\lambda^+} (G, D)$ if, for every $F \in C_c(\mathbb{C})$ with $F(\overline{z}) = F(z)$,

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F\big(\lambda_j^+(A_n)\big) = \frac{1}{|D|} \int_D \frac{1}{q} \sum_{i=1}^{q} F\big(\lambda_i(G(x))\big)\,\mathrm{d}x.$$

Remark 4.21. By (4.11), if $\{A_n\}_n \sim_\lambda (G, D)$ in the sense of Definition 4.19, then $\{A_n\}_n \sim_{\lambda^+} (G, D)$ in the sense of Definition 4.20. Conversely, if for a.e. $x \in D$ the multiset $\{\lambda_i(G(x))\}_{i=1}^q$ is invariant under complex conjugation, then $\{A_n\}_n \sim_{\lambda^+} (G, D)$ implies $\{A_n\}_n \sim_\lambda (G, D)$.

Remark 4.22 (Measure-preserving reparametrizations). By Remark 2.4, a measure-preserving change of variables leaves embedding-based and canonical-eigenvalue distributions unchanged. The same argument applies to conjugation-symmetric test functions.

Singular-value distributions via the symplectic representation

Definition 4.23 (Embedding-based singular-value symbol). Let $A_n \in \mathbb{H}^{m_n\times n_n}$ be rectangular quaternion matrices and set $r_n := \min\{m_n, n_n\} \to \infty$. Let $D \subset \mathbb{R}^k$ be measurable with $0 < |D| < \infty$, and let $H \in L^0(D; \mathbb{C}^{s\times t})$ be measurable. We say that $\{A_n\}_n$ is distributed in singular values as H over D, and write

$$\{A_n\}_n \sim_\sigma (H, D),$$

if and only if $\{\Phi_{\mathrm{ent}}(A_n)\}_n \sim_\sigma (H, D)$ in the sense of Definition 2.3, i.e., for every $\psi \in C_c([0,\infty))$,

$$\lim_{n\to\infty} \frac{1}{2r_n} \sum_{i=1}^{2r_n} \psi\big(\sigma_i(\Phi_{\mathrm{ent}}(A_n))\big) = \frac{1}{|D|} \int_D \frac{1}{\min\{s,t\}} \sum_{i=1}^{\min\{s,t\}} \psi\big(\sigma_i(H(x))\big)\,\mathrm{d}x. \tag{4.12}$$

Remark 4.24 (Duplication of singular values under the embedding). By Proposition 4.8, the singular values of $\Phi_{\mathrm{ent}}(A_n)$ are exactly those of A_n, each duplicated. Consequently, for every $\psi \in C_c([0,\infty))$,

$$\frac{1}{2r_n} \sum_{i=1}^{2r_n} \psi\big(\sigma_i(\Phi_{\mathrm{ent}}(A_n))\big) = \frac{1}{r_n} \sum_{i=1}^{r_n} \psi\big(\sigma_i(A_n)\big), \qquad \forall\, \psi \in C_c([0,\infty)). \tag{4.13}$$

Thus the empirical singular-value distribution may be evaluated directly on the quaternion matrices.

4.3.1 Quaternion approximating classes and their calculus

The quaternion formulation uses the sparse-tail conditions and a.c.s. calculus of Subsections 2.2.4 and 2.2.5, with quaternion ranks. The result below justifies passage between the two scalar fields through the symplectic embedding.

Definition 4.25 (Quaternion approximating classes and sparse tails). For $A_n, B_{n,m} \in \mathbb{H}^{d_n\times e_n}$, with $r_n = \min\{d_n, e_n\} \to \infty$, use Definition 2.16 with quaternion error matrices and $\mathrm{rank}_{\mathbb{H}}$. Thus $A_n - B_{n,m}$ splits into a low-rank term and a small-operator-norm term: their rank fraction and norm bound tend to zero as $m \to \infty$, after the limit in n. We write $\{B_{n,m}\}_n \xrightarrow[\mathbb{H}]{\text{a.c.s.}} \{A_n\}_n$. Sparse unboundedness and sparse vanishing are the large- and near-zero singular-value tail conditions of Definition 2.13, respectively, evaluated on quaternion singular values and normalized by r_n.

For multilevel families, these notions are required along every scalar extraction $\boldsymbol{n} = \boldsymbol{n}(n)$ with $\min_j n_j(n) \to \infty$, exactly as in Chapter 2. The external approximation index m remains scalar, and the product-compatibility condition is Definition 2.15. By singular-value duplication, a quaternion sequence is sparsely unbounded or sparsely vanishing if and only if its complex symplectic realization has the corresponding property.

Corollary 4.9 gives the forward implication: rank and active dimension both double, while the operator norm is unchanged. For the converse, arbitrary complex errors need not be symplectic images; the proof projects them onto that image to recover quaternion errors.

Theorem 4.26 (Embedding equivalence). Let $\{A_n\}_n$ be a matrix-sequence with $A_n \in \mathbb{H}^{d_n\times e_n}$ and $r_n := d_n \wedge e_n \to \infty$, and let $\big\{\{B_{n,m}\}_n\big\}_{m\in\mathbb{N}}$ be given, with $B_{n,m} \in \mathbb{H}^{d_n\times e_n}$. Then the following are equivalent:

(a) $\big\{\{B_{n,m}\}_n\big\}_{m\in\mathbb{N}}$ is a quaternion a.c.s. for $\{A_n\}_n$.

(b) $\big\{\{\Phi_{\mathrm{ent}}(B_{n,m})\}_n\big\}_{m\in\mathbb{N}}$ is a complex a.c.s. for $\{\Phi_{\mathrm{ent}}(A_n)\}_n$.

A proof is given in Section 4.7.2.

Quaternion a.c.s. calculus. The following rules are the quaternion forms of ACS 1–5 in Subsection 2.2.5. They use quaternion ranks and the active dimension $r_n = d_n \wedge e_n$. Symbols in the distribution rules have fixed matrix sizes on one domain of finite positive measure.

ACS$_{\mathbb{H}}$1. *Singular-value transfer.* If quaternion approximants have complex symbols f_m converging in measure to f, then a.c.s. convergence transfers their singular-value distribution to the target sequence. The full transfer/realization equivalence is ACS 1, with the distribution understood as in Definition 4.23.

ACS$_{\mathbb{H}}$2. *Hermitian eigenvalue transfer.* The same transfer and realization rule holds for quaternion Hermitian targets and approximants, with complex Hermitian symbols (ACS 2). The embedding duplicates real eigenvalues, so the normalized averages agree.

ACS$_{\mathbb{H}}$3. *Algebraic operations.* Adjoints and compatible sums preserve a.c.s. convergence; quaternion coefficients in sums act on the left. For products of two approximants, both target sequences must be sparsely unbounded. Multiplication by a fixed sequence requires sparse unboundedness only of that multiplier; fixed means independent of m, not of n. Every product must satisfy the corresponding PC condition of Definition 2.15, so that factor-rank errors remain small relative to the minimum output dimension. This is ACS 3 over $\mathbb{H}$.

ACS$_{\mathbb{H}}$4. *Block assembly.* Finitely many conformal blockwise approximations give an a.c.s. for the assembled quaternion sequence, as in ACS 4.

ACS$_{\mathbb{H}}$5. *Schatten test.* An error bound for $\|A_n - B_{n,m}\|_{S_p}/r_n^{1/p}$ that vanishes as $m \to \infty$, after taking the upper limit in n, suffices for a.c.s. convergence (ACS 5). For $p = \infty$, the denominator is one.

For the fixed block sizes over a common Toeplitz grid used below, PC is automatic. The quaternion-specific transfer steps, including left scalar multiplication and the doubled rank and Schatten normalizations, are verified in Subsection 4.7.2.

4.4 Single-axis quaternion Toeplitz sequences

This section develops the Toeplitz theory for multilevel block sequences generated by functions in $L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$. The scalar unilevel Hermitian case was studied in [118]. Here the generating functions are matrix valued and multivariate, and the Fourier kernel may act on the left, on the right, or on both sides of a coefficient. Under the single-axis assumption, these constructions are related by exact reduction formulas, although they generally produce different finite matrices. The subsequent analysis derives their complex embedded symbols and transfers the standard block multilevel Toeplitz distribution, Schatten, and localization results; see [77, Chapter 6], [79, Chapter 3], and [16].

4.4.1 d-level quaternion Toeplitz matrices

Let L_{S_L} and R_{S_R} be the kernels in Definition 4.13, and let the Fourier coefficients be those of Definition 4.14. For $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathbb{N}^d$, define the multi-index set

$$\Lambda_{\boldsymbol{n}} := \{0, \dots, n_1 - 1\} \times \cdots \times \{0, \dots, n_d - 1\}, \tag{4.14}$$

endowed with the standard lexicographic order. The d-level block Toeplitz matrix is

$$T_{\boldsymbol{n}}^{(S_L,S_R)}\big(F\big) := \big[\widehat{F}^{(S_L,S_R)}(\alpha - \beta)\big]_{\alpha,\beta\in\Lambda_{\boldsymbol{n}}} \in \mathbb{H}^{(N_{\boldsymbol{n}}s)\times(N_{\boldsymbol{n}}t)}, \qquad N_{\boldsymbol{n}} = \prod_{\ell=1}^{d} n_\ell. \tag{4.15}$$

For the purely left and purely right kernels, respectively, we use the abbreviations

$$T_{\boldsymbol{n}}^{(\mathrm{L})}(F) := T_{\boldsymbol{n}}^{([d],\varnothing)}(F), \qquad T_{\boldsymbol{n}}^{(\mathrm{R})}(F) := T_{\boldsymbol{n}}^{(\varnothing,[d])}(F).$$

The following proposition reduces every sandwich construction to a right-kernel Toeplitz matrix with a transformed generating function.

Proposition 4.27 (Right-kernel reduction for Toeplitz matrices). Let $F = Z + W\mathbf{j} \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$ with $Z, W : \mathbb{T}^d \to \mathbb{C}_{\mathbf{i}}^{s\times t}$ and fix an ordered partition $S_L \dot\cup S_R = [d]$. Consider

$$(r_{S_L}(\boldsymbol{\theta}))_j = \begin{cases} -\theta_j, & j \in S_L, \\ \theta_j, & j \in S_R, \end{cases}$$

and define

$$H(\boldsymbol{\theta}) := Z(\boldsymbol{\theta}) + \big(W \circ r_{S_L}\big)(\boldsymbol{\theta})\,\mathbf{j}. \tag{4.16}$$

Then, for every $\boldsymbol{n} \in \mathbb{N}^d$,

$$T_{\boldsymbol{n}}^{(S_L,S_R)}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(H). \tag{4.17}$$

In the left-kernel case ($S_L = [d], S_R = \varnothing$), we obtain

$$T_{\boldsymbol{n}}^{(\mathrm{L})}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}\Big(Z + \big(W \circ (-\mathrm{id})\big)\mathbf{j}\Big). \tag{4.18}$$

Moreover, for a fixed $\boldsymbol{n} \in \mathbb{N}^d$ and the Cartesian set

$$\Delta(\boldsymbol{n}) := \{\, \boldsymbol{k} \in \mathbb{Z}^d :\ -n_j + 1 \le k_j \le n_j - 1,\ j \in [d] \,\},$$

we have

$$T_{\boldsymbol{n}}^{(\mathrm{L})}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(F) \iff \widehat{W}(\boldsymbol{k}) = \widehat{W}(-\boldsymbol{k}) \quad \forall\, \boldsymbol{k} \in \Delta(\boldsymbol{n}). \tag{4.19}$$

The identity holds for every matrix size precisely when

$$T_{\boldsymbol{n}}^{(\mathrm{L})}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(F) \quad \text{for every } \boldsymbol{n} \in \mathbb{N}^d \iff \widehat{W}(\boldsymbol{k}) = \widehat{W}(-\boldsymbol{k}) \quad \forall\, \boldsymbol{k} \in \mathbb{Z}^d, \tag{4.20}$$

and the latter is equivalent to $W(\boldsymbol{\theta}) = W(-\boldsymbol{\theta})$ a.e. on $\mathbb{T}^d$.

A proof is given in Section 4.7.3.

The reduction changes the generating function: $T_{\boldsymbol{n}}^{(S_L,S_R)}(F)$ is represented by $T_{\boldsymbol{n}}^{(\mathrm{R})}(H)$, not in general by $T_{\boldsymbol{n}}^{(\mathrm{R})}(F)$. The final conditions in the proposition distinguish equality at a fixed size from equality for every size. We next examine how the kernel placement interacts with adjoints and Hermitian structure.

4.4.2 Linearity, adjoints, and Hermitian structure

We first record linearity over the fixed complex slice. Taking adjoints interchanges the left and right kernel factors, leading to the Hermitian criterion that will also be used in the numerical examples.

Lemma 4.28 ($\mathbb{C}_{\mathbf{i}}$-linearity of sandwich Fourier coefficients). Fix an ordered partition $S_L \dot\cup S_R = [d]$. For the Fourier coefficients of Definition 4.14, every $F, G \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$, $a, b \in \mathbb{C}_{\mathbf{i}}$, and $\boldsymbol{k} \in \mathbb{Z}^d$ satisfy

$$\widehat{aF + bG}^{(S_L,S_R)}(\boldsymbol{k}) = a\,\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) + b\,\widehat{G}^{(S_L,S_R)}(\boldsymbol{k}). \tag{4.21}$$

A proof is given in Section 4.7.3.

Proposition 4.29 ($\mathbb{C}_{\mathbf{i}}$-linearity of quaternion Toeplitz matrices). For the Toeplitz matrices of (4.15), all $a, b \in \mathbb{C}_{\mathbf{i}}$ and $F, G \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$ satisfy

$$T_{\boldsymbol{n}}^{(S_L,S_R)}(aF + bG) = a\,T_{\boldsymbol{n}}^{(S_L,S_R)}(F) + b\,T_{\boldsymbol{n}}^{(S_L,S_R)}(G). \tag{4.22}$$

Proof of Proposition 4.29.

By definition, each block entry of $T_{\boldsymbol{n}}^{(S_L,S_R)}(H)$ equals $\widehat{H}^{(S_L,S_R)}(\boldsymbol{k})$ with $\boldsymbol{k} = \alpha - \beta$. Applying Lemma 4.28 entrywise with $H = aF + bG$ gives (4.22). □

We now characterize adjoints and give a criterion for Hermitian Toeplitz matrices. Consider

$$F : \mathbb{T}^d \to \mathbb{H}^{s\times t}, \qquad F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j},$$

with $Z, W \in L^1(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{s\times t})$. For $\boldsymbol{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d$, define its coordinate projections onto S_L and S_R by

$$(\boldsymbol{k}_L)_j := \begin{cases} k_j, & j \in S_L, \\ 0, & j \notin S_L, \end{cases} \qquad (\boldsymbol{k}_R)_j := \begin{cases} k_j, & j \in S_R, \\ 0, & j \notin S_R. \end{cases}$$

Thus $\boldsymbol{k} = \boldsymbol{k}_L + \boldsymbol{k}_R$. With this notation, for any ordered partition $S_L \dot\cup S_R = [d]$ we have

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) = \widehat{Z}(\,\boldsymbol{k}_L + \boldsymbol{k}_R\,) + \widehat{W}(\,\boldsymbol{k}_L - \boldsymbol{k}_R\,)\,\mathbf{j}, \tag{4.23}$$

where the hats denote classical complex Fourier coefficients.

Theorem 4.30 (Adjoint identities). Let $F : \mathbb{T}^d \to \mathbb{H}^{s\times t}$ with $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$, $Z, W \in L^1(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{s\times t})$, and let $\boldsymbol{n} \in \mathbb{N}^d$. Then

$$\big(T_{\boldsymbol{n}}^{(\mathrm{L})}(F)\big)^* = T_{\boldsymbol{n}}^{(\mathrm{R})}(F^*), \qquad \big(T_{\boldsymbol{n}}^{(\mathrm{R})}(F)\big)^* = T_{\boldsymbol{n}}^{(\mathrm{L})}(F^*), \qquad \big(T_{\boldsymbol{n}}^{(S_L,S_R)}(F)\big)^* = T_{\boldsymbol{n}}^{(S_R,S_L)}(F^*). \tag{4.24}$$

A proof is given in Section 4.7.3.

The next theorem extends the Hermitian criterion of [118] from scalar unilevel right Toeplitz matrices to the block multilevel sandwich setting.

Theorem 4.31 (Hermitian criterion for single-axis quaternion Toeplitz matrices). Assume $s = t$ and let $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ with $Z, W \in L^1(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{s\times s})$. Fix an ordered partition $S_L \dot\cup S_R = [d]$. For $\boldsymbol{n} \in \mathbb{N}^d$ set

$$T_{\boldsymbol{n}}^{(S_L,S_R)}(F) = \big[\widehat{F}^{(S_L,S_R)}(\alpha - \beta)\big]_{\alpha,\beta\in\Lambda_{\boldsymbol{n}}}.$$

Then the following are equivalent:

(i) $T_{\boldsymbol{n}}^{(S_L,S_R)}(F)$ is Hermitian for all $\boldsymbol{n}$;

(ii) $Z(\boldsymbol{\theta})$ is essentially Hermitian on $\mathbb{T}^d$ and $W(-\boldsymbol{\theta}) = -W(\boldsymbol{\theta})^{\mathsf{T}}$ a.e. on $\mathbb{T}^d$.

In particular, the claim does not depend on the choice of the ordered partition.

A proof is given in Section 4.7.3.

This criterion concerns the Hermitian character of the Toeplitz matrices and uses the slice components at reflected arguments. It should not be replaced by pointwise Hermitian symmetry of the quaternion-valued generator.

4.4.3 Embedded Toeplitz symbols and asymptotic distributions

The exact identities in this subsection connect the quaternion constructions to the complex asymptotic theory. We first identify the complex block Toeplitz symbols, then use them to obtain the singular-value and eigenvalue distributions.

Left/right and sandwich embeddings

Throughout this section, let $F = Z + W\mathbf{j} \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$ with $Z, W : \mathbb{T}^d \to \mathbb{C}_{\mathbf{i}}^{s\times t}$ slice-valued. Recall from Subsection 4.2.2 that the grouped realization of one quaternion block is

$$\Phi_{\mathrm{grp}}(Z + W\mathbf{j}) = \begin{bmatrix} Z & W \\ -\overline{W} & \overline{Z} \end{bmatrix}, \qquad Z, W \in \mathbb{C}_{\mathbf{i}}^{s\times t}.$$

For a Toeplitz matrix, we apply this grouped realization within each quaternion block while retaining the outer block indices. If $A = [A_{\alpha\beta}]_{\alpha,\beta\in\Lambda_{\boldsymbol{n}}} \in \mathbb{H}^{N_{\boldsymbol{n}}s\times N_{\boldsymbol{n}}t}$ is partitioned into $s\times t$ quaternion blocks, its block-ordered realization is

$$\Phi_{\mathrm{blk}}^{(s,t)}(A) := \big[\Phi_{\mathrm{grp}}(A_{\alpha\beta})\big]_{\alpha,\beta\in\Lambda_{\boldsymbol{n}}} \in \mathbb{C}^{(2sN_{\boldsymbol{n}})\times(2tN_{\boldsymbol{n}})}.$$

Suitable fixed permutation matrices $Q_{\boldsymbol{n},s}$ and $Q_{\boldsymbol{n},t}$ satisfy

$$\Phi_{\mathrm{blk}}^{(s,t)}(A) = Q_{\boldsymbol{n},s}^{*}\,\Phi_{\mathrm{ent}}(A)\,Q_{\boldsymbol{n},t}. \tag{4.25}$$

Thus the entrywise and block-ordered realizations have identical singular values and, in the square case, identical eigenvalues. For compatible block products, the common inner ordering cancels; in particular,

$$\Phi_{\mathrm{blk}}^{(s,u)}(AB) = \Phi_{\mathrm{blk}}^{(s,t)}(A)\Phi_{\mathrm{blk}}^{(t,u)}(B). \tag{4.26}$$

For a single $s\times t$ block, we will use the covariance identity

$$\Phi_{\mathrm{grp}}(\lambda A\rho) = \mathrm{diag}(\lambda,\overline{\lambda})\,\Phi_{\mathrm{grp}}(A)\,\mathrm{diag}(\rho,\overline{\rho}) \quad \big(\lambda\in\mathbb{C}_{\mathbf{i}}^{s\times s},\ \rho\in\mathbb{C}_{\mathbf{i}}^{t\times t}\big), \tag{4.27}$$

where $\mathrm{diag}(\lambda,\overline{\lambda})$ and $\mathrm{diag}(\rho,\overline{\rho})$ are 2×2 block-diagonal matrices with blocks of sizes $s\times s$ and $t\times t$, respectively. When $\lambda,\rho\in\mathbb{C}_{\mathbf{i}}$ are scalars we read them as λI_s and ρI_t. For a complex symbol $G:\mathbb{T}^d\to\mathbb{C}^{(2s)\times(2t)}$, $T_{\boldsymbol{n}}(G)$ is the complex Toeplitz matrix of Definition 2.22: Fourier coefficients use the negative exponential, and the block (α,β) has coefficient index $\alpha-\beta$.

Theorem 4.32 (Left/right embedding). With the decomposition $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ from Lemma 4.15, define the complex symbols

$$G_{\mathrm{L}}(\boldsymbol{\theta}) = \begin{pmatrix} Z(\boldsymbol{\theta}) & W(\boldsymbol{\theta}) \\ -\overline{W(-\boldsymbol{\theta})} & \overline{Z(-\boldsymbol{\theta})} \end{pmatrix}, \qquad G_{\mathrm{R}}(\boldsymbol{\theta}) = \begin{pmatrix} Z(\boldsymbol{\theta}) & W(-\boldsymbol{\theta}) \\ -\overline{W(\boldsymbol{\theta})} & \overline{Z(-\boldsymbol{\theta})} \end{pmatrix}.$$

Then, for every $\boldsymbol{n}\in\mathbb{N}^d$,

$$\Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(\mathrm{L})}(F)\big) = T_{\boldsymbol{n}}(G_{\mathrm{L}}), \qquad \Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(\mathrm{R})}(F)\big) = T_{\boldsymbol{n}}(G_{\mathrm{R}}).$$

A proof is given in Section 4.7.3.

We now treat the two-sided sandwich case using the involution $r_{S_L}:\mathbb{T}^d\to\mathbb{T}^d$ from Proposition 4.27. Note that $r_{S_L}\circ(-\mathrm{id})$ flips the S_R-coordinates and that $-r_{S_L}(-\boldsymbol{\theta}) = r_{S_L}(\boldsymbol{\theta})$.

Theorem 4.33 (Sandwich embedding). Let $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ with $Z, W:\mathbb{T}^d\to\mathbb{C}_{\mathbf{i}}^{s\times t}$ and fix an ordered partition $S_L\dot\cup S_R = [d]$. Define the complex symbol

$$G_{S_L,S_R}(\boldsymbol{\theta}) = \begin{pmatrix} Z(\boldsymbol{\theta}) & W\big(r_{S_L}(-\boldsymbol{\theta})\big) \\ -\overline{W\big(r_{S_L}(\boldsymbol{\theta})\big)} & \overline{Z(-\boldsymbol{\theta})} \end{pmatrix}.$$

Then, for every $\boldsymbol{n}\in\mathbb{N}^d$,

$$\Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(S_L,S_R)}(F)\big) = T_{\boldsymbol{n}}\big(G_{S_L,S_R}\big). \tag{4.28}$$

For $S_L = [d]$ and $S_L = \varnothing$, this formula reduces, respectively, to the symbols G_{L} and G_{R} in Theorem 4.32.

A proof is given in Section 4.7.3.

Singular-value and eigenvalue distributions

The embedding identities transfer the complex block Toeplitz singular-value and eigenvalue distribution theorems to the quaternion sequences. Throughout, write $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathbb{N}^d$ and $N_{\boldsymbol{n}} = \prod_{\ell=1}^{d} n_\ell$. For $\tau \in \{\mathrm{L}, \mathrm{R}, (S_L, S_R)\}$ we define the complex symbol

$$G_\tau(\boldsymbol{\theta}) := \begin{cases} G_{\mathrm{L}}(\boldsymbol{\theta}), & \tau = \mathrm{L}, \\ G_{\mathrm{R}}(\boldsymbol{\theta}), & \tau = \mathrm{R}, \\ G_{S_L,S_R}(\boldsymbol{\theta}), & \tau = (S_L, S_R), \end{cases}$$

where $G_{\mathrm{L}}, G_{\mathrm{R}}, G_{S_L,S_R}$ are those in Theorems 4.32 and 4.33. It is G_τ, not in general the pointwise matrix $\Phi_{\mathrm{grp}}(F)$, that enters the complex Toeplitz theorems. The reflected arguments in G_τ account for the Fourier-kernel placement.

Theorem 4.34 (Singular-value distribution and Hermitian eigenvalue distribution). Let $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ with $Z, W \in L^1(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{s\times t})$. For each $\tau \in \{\mathrm{L}, \mathrm{R}, (S_L, S_R)\}$,

$$\big\{ T_{\boldsymbol{n}}^{(\tau)}(F) \big\}_{\boldsymbol{n}} \sim_\sigma G_\tau. \tag{4.29}$$

If, in addition, $s = t$ and $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian for all $\boldsymbol{n}$, then also

$$\big\{ T_{\boldsymbol{n}}^{(\tau)}(F) \big\}_{\boldsymbol{n}} \sim_\lambda G_\tau.$$

Proof of Theorem 4.34. For each kernel choice τ, the embedding theorems give

$$\Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) = T_{\boldsymbol{n}}(G_\tau).$$

Hence it suffices to analyze $T_{\boldsymbol{n}}(G_\tau)$ on the complex side. Theorem 2.23 gives $\{T_{\boldsymbol{n}}(G_\tau)\}_{\boldsymbol{n}} \sim_\sigma G_\tau$. By (4.25), the block-ordered and entrywise representations are related by fixed unitary permutations. Corollary 4.9 then shows that their singular values are those of the quaternion matrix, each duplicated. The quaternion and complex normalizations therefore agree, proving (4.29). If $s = t$ and every $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian, then $\Phi_{\mathrm{blk}}^{(s,s)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big)$ is Hermitian as well. Theorem 2.23 gives $\{T_{\boldsymbol{n}}(G_\tau)\}_{\boldsymbol{n}} \sim_\lambda G_\tau$. The permutation relation (4.25) and the duplication of real Hermitian eigenvalues transfer the claim back to the quaternion sequence. □

For non-Hermitian sequences, singular-value convergence alone does not give an eigenvalue distribution. We instead apply Theorem 2.24 to the bounded complex symbol G_τ. Its spectral essential range (Definition 2.6) collects the complex values approached by its eigenvalues on sets of positive measure. The required geometric conditions are stated explicitly below.

Theorem 4.35 (Eigenvalue distribution for Tilli-class symbols). Assume $s = t$ and $F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}$ with $Z, W \in L^\infty(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{s\times s})$. For $\tau \in \{\mathrm{L}, \mathrm{R}, (S_L, S_R)\}$, consider the spectral essential range $\mathrm{ER}_\lambda(G_\tau)$. Suppose that G_τ belongs to the Tilli class, i.e.,

$$\operatorname{int} \mathrm{ER}_\lambda(G_\tau) = \varnothing \quad \text{and} \quad \mathbb{C} \setminus \mathrm{ER}_\lambda(G_\tau) \text{ is connected.}$$

Then

$$\big\{ T_{\boldsymbol{n}}^{(\tau)}(F) \big\}_{\boldsymbol{n}} \sim_\lambda G_\tau. \tag{4.30}$$

A proof is given in Section 4.7.3.

4.4.4 Schatten norm estimates

Let $F = Z + W\mathbf{j} \in L^p(\mathbb{T}^d; \mathbb{H}^{s\times t})$, with $1 \le p \le \infty$, and use the embedded symbols G_τ from the preceding subsection. The Schatten estimate combines their four-block structure with the complex Toeplitz bound. We recall the scaling from Corollary 4.9:

$$\|\Phi_{\mathrm{grp}}(X)\|_{S_p} = 2^{1/p}\,\|X\|_{S_p} \quad \text{(with the convention } 2^{1/\infty} = 1),$$

and integrating this identity gives

$$\|\Phi_{\mathrm{grp}} \circ F\|_{L^p(\mathbb{T}^d;\mathbb{C}^{(2s)\times(2t)})} = 2^{1/p}\,\|F\|_{L^p(\mathbb{T}^d;\mathbb{H}^{s\times t})}. \tag{4.31}$$

The complex Toeplitz–Schatten estimate used below is Theorem 2.25 in Chapter 2.

Theorem 4.36 (Quaternion matrix-valued Toeplitz–Schatten bound). Let $1 \le p \le \infty$ and $F \in L^p(\mathbb{T}^d; \mathbb{H}^{s\times t})$. For any $\tau \in \{\mathrm{L}, \mathrm{R}, (S_L, S_R)\}$ and $\boldsymbol{n} \in \mathbb{N}^d$, where $S_L \dot\cup S_R = [d]$ in the sandwich case, and with $N_{\boldsymbol{n}} := \prod_{\ell=1}^d n_\ell$,

$$\left\|T_{\boldsymbol{n}}^{(\tau)}(F)\right\|_{S_p} \le 4\left(\frac{N_{\boldsymbol{n}}}{(2\pi)^d}\right)^{1/p} \|F\|_{L^p(\mathbb{T}^d;\mathbb{H}^{s\times t})}. \tag{4.32}$$

The constant 4 is not sharp. The corresponding complex bound in Theorem 2.25 has constant 1, and the additional factor arises from the four-block projection estimate after the symplectic reduction. This constant is sufficient for the a.c.s. application.

A proof is given in Section 4.7.3.

4.4.5 Hermitian spectral localization and positivity

The distribution theorems describe asymptotic spectral averages. Here we obtain bounds for each finite Hermitian Toeplitz matrix and a uniform positive-definiteness criterion. Lin, Pan, and Ng [118] establish corresponding results for scalar unilevel right-quaternion Toeplitz matrices; the symplectic embedding extends the argument to the multilevel block setting.

Theorem 4.37 (Spectral localization of Hermitian Toeplitz matrices). Let $F \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times s})$, fix $\tau \in \{\mathrm{L}, \mathrm{R}, (S_L, S_R)\}$, and let $G_\tau \in L^1(\mathbb{T}^d; \mathbb{C}^{2s\times 2s})$ be the corresponding complex symbol defined in Theorems 4.32 and 4.33. Assume that $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian for every $\boldsymbol{n} \in \mathbb{N}^d$. Set

$$m_\tau := \operatorname*{ess\,inf}_{\boldsymbol{\theta}\in\mathbb{T}^d} \lambda_{\min}\big(G_\tau(\boldsymbol{\theta})\big), \qquad M_\tau := \operatorname*{ess\,sup}_{\boldsymbol{\theta}\in\mathbb{T}^d} \lambda_{\max}\big(G_\tau(\boldsymbol{\theta})\big).$$

The endpoints are understood in the extended-real sense; in particular, $m_\tau \in [-\infty, \infty)$ and $M_\tau \in (-\infty, \infty]$. If $G_\tau \in L^\infty$, both endpoints are finite. Then, for every $\boldsymbol{n} \in \mathbb{N}^d$,

$$\operatorname{spec}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) \;\subset\; [\,m_\tau,\, M_\tau\,]. \tag{4.33}$$

A proof is given in Section 4.7.3.

The multilevel block positivity mechanism used here is the matrix-valued linear-positive-operator property in [155, Sec. 3.1, Prop. 3.1]. For related unilevel block Toeplitz spectral analyses, see [151, 146]. For Hermitian matrices A and B of the same size, $A \succeq B$ means that $A - B$ is positive semidefinite.

Corollary 4.38 (Positive-definiteness criterion). Let $F \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times s})$, and fix a left, right, or sandwich kernel τ. Let $G_\tau \in L^1(\mathbb{T}^d; \mathbb{C}^{2s\times 2s})$ be the corresponding complex symbol defined in Theorems 4.32 and 4.33. Assume that $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian for every $\boldsymbol{n} \in \mathbb{N}^d$, and set

$$m_\tau := \operatorname*{ess\,inf}_{\boldsymbol{\theta}\in\mathbb{T}^d} \lambda_{\min}\big(G_\tau(\boldsymbol{\theta})\big).$$

Then:

(a) If $m_\tau > 0$, then $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian positive definite for every $\boldsymbol{n}$, with $\lambda_{\min}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) \geq m_\tau$.

(b) Conversely, if there exists $c > 0$ such that $T_{\boldsymbol{n}}^{(\tau)}(F) \geq c\,I$ for all $\boldsymbol{n}$, then $m_\tau \geq c$.

A proof is given in Section 4.7.3.

The converse requires one positive lower bound valid for all matrix sizes, rather than positive definiteness at each size without a common bound.

4.4.6 Circulant approximation of Toeplitz sequences

This subsection constructs circulant approximations of right quaternion Toeplitz sequences and uses them to recover the singular-value distribution through a.c.s. theory. Circulant approximations are classical in Toeplitz analysis and underlie many fast preconditioners; see [131]. A quaternion Strang-type preconditioner is analyzed in [118]. The simple spectral structure of circulants also provides an alternative route to distribution results for normal GLT sequences [14].

Following Pan and Ng [137], we apply the quaternion discrete Fourier transform (QDFT) on the slice $\mathbb{C}_{\mathbf{i}}$ to multilevel $s \times t$ block circulants. The components in $\mathbb{C}_{\mathbf{i}}$ are diagonalized frequency by frequency, whereas the components proportional to $\mathbf{j}$ or $\mathbf{ij}$ couple the frequencies $\boldsymbol{k}$ and $-\boldsymbol{k}$. Grouping fixed frequencies and paired frequencies by permutations yields blocks of size $s \times t$ and $2s \times 2t$. The argument proceeds in three stages: Fourier decomposition, spectral sampling of the circulants, and approximation of right Toeplitz sequences.

Fourier decomposition of block circulants

Let A_n and $A_{\boldsymbol{n}}$ be the modular-reversal permutations of Lemma 2.33. The single-axis QDFT is the complex DFT of Proposition 2.32, interpreted in $\mathbb{C}_{\mathbf{i}}$ by replacing i with $\mathbf{i}$ [64]. We denote its one- and multilevel matrices by $F_{\mathbf{i},n}$ and $F_{\mathbf{i}}^{(\boldsymbol{n})} = F_{\mathbf{i},n_1} \otimes \cdots \otimes F_{\mathbf{i},n_d}$, respectively. In particular, $F_{\mathbf{i},n}$ uses the negative exponential and the unitary normalization $n^{-1/2}$; multilevel indices retain the lexicographic ordering of Chapter 2.

For block sizes $s, t \in \mathbb{N}$, set

$$U_L \coloneqq F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_s, \qquad U_R \coloneqq F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_t.$$

Because every single-axis sandwich construction reduces to a right-kernel form, it is sufficient to formulate the circulant approximation for right trigonometric polynomials.

Polynomial circulants. Use the periodic shifts $P_{\boldsymbol{n}}^{(\boldsymbol{\rho})}$ and the grid $\Lambda_{\boldsymbol{n}}$ from Definition 2.31. The shifts are real matrices and therefore act unchanged on quaternion blocks.

The shift representation gives the following definition of a multilevel block circulant; an equivalent entrywise formula is included for later use.

Definition 4.39 (d-level $s\times t$ quaternion block-circulant generated by a right polynomial). Let p_{R} be a right trigonometric polynomial with block coefficients $\{p_{\mathrm{R}}(\boldsymbol{\rho})\}_{\boldsymbol{\rho}\in\mathcal{F}} \subset \mathbb{H}^{s\times t}$. The associated d-level $s \times t$ block-circulant matrix is

$$C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big) \coloneqq \sum_{\boldsymbol{\rho}\in\mathcal{F}} \big(P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})\big) \;\in\; \mathbb{H}^{(N_{\boldsymbol{n}}s)\times(N_{\boldsymbol{n}}t)}, \qquad N_{\boldsymbol{n}} = \prod_{\ell=1}^{d} n_\ell.$$

Equivalently, its blocks are

$$\big[C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big)\big]_{\alpha,\beta} = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\mathrm{R}}(\boldsymbol{\rho})\,\mathbf{1}_{\{\alpha\equiv\beta+\boldsymbol{\rho} \pmod{\boldsymbol{n}}\}}.$$

For a polynomial $p : \mathbb{T}^d \to \mathbb{C}_{\mathbf{i}}^{s\times t}$, identify the slice $\mathbb{C}_{\mathbf{i}}$ with $\mathbb{C}$. Definition 4.39 uses the negative-exponent polynomial convention of Definition 4.13, whereas Definition 2.34 uses positive exponents. For the same function p, these conventions therefore satisfy

$$C_{\boldsymbol{n}}^{\text{here}}(p) = C_{\boldsymbol{n}}^{\text{toolbox}}\big(p(-\cdot)\big),$$

where the superscripts distinguish the two definitions only in this comparison. The QDFT factorization follows the splitting argument of Pan and Ng [137]; the multilevel block extension uses the same pairing of the frequencies $\boldsymbol{k}$ and $-\boldsymbol{k}$. Zheng and Ni [176] obtain a related full-block diagonalization for unilevel block circulants through an octonionic representation.

Write each polynomial coefficient in Cartesian form $p_{\mathrm{R}}(\boldsymbol{\rho}) = p_{\mathrm{R}}(\boldsymbol{\rho})^{(0)} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(1)}\mathbf{i} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\mathbf{j} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(3)}(\mathbf{ij})$ with $p_{\mathrm{R}}(\boldsymbol{\rho})^{(\ell)} \in \mathbb{R}^{s\times t}$. By expanding the shift representation, we obtain

$$\begin{aligned} C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} \big(P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(0)}\big) \;+\; \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(1)}\Big)\mathbf{i} \\ +\; \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\Big)\mathbf{j} \;+\; \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(3)}\Big)(\mathbf{ij}). \end{aligned}$$

Thus $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ is the sum of four real multilevel block circulants multiplied on the right by 1, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{ij}$. The QDFT diagonalizes the first two components. The last two components couple opposite frequencies and produce the multilevel X-shaped pattern of [137]. The following identities make this coupling explicit.

The unitarity, symmetry, and reversal identities of Lemma 2.33 apply unchanged on $\mathbb{C}_{\mathbf{i}}$. Their interaction with $\mathbf{j}$ gives the quaternion-specific flip relations below.

Lemma 4.40 (QDFT flip identities). Let $N_{\boldsymbol{n}} := \prod_{\ell=1}^{d} n_\ell$. Then

$$F_{\mathbf{i}}^{(\boldsymbol{n})}\,(\mathbf{j}I_{N_{\boldsymbol{n}}})\,\big(F_{\mathbf{i}}^{(\boldsymbol{n})}\big)^{*} = A_{\boldsymbol{n}}\,(\mathbf{j}I_{N_{\boldsymbol{n}}}), \tag{4.34}$$

and

$$F_{\mathbf{i}}^{(\boldsymbol{n})}\,((\mathbf{ij})I_{N_{\boldsymbol{n}}})\,\big(F_{\mathbf{i}}^{(\boldsymbol{n})}\big)^{*} = A_{\boldsymbol{n}}\,((\mathbf{ij})I_{N_{\boldsymbol{n}}}). \tag{4.35}$$

The complex DFT diagonalization used below is Proposition 2.32 in Chapter 2. To expose the block-diagonal pattern, we group frequencies that equal their own modular negatives separately from the remaining pairs $\{\boldsymbol{k}, -\boldsymbol{k}\}$. The following definitions make this ordering precise and extend that of [137].

Definition 4.41. Let $d \in \mathbb{N}$, let $\boldsymbol{n} = (n_1, \dots, n_d) \in \mathbb{N}^d$, and let $\Lambda_{\boldsymbol{n}}$ be the index set defined in (4.14). If $\boldsymbol{k} = (k_1, \dots, k_d) \in \Lambda_{\boldsymbol{n}}$, set $-\boldsymbol{k} := ((-k_1) \bmod n_1, \dots, (-k_d) \bmod n_d)$. Define the fixed-point set

$$\mathrm{Fix} := \{\boldsymbol{k} \in \Lambda_{\boldsymbol{n}} :\ \boldsymbol{k} \equiv -\boldsymbol{k} \pmod{\boldsymbol{n}}\} = \big\{\boldsymbol{k} \in \Lambda_{\boldsymbol{n}} :\ 2k_\ell \equiv 0 \pmod{n_\ell}\ \forall\, \ell\big\}.$$

Definition 4.42. We denote the standard lexicographic order on $\Lambda_{\boldsymbol{n}}$ by $\prec$, and set

$$\mathcal{K} := \{\boldsymbol{k} \in \Lambda_{\boldsymbol{n}} \setminus \mathrm{Fix} :\ \boldsymbol{k} \prec -\boldsymbol{k}\}.$$

Set $f := \#\mathrm{Fix}$ and $q := \#\mathcal{K}$, and enumerate the two sets in lexicographic order as

$$\mathrm{Fix} = \{\boldsymbol{k}^{(1)}, \dots, \boldsymbol{k}^{(f)}\}, \qquad \mathcal{K} = \{\boldsymbol{k}_1, \dots, \boldsymbol{k}_q\}.$$

Define the complete ordered list

$$\mathcal{L} = \big(\boldsymbol{k}^{(1)}, \dots, \boldsymbol{k}^{(f)}, \boldsymbol{k}_1, -\boldsymbol{k}_1, \dots, \boldsymbol{k}_q, -\boldsymbol{k}_q\big) =: (\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_{N_{\boldsymbol{n}}}).$$

If $e_{\boldsymbol{k}}$ denotes the standard basis vector indexed by $\boldsymbol{k} \in \Lambda_{\boldsymbol{n}}$ in the original lexicographic ordering, let P be the unique permutation matrix satisfying

$$P e_{\boldsymbol{\ell}_j} = e_j, \qquad j = 1, \dots, N_{\boldsymbol{n}}.$$

Thus P maps the original lexicographic ordering to $\mathcal{L}$. For the multilevel $s \times t$ block case, define $\Pi_L := P \otimes I_s$ and $\Pi_R := P \otimes I_t$.

The construction gives the identities:

Lemma 4.43. Let $A_{\boldsymbol{n}} = \bigotimes_{\ell=1}^{d} A_{n_\ell}$ be the multilevel exchange permutation as in Lemma 2.33. With P from Definition 4.42 and $\mathcal{K}$ as above,

$$P\, A_{\boldsymbol{n}}\, P^{*} = I_{\#\mathrm{Fix}} \;\oplus\; \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Consequently,

$$\Pi_R\,(A_{\boldsymbol{n}} \otimes I_t)\,\Pi_R^{*} = I_{\#\mathrm{Fix}} \otimes I_t \;\oplus\; \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} 0 & I_t \\ I_t & 0 \end{bmatrix}.$$

The preceding identities yield the canonical QDFT decomposition of multilevel block-circulant matrices.

Proposition 4.44 (Canonical QDFT form for multilevel rectangular block circulants). Let $U_L := F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_s$ and $U_R := F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_t$. Define, for $\ell = 0, 1, 2, 3$,

$$\Lambda_\ell := \mathrm{Bdiag}\big(\widehat{S}_\ell(\boldsymbol{k})\big)_{\boldsymbol{k}\in\Lambda_{\boldsymbol{n}}},$$

where

$$\widehat{S}_\ell(\boldsymbol{k}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\mathrm{R}}(\boldsymbol{\rho})^{(\ell)}\, e^{-2\pi\mathbf{i}\sum_{\ell'=1}^{d} k_{\ell'}\rho_{\ell'}/n_{\ell'}}.$$

Then

$$U_L\, C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big)\, U_R^{*} = \Lambda_0 + \Lambda_1\,\mathbf{i} + \Lambda_2\,(A_{\boldsymbol{n}} \otimes I_t)\,\mathbf{j} + \Lambda_3\,(A_{\boldsymbol{n}} \otimes I_t)\,(\mathbf{ij}). \tag{4.36}$$

Let P be the permutation defined in Definition 4.42, and set $\Pi_L := P \otimes I_s$ and $\Pi_R := P \otimes I_t$. Applying Π_L to the block rows and Π_R^{*} to the block columns gives

$$\Pi_L U_L\, C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big)\, U_R^{*}\Pi_R^{*} = \bigoplus_{\boldsymbol{k}\in\mathrm{Fix}} \big(D_1(\boldsymbol{k}) + D_2(\boldsymbol{k})\big) \;\oplus\; \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} D_1(\boldsymbol{k}) & D_2(\boldsymbol{k}) \\ D_2(-\boldsymbol{k}) & D_1(-\boldsymbol{k}) \end{bmatrix},$$

where

$$D_1(\boldsymbol{k}) := \widehat{S}_0(\boldsymbol{k}) + \widehat{S}_1(\boldsymbol{k})\,\mathbf{i}, \qquad D_2(\boldsymbol{k}) := \big(\widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\,\mathbf{i}\big)\,\mathbf{j}.$$

A proof is given in Section 4.7.3.

Spectral sampling of circulants

We now identify the canonical eigenvalues, in the square case, and the singular values of the circulants through samples of their embedded symbols. Let $p_{\mathrm{R}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\mathrm{R}}(\boldsymbol{\rho}) e^{-\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}$ be an $s \times t$ trigonometric polynomial with finite $\mathcal{F} \subset \mathbb{Z}^d$. Let $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ be the block circulant of Definition 4.39, and let U_L, U_R, and $A_{\boldsymbol{n}}$ be as in Proposition 4.44. The small blocks associated with fixed frequencies and frequency pairs are called *fibers*. Their symplectic representations, with the indicated permutations, give the sampling relation for the embedded circulant symbol. For $\boldsymbol{k} \in \Lambda_{\boldsymbol{n}}$, set $\boldsymbol{\theta}_{\boldsymbol{k}} = 2\pi(\boldsymbol{k}/\boldsymbol{n})$. Periodicity gives $-\boldsymbol{\theta}_{\boldsymbol{k}} \equiv \boldsymbol{\theta}_{-\boldsymbol{k}} \pmod{2\pi}$. Write

$$p_{\mathrm{R}}(\boldsymbol{\rho}) = z_{\boldsymbol{\rho}} + w_{\boldsymbol{\rho}}\mathbf{j},$$

where

$$z_{\boldsymbol{\rho}} := p_{\mathrm{R}}(\boldsymbol{\rho})^{(0)} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(1)}\mathbf{i}, \qquad w_{\boldsymbol{\rho}} := p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(3)}\mathbf{i},$$

and $p_{\mathrm{R}}(\boldsymbol{\rho})^{(\ell)} \in \mathbb{R}^{s\times t}$. Define

$$Z(\boldsymbol{\theta}) := \sum_{\boldsymbol{\rho}\in\mathcal{F}} z_{\boldsymbol{\rho}} e^{-\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}, \qquad W(\boldsymbol{\theta}) := \sum_{\boldsymbol{\rho}\in\mathcal{F}} w_{\boldsymbol{\rho}} e^{+\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}. \tag{4.37}$$

The identity $\mathbf{j}e^{-\mathbf{i}t} = e^{+\mathbf{i}t}\mathbf{j}$ implies

$$p_{\mathrm{R}}(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}, \qquad W(-\boldsymbol{\theta}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} w_{\boldsymbol{\rho}} e^{-\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}.$$

Thus $W(-\boldsymbol{\theta})$ has the same negative Fourier phase as $Z(\boldsymbol{\theta})$, which is the form occurring in the embedded right symbol. For any right symbol $H = Z_H + W_H\mathbf{j}$, we write

$$G_{\mathrm{R}}[H](\boldsymbol{\theta}) := \begin{bmatrix} Z_H(\boldsymbol{\theta}) & W_H(-\boldsymbol{\theta}) \\ -\overline{W_H(\boldsymbol{\theta})} & \overline{Z_H(-\boldsymbol{\theta})} \end{bmatrix}. \tag{4.38}$$

The following proposition makes this sampling relation explicit in the grouped realization $\Phi_{\mathrm{grp}} : \mathbb{H}^{s\times t} \to \mathbb{C}^{2s\times 2t}$.

Proposition 4.45. Let $p_{\mathrm{R}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\mathrm{R}}(\boldsymbol{\rho})\, e^{-\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}$ be an $s\times t$ trigonometric polynomial with finite $\mathcal{F}\subset\mathbb{Z}^d$ and coefficients $p_{\mathrm{R}}(\boldsymbol{\rho})\in\mathbb{H}^{s\times t}$, and let Z and W be defined by (4.37). Let $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ be the associated d-level $s\times t$ block-circulant. The QDFT canonical form of $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ is

$$\Pi_L\, U_L\, C_{\boldsymbol{n}}(p_{\mathrm{R}})\, U_R^*\, \Pi_R^* = \bigoplus_{\boldsymbol{k}\in\mathrm{Fix}} \big(Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) \;\oplus\; \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix}. \tag{4.39}$$

For $\boldsymbol{k}\in\mathcal{K}$, the symplectic representation of the paired block is

$$\Phi_{\mathrm{blk}}^{(s,t)}\left(\begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix}\right) = \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & 0 & 0 & W(-\boldsymbol{\theta}_{\boldsymbol{k}}) \\ 0 & \overline{Z(\boldsymbol{\theta}_{\boldsymbol{k}})} & -\overline{W(-\boldsymbol{\theta}_{\boldsymbol{k}})} & 0 \\ 0 & W(-\boldsymbol{\theta}_{-\boldsymbol{k}}) & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) & 0 \\ -\overline{W(-\boldsymbol{\theta}_{-\boldsymbol{k}})} & 0 & 0 & \overline{Z(\boldsymbol{\theta}_{-\boldsymbol{k}})} \end{bmatrix}, \tag{4.40}$$

when $\boldsymbol{k}\in\mathcal{K}$, while for $\boldsymbol{k}\in\mathrm{Fix}$ we have

$$\Phi_{\mathrm{grp}}\big(Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) = \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}}) \\ -\overline{W(-\boldsymbol{\theta}_{\boldsymbol{k}})} & \overline{Z(\boldsymbol{\theta}_{\boldsymbol{k}})} \end{bmatrix}. \tag{4.41}$$

A proof is given in Section 4.7.3.

The resulting singular-value and eigenvalue distributions are stated next.

Corollary 4.46 (Asymptotic distributions for right polynomial quaternion circulants). Let

$$p_{\mathrm{R}}(\boldsymbol{\theta}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} p_{\mathrm{R}}(\boldsymbol{\rho})\, e^{-\mathbf{i}\langle\boldsymbol{\rho},\boldsymbol{\theta}\rangle}, \qquad \mathcal{F}\subset\mathbb{Z}^d \text{ finite}, \quad p_{\mathrm{R}}(\boldsymbol{\rho})\in\mathbb{H}^{s\times t}.$$

Let Z and W be defined by (4.37), and let $G_{\mathrm{R}}[p_{\mathrm{R}}]$ be the associated complex symbol from (4.38). Let $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ be the corresponding d-level $s\times t$ block-circulant. Then

$$\{C_{\boldsymbol{n}}(p_{\mathrm{R}})\}_{\boldsymbol{n}} \;\sim_\sigma\; G_{\mathrm{R}}[p_{\mathrm{R}}].$$

If $s = t$, then also

$$\{C_{\boldsymbol{n}}(p_{\mathrm{R}})\}_{\boldsymbol{n}} \;\sim_\lambda\; G_{\mathrm{R}}[p_{\mathrm{R}}].$$

A proof is given in Section 4.7.3.

Toeplitz approximation by circulants

The polynomial circulant distributions now provide approximants for integrable right Toeplitz generators. The construction combines polynomial approximation of the generator with a low-rank comparison between the polynomial Toeplitz matrices and their periodic counterparts.

Theorem 4.47 (Right Toeplitz sequences admit a circulant a.c.s.). Let $d \in \mathbb{N}$, $s, t \in \mathbb{N}$, and $F \in L^1(\mathbb{T}^d; \mathbb{H}^{s\times t})$. Consider the right Toeplitz family

$$\{T_{\boldsymbol{n}}^{(\mathrm{R})}(F)\}_{\boldsymbol{n}\in\mathbb{N}^d}.$$

Then there exist finite sets $\mathcal{S}_m \subset \mathbb{Z}^d$ and coefficients $A_{m,\boldsymbol{k}} \in \mathbb{H}^{s\times t}$ defining right circulants

$$B_{\boldsymbol{n},m} := \sum_{\boldsymbol{k}\in\mathcal{S}_m} P_{\boldsymbol{n}}^{(-\boldsymbol{k})} \otimes A_{m,\boldsymbol{k}}$$

such that, for every scalar-indexed sequence $\boldsymbol{n} = \boldsymbol{n}(n) \to \infty$,

$$\left\{\{B_{\boldsymbol{n}(n),m}\}_n\right\}_{m\in\mathbb{N}}$$

is an a.c.s. for $\{T_{\boldsymbol{n}(n)}^{(\mathrm{R})}(F)\}_n$. The external approximation index $m \in \mathbb{N}$ remains scalar.

A proof is given in Section 4.7.3.

Alternative a.c.s.-based derivation. The circulant approximation gives a second proof of the singular-value assertion in Theorem 4.34. The Hermitian eigenvalue assertion follows from the embedding proof in Subsection 4.4.3.

A proof is given in Section 4.7.3.

4.5 Numerical results

The experiments reported in [112] illustrate Theorem 4.34 for two-level quaternion Toeplitz matrices with 2×2 blocks. For Hermitian Toeplitz matrices, the ordered eigenvalues are compared with samples of the monotone rearrangements of the embedded-symbol eigenvalue functions. For non-Hermitian matrices, the comparison uses singular values. This is the quantile representation of Weyl distributions; see [79]. The relation with L^1-convergence of quantile functions is discussed, under the support and regularity hypotheses of the respective results, in [23, 13].

The multilevel dimensions are $\boldsymbol{n} = (n, n)$, with $n = 2, 8, 32, 64$. Thus the quaternion matrices have order $2n^2$, and their complex embeddings have order $4n^2$. We first give the common embedded-symbol formulas, then present each generator together with its slice decomposition and numerical comparison.

4.5.1 Embedded symbols and comparison procedure

The four examples use the same slice decomposition and reflection conventions. Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ and write

$$F(\boldsymbol{\theta}) = Z(\boldsymbol{\theta}) + W(\boldsymbol{\theta})\mathbf{j}, \qquad Z(\boldsymbol{\theta}), W(\boldsymbol{\theta}) \in \mathbb{C}_{\mathbf{i}}^{2\times 2}.$$

For brevity, write

$$S_{12} := (S_L, S_R) = (\{1\}, \{2\}), \qquad S_{21} := (S_L, S_R) = (\{2\}, \{1\}).$$

For the four kernel classes used in the experiments, the general embedded symbol formula gives

$$G_{\mathrm{L}}(\boldsymbol{\theta}) = \begin{bmatrix} Z(\theta_1,\theta_2) & W(\theta_1,\theta_2) \\ -\overline{W(-\theta_1,-\theta_2)} & \overline{Z(-\theta_1,-\theta_2)} \end{bmatrix},$$

$$G_{\mathrm{R}}(\boldsymbol{\theta}) = \begin{bmatrix} Z(\theta_1,\theta_2) & W(-\theta_1,-\theta_2) \\ -\overline{W(\theta_1,\theta_2)} & \overline{Z(-\theta_1,-\theta_2)} \end{bmatrix},$$

$$G_{S_{12}}(\boldsymbol{\theta}) = \begin{bmatrix} Z(\theta_1,\theta_2) & W(\theta_1,-\theta_2) \\ -\overline{W(-\theta_1,\theta_2)} & \overline{Z(-\theta_1,-\theta_2)} \end{bmatrix},$$

$$G_{S_{21}}(\boldsymbol{\theta}) = \begin{bmatrix} Z(\theta_1,\theta_2) & W(-\theta_1,\theta_2) \\ -\overline{W(\theta_1,-\theta_2)} & \overline{Z(-\theta_1,-\theta_2)} \end{bmatrix}.$$

The two sandwich cases follow from

$$r_{\{1\}}(\theta_1,\theta_2) = (-\theta_1,\theta_2), \qquad r_{\{2\}}(\theta_1,\theta_2) = (\theta_1,-\theta_2).$$

For each example, substitution of its slice functions into these formulas gives the four complex symbols to be sampled. The figure markers represent ordered finite-matrix spectral values; the solid curves represent the corresponding rearranged-symbol values. The horizontal axis is the quantile coordinate, not a spatial coordinate of the generator. Within each figure, rows identify the kernel placement and columns identify n. The comparisons retain the normalization conventions of Section 4.3.

4.5.2 Continuous symbols

Continuous Hermitian Toeplitz example. We consider a smooth quaternion-valued generator whose associated Toeplitz matrices are Hermitian:

$$F(\theta_1,\theta_2) = \begin{bmatrix} 2+\cos\theta_1+\cos\theta_2 & \frac{1}{2}\big(\sin\theta_1\,\mathbf{i}+\sin\theta_2\,\mathbf{j}\big)+\frac{1}{4}\sin(\theta_1-\theta_2)\,\mathbf{k} \\ \frac{1}{2}\big(\sin\theta_2\,\mathbf{j}-\sin\theta_1\,\mathbf{i}\big)+\frac{1}{4}\sin(\theta_1-\theta_2)\,\mathbf{k} & 2+\cos\theta_1+\cos\theta_2 \end{bmatrix}.$$

Set

$$a(\boldsymbol{\theta}) := 2+\cos\theta_1+\cos\theta_2.$$

The slice decomposition of this generator is

$$Z_{\mathrm{cH}}(\boldsymbol{\theta}) = \begin{bmatrix} a(\boldsymbol{\theta}) & \frac{1}{2}\sin\theta_1\,\mathbf{i} \\ -\frac{1}{2}\sin\theta_1\,\mathbf{i} & a(\boldsymbol{\theta}) \end{bmatrix}.$$

$$W_{\mathrm{cH}}(\boldsymbol{\theta}) = \begin{bmatrix} 0 & \frac{1}{2}\sin\theta_2+\frac{1}{4}\sin(\theta_1-\theta_2)\mathbf{i} \\ \frac{1}{2}\sin\theta_2+\frac{1}{4}\sin(\theta_1-\theta_2)\mathbf{i} & 0 \end{bmatrix}.$$

Substitution of Z_{cH} and W_{cH} into the four formulas in Subsection 4.5.1 gives the embedded symbols for all four kernel classes. The slice functions satisfy Theorem 4.31, so the Toeplitz matrices are Hermitian for every placement.

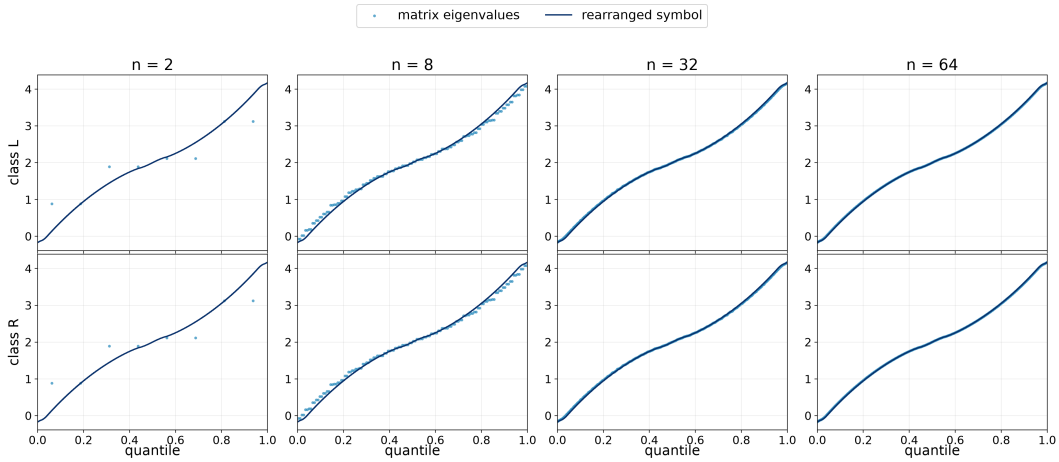


Figure 4.1: Continuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$.

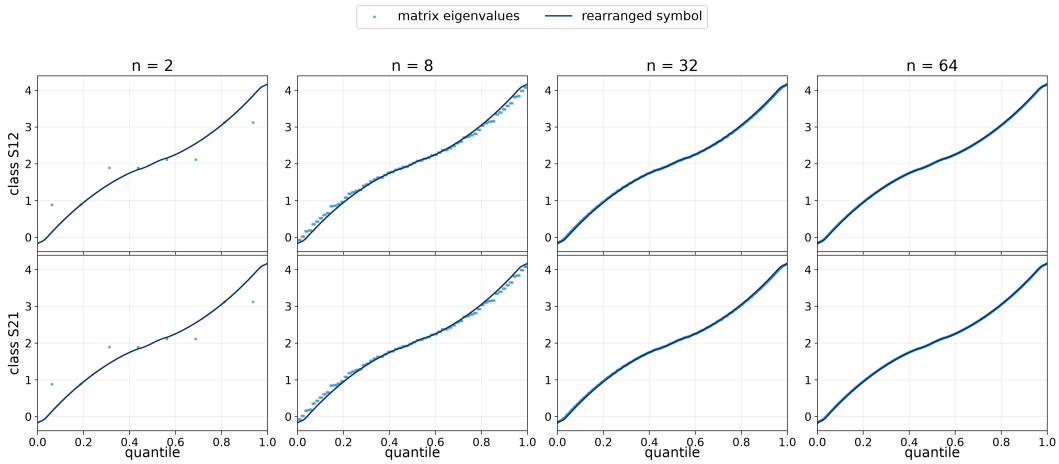


Figure 4.2: Continuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$.

Figures 4.1 and 4.2 show increasingly close agreement between the ordered eigenvalues and the rearranged-symbol curves. The coarse samples at the smallest size become dense along the curves as n grows. Each kernel placement is compared with its own embedded symbol.

Continuous non-Hermitian Toeplitz example. The second smooth generator produces non-Hermitian Toeplitz matrices:

$$F(\theta_1, \theta_2) = \begin{bmatrix} e^{\mathbf{i}\theta_1} + e^{\mathbf{j}\theta_2} & (\cos\theta_2 + \sin\theta_1)\,\mathbf{k} \\ (\cos\theta_1 - \sin\theta_2)\,\mathbf{i} & 1 + \frac{1}{2}\left(e^{\mathbf{j}\theta_1} - e^{\mathbf{k}\theta_2}\right) \end{bmatrix}.$$

The two quaternion exponentials that contribute to the second slice satisfy

$$e^{\mathbf{j}\theta_2} = \cos\theta_2 + \sin\theta_2\,\mathbf{j}, \qquad e^{\mathbf{k}\theta_2} = \cos\theta_2 + (\mathbf{i}\sin\theta_2)\mathbf{j}.$$

Consequently,

$$Z_{\mathrm{cN}}(\boldsymbol{\theta}) = \begin{bmatrix} e^{\mathbf{i}\theta_1} + \cos\theta_2 & 0 \\ (\cos\theta_1 - \sin\theta_2)\mathbf{i} & 1 + \frac{1}{2}(\cos\theta_1 - \cos\theta_2) \end{bmatrix},$$

$$W_{\mathrm{cN}}(\boldsymbol{\theta}) = \begin{bmatrix} \sin\theta_2 & (\cos\theta_2 + \sin\theta_1)\mathbf{i} \\ 0 & \frac{1}{2}(\sin\theta_1 - \mathbf{i}\sin\theta_2) \end{bmatrix}.$$

Substituting these components into the formulas in Subsection 4.5.1 gives the four embedded symbols. The appearances of $\mathbf{j}$ and $\mathbf{k}$ in the generator do not change the single-axis Fourier-kernel convention. The numerical comparison concerns the singular-value assertion, not the separate non-Hermitian eigenvalue theorem.

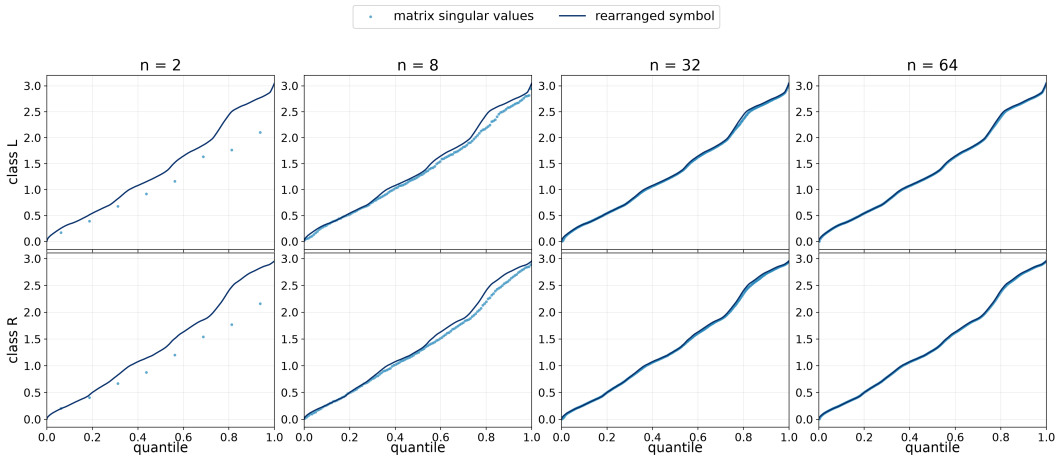


Figure 4.3: Continuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$.

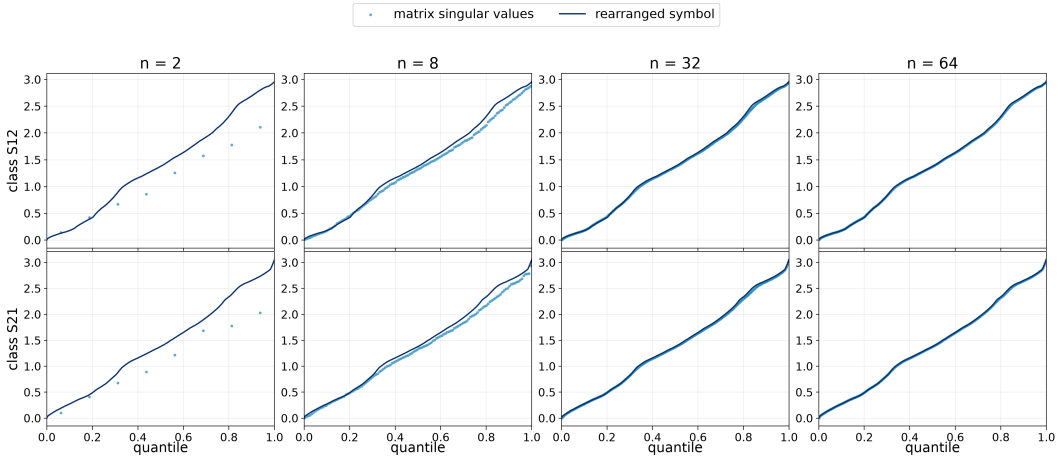


Figure 4.4: Continuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$.

Figures 4.3 and 4.4 show the analogous agreement for singular values. The finite-matrix values follow the shape of the corresponding rearranged symbols more closely at the larger sizes.

4.5.3 Discontinuous L^1 symbols

We next consider bounded discontinuous generators in the L^1 framework. Discontinuities are described by

$$\operatorname{sgn}(x) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases}$$

For both examples, set

$$s_j := \operatorname{sgn}(\theta_j), \qquad j = 1, 2.$$

Discontinuous Hermitian Toeplitz example. We use the following piecewise-constant generator of Hermitian Toeplitz matrices:

$$F(\theta_1, \theta_2) = \begin{bmatrix} 2 & \operatorname{sgn}(\theta_1)\,\mathbf{i} + \operatorname{sgn}(\theta_2)(\mathbf{j} + \mathbf{k}) \\ -\operatorname{sgn}(\theta_1)\,\mathbf{i} + \operatorname{sgn}(\theta_2)(\mathbf{j} + \mathbf{k}) & 2 \end{bmatrix}.$$

Since $\mathbf{j} + \mathbf{k} = (1 + \mathbf{i})\mathbf{j}$, we obtain

$$Z_{1\mathrm{H}}(\boldsymbol{\theta}) = \begin{bmatrix} 2 & s_1\mathbf{i} \\ -s_1\mathbf{i} & 2 \end{bmatrix}, \qquad W_{1\mathrm{H}}(\boldsymbol{\theta}) = \begin{bmatrix} 0 & s_2(1 + \mathbf{i}) \\ s_2(1 + \mathbf{i}) & 0 \end{bmatrix}.$$

These slice components satisfy the Hermitian criterion in Theorem 4.31. Their substitutions into the common formulas therefore yield Hermitian embedded symbols for all four kernel classes.

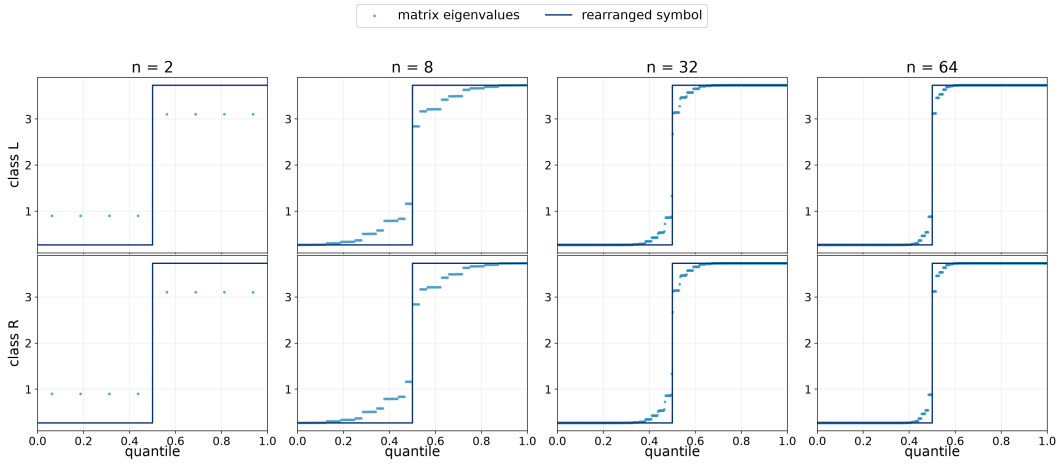


Figure 4.5: Discontinuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$.

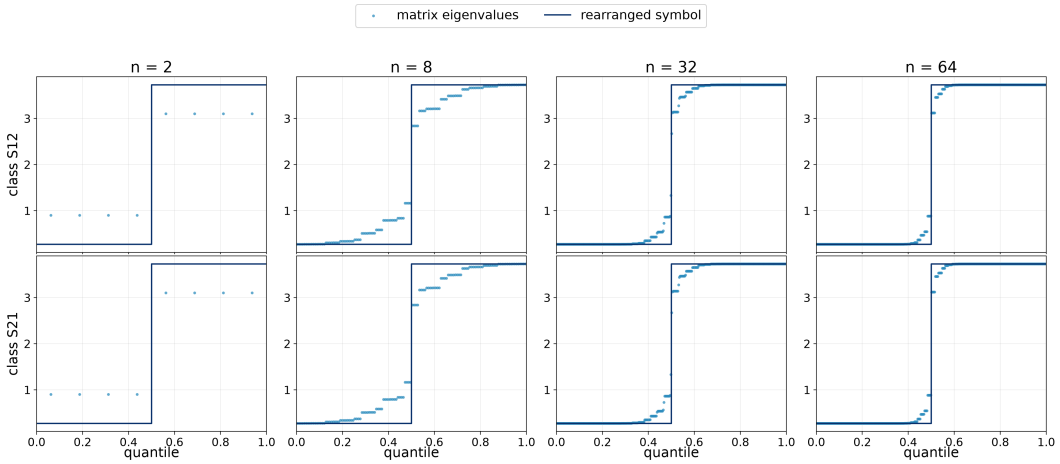


Figure 4.6: Discontinuous Hermitian Toeplitz example: ordered eigenvalues of the finite Toeplitz matrices and monotone rearrangements of the eigenvalue functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$.

In Figures 4.5 and 4.6, the rearranged-symbol curves have two plateaus separated by a jump. As n grows, the ordered eigenvalues accumulate near the plateaus, while the transition occupies a smaller portion of the quantile axis. This illustrates the distributional comparison rather than uniform convergence across the jump.

Discontinuous non-Hermitian Toeplitz example. The non-Hermitian discontinuous example uses the generator

$$F(\theta_1,\theta_2) = \begin{bmatrix} e^{\mathbf{i}\theta_1} + \operatorname{sgn}(\theta_1) + \mathbf{j}\,\operatorname{sgn}(\theta_2) & \mathbf{k}\,\operatorname{sgn}\big(\sin\theta_1 + \frac{1}{4}\cos\theta_2\big) \\ \mathbf{i}\,\operatorname{sgn}\big(\cos\theta_1 - \sin\theta_2\big) & 1 + \frac{1}{4}\big(e^{\mathbf{j}\,\operatorname{sgn}(\theta_1)\pi/2} - e^{\mathbf{k}\,\operatorname{sgn}(\theta_2)\pi/2}\big) \end{bmatrix}.$$

Define

$$h(\boldsymbol{\theta}) := \operatorname{sgn}\left(\sin\theta_1 + \frac{1}{4}\cos\theta_2\right), \qquad g(\boldsymbol{\theta}) := \operatorname{sgn}(\cos\theta_1 - \sin\theta_2),$$

$$c_j := \cos\left(\frac{\pi s_j}{2}\right), \qquad u_j := \sin\left(\frac{\pi s_j}{2}\right), \qquad j = 1,2.$$

The corresponding slice functions are

$$Z_{1\mathrm{N}}(\boldsymbol{\theta}) = \begin{bmatrix} e^{\mathbf{i}\theta_1} + s_1 & 0 \\ g(\boldsymbol{\theta})\mathbf{i} & 1 + \frac{1}{4}(c_1 - c_2) \end{bmatrix}, \qquad W_{1\mathrm{N}}(\boldsymbol{\theta}) = \begin{bmatrix} s_2 & h(\boldsymbol{\theta})\mathbf{i} \\ 0 & \frac{1}{4}(u_1 - \mathbf{i}u_2) \end{bmatrix}.$$

Away from the measure-zero sets defined by the equations $\theta_j = 0$, $j = 1,2$, one has $c_1 = c_2 = 0$ and $u_j = s_j$, $j = 1,2$. Thus the lower-right entries simplify almost everywhere to the quantities

$$(Z_{1\mathrm{N}})_{22} = 1, \qquad (W_{1\mathrm{N}})_{22} = \frac{1}{4}(s_1 - \mathbf{i}s_2).$$

The common reflected-argument formulas give the four embedded symbols. As in the continuous non-Hermitian example, the numerical comparison uses singular values.

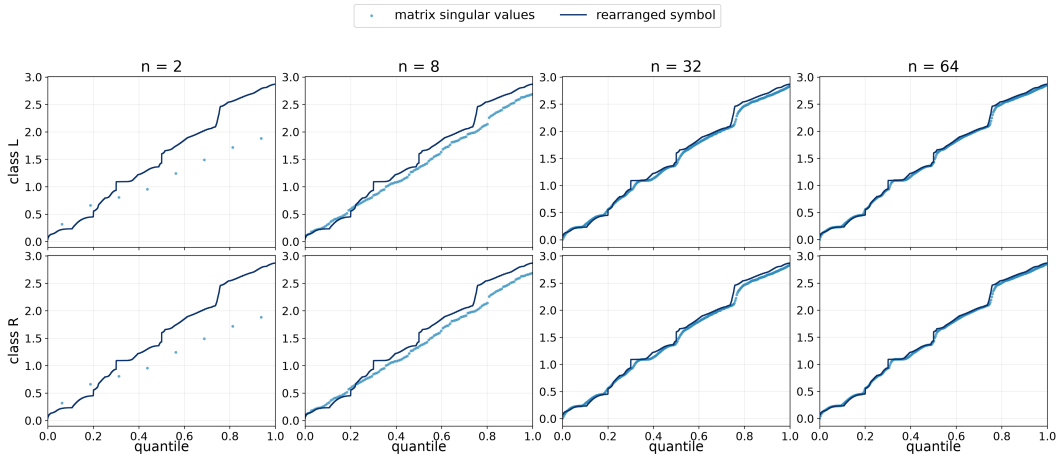


Figure 4.7: Discontinuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to the left and right kernels; columns correspond to $n = 2, 8, 32, 64$.

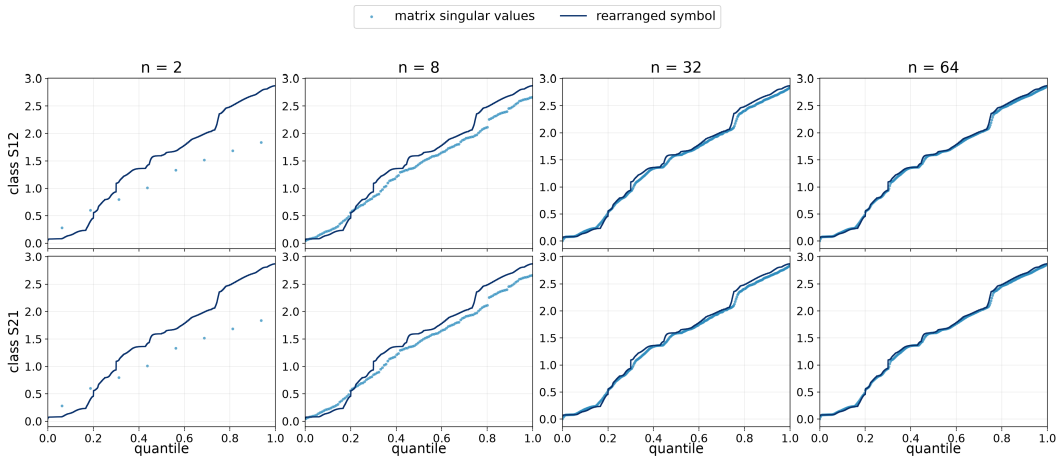


Figure 4.8: Discontinuous non-Hermitian Toeplitz example: ordered singular values of the finite Toeplitz matrices and monotone rearrangements of the singular-value functions of the embedded symbols. Rows correspond to S_{12} and S_{21}; columns correspond to $n = 2, 8, 32, 64$.

Figures 4.7 and 4.8 compare singular values for bounded discontinuous non-Hermitian data. The larger matrices resolve the changes in the rearranged-symbol curves more closely than the smallest matrices.

4.6 Conclusions

This chapter identifies explicit complex block Toeplitz symbols for the left, right, and sandwich single-axis quaternion constructions. The symplectic representation preserves the algebraic operations needed for the transfer, while the spectral normalizations account for doubled dimensions, duplicated singular values, and conjugate-paired eigenvalues.

For integrable generators, the embeddings yield singular-value distributions and, for Hermitian sequences, eigenvalue distributions. The non-Hermitian eigenvalue result requires boundedness and the stated Tilli-class hypotheses. The same framework gives Schatten norm estimates and finite-dimensional Hermitian localization and positivity criteria. Quaternion

a.c.s. theory, including the product-compatible rectangular extension, supports the circulant approximation of right Toeplitz sequences. The numerical comparisons illustrate the distributions for continuous and bounded discontinuous two-level generators; the supplementary calculations evaluate their embedded symbols at a representative point.

4.7 Supplementary calculations and proofs

The worked evaluations come first. Subsection 4.7.2 then proves the quaternion spectral and a.c.s. transfer steps, including canonical eigenvalue selection and recovery of quaternion approximation errors. Subsection 4.7.3 gives the Toeplitz and circulant proofs, applying the shared complex results at the indicated steps.

4.7.1 Worked embedded-symbol computations

We evaluate the four formulas from Subsection 4.5.1 at the interior point $\boldsymbol{\theta}_0 = (\pi/3, \pi/4)$, away from the discontinuity sets of the examples in Subsection 4.5.3. The slice decompositions are given beside the generators in Subsections 4.5.2 and 4.5.3. Each calculation uses Z, W at $\boldsymbol{\theta}_0$ and at the reflected arguments occurring in $G_{\mathrm{L}}, G_{\mathrm{R}}, G_{S_{12}}, G_{S_{21}}$. In particular,

$$F(\boldsymbol{\theta}_0) = Z(\boldsymbol{\theta}_0) + W(\boldsymbol{\theta}_0)\mathbf{j}, \qquad Z(\boldsymbol{\theta}_0), W(\boldsymbol{\theta}_0) \in \mathbb{C}_{\mathbf{i}}^{s\times t}.$$

Substituting these values into the four formulas in Subsection 4.5.1 gives the corresponding 4×4 complex matrices. These evaluations illustrate the pointwise spectral functions whose averages define the limiting distributions: eigenvalues in the Hermitian Toeplitz examples and singular values otherwise. Decimal matrix entries and spectral values below are rounded to four decimal places. The abbreviated family labels in the displays refer to the associated Toeplitz matrices, not to pointwise Hermitian symmetry of the quaternion generators.

Worked example: the continuous Hermitian Toeplitz case. For the continuous Hermitian Toeplitz example in Subsection 4.5.2, evaluation of its slice components at $\boldsymbol{\theta}_0$ gives

$$Z = \begin{bmatrix} 3.2071 & 0.4330\,\mathbf{i} \\ -0.4330\,\mathbf{i} & 3.2071 \end{bmatrix}, \qquad W = \begin{bmatrix} 0 & 0.3536 + 0.0647\,\mathbf{i} \\ 0.3536 + 0.0647\,\mathbf{i} & 0 \end{bmatrix},$$

and the four embedded 4×4 symbols are

$$G_{\mathrm{L}} = \begin{bmatrix} 3.2071 & 0.4330\,\mathbf{i} & 0 & 0.3536 + 0.0647\,\mathbf{i} \\ -0.4330\,\mathbf{i} & 3.2071 & 0.3536 + 0.0647\,\mathbf{i} & 0 \\ 0 & 0.3536 - 0.0647\,\mathbf{i} & 3.2071 & 0.4330\,\mathbf{i} \\ 0.3536 - 0.0647\,\mathbf{i} & 0 & -0.4330\,\mathbf{i} & 3.2071 \end{bmatrix},$$

$$G_{\mathrm{R}} = \begin{bmatrix} 3.2071 & 0.4330\,\mathbf{i} & 0 & -0.3536 - 0.0647\,\mathbf{i} \\ -0.4330\,\mathbf{i} & 3.2071 & -0.3536 - 0.0647\,\mathbf{i} & 0 \\ 0 & -0.3536 + 0.0647\,\mathbf{i} & 3.2071 & 0.4330\,\mathbf{i} \\ -0.3536 + 0.0647\,\mathbf{i} & 0 & -0.4330\,\mathbf{i} & 3.2071 \end{bmatrix}.$$

$$G_{S_{12}} = \begin{bmatrix} 3.2071 & 0.4330\,\mathbf{i} & 0 & -0.3536 + 0.2415\,\mathbf{i} \\ -0.4330\,\mathbf{i} & 3.2071 & -0.3536 + 0.2415\,\mathbf{i} & 0 \\ 0 & -0.3536 - 0.2415\,\mathbf{i} & 3.2071 & 0.4330\,\mathbf{i} \\ -0.3536 - 0.2415\,\mathbf{i} & 0 & -0.4330\,\mathbf{i} & 3.2071 \end{bmatrix},$$

$$G_{S_{21}} = \begin{bmatrix} 3.2071 & 0.4330\,\mathbf{i} & 0 & 0.3536 - 0.2415\,\mathbf{i} \\ -0.4330\,\mathbf{i} & 3.2071 & 0.3536 - 0.2415\,\mathbf{i} & 0 \\ 0 & 0.3536 + 0.2415\,\mathbf{i} & 3.2071 & 0.4330\,\mathbf{i} \\ 0.3536 + 0.2415\,\mathbf{i} & 0 & -0.4330\,\mathbf{i} & 3.2071 \end{bmatrix}.$$

The embedded Hermitian symbols have eigenvalues $\{2.6444, 2.6444, 3.7699, 3.7699\}$ for $\tau \in \{\mathrm{L}, \mathrm{R}\}$ and $\{2.5982, 2.5982, 3.8161, 3.8161\}$ for $\tau \in \{S_{12}, S_{21}\}$. These pointwise values illustrate the eigenvalue functions underlying the rearranged-symbol curves in Figures 4.1 and 4.2.

The remaining symbols. Evaluating the remaining slice decompositions from Subsections 4.5.2 and 4.5.3 at $\boldsymbol{\theta}_0$ gives

$$\text{(cont. non-Herm.)} \quad Z = \begin{bmatrix} 1.2071 + 0.8660\,\mathbf{i} & 0 \\ -0.2071\,\mathbf{i} & 0.8964 \end{bmatrix},\ W = \begin{bmatrix} 0.7071 & 1.5731\,\mathbf{i} \\ 0 & 0.4330 - 0.3536\,\mathbf{i} \end{bmatrix},$$

$$(L^1 \text{ Herm.}) \quad Z = \begin{bmatrix} 2 & \mathbf{i} \\ -\mathbf{i} & 2 \end{bmatrix},\ W = \begin{bmatrix} 0 & 1+\mathbf{i} \\ 1+\mathbf{i} & 0 \end{bmatrix};$$

$$(L^1 \text{ non-Herm.}) \quad Z = \begin{bmatrix} 1.5000 + 0.8660\,\mathbf{i} & 0 \\ -\mathbf{i} & 1 \end{bmatrix},\ W = \begin{bmatrix} 1 & \mathbf{i} \\ 0 & 0.2500 - 0.2500\,\mathbf{i} \end{bmatrix}.$$

The four G_τ are then obtained from the reflected-argument formulas in Subsection 4.5.1. Table 4.1 collects the eigenvalues of the Hermitian embedded symbols and the singular values in the other cases. Its family labels refer to these embedded symbols.

Table 4.1: Eigenvalues for the Hermitian embedded symbols and singular values for the non-Hermitian embedded symbols at $\boldsymbol{\theta}_0 = (\pi/3, \pi/4)$.

Symbol	Kernel class τ	Spectral data at $\boldsymbol{\theta}_0$
Continuous Hermitian	L, R	2.6444, 2.6444, 3.7699, 3.7699
	S_{12}, S_{21}	2.5982, 2.5982, 3.8161, 3.8161
Continuous non-Hermitian	L	2.8029, 1.6003, 1.0385, 0.3881
	R	2.6603, 1.7932, 1.1078, 0.3560
	S_{12}	2.5117, 2.0536, 1.0360, 0.2180
	S_{21}	2.5820, 1.7950, 1.3230, 0.0840
L^1 Hermitian	L, R, S_{12}, S_{21}	0.2679, 0.2679, 3.7321, 3.7321
L^1 non-Hermitian	L, R	2.6670, 1.9417, 1.0932, 0.4147
	S_{12}, S_{21}	2.6204, 1.8711, 1.3568, 0.2034

4.7.2 Proofs of preliminary and a.c.s. results

The first group proves the spectral and Fourier preliminaries, including the measurability lemma used in the distribution definitions. The second group proves the embedding equivalence and the product rules for approximating classes.

Preliminary tools

Lemma 4.48 (Borel enumeration of canonical quaternion eigenvalues). For every $r \in \mathbb{N}$, there exist Borel maps

$$\lambda_1^+, \dots, \lambda_r^+ : \mathbb{H}^{r\times r} \longrightarrow \{z \in \mathbb{C} : \operatorname{Im} z \geq 0\}$$

whose values enumerate the canonical eigenvalues with multiplicity.

Proof. Let $\mu_1, \dots, \mu_{2r}$ be the complex Borel maps of Lemma 2.2. Only the quaternion conjugate-pairing and the selection of canonical representatives remain to be checked. Now let $A \in \mathbb{H}^{r\times r}$, set $B := \Phi_{\mathrm{grp}}(A)$, and define

$$\rho(z) := \operatorname{Re} z + \mathbf{i}|\operatorname{Im} z|.$$

Let

$$\eta_1(A) \preceq \cdots \preceq \eta_{2r}(A)$$

be the lexicographically ordered values $\rho(\mu_1(B)), \dots, \rho(\mu_{2r}(B))$. Since finite lexicographic sorting is Borel, the maps η_j are Borel. The next symmetry argument follows from the paired Jordan structure; see [120, Thm. 3.2]. The exact conjugate-pairing and real-parity formulation can be found in [94, Lem. 3.2]. For

$$J_r := \begin{bmatrix} 0 & I_r \\ -I_r & 0 \end{bmatrix},$$

the block form of B gives $BJ_r = J_r\overline{B}$. Hence the conjugate-linear map

$$T(v) := J_r\overline{v}$$

satisfies $T^2 = -I$ and

$$(B - \overline{\lambda}I)^k T = T(B - \lambda I)^k, \qquad k \geq 1.$$

Thus T maps the generalized eigenspace of λ bijectively onto that of $\overline{\lambda}$, so conjugate eigenvalues have equal algebraic multiplicity. For real λ, the generalized eigenspace has even dimension d in the complex field: in a basis, the restriction of T is $v \mapsto C\overline{v}$, so $C\overline{C} = -I_d$ and $|\det C|^2 = (-1)^d = 1$. Therefore every distinct value in the folded list occurs with even multiplicity. Since each preceding constant block also has even length, the entries in the even positions retain exactly half of every block. Consequently,

$$\lambda_j^{+}(A) := \eta_{2j}(A), \qquad j = 1, \dots, r,$$

is a Borel enumeration of the canonical eigenvalues. □

Proof of Lemma 4.10. The field $F = \Phi_{\mathrm{ent}} \circ f$ is measurable. Lemma 2.2 and the duplication identities

$$\|f(x)\|_{S_p} = 2^{-1/p}\|F(x)\|_{S_p}, \qquad \sigma_j(f(x)) = \sigma_{2j-1}(F(x)) = \sigma_{2j}(F(x))$$

give the norm and singular-value assertions, with $2^{-1/\infty} = 1$. For canonical eigenvalues, compose the Borel maps in Lemma 4.48 with f, then with g. Symmetry of g makes the value independent of the enumeration. □

Proof of Lemma 4.15. Use $\mathbf{j}\alpha = \overline{\alpha}\,\mathbf{j}$ for all $\alpha \in \mathbb{C}_{\mathbf{i}}$ and the identities $\mathbf{j}e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle} = e^{+\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle}\mathbf{j}$, $e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle}\mathbf{j} = \mathbf{j}e^{+\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle}$, to rewrite

$$L_{S_L}(\boldsymbol{\theta},\boldsymbol{m})\,F(\boldsymbol{\theta})\,R_{S_R}(\boldsymbol{\theta},\boldsymbol{m}) = Z(\boldsymbol{\theta})\,e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle} + W(\boldsymbol{\theta})\,e^{-\mathbf{i}\langle \tilde{\boldsymbol{m}},\boldsymbol{\theta}\rangle}\,\mathbf{j}.$$

Then apply Φ_{grp} and integrate entrywise. □

Proof of Proposition 4.16. Write $F = Z + W\mathbf{j}$, where $Z, W \in L^p(\mathbb{T}^d; \mathbb{C}_{\mathbf{i}}^{m\times n})$. By the density of complex matrix-valued trigonometric polynomials, choose sequences

$$P_N(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}_N} Z_{\boldsymbol{m},N} e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle}, \qquad Q_N(\boldsymbol{\theta}) = \sum_{\boldsymbol{m}\in\mathcal{F}_N} W_{\boldsymbol{m},N} e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle},$$

converging to Z and W, respectively, in L^p. We may take the finite support $\mathcal{F}_N$ invariant under the involution $\boldsymbol{m} \mapsto \widetilde{\boldsymbol{m}}$ from Lemma 4.15, setting missing coefficients to zero. Define

$$A_{\boldsymbol{m},N} := Z_{\boldsymbol{m},N} + W_{\widetilde{\boldsymbol{m}},N}\mathbf{j}.$$

Using $\mathbf{j}e^{-\mathbf{i}b} = e^{\mathbf{i}b}\mathbf{j}$, we obtain

$$\sum_{\boldsymbol{m}\in\mathcal{F}_N} L_{S_L}(\boldsymbol{\theta},\boldsymbol{m}) A_{\boldsymbol{m},N} R_{S_R}(\boldsymbol{\theta},\boldsymbol{m}) = P_N(\boldsymbol{\theta}) + Q_N(\boldsymbol{\theta})\mathbf{j}.$$

Thus each pair of complex trigonometric approximants recombines as a sandwich trigonometric polynomial. The slice decomposition is a continuous real-linear isomorphism, so convergence of P_N and Q_N gives convergence to F in $L^p(\mathbb{T}^d; \mathbb{H}^{m\times n})$. □

Proof of Proposition 4.17. Fix $\boldsymbol{m} \in \mathbb{Z}^d$ and write $A_{\boldsymbol{m}} = Z_{\boldsymbol{m}} + W_{\boldsymbol{m}}\mathbf{j}$ with $Z_{\boldsymbol{m}}, W_{\boldsymbol{m}} \in \mathbb{C}_{\mathbf{i}}^{m\times n}$. Set $a = \langle \boldsymbol{m},\boldsymbol{\theta}\rangle_{S_L}$ and $b = \langle \boldsymbol{m},\boldsymbol{\theta}\rangle_{S_R}$. Since $e^{-\mathbf{i}a}, e^{-\mathbf{i}b} \in \mathbb{C}_{\mathbf{i}}$ commute with $Z_{\boldsymbol{m}}, W_{\boldsymbol{m}}$ and $\mathbf{j}\alpha = \overline{\alpha}\mathbf{j}$ for all $\alpha \in \mathbb{C}_{\mathbf{i}}$, we compute

$$L_{S_L} A_{\boldsymbol{m}} R_{S_R} = e^{-\mathbf{i}a}(Z_{\boldsymbol{m}} + W_{\boldsymbol{m}}\mathbf{j})e^{-\mathbf{i}b} = Z_{\boldsymbol{m}} e^{-\mathbf{i}(a+b)} + W_{\boldsymbol{m}}\,(e^{-\mathbf{i}a}\mathbf{j})e^{-\mathbf{i}b} = Z_{\boldsymbol{m}} e^{-\mathbf{i}(a+b)} + W_{\boldsymbol{m}}\, e^{-\mathbf{i}(a-b)}\mathbf{j}.$$

Because $a+b = \langle \boldsymbol{m},\boldsymbol{\theta}\rangle$ and $a-b = \langle \tilde{\boldsymbol{m}},\boldsymbol{\theta}\rangle$, this yields (4.7). Now let $F = Z + W\mathbf{j} \in L^1(\mathbb{T}^d; \mathbb{H}^{m\times n})$ with $Z, W : \mathbb{T}^d \to \mathbb{C}_{\mathbf{i}}^{m\times n}$. Since the decomposition $F \mapsto (Z, W)$ is given by a bounded linear map $\mathbb{H}^{m\times n} \to \mathbb{C}_{\mathbf{i}}^{m\times n} \times \mathbb{C}_{\mathbf{i}}^{m\times n}$, we infer $Z, W \in L^1$. Applying the previous algebra pointwise and integrating entrywise,

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{m}) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \big(Z(\boldsymbol{\theta})e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle} + W(\boldsymbol{\theta})e^{-\mathbf{i}\langle \tilde{\boldsymbol{m}},\boldsymbol{\theta}\rangle}\mathbf{j}\big)\,\mathrm{d}\boldsymbol{\theta} = \widehat{Z}(\boldsymbol{m}) + \widehat{W}(\tilde{\boldsymbol{m}})\,\mathbf{j},$$

which is (4.8). For (4.9), by definition $\widehat{F}^{(\mathrm{L})}(\boldsymbol{m}) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} e^{-\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle} F(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta}$. Taking conjugates entrywise, using the scalar quaternion identity $\overline{ab} = \overline{b}\,\overline{a}$ and $\overline{e^{-\mathbf{i}t}} = e^{\mathbf{i}t}$,

$$\overline{\widehat{F}^{(\mathrm{L})}(\boldsymbol{m})} = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \overline{F(\boldsymbol{\theta})}\, e^{\mathbf{i}\langle \boldsymbol{m},\boldsymbol{\theta}\rangle}\,\mathrm{d}\boldsymbol{\theta} = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \overline{F(\boldsymbol{\theta})}\, e^{-\mathbf{i}\langle -\boldsymbol{m},\boldsymbol{\theta}\rangle}\,\mathrm{d}\boldsymbol{\theta} = \widehat{\overline{F}}^{(\mathrm{R})}(-\boldsymbol{m}),$$

and conjugating both sides gives the claim. □

Approximating classes of sequences

Proof of Theorem 4.26. We use Corollary 4.9.

(a)⇒(b). Fix m and $n \ge n_m$. From (2.4), $A_n = B_{n,m} + R_{n,m} + N_{n,m}$ with $\operatorname{rank}_{\mathbb{H}}(R_{n,m})/r_n \le c(m)$ and $\|N_{n,m}\| \le \omega(m)$. Set

$$C_{n,m} := \Phi_{\mathrm{ent}}(B_{n,m}), \quad S_{n,m} := \Phi_{\mathrm{ent}}(R_{n,m}), \quad T_{n,m} := \Phi_{\mathrm{ent}}(N_{n,m}).$$

Then

$$\begin{aligned} \Phi_{\mathrm{ent}}(A_n) &= C_{n,m} + S_{n,m} + T_{n,m}, \\ \frac{\operatorname{rank}_{\mathbb{C}}(S_{n,m})}{2r_n} &= \frac{2\operatorname{rank}_{\mathbb{H}}(R_{n,m})}{2r_n} \le c(m), \\ \|T_{n,m}\| &= \|N_{n,m}\| \le \omega(m). \end{aligned}$$

Thus $\{\{C_{n,m}\}_n\}_{m\in\mathbb{N}}$ is a complex rectangular a.c.s. with the same n_m, $c(m)$, and $\omega(m)$.

(b)⇒(a). The complex error terms need not lie in the symplectic image. We first project them into that image, control the resulting norms and ranks, and then recover quaternion error terms.

Projection into the symplectic image. Set

$$\mathsf{J} := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \mathsf{J}_k := I_k \otimes \mathsf{J}.$$

For each n, define

$$\tau_{d_n,e_n}\colon \mathbb{C}^{2d_n\times 2e_n} \longrightarrow \mathbb{C}^{2d_n\times 2e_n}, \qquad \tau_{d_n,e_n}(X) := -\mathsf{J}_{d_n}\,\overline{X}\,\mathsf{J}_{e_n},$$
$$\Pi_{d_n,e_n}\colon \mathbb{C}^{2d_n\times 2e_n} \longrightarrow \mathbb{C}^{2d_n\times 2e_n}, \qquad \Pi_{d_n,e_n}(X) := \tfrac{1}{2}\big(X + \tau_{d_n,e_n}(X)\big).$$

The map τ_{d_n,e_n} is conjugate-linear, whereas Π_{d_n,e_n} is real-linear. Since J_k is real unitary and $\mathsf{J}_k^2 = -I_{2k}$,

$$\tau_{d_n,e_n}^2 = \mathrm{id}, \qquad \|\tau_{d_n,e_n}(X)\| = \|X\|, \qquad \mathrm{rank}_{\mathbb{C}}\big(\tau_{d_n,e_n}(X)\big) = \mathrm{rank}_{\mathbb{C}}(X).$$

To identify the fixed-point space, write $X = (X_{ij})_{i,j}$ as a $d_n \times e_n$ block matrix with $X_{ij} \in \mathbb{C}^{2\times 2}$. If

$$X_{ij} = \begin{bmatrix} z & w \\ u & v \end{bmatrix},$$

then

$$-\mathsf{J}\,\overline{X_{ij}}\,\mathsf{J} = \begin{bmatrix} \overline{v} & -\overline{u} \\ -\overline{w} & \overline{z} \end{bmatrix}.$$

Hence $\tau_{d_n,e_n}(X) = X$ if and only if every block has the form

$$X_{ij} = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} = \Phi(z + w\mathbf{j}).$$

Therefore

$$\mathrm{Fix}(\tau_{d_n,e_n}) = \Phi_{\mathrm{ent}}\big(\mathbb{H}^{d_n\times e_n}\big).$$

Since $\tau_{d_n,e_n}^2 = \mathrm{id}$, the map Π_{d_n,e_n} is the real-linear projector onto this fixed-point space, and

$$\|\Pi_{d_n,e_n}(X)\| \le \tfrac{1}{2}\big(\|X\| + \|\tau_{d_n,e_n}(X)\|\big) = \|X\|.$$

Norm and rank estimates. For $n \ge n_m$, assumption (b) gives

$$\Phi_{\mathrm{ent}}(A_n) = \Phi_{\mathrm{ent}}(B_{n,m}) + S_{n,m} + T_{n,m}, \qquad \frac{\mathrm{rank}_{\mathbb{C}}(S_{n,m})}{2r_n} \le c(m), \quad \|T_{n,m}\| \le \omega(m). \tag{4.42}$$

Since $\Phi_{\mathrm{ent}}(B_{n,m}) \in \mathrm{Ran}(\Phi_{\mathrm{ent}})$, projecting (4.42) onto $\mathrm{Ran}(\Phi_{\mathrm{ent}})$ gives

$$\Phi_{\mathrm{ent}}(A_n) = \Phi_{\mathrm{ent}}(B_{n,m}) + \Pi_{d_n,e_n}(S_{n,m}) + \Pi_{d_n,e_n}(T_{n,m}).$$

with

$$\|\Pi_{d_n,e_n}(T_{n,m})\| \le \omega(m),$$
$$\mathrm{rank}_{\mathbb{C}}\big(\Pi_{d_n,e_n}(S_{n,m})\big) \le \mathrm{rank}_{\mathbb{C}}(S_{n,m}) + \mathrm{rank}_{\mathbb{C}}\big(\tau_{d_n,e_n}(S_{n,m})\big)$$
$$\le 2\,\mathrm{rank}_{\mathbb{C}}(S_{n,m}).$$

Recovery of the quaternion decomposition. Since the projected error terms lie in $\mathrm{Ran}(\Phi_{\mathrm{ent}})$, set

$$R_{n,m} := \Phi_{\mathrm{ent}}^{-1}\big(\Pi_{d_n,e_n}(S_{n,m})\big), \qquad N_{n,m} := \Phi_{\mathrm{ent}}^{-1}\big(\Pi_{d_n,e_n}(T_{n,m})\big).$$

Then $A_n = B_{n,m} + R_{n,m} + N_{n,m}$ and

$$\frac{\mathrm{rank}_{\mathbb{H}}(R_{n,m})}{r_n} = \frac{\mathrm{rank}_{\mathbb{C}}(\Pi_{d_n,e_n}(S_{n,m}))}{2r_n} \le 2\,c(m), \qquad \|N_{n,m}\| = \|\Pi_{d_n,e_n}(T_{n,m})\| \le \omega(m).$$

Thus $\{\{B_{n,m}\}_n\}_{m\in\mathbb{N}}$ is a quaternion a.c.s. with the same n_m, the same $\omega(m)$, and rank control at most $2c(m)$. □

Quaternion a.c.s. calculus. For distributional transfer, apply the corresponding complex ACS rule to the embedded approximants, use the distribution definitions, and return through Theorem 4.26. In either realization statement, the converse existence direction also follows by taking the constant class $B_{n,m} = A_n$ and the unchanged symbol. This choice preserves Hermitian structure in $\mathrm{ACS}_{\mathbb{H}}2$.

Adjoints and products commute with the embedding by Lemma 4.1. A fixed quaternion scalar acts on the left through $\Phi_{\mathrm{ent}}(\alpha I)$; its operator norm is $|\alpha|$, and its action cannot increase rank. Thus direct rank–norm decompositions give the sum rule with quaternion coefficients. For products of two approximants, the hypotheses of ACS 3 hold for the embedded sequences: sparse tails are unchanged by singular-value duplication, and

$$\mathrm{PC}(d, e, f) \quad \Longleftrightarrow \quad \mathrm{PC}(2d, 2e, 2f).$$

The complex product is the embedding of the quaternion product, so Theorem 4.26 transfers the conclusion back. The same argument gives fixed multiplication under the corresponding hypotheses, without requiring sparse unboundedness of the other target.

The entrywise embedding respects conformal block assembly, which proves $\mathrm{ACS}_{\mathbb{H}}4$. Finally, the Schatten scaling and doubled active dimension cancel:

$$\frac{\|\Phi_{\mathrm{ent}}(A_n - B_{n,m})\|_{S_p}}{(2r_n)^{1/p}} = \frac{\|A_n - B_{n,m}\|_{S_p}}{r_n^{1/p}},$$

with the usual convention for $p = \infty$. Applying ACS 5 proves $\mathrm{ACS}_{\mathbb{H}}5$. □

4.7.3 Proofs of Toeplitz and circulant results

The proofs are grouped into structural identities, embedded symbols and distributions, norm estimates and localization, and circulant approximation. The short Toeplitz-linearity and distribution-transfer proofs are given beside their statements in Subsections 4.4.2 and 4.4.3.

Toeplitz structural identities

Proof of Proposition 4.27. Fix $\boldsymbol{k} \in \mathbb{Z}^d$. By (4.8),

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) = \widehat{Z}(\boldsymbol{k}) + \widehat{W}(\tilde{\boldsymbol{k}})\,\mathbf{j},$$

where $\tilde{k}_j = k_j$ for $j \in S_L$ and $\tilde{k}_j = -k_j$ for $j \in S_R$. Given $H = Z + (W \circ r_{S_L})\mathbf{j}$, applying (4.8) with the right kernel ($S_L = \varnothing, S_R = [d]$) yields

$$\widehat{H}^{(\mathrm{R})}(\boldsymbol{k}) = \widehat{Z}(\boldsymbol{k}) + \widehat{W \circ r_{S_L}}\,(-\boldsymbol{k})\,\mathbf{j},$$

and since $\langle -\boldsymbol{k}, r_{S_L}(\boldsymbol{\theta})\rangle = \langle \tilde{\boldsymbol{k}}, \boldsymbol{\theta}\rangle$ and Lebesgue measure on $\mathbb{T}^d$ is invariant under r_{S_L},

$$\widehat{W \circ r_{S_L}}\,(-\boldsymbol{k}) = \widehat{W}(\tilde{\boldsymbol{k}}).$$

Thus $\widehat{H}^{(\mathrm{R})}(\boldsymbol{k}) = \widehat{F}^{(S_L,S_R)}(\boldsymbol{k})$ for all $\boldsymbol{k}$, which implies

$$\big[T_{\boldsymbol{n}}^{(S_L,S_R)}(F)\big]_{\alpha,\beta} = \widehat{F}^{(S_L,S_R)}(\alpha - \beta) = \widehat{H}^{(\mathrm{R})}(\alpha - \beta) = \big[T_{\boldsymbol{n}}^{(\mathrm{R})}(H)\big]_{\alpha,\beta},$$

proving (4.17). Taking $S_L = [d]$ gives (4.18). For the fixed-size criterion, $T_{\boldsymbol{n}}^{(\mathrm{L})}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(F)$ if and only if $\widehat{F}^{(\mathrm{L})}(\boldsymbol{k}) = \widehat{F}^{(\mathrm{R})}(\boldsymbol{k})$ for all difference indices $\boldsymbol{k} = \alpha - \beta$ that occur in $T_{\boldsymbol{n}}$, i.e., for all $\boldsymbol{k} \in \Delta(\boldsymbol{n})$; namely

$$\widehat{Z}(\boldsymbol{k}) + \widehat{W}(\boldsymbol{k})\mathbf{j} = \widehat{Z}(\boldsymbol{k}) + \widehat{W}(-\boldsymbol{k})\mathbf{j} \quad \text{for all } \boldsymbol{k} \in \Delta(\boldsymbol{n}),$$

which is equivalent to (4.19). The equality $T_{\boldsymbol{n}}^{(\mathrm{L})}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(F)$ for every $\boldsymbol{n} \in \mathbb{N}^d$ holds if and only if the above identity holds for all $\boldsymbol{k} \in \mathbb{Z}^d$, yielding (4.20), which is equivalent to $W(\boldsymbol{\theta}) = W(-\boldsymbol{\theta})$ a.e. □

Proof of Lemma 4.28. Since the kernels $e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_L}}$ and $e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_R}}$ take values in $\mathbb{C}_{\mathbf{i}}$, they commute with a, b. Using associativity in $\mathbb{H}$, commutation with the $\mathbb{C}_{\mathbf{i}}$-valued kernels, and the $\mathbb{R}$-linearity of the Lebesgue integral, we obtain

$$\begin{aligned}
\widehat{aF+bG}^{(S_L,S_R)}(\boldsymbol{k}) &= (2\pi)^{-d}\int e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_L}}\,(aF+bG)\,e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_R}}\,\mathrm{d}\boldsymbol{\theta}\\
&= (2\pi)^{-d}\int \left(a\,e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_L}}Fe^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_R}} + b\,e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_L}}Ge^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle_{S_R}}\right)\mathrm{d}\boldsymbol{\theta}\\
&= a\,\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) + b\,\widehat{G}^{(S_L,S_R)}(\boldsymbol{k}),
\end{aligned}$$

which proves (4.21). □

Proof of Theorem 4.30. By the $*$-closure of integrals,

$$\left(\widehat{F}^{(\mathrm{L})}(-\boldsymbol{k})\right)^* = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d}\left(F(\boldsymbol{\theta})\right)^* e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}\,\mathrm{d}\boldsymbol{\theta} = \widehat{(F^*)}^{(\mathrm{R})}(\boldsymbol{k}),$$

and similarly

$$\left(\widehat{F}^{(\mathrm{R})}(-\boldsymbol{k})\right)^* = \widehat{(F^*)}^{(\mathrm{L})}(\boldsymbol{k}).$$

For the two-sided case,

$$\left(\widehat{F}^{(S_L,S_R)}(-\boldsymbol{k})\right)^* = \widehat{(F^*)}^{(S_R,S_L)}(\boldsymbol{k}).$$

Hence, for $\alpha,\beta\in\Lambda_{\boldsymbol{n}}$,

$$\left(T_{\boldsymbol{n}}^{(\tau)}(F)\right)^*_{\alpha\beta} = \left(\widehat{F}^{(\tau)}(\beta-\alpha)\right)^* = \widehat{(F^*)}^{(\tau')}(\alpha-\beta) = \left(T_{\boldsymbol{n}}^{(\tau')}(F^*)\right)_{\alpha\beta},$$

where

$$(\tau,\tau')\in\left\{(\mathrm{L},\mathrm{R}),\,(\mathrm{R},\mathrm{L}),\,\big((S_L,S_R),(S_R,S_L)\big)\right\}.$$

This proves the claim. □

Proof of Theorem 4.31. For fixed $\boldsymbol{n}$, the matrix $T_{\boldsymbol{n}}^{(S_L,S_R)}(F)$ is Hermitian if and only if

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) = \left(\widehat{F}^{(S_L,S_R)}(-\boldsymbol{k})\right)^* \qquad \text{for all } \boldsymbol{k}\in\Delta(\boldsymbol{n}), \tag{4.43}$$

where $\Delta(\boldsymbol{n}) = \{\alpha-\beta:\ \alpha,\beta\in\Lambda_{\boldsymbol{n}}\}$. Since the statement holds for all $\boldsymbol{n}$, (4.43) is equivalent to the same identity for all $\boldsymbol{k}\in\mathbb{Z}^d$. Let $\boldsymbol{k}_L$ and $\boldsymbol{k}_R$ be the coordinate projections defined before (4.23). By (4.23),

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) = \widehat{Z}(\boldsymbol{k}_L+\boldsymbol{k}_R) + \widehat{W}(\boldsymbol{k}_L-\boldsymbol{k}_R)\,\mathbf{j}.$$

Set

$$\boldsymbol{p} := \boldsymbol{k}_L+\boldsymbol{k}_R = \boldsymbol{k}, \qquad \boldsymbol{q} := \boldsymbol{k}_L-\boldsymbol{k}_R.$$

The maps $\boldsymbol{k}\mapsto\boldsymbol{p}$ and $\boldsymbol{k}\mapsto\boldsymbol{q}$ are bijections of $\mathbb{Z}^d$; the latter changes only the signs of selected coordinates. For each fixed $\boldsymbol{k}$, equation (4.43) becomes

$$\widehat{Z}(\boldsymbol{p}) + \widehat{W}(\boldsymbol{q})\,\mathbf{j} = \left(\widehat{Z}(-\boldsymbol{p}) + \widehat{W}(-\boldsymbol{q})\,\mathbf{j}\right)^* = \widehat{Z}(-\boldsymbol{p})^* + (\widehat{W}(-\boldsymbol{q})\,\mathbf{j})^*.$$

Using $\mathbf{j}^* = -\mathbf{j}$ and $\mathbf{j}A = \overline{A}\,\mathbf{j}$ for $A\in\mathbb{C}_{\mathbf{i}}^{s\times s}$, together with $\overline{A^*} = A^{\mathsf{T}}$, we obtain

$$(\widehat{W}(-\boldsymbol{q})\,\mathbf{j})^* = \mathbf{j}^*\widehat{W}(-\boldsymbol{q})^* = -\mathbf{j}\,\widehat{W}(-\boldsymbol{q})^* = -\overline{\widehat{W}(-\boldsymbol{q})^*}\,\mathbf{j} = -\left(\widehat{W}(-\boldsymbol{q})\right)^{\mathsf{T}}\mathbf{j}.$$

Hence

$$\widehat{Z}(\boldsymbol{p}) + \widehat{W}(\boldsymbol{q})\,\mathbf{j} = \widehat{Z}(-\boldsymbol{p})^* - \left(\widehat{W}(-\boldsymbol{q})\right)^{\mathsf{T}}\mathbf{j}.$$

Since 1 and $\mathbf{j}$ are linearly independent over the slice $\mathbb{C}_\mathbf{i}$, and since both maps above are bijective, this is equivalent to

$$\widehat{Z}(\boldsymbol{p}) = \widehat{Z}(-\boldsymbol{p})^* \qquad \text{and} \qquad \widehat{W}(\boldsymbol{q}) = -\big(\widehat{W}(-\boldsymbol{q})\big)^{\mathsf{T}} \quad \text{for all } \boldsymbol{p}, \boldsymbol{q} \in \mathbb{Z}^d. \tag{4.44}$$

By standard properties of complex Fourier series, (4.44) is equivalent to $Z(\boldsymbol{\theta})^* = Z(\boldsymbol{\theta})$ and $W(-\boldsymbol{\theta}) = -W(\boldsymbol{\theta})^{\mathsf{T}}$ a.e. on $\mathbb{T}^d$, which proves (i)⇒(ii). Conversely, if (ii) holds, then (4.44) holds and thus (4.43) follows for every $\boldsymbol{n}$, so $T_{\boldsymbol{n}}^{(S_L,S_R)}(F)$ is Hermitian for all $\boldsymbol{n}$. Because both maps above are bijections for every ordered partition, the conclusion does not depend on the partition. □

Embeddings and distribution theorems

Proof of Theorem 4.32. Let $\boldsymbol{k} = \alpha - \beta \in \mathbb{Z}^d$. Using (4.27) with the left multiplier $\lambda(\boldsymbol{\theta}) = e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_s$ for the left case (respectively the right multiplier $\rho(\boldsymbol{\theta}) = e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_t$ for the right case), we obtain

$$\Phi_{\mathrm{grp}}\big(\widehat{F}^{(\mathrm{L})}(\boldsymbol{k})\big) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \operatorname{diag}\big(e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_s,\, e^{+\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_s\big)\, \Phi_{\mathrm{grp}}\big(F(\boldsymbol{\theta})\big)\, \mathrm{d}\boldsymbol{\theta},$$

and

$$\Phi_{\mathrm{grp}}\big(\widehat{F}^{(\mathrm{R})}(\boldsymbol{k})\big) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \Phi_{\mathrm{grp}}\big(F(\boldsymbol{\theta})\big)\, \operatorname{diag}\big(e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_t,\, e^{+\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle} I_t\big)\, \mathrm{d}\boldsymbol{\theta}.$$

Since $\Phi_{\mathrm{grp}}(F) = \left[\begin{smallmatrix} Z & W \\ -\overline{W} & \overline{Z} \end{smallmatrix}\right]$, a direct multiplication of the integrands yields the phase factors $Z\, e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}$ and $W\, e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}$ in the left case, and $Z\, e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}$ and $W\, e^{+\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}$ in the right case; the lower row carries the corresponding conjugate phases. Therefore

$$\Phi_{\mathrm{grp}}\big(\widehat{F}^{(\mathrm{L})}(\boldsymbol{k})\big) = \begin{pmatrix} \widehat{Z}(\boldsymbol{k}) & \widehat{W}(\boldsymbol{k}) \\ \overline{-\widehat{W}(\boldsymbol{k})} & \overline{\widehat{Z}(\boldsymbol{k})} \end{pmatrix}, \qquad \Phi_{\mathrm{grp}}\big(\widehat{F}^{(\mathrm{R})}(\boldsymbol{k})\big) = \begin{pmatrix} \widehat{Z}(\boldsymbol{k}) & \widehat{W}(-\boldsymbol{k}) \\ \overline{-\widehat{W}(-\boldsymbol{k})} & \overline{\widehat{Z}(\boldsymbol{k})} \end{pmatrix}.$$

These are precisely the $\boldsymbol{k}$-th Fourier coefficients of G_{L} and G_{R}, respectively. Setting $\boldsymbol{k} = \alpha - \beta$ gives the Toeplitz identities. □

Proof of Theorem 4.33. Let $\boldsymbol{k} = \alpha - \beta \in \mathbb{Z}^d$. Let $\boldsymbol{k}_L$ and $\boldsymbol{k}_R$ be the coordinate projections defined before (4.23). By the definition of the sandwich Fourier coefficients,

$$\widehat{F}^{(S_L,S_R)}(\boldsymbol{k}) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} e^{-\mathbf{i}\langle \boldsymbol{k}_L,\boldsymbol{\theta}\rangle}\, (Z + W\mathbf{j})\, e^{-\mathbf{i}\langle \boldsymbol{k}_R,\boldsymbol{\theta}\rangle}\, \mathrm{d}\boldsymbol{\theta}.$$

Applying the covariance identity (4.27) with

$$\lambda(\boldsymbol{\theta}) = e^{-\mathbf{i}\langle \boldsymbol{k}_L,\boldsymbol{\theta}\rangle} I_s, \qquad \rho(\boldsymbol{\theta}) = e^{-\mathbf{i}\langle \boldsymbol{k}_R,\boldsymbol{\theta}\rangle} I_t,$$

we obtain

$$\Phi_{\mathrm{grp}}\big(\widehat{F}^{(S_L,S_R)}(\boldsymbol{k})\big) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \begin{bmatrix} e^{-\mathbf{i}\langle \boldsymbol{k}_L,\boldsymbol{\theta}\rangle} I_s & 0 \\ 0 & e^{\mathbf{i}\langle \boldsymbol{k}_L,\boldsymbol{\theta}\rangle} I_s \end{bmatrix} \begin{bmatrix} Z & W \\ -\overline{W} & \overline{Z} \end{bmatrix} \times \begin{bmatrix} e^{-\mathbf{i}\langle \boldsymbol{k}_R,\boldsymbol{\theta}\rangle} I_t & 0 \\ 0 & e^{\mathbf{i}\langle \boldsymbol{k}_R,\boldsymbol{\theta}\rangle} I_t \end{bmatrix} \mathrm{d}\boldsymbol{\theta}.$$

Multiplication of the integrand gives the four entries

$$Z\, e^{-\mathbf{i}\langle \boldsymbol{k}_L+\boldsymbol{k}_R,\boldsymbol{\theta}\rangle}, \quad W\, e^{-\mathbf{i}\langle \boldsymbol{k}_L-\boldsymbol{k}_R,\boldsymbol{\theta}\rangle}, \quad -\overline{W}\, e^{+\mathbf{i}\langle \boldsymbol{k}_L-\boldsymbol{k}_R,\boldsymbol{\theta}\rangle}, \quad \overline{Z}\, e^{+\mathbf{i}\langle \boldsymbol{k}_L+\boldsymbol{k}_R,\boldsymbol{\theta}\rangle}.$$

Hence

$$\Phi_{\mathrm{grp}}\big(\widehat{F}^{(S_L,S_R)}(\boldsymbol{k})\big) = \begin{pmatrix} \widehat{Z}(\boldsymbol{k}_L + \boldsymbol{k}_R) & \widehat{W}(\boldsymbol{k}_L - \boldsymbol{k}_R) \\ \overline{-\widehat{W}(\boldsymbol{k}_L - \boldsymbol{k}_R)} & \overline{\widehat{Z}(\boldsymbol{k}_L + \boldsymbol{k}_R)} \end{pmatrix}.$$

By standard changes of variables on $\mathbb{T}^d$, we have

$$\widehat{H\circ(-\mathrm{id})}(\boldsymbol{k}) = \widehat{H}(-\boldsymbol{k}), \qquad \widehat{H\circ r_{S_L}}(\boldsymbol{k}) = \widehat{H}\big(r_{S_L}(\boldsymbol{k})\big),$$

and since $r_{S_L}(-\boldsymbol{k}) = \boldsymbol{k}_L - \boldsymbol{k}_R$, we obtain

$$\widehat{W\circ r_{S_L}\circ(-\mathrm{id})}(\boldsymbol{k}) = \widehat{W}\big(r_{S_L}(-\boldsymbol{k})\big) = \widehat{W}(\boldsymbol{k}_L - \boldsymbol{k}_R), \qquad \widehat{\overline{Z(-\cdot)}}(\boldsymbol{k}) = \overline{\widehat{Z}(\boldsymbol{k})}.$$

Therefore the last displayed matrix equals $\widehat{G_{S_L,S_R}}(\boldsymbol{k})$. Setting $\boldsymbol{k} = \alpha - \beta$ and assembling by blocks yields $\Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(S_L,S_R)}(F)\big) = T_{\boldsymbol{n}}\big(G_{S_L,S_R}\big)$, which is (4.28). □

Proof of Theorem 4.35. By the embedding identities,

$$\Phi_{\mathrm{blk}}^{(s,s)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) = T_{\boldsymbol{n}}(G_\tau).$$

Since $G_\tau \in L^\infty(\mathbb{T}^d;\mathbb{C}^{(2s)\times(2s)})$ and belongs to the complex Tilli class, Theorem 2.24 gives $\{T_{\boldsymbol{n}}(G_\tau)\}_{\boldsymbol{n}} \sim_\lambda G_\tau$. The permutation relation (4.25) identifies this with the entrywise symplectic representation, so Definition 4.19 transfers the eigenvalue distribution to the quaternion sequence. □

Schatten estimates and localization

Proof of Theorem 4.36. It is sufficient to consider the general sandwich-kernel symbol

$$G_{S_L,S_R}(\boldsymbol{\theta}) = \begin{pmatrix} Z(\boldsymbol{\theta}) & W\big(r_{S_L}(-\boldsymbol{\theta})\big) \\ -\overline{W\big(r_{S_L}(\boldsymbol{\theta})\big)} & \overline{Z(-\boldsymbol{\theta})} \end{pmatrix},$$

where $r_{S_L}:\mathbb{T}^d\to\mathbb{T}^d$ is the reflection of the left-kernel variables. Define the row projections

$$P_1 := \mathrm{diag}(I_s,0)\in\mathbb{C}^{(2s)\times(2s)}, \qquad P_2 := \mathrm{diag}(0,I_s),$$

and the column projections

$$E_1 := \mathrm{diag}(I_t,0)\in\mathbb{C}^{(2t)\times(2t)}, \qquad E_2 := \mathrm{diag}(0,I_t).$$

Then $\|P_j\|_{S_\infty} = \|E_j\|_{S_\infty} = 1$ for $j = 1,2$. Using these projections, we decompose the symbol into four terms:

$$\begin{aligned} G_{S_L,S_R}(\boldsymbol{\theta}) = {} & P_1\Phi_{\mathrm{grp}}\big(F(\boldsymbol{\theta})\big)E_1 + P_1\Phi_{\mathrm{grp}}\big(F(r_{S_L}(-\boldsymbol{\theta}))\big)E_2 \\ & + P_2\Phi_{\mathrm{grp}}\big(F(r_{S_L}(\boldsymbol{\theta}))\big)E_1 + P_2\Phi_{\mathrm{grp}}\big(F(-\boldsymbol{\theta})\big)E_2. \end{aligned}$$

Explicitly, the four summands are

$$\begin{pmatrix} Z(\boldsymbol{\theta}) & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & W(r_{S_L}(-\boldsymbol{\theta})) \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ -\overline{W(r_{S_L}(\boldsymbol{\theta}))} & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & \overline{Z(-\boldsymbol{\theta})} \end{pmatrix}.$$

For the contractions A and B used above and a compatible matrix X, the ideal property gives

$$\|AXB\|_{S_p} \le \|A\|\,\|X\|_{S_p}\,\|B\| \le \|X\|_{S_p}.$$

Consequently,

$$\|G_{S_L,S_R}(\boldsymbol{\theta})\|_{S_p} \le \sum_{j=1}^{4}\big\|\Phi_{\mathrm{grp}}\big(F(\varphi_j(\boldsymbol{\theta}))\big)\big\|_{S_p},$$

where each φ_j is one of

$$\mathrm{id}, \qquad -\mathrm{id}, \qquad r_{S_L}, \qquad r_{S_L}\circ(-\mathrm{id}).$$

These maps preserve normalized Lebesgue measure on $\mathbb{T}^d$. Hence

$$\big\|\Phi_{\mathrm{grp}} \circ F \circ \varphi_j\big\|_{L^p} = \big\|\Phi_{\mathrm{grp}} \circ F\big\|_{L^p}.$$

The same formula covers the pure left and right placements when one part of the ordered partition is empty. Therefore, for every supported placement τ,

$$\|G_\tau\|_{L^p(\mathbb{T}^d)} \le 4\big\|\Phi_{\mathrm{grp}} \circ F\big\|_{L^p(\mathbb{T}^d)} = 4 \cdot 2^{1/p}\|F\|_{L^p(\mathbb{T}^d;\mathbb{H}^{s\times t})},$$

where the last equality is (4.31).

For the corresponding Toeplitz matrix, the embedding identity gives

$$\Phi_{\mathrm{blk}}^{(s,t)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) = T_{\boldsymbol{n}}(G_\tau).$$

By the Schatten scaling under the block symplectic embedding,

$$\big\|T_{\boldsymbol{n}}^{(\tau)}(F)\big\|_{S_p} = 2^{-1/p}\big\|T_{\boldsymbol{n}}(G_\tau)\big\|_{S_p}.$$

Applying Theorem 2.25 and the preceding L^p bound for G_τ, we obtain

$$\begin{aligned}\big\|T_{\boldsymbol{n}}^{(\tau)}(F)\big\|_{S_p} &\le 2^{-1/p}\left(\frac{N_{\boldsymbol{n}}}{(2\pi)^d}\right)^{1/p}\|G_\tau\|_{L^p(\mathbb{T}^d)}\\ &\le 4\left(\frac{N_{\boldsymbol{n}}}{(2\pi)^d}\right)^{1/p}\|F\|_{L^p(\mathbb{T}^d;\mathbb{H}^{s\times t})},\end{aligned}$$

which proves the claim. □

Proof of Theorem 4.37. The embedding identities give

$$\Phi_{\mathrm{blk}}^{(s,s)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big) = T_{\boldsymbol{n}}(G_\tau).$$

Since $T_{\boldsymbol{n}}^{(\tau)}(F)$ is Hermitian, $\Phi_{\mathrm{blk}}^{(s,s)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big)$ is Hermitian and hence $T_{\boldsymbol{n}}(G_\tau)$ is Hermitian as well. Because this holds for every $\boldsymbol{n} \in \mathbb{N}^d$, every $\boldsymbol{k} \in \mathbb{Z}^d$ occurs as a Toeplitz difference index for some $\boldsymbol{n}$, and therefore $\widehat{G}_\tau(\boldsymbol{k}) = \widehat{G}_\tau(-\boldsymbol{k})^*$ for every $\boldsymbol{k}$. Uniqueness of Fourier coefficients gives $G_\tau(\boldsymbol{\theta}) = G_\tau(\boldsymbol{\theta})^*$ for a.e. $\boldsymbol{\theta} \in \mathbb{T}^d$. Hence Theorem 2.26 gives

$$\operatorname{spec}\big(T_{\boldsymbol{n}}(G_\tau)\big) \subset [\, m_\tau, M_\tau \,].$$

By (4.25), the block-ordered and entrywise realizations are unitarily similar. The entrywise representation duplicates the real eigenvalues of a Hermitian quaternion matrix, so (4.33) follows. □

Proof of Corollary 4.38. (a) Apply Theorem 4.37. For (b), the Hermitian matrices $T_{\boldsymbol{n}}(G_\tau) = \Phi_{\mathrm{blk}}^{(s,s)}\big(T_{\boldsymbol{n}}^{(\tau)}(F)\big)$ have all their real eigenvalues at least c. Suppose, to the contrary, that $m_\tau < c$. Then there exists $\varepsilon > 0$ such that the set

$$E_\varepsilon := \big\{\boldsymbol{\theta} \in \mathbb{T}^d : \lambda_{\min}(G_\tau(\boldsymbol{\theta})) \le c - \varepsilon\big\}$$

has positive measure. Since the eigenvalues of the finite matrix $G_\tau(\boldsymbol{\theta})$ are finite for a.e. $\boldsymbol{\theta}$, there exists $M > 0$ such that

$$E_{\varepsilon,M} := E_\varepsilon \cap \big\{\boldsymbol{\theta} : \lambda_{\min}(G_\tau(\boldsymbol{\theta})) \ge -M\big\}$$

still has positive measure. Choose a nonnegative $\varphi \in C_c(\mathbb{R})$ that vanishes on $[c, \infty)$ and is strictly positive on $[-M, c-\varepsilon]$. Every finite-matrix average of φ is zero, whereas the symbol average in Theorem 2.23 is positive because, on $E_{\varepsilon,M}$, at least one eigenvalue contributes a positive term. This contradiction proves

$$\lambda_{\min}\big(G_\tau(\boldsymbol{\theta})\big) \ge c \quad \text{for a.e. } \boldsymbol{\theta} \in \mathbb{T}^d,$$

and hence $m_\tau \ge c$. □

Circulants and a.c.s. approximation

Proof of Proposition 4.44. Write each coefficient of p_{R} in Cartesian form

$$p_{\mathrm{R}}(\boldsymbol{\rho}) = p_{\mathrm{R}}(\boldsymbol{\rho})^{(0)} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(1)}\mathbf{i} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\mathbf{j} + p_{\mathrm{R}}(\boldsymbol{\rho})^{(3)}(\mathbf{ij}),$$

with real blocks $p_{\mathrm{R}}(\boldsymbol{\rho})^{(\ell)} \in \mathbb{R}^{s\times t}$. Expanding the shift representation,

$$\begin{aligned} C_{\boldsymbol{n}}(p_{\mathrm{R}}) = \sum_{\boldsymbol{\rho}\in\mathcal{F}} \big(P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(0)}\big) + \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(1)}\Big)\mathbf{i} \\ + \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\Big)\mathbf{j} + \Big(\sum_{\boldsymbol{\rho}\in\mathcal{F}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(3)}\Big)(\mathbf{ij}). \end{aligned}$$

For $\ell = 0, 1$, Proposition 2.32 applied componentwise gives

$$U_L\Big(\sum_{\boldsymbol{\rho}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(\ell)}\Big)U_R^* = \mathrm{Bdiag}\big(\widehat{S}_\ell(\boldsymbol{k})\big)_{\boldsymbol{k}\in\Lambda_{\boldsymbol{n}}} =: \Lambda_\ell.$$

For the $\mathbf{j}$-component, use $\mathbf{j}\alpha = \overline{\alpha}\mathbf{j}$ for $\alpha \in \mathbb{C}_{\mathbf{i}}$ and the QDFT flip $F_{\mathbf{i}}^{(\boldsymbol{n})}(\mathbf{j}I_{N_{\boldsymbol{n}}})(F_{\mathbf{i}}^{(\boldsymbol{n})})^* = A_{\boldsymbol{n}}(\mathbf{j}I_{N_{\boldsymbol{n}}})$ from Lemma 4.40. Then

$$\begin{aligned} &U_L\Big(\sum_{\boldsymbol{\rho}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\Big)(\mathbf{j}I_{N_{\boldsymbol{n}}t})\, U_R^* \\ &= \Big[U_L\Big(\sum_{\boldsymbol{\rho}} P_{\boldsymbol{n}}^{(\boldsymbol{\rho})} \otimes p_{\mathrm{R}}(\boldsymbol{\rho})^{(2)}\Big)U_R^*\Big] \cdot \big(U_R(\mathbf{j}I_{N_{\boldsymbol{n}}t})U_R^*\big) \\ &= \Lambda_2 \cdot \big((F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_t)((\mathbf{j}I_{N_{\boldsymbol{n}}}) \otimes I_t)(F_{\mathbf{i}}^{(\boldsymbol{n})} \otimes I_t)^*\big) \\ &= \Lambda_2 \cdot \big((F_{\mathbf{i}}^{(\boldsymbol{n})}(\mathbf{j}I_{N_{\boldsymbol{n}}})(F_{\mathbf{i}}^{(\boldsymbol{n})})^*) \otimes I_t\big) = \Lambda_2\,(A_{\boldsymbol{n}} \otimes I_t)(\mathbf{j}I_{N_{\boldsymbol{n}}t}). \end{aligned}$$

The case $\ell = 3$ is identical, and (4.36) follows. Consider now the permutations P of Definition 4.42 and set $\Pi_L := P \otimes I_s$ and $\Pi_R := P \otimes I_t$. Define

$$\widetilde{\Lambda}_\ell := \Pi_L \Lambda_\ell \Pi_R^* \qquad (\ell = 0, 1, 2, 3).$$

Applying Π_L to the block rows and Π_R^* to the block columns in (4.36), we obtain

$$\Pi_L U_L\, C_{\boldsymbol{n}}(p_{\mathrm{R}})\, U_R^*\Pi_R^* = \widetilde{\Lambda}_0 + \widetilde{\Lambda}_1\,\mathbf{i} + \Pi_L\Lambda_2(A_{\boldsymbol{n}} \otimes I_t)\Pi_R^*\mathbf{j} + \Pi_L\Lambda_3(A_{\boldsymbol{n}} \otimes I_t)\Pi_R^*(\mathbf{ij}).$$

Insert $\Pi_R^*\Pi_R = I$ between Λ_2 and $A_{\boldsymbol{n}} \otimes I_t$:

$$\Pi_L\Lambda_2(A_{\boldsymbol{n}} \otimes I_t)\Pi_R^* = (\Pi_L\Lambda_2\Pi_R^*) \cdot \big(\Pi_R(A_{\boldsymbol{n}} \otimes I_t)\Pi_R^*\big),$$

and the same calculation applies to Λ_3. The block identity in Lemma 4.43 then gives

$$Z := \Pi_R(A_{\boldsymbol{n}} \otimes I_t)\Pi_R^* = (P \otimes I_t)(A_{\boldsymbol{n}} \otimes I_t)(P \otimes I_t)^* = I_{\#\mathrm{Fix}} \otimes I_t \;\oplus\; \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} 0 & I_t \\ I_t & 0 \end{bmatrix}.$$

Therefore,

$$\Pi_L U_L\, C_{\boldsymbol{n}}(p_{\mathrm{R}})\, U_R^*\Pi_R^* = \widetilde{\Lambda}_0 + \widetilde{\Lambda}_1\,\mathbf{i} + \big(\widetilde{\Lambda}_2 Z\big)\mathbf{j} + \big(\widetilde{\Lambda}_3 Z\big)(\mathbf{ij}). \tag{4.45}$$

Equation (4.45) has the following block form.

- If $\boldsymbol{k} \in \mathrm{Fix}$, the corresponding block of Z is I_t, and the corresponding diagonal block in (4.45) is
$$D_1(\boldsymbol{k}) + D_2(\boldsymbol{k}) := \widehat{S}_0(\boldsymbol{k}) + \widehat{S}_1(\boldsymbol{k})\,\mathbf{i} + \big(\widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\,\mathbf{i}\big)\mathbf{j}.$$
- For $\boldsymbol{k} \in \mathcal{K}$, the corresponding paired block of Z is the 2×2 exchange matrix with $t \times t$ blocks, and the corresponding block in (4.45) is
$$\begin{bmatrix} D_1(\boldsymbol{k}) & D_2(\boldsymbol{k}) \\ D_2(-\boldsymbol{k}) & D_1(-\boldsymbol{k}) \end{bmatrix}, \qquad D_2(\boldsymbol{k}) := \big(\widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\,\mathbf{i}\big)\mathbf{j}.$$

Collecting the blocks in the order induced by P proves the proposition. □

Proof of Proposition 4.45. Proposition 4.44, in the notation of (4.39), gives the canonical QDFT form

$$\Pi_L\, U_L\, C_{\boldsymbol{n}}\big(p_{\mathrm{R}}\big)\, U_R^{*}\, \Pi_R^{*} = \bigoplus_{\boldsymbol{k}\in \mathrm{Fix}} \Big(\widehat{S}_0(\boldsymbol{k}) + \widehat{S}_1(\boldsymbol{k})\mathbf{i} + \big(\widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\mathbf{i}\big)\mathbf{j}\Big)$$

$$\oplus \bigoplus_{\boldsymbol{k}\in\mathcal{K}} \begin{bmatrix} \widehat{S}_0(\boldsymbol{k}) + \widehat{S}_1(\boldsymbol{k})\mathbf{i} & \big(\widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\mathbf{i}\big)\mathbf{j} \\ \big(\widehat{S}_2(-\boldsymbol{k}) + \widehat{S}_3(-\boldsymbol{k})\mathbf{i}\big)\mathbf{j} & \widehat{S}_0(-\boldsymbol{k}) + \widehat{S}_1(-\boldsymbol{k})\mathbf{i} \end{bmatrix}.$$

By the definitions of Z and W,

$$Z(\boldsymbol{\theta}_{\boldsymbol{k}}) = \widehat{S}_0(\boldsymbol{k}) + \widehat{S}_1(\boldsymbol{k})\,\mathbf{i}, \qquad W(-\boldsymbol{\theta}_{\boldsymbol{k}}) = \widehat{S}_2(\boldsymbol{k}) + \widehat{S}_3(\boldsymbol{k})\,\mathbf{i}.$$

Thus the fixed blocks are $Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}$ and the paired blocks are

$$\begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix},$$

which confirms (4.39). For the complex embedding, use the identities, valid for every $X \in \mathbb{C}_{\mathbf{i}}^{s\times t}$,

$$\Phi_{\mathrm{grp}}(X) = \begin{bmatrix} X & 0 \\ 0 & \overline{X} \end{bmatrix}, \qquad \Phi_{\mathrm{grp}}(X\mathbf{j}) = \begin{bmatrix} 0 & X \\ -\overline{X} & 0 \end{bmatrix},$$

together with the $\mathbb{R}$-linearity of Φ_{grp} and the block-ordered action of $\Phi_{\mathrm{blk}}^{(s,t)}$ on real 2×2 partitions. For a fixed index $\boldsymbol{k} \in \mathrm{Fix}$ we obtain

$$\Phi_{\mathrm{grp}}\big(Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) = \Phi_{\mathrm{grp}}\big(Z(\boldsymbol{\theta}_{\boldsymbol{k}})\big) + \Phi_{\mathrm{grp}}\big(W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) = \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}}) \\ -\overline{W(-\boldsymbol{\theta}_{\boldsymbol{k}})} & \overline{Z(\boldsymbol{\theta}_{\boldsymbol{k}})} \end{bmatrix},$$

which is (4.41). For the paired blocks, split them as sums of their diagonal and off-diagonal parts,

$$\begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix} = \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & 0 \\ 0 & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix} + \begin{bmatrix} 0 & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & 0 \end{bmatrix}.$$

Applying $\Phi_{\mathrm{blk}}^{(s,t)}$ to the paired block and summing, using the two identities above, yields exactly the 4×4 complex block in (4.40). This completes the proof. □

Proof of Corollary 4.46. Let $\boldsymbol{\theta}_{\boldsymbol{k}} = 2\pi(\boldsymbol{k}/\boldsymbol{n})$ for $\boldsymbol{k} \in \Lambda_{\boldsymbol{n}}$ and $\mathrm{Fix} = \{\boldsymbol{k} : 2\boldsymbol{k} \equiv 0 \pmod{\boldsymbol{n}}\}$ the usual fixed-point set. By Proposition 4.45 (see (4.39)), after applying the QDFT and the permutations, the matrix $C_{\boldsymbol{n}}(p_{\mathrm{R}})$ is block diagonal with blocks

$$\big(Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) \quad (\boldsymbol{k} \in \mathrm{Fix}), \qquad \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j} \\ W(-\boldsymbol{\theta}_{-\boldsymbol{k}})\mathbf{j} & Z(\boldsymbol{\theta}_{-\boldsymbol{k}}) \end{bmatrix} \quad (\{\boldsymbol{k}, -\boldsymbol{k}\}),$$

After the symplectic embedding, paired frequencies give the 4×4 block in (4.40), whereas fixed frequencies give the 2×2 block in (4.41). Applying the block permutation $(1, 4, 3, 2)$ to the rows and columns of the 4×4 blocks in (4.40) gives

$$\mathrm{diag}\Big(G_{\mathrm{R}}[p_{\mathrm{R}}](\boldsymbol{\theta}_{\boldsymbol{k}}),\ G_{\mathrm{R}}[p_{\mathrm{R}}](\boldsymbol{\theta}_{-\boldsymbol{k}})\Big),$$

while for $\boldsymbol{k} \in \mathrm{Fix}$ we have the periodic congruence $-\boldsymbol{\theta}_{\boldsymbol{k}} \equiv \boldsymbol{\theta}_{\boldsymbol{k}} \pmod{2\pi}$, so

$$\Phi_{\mathrm{grp}}\big(Z(\boldsymbol{\theta}_{\boldsymbol{k}}) + W(-\boldsymbol{\theta}_{\boldsymbol{k}})\mathbf{j}\big) = \begin{bmatrix} Z(\boldsymbol{\theta}_{\boldsymbol{k}}) & W(-\boldsymbol{\theta}_{\boldsymbol{k}}) \\ -\overline{W(-\boldsymbol{\theta}_{\boldsymbol{k}})} & \overline{Z(\boldsymbol{\theta}_{\boldsymbol{k}})} \end{bmatrix} = G_{\mathrm{R}}[p_{\mathrm{R}}](\boldsymbol{\theta}_{\boldsymbol{k}}).$$

Set $\mathcal{A}_{\boldsymbol{n}} \coloneqq \Phi_{\mathrm{blk}}^{(s,t)}(C_{\boldsymbol{n}}(p_{\mathrm{R}}))$ and $\mathcal{G}_{\boldsymbol{k}} \coloneqq G_{\mathrm{R}}[p_{\mathrm{R}}](\boldsymbol{\theta}_{\boldsymbol{k}})$. Hence, for every $\boldsymbol{n}$,

$$\big\{\{\sigma_j(\mathcal{A}_{\boldsymbol{n}}) : 1 \le j \le 2\min\{s,t\}N_{\boldsymbol{n}}\}\big\} = \big\{\{\sigma_j(\mathcal{G}_{\boldsymbol{k}}) : \boldsymbol{k} \in \Lambda_{\boldsymbol{n}},\ 1 \le j \le 2\min\{s,t\}\}\big\}.$$

If $s = t$, the left and right QDFT factors and the corresponding permutations coincide. The factorization is then a unitary similarity, and the same identity holds for eigenvalues:

$$\big\{\{\lambda_j(\mathcal{A}_{\boldsymbol{n}}) : 1 \le j \le 2sN_{\boldsymbol{n}}\}\big\} = \big\{\{\lambda_j(\mathcal{G}_{\boldsymbol{k}}) : \boldsymbol{k} \in \Lambda_{\boldsymbol{n}},\ 1 \le j \le 2s\}\big\}.$$

For every continuous compactly supported test function ψ, the corresponding sums over $\{\boldsymbol{\theta}_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda_{\boldsymbol{n}}}$ are multidimensional trapezoidal Riemann sums. Their convergence yields

$$\{C_{\boldsymbol{n}}(p_{\mathrm{R}})\}_{\boldsymbol{n}} \sim_\sigma G_{\mathrm{R}}[p_{\mathrm{R}}], \qquad \{C_{\boldsymbol{n}}(p_{\mathrm{R}})\}_{\boldsymbol{n}} \sim_\lambda G_{\mathrm{R}}[p_{\mathrm{R}}] \quad \text{when } s = t.$$

□

Proof of Theorem 4.47. Polynomial approximation. By Proposition 4.16, right trigonometric polynomials are dense in $L^1(\mathbb{T}^d;\mathbb{H}^{s\times t})$. Hence, there exist finite sets $\mathcal{S}_m \subset \mathbb{Z}^d$ and coefficients $A_{m,\boldsymbol{k}} \in \mathbb{H}^{s\times t}$ such that

$$F_m(\boldsymbol{\theta}) \;=\; \sum_{\boldsymbol{k}\in\mathcal{S}_m} A_{m,\boldsymbol{k}}\, e^{-\mathbf{i}\langle \boldsymbol{k},\boldsymbol{\theta}\rangle}, \qquad \|F - F_m\|_{L^1(\mathbb{T}^d;\mathbb{H}^{s\times t})} \xrightarrow[m\to\infty]{} 0.$$

Boundary-rank estimate. Fix m and write $\mathcal{S}_m \subset \prod_{\ell=1}^d \{-r_\ell(m),\dots,r_\ell(m)\}$. For all $\boldsymbol{n}$ with $n_\ell \ge r_\ell(m)+1$,

$$T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m) - B_{\boldsymbol{n},m} = \sum_{\boldsymbol{k}\in\mathcal{S}_m} \big(J_{\boldsymbol{n}}^{(-\boldsymbol{k})} - P_{\boldsymbol{n}}^{(-\boldsymbol{k})}\big) \otimes A_{m,\boldsymbol{k}}.$$

For each $\boldsymbol{k} = (k_1,\dots,k_d)$, Lemma 2.36 gives

$$\operatorname{rank}\big(J_{\boldsymbol{n}}^{(-\boldsymbol{k})} - P_{\boldsymbol{n}}^{(-\boldsymbol{k})}\big) \le \sum_{\ell=1}^d 2|k_\ell|\,\frac{N_{\boldsymbol{n}}}{n_\ell}.$$

For a real matrix X and a quaternion matrix Y, we have $\operatorname{rank}_{\mathbb{H}}(X\otimes Y) = \operatorname{rank}_{\mathbb{R}}(X)\operatorname{rank}_{\mathbb{H}}(Y)$. Apply this identity with $X = J_{\boldsymbol{n}}^{(-\boldsymbol{k})} - P_{\boldsymbol{n}}^{(-\boldsymbol{k})}$ and $Y = A_{m,\boldsymbol{k}}$. Together with $\operatorname{rank}_{\mathbb{H}}(A_{m,\boldsymbol{k}}) \le \min\{s,t\}$ and $r_{\boldsymbol{n}} \coloneqq N_{\boldsymbol{n}}\min\{s,t\}$, this gives

$$\frac{\operatorname{rank}_{\mathbb{H}}\big(T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m) - B_{\boldsymbol{n},m}\big)}{r_{\boldsymbol{n}}} \le 2\sum_{\ell=1}^d \frac{1}{n_\ell} \sum_{\boldsymbol{k}\in\mathcal{S}_m} |k_\ell| \xrightarrow[\boldsymbol{n}\to\infty]{} 0 \qquad \text{(for each fixed } m\text{)}.$$

For every scalar $m \in \mathbb{N}$, choose $\boldsymbol{n}_m \in \mathbb{N}^d$ componentwise large enough that $n_\ell \ge r_\ell(m)+1$ for every ℓ and

$$\frac{\operatorname{rank}_{\mathbb{H}}\big(T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m) - B_{\boldsymbol{n},m}\big)}{r_{\boldsymbol{n}}} \le \frac{1}{m} \qquad \text{for every } \boldsymbol{n} \ge \boldsymbol{n}_m.$$

Thus the circulant-versus-polynomial-Toeplitz difference satisfies the rank bound required for a.c.s. convergence, with $c(m) = 1/m \to 0$. *Trace-norm estimate.* By Theorem 4.36 with $p = 1$,

$$\|T_{\boldsymbol{n}}^{(\mathrm{R})}(F) - T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m)\|_{S_1} \;\le\; 4\,\frac{N_{\boldsymbol{n}}}{(2\pi)^d}\,\|F - F_m\|_{L^1(\mathbb{T}^d)}.$$

Dividing by $r_{\boldsymbol{n}} = N_{\boldsymbol{n}}\min\{s,t\}$ gives

$$\frac{\|T_{\boldsymbol{n}}^{(\mathrm{R})}(F) - T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m)\|_{S_1}}{r_{\boldsymbol{n}}} \;\le\; \frac{4}{(2\pi)^d\min\{s,t\}}\,\|F - F_m\|_{L^1(\mathbb{T}^d)} \;=:\; \omega_1(m) \xrightarrow[m\to\infty]{} 0.$$

Assembly of the approximating class. Finally,

$$T_{\boldsymbol{n}}^{(\mathrm{R})}(F) - B_{\boldsymbol{n},m} = \big[T_{\boldsymbol{n}}^{(\mathrm{R})}(F) - T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m)\big] + \big[T_{\boldsymbol{n}}^{(\mathrm{R})}(F_m) - B_{\boldsymbol{n},m}\big].$$

The first bracket satisfies the norm criterion $\mathrm{ACS}_{\mathbb{H}}5$, while the second has normalized rank at most $1/m$. Hence

$$\big\{\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}}\big\}_{m\in\mathbb{N}}$$

is an a.c.s. for $\{T_{\boldsymbol{n}}^{(\mathrm{R})}(F)\}_{\boldsymbol{n}}$. □

Alternative a.c.s.-based proof of the singular-value distribution in Theorem 4.34. Set

$$H(\boldsymbol{\theta}) \coloneqq Z(\boldsymbol{\theta}) + \big(W \circ r_{S_L}\big)(\boldsymbol{\theta})\mathbf{j}.$$

By Proposition 4.27, $H \in L^1$ and

$$T_{\boldsymbol{n}}^{(S_L,S_R)}(F) = T_{\boldsymbol{n}}^{(\mathrm{R})}(H) \qquad \text{for every } \boldsymbol{n}.$$

Theorem 4.47 provides an a.c.s. of circulants from right trigonometric polynomials H_m satisfying

$$H_m \longrightarrow H \qquad \text{in } L^1(\mathbb{T}^d;\mathbb{H}^{s\times t}).$$

Set

$$\widetilde{H}_m(\boldsymbol{\theta}) \coloneqq H_m(-\boldsymbol{\theta}), \qquad \widetilde{H}(\boldsymbol{\theta}) \coloneqq H(-\boldsymbol{\theta}).$$

By construction, $B_{\boldsymbol{n},m} = C_{\boldsymbol{n}}(\widetilde{H}_m)$. Hence Corollary 4.46 gives

$$\{B_{\boldsymbol{n},m}\}_{\boldsymbol{n}} \sim_\sigma G_{\mathrm{R}}[\widetilde{H}_m].$$

Reflection preserves the L^1 norm, and (4.38) yields

$$G_{\mathrm{R}}[\widetilde{H}_m] \longrightarrow G_{\mathrm{R}}[\widetilde{H}] \qquad \text{in } L^1.$$

Therefore $\mathrm{ACS}_{\mathbb{H}}1$ gives

$$\{T_{\boldsymbol{n}}^{(\mathrm{R})}(H)\}_{\boldsymbol{n}} \sim_\sigma G_{\mathrm{R}}[\widetilde{H}].$$

Moreover,

$$G_{\mathrm{R}}[\widetilde{H}](\boldsymbol{\theta}) = G_{\mathrm{R}}[H](-\boldsymbol{\theta}).$$

Since reflection preserves Haar measure, we have

$$\{T_{\boldsymbol{n}}^{(\mathrm{R})}(H)\}_{\boldsymbol{n}} \sim_\sigma G_{\mathrm{R}}[H].$$

Moreover, for a.e. $\boldsymbol{\theta} \in \mathbb{T}^d$, (4.38) and the definition of G_{S_L,S_R} in Theorem 4.33 lead to

$$G_{\mathrm{R}}[H](\boldsymbol{\theta}) = \begin{bmatrix} Z(\boldsymbol{\theta}) & W\big(r_{S_L}(-\boldsymbol{\theta})\big) \\ -\overline{W\big(r_{S_L}(\boldsymbol{\theta})\big)} & \overline{Z(-\boldsymbol{\theta})} \end{bmatrix} = G_{S_L,S_R}(\boldsymbol{\theta}).$$

Therefore

$$\{T_{\boldsymbol{n}}^{(S_L,S_R)}(F)\}_{\boldsymbol{n}} \sim_\sigma G_{S_L,S_R}.$$

This completes the alternative a.c.s.-based proof of the singular-value assertion. The Hermitian eigenvalue assertion follows instead from the exact block-ordered embedding and the Hermitian complex Toeplitz distribution theorem. □

Part II

Structured Approximation Theory

Chapter 5

Approximation of the Hilbert Transform on the Unit Circle

We study pointwise approximation of the circular Hilbert transform by paired Szegő and anti-Szegő rules, following [68]. Singularity subtraction makes the integrand integrable under the stated regularity assumptions, but it does not prevent a quadrature node from coinciding with or approaching the evaluation point.

The construction makes the para-orthogonal parameter depend on that point, keeping both interlaced node sets separated from the singularity. The analysis then combines discrete stability, a continuous-transform bound, and Laurent exactness to obtain the pointwise error estimate. Averaging and the paired half-difference are treated separately: the opposite-error identity gives exact cancellation on its stated Laurent space, whereas its use for general integrands is assessed numerically.

Chapter 2 supplies the Laurent approximation and paired quadrature results. After fixing their local notation, we develop the prescribed-node geometry and the Dini-scale estimates needed for the Hilbert kernel. The numerical examples compare the direct and prescribed formulas, then examine averaging and error indicators for data of different regularity.

5.1 Hilbert transform and numerical obstruction

The Hilbert transform occurs in signal analysis [121], the Benjamin–Ono equation [19, 135], mechanical-vibration diagnostics, and matrix-valued Riemann–Hilbert problems [134]. On the unit circle, it has the form

$$(If)(z) = \frac{1}{\pi}\fint_{\Gamma} \frac{f(t)}{t-z}\,\mathrm{d}t, \qquad z \in \Gamma, \tag{5.1}$$

Here the integral is understood in the Cauchy principal-value sense. The contour $\Gamma = e^{\mathrm{i}\mathbb{T}} = \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle, $\mathrm{i} = \sqrt{-1}$. Angular variables are represented in $[-\pi, \pi)$, with the endpoints identified on $\mathbb{T}$. The function $f : \Gamma \to \mathbb{C}$ is given.

With $t = e^{\mathrm{i}\theta}$, $z = e^{\mathrm{i}\phi}$, and $\theta, \phi \in [-\pi, \pi)$ representing points of $\mathbb{T}$, the identity [109, p. 104]

$$\frac{\mathrm{i}e^{\mathrm{i}\theta}}{e^{\mathrm{i}\theta} - e^{\mathrm{i}\phi}} = \mathrm{i}\left[\frac{e^{(\frac{\theta-\phi}{2})\mathrm{i}}}{e^{(\frac{\theta-\phi}{2})\mathrm{i}} - e^{-(\frac{\theta-\phi}{2})\mathrm{i}}}\right] = \frac{1}{2}\cot\left(\frac{\theta-\phi}{2}\right) + \frac{\mathrm{i}}{2},$$

rewrites the defining integral as

$$(If)(e^{\mathrm{i}\phi}) = \frac{1}{2\pi}\fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right) f(e^{\mathrm{i}\theta})\,\mathrm{d}\theta + \frac{\mathrm{i}}{2\pi}\int_{-\pi}^{\pi} f(e^{\mathrm{i}\theta})\,\mathrm{d}\theta, \quad \phi \in \mathbb{T}. \tag{5.2}$$

From a numerical perspective, the principal-value term

$$(Hf)(\phi) = \frac{1}{2\pi} \fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right) f(e^{\mathrm{i}\theta})\,\mathrm{d}\theta, \quad \phi \in \mathbb{T}, \tag{5.3}$$

is the more delicate component because of the singular Hilbert kernel. The integral in (5.3) is commonly called the circular Hilbert transform. Several numerical methods have been developed for its approximation; see [37, 60, 59, 126, 134, 165].

We approximate (5.3) with Szegő and anti-Szegő quadrature rules adapted to the kernel singularity. Anti-Szegő rules were introduced in [108] as unit-circle counterparts of Laurie's anti-Gauss rule [114]. Their construction and the balanced pairing are recalled in Sections 2.6.2 and 2.7.

The issue here is node placement. For a fixed quadrature parameter, a moving evaluation point can meet a node even though the rule has the required polynomial exactness. Before constructing evaluation-dependent rules, we fix the periodic smoothness and approximation notation used in their analysis.

5.2 Function spaces and approximation

The modulus of smoothness controls the regularized Hilbert kernel, while best Laurent approximation measures the part of the data not removed by exactness. The global Dini hypothesis will be stated explicitly in the continuous-transform and error theorems.

5.2.1 Periodic smoothness

Let $C(\Gamma)$ denote the continuous functions on the unit circle, equipped with the uniform norm

$$\|f\|_\infty = \sup_{z\in\Gamma} |f(z)|.$$

The pullback $f^\sharp(\theta) := f(e^{\mathrm{i}\theta})$ identifies this space isometrically with the space $C_{2\pi}^0(\mathbb{T})$ of continuous 2π-periodic functions. Below we suppress the superscript and use the same letter f for a circle function and its periodic pullback; angular derivatives and smoothness quantities always refer to the pullback.

For $f \in C_{2\pi}^0(\mathbb{T})$, extended 2π-periodically to $\mathbb{R}$, the kth modulus of smoothness is

$$\omega^k(f,t) = \sup_{0<h\le t} \sup_{\theta\in\mathbb{R}} |\Delta_h^k f(\theta)|,$$

where $k \in \mathbb{N}$, $t > 0$, and

$$\Delta_h f(\theta) = f(\theta+h) - f(\theta), \qquad \Delta_h^k f = \Delta_h(\Delta_h^{k-1} f).$$

Equivalently, the inner supremum may be taken over any complete period. We write $\omega(f,t) := \omega^1(f,t)$.

For $r \in \mathbb{N}$, $r \ge 1$, let

$$W_\infty^r(\mathbb{T}) = \left\{ f \in C_{2\pi}^0(\mathbb{T}) : f^{(r-1)} \in AC_{\mathrm{per}}[-\pi,\pi],\ f^{(r)} \in L^\infty(\mathbb{T}) \right\},$$

where $AC_{\mathrm{per}}[-\pi,\pi]$ denotes absolute continuity on a complete period with the periodic endpoint compatibility understood. We abbreviate this space by W^r and equip it with the norm

$$\|f\|_{W^r} = \|f\|_\infty + \|f^{(r)}\|_\infty.$$

The associated K-functional is

$$K^k(f,t) = \inf_{g\in W^k} \left\{ \|f-g\|_\infty + t^k \|g^{(k)}\|_\infty \right\}, \qquad K(f,t) := K^1(f,t).$$

For $f \in C^0_{2\pi}(\mathbb{T})$ it is equivalent to the modulus of smoothness [122, pp. 33–34]:

$$\mathcal{C}^{-1}\omega^k(f,t) \le K^k(f,t) \le \mathcal{C}\omega^k(f,t), \tag{5.4}$$

where $\mathcal{C}$ is a positive constant depending only on k.

5.2.2 Laurent approximation and best error

For $n \ge 0$, let $\Lambda_{-n,n}$ be the symmetric Laurent space introduced in Section 2.6.2. Its elements have the form

$$L_n(z) = \sum_{k=-n}^{n} c_k z^k, \qquad c_k \in \mathbb{C}, \quad z \in \Gamma. \tag{5.5}$$

We use the local abbreviation

$$E_n(f) := E_n^{\Gamma}(f) = \inf_{L_n \in \Lambda_{-n,n}} \|f - L_n\|_\infty.$$

The density statement recalled in Chapter 2 gives $E_n(f) \to 0$ for every $f \in C(\Gamma)$.

Under the substitution $z = e^{i\theta}$ in (5.5), Laurent polynomials correspond to trigonometric polynomials t_n:

$$t_n(\theta) = \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(k\theta) + b_k \sin(k\theta)],$$

with $a_0 = 2c_0 \in \mathbb{C}$, $a_k = c_k + c_{-k} \in \mathbb{C}$, and $b_k = i(c_k - c_{-k}) \in \mathbb{C}$, for each $k \ge 1$. Hence, these polynomials are well suited to approximating periodic functions.

The first estimate below is the standard periodic Jackson-kernel realization: it follows from the kernel and associated direct theorem in [44, Ch. 7, Sec. 2, (2.4) and Theorem 2.2, pp. 203–205], with the derivative term obtained by differentiating the same convolution. For the simultaneous-derivative estimate, combine Garkavi's definition of the simultaneous approximation constant, his inequalities (13) and (22), and Theorem I [74, pp. 103–104, 108, 110].

Proposition 5.1 (Periodic realization and simultaneous approximation). Under the identification above, let $f \in C^0_{2\pi}(\mathbb{T})$. All derivatives below are angular derivatives of the periodic pullback, and $T_m'(\theta) = \frac{d}{d\theta} T_m(e^{i\theta})$. For every $m \ge 1$, there exists $T_m \in \Lambda_{-m,m}$ such that

$$\|f - T_m\|_\infty + (m+1)^{-1}\|T_m'\|_\infty \le \mathcal{C}\,\omega\big(f, (m+1)^{-1}\big).$$

Consequently,

$$E_m(f) \le \mathcal{C}\,\omega\big(f, (m+1)^{-1}\big).$$

If $f \in W^r_\infty(\mathbb{T})$, $r \in \mathbb{N}$, $r \ge 1$, T_m may instead be chosen so that, for $j = 0, 1$,

$$\|f^{(j)} - T_m^{(j)}\|_\infty \le \mathcal{C}(m+1)^{j-r}\|f^{(r)}\|_\infty.$$

The constant depends on r but not on f or m.

5.3 Paired quadrature schemes

We recall the normalized-Lebesgue Szegő/anti-Szegő pair from Chapter 2. The explicit parameter, node grids, and error identities introduced here will be used in the prescribed Hilbert formulas.

5.3.1 Szegő rules

Throughout this subsection the measure is normalized Lebesgue measure, so the monic and orthonormal polynomials on the unit circle are z^j, the para-orthogonal polynomial in the convention of (2.48) is $z^n + \tau$, and all quadrature weights are $1/n$. Consider

$$\mathcal{I}(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{\mathrm{i}\theta})\, \mathrm{d}\theta. \tag{5.6}$$

The corresponding n-point Szegő quadrature rule is

$$\mathcal{I}(f) = \frac{1}{n} \sum_{k=1}^{n} f(z_k) + e_n(f) \coloneqq S_n(f) + e_n(f), \tag{5.7}$$

where $z_k \in \Gamma$ and $e_n(f)$ is the remainder term.

By the normalized-Lebesgue specialization in Corollary 2.69, the rule is exact on the Laurent polynomials of degree at most $n-1$:

$$e_n(f) = 0, \qquad \forall f \in \Lambda_{-(n-1),n-1}. \tag{5.8}$$

Its quadrature nodes z_k are the zeros of the following para-orthogonal polynomial [92, 105].

$$\psi_n(z;\tau) = \Psi_n(z;\tau) = z^n + \tau, \qquad n = 1, 2, \ldots \tag{5.9}$$

Fix $\tau \in \Gamma$. The nodes are given by

$$z_k = e^{\mathrm{i}(\theta + 2k\pi)/n}, \qquad \tau = -e^{\mathrm{i}\theta}, \quad k = 1, \ldots, n$$

and are the eigenvalues of the following unitary upper Hessenberg matrix

$$H_n(\tau) = \begin{bmatrix} \mathbf{0}_{n-1}^T & -\tau \\ I_{n-1} & \mathbf{0}_{n-1} \end{bmatrix} \in \mathbb{C}^{n\times n},$$

where I_k is the identity matrix of order k and $\mathbf{0}_k^T$ is the zero row vector of length k, with $k = 1, 2, \ldots$. This specializes the unitary Hessenberg representation in Definition 2.64. Here the weights are explicitly $1/n$; their general eigenvector formula is (2.52).

For later use, the general prescribed-node formula (2.55) reduces here to

$$\tau = -z_p^n = -e^{\mathrm{i}n\theta_p}, \qquad z_p = e^{\mathrm{i}\theta_p} \in \Gamma. \tag{5.10}$$

Choosing a prescribed Szegő node fixes the parameter and hence the paired anti-Szegő grid. We describe their interlacing with lifted angles, so that the convention also covers the arc crossing the branch cut.

For a cyclically ordered node set, choose real lifts

$$\theta_1 < \theta_2 < \cdots < \theta_n < \theta_1 + 2\pi, \qquad z_k = e^{\mathrm{i}\theta_k},$$

and set $\theta_{n+1} = \theta_1 + 2\pi$. The positively oriented open arc from z_k to z_{k+1} is

$$(z_k, z_{k+1})_+ = \{e^{\mathrm{i}\theta} : \theta_k < \theta < \theta_{k+1}\}.$$

The nodes $\{z_k\}_{k=1}^n$ strictly interlace $\{\widetilde{z}_k\}_{k=1}^n$ if, after a cyclic relabeling, compatible lifts $\widetilde{z}_k = e^{\mathrm{i}\widetilde{\theta}_k}$ satisfy

$$\theta_k < \widetilde{\theta}_k < \theta_{k+1}, \qquad k = 1, \ldots, n.$$

This lifted-angle convention includes the arc crossing the principal branch cut.

For analytic functions, [32] gives sharper error bounds. For continuous functions, Corollary 2.67 gives the local specialization

$$|e_n(f)| \le 2E_{n-1}(f), \qquad f \in C_{2\pi}^0(\mathbb{T}).$$

5.3.2 Anti-Szegő rules

Corollaries 2.75 and 2.76 give the balanced anti-Szegő rule for normalized Lebesgue measure:

$$\mathcal{I}(f) = \frac{1}{n}\sum_{k=1}^{n} f(\tilde{z}_k) + \tilde{e}_n(f) := \tilde{S}_n(f) + \tilde{e}_n(f), \tag{5.11}$$

where the nodes

$$\tilde{z}_k = e^{\mathrm{i}(\theta+(2k+1)\pi)/n} \in \Gamma, \qquad -\tau = e^{\mathrm{i}\theta} \tag{5.12}$$

are the zeros of the polynomial $\psi_n(z;-\tau) = z^n - \tau$ and the remainder term is such that

$$\tilde{e}_n(p) = -e_n(p), \qquad \forall p \in \Lambda_{-n,n}. \tag{5.13}$$

The anti-Szegő formula is exact for Laurent polynomials of degree at most $n-1$, that is,

$$\tilde{S}_n(p) = \mathcal{I}(p), \qquad \forall p \in \Lambda_{-(n-1),n-1}.$$

For every complex Laurent polynomial in $\Lambda_{-n,n}$, identity (5.13) implies

$$\mathcal{I}(p) = \frac{1}{2}[\tilde{S}_n(p) + S_n(p)], \qquad \forall p \in \Lambda_{-n,n} \tag{5.14}$$

for complex-valued p. If p is real-valued on Γ, all three quantities are real and the midpoint identity places $\mathcal{I}(p)$ between the two rule values. This bracketing statement is restricted to the stated Laurent space.

By Corollary 2.76, the anti-Szegő nodes strictly interlace the Szegő nodes and, with the compatible lifts and cyclic labeling fixed above,

$$|\theta_k - \widetilde{\theta}_k| = \frac{\pi}{n}.$$

The corresponding first-kind Chebyshev Gauss/anti-Gauss lift is Theorem 2.77 in Chapter 2.

5.3.3 Averaging and the paired difference

The midpoint identity (5.14) motivates the averaged quadrature formula

$$\hat{\mathcal{S}}_{2n}(f) = \frac{1}{2}[\tilde{S}_n(f) + S_n(f)], \tag{5.15}$$

By (5.14), this average is exact on $\Lambda_{-n,n}$; for more general integrands it may exhibit error cancellation without a theorem of uniform improvement.

The paired-rule half-difference is the computable quantity

$$R_n(f) := \hat{\mathcal{S}}_{2n}(f) - S_n(f) = \frac{1}{2}\big[\widetilde{S}_n(f) - S_n(f)\big]. \tag{5.16}$$

For every $p \in \Lambda_{-n,n}$, identity (5.14) gives the exact relation

$$R_n(p) = \mathcal{I}(p) - S_n(p) = e_n(p).$$

For general integrands, we use $R_n(f)$ as a numerical error indicator. The identity above does not establish asymptotic equivalence to the error or provide general reliability and efficiency constants.

Cost at a fixed total node count. Formula (5.15) uses $2n$ nodes, with explicit weights and locations for the present measure. The following cost comparison instead concerns a different measure, in an integral of the form

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{\mathrm{i}\theta})\,\mathrm{d}\mu(\theta),$$

where μ is real-valued, bounded, and nondecreasing, with infinitely many points of increase in $[-\pi, \pi]$. The nodes and weights must be computed. They are obtained from the eigenvalues of two unitary upper Hessenberg matrices, one for each rule; see [101]. With the QR algorithm in [92], the leading operation count is $\frac{3}{2}n^2$ flops, plus lower-order terms, for constructing the pair, compared with $4n^2$ flops, plus lower-order terms, for a single Szegő formula with the same total number $2n$ of points.

This comparison assumes the stated Hessenberg eigenvalue model and equal total node counts. For the Hilbert transform, the next question is whether the chosen nodes permit stable evaluation of the regularized kernel.

5.4 Pointwise approximation of Hf

We first derive the direct quadrature formulas and exhibit their near-node instability. We then choose the parameter as a function of the evaluation point and establish the discrete, continuous, and error estimates. The construction follows [68, Section 4]; the arguments below retain the periodic conventions and regularized Laurent exactness used in this chapter.

5.4.1 Direct formulas and near-node instability

Singularity subtraction and node separation address different issues. The former controls the integrability of the kernel quotient; the latter prevents its direct evaluation at or too close to the singularity. Whenever the regularized quotient below is integrable—in particular, under the Dini hypothesis used in Theorems 5.4 and 5.6—adding and subtracting $f(e^{\mathrm{i}\phi})$ gives

$$\begin{aligned}(Hf)(\phi) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(e^{\mathrm{i}\theta}) - f(e^{\mathrm{i}\phi})}{\tan\left((\theta-\phi)/2\right)}\,\mathrm{d}\theta + \frac{f(e^{\mathrm{i}\phi})}{2\pi}\fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right)\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(e^{\mathrm{i}\theta}) - f(e^{\mathrm{i}\phi})}{\tan\left((\theta-\phi)/2\right)}\,\mathrm{d}\theta,\end{aligned} \tag{5.17}$$

because

$$\fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right)\mathrm{d}\theta = 0.$$

Set $F(\theta) = f(e^{\mathrm{i}\theta})$. If F is differentiable at ϕ, then

$$\lim_{\theta\to\phi} \frac{F(\theta) - F(\phi)}{\tan((\theta-\phi)/2)} = 2F'(\phi).$$

Thus, if $F \in C^1_{2\pi}$, the quotient in (5.17) has a continuous extension across the diagonal. Under the weaker Dini condition $\int_0^{\pi} \omega(f,u)u^{-1}\,\mathrm{d}u < \infty$, the subtraction instead guarantees local integrability of the quotient for each fixed ϕ; it does not by itself imply a pointwise limit at the diagonal.

Applying (5.7) and (5.11) to the regularized integral, without adapting the parameter to ϕ, gives

$$(H_n f)(\phi) = \frac{1}{n}\sum_{k=1}^{n} \frac{f(z_k) - f(e^{\mathrm{i}\phi})}{\tan\left((\theta_k-\phi)/2\right)}, \qquad z_k = e^{\mathrm{i}\theta_k}, \tag{5.18}$$

and

$$(\tilde{H}_n f)(\phi) = \frac{1}{n}\sum_{k=1}^{n} \frac{f(\tilde{z}_k) - f(e^{\mathrm{i}\phi})}{\tan\left((\tilde{\theta}_k-\phi)/2\right)}, \qquad \tilde{z}_k = e^{\mathrm{i}\tilde{\theta}_k}, \tag{5.19}$$

respectively. With a fixed parameter, neither exactness nor interlacing prevents a moving evaluation point from meeting a node. Coincidence leaves a quotient undefined as written, while a near-coincidence can amplify rounding and data errors. The following example concerns these direct formulas, not the prescribed-node construction introduced next.

Example 5.2. Consider the integral

$$(Hf_0)(\phi) = \frac{1}{2\pi} \fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right) f_0(\theta)\,\mathrm{d}\theta, \quad f_0(\theta) = e^{2\cos\theta}, \qquad \phi \in \mathbb{T}.$$

Table 5.1 reproduces the approximations reported in [68, Table 1], produced by (5.18) and (5.19) for increasing n at $\phi = \pi/16$ and $\phi = \pi/32$. At $n = 16$ and $n = 32$, respectively, the direct Szegő formula suffers a large loss of accuracy because a node is close to the evaluation point. The displayed anti-Szegő values recover only slowly after the corresponding near-collision and do not show monotone error decay over the reported range. These finite data demonstrate severe near-node deterioration; they do not establish nonconvergence of the anti-Szegő sequence. Since the anti-Szegő nodes satisfy $\widetilde{z}_j^n = \tau$, their periodic angular distance from ϕ is

$$\widetilde{d}_n(\phi;\tau) := \min_{1\le j\le n} \mathrm{dist}_{\mathbb{T}}(\phi,\widetilde{\theta}_j) = \frac{1}{n}\,\mathrm{dist}_{\mathbb{T}}(n\phi, \arg\tau),$$
$$\mathrm{dist}_{\mathbb{T}}(a,b) := \min_{\ell\in\mathbb{Z}} |a-b+2\pi\ell|.$$

The added column uses the reference choice $\tau = 1$.

Table 5.1: Direct Szegő and anti-Szegő approximations for Example 5.2; a node approaching the evaluation point produces the reported near-singular behavior. The last column gives exact distances for the reference grid $\tau = 1$.

ϕ	n	$(H_n f_0)(\phi)$	$(\tilde{H}_n f_0)(\phi)$	$\widetilde{d}_n(\phi;1)$
	4	-1.465154871646335	-1.486566736150962	$\pi/16$
	8	-1.475854994149514	-1.475860803898647	$\pi/16$
	16	-1.129061160339895	-1.475857899024084	$\pi/16$
$\pi/16$	32	-1.475857899024079	-1.302459529681990	0
	64	-1.475857899024079	-1.389158714353031	0
	128	-1.475857899024163	-1.432508306688584	0
	256	-1.475857899024329	-1.454183102856528	0
	4	$-7.489317111388031\times10^{-1}$	$-7.597532661507003\times10^{-1}$	$\pi/32$
	8	$-7.543395598939024\times10^{-1}$	$-7.543424886447505\times10^{-1}$	$\pi/32$
	16	$-7.543410242693274\times10^{-1}$	$-7.543410242693283\times10^{-1}$	$\pi/32$
$\pi/32$	32	$-6.646769412718797\times10^{-1}$	$-7.543410242693276\times10^{-1}$	$\pi/32$
	64	$-7.543410242693263\times10^{-1}$	$-7.095089827706016\times10^{-1}$	0
	128	$-7.543410242693689\times10^{-1}$	$-7.319250035199811\times10^{-1}$	0
	256	$-7.543410242694539\times10^{-1}$	$-7.431330138947536\times10^{-1}$	0

5.4.2 Prescribed-node construction

We now place a prescribed Szegő node $z_p = e^{\mathrm{i}\theta_p}$ at a controlled offset from $e^{\mathrm{i}\phi}$ and set τ by (5.10); see [101]. Since this parameter also fixes the interlacing anti-Szegő grid, choosing θ_p controls both node sets.

With compatible lifted angles and the cyclic extensions $\theta_{n+1} = \theta_1 + 2\pi$ and $\widetilde{\theta}_{n+1} = \widetilde{\theta}_1 + 2\pi$, the Szegő and anti-Szegő nodes satisfy

$$\widetilde{\theta}_{k+1} - \widetilde{\theta}_k = \theta_{k+1} - \theta_k = \frac{2\pi}{n}, \qquad k = 1,\dots,n,$$

and, after the compatible cyclic labeling,

$$|\theta_k - \widetilde{\theta}_k| = \frac{\pi}{n}, \qquad k = 1,\dots,n.$$

Using these spacings, the construction in [68, Section 4.1] sets

$$\theta_p = \phi + \frac{\pi}{4n}.$$

We therefore introduce the approximations

$$(H_n^{\mathrm{pr}} f)(\phi) = \frac{1}{n} \sum_{k=1}^{n} \frac{f(z_k) - f(e^{\mathrm{i}\phi})}{\tan((\theta_k - \phi)/2)}, \qquad z_k = e^{\mathrm{i}\theta_k}, \tag{5.20}$$

and

$$(\widetilde{H}_n^{\mathrm{pr}} f)(\phi) = \frac{1}{n} \sum_{k=1}^{n} \frac{f(\widetilde{z}_k) - f(e^{\mathrm{i}\phi})}{\tan\big((\widetilde{\theta}_k - \phi)/2\big)}, \qquad \widetilde{z}_k = e^{\mathrm{i}\widetilde{\theta}_k}. \tag{5.21}$$

Here one chooses a real lift ϑ_τ such that $-\tau = e^{\mathrm{i}\vartheta_\tau}$ and sets

$$z_k = e^{\mathrm{i}(\vartheta_\tau + 2k\pi)/n}, \quad \tilde{z}_k = e^{\mathrm{i}(\vartheta_\tau + (2k+1)\pi)/n}, \quad \tau = -e^{\mathrm{i}n(\phi + \frac{\pi}{4n})}, \quad k = 1, \dots, n.$$

Changing ϑ_τ by an integer multiple of 2π only cyclically relabels the nodes. Averaging the two formulas defines

$$(\widehat{H}_n^{\mathrm{pr}} f)(\phi) = \frac{1}{2}\big[(H_n^{\mathrm{pr}} f)(\phi) + (\widetilde{H}_n^{\mathrm{pr}} f)(\phi)\big]. \tag{5.22}$$

The half-difference provides an error indicator for the rule (5.20):

$$r_n(f) := \frac{1}{2}\big[\widetilde{H}_n^{\mathrm{pr}} f - H_n^{\mathrm{pr}} f\big]. \tag{5.23}$$

The estimates below concern the two prescribed operators. Their arithmetic mean gives the averaged approximation, whereas the half-difference is the Hilbert-transform counterpart of the indicator in Section 5.3. Its dependence on the evaluation point is understood in (5.23).

5.4.3 Stability and pointwise error analysis

The proof proceeds from discrete stability to a continuous-transform bound, then combines both with regularized Laurent exactness. Throughout, the constants have the independence stated in the corresponding results.

Discrete stability. The first theorem bounds the prescribed operators through their cotangent weights. It is an operator estimate, not a floating-point rounding-error analysis. The averaged operator inherits the same logarithmic bound.

Theorem 5.3. Let $f \in C^0_{2\pi}(\mathbb{T})$ and $n \geq 2$. Then, for each $\phi \in \mathbb{T}$, formulas (5.20) and (5.21) satisfy

$$|(H_n^{\mathrm{pr}} f)(\phi)| \leq \mathcal{C} \log n \, \|f\|_\infty, \tag{5.24}$$

and

$$|(\widetilde{H}_n^{\mathrm{pr}} f)(\phi)| \leq \mathcal{C} \log n \, \|f\|_\infty, \tag{5.25}$$

where $\mathcal{C}$ is a positive constant independent of n, f, and ϕ.

Proof. We prove (5.24) by separating the term involving the value at the evaluation point from the sampled values. By definition, we have

$$(H_n^{\mathrm{pr}} f)(\phi) = -\frac{1}{n} f(e^{\mathrm{i}\phi}) \sum_{k=1}^{n} \cot\left(\frac{\theta_k - \phi}{2}\right) + \frac{1}{n} \sum_{k=1}^{n} \frac{f(z_k)}{\tan\left((\theta_k - \phi)/2\right)},$$

which gives

$$\begin{aligned} |(H_n^{\mathrm{pr}} f)(\phi)| &\leq \frac{1}{n}\left[|f(e^{\mathrm{i}\phi})| \left|\sum_{k=1}^{n} \cot\left(\frac{\theta_k - \phi}{2}\right)\right| + \sum_{k=1}^{n} \left|f(z_k) \cot\left(\frac{\theta_k - \phi}{2}\right)\right|\right] \\ &=: \frac{1}{n}\left[S_1 + S_2\right]. \end{aligned} \tag{5.26}$$

For S_1, the prescribed angles satisfy $\theta_k = \phi + \frac{\pi}{4n} + \frac{2k\pi}{n}$, so

$$\frac{\theta_k - \phi}{2} = \frac{9\pi}{8n} + (k-1)\frac{\pi}{n}.$$

Using the relation

$$\frac{1}{n}\sum_{k=1}^{n} \cot\left(x + \frac{(k-1)\pi}{n}\right) = \cot(nx),$$

we get for $x = \dfrac{9\pi}{8n}$

$$S_1 = |f(e^{\mathrm{i}\phi})| \left|\sum_{k=1}^{n} \cot\left(\frac{\theta_k - \phi}{2}\right)\right| = n|f(e^{\mathrm{i}\phi})| \cot\left(\frac{9\pi}{8}\right) \le n\|f\|_\infty(\sqrt{2}+1). \tag{5.27}$$

We next consider S_2. Using the relation [91, Formula 1.421],

$$\cot(\pi x) = \frac{1}{\pi x} + \frac{2x}{\pi}\sum_{i=1}^{\infty} \frac{1}{x^2 - i^2},$$

we have

$$\begin{aligned}
S_2 &= \sum_{k=1}^{n} \left|f(z_k)\cot\left(\frac{\theta_k - \phi}{2}\right)\right| \\
&\le \|f\|_\infty \sum_{k=1}^{n} \left|\cot\left(\pi\left(\frac{1+8k}{8n}\right)\right)\right| \\
&= \frac{1}{\pi}\|f\|_\infty \sum_{k=1}^{n} \left|\frac{8n}{(1+8k)} + \frac{1+8k}{4n}\sum_{i=1}^{\infty} \frac{1}{\left(\frac{1+8k}{8n}\right)^2 - i^2}\right| \\
&\le \mathcal{C}\|f\|_\infty \left(n\sum_{k=1}^{n}\frac{1}{k} + \sum_{k=1}^{n}\frac{1+8k}{4n}\sum_{i=1}^{\infty}\frac{1}{\left|\left(\frac{1+8k}{8n}\right)^2 - i^2\right|}\right) \\
&\le \mathcal{C}\|f\|_\infty \left(n\log n + \sum_{k=1}^{n}\frac{1+8k}{4n}\left(\sum_{i=1}^{\lfloor\frac{1+8k}{8n}\rfloor}\frac{1}{\left(\frac{1+8k}{8n}\right)^2 - i^2} + \sum_{i=\lfloor\frac{1+8k}{8n}\rfloor+1}^{\infty}\frac{1}{i^2 - \left(\frac{1+8k}{8n}\right)^2}\right)\right) \\
&\le \mathcal{C}\|f\|_\infty\, n\log n
\end{aligned}$$

Combining the preceding estimate with (5.27) in (5.26) proves (5.24). Estimate (5.25) follows analogously.

□

Continuous-transform bound. The logarithmic factor above comes from summing reciprocal angular distances. To control the continuous target, we instead use the Dini integral of the modulus of smoothness. The next theorem states the required global assumption explicitly.

Theorem 5.4. Let $f \in C^0_{2\pi}(\mathbb{T})$ satisfy the global Dini condition

$$\int_0^{\pi} \frac{\omega(f,u)}{u}\,\mathrm{d}u < \infty.$$

Then, uniformly for $\phi \in \mathbb{T}$, we have

$$|(Hf)(\phi)| \le \mathcal{C}\int_0^{\pi} \frac{\omega(f,u)}{u}\,\mathrm{d}u,$$

where $\mathcal{C}$ denotes a positive constant independent of f and ϕ.

Proof. Set $F(\theta) = f(e^{i\theta})$ and extend F 2π-periodically. Translating the angular variable in (5.3) and using the 2π-periodicity of both F and the cotangent kernel gives, for every $\phi \in \mathbb{T}$,

$$(Hf)(\phi) = \frac{1}{2\pi} \fint_{-\pi}^{\pi} \cot\left(\frac{u}{2}\right) F(\phi + u)\, \mathrm{d}u.$$

The principal value of the odd kernel over $[-\pi, \pi]$ is zero. Hence

$$(Hf)(\phi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cot\left(\frac{u}{2}\right) \big(F(\phi + u) - F(\phi)\big)\, \mathrm{d}u.$$

The last integral is absolutely convergent under the Dini hypothesis. Indeed, $|\cot(u/2)| \le \mathcal{C}/|u|$ for $0 < |u| \le \pi$, and periodicity gives

$$\big|F(\phi + u) - F(\phi)\big| \le \omega(f, |u|).$$

Consequently,

$$|(Hf)(\phi)| \le \mathcal{C} \int_{-\pi}^{\pi} \frac{\omega(f, |u|)}{|u|}\, \mathrm{d}u \le \mathcal{C} \int_{0}^{\pi} \frac{\omega(f, u)}{u}\, \mathrm{d}u,$$

with a constant independent of both f and ϕ.

□

Regularized exactness and quadrature error. Before inserting a Laurent approximant into the error, we must check that singularity subtraction preserves the exactness space. The following lemma supplies that step; it is not enough to invoke exactness for the original polynomial without examining the regularized integrand. For fixed ϕ, put $w = e^{i\phi}$ and define the regularized kernel action

$$\mathcal{R}_\phi q(z) = i \frac{z + w}{z - w} \big(q(z) - q(w)\big), \qquad z \in \Gamma \setminus \{w\}.$$

Lemma 5.5 (Regularized Laurent exactness). If $Q \in \Lambda_{-(n-1),n-1}$, then

$$(HQ)(\phi) = (H_n^{\mathrm{pr}} Q)(\phi).$$

Proof. It is enough to consider a monomial. For $q(z) = z^k$ with $1 \le k \le n - 1$,

$$\mathcal{R}_\phi q(z) = i(z + w) \sum_{j=0}^{k-1} z^{k-1-j} w^j \in \Lambda_{0,k}.$$

For $q(z) = z^{-m}$ with $1 \le m \le n - 1$,

$$\mathcal{R}_\phi q(z) = -\frac{i(z + w)}{z^m w^m} \sum_{j=0}^{m-1} z^{m-1-j} w^j \in \Lambda_{-m,0},$$

and the assertion is immediate for $q = 1$. Hence $\mathcal{R}_\phi Q \in \Lambda_{-(n-1),n-1}$. The Szegő exactness (5.8) therefore gives $\mathcal{I}(\mathcal{R}_\phi Q) = S_n(\mathcal{R}_\phi Q)$, which is precisely the claimed identity after singularity subtraction. The boundary monomials $z^{\pm n}$ produce powers outside the exactness space, so they are not included in the lemma. □

We write

$$e_n^{\mathrm{pr}}(f, \phi) := (Hf)(\phi) - (H_n^{\mathrm{pr}} f)(\phi).$$

Theorem 5.6. Let $\phi \in \mathbb{T}$ be fixed, let $n \ge 2$, and let $f \in C_{2\pi}^0(\mathbb{T})$ satisfy

$$\int_{0}^{\pi} \frac{\omega(f, u)}{u}\, \mathrm{d}u < \infty.$$

Then the quadrature error of the prescribed Szegő rule defined in (5.20) satisfies

$$|e_n^{\mathrm{pr}}(f, \phi)| \le \mathcal{C} \left[\log n\, E_{n-1}(f) + \int_{0}^{\pi/n} \frac{\omega(f, u)}{u}\, \mathrm{d}u \right],$$

where $\mathcal{C}$ is a positive constant independent of f, n, and ϕ.

Proof. Let $Q \in \Lambda_{-(n-1),n-1}$ be a Laurent polynomial of best approximation to f. By Lemma 5.5,

$$|e_n^{\mathrm{pr}}(f,\phi)| = |e_n^{\mathrm{pr}}(f-Q,\phi)| \le |(H(f-Q))(\phi)| + |(H_n^{\mathrm{pr}}(f-Q))(\phi)|,$$

and Theorems 5.3 and 5.4 give

$$|e_n^{\mathrm{pr}}(f,\phi)| \le \mathcal{C}\left[\|f-Q\|_\infty + \int_0^\pi \frac{\omega(f-Q,u)}{u}\,\mathrm{d}u + \log n\, E_{n-1}(f)\right].$$

Splitting the Dini integral. Subadditivity and homogeneity of the modulus of smoothness give

$$\begin{aligned}\int_0^\pi \frac{\omega(f-Q,u)}{u}\,\mathrm{d}u &\le \int_0^{\pi/n} \frac{\omega(f-Q,u)}{u}\,\mathrm{d}u + 2\|f-Q\|_\infty \int_{\pi/n}^\pi \frac{\mathrm{d}u}{u} \\ &\le \mathcal{C}\left[\int_0^{\pi/n} \frac{\omega(f,u)}{u}\,\mathrm{d}u + \int_0^{\pi/n} \frac{\omega(Q,u)}{u}\,\mathrm{d}u + \log n\|f-Q\|_\infty\right] \\ &\le \mathcal{C}\left[\int_0^{\pi/n} \frac{\omega(f,u)}{u}\,\mathrm{d}u + \frac{\pi}{n}\|Q'\|_\infty + \log n\|f-Q\|_\infty\right], \end{aligned} \tag{5.28}$$

where $\omega(Q,u)$ is estimated using the first inequality in (5.4) and the definition of the K-functional.

Controlling the derivative of the best approximant. Proposition 5.1, with $m = n-1$, provides $T \in \Lambda_{-(n-1),n-1}$ such that

$$\|f-T\|_\infty + n^{-1}\|T'\|_\infty \le \mathcal{C}\,\omega\left(f,n^{-1}\right).$$

Since Q is a best approximant, $\|f-Q\|_\infty \le \|f-T\|_\infty$. Bernstein's inequality for trigonometric polynomials therefore yields

$$\begin{aligned} n^{-1}\|Q'\|_\infty &\le n^{-1}\|T'\|_\infty + n^{-1}\|(Q-T)'\|_\infty \\ &\le n^{-1}\|T'\|_\infty + \|Q-T\|_\infty \\ &\le n^{-1}\|T'\|_\infty + \|f-Q\|_\infty + \|f-T\|_\infty \\ &\le \mathcal{C}\,\omega\left(f,n^{-1}\right). \end{aligned}$$

The elementary modulus inequality $\omega(f,2t) \le 2\omega(f,t)$ then gives

$$\frac{\pi}{n}\|Q'\|_\infty \le \mathcal{C}\omega\left(f,n^{-1}\right) \le 2\mathcal{C}\omega\left(f,(2n)^{-1}\right) \le 4\mathcal{C}\int_{1/(2n)}^{1/n} \frac{\omega(f,t)}{t}\,\mathrm{d}t.$$

Substituting this estimate into (5.28) proves the assertion. □

Remark 5.7. Dini convergence. For $n \ge 2$, the same regularized-exactness argument and Theorems 5.3 and 5.4 give the same error estimate for the anti-Szegő rule (5.21). Consequently, for fixed $\phi \in \mathbb{T}$, the prescribed rules converge whenever f satisfies the global Dini condition

$$\int_0^\pi \frac{\omega(f,u)}{u}\,\mathrm{d}u < \infty.$$

Indeed, Proposition 5.1 gives $E_{n-1}(f) \le \mathcal{C}\omega(f,n^{-1})$, and monotonicity of the modulus shows that

$$\tfrac{1}{2}\log n\,\omega(f,n^{-1}) \le \int_{1/n}^{1/\sqrt{n}} \frac{\omega(f,u)}{u}\,\mathrm{d}u \longrightarrow 0.$$

The short Dini integral in Theorem 5.6 also tends to zero.

Sobolev upper rate. If $f \in W^r_\infty(\mathbb{T})$, the simultaneous approximant in Proposition 5.1 instead gives

$$\|f - T_{n-1}\|_\infty \le \mathcal{C} n^{-r} \|f^{(r)}\|_\infty, \qquad \|f' - T'_{n-1}\|_\infty \le \mathcal{C} n^{1-r} \|f^{(r)}\|_\infty.$$

Applying regularized exactness to $f - T_{n-1}$ and splitting its Dini integral at π/n yields

$$g := f - T_{n-1}, \qquad \|g\|_\infty = O(n^{-r}), \qquad \|g'\|_\infty = O(n^{1-r}),$$

and therefore

$$\int_0^{\pi/n} \frac{\omega(g,u)}{u}\,\mathrm{d}u \le \int_0^{\pi/n} \|g'\|_\infty\,\mathrm{d}u = \frac{\pi}{n}\|g'\|_\infty = O(n^{-r}),$$

whereas

$$\int_{\pi/n}^{\pi} \frac{\omega(g,u)}{u}\,\mathrm{d}u \le 2\|g\|_\infty \int_{\pi/n}^{\pi} \frac{\mathrm{d}u}{u} = O(n^{-r}\log n).$$

Combining these estimates with the discrete stability bound gives

$$|e_n^{\mathrm{pr}}(f,\phi)| \le \mathcal{C}\frac{\log n}{n^r}\|f^{(r)}\|_\infty.$$

The anti-Szegő rule satisfies the same upper estimate.

The prescribed geometry, Dini control, and Laurent exactness establish the stated convergence results. We next examine their finite-degree behavior and the cancellation obtained by averaging the two rules.

5.5 Numerical tests

The experiments reported in [68, Section 5] compare the direct and prescribed formulas, examine averaging and paired differences, and then vary the regularity of the data. The tables and figure below reproduce those results; their interpretation is limited to the reported degrees and error quantities.

5.5.1 Protocol and error quantities

The numerical comparison separates the circular Hilbert transform from the nonsingular mean in (5.1):

$$(If)(z) = (Hf)(z) + \mathrm{i}\mathcal{I}(f),$$

Here H and $\mathcal{I}$ are defined in (5.3) and (5.6). In this circle-variable notation, Hf is evaluated at the angle corresponding to z. For the Hilbert component, the reported protocol evaluates (5.20), (5.21), and (5.22) at 10^2 equidistant angles $\phi_i \in [-\pi, \pi]$ and records the discrete errors

$$\epsilon_n(f) = \|Hf - H_n^{\mathrm{pr}} f\|, \qquad \tilde{\epsilon}_n(f) = \|Hf - \tilde{H}_n^{\mathrm{pr}} f\|, \qquad \hat{\epsilon}_n(f) = \|Hf - \hat{H}_n^{\mathrm{pr}} f\|, \tag{5.29}$$

where $\|\cdot\|$ is the maximum over the sampled angles, not the continuous uniform norm. When the exact transform is unavailable, the numerical reference described below replaces it in the reported differences. For the nonsingular mean, the Szegő, anti-Szegő, and averaged rules give the signed errors

$$e_n(f) = \mathcal{I}f - S_n f, \qquad \tilde{e}_n(f) = \mathcal{I}f - \tilde{S}_n f, \qquad \hat{e}_n(f) = \mathcal{I}f - \hat{S}_n f. \tag{5.30}$$

In (5.30), $\hat{S}_n$ denotes the averaged rule $\hat{\mathcal{S}}_{2n}$ of (5.15); the subscript n refers to the size of each member of the pair. Thus the ϵ columns are nonnegative discrete maxima, whereas the e columns retain their signs.

The reference degrees specified in Examples 5.8 and 5.9 are $n = 128$ and $n = 64$, respectively. A larger-degree computation is a numerical reference, not an exact value. No corresponding reference degree is specified for the remaining examples in the source's numerical discussion.

The source reports Matlab R2023b computations in double precision on an Intel Core i7-2600 system running Debian GNU/Linux [68, Section 5].

5.5.2 Effect of prescribed-node separation

The first experiment returns to the diagnostic function used before the prescribed-node construction. Its purpose is a controlled before-and-after comparison at the same evaluation points.

Example 5.8. We return to the circular Hilbert transform of $f_0(\phi) = e^{2\cos\phi}$ from Example 5.2. Table 5.2 gives the prescribed-rule values at $\phi = \pi/16, \pi/32$, for comparison with Table 5.1. The isolated deterioration visible in the direct formulas is absent from the prescribed results at the reported degrees.

Table 5.3 reports the discrete errors using the $n = 128$ computation as reference. It also shows the smaller errors of the average. These finite results illustrate the node construction; the general operator bound is supplied by Theorem 5.3.

Table 5.2: Prescribed-node Szegő and anti-Szegő approximations for Example 5.8. The column headings retain the source notation H_n and $\tilde{H}_n$ for the prescribed operators in this table.

n	ϕ	$(H_n f_0)(\phi)$	$(\tilde{H}_n f_0)(\phi)$
4		-1.622605841221501	-1.329104147077534
8		-1.475904319788829	-1.475811478259103
16		-1.475857899023998	-1.475857899024163
32	$\frac{\pi}{16}$	-1.475857899024072	-1.475857899024079
64		-1.475857899024066	-1.475857899024052
128		-1.475857899024082	-1.475857899024035
256		-1.475857899024136	-1.475857899024142
4		$-8.930293238806029 \times 10^{-1}$	$-6.157479708830708 \times 10^{-1}$
8		$-7.544098378965085 \times 10^{-1}$	$-7.542722106421451 \times 10^{-1}$
16		$-7.543410242694044 \times 10^{-1}$	$-7.543410242692503 \times 10^{-1}$
32	$\frac{\pi}{32}$	$-7.543410242693269 \times 10^{-1}$	$-7.543410242693250 \times 10^{-1}$
64		$-7.543410242693245 \times 10^{-1}$	$-7.543410242693196 \times 10^{-1}$
128		$-7.543410242692896 \times 10^{-1}$	$-7.543410242693149 \times 10^{-1}$
256		$-7.543410242693608 \times 10^{-1}$	$-7.543410242692843 \times 10^{-1}$

Table 5.3: Discrete maximum errors of the prescribed rules and their average in Example 5.8, using the $n = 128$ computation as reference.

n	$\epsilon_n(f)$	$\tilde{\epsilon}_n(f)$	$\hat{\epsilon}_n(f)$
4	1.47×10^{-1}	1.47×10^{-1}	6.66×10^{-5}
8	6.88×10^{-5}	6.88×10^{-5}	2.02×10^{-13}
16	1.77×10^{-13}	1.41×10^{-13}	9.57×10^{-14}

5.5.3 Averaging and paired-rule indicators

The next two examples examine cancellation for smooth nonpolynomial data. The first compares the individual and averaged errors; the second also reports the paired half-differences.

The contour formulas follow the stated circle functions and operator normalization.

Example 5.9. Consider the integral

$$(If_1)(z) = \frac{1}{\pi}\fint_\Gamma \frac{f_1(t)}{t-z}\,\mathrm{d}t, \qquad z \in \Gamma.$$

Here the angular data are $f_1(e^{\mathrm{i}\theta}) = \ln\left(\frac{3}{2} + \frac{e^{\mathrm{i}\theta}}{4} + \frac{e^{-\mathrm{i}\theta}}{4}\right)$ and the corresponding decomposition is

$$(If_1)(e^{\mathrm{i}\phi}) = \frac{1}{2\pi}\fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right) f_1(e^{\mathrm{i}\theta})\,\mathrm{d}\theta + \frac{\mathrm{i}}{2\pi}\int_{-\pi}^{\pi} f_1(e^{\mathrm{i}\theta})\,\mathrm{d}\theta \coloneqq (Hf_1)(e^{\mathrm{i}\phi}) + \mathrm{i}\mathcal{I}(f_1).$$

Table 5.4 reports the discrete Hilbert-transform errors (5.29) and the signed quadrature errors (5.30), using the $n = 64$ computation as reference. The signed quadrature errors have opposite signs, and the averaged errors are smaller at the displayed degrees. The nonnegative Hilbert-error columns alone do not show the signs of the pointwise errors.

Figure 5.1 supplies that pointwise comparison for $n = 4$: the two error curves largely oppose one another, while the averaged curve lies close to zero on the plotted scale. For these nonpolynomial data, the cancellation is observed numerically.

Table 5.4: Errors of the paired quadrature rules and the corresponding Hilbert-transform approximations in Example 5.9.

n	$\epsilon_n(f_1)$	$\tilde{\epsilon}_n(f_1)$	$\hat{\epsilon}_n(f_1)$	$e_n(f_1)$	$\tilde{e}_n(f_1)$	$\hat{e}_n(f_1)$
4	5.69×10^{-4}	5.69×10^{-4}	2.55×10^{-7}	4.33×10^{-4}	-4.33×10^{-4}	-1.88×10^{-7}
8	2.47×10^{-7}	2.47×10^{-7}	9.84×10^{-14}	1.88×10^{-7}	-1.88×10^{-7}	-7.07×10^{-14}
16	9.52×10^{-14}	9.53×10^{-14}	4.91×10^{-15}	7.13×10^{-14}	-7.01×10^{-14}	6.11×10^{-16}

Example 5.10. Consider the integral

$$(If_2)(z) = \frac{1}{\pi}\fint_\Gamma \frac{f_2(t)}{t-z}\,\mathrm{d}t, \qquad z \in \Gamma.$$

The corresponding angular form is

$$(If_2)(e^{\mathrm{i}\phi}) = \frac{1}{2\pi}\fint_{-\pi}^{\pi} \cot\left(\frac{\theta-\phi}{2}\right) f_2(e^{\mathrm{i}\theta})\,\mathrm{d}\theta + \frac{\mathrm{i}}{2\pi}\int_{-\pi}^{\pi} f_2(e^{\mathrm{i}\theta})\,\mathrm{d}\theta,$$

with $f_2(e^{\mathrm{i}\theta}) = \ln\left(5 + 2e^{\mathrm{i}\theta} + 2e^{-\mathrm{i}\theta}\right)$. Table 5.5 reports the discrete errors and paired differences from [68, Table 5]. The indicator magnitudes are close to the corresponding individual errors at the reported degrees, and averaging reduces the displayed errors.

The source reports the negative of $R_n(f_2)$ under the convention in (5.16); the quadrature-indicator column is therefore labeled $-R_n(f_2)$, with all numerical values retained. Comparisons using the indicator remain restricted to magnitudes. The scalar Hilbert indicator is denoted by $r_n^{\mathrm{src}}(f_2)$. The source does not specify how the angle-dependent function $r_n(f_2)$ is reduced to that scalar, so no pointwise sign relation is inferred.

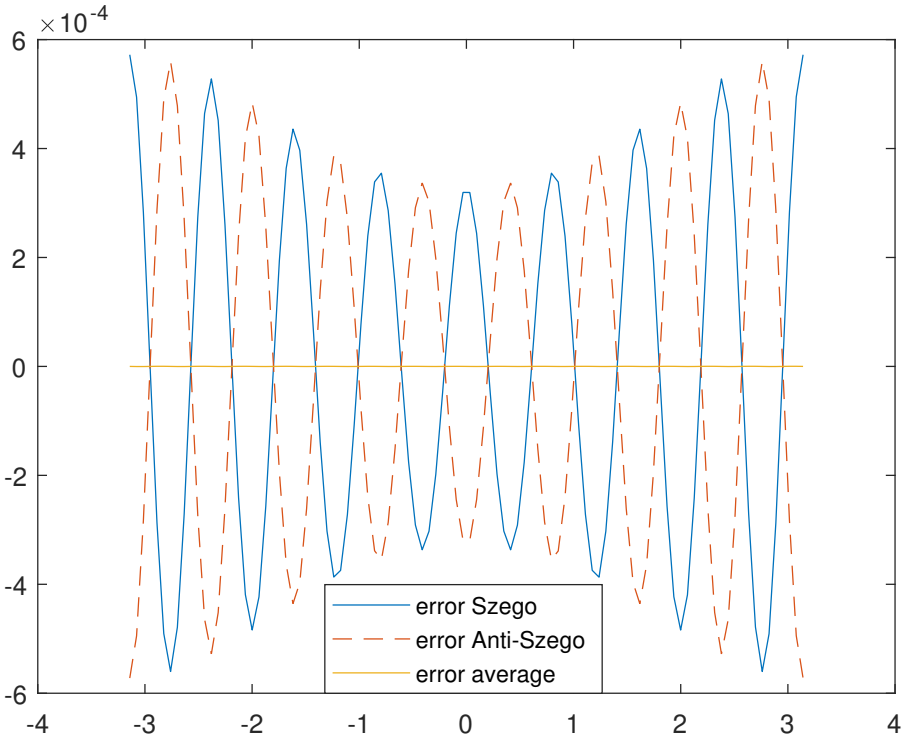


Figure 5.1: Pointwise errors of the two prescribed rules and their average in Example 5.9 for $n = 4$.

Table 5.5: Discrete Hilbert-transform errors, signed quadrature errors, and reported indicators for Example 5.10. The indicator sign convention is discussed in the text.

Hilbert-transform errors and indicator

n	$\epsilon_n(f_2)$	$\tilde{\epsilon}_n(f_2)$	$\hat{\epsilon}_n(f_2)$	$r_n^{\mathrm{src}}(f_2)$
4	3.69×10^{-2}	3.60×10^{-2}	1.28×10^{-3}	3.64×10^{-2}
8	1.25×10^{-3}	1.25×10^{-3}	2.66×10^{-6}	1.25×10^{-3}
16	2.58×10^{-6}	2.58×10^{-6}	2.10×10^{-11}	2.58×10^{-6}
32	2.03×10^{-11}	2.03×10^{-11}	3.45×10^{-14}	2.03×10^{-11}

Quadrature errors and indicator

n	$e_n(f_2)$	$\tilde{e}_n(f_2)$	$\hat{e}_n(f_2)$	$-R_n(f_2)$
4	1.52×10^{-2}	-1.61×10^{-2}	-4.89×10^{-4}	-1.56×10^{-2}
8	4.87×10^{-4}	-4.89×10^{-4}	-9.54×10^{-7}	-4.88×10^{-4}
16	9.54×10^{-7}	-9.54×10^{-7}	-7.28×10^{-12}	-9.54×10^{-7}
32	7.28×10^{-12}	-7.28×10^{-12}	2.22×10^{-16}	-7.28×10^{-12}

5.5.4 Reduced smoothness and observed rates

The final examples use data of finite smoothness. We compare the reported error decay with the preceding smooth examples and examine whether averaging remains beneficial, without inferring a sharp rate from finitely many rows.

Example 5.11. We now consider the finite-smoothness example reported in the source:

$$(If_3)(z) = \frac{1}{\pi} \fint_{\Gamma} \frac{f_3(t)}{t - z}\, \mathrm{d}t = \frac{1}{2\pi} \fint_{-\pi}^{\pi} \cot\left(\frac{\theta - \phi}{2}\right) f_3(e^{\mathrm{i}\theta})\, \mathrm{d}\theta + \frac{\mathrm{i}}{2\pi} \int_{-\pi}^{\pi} f_3(e^{\mathrm{i}\theta})\, \mathrm{d}\theta,$$

where $f_3(e^{i\theta}) = |1 + \frac{e^{i\theta}}{2} + \frac{1}{2e^{i\theta}}|^{5/2}$. The source discusses this function in the class W^2. Table 5.6 shows slower error decay than the preceding smooth examples. The averaged Hilbert-transform error is smaller in every displayed row. For the nonsingular quadrature error, the last row no longer shows improvement over both individual rules. Thus the table does not support an unrestricted improvement claim at the smallest reported error levels.

The source's quadrature-indicator values are again labeled $-R_n(f_3)$, and the reported scalar Hilbert indicator is denoted by $r_n^{\mathrm{src}}(f_3)$, with the same unspecified reduction. At the largest degree, the quadrature indicator's magnitude is substantially smaller than either individual quadrature error, illustrating a limitation of the paired difference in these data.

Table 5.6: Discrete Hilbert-transform errors, signed quadrature errors, and reported indicators for the finite-smoothness data in Example 5.11.

Hilbert-transform errors and indicator

n	$\epsilon_n(f_3)$	$\tilde{\epsilon}_n(f_3)$	$\hat{\epsilon}_n(f_3)$	$r_n^{\mathrm{src}}(f_3)$
4	9.97×10^{-3}	1.00×10^{-2}	1.89×10^{-4}	9.97×10^{-3}
8	1.86×10^{-4}	1.83×10^{-4}	5.31×10^{-6}	1.85×10^{-4}
16	4.64×10^{-6}	4.67×10^{-6}	1.62×10^{-7}	4.64×10^{-6}
32	1.41×10^{-7}	1.41×10^{-7}	5.02×10^{-9}	1.41×10^{-7}
64	3.73×10^{-9}	3.42×10^{-9}	1.57×10^{-10}	3.58×10^{-9}
128	1.16×10^{-10}	1.07×10^{-10}	4.79×10^{-12}	1.12×10^{-10}
256	3.53×10^{-12}	3.51×10^{-12}	8.11×10^{-13}	3.49×10^{-12}

Quadrature errors and indicator

n	$e_n(f_3)$	$\tilde{e}_n(f_3)$	$\hat{e}_n(f_3)$	$-R_n(f_3)$
4	6.34×10^{-3}	-6.46×10^{-3}	-6.03×10^{-5}	-6.40×10^{-3}
8	5.86×10^{-5}	-6.03×10^{-5}	-8.47×10^{-7}	-5.95×10^{-5}
16	8.21×10^{-7}	-8.47×10^{-7}	-1.29×10^{-8}	-8.34×10^{-7}
32	1.25×10^{-8}	-1.29×10^{-8}	-2.02×10^{-10}	-1.27×10^{-8}
64	1.92×10^{-10}	-2.02×10^{-10}	-4.74×10^{-12}	-1.97×10^{-10}
128	1.55×10^{-12}	-4.67×10^{-12}	-1.56×10^{-12}	-3.11×10^{-12}
256	-1.20×10^{-12}	-1.36×10^{-12}	-1.28×10^{-12}	-8.13×10^{-14}

Example 5.12. We test the quadrature schemes on the following integral:

$$(If_4)(z) = \frac{1}{\pi} \fint_\Gamma \frac{f_4(t)}{t - z}\, \mathrm{d}t, \qquad z \in \Gamma,$$

where $f_4(e^{\mathrm{i}\theta}) = \left|\frac{e^{\mathrm{i}\theta} - e^{-\mathrm{i}\theta}}{2\mathrm{i}}\right|^{7/2} \in W^3$. Table 5.7 shows decreasing errors for the individual rules and smaller averaged errors at the reported degrees. The observed decay is compatible with the general Sobolev upper estimate, which need not be attained by every function.

Table 5.7: Discrete Hilbert-transform errors and signed quadrature errors for the individual and averaged rules in Example 5.12.

Hilbert-transform errors

n	$\epsilon_n(f_4)$	$\tilde{\epsilon}_n(f_4)$	$\hat{\epsilon}_n(f_4)$
8	2.55×10^{-3}	2.48×10^{-3}	1.90×10^{-4}
16	1.86×10^{-4}	1.81×10^{-4}	1.61×10^{-5}
32	1.32×10^{-5}	1.34×10^{-5}	1.40×10^{-6}
64	1.15×10^{-6}	1.08×10^{-6}	1.23×10^{-7}
128	1.01×10^{-7}	8.06×10^{-8}	1.01×10^{-8}
256	8.19×10^{-9}	7.85×10^{-9}	1.73×10^{-10}

Quadrature errors

n	$e_n(f_4)$	$\tilde{e}_n(f_4)$	$\hat{e}_n(f_4)$
8	-1.12×10^{-3}	1.21×10^{-3}	4.45×10^{-5}
16	-4.08×10^{-5}	4.45×10^{-5}	1.89×10^{-6}
32	-1.72×10^{-6}	1.89×10^{-6}	8.26×10^{-8}
64	-7.53×10^{-8}	8.26×10^{-8}	3.65×10^{-9}
128	-3.31×10^{-9}	3.65×10^{-9}	1.66×10^{-10}
256	-1.41×10^{-10}	1.67×10^{-10}	1.29×10^{-11}

5.6 Conclusions

The prescribed-node construction ties the para-orthogonal parameter to the evaluation point of the circular Hilbert transform. Its offset separates both interlaced node sets from the singularity while retaining the paired quadrature structure. The resulting logarithmic operator bounds combine with global Dini control and regularized Laurent exactness to give the pointwise error estimate. Simultaneous approximation yields the stated Sobolev upper rate.

The numerical results illustrate the removal of the direct-rule breakdown and substantial cancellation through averaging. They also delimit what can be inferred from the paired difference: the source-reported quadrature indicator is $-R_n$, the scalar reduction defining r_n^{src} is unspecified, and the smallest-error rows do not support universal improvement or indicator reliability. These observations remain separate from the exact opposite-error identity on its Laurent space.

The analysis concerns normalized Lebesgue measure on the unit circle. The next chapter instead studies interpolation at anti-Gauss nodes, where endpoint location, node spacing, Christoffel quantities, and weighted operator estimates replace separation from a moving kernel singularity as the central geometric ingredients.

Chapter 6

Anti-Gauss Lagrange and Jacobi Interpolation

The present chapter examines two aspects of interpolation at anti-Gauss nodes. The first concerns a fixed interpolant: a modified Christoffel–Darboux formula, a barycentric formula, and an orthogonal expansion give three exact representations of the same polynomial. The second concerns a sequence of interpolation operators at anti-Gauss Jacobi nodes: local node estimates and weighted polynomial bounds lead to a logarithmic upper bound for the weighted Lebesgue constants and corresponding approximation estimates.

The first study follows [46]; the Jacobi study follows [47], with the normalizations and parameter restrictions stated below. The finite-precision comparison of the exact formulas is separate from the operator-norm analysis.

In contrast to Chapter 5, the target here is a polynomial interpolant rather than a singular integral. Chapter 2 provides the shared quadrature constructions, perturbation principle, and exactness–stability framework. We retain the two-study organization: representations and their numerical comparison come first, followed by the weighted Jacobi setting, node estimates, interpolation bounds with their proofs, and numerical comparisons.

6.1 Representations of the anti-Gauss Lagrange interpolant

6.1.1 Anti-Gauss nodes and the interpolation problem

We begin with interpolation at the zeros of a fixed anti-Gauss polynomial. The direct nodal product defines the interpolant; the representations developed below express the same polynomial through orthogonal-polynomial data. For Lagrange interpolation [86], the choice of nodes affects both approximation and evaluation. Lebesgue constants describe the operator norms; logarithmic growth is the benchmark in the classical settings discussed in [65, 173, 122]. This motivates the use of nodes associated with orthogonal polynomials. Let $\{p_n\}_n$ be the monic family for the weight w, with squared norms c_n:

$$\langle p_n, p_m \rangle_w = \int_a^b p_n(x)p_m(x)w(x)\,\mathrm{d}x = \begin{cases} 0, & n \neq m, \\ c_n > 0, & n = m. \end{cases}$$

We write $p_n(w,x) = p_n(x)/\sqrt{c_n}$ for the orthonormal counterpart. This distinction will be used in the kernel formulas. Barycentric interpolation is discussed in [20, 95]; Christoffel–Darboux formulas for interpolation at orthogonal-polynomial zeros are discussed in [130].

Laurie's anti-Gauss polynomial [114] is defined from the monic orthogonal family by

$$\tilde{p}_{n+1}(x) := 2p_{n+1}(x) + (a_n - x)p_n(x), \qquad a_n = \frac{\langle xp_n, p_n\rangle_w}{\langle p_n, p_n\rangle_w}. \tag{6.1}$$

This is the equivalent Laurie form (2.62) of the generic anti-Gauss polynomial in Chapter 2. Its zeros are real and distinct and interlace the zeros of p_n; their inclusion in $[a, b]$ depends on the weight. They are the nodes of the anti-Gauss quadrature rule. Extensions of that rule are described in [49, 52, 139, 140, 142, 164], with applications in [48, 50, 53, 68, 70, 141]. Here we use the nodes for interpolation. When f is specified only on $[a, b]$, its values must be available at every interpolation node; the Jacobi part below states the relevant internality restrictions explicitly.

Interpolation processes on the real half-lines are studied in [69]. For the present fixed-degree problem, define the *anti-Gauss Lagrange interpolant* by

$$\tilde{L}_{n+1}(f, x) := \sum_{k=1}^{n+1} \tilde{\ell}_k(x) f(\tilde{x}_k), \qquad \tilde{\ell}_k(x) = \prod_{\substack{j=1 \\ j\neq k}}^{n+1} \frac{x - \tilde{x}_j}{\tilde{x}_k - \tilde{x}_j} = \frac{\tilde{p}_{n+1}(x)}{\tilde{p}'_{n+1}(\tilde{x}_k)(x - \tilde{x}_k)}. \tag{6.2}$$

We derive the kernel, barycentric, and orthogonal-expansion forms in that order, then compare all three with the nodal formula in finite precision. The kernel identities also supply the discrete orthogonality needed for the coefficient formula in the third representation.

6.1.2 Modified Christoffel–Darboux representation

We modify the Christoffel–Darboux kernel so that each fundamental polynomial is a normalized kernel section at an anti-Gauss node. The diagonal values will identify the anti-Gauss Christoffel numbers.

Definition 6.1. Let $\{p_n(w, x)\}_n$ be the sequence of orthonormal polynomials with respect to w, where $p_n(w, x) = \gamma_n x^n + \dots$ and $\gamma_n > 0$. We define the modified Christoffel–Darboux kernel by

$$\tilde{K}_n(w, x, y) := K_n(w, x, y) - \frac{1}{2} p_n(w, x)p_n(w, y), \qquad n \geq 0, \tag{6.3}$$

where

$$K_n(w, x, y) := \sum_{j=0}^{n} p_j(w, x)p_j(w, y)$$

is the Christoffel–Darboux kernel [166, Eq. (3.1.9)].

Proposition 6.2. For every $x \neq y$, the modified Christoffel–Darboux kernel (6.3) can be expressed as

$$\tilde{K}_n(w, x, y) = \frac{p_n(y)\tilde{p}_{n+1}(x) - p_n(x)\tilde{p}_{n+1}(y)}{2c_n(x - y)}, \quad n \geq 0. \tag{6.4}$$

Proof. By [122, Theorem 2.2.3],

$$\begin{aligned} K_n(w, x, y) &= \frac{\gamma_n}{\gamma_{n+1}} \frac{p_{n+1}(w, x)p_n(w, y) - p_n(w, x)p_{n+1}(w, y)}{x - y} \\ &= \frac{p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)}{c_n(x - y)}. \end{aligned} \tag{6.5}$$

By combining (6.3) with the second identity of (6.5), we obtain

$$2c_n(x - y)\tilde{K}_n(w, x, y) = 2\big[p_{n+1}(x)p_n(y) - p_n(x)p_{n+1}(y)\big] - (x - y)p_n(x)p_n(y).$$

The assertion follows by noting that the last term can be written as

$$-(x - y)p_n(x)p_n(y) = \big[(a_n - x) - (a_n - y)\big]p_n(x)p_n(y),$$

and using (6.1). □

For $n \geq 1$, recall the Christoffel numbers of the anti-Gauss formula [132]:

$$\tilde{\lambda}_k := \frac{2c_n}{p_n(\tilde{x}_k)\,\tilde{p}'_{n+1}(\tilde{x}_k)}, \qquad k = 1, \ldots, n+1. \tag{6.6}$$

The same quotient has an algebraic extension to $n = 0$, for which the sole node is indexed by $k = 1$, $\tilde{\lambda}_1 = 2c_0$ and $\tilde{K}_0 = 1/(2c_0)$. Their product still gives the constant cardinal polynomial, but $2c_0$ is not the weight of a one-node quadrature rule preserving the moment of order zero. Whenever these coefficients are identified with anti-Gauss quadrature weights, $n \geq 1$ is understood.

Theorem 6.3. The following identities hold for $n \geq 0$ and $k = 1, \ldots, n+1$,

$$\tilde{\ell}_k(x) = \tilde{\lambda}_k\, \tilde{K}_n(w, x, \tilde{x}_k), \qquad \tilde{\lambda}_k = \big[\tilde{K}_n(w, \tilde{x}_k, \tilde{x}_k)\big]^{-1}, \qquad \text{and} \qquad \tilde{\ell}_k(x) = \frac{\tilde{K}_n(w, x, \tilde{x}_k)}{\tilde{K}_n(w, \tilde{x}_k, \tilde{x}_k)}. \tag{6.7}$$

Consequently, for a function f whose values are available at every interpolation node, the anti-Gauss Lagrange polynomial is given by

$$\tilde{L}_{n+1}(f, x) = \sum_{k=1}^{n+1} \tilde{\lambda}_k \tilde{K}_n(w, x, \tilde{x}_k) f(\tilde{x}_k). \tag{6.8}$$

Proof. We prove (6.7). Evaluating (6.4) at $y = \tilde{x}_k$ and using $\tilde{p}_{n+1}(\tilde{x}_k) = 0$ yields

$$\tilde{K}_n(w, x, \tilde{x}_k) = \frac{p_n(\tilde{x}_k)\tilde{p}_{n+1}(x)}{2c_n(x - \tilde{x}_k)}.$$

Therefore, by (6.6) and (6.2), for $x \neq \tilde{x}_k$ we have $\tilde{\lambda}_k\, \tilde{K}_n(w, x, \tilde{x}_k) = \dfrac{\tilde{p}_{n+1}(x)}{\tilde{p}'_{n+1}(\tilde{x}_k)(x - \tilde{x}_k)} = \tilde{\ell}_k(x)$, which proves the first identity. The second identity follows by evaluating the expression above at $x = \tilde{x}_k$ and using $\tilde{\ell}_k(\tilde{x}_k) = 1$. The last identity of (6.7) follows by combining the first two. Relation (6.8) is a consequence of the first identity. □

The reciprocal diagonal identity will be used again in the Jacobi operator analysis. It connects the interpolation basis with the positive anti-Gauss weights.

6.1.3 Barycentric representation

The barycentric form reorganizes the same interpolant as a quotient of two sums. Its weights are reciprocals of the derivatives of the nodal polynomial, as in the classical formula [20, 95]. Remark 6.5 relates those weights to the Christoffel data.

Theorem 6.4. The anti-Gauss Lagrange interpolant can be written in the following barycentric form for a function f whose values are available at every interpolation node:

$$\tilde{L}_{n+1}(f, x) = \frac{\displaystyle\sum_{k=1}^{n+1} \frac{\omega_k}{x - \tilde{x}_k} f(\tilde{x}_k)}{\displaystyle\sum_{k=1}^{n+1} \frac{\omega_k}{x - \tilde{x}_k}}, \qquad x \neq \tilde{x}_k, \qquad \omega_k := \frac{1}{\tilde{p}'_{n+1}(\tilde{x}_k)}. \tag{6.9}$$

At an interpolation node $x = \tilde{x}_j$, the formula is defined by continuity and equals $f(\tilde{x}_j)$.

Proof. Apply Proposition 2.56 to the nodes $\tilde{x}_1, \ldots, \tilde{x}_{n+1}$, whose nodal polynomial is $\tilde{p}_{n+1}$. The generic barycentric weights are precisely $\omega_k = 1/\tilde{p}'_{n+1}(\tilde{x}_k)$, which gives (6.9). □

Remark 6.5. The barycentric weights ω_k involve the first derivative of $\tilde{p}_{n+1}$. If the anti-Gauss Christoffel numbers are obtained independently, for example from normalized eigenvectors of

the anti-Gauss Jacobi matrix, then (6.6) avoids an explicit evaluation of that derivative and gives

$$\omega_k = \frac{\tilde{\lambda}_k\, p_n(\tilde{x}_k)}{2c_n}.$$

6.1.4 Orthogonal expansion

We now express the interpolant through orthogonal-polynomial coefficients. The modified family below is orthonormal for the discrete anti-Gauss inner product, as the kernel identities will show. Quadrature exactness alone does not establish the coefficient formula.

Discrete orthogonality. Introduce the modified family

$$\phi_j(x) = \begin{cases} p_j(w,x), & j = 0,\dots,n-1, \\ \frac{1}{\sqrt{2}} p_n(w,x), & j = n, \end{cases}$$

and note that the modified Christoffel–Darboux kernel introduced in Definition 6.1 can be expressed as

$$\tilde{K}_n(w,x,y) = \sum_{j=0}^{n} \phi_j(x)\phi_j(y) = \boldsymbol{\phi}(x)\boldsymbol{\phi}(y)^T, \qquad \boldsymbol{\phi}(x) = [\phi_0(x),\dots,\phi_n(x)]. \tag{6.10}$$

Proposition 6.6. For real-valued functions f and g, define the discrete anti-Gauss inner product by

$$\langle f, g\rangle_{\mathrm{AG}} := \sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) g(\tilde{x}_k).$$

Then the family $\{\phi_0,\dots,\phi_n\}$ is orthonormal: $\langle \phi_i, \phi_j\rangle_{\mathrm{AG}} = \delta_{ij}$ for $0 \le i, j \le n$, where δ_{ij} is the Kronecker delta.

Proof. First, combining (6.7) with (6.10) gives

$$\tilde{\ell}_k(x) = \tilde{\lambda}_k \sum_{j=0}^{n} \phi_j(x)\phi_j(\tilde{x}_k), \qquad k = 1,\dots,n+1. \tag{6.11}$$

For each fixed $i = 0,\dots,n$, since $\phi_i \in \mathbb{P}_n$, the Lagrange interpolation formula gives $\phi_i(x) = \sum_{k=1}^{n+1} \tilde{\ell}_k(x)\phi_i(\tilde{x}_k)$, and substituting (6.11) into this identity yields

$$\phi_i(x) = \sum_{j=0}^{n} \phi_j(x) \left(\sum_{k=1}^{n+1} \tilde{\lambda}_k \phi_j(\tilde{x}_k)\phi_i(\tilde{x}_k) \right) = \sum_{j=0}^{n} \phi_j(x) \langle \phi_i, \phi_j\rangle_{\mathrm{AG}}, \qquad i = 0,\dots,n.$$

Then, a comparison of the coefficients gives the assertion. □

Matrix form and expansion coefficients. The discrete orthogonality can be recorded in a square matrix. This also recovers the nodal kernel normalization in (6.7). Define Q of order $n+1$ by

$$Q_{k,j} := \sqrt{\tilde{\lambda}_k}\, \phi_j(\tilde{x}_k), \quad k = 1,\dots,n+1, \quad j = 0,\dots,n. \tag{6.12}$$

By Proposition 6.6, $(Q^TQ)_{ij} = \delta_{ij}$ for $0 \le i, j \le n$, and hence $Q^TQ = I_{n+1} = QQ^T$, where I_{n+1} is the identity matrix of order $n+1$. Thus, Q is an orthogonal matrix. Moreover,

$$\begin{aligned} (QQ^T)_{ik} &= \sum_{j=0}^{n} Q_{i,j} Q_{k,j} \\ &= \sqrt{\tilde{\lambda}_i \tilde{\lambda}_k} \sum_{j=0}^{n} \phi_j(\tilde{x}_i)\phi_j(\tilde{x}_k) \\ &= \sqrt{\tilde{\lambda}_i \tilde{\lambda}_k}\, \tilde{K}_n(w, \tilde{x}_i, \tilde{x}_k), \qquad 1 \le i, k \le n+1. \end{aligned}$$

Then, using $QQ^T = I_{n+1}$, we obtain

$$\tilde{K}_n(w, \tilde{x}_i, \tilde{x}_k) = \frac{\delta_{ik}}{\sqrt{\tilde{\lambda}_k \tilde{\lambda}_i}}, \qquad 1 \le i, k \le n+1. \tag{6.13}$$

Theorem 6.7. The anti-Gauss Lagrange interpolant of a real-valued function f defined at all interpolation nodes admits the orthogonal expansion

$$\tilde{L}_{n+1}(f, x) = \sum_{j=0}^{n} c_j \phi_j(x), \tag{6.14}$$

where

$$c_j := \sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) \phi_j(\tilde{x}_k), \qquad j = 0, \ldots, n. \tag{6.15}$$

Equivalently, with

$$\mathbf{c} := (c_0, \ldots, c_n)^T, \qquad \mathbf{f} := (f(\tilde{x}_1), \ldots, f(\tilde{x}_{n+1}))^T,$$

one has

$$\mathbf{c} = Q^T \tilde{\Lambda}^{1/2} \mathbf{f}, \qquad \tilde{\Lambda} := \operatorname{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{n+1}), \tag{6.16}$$

where Q is given in (6.12).

Proof. Let f be a function, and define $P_f(x) := \sum_{j=0}^{n} c_j \phi_j(x)$, with the coefficients c_j given in (6.15). Then, for every $i = 1, \ldots, n+1$, we have

$$\begin{aligned} P_f(\tilde{x}_i) &= \sum_{j=0}^{n} \left[\sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) \phi_j(\tilde{x}_k) \right] \phi_j(\tilde{x}_i) \\ &= \sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) \left[\sum_{j=0}^{n} \phi_j(\tilde{x}_i) \phi_j(\tilde{x}_k) \right] \\ &= \sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) \tilde{K}_n(w, \tilde{x}_i, \tilde{x}_k), \end{aligned}$$

where the last identity follows from (6.10). By (6.13), $P_f(\tilde{x}_i) = \sum_{k=1}^{n+1} \tilde{\lambda}_k f(\tilde{x}_k) \frac{\delta_{ik}}{\sqrt{\tilde{\lambda}_k \tilde{\lambda}_i}} = f(\tilde{x}_i)$. Thus, $P_f \in \mathbb{P}_n$ interpolates f at the $n+1$ distinct nodes $\tilde{x}_1, \ldots, \tilde{x}_{n+1}$. By uniqueness of polynomial interpolation, we have $P_f(x) = \tilde{L}_{n+1}(f, x)$, which proves (6.14). □

The three structured formulas represent the nodal interpolant exactly.

6.1.5 Finite-precision comparison

The experiment reported in [46, Section 5] compares the four representations for a Chebyshev anti-Gauss node set. The direct product of node differences is compared with the kernel, barycentric, and orthogonal-expansion evaluations. The test records sampled approximation errors and nonfinite outputs; it is not a forward- or backward-error analysis.

Nodes and test functions. For the weight $w(x) = (1 - x^2)^{-1/2}$, use the monic Chebyshev family $p_0(x) = 1$ and $p_n(x) = 2^{1-n} \cos(n \arccos(x))$, $n \ge 1$. The associated anti-Gauss polynomial has the form [48, Theorem 2]

$$\tilde{p}_{n+1}(x) = 2^{1-n} U_{n-1}(x)(x^2 - 1), \qquad U_{n-1}(x) = \frac{\sin(n \arccos x)}{\sin(\arccos x)}, \qquad n \ge 1.$$

The zeros of this polynomial are therefore $\tilde{x}_k = \cos\left((n - k + 1)\frac{\pi}{n}\right)$, $k = 1, \ldots, n+1$. The convergence properties of this node set are analyzed in [123]. The test functions are $f_1(x) = |x| + \frac{x}{2} + x^2$ and $f_2(x) = e^{|x-0.5|^{5/2}}$. For each function, the same nodal data are used in all four representations. Here $n+1$ is the number of interpolation nodes and the interpolant has degree at most n.

Sampled errors. Let $X_{100} = \{x_r\}_{r=1}^{100} \subset (-1,1)$ denote this equidistant test grid, and let $\tilde{L}_{n+1}^{(i)}(f_j,x)$ denote the floating-point value returned by representation i. Then

$$\epsilon_i(f_j) = \max_{x\in X_{100}} \left| f_j(x) - \tilde{L}_{n+1}^{(i)}(f_j,x) \right|, \qquad i \in \{N,K,B,O\}, \quad j = 1,2.$$

The subscripts N, K, B, and O denote the nodal formula (6.2), the kernel formula (6.8), the barycentric formula (6.9), and the orthogonal expansion (6.14), respectively.

Scaling of the barycentric weights. For this Chebyshev case, differentiating the nodal polynomial in the proof of [48, Theorem 2] and taking reciprocal derivatives gives the barycentric weights

$$\omega_1 = \frac{2^{n-2}}{n}(-1)^n, \qquad \omega_k = \frac{2^{n-1}}{n}\cos\big(\pi(n-k+1)\big), \quad k = 2,\dots,n, \qquad \omega_{n+1} = \frac{2^{n-2}}{n}.$$

The endpoint sign $(-1)^n$ follows directly from $\omega_1 = 1/\tilde{p}'_{n+1}(-1)$. The common factor $2^{n-1}/n$ cancels from the numerator and denominator of (6.9). The reported barycentric computation uses this cancellation. The source records `NaN` at $n = 1024$ and 2048 when the common factor is retained. Thus the scaling of the weights is part of the implementation being compared.

Table 6.1: Maximum absolute errors on X_{100} for the direct nodal (N), kernel (K), scaled barycentric (B), and orthogonal-expansion (O) evaluations. The interpolants use $n+1$ nodes. Values are reproduced from [46, Table 1]; `NaN` denotes a nonfinite numerical output.

Errors for f_1

n	$\epsilon_N(f_1)$	$\epsilon_K(f_1)$	$\epsilon_B(f_1)$	$\epsilon_O(f_1)$
32	1.77×10^{-2}	1.77×10^{-2}	1.77×10^{-2}	1.77×10^{-2}
64	7.59×10^{-3}	7.59×10^{-3}	7.59×10^{-3}	7.59×10^{-3}
128	4.66×10^{-3}	4.66×10^{-3}	4.66×10^{-3}	4.66×10^{-3}
256	9.45×10^{-4}	9.45×10^{-4}	9.45×10^{-4}	9.45×10^{-4}
512	4.83×10^{-4}	4.83×10^{-4}	4.83×10^{-4}	4.83×10^{-4}
1024	NaN	1.05×10^{-4}	1.05×10^{-4}	1.05×10^{-4}
2048	NaN	3.70×10^{-5}	3.70×10^{-5}	3.70×10^{-5}

Errors for f_2

n	$\epsilon_N(f_2)$	$\epsilon_K(f_2)$	$\epsilon_B(f_2)$	$\epsilon_O(f_2)$
32	9.95×10^{-5}	9.95×10^{-5}	9.95×10^{-5}	9.95×10^{-5}
64	1.47×10^{-5}	1.47×10^{-5}	1.47×10^{-5}	1.47×10^{-5}
128	1.35×10^{-6}	1.35×10^{-6}	1.35×10^{-6}	1.35×10^{-6}
256	3.82×10^{-7}	3.82×10^{-7}	3.82×10^{-7}	3.82×10^{-7}
512	4.09×10^{-8}	4.09×10^{-8}	4.09×10^{-8}	4.09×10^{-8}
1024	NaN	3.37×10^{-9}	3.37×10^{-9}	3.37×10^{-9}
2048	NaN	1.39×10^{-10}	1.40×10^{-10}	1.39×10^{-10}

In Table 6.1, the direct nodal computation returns `NaN` at $n = 1024$ and 2048, while the three structured evaluations return finite sampled errors. Their reported errors agree closely, although not in every last digit: the barycentric entry for f_2 at $n = 2048$ differs slightly from the kernel and orthogonal-expansion entries. Agreement of these sampled maxima does not measure the rounding error of each evaluation separately.

We now turn from representations of one polynomial to a sequence of weighted interpolation operators. The Chebyshev test above includes endpoint nodes; the weighted Jacobi process below is restricted to internal nodes. Specialize the construction to

$$w = v^{\alpha,\beta}, \qquad \tilde{x}_k = \tilde{x}_{n+1,k}, \qquad \tilde{\lambda}_k = \tilde{\lambda}_{n+1,k},$$

and restore the degree index whenever the interpolation operators form a sequence. Thus the interpolant, fundamental polynomials, Christoffel numbers, and modified kernel below are the Jacobi specializations of (6.2), (6.6), and Definition 6.1.

6.2 Weighted interpolation at anti-Gauss Jacobi nodes

6.2.1 Scope of the weighted Jacobi study

For Jacobi weights, we study interpolation at the zeros of $\tilde{p}_{n+1}$ as n grows. Internality ensures that the nodes belong to the domain of the weighted functions. Endpoint location, adjacent spacings, and polynomial estimates then control the interpolation operator; the global zero distribution alone does not provide those bounds.

The Jacobi results follow [47], with the sufficient restriction $\alpha, \beta > -1/2$ and the admissible degrees specified below. We first introduce the weighted spaces and Jacobi formulas, then state the node estimates and the logarithmic weighted Lebesgue bound. Best approximation gives the error estimate, including convergence for the stated Sobolev classes.

The anti-Gauss and classical Gauss–Jacobi bounds use different endpoint-weight windows. Their comparison must therefore distinguish the range of a theorem from the finite errors of a particular example.

6.2.2 Weighted spaces and Jacobi preliminaries

We distinguish the Jacobi weight $v^{\alpha,\beta}$, which determines the nodes, from the weight $u = v^{\gamma,\delta}$, which determines the approximation norm. After defining the function spaces, we recall the Jacobi formulas and specialize the anti-Gauss construction.

Here and below, $\mathcal{C}$ denotes a positive constant that may vary from one formula to another. The notation $\mathcal{C} = \mathcal{C}(a, b, \dots)$ records possible dependence on the parameters $a, b, \dots$; independence from other variables is stated explicitly.

For positive functions $A(n, f, x, \dots)$ and $B(n, f, x, \dots)$, we write $A \asymp B$ if there exists a positive constant $\mathcal{C}$, independent of $n, f, x, \dots$, such that $\mathcal{C}^{-1}B(n, f, x, \dots) \le A(n, f, x, \dots) \le \mathcal{C}B(n, f, x, \dots)$. The unqualified relation $A \sim B$ is reserved for ratio asymptotics $A/B \to 1$. For $\rho, \sigma \in \mathbb{R}$, set

$$v^{\rho,\sigma}(x) := (1-x)^{\rho}(1+x)^{\sigma}, \qquad x \in (-1, 1).$$

For any $z \in \mathbb{R}$ the notation $(z)_n$ denotes the Pochhammer symbol, defined for $n \in \mathbb{N}$ by $(z)_0 := 1$, $(z)_n := \prod_{k=1}^{n}(z + k - 1)$.

Function spaces

The operator result is formulated in a weighted uniform space. The endpoint conditions in the definition of C_u ensure that polynomial best approximation converges in the stated norm for $f \in C_u$, while the weighted Sobolev spaces provide the smoothness scale used in the final error estimate. Define the Jacobi weight u by

$$u(x) = v^{\gamma,\delta}(x), \qquad x \in [-1, 1], \quad \gamma, \delta \ge 0. \tag{6.17}$$

Let C_u denote the space of functions f that are continuous on every compact subinterval of $(-1, 1)$ and for which the product fu admits a continuous extension to $[-1, 1]$. The extension is required to satisfy

$$\lim_{x \to +1} f(x)u(x) = 0, \quad \text{if } \gamma > 0, \qquad \text{and} \qquad \lim_{x \to -1} f(x)u(x) = 0, \quad \text{if } \delta > 0. \tag{6.18}$$

We equip C_u with the norm of this continuous extension:

$$\|f\|_{C_u} := \|fu\|_\infty = \max_{x\in[-1,1]} |(fu)(x)|.$$

Let $\mathbb{P}_n$ be the algebraic polynomials of degree at most n. Define

$$E_n(f)_u := \inf_{P\in\mathbb{P}_n} \|f - P\|_{C_u} \tag{6.19}$$

The endpoint conditions (6.18) ensure the approximation property [122]

$$\lim_{n\to\infty} E_n(f)_u = 0, \quad \forall f \in C_u.$$

In the weighted setting under consideration, the corresponding characterization is [51]

$$\lim_{n\to\infty} E_n(f)_u = 0 \Longleftrightarrow f \in C_u.$$

For smoother functions, let $W_r(u)$ denote the Sobolev-type subspace of C_u of order $r \in \mathbb{N}$, $r \geq 1$, defined by

$$W_r(u) = \left\{ f \in C_u : f^{(r-1)} \in AC((-1,1)), \|f^{(r)}\varphi^r u\|_\infty < \infty \right\},$$

where $\varphi(x) = \sqrt{1-x^2}$ and $AC((-1,1))$ denotes the functions that are absolutely continuous on every compact subinterval of $(-1,1)$. We equip $W_r(u)$ with the norm

$$\|f\|_{W_r(u)} := \|fu\|_\infty + \|f^{(r)}\varphi^r u\|_\infty.$$

The following Favard inequality provides an estimate of $E_n(f)_u$ for functions $f \in W_r(u)$

$$E_n(f)_u \leq \frac{\mathcal{C}}{n^r}\|f\|_{W_r(u)}, \quad \forall f \in W_r(u), \tag{6.20}$$

where $\mathcal{C}$ is independent of n and f.

The approximation error is measured in C_u. The norm of the interpolation operator in that space will be estimated separately through its weighted Lebesgue constant.

Jacobi polynomials

We now specialize the Gauss–Christoffel material from Section 2.6.1. As in the first part of this chapter, p_n is monic and $p_n(v^{\alpha,\beta},\cdot)$ is orthonormal. This differs from Chapter 2, where the monic family is denoted by π_n. The recurrence, endpoint values, spacing estimates, and Christoffel numbers below retain the source notation. Let $v^{\alpha,\beta}$ be a Jacobi weight with exponents $\alpha, \beta > -1$, and let $\{p_n\}_{n=0}^\infty$ be the corresponding sequence of monic orthogonal polynomials on $[-1, 1]$, i.e.,

$$\langle p_n, p_m\rangle_{v^{\alpha,\beta}} = \int_{-1}^{1} p_n(x)p_m(x)v^{\alpha,\beta}(x)\,\mathrm{d}x = \begin{cases} 0, & m \neq n, \\ c_n, & m = n, \end{cases}$$

where

$$c_n = \frac{2^{2n+\alpha+\beta+1}}{2n+\alpha+\beta+1} \cdot \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!\,\Gamma(n+\alpha+\beta+1)} \left[\frac{n!}{(n+\alpha+\beta+1)_n}\right]^2, \tag{6.21}$$

for $n \geq 1$, with $c_0 = b_0$ given by (6.25); Γ is the Gamma function.

This sequence satisfies the three-term recurrence relation [85]

$$\begin{cases} p_{-1}(x) = 0, \quad p_0(x) = 1, \\ p_{n+1}(x) = (x - a_n)p_n(x) - b_n p_{n-1}(x), \quad n = 0, 1, 2, \ldots, \end{cases} \tag{6.22}$$

where the coefficients a_n and b_n are given by

$$a_0 = \frac{\beta - \alpha}{\alpha + \beta + 2}, \tag{6.23}$$

$$a_n = \frac{\beta^2 - \alpha^2}{(2n + \alpha + \beta)(2n + \alpha + \beta + 2)}, \qquad n \geq 1, \tag{6.24}$$

$$b_0 = \frac{2^{\alpha+\beta+1}\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(\alpha + \beta + 2)}, \tag{6.25}$$

$$b_n = \frac{4n(n + \alpha)(n + \beta)(n + \alpha + \beta)}{(2n + \alpha + \beta)^2((2n + \alpha + \beta)^2 - 1)}, \qquad n \geq 2. \tag{6.26}$$

The separate expression for a_0 is the removable extension of the quotient in (6.24); in particular, it covers the case $\alpha + \beta = 0$. The remaining low-degree coefficient is

$$b_1 = \frac{4(1 + \alpha)(1 + \beta)}{(\alpha + \beta + 2)^2(\alpha + \beta + 3)}.$$

This is the removable extension of (6.26) at $n = 1$ when $\alpha + \beta = -1$; it is well defined for every $\alpha, \beta > -1$. The values of p_n at the endpoints ± 1 depend on the degree n and the parameters α and β as follows [166, Chapter IV]:

$$p_n(-1) = \frac{(-2)^n(\beta + 1)_n}{(n + \alpha + \beta + 1)_n}, \quad p_n(1) = \frac{2^n(\alpha + 1)_n}{(n + \alpha + \beta + 1)_n}. \tag{6.27}$$

Let $\{x_{n,k}\}_{k=1}^n$ denote the zeros of p_n in increasing order and adjoin the endpoints:

$$-1 =: x_{n,0} < x_{n,1} < \cdots < x_{n,n} < x_{n,n+1} := 1.$$

Their normalized zero-counting measures converge weakly to the arc-sine measure. Writing $x_{n,k} = \cos\theta_{n,k}$ with $\pi = \theta_{n,0} > \cdots > \theta_{n,n+1} = 0$, the adjacent angular spacings obey

$$\theta_{n,k} - \theta_{n,k+1} \asymp \frac{1}{n}, \qquad k = 0, 1, \ldots, n.$$

The corresponding uniform spacing estimate in the x variable is

$$\Delta_{n,k} := x_{n,k+1} - x_{n,k} \asymp \frac{\sqrt{1 - x_{n,k}^2} + n^{-1}}{n}, \qquad k = 0, \ldots, n, \tag{6.28}$$

where the constants in $\asymp$ may depend on the fixed Jacobi parameters but are independent of n and k. For the angular spacing, see [122, Eq. (2.3.35), p. 136]; see also [166, Chapter VI]. The x-spacing follows from the cosine-difference identity and the adjacent angular estimates, including the endpoint intervals. By Theorem 2.61, the zeros $\{x_{n,k}\}_{k=1}^n$ are the eigenvalues of the Jacobi matrix. With the explicit Jacobi coefficients above, this matrix is

$$J_n = \begin{bmatrix} a_0 & \sqrt{b_1} & & & \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & & \\ & \sqrt{b_2} & a_2 & \ddots & \\ & & \ddots & \ddots & \sqrt{b_{n-1}} \\ & & & \sqrt{b_{n-1}} & a_{n-1} \end{bmatrix}, \tag{6.29}$$

where a_0 is defined in (6.23), the remaining a_j in (6.24), and the b_j in (6.26), with the separate value of b_1 given above.

We also use the orthonormal-polynomial sequence $\{p_n(v^{\alpha,\beta})\}_{n=0}^\infty$ for the weight $v^{\alpha,\beta}$, with positive leading coefficients:

$$p_n(v^{\alpha,\beta}, x) = \gamma_n x^n + \text{lower degree terms}, \quad \gamma_n > 0,$$

with

$$p_n(v^{\alpha,\beta},x)=\frac{p_n(x)}{\sqrt{c_n}},$$

where c_n is given in (6.21) for $n\ge 1$, and $c_0=b_0$. Fix a constant $\eta_0>0$, distinct from the norm c_0 and depending only on the Jacobi parameters, sufficiently small for the endpoint estimates below, and set

$$I_n:=[-1+\eta_0 n^{-2},1-\eta_0 n^{-2}].$$

The global estimate used below is the upper bound [129, 166]

$$|p_n(v^{\alpha,\beta},x)|\le \frac{\mathcal{C}}{\left(\sqrt{1-x}+\frac{1}{n}\right)^{\alpha+\frac12}\left(\sqrt{1+x}+\frac{1}{n}\right)^{\beta+\frac12}},\qquad |x|\le 1.$$

In particular (see, for instance, [166, 129, 122]),

$$|p_n(v^{\alpha,\beta},x)|\le \mathcal{C}\begin{cases}\dfrac{1}{\sqrt{v^{\alpha,\beta}(x)\varphi(x)}} & x\in I_n,\\ n^{\beta+\frac12} & x\in[-1,-1+\eta_0 n^{-2}]\\ n^{\alpha+\frac12} & x\in[1-\eta_0 n^{-2},1]\end{cases}\tag{6.30}$$

where $\varphi(x)=\sqrt{1-x^2}$ and the constants represented by $\mathcal{C}$ may depend on α and β but are independent of n. At the endpoints, the exact values in (6.27) and the normalization in (6.21) give the separate two-sided asymptotics

$$|p_n(v^{\alpha,\beta},-1)|\asymp n^{\beta+\frac12},\qquad |p_n(v^{\alpha,\beta},1)|\asymp n^{\alpha+\frac12}.\tag{6.31}$$

The endpoint comparison is separate from the global upper bound: a uniform lower bound of that form would fail at the polynomial zeros.

The Jacobi specialization of the Christoffel formula in Theorem 2.61 is

$$\lambda_{n,k}=\left[\sum_{j=0}^{n-1}[p_j(v^{\alpha,\beta},x_{n,k})]^2\right]^{-1},\qquad k=1,\dots,n.$$

The same weights are related to the Jacobi matrix by $\lambda_{n,k}=b_0 v_{k,1}^2$, where b_0 is given in (6.25) and $v_{k,1}$ is the first component of a normalized eigenvector corresponding to $x_{n,k}$; see [174, 122]. The following estimate also holds [129, 122]:

$$\lambda_{n,k}\asymp v^{\alpha,\beta}(x_{n,k})\frac{\sqrt{1-x_{n,k}^2}}{n},$$

where the constants in $\asymp$ do not depend on n and k.

Anti-Gauss Jacobi polynomials

The anti-Gauss construction changes the terminal Jacobi coefficient and produces $n+1$ nodes. We first recall its quadrature and interlacing properties, then specify when the nodes are internal. Laurie [114] introduced the monic anti-Gauss polynomials $\{\tilde p_n\}_n$ by

$$\tilde p_{n+1}(x)=p_{n+1}(x)-b_n p_{n-1}(x),\qquad n=1,2,\dots,\tag{6.32}$$

where p_n is the monic Jacobi polynomial in (6.22) and b_n is the recurrence coefficient in (6.26), with the separate value of b_1 given above. This is exactly the monic normalization of the common anti-Gauss construction in (2.61): in the notation of Chapter 2,

$$\pi_n=p_n,\qquad \alpha_n=a_n,\qquad \beta_n^2=b_n.$$

Thus the terminal modification $\beta_n \mapsto \sqrt{2}\,\beta_n$ produces the off-diagonal entry $\sqrt{2b_n}$ in (6.35).

For later notation, write the local Gauss–Jacobi and anti-Gauss rules as

$$G_n(f) = \sum_{k=1}^{n} \lambda_{n,k} f(x_{n,k})$$

and

$$\tilde{G}_{n+1}(f) = \sum_{k=1}^{n+1} \tilde{\lambda}_{n+1,k} f(\tilde{x}_{n+1,k}). \tag{6.33}$$

Theorem 2.71 gives the positivity of the weights and the opposite-error identity

$$\int_{-1}^{1} f(x) v^{\alpha,\beta}(x)\,\mathrm{d}x - \tilde{G}_{n+1}(f) = -\left(\int_{-1}^{1} f(x) v^{\alpha,\beta}(x)\,\mathrm{d}x - G_n(f)\right), \qquad f \in \mathbb{P}_{2n+1}. \tag{6.34}$$

Thus the errors are exact negatives of one another on $\mathbb{P}_{2n+1}$; this identity motivates the name *anti-Gauss rule*. The polynomials $\tilde{p}_{n+1}$ will be called *anti-Gauss Jacobi polynomials*.

Specializing (2.60), the zeros of $\tilde{p}_{n+1}$ are the eigenvalues of the modified Jacobi matrix of order $n+1$

$$\tilde{J}_{n+1} = \begin{bmatrix} J_n & \sqrt{2b_n}\,\mathbf{e}_n \\ \sqrt{2b_n}\,\mathbf{e}_n^T & a_n \end{bmatrix}, \tag{6.35}$$

with $\mathbf{e}_n = (0, 0, \ldots, 1)^T \in \mathbb{R}^n$ and J_n, a_n, and b_n defined in (6.29), (6.24), and (6.26), respectively, with the separate value of b_1 when $n = 1$. The interlacing relation (2.63) becomes

$$\tilde{x}_{n+1,1} < x_{n,1} < \tilde{x}_{n+1,2} < \cdots < \tilde{x}_{n+1,n} < x_{n,n} < \tilde{x}_{n+1,n+1}. \tag{6.36}$$

The anti-Gauss nodes $\tilde{x}_{n+1,k}$ with $k = 2, \ldots, n$ therefore belong to the interval $(-1, 1)$, whereas the first and last nodes may lie outside this interval. More specifically, the fixed-n conditions are stated in [49, Theorem 2.1, p. 20]; a direct verification from the endpoint identities is given below. For $n \geq 1$ and $\alpha, \beta > -1$, the smallest node $\tilde{x}_{n+1,1}$ lies in $[-1, 1]$ if and only if

$$(2\beta+1)n^2 + (2\beta+1)(\alpha+\beta+1)n + \frac{1}{2}(\beta+1)(\alpha+\beta)(\alpha+\beta+1) \geq 0, \tag{6.37}$$

and the largest node $\tilde{x}_{n+1,n+1} \in [-1, 1]$ if and only if

$$(2\alpha+1)n^2 + (2\alpha+1)(\alpha+\beta+1)n + \frac{1}{2}(\alpha+1)(\alpha+\beta)(\alpha+\beta+1) \geq 0. \tag{6.38}$$

Endpoint identities must distinguish one fixed degree from all degrees. For fixed n, equality in (6.37) is equivalent to $\tilde{x}_{n+1,1} = -1$, and equality in (6.38) is equivalent to $\tilde{x}_{n+1,n+1} = 1$. The identity $\tilde{x}_{n+1,1} = -1$ for *every* $n \geq 1$ holds if and only if $\beta = -1/2$ and $\alpha = \pm 1/2$; the identity $\tilde{x}_{n+1,n+1} = 1$ for every $n \geq 1$ holds if and only if $\alpha = -1/2$ and $\beta = \pm 1/2$. In the Gegenbauer case $\alpha = \beta = t > -1$, the condition polynomial factors as $(2t+1)(n+t)(n+t+1)$. Thus all nodes are in the closed interval $[-1, 1]$ exactly when $t \geq -1/2$, and in the open interval $(-1, 1)$ exactly when $t > -1/2$.

For interpolation on C_u, let $F_\beta(n)$ and $F_\alpha(n)$ denote the left-hand sides of (6.37) and (6.38), respectively, and define

$$\mathcal{N}_{\alpha,\beta} := \{n \in \mathbb{N} : n \geq 2,\ F_\alpha(n) > 0,\ F_\beta(n) > 0\}.$$

In the operator estimates below, $\alpha, \beta > -1/2$ are fixed and $n \in \mathcal{N}_{\alpha,\beta}$. This set contains all sufficiently large integers. The strict inequalities exclude endpoint nodes; the non-strict inequalities above characterize inclusion in the closed interval. Constants may depend on the fixed Jacobi and endpoint-weight parameters, but not on n, the node index, the evaluation point, or the interpolated function.

The equivalent Laurie form (2.62), specialized to the Jacobi recurrence, is

$$\tilde{p}_{n+1}(x) = 2p_{n+1}(x) + (a_n - x)p_n(x).$$

With (6.27), this gives the identities

$$\tilde{p}_{n+1}(-1) = \left[-\frac{4(n+\beta+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} + a_n + 1\right] p_n(-1), \tag{6.39}$$
$$\tilde{p}_{n+1}(1) = \left[\frac{4(n+\alpha+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} + a_n - 1\right] p_n(1).$$

Here is the direct finite-degree verification. Set $s = \alpha+\beta$ and $D_n = (2n+s)(2n+s+1)(2n+s+2) > 0$ for $n \geq 1$. Substitution of $a_n = (\beta^2-\alpha^2)/((2n+s)(2n+s+2))$ into (6.39) and factorization give

$$\frac{\tilde{p}_{n+1}(-1)}{p_n(-1)} = -\frac{4F_\beta(n)}{D_n}, \qquad \frac{\tilde{p}_{n+1}(1)}{p_n(1)} = \frac{4F_\alpha(n)}{D_n}.$$

Because the polynomials are monic and their zeros interlace, the first quotient is nonpositive exactly when the smallest anti-Gauss zero is at least -1, and the second is nonnegative exactly when the largest anti-Gauss zero is at most 1. Equality characterizes the respective endpoint zero. This proves (6.37)–(6.38), including their strict and equality cases. Finally, $F_\beta(n)$ vanishes for every integer $n \geq 1$ precisely when it is the zero polynomial in n. Its leading coefficient forces $\beta = -1/2$, and its constant coefficient then forces $(\alpha - 1/2)(\alpha + 1/2) = 0$. Interchanging α and β gives the other all-degree identity. These all-degree conditions are not necessary for an endpoint zero at a single fixed degree. Consequently, if $\alpha, \beta > -1/2$, then

$$\frac{\tilde{p}_{n+1}(1)}{p_n(1)} = \frac{\alpha+\frac{1}{2}}{n} + O(n^{-2}), \qquad \frac{\tilde{p}_{n+1}(-1)}{p_n(-1)} = -\frac{\beta+\frac{1}{2}}{n} + O(n^{-2}).$$

Moreover,

$$\begin{aligned}\tilde{p}_{n+1}(x) &= 2\sqrt{c_{n+1}}\,p_{n+1}(v^{\alpha,\beta},x) + \sqrt{c_n}(a_n - x)p_n(v^{\alpha,\beta},x)\\ &= \sqrt{c_n}\left[2\rho_n\, p_{n+1}(v^{\alpha,\beta},x) + (a_n - x)p_n(v^{\alpha,\beta},x)\right],\end{aligned}$$

where

$$\rho_n := \sqrt{\frac{c_{n+1}}{c_n}} = \sqrt{b_{n+1}} \sim \frac{1}{2}. \tag{6.40}$$

The identity $c_{n+1} = b_{n+1}c_n$ is the standard norm relation for monic orthogonal polynomials. Hence the normalized anti-Gauss polynomials have the form

$$\frac{\tilde{p}_{n+1}(x)}{\sqrt{c_n}} = 2\rho_n p_{n+1}(v^{\alpha,\beta},x) + (a_n - x)p_n(v^{\alpha,\beta},x), \qquad 2\rho_n = 1 + O(n^{-2}). \tag{6.41}$$

This is the form used below for the local estimates. Using the Jacobi derivative identity [166, Chapter IV],

$$p_n'(v^{\alpha,\beta},x) = \sqrt{n(n+\alpha+\beta+1)}\,p_{n-1}(v^{\alpha+1,\beta+1},x),$$

in (6.32), together with $c_m = b_m c_{m-1}$ and $b_n \to 1/4$, gives the two-term derivative bound

$$\frac{|\tilde{p}_{n+1}'(x)|}{\sqrt{c_n}} \leq \mathcal{C}n\left(|p_n(v^{\alpha+1,\beta+1},x)| + |p_{n-2}(v^{\alpha+1,\beta+1},x)|\right), \qquad n \geq 2, \tag{6.42}$$

where $\mathcal{C}$ is independent of n and x.

The eigenvector-weight formula in Theorem 2.71 specializes to

$$\tilde{\lambda}_{n+1,k} = b_0\,\tilde{v}_{k,1}^2, \tag{6.43}$$

where b_0 is defined by (6.25) and $\tilde{v}_{k,1}$ is the first component of the normalized eigenvector associated with the eigenvalue $\tilde{x}_{n+1,k}$. In particular, the weights are strictly positive. Specializing (6.6) to $w = v^{\alpha,\beta}$ gives

$$\tilde{\lambda}_{n+1,k} = \frac{2\|p_n\|_{v^{\alpha,\beta}}^2}{p_n(\tilde{x}_{n+1,k})\,\tilde{p}_{n+1}'(\tilde{x}_{n+1,k})}, \qquad \|p_n\|_{v^{\alpha,\beta}}^2 = c_n,$$

or equivalently

$$\tilde{\lambda}_{n+1,k} = \frac{2\sqrt{c_n}}{p_n(v^{\alpha,\beta}, \tilde{x}_{n+1,k})\, \tilde{p}'_{n+1}(\tilde{x}_{n+1,k})}. \tag{6.44}$$

In the Jacobi notation, the modified kernel is

$$\tilde{K}_n(v^{\alpha,\beta}, x, y) := \sum_{j=0}^{n} p_j(v^{\alpha,\beta}, x) p_j(v^{\alpha,\beta}, y) - \frac{1}{2} p_n(v^{\alpha,\beta}, x) p_n(v^{\alpha,\beta}, y).$$

Equivalently, if

$$K_n(v^{\alpha,\beta}, x, y) := \sum_{j=0}^{n} p_j(v^{\alpha,\beta}, x) p_j(v^{\alpha,\beta}, y),$$

then

$$\tilde{K}_n(v^{\alpha,\beta}, x, y) = \frac{1}{2}\left(K_n(v^{\alpha,\beta}, x, y) + K_{n-1}(v^{\alpha,\beta}, x, y)\right). \tag{6.45}$$

Since the diagonal kernel values are sums of squares,

$$\frac{1}{2} K_n(v^{\alpha,\beta}, x, x) \le \tilde{K}_n(v^{\alpha,\beta}, x, x) \le K_n(v^{\alpha,\beta}, x, x). \tag{6.46}$$

The diagonal comparison supplies the kernel estimate used in the proofs of the derivative and Lebesgue bounds. Those arguments do not invoke Proposition 6.11.

We next state and prove the node and polynomial estimates used by the weighted interpolation argument.

6.2.3 Node geometry and polynomial estimates

The estimates are presented in the order used later: endpoint location, global distribution and adjacent spacing, Christoffel numbers, and the normalized polynomial and its derivative. The distinction between global distribution and local bounds is essential for the Lebesgue estimate. For the remaining analysis, set

$$x_k := x_{n,k}, \quad \lambda_k := \lambda_{n,k}, \quad k = 1, 2, \dots, n,$$

$$\tilde{x}_k := \tilde{x}_{n+1,k}, \quad \tilde{\lambda}_k := \tilde{\lambda}_{n+1,k}, \quad k = 1, 2, \dots, n+1.$$

Define the endpoint-augmented spacings by

$$\Delta\tilde{x}_0 := 1 + \tilde{x}_1, \qquad \Delta\tilde{x}_k := \tilde{x}_{k+1} - \tilde{x}_k, \quad k = 1, \dots, n, \qquad \Delta\tilde{x}_{n+1} := 1 - \tilde{x}_{n+1}. \tag{6.47}$$

The first proposition locates the extreme anti-Gauss zeros on the same endpoint scale as the classical Jacobi zeros, under the stated parameter and strict-internality assumptions.

Proposition 6.8. Assume that $\alpha, \beta > -1/2$ are fixed and that $n \in \mathcal{N}_{\alpha,\beta}$. Then, $\tilde{x}_k \in I_n$ for every $k = 1, \dots, n+1$, where I_n is defined in (6.30). Moreover,

$$1 + \tilde{x}_1 \asymp n^{-2}, \qquad 1 - \tilde{x}_{n+1} \asymp n^{-2}, \tag{6.48}$$

with constants independent of n.

Proof. Under the stated assumptions, it suffices to prove

$$1 + \tilde{x}_1 \asymp \frac{1}{n^2} \qquad \text{and} \qquad 1 - \tilde{x}_{n+1} \asymp \frac{1}{n^2},$$

with constants independent of n. Indeed, choosing the fixed constant η_0 in the definition of I_n sufficiently small makes this implication uniform in n, and hence $\tilde{x}_k \in I_n$, $k = 1, \dots, n+1$.

Endpoint upper bounds. By interlacing and the classical endpoint estimates for the Gauss–Jacobi nodes,

$$1+\tilde{x}_1 < 1+x_1 \le \frac{\mathcal{C}}{n^2}, \qquad 1-\tilde{x}_{n+1} < 1-x_n \le \frac{\mathcal{C}}{n^2}.$$

Lower bound at the left endpoint. It remains to prove the reverse inequalities. By the mean value theorem,

$$\begin{aligned}|\tilde{p}_{n+1}(-1)| &= |\tilde{p}_{n+1}(-1)-\tilde{p}_{n+1}(\tilde{x}_1)| \\ &= (1+\tilde{x}_1)|\tilde{p}'_{n+1}(\zeta)|, \qquad \zeta\in(-1,\tilde{x}_1).\end{aligned}$$

Hence

$$1+\tilde{x}_1 = \frac{|\tilde{p}_{n+1}(-1)|}{|\tilde{p}'_{n+1}(\zeta)|}, \qquad \zeta\in(-1,\tilde{x}_1). \tag{6.49}$$

From (6.39), since $\beta > -1/2$,

$$\frac{\tilde{p}_{n+1}(-1)}{p_n(-1)} = -\frac{\beta+\frac{1}{2}}{n}+O(n^{-2}).$$

Therefore, switching to orthonormal polynomials and using the endpoint asymptotic (6.31),

$$|\tilde{p}_{n+1}(-1)| \asymp \frac{|p_n(-1)|}{n} = \sqrt{c_n}\frac{|p_n(v^{\alpha,\beta},-1)|}{n} \asymp \sqrt{c_n}n^{\beta-\frac{1}{2}}. \tag{6.50}$$

Derivative estimate. Set

$$P_n(x) := p_n(v^{\alpha,\beta},x), \qquad s := \alpha+\beta.$$

From the Jacobi structure relation and (6.32), one has

$$\frac{\tilde{p}_{n+1}(x)}{\sqrt{c_n}} = -\frac{2(1-x^2)}{2n+s+1}P'_n(x)+A_n(x)P_n(x), \tag{6.51}$$

where

$$A_n(x) = x-a_n+\frac{2n(\alpha-\beta-(2n+s)x)}{(2n+s)(2n+s+1)}$$

satisfies

$$|A_n(x)|+|A'_n(x)| \le \frac{\mathcal{C}}{n}, \qquad |x|\le 1.$$

Differentiating (6.51) and using the Jacobi differential equation yields

$$\frac{|\tilde{p}'_{n+1}(x)|}{\sqrt{c_n}} \le \mathcal{C}n|P_n(x)|+\frac{\mathcal{C}}{n}|P'_n(x)|. \tag{6.52}$$

Since the previous upper bound gives $\zeta\in[-1,-1+\mathcal{C}n^{-2}]$, by (6.30),

$$|P_n(\zeta)| \le \mathcal{C}n^{\beta+\frac{1}{2}}.$$

Moreover, by the derivative formula recalled before (6.42) and by (6.30) with parameters $\alpha+1,\beta+1$,

$$|P'_n(\zeta)| \le \mathcal{C}n^{\beta+\frac{5}{2}}.$$

Thus (6.52) gives

$$|\tilde{p}'_{n+1}(\zeta)| \le \mathcal{C}\sqrt{c_n}n^{\beta+\frac{3}{2}}. \tag{6.53}$$

Combining (6.49), (6.50), and (6.53), we obtain

$$1+\tilde{x}_1 \ge \mathcal{C}\frac{\sqrt{c_n}n^{\beta-\frac{1}{2}}}{\sqrt{c_n}n^{\beta+\frac{3}{2}}} = \frac{\mathcal{C}}{n^2}.$$

Together with the already proved upper bound, this gives

$$1 + \tilde{x}_1 \asymp \frac{1}{n^2}.$$

Lower bound at the right endpoint. The argument is analogous. By the mean value theorem,

$$1 - \tilde{x}_{n+1} = \frac{|\tilde{p}_{n+1}(1)|}{|\tilde{p}'_{n+1}(\eta)|}, \qquad \eta \in (\tilde{x}_{n+1}, 1).$$

From (6.39), since $\alpha > -1/2$,

$$\frac{\tilde{p}_{n+1}(1)}{p_n(1)} = \frac{\alpha + \frac{1}{2}}{n} + O(n^{-2}),$$

and hence, by the endpoint asymptotic (6.31),

$$|\tilde{p}_{n+1}(1)| \asymp \sqrt{c_n} n^{\alpha - \frac{1}{2}}.$$

Moreover, since $\eta \in [1 - \mathcal{C}n^{-2}, 1]$, the same argument leading to (6.53) gives

$$|\tilde{p}'_{n+1}(\eta)| \le \mathcal{C}\sqrt{c_n} n^{\alpha + \frac{3}{2}}.$$

Therefore,

$$1 - \tilde{x}_{n+1} \ge \mathcal{C} \frac{\sqrt{c_n} n^{\alpha - \frac{1}{2}}}{\sqrt{c_n} n^{\alpha + \frac{3}{2}}} = \frac{\mathcal{C}}{n^2}.$$

Combining this with the upper bound gives

$$1 - \tilde{x}_{n+1} \asymp \frac{1}{n^2}.$$

Thus $\tilde{x}_k \in I_n$, $k = 1, \dots, n+1$, and the proposition follows. □

The next result states both weak convergence to the arc-sine measure and a uniform two-sided estimate for adjacent node spacings.

Proposition 6.9. Assume that $\alpha, \beta > -1/2$ are fixed and that $n \in \mathcal{N}_{\alpha,\beta}$. Then the anti-Gauss nodes have the same asymptotic arc-sine distribution as the zeros of the Jacobi polynomial p_n, equivalently as the spectra of the Jacobi matrices J_n defined in (6.29).

Moreover, with $\varphi(x) = \sqrt{1 - x^2}$, one has

$$\Delta\tilde{x}_k := \tilde{x}_{k+1} - \tilde{x}_k \asymp \frac{\varphi(\tilde{x}_k)}{n}, \quad k = 1, \dots, n,$$

where the constants in $\asymp$ are independent of n and k.

Proof.
Global zero distribution. Write $\tilde{J}_{n+1}$ as a perturbation of J_{n+1}:

$$\tilde{J}_{n+1} = \begin{bmatrix} J_n & \sqrt{b_n}\mathbf{e}_n \\ \sqrt{b_n}\mathbf{e}_n^T & a_n \end{bmatrix} + \begin{bmatrix} \mathbf{0} & (\sqrt{2}-1)\sqrt{b_n}\mathbf{e}_n \\ (\sqrt{2}-1)\sqrt{b_n}\mathbf{e}_n^T & 0 \end{bmatrix} = J_{n+1} + \Upsilon_{n+1},$$

where $\mathbf{0}$ represents the null matrix of order n.

The eigenvalues of J_{n+1} are the Gauss nodes. Their arc-sine distribution was recalled in Subsection 6.2.2; (6.28) is the separate local spacing estimate. The perturbation Υ_{n+1} is Hermitian and has rank two, with nonzero eigenvalues $\pm(\sqrt{2}-1)\sqrt{b_n}$. Hence

$$\frac{\operatorname{rank}(\Upsilon_{n+1})}{n+1} = \frac{2}{n+1} \longrightarrow 0,$$

so $\{\Upsilon_{n+1}\}_n$ is zero-distributed by Theorem 2.11. Corollary 2.12 therefore transfers the Hermitian eigenvalue distribution of $\{J_{n+1}\}_n$ to $\{\tilde{J}_{n+1}\}_n$, proving the first assertion.

Adjacent-node spacing. Interlacing gives

$$\tilde{x}_{k+1} - \tilde{x}_k \le x_{k+1} - x_{k-1}, \qquad k = 1, \dots, n,$$

where $x_0 = -1$ and $x_{n+1} = 1$. For $2 \le k \le n$, the adjacent Gauss angular spacings imply $\varphi(x_{k-1}) \asymp \varphi(x_k) \asymp \varphi(\tilde{x}_k)$. For $k = 1$, the endpoint estimate in Proposition 6.8 gives $\varphi(\tilde{x}_1) \asymp n^{-1}$. Thus (6.28) yields, also at the two extreme gaps,

$$\Delta\tilde{x}_k \le \mathcal{C}\frac{\varphi(\tilde{x}_k)}{n}, \qquad k = 1, \dots, n.$$

It remains to prove the reverse inequality:

$$\Delta\tilde{x}_k \ge \mathcal{C}\frac{\varphi(\tilde{x}_k)}{n}.$$

Set

$$P_m(x) := p_m(v^{\alpha,\beta}, x) = \frac{p_m(x)}{\sqrt{c_m}}, \qquad v(x) := v^{\alpha,\beta}(x).$$

We shall use the following two estimates:

$$|\tilde{p}_{n+1}(x_k)| \ge \mathcal{C}\sqrt{c_n}\,[v(x_k)]^{-\frac{1}{2}}\varphi(x_k)^{\frac{1}{2}}, \tag{6.54}$$

$$|\tilde{p}'_{n+1}(x)| \le \mathcal{C}\sqrt{c_n}\,n[v(x)\varphi(x)]^{-\frac{1}{2}}, \qquad x \in I_n. \tag{6.55}$$

Polynomial value at a Gauss node. Since $p_n(x_k) = 0$, the recurrences (6.22) and (6.32) give

$$\tilde{p}_{n+1}(x_k) = p_{n+1}(x_k) - b_n p_{n-1}(x_k) = 2p_{n+1}(x_k) = -2b_n p_{n-1}(x_k).$$

Hence, using $c_n = b_n c_{n-1}$,

$$\frac{|\tilde{p}_{n+1}(x_k)|}{\sqrt{c_n}} = 2\sqrt{b_n}\,|P_{n-1}(x_k)|.$$

Furthermore, the Christoffel–Darboux identity at the zero x_k of P_n gives

$$\lambda_k^{-1} = \sum_{j=0}^{n-1} P_j^2(x_k) = \sqrt{b_n}\,\left|P'_n(x_k)P_{n-1}(x_k)\right|.$$

Using the Christoffel number estimate recalled above,

$$\lambda_k \asymp v(x_k)\frac{\varphi(x_k)}{n},$$

together with

$$|P'_n(x_k)| \le \mathcal{C}n[v(x_k)]^{-\frac{1}{2}}\varphi(x_k)^{-\frac{3}{2}},$$

we obtain (6.54).

Derivative bound on the interior interval. We use the same Jacobi structure relation as in (6.51). With $s = \alpha + \beta$,

$$\frac{\tilde{p}_{n+1}(x)}{\sqrt{c_n}} = -\frac{2(1-x^2)}{2n+s+1}P'_n(x) + A_n(x)P_n(x),$$

where

$$A_n(x) = x - a_n + \frac{2n(\alpha - \beta - (2n+s)x)}{(2n+s)(2n+s+1)}$$

satisfies

$$|A_n(x)| + |A'_n(x)| \le \frac{\mathcal{C}}{n}, \qquad |x| \le 1.$$

Differentiating the last identity and using the Jacobi differential equation in the form

$$\big[(1-x^2)P_n'(x)\big]' = (\alpha-\beta+(\alpha+\beta)x)\,P_n'(x) - n(n+\alpha+\beta+1)P_n(x),$$

we obtain

$$\frac{|\tilde{p}_{n+1}'(x)|}{\sqrt{c_n}} \le \mathcal{C}n|P_n(x)| + \frac{\mathcal{C}}{n}|P_n'(x)|.$$

By (6.30),

$$|P_n(x)| \le \mathcal{C}[v(x)\varphi(x)]^{-\frac{1}{2}}, \qquad x \in I_n.$$

Moreover, by the derivative formula recalled before (6.42) and again by (6.30), now with parameters $\alpha+1, \beta+1$,

$$|P_n'(x)| \le \mathcal{C}n[v(x)]^{-\frac{1}{2}}\varphi(x)^{-\frac{3}{2}}.$$

Since $x \in I_n$, one has $\varphi(x) \ge \mathcal{C}/n$, and therefore

$$\frac{1}{n}|P_n'(x)| \le \mathcal{C}[v(x)]^{-\frac{1}{2}}\varphi(x)^{-\frac{3}{2}} \le \mathcal{C}n[v(x)\varphi(x)]^{-\frac{1}{2}}.$$

This proves (6.55).

Lower spacing bound. Apply the mean value theorem. By (6.36),

$$\tilde{x}_k < x_k < \tilde{x}_{k+1}, \qquad k = 1, \dots, n.$$

Hence, for some $\zeta \in (\tilde{x}_k, x_k)$,

$$|\tilde{p}_{n+1}(x_k) - \tilde{p}_{n+1}(\tilde{x}_k)| = (x_k - \tilde{x}_k)|\tilde{p}_{n+1}'(\zeta)|.$$

Since $\tilde{p}_{n+1}(\tilde{x}_k) = 0$, we have

$$\tilde{x}_{k+1} - \tilde{x}_k > x_k - \tilde{x}_k = \frac{|\tilde{p}_{n+1}(x_k)|}{|\tilde{p}_{n+1}'(\zeta)|}.$$

By Proposition 6.8 and the classical endpoint estimates for the Gauss-Jacobi nodes, both $\tilde{x}_k$ and x_k belong to I_n; hence $\zeta \in I_n$. Moreover, by (6.36), (6.28), Proposition 6.8, and the classical endpoint estimates for the Gauss-Jacobi nodes, one has

$$1 \pm \zeta \asymp 1 \pm x_k, \qquad \varphi(\zeta) \asymp \varphi(x_k) \asymp \varphi(\tilde{x}_k),$$

uniformly with respect to n and k. Consequently,

$$v(\zeta) \asymp v(x_k), \qquad \varphi(\zeta) \asymp \varphi(x_k).$$

Using (6.54) and (6.55), we deduce

$$\tilde{x}_{k+1} - \tilde{x}_k \ge \mathcal{C}\frac{[v(x_k)]^{-\frac{1}{2}}\varphi(x_k)^{\frac{1}{2}}}{n[v(\zeta)\varphi(\zeta)]^{-\frac{1}{2}}} \ge \mathcal{C}\frac{\varphi(x_k)}{n} \asymp \mathcal{C}\frac{\varphi(\tilde{x}_k)}{n}.$$

This proves the lower estimate. Combining it with the upper estimate gives

$$\Delta\tilde{x}_k = \tilde{x}_{k+1} - \tilde{x}_k \asymp \frac{\varphi(\tilde{x}_k)}{n}, \qquad k = 1, \dots, n,$$

where the constants in $\asymp$ are independent of n and k. □

Remark 6.10. The density result proved above is classical and it is treated for a general weight in [171]. A general solution to the nonsymmetric perturbation problem for Jacobi sequences is given in [88]; see also [110] for the asymptotic zero distribution of orthogonal polynomials with discontinuously varying recurrence coefficients. Finally, the stronger arcsine spacing result needed below requires a separate argument.

Combining Propositions 6.8 and 6.9 gives the endpoint-compatible estimates

$$\Delta\tilde{x}_k \asymp \frac{\varphi(\tilde{x}_k)}{n}, \qquad k = 1, \dots, n+1. \tag{6.56}$$

The operator analysis also requires Christoffel-number and polynomial bounds, which we state next.

The Christoffel-number comparison follows directly from the exact modified-kernel identity and an all-point diagonal Jacobi kernel estimate. We first establish the kernel estimate.

The diagonal Jacobi kernel estimate. For fixed $\alpha, \beta > -1/2$ and $n \geq 2$,

$$K_n(v^{\alpha,\beta}, t, t) \asymp \frac{n}{v^{\alpha,\beta}(t)\varphi(t)}, \qquad t \in I_n.$$

Proof. Here is a direct justification. Write $v = v^{\alpha,\beta}$ and $P_j = p_j/\sqrt{c_j}$. The global upper estimate for P_j, with its positive endpoint exponents, gives

$$K_n(t,t) = \sum_{j=0}^{n} P_j(t)^2 \leq \frac{\mathcal{C}n}{v(t)\varphi(t)}.$$

For the reverse inequality, expansion in the orthonormal basis and Cauchy–Schwarz give the extremal identity

$$K_n(t,t)^{-1} = \min_{\substack{R \in \mathbb{P}_n \\ R(t)=1}} \int_{-1}^{1} R(x)^2 v(x)\,\mathrm{d}x.$$

Equality is attained by $R(x) = K_n(x,t)/K_n(t,t)$. Put $t = \cos\theta_0$, $a = 2\alpha + 1 > 0$, and $b = 2\beta + 1 > 0$. Choose a fixed positive integer L such that $4L > a + b + 1$, and set $m = \lfloor n/L \rfloor + 1$. Define

$$F_m(\psi) = \left(\frac{\sin(m\psi/2)}{m\sin(\psi/2)}\right)^2, \qquad R(\cos\theta) = \left[\frac{F_m(\theta - \theta_0) + F_m(\theta + \theta_0)}{1 + F_m(2\theta_0)}\right]^L,$$

with the continuous values of F_m at multiples of 2π. The symmetric sum is a polynomial in $\cos\theta$ of degree $m - 1$, so $\deg R \leq L(m-1) \leq n$ and $R(t) = 1$. For $0 \leq \theta \leq \pi$,

$$|R(\cos\theta)| \leq \mathcal{C}(1 + m|\theta - \theta_0|)^{-2L}.$$

For the reflected term this follows from $\operatorname{dist}(\theta + \theta_0, 2\pi\mathbb{Z}) \geq |\theta - \theta_0|$. Moreover,

$$\begin{aligned} v(\cos\theta)\sin\theta &\asymp \theta^a(\pi - \theta)^b \\ &\leq \mathcal{C}(\theta_0 + m^{-1})^a(\pi - \theta_0 + m^{-1})^b(1 + m|\theta - \theta_0|)^{a+b}. \end{aligned}$$

Since $4L > a + b + 1$, integration gives

$$\int_{-1}^{1} R^2 v \leq \frac{\mathcal{C}}{m}(\theta_0 + m^{-1})^a(\pi - \theta_0 + m^{-1})^b.$$

For $t \in I_n$, both θ_0 and $\pi - \theta_0$ are bounded below by c/n. Since $m \asymp n$ for fixed L, the last expression is bounded by $\mathcal{C}v(t)\varphi(t)/n$. The extremal identity proves the lower kernel bound. Finitely many small degrees are absorbed into the constants. □

Proposition 6.11. Let $\{\tilde{x}_k\}_{k=1}^{n+1}$ be the zeros of the polynomial $\tilde{p}_{n+1}$ defined in (6.32), assume that $\alpha, \beta > -1/2$ are fixed and that $n \in \mathcal{N}_{\alpha,\beta}$, and let $\{\tilde{\lambda}_k\}_{k=1}^{n+1}$ be defined in (6.44). Then,

$$\tilde{\lambda}_k \asymp \frac{v^{\alpha,\beta}(\tilde{x}_k)\varphi(\tilde{x}_k)}{n} \asymp v^{\alpha,\beta}(\tilde{x}_k)\Delta\tilde{x}_k, \qquad k = 1, \dots, n+1,$$

where the constants in $\asymp$ are independent of n and k.

Proof. The exact kernel representation and the cardinal identity give

$$\tilde{\lambda}_k = \frac{1}{\tilde{K}_n(v^{\alpha,\beta}, \tilde{x}_k, \tilde{x}_k)}, \qquad \frac{1}{2} K_n(v^{\alpha,\beta}, t, t) \le \tilde{K}_n(v^{\alpha,\beta}, t, t) \le K_n(v^{\alpha,\beta}, t, t).$$

By Proposition 6.8, every anti-Gauss node belongs to I_n. The preceding all-point diagonal kernel estimate therefore gives

$$\tilde{\lambda}_k \asymp \frac{v^{\alpha,\beta}(\tilde{x}_k)\varphi(\tilde{x}_k)}{n}.$$

The spacing estimate (6.56) proves the second comparison. □

Introduce the normalized anti-Gauss Jacobi polynomial

$$\hat{p}_{n+1}(x) := \frac{\tilde{p}_{n+1}(x)}{\sqrt{\langle \tilde{p}_{n+1}, \tilde{p}_{n+1} \rangle_{v^{\alpha,\beta}}}} = \frac{\tilde{p}_{n+1}(x)}{\sqrt{c_{n+1} + b_n^2 c_{n-1}}}. \tag{6.57}$$

The denominator follows from $\tilde{p}_{n+1} = p_{n+1} - b_n p_{n-1}$ and the orthogonality of p_{n+1} and p_{n-1}. In particular, using $c_{n+1} = b_{n+1} c_n$, we shall use throughout the explicit norm equivalence

$$\|\tilde{p}_{n+1}\|_{v^{\alpha,\beta}}^2 = c_{n+1} + b_n^2 c_{n-1} = (b_{n+1} + b_n) c_n \asymp c_n, \qquad \frac{\|\tilde{p}_{n+1}\|_{v^{\alpha,\beta}}}{\sqrt{c_n}} \asymp 1, \tag{6.58}$$

since $b_n \to 1/4$. Here $\|\cdot\|_{v^{\alpha,\beta}}$ denotes the norm induced by the scalar product.

Proposition 6.12. Assume that $\alpha, \beta > -1/2$. Let I_n be as in (6.30), let $u(x) = (1-x)^\gamma (1+x)^\delta$, and set

$$A := \gamma - \frac{\alpha}{2} + \frac{1}{4}, \qquad B := \delta - \frac{\beta}{2} + \frac{1}{4}. \tag{6.59}$$

Then

$$|\hat{p}_{n+1}(x) u(x)| \le \mathcal{C} (1-x)^A (1+x)^B, \qquad x \in I_n, \tag{6.60}$$

where $\mathcal{C}$ is independent of n.

Moreover, if $\tilde{x}_k \in (-1, 1)$ for each $k = 1, \dots, n+1$, then

$$\frac{1}{|\hat{p}'_{n+1}(\tilde{x}_k)| u(\tilde{x}_k)} \le \mathcal{C}\, \Delta\tilde{x}_k\, (1-\tilde{x}_k)^{-A} (1+\tilde{x}_k)^{-B}, \tag{6.61}$$

where $\mathcal{C}$ is independent of n.

Proof. Throughout the proof the parameters α, β are fixed and satisfy $\alpha, \beta > -1/2$. The constants denoted by $\mathcal{C}$ may change from line to line, but are independent of n, of x, and of the node index k.

Set

$$P_m(x) := p_m(v^{\alpha,\beta}, x) = \frac{p_m(x)}{\sqrt{c_m}}, \qquad m \ge 0.$$

Polynomial envelope. To prove (6.60), use the structure relation for monic Jacobi polynomials,

$$(1-x^2) p'_n(x) = -n p_{n+1}(x) + d_n p_n(x) + (n + \alpha + \beta + 1) b_n p_{n-1}(x), \qquad d_n = O(1). \tag{6.62}$$

Since

$$\tilde{p}_{n+1}(x) = p_{n+1}(x) - b_n p_{n-1}(x),$$

relation (6.62) gives

$$\tilde{p}_{n+1}(x) = -\frac{\varphi(x)^2}{n} p'_n(x) + \frac{d_n}{n} p_n(x) + \frac{\alpha + \beta + 1}{n} b_n p_{n-1}(x). \tag{6.63}$$

Dividing by $\sqrt{c_n}$ and using $c_n = b_n c_{n-1}$, we obtain

$$\frac{\tilde{p}_{n+1}(x)}{\sqrt{c_n}} = -\frac{\varphi(x)^2}{n} P_n'(x) + O(n^{-1})P_n(x) + O(n^{-1})P_{n-1}(x). \tag{6.64}$$

We shall use only the upper estimates for Jacobi orthonormal polynomials. Namely, uniformly for $x \in I_n$ and for $m = n, n-1$,

$$|P_m(x)| \le \mathcal{C}[v^{\alpha,\beta}(x)\varphi(x)]^{-1/2}, \tag{6.65}$$

and, uniformly for $x \in I_n$,

$$|P_n'(x)| \le \mathcal{C}n\,[v^{\alpha,\beta}(x)]^{-1/2}\varphi(x)^{-3/2}. \tag{6.66}$$

Therefore,

$$\frac{\varphi(x)^2}{n}|P_n'(x)| \le \mathcal{C}[v^{\alpha,\beta}(x)]^{-1/2}\varphi(x)^{1/2}.$$

Moreover, since $x \in I_n$, one has $\varphi(x) \ge c/n$; hence, by (6.65),

$$\frac{1}{n}|P_m(x)| \le \mathcal{C}[v^{\alpha,\beta}(x)]^{-1/2}\varphi(x)^{1/2}, \qquad m = n, n-1.$$

Substitution in (6.64) yields

$$\left|\frac{\tilde{p}_{n+1}(x)}{\sqrt{c_n}}\right| \le \mathcal{C}[v^{\alpha,\beta}(x)]^{-1/2}\varphi(x)^{1/2}, \qquad x \in I_n. \tag{6.67}$$

By the orthogonality of p_{n+1} and p_{n-1},

$$\|\tilde{p}_{n+1}\|^2_{v^{\alpha,\beta}} = c_{n+1} + b_n^2 c_{n-1}.$$

Since $c_{n+1} = b_{n+1}c_n$ and $c_n = b_n c_{n-1}$, we have

$$\|\tilde{p}_{n+1}\|^2_{v^{\alpha,\beta}} = (b_{n+1} + b_n)c_n \asymp c_n.$$

Consequently, from (6.67),

$$|\hat{p}_{n+1}(x)| \le \mathcal{C}[v^{\alpha,\beta}(x)]^{-1/2}\varphi(x)^{1/2}, \qquad x \in I_n. \tag{6.68}$$

Multiplying by $u(x) = (1-x)^\gamma(1+x)^\delta$ gives

$$|\hat{p}_{n+1}(x)u(x)| \le \mathcal{C}(1-x)^{\gamma-\alpha/2+1/4}(1+x)^{\delta-\beta/2+1/4}.$$

By the definitions

$$A = \gamma - \frac{\alpha}{2} + \frac{1}{4}, \qquad B = \delta - \frac{\beta}{2} + \frac{1}{4},$$

this is exactly (6.60).

Derivative at an anti-Gauss node. Let $\xi = \tilde{x}_k$ be a zero of $\tilde{p}_{n+1}$. Since the nodes are internal and $\alpha, \beta > -1/2$, Proposition 6.8 gives $\xi \in I_n$. Moreover, by interlacing, $P_n(\xi) \ne 0$.

Using the modified Christoffel–Darboux representation, or equivalently (6.44) together with $\tilde{\ell}_k(\tilde{x}_k) = 1$, we have

$$\tilde{K}_n(v^{\alpha,\beta}, \xi, \xi) = \frac{\tilde{p}_{n+1}'(\xi)P_n(\xi)}{2\sqrt{c_n}}. \tag{6.69}$$

On the other hand, by the diagonal comparison (6.46) and the all-point diagonal Jacobi kernel estimate proved above,

$$K_n(v^{\alpha,\beta}, t, t) \asymp \frac{n}{v^{\alpha,\beta}(t)\varphi(t)}, \qquad t \in I_n,$$

we get

$$\tilde{K}_n(v^{\alpha,\beta},\xi,\xi) \geq \mathcal{C}\frac{n}{v^{\alpha,\beta}(\xi)\varphi(\xi)}. \tag{6.70}$$

Furthermore, by (6.65),

$$|P_n(\xi)| \leq \mathcal{C}[v^{\alpha,\beta}(\xi)\varphi(\xi)]^{-1/2}.$$

Taking absolute values in (6.69) and using (6.70), we obtain

$$|\tilde{p}'_{n+1}(\xi)| = \frac{2\sqrt{c_n}\,\tilde{K}_n(v^{\alpha,\beta},\xi,\xi)}{|P_n(\xi)|} \geq \mathcal{C}\sqrt{c_n}\,n[v^{\alpha,\beta}(\xi)\varphi(\xi)]^{-1/2}.$$

Since $\|\tilde{p}_{n+1}\|_{v^{\alpha,\beta}} \asymp \sqrt{c_n}$, this gives

$$|\hat{p}'_{n+1}(\xi)| \geq \mathcal{C}n[v^{\alpha,\beta}(\xi)\varphi(\xi)]^{-1/2}. \tag{6.71}$$

Therefore,

$$\frac{1}{|\hat{p}'_{n+1}(\xi)|u(\xi)} \leq \mathcal{C}\frac{1}{n}\frac{[v^{\alpha,\beta}(\xi)\varphi(\xi)]^{1/2}}{u(\xi)}.$$

Since

$$\frac{[v^{\alpha,\beta}(\xi)\varphi(\xi)]^{1/2}}{u(\xi)} = \varphi(\xi)(1-\xi)^{-A}(1+\xi)^{-B},$$

we obtain

$$\frac{1}{|\hat{p}'_{n+1}(\xi)|u(\xi)} \leq \mathcal{C}\frac{\varphi(\xi)}{n}(1-\xi)^{-A}(1+\xi)^{-B}. \tag{6.72}$$

Estimate (6.56) replaces $\varphi(\tilde{x}_k)/n$ by the endpoint-compatible spacing $\Delta\tilde{x}_k$. Combining it with (6.72), with $\xi = \tilde{x}_k$, gives

$$\frac{1}{|\hat{p}'_{n+1}(\tilde{x}_k)|u(\tilde{x}_k)} \leq \mathcal{C}\Delta\tilde{x}_k(1-\tilde{x}_k)^{-A}(1+\tilde{x}_k)^{-B}, \qquad k = 1,\dots,n+1.$$

This proves (6.61). □

6.2.4 Weighted Lebesgue and approximation bounds

We assemble the local estimates into a bound for the interpolation operator. The nodal formula is suited to the terms away from the evaluation point; the modified kernel controls the nearest-node term. Together they bound the weighted Lebesgue function. For each admissible degree, specialize the interpolant (6.2) to $w = v^{\alpha,\beta}$ and write

$$\tilde{L}_{n+1}(v^{\alpha,\beta},f,x) = \sum_{k=1}^{n+1}\tilde{\ell}_k(x)f(\tilde{x}_k), \tag{6.73}$$

where the fundamental Lagrange polynomials are given by

$$\tilde{\ell}_k(x) = \prod_{\substack{j=1\\ j\neq k}}^{n+1}\frac{(x-\tilde{x}_j)}{(\tilde{x}_k-\tilde{x}_j)} = \frac{\tilde{p}_{n+1}(x)}{\tilde{p}'_{n+1}(\tilde{x}_k)(x-\tilde{x}_k)}. \tag{6.74}$$

Specializing Definition 6.1 and Theorem 6.3 gives the modified Christoffel–Darboux representation

$$\tilde{\ell}_k(x) = \tilde{\lambda}_k\,\tilde{K}_n(v^{\alpha,\beta},x,\tilde{x}_k), \tag{6.75}$$

The norm of the operator $\tilde{L}_{n+1}(v^{\alpha,\beta}) : C_u \to C_u$ is given by

$$\begin{aligned}\tilde{\Lambda}_{n+1}(v^{\alpha,\beta}) &:= \left\|\tilde{L}_{n+1}(v^{\alpha,\beta})\right\|_{C_u\to C_u} = \sup_{\|f\|_{C_u}\leq 1}\left\|\tilde{L}_{n+1}(v^{\alpha,\beta},f)\right\|_{C_u}\\ &= \max_{x\in[-1,1]}\left\{u(x)\sum_{k=1}^{n+1}\frac{|\tilde{\ell}_k(x)|}{u(\tilde{x}_k)}\right\}.\end{aligned} \tag{6.76}$$

The function

$$\tilde{\Lambda}_{n+1}(x) := u(x) \sum_{k=1}^{n+1} \frac{|\tilde{\ell}_k(x)|}{u(\tilde{x}_k)} \tag{6.77}$$

is the *weighted Lebesgue function* associated with the $(n+1)$ nodes of $\tilde{p}_{n+1}$, whereas the quantities $\{\tilde{\Lambda}_{n+1}(v^{\alpha,\beta})\}_n$ are the *weighted Lebesgue constants*.

The behavior of the sequence $\{\tilde{\Lambda}_{n+1}(v^{\alpha,\beta})\}_n$ controls convergence of the interpolation sequence $\{\tilde{L}_{n+1}(v^{\alpha,\beta})\}_n$ in subspaces of C_u. Applying Theorem 2.50 with the identity operator as target gives, for every $f \in C_u$, the weighted projector estimate

$$\left\| f - \tilde{L}_{n+1}(v^{\alpha,\beta}, f) \right\|_{C_u} \le \left(1 + \tilde{\Lambda}_{n+1}(v^{\alpha,\beta})\right) E_n(f)_u, \tag{6.78}$$

with $E_n(f)_u$ defined in (6.19).

For the Lebesgue bound, we use the following elementary summation estimate, proved here for the endpoint-inclusive exponent range needed below. It isolates the estimate for all nodes except the one nearest the evaluation point.

Lemma 6.13. Let $N \ge 2$, and let

$$-1 = z_0 < z_1 < \cdots < z_N < z_{N+1} = 1$$

be such that

$$1 + z_1 \asymp \frac{1}{N^2}, \qquad 1 - z_N \asymp \frac{1}{N^2}, \qquad \Delta z_k^* \asymp \frac{\varphi(z_k)}{N}, \qquad k = 1, \dots, N,$$

where

$$\Delta z_k^* = \begin{cases} z_{k+1} - z_k, & k = 1, \dots, N-1, \\ z_N - z_{N-1}, & k = N. \end{cases}$$

Let $z \in I_N$, and let d be the index of a point z_d closest to z, i.e.,

$$|z - z_d| = \min_{1 \le k \le N} |z - z_k|.$$

Then, for every $A, B \in [0,1]$,

$$\sum_{\substack{k=1 \\ k \ne d}}^{N} \left(\frac{1-z}{1-z_k}\right)^A \left(\frac{1+z}{1+z_k}\right)^B \frac{\Delta z_k^*}{|z - z_k|} \le \mathcal{C} \log N.$$

Here $\mathcal{C} > 0$ is independent of N.

Proof. Set $r_k = (1-z)/(1-z_k)$ and $s_k = (1+z)/(1+z_k)$. If $z_k > z$, then $r_k \ge 1$, $s_k \le 1$, and $r_k^A s_k^B \le r_k$ for $A, B \in [0,1]$. If $z_k < z$, the corresponding bound is $r_k^A s_k^B \le s_k$. Consequently,

$$\frac{r_k^A s_k^B}{|z - z_k|} \le \frac{1}{|z - z_k|} + \begin{cases} (1 - z_k)^{-1}, & z_k > z, \\ (1 + z_k)^{-1}, & z_k < z. \end{cases}$$

It remains to bound the resulting three sums after multiplication by Δz_k^*.

Write $z_k = -\cos\theta_k$, with θ_k increasing. The endpoint assumptions give $\varphi(z_k) \ge c/N$ for all k. Since $|\varphi(z_{k+1})^2 - \varphi(z_k)^2| \le 2\Delta z_k^*$ and $\Delta z_k^* \asymp \varphi(z_k)/N$, consecutive values of $\varphi(z_k)$ are comparable. Integrating $1/\varphi$ over consecutive node intervals then yields $\theta_{k+1} - \theta_k \asymp 1/N$. Together with the endpoint assumptions, this gives

$$\theta_k \asymp k/N, \qquad \pi - \theta_k \asymp (N + 1 - k)/N.$$

Therefore

$$\frac{\Delta z_k^*}{1 + z_k} \le \frac{\mathcal{C}}{k}, \qquad \frac{\Delta z_k^*}{1 - z_k} \le \frac{\mathcal{C}}{N + 1 - k},$$

and both endpoint sums are $O(\log N)$.

For $z = -\cos\theta$, a difference-of-cosines identity and the inequality $\sin((\theta+\theta_k)/2) \geq \frac{1}{2}\sin\theta_k$ and $|\sin((\theta-\theta_k)/2)| \geq |\theta-\theta_k|/\pi$ give

$$\frac{\Delta z_k^*}{|z-z_k|} \leq \frac{\mathcal{C}}{N|\theta-\theta_k|}.$$

When $|k-d| \geq 2$, the nearest-node property and the angular spacing bound make the right-hand side at most $\mathcal{C}/(|k-d|-1)$. For the at most two adjacent indices, the nearest-node property gives $|z-z_k| \geq |z_k - z_d|/2$; comparability of consecutive gaps bounds their contributions by a constant. Thus

$$\sum_{k\neq d} \frac{\Delta z_k^*}{|z-z_k|} \leq \mathcal{C}\log N.$$

Combining the three estimates proves the result, including $A = 1$ and $B = 1$. □

It remains to bound the Lebesgue constants. The following theorem gives a logarithmic upper bound in the stated endpoint-weight window; (6.78) then gives the approximation estimate.

The operator analysis is restricted to $\alpha, \beta > -1/2$. This restriction is used in the two-sided endpoint estimates and hence in the local mesh bounds; strict internality alone does not imply the required n^{-2} endpoint scale. The boundary cases $\alpha = -1/2$ or $\beta = -1/2$ are not included in the following theorem.

Theorem 6.14. Let $v^{\alpha,\beta}$ be a Jacobi weight with fixed $\alpha, \beta > -1/2$, and let $n \in \mathcal{N}_{\alpha,\beta}$. Let u be defined in (6.17). If

$$\max\left\{0, \frac{\alpha}{2} - \frac{1}{4}\right\} \leq \gamma \leq \frac{\alpha}{2} + \frac{3}{4}, \qquad \max\left\{0, \frac{\beta}{2} - \frac{1}{4}\right\} \leq \delta \leq \frac{\beta}{2} + \frac{3}{4}, \tag{6.79}$$

then, for any $f \in C_u$, one has

$$\left\|\tilde{L}_{n+1}(v^{\alpha,\beta}, f)\right\|_{C_u} \leq \mathcal{C}\log n \, \|f\|_{C_u},$$

where $\mathcal{C}$ is independent of n and f.

Proof.

Reduction to the interior interval. By the Remez inequality [122] (see also [67]),

$$\left\|\tilde{L}_{n+1}(v^{\alpha,\beta}, f)\right\|_{C_u} \leq \mathcal{C} \max_{x\in I_n} \left|\tilde{L}_{n+1}(v^{\alpha,\beta}, f, x)u(x)\right|.$$

Here $\mathcal{C} > 0$ is independent of n. Set

$$A := \gamma - \frac{\alpha}{2} + \frac{1}{4}, \qquad B := \delta - \frac{\beta}{2} + \frac{1}{4}.$$

Under assumption (6.79), one has

$$0 \leq A \leq 1, \qquad 0 \leq B \leq 1.$$

Now, by definition (6.73), and denoting by d the index of the node $\tilde{x}_d$ closest to x, we have

$$\left|\tilde{L}_{n+1}(v^{\alpha,\beta}, f, x)u(x)\right| \leq \|fu\|_\infty \left[\sum_{\substack{k=1\\k\neq d}}^{n+1} |\tilde{\ell}_k(x)| \frac{u(x)}{u(\tilde{x}_k)} + |\tilde{\ell}_d(x)| \frac{u(x)}{u(\tilde{x}_d)}\right]. \tag{6.80}$$

Terms away from the nearest node. Since the fundamental polynomials are invariant under normalization, by (6.74) and Proposition 6.12, for $x \in I_n$,

$$\sum_{\substack{k=1\\k\neq d}}^{n+1} |\tilde{\ell}_k(x)| \frac{u(x)}{u(\tilde{x}_k)} = \sum_{\substack{k=1\\k\neq d}}^{n+1} \frac{|\hat{p}_{n+1}(x)|u(x)}{|\hat{p}'_{n+1}(\tilde{x}_k)|u(\tilde{x}_k)|x-\tilde{x}_k|} \leq \mathcal{C} \sum_{\substack{k=1\\k\neq d}}^{n+1} \left(\frac{1-x}{1-\tilde{x}_k}\right)^A \left(\frac{1+x}{1+\tilde{x}_k}\right)^B \frac{\Delta\tilde{x}_k}{|x-\tilde{x}_k|}.$$

Propositions 6.8 and 6.9 provide the hypotheses of Lemma 6.13, with $N = n+1$. For the last node, the endpoint gap $\Delta\tilde{x}_{n+1}$ used above and the adjacent-node gap in that lemma are comparable by (6.56) and the local spacing estimates. Hence

$$\sum_{\substack{k=1\\k\neq d}}^{n+1} |\tilde{\ell}_k(x)| \frac{u(x)}{u(\tilde{x}_k)} \leq \mathcal{C} \log n.$$

Nearest-node term. By the modified kernel representation (6.75) and by the identity $\tilde{\ell}_d(\tilde{x}_d) = 1$, we have

$$\tilde{\ell}_d(x) = \frac{\tilde{K}_n(v^{\alpha,\beta}, x, \tilde{x}_d)}{\tilde{K}_n(v^{\alpha,\beta}, \tilde{x}_d, \tilde{x}_d)}.$$

Applying Cauchy–Schwarz to the finite-sum kernel representation gives

$$|\tilde{K}_n(v^{\alpha,\beta}, x, \tilde{x}_d)|^2 \leq \tilde{K}_n(v^{\alpha,\beta}, x, x)\tilde{K}_n(v^{\alpha,\beta}, \tilde{x}_d, \tilde{x}_d).$$

Thus

$$|\tilde{\ell}_d(x)| \leq \left(\frac{\tilde{K}_n(v^{\alpha,\beta}, x, x)}{\tilde{K}_n(v^{\alpha,\beta}, \tilde{x}_d, \tilde{x}_d)}\right)^{1/2}.$$

By (6.46) and the all-point diagonal Jacobi kernel estimate proved above,

$$\tilde{K}_n(v^{\alpha,\beta}, t, t) \asymp \frac{n}{v^{\alpha,\beta}(t)\varphi(t)}, \qquad t \in I_n.$$

Let Δ_d^* denote the adjacent-node gap corresponding to the convention for Δz_d^* in Lemma 6.13. Since $\tilde{x}_d$ is the closest node to x, the endpoint and spacing estimates in Propositions 6.8 and 6.9 give

$$|x - \tilde{x}_d| \leq \mathcal{C}\Delta_d^* \leq \mathcal{C}\frac{\varphi(\tilde{x}_d)}{n}.$$

Together with $x, \tilde{x}_d \in I_n$, this implies

$$1 \pm x \asymp 1 \pm \tilde{x}_d, \qquad v^{\alpha,\beta}(x)\varphi(x) \asymp v^{\alpha,\beta}(\tilde{x}_d)\varphi(\tilde{x}_d), \qquad u(x) \asymp u(\tilde{x}_d).$$

Consequently,

$$|\tilde{\ell}_d(x)| \frac{u(x)}{u(\tilde{x}_d)} \leq \mathcal{C}.$$

Conclusion. Combining the two estimates in (6.80), we obtain

$$\max_{x\in I_n} \left|\tilde{L}_{n+1}(v^{\alpha,\beta}, f, x)u(x)\right| \leq \mathcal{C} \log n \, \|f\|_{C_u}.$$

The Remez reduction at the beginning of the proof then yields

$$\left\|\tilde{L}_{n+1}(v^{\alpha,\beta}, f)\right\|_{C_u} \leq \mathcal{C} \log n \, \|f\|_{C_u}.$$

This proves the assertion. □

Theorem 6.14 gives sufficient conditions for a logarithmic upper bound.

Combining Theorem 6.14 with the projector estimate gives the following near-best bound. The Sobolev estimate follows from the Favard inequality.

Theorem 6.15. Under the assumptions of Theorem 6.14, for any $f \in C_u$, one has

$$\left\| f - \tilde{L}_{n+1}(v^{\alpha,\beta}, f) \right\|_{C_u} \le \mathcal{C} \log n \, E_n(f)_u,$$

where $\mathcal{C}$ is independent of n and f. Moreover, for any $f \in W_r(u)$, $r \ge 1$, the following estimate holds

$$\left\| f - \tilde{L}_{n+1}(v^{\alpha,\beta}, f) \right\|_{C_u} \le \mathcal{C} \frac{\log n}{n^r} \|f\|_{W_r(u)}, \tag{6.81}$$

with $\mathcal{C}$ independent of n and f.

Proof. Combine the weighted projector estimate (6.78) with Theorem 6.14. The Sobolev estimate (6.81) then follows from the Favard inequality (6.20). □

Remark 6.16. In view of the previous result, the Lagrange interpolation sequence

$$\{\tilde{L}_{n+1}(v^{\alpha,\beta}, f)\}_n$$

has the same approximation behavior as the sequence of best polynomial approximants in C_u, up to the slowly diverging factor $\log n$. This is the same near-best phenomenon observed for Lagrange interpolation at the zeros of the Jacobi polynomials; see [124, 122]. The endpoint-weight assumptions do not coincide, however: the anti-Gauss process is governed by the shifted window (6.79), whereas the classical Gauss–Jacobi window recalled in the numerical comparison below is different. This shift reflects the different placement of the extreme nodes.

For an individual $f \in C_u$, the near-best estimate yields convergence when $(\log n)E_n(f)_u \to 0$, in particular for the Sobolev classes in Theorem 6.15. The fact that $E_n(f)_u \to 0$ alone does not remove the logarithmic factor. The experiments below compare the finite errors and the parameter windows of the two interpolation processes.

6.2.5 Numerical experiments

We discuss the experiments in [47, Section 5]. The first comparison concerns weighted Lebesgue functions and constants; its plots are in the source and are not reproduced here. The second concerns sampled interpolation errors for three functions, reproduced in Table 6.2. The source reports MATLAB R2024b computations.

Error measure and comparison convention. The reported anti-Gauss error is

$$\tilde{\epsilon}_n(f_i) = \max_{x \in X} |[f_i(x) - \tilde{L}_{n+1}(v^{\alpha,\beta}, f_i, x)]u(x)|, \qquad i = 1, 2, 3, \tag{6.82}$$

Here X is the equispaced grid in $[-1, 1]$ described in the source. The reported maximum is a discrete approximation to the weighted uniform error. The source does not specify the number of grid points.

The comparison uses the classical Gauss–Jacobi interpolant $\mathcal{L}_n(v^{\alpha,\beta}, f)$ at the n zeros of $p_n(v^{\alpha,\beta})$. Thus a row indexed by n compares n Gauss nodes with $n+1$ anti-Gauss nodes, not equal numbers of function values:

$$\mathcal{L}_n(v^{\alpha,\beta}, f, x) = \sum_{k=1}^{n} \ell_k(x) f(x_k), \quad \ell_k(x) = \lambda_k \left[\sum_{j=0}^{n-1} p_j(v^{\alpha,\beta}, x) p_j(v^{\alpha,\beta}, x_k) \right]. \tag{6.83}$$

The weighted Lebesgue constants associated with this interpolation process in C_u are

$$\Lambda_n(v^{\alpha,\beta}) = \max_{x\in[-1,1]} u(x) \sum_{k=1}^{n} \frac{|\ell_k(x)|}{u(x_k)}. \tag{6.84}$$

For the classical Jacobi interpolation process, $\Lambda_n(v^{\alpha,\beta}) = O(\log n)$ if and only if the classical Jacobi window

$$\frac{\alpha}{2} + \frac{1}{4} \le \gamma \le \frac{\alpha}{2} + \frac{5}{4}, \qquad \frac{\beta}{2} + \frac{1}{4} \le \delta \le \frac{\beta}{2} + \frac{5}{4}$$

is satisfied. The anti-Gauss process is governed instead by the shifted window (6.79). Therefore, in the comparisons below, the behavior of the Gauss and anti-Gauss Lebesgue constants should be read according to their respective windows; see, for instance, [122, 124] for the classical Jacobi case.

Example 6.17. Consider the Legendre weight ($\alpha = \beta = 0$) and the second-kind Chebyshev weight ($\alpha = \beta = 1/2$). The source compares their Gauss and anti-Gauss Lebesgue constants in both weighted and unweighted norms; see [47, Figures 2–3]. We summarize the parameter comparison rather than reproduce those plots.

The case $\gamma = \delta = 0$ is covered by Theorem 6.14: for the Legendre weight ($\alpha = \beta = 0$) the shifted window gives $0 \le \gamma, \delta \le 3/4$, while for the second-kind Chebyshev weight ($\alpha = \beta = 1/2$) it gives $0 \le \gamma, \delta \le 1$. Thus the $O(\log n)$ upper bound for the anti-Gauss Lebesgue constants in the unweighted norm follows from the shifted endpoint window. In contrast, the corresponding classical Jacobi interpolation process is outside its classical weighted window when $u = 1$, which explains the different behavior of the two processes in this comparison.

Example 6.18. Consider the functions

$$f_1(x) = (1+|x|)^{\frac{3}{2}}(x^2+\cos(x)), \qquad f_2(x) = e^{-|x|^{\frac{5}{2}}}, \qquad f_3(x) = |x+0.5|^{\frac{7}{2}}\frac{\sin(x)}{1+x^2},$$

and approximate them with the interpolating polynomials (6.73) and (6.83) associated with the weights

$$v(x) = \begin{cases} (1+x)^{-1/3}, & i=1, \\ \sqrt{1-x^2}, & i=2, \\ (1-x)^{-1/4}(1+x)^{1/3}, & i=3. \end{cases}$$

The left-hand sides of (6.37) and (6.38) are strictly positive for these configurations and the tabulated degrees, so all anti-Gauss nodes belong to $(-1,1)$.

Table 6.2 reports the errors (6.82) and the errors

$$\epsilon_n(f_i) = \max_{x\in X} |[f_i(x) - \mathcal{L}_n(v^{\alpha,\beta}, f_i, x)]u(x)| \tag{6.85}$$

with

$$u(x) = \begin{cases} (1-x)^{1/4}, & i=1, \\ \sqrt{1-x^2}, & i=2, \\ (1-x)^{1/8}(1+x)^{5/12}, & i=3. \end{cases}$$

The source records $f_1 \in W_1(u)$, $f_2 \in W_2(u)$, and $f_3 \in W_3(u)$. These memberships give the corresponding asymptotic upper estimates in Theorem 6.15. The constant is not evaluated.

The displayed weights belong to the anti-Gauss window. For f_1, however, $\beta = -1/3$ and $\delta = 0$, whereas the recalled classical Gauss window requires $\delta \geq 1/12$. Thus that Gauss logarithmic bound does not apply to the first comparison. The other two choices satisfy both displayed windows.

The computations use the Christoffel–Darboux representations (6.83) and (6.75). In Table 6.2, the anti-Gauss errors are smaller for f_1 and f_2 in every displayed row. For f_3, the comparison changes with n: the Gauss errors are smaller at $n = 32$, 128, and 512.

Table 6.2: Weighted maximum errors on the sampled grid X in Example 6.18. Each row compares n Gauss nodes with $n+1$ anti-Gauss nodes. Values are reproduced from [47, Table 1].

n	f_1		f_2		f_3	
	ϵ_n	$\tilde{\epsilon}_n$	ϵ_n	$\tilde{\epsilon}_n$	ϵ_n	$\tilde{\epsilon}_n$
16	8.90×10^{-2}	4.75×10^{-2}	1.01×10^{-3}	4.50×10^{-4}	2.95×10^{-5}	2.31×10^{-5}
32	4.49×10^{-2}	2.41×10^{-2}	1.84×10^{-4}	8.47×10^{-5}	1.33×10^{-6}	2.76×10^{-6}
64	2.25×10^{-2}	1.18×10^{-2}	3.34×10^{-5}	1.46×10^{-5}	2.26×10^{-7}	1.45×10^{-7}
128	1.13×10^{-2}	5.92×10^{-3}	6.01×10^{-6}	2.58×10^{-6}	1.12×10^{-8}	2.14×10^{-8}
256	5.65×10^{-3}	1.78×10^{-3}	1.07×10^{-6}	3.62×10^{-7}	1.81×10^{-9}	5.38×10^{-10}
512	2.83×10^{-3}	7.58×10^{-4}	1.90×10^{-7}	6.46×10^{-8}	4.41×10^{-11}	1.68×10^{-10}
1024	1.41×10^{-3}	3.79×10^{-4}	3.38×10^{-8}	5.81×10^{-9}	1.43×10^{-11}	3.62×10^{-12}
2048	7.07×10^{-4}	1.90×10^{-4}	5.97×10^{-9}	7.61×10^{-10}	4.95×10^{-12}	1.30×10^{-12}

6.3 Conclusions

The first study gives three exact structured representations of the anti-Gauss Lagrange interpolant. In the reported Chebyshev implementation, their scaled or kernel-based evaluations return finite sampled errors at degrees where the direct nodal computation returns `NaN`.

The Jacobi study concerns the interpolation operators rather than a single evaluation formula. Under the fixed-parameter restriction $\alpha, \beta > -1/2$, with $n \in \mathcal{N}_{\alpha,\beta}$ and the stated endpoint-weight assumptions, the local node and polynomial estimates lead to a logarithmic weighted Lebesgue bound. The resulting near-best estimate retains the factor $\log n$; the weighted Sobolev estimate supplies a class for which the error tends to zero. The numerical comparisons must be interpreted within the respective Gauss and anti-Gauss parameter windows.

Together with Chapter 5, this chapter shows how node geometry enters approximation in two different ways: separation from a moving singularity for quadrature, and local spacing and endpoint control for weighted interpolation.

Chapter 7

Conclusions and Perspectives

7.1 Summary of the main results

Across the four core chapters, a structured representation preserves a specified asymptotic or exact property. Its computational use depends on the spectral normalization and on further clustering, separation, or local node estimates. Part I develops unequal-block Toeplitz approximations and single-axis quaternion representations; Part II studies prescribed-node quadrature for the circular Hilbert transform and anti-Gauss interpolation.

Unequal-block Toeplitz approximation and preconditioning. Chapter 3 constructs a structured approximation S_n for square and rectangular unilevel Toeplitz block matrices with unequal external dimensions. For a fixed block count, rational limiting ratios, and the stated principal and residual approximation hypotheses, the error sequence $\{S_n - A_n\}_n$ is zero-distributed. Strong error clustering requires continuity and strong clustering of the blockwise errors. Weak preconditioned singular-value clustering requires the stated rank hypotheses; the eigenvalue conclusion additionally requires the specified Hermitian setting and a positive-definite preconditioner.

Quaternion block multilevel Toeplitz sequences. Chapter 4 identifies complex symbols for the left, right, and sandwich single-axis quaternion Toeplitz constructions. The symplectic representation and its doubled spectral normalization allow the stated distribution, localization, and approximation results to be interpreted in terms of quaternion spectra. Singular-value distributions hold for integrable generators; eigenvalue results retain the Hermitian criterion or, in the non-Hermitian case, boundedness and the stated Tilli-class hypotheses.

Prescribed-node quadrature for the circular Hilbert transform. Chapter 5 combines paired Szegő and anti-Szegő quadrature with evaluation-dependent node placement. Cancellation on the stated Laurent-polynomial space and separation from the kernel singularity control different parts of the construction. For normalized Lebesgue measure on the unit circle, under the stated regularity and approximation hypotheses, the prescribed geometry gives $O(\log n)$ pointwise operator bounds and the quadrature-error estimate. The specified Sobolev class has the upper rate $O(\log n/n^r)$.

Anti-Gauss representations and weighted Jacobi interpolation. Chapter 6 gives three exact representations of the same finite interpolant: modified Christoffel–Darboux, barycentric, and orthogonal-polynomial expansion forms.

Local Jacobi-node and polynomial estimates control the growing interpolation process. For fixed $\alpha, \beta > -1/2$, admissible degrees $n \in \mathcal{N}_{\alpha,\beta}$, and the stated endpoint-weight assumptions, the weighted Lebesgue constants satisfy a logarithmic $O(\log n)$ upper bound, which leads to the near best-approximation and weighted Sobolev convergence estimates. The limiting node distribution alone does not supply these local estimates.

7.2 Scope and limitations

Normalized eigenvalue or singular-value distributions do not by themselves control extremal values, pseudospectra, or solver convergence. Clustering and localization require separate results. Weak singular-value clustering still permits $o(r_n)$ outliers, where r_n is the active singular-value dimension; in the square eigenvalue setting, the corresponding count is $o(d_n)$. The reported Krylov comparisons therefore remain finite observations rather than general convergence guarantees. Similarly, the paired quadrature difference has no general reliability or efficiency theorem, and the finite-precision interpolation comparisons do not constitute a general rounding-error analysis.

The approximation results retain the measure, function-space, regularity, and geometric assumptions stated in their chapters. In particular, the quadrature rate is an upper estimate. The anti-Gauss interpolation theory gives sufficient conditions for the logarithmic upper estimate. It does not prove a matching lower bound, necessity of the endpoint-weight restrictions, or a classification of the parameter region for logarithmic anti-Gauss Jacobi interpolation.

7.3 Directions for future research

Structured matrix-sequences and solvers. For unequal-block systems, open problems include preconditioning for irrational limiting dimension ratios and multilevel extensions, as discussed in [9, Section 6]. An extension to block multilevel Toeplitz and GLT structures must also account for the barriers discussed in [133, 154]. The multi-iterative strategy of [145], which combines stationary or preconditioned Krylov solvers with multigrid methods, is one computational approach; see also [61, 18]. Nested irregular blocks, such as those arising in [72], further require an analysis of interacting compressions and their effect on the symbol. Discretization accuracy, matrix structure, and computational cost must be assessed together. The optimal and reduced spline spaces studied in [113] provide one example: for the stated Laplace eigenvalue problems on bounded intervals and tensor-product boxes, their mass and stiffness matrices have Toeplitz-minus-Hankel or Toeplitz-plus-Hankel structure and closed-form eigenpairs.

Quaternion extensions. Extending the single-axis quaternion setting to variable coefficients and multiaxial Fourier kernels calls for a GLT-type $*$-algebra beyond the reductions proved here. Quantitative structure-preserving preconditioning and quaternion Krylov analysis are also open, with Korovkin-type constructions [150, 149, 111] and a quaternionic multigrid approximation theory as possible approaches. Finer spectral questions concern extremal and pseudospectral behavior, distribution rates, and asymptotic expansions for matrix-free computation. The representation technique used here and slice quaternion function theory [87] offer complementary tools. Applications in imaging and low-rank approximation [104, 38, 136] and structure-preserving algorithms [103, 119, 137] motivate these questions.

Quadrature, interpolation, and finite precision. For the circular Hilbert transform, broader measures, including singular weights, other strongly singular kernels, and periodic Cauchy integral equations require stability and convergence estimates adapted to the measure, kernel, and function space. Certifying the paired-rule difference for adaptive use would require reliability and efficiency bounds under explicit conditions. For anti-Gauss interpolation, further tasks include other polynomial families, multidimensional settings, and matching lower bounds. Whether combinations of the Gauss–Jacobi and anti-Gauss processes improve weighted approximation requires a separate analysis; quadrature error cancellation alone does not establish such an improvement. Extended interpolation was studied in [41], and processes involving zeros of second-kind Jacobi polynomials in [40]. These results motivate studying extended processes that combine Jacobi and anti-Jacobi node sets. For finite-precision evaluation, the concrete task is to derive forward- or backward-error bounds for the three representations and their scaled or kernel-based evaluations, accounting for degree, node geometry, and cancellation.

Structured descriptions of the Part II operators. The Jacobi matrices underlying real-line quadrature and interpolation and the Hessenberg matrices used on the unit circle connect Part II to the structured-matrix analysis of Part I. Here local node and weight information is essential, so a global distribution theorem is only part of the required analysis. When a matrix sequence admits a suitable asymptotic eigenvalue expansion, smaller matrices can provide spectral approximations at larger dimensions [63, 62, 24]. Applying this approach to Gaussian and Szegő quadrature and related interpolation processes requires the relevant expansion, control of the quantities defining nodes and weights, and an analysis of topological obstructions and numerical stability. Fixed structure alone is insufficient; extrapolation remains conditional where the expansion is conjectural.

The representations and estimates established here provide a basis for further work on structured solvers, quadrature, and interpolation.

Bibliography

[1] A. Adriani, I. Furci, C. Garoni, and S. Serra-Capizzano. Spectral and singular value distribution of sequences of block matrices with rectangular Toeplitz blocks. part I: asymptotically rational block size ratios. *J. Numer. Math.*, 2026. Published online 13 March 2026.

[2] A. Adriani, C. Garoni, and S. Serra-Capizzano. Spectral and singular value distribution of sequences of block matrices with rectangular Toeplitz blocks. part II: asymptotically irrational block size ratios. *J. Numer. Math.*, 2026. Published online 19 March 2026.

[3] A. Adriani and A. J. A. Schiavoni-Piazza. The $*$-algebra of unbounded GLT: construction and theoretical foundations, 2026. arXiv:2602.23879v1 [math.NA].

[4] A. Adriani, A. J. A. Schiavoni-Piazza, and S. Serra-Capizzano. Block structures, g.a.c.s. approximation, and distributions. *Bol. Soc. Mat. Mex.*, 31(2), 2025. Article 41.

[5] A. Adriani, A. J. A. Schiavoni-Piazza, S. Serra-Capizzano, and C. Tablino-Possio. Revisiting the notion of approximating class of sequences for handling approximated PDEs on moving or unbounded domains. *Electron. Trans. Numer. Anal.*, 63:424–451, 2025.

[6] F. Avram. On bilinear forms in Gaussian random variables and Toeplitz matrices. *Probab. Theory Relat. Fields*, 79(1):37–45, 1988.

[7] O. Axelsson and G. Lindskog. On the rate of convergence of the preconditioned conjugate gradient method. *Numer. Math.*, 48(5):499–523, 1986.

[8] E. A. Azoff. Borel measurability in linear algebra. *Proc. Am. Math. Soc.*, 42(2):346–350, 1974.

[9] N. Barakitis, M. Donatelli, S. Ferri, V. Loi, S. Serra-Capizzano, and R. L. Sormani. Blocking structures, approximation, and preconditioning. *Numer. Algorithms*, 100(4):1893–1927, 2025. Published online 8 July 2025; arXiv:2501.14874v1 [math.NA].

[10] N. Barakitis, S.-E. Ekström, and P. Vassalos. Preconditioners for fractional diffusion equations based on the spectral symbol. *Numer. Linear Algebra Appl.*, 29(5):e2441, 2022.

[11] N. Barakitis, P. Ferrari, I. Furci, and S. Serra-Capizzano. An extradimensional approach for distributional results: the case of 2×2 block Toeplitz structures. In K. Jbilou, M. Hached, L. Maniar, and A. Ratnani, editors, *Mathematical Modeling with Modern Applications: M3A24, Istanbul, Türkiye, June 4–6, 2024*, volume 497 of *Springer Proceedings in Mathematics and Statistics*, pages 61–78. Springer Cham, Cham, 2025.

[12] N. Barakitis, V. Loi, and S. Serra-Capizzano. A note on eigenvalues and singular values of variable Toeplitz matrices and matrix-sequences, with application to variable two-step BDF approximations to parabolic equations. In N. Asharaf, W. Bauer, B. V. R. Bhat, and J. Sarkar, editors, *Recent Developments in Spectral and Approximation Theory*, Trends in Mathematics, pages 37–67. Birkhäuser, Cham, 2025. First published online 27 July 2025; arXiv:2407.00792v3 [math.NA].

[13] G. Barbarino, S. E. Ekström, C. Garoni, D. Meadon, S. Serra-Capizzano, and P. Vassalos. From asymptotic distribution and vague convergence to uniform convergence, with numerical applications. *J. Numer. Math.*, 34(1):21–60, 2026.

[14] G. Barbarino and C. Garoni. GLT sequences and normal matrices. *Electron. J. Linear Algebra*, 41:1–20, 2025.

[15] G. Barbarino, C. Garoni, M. Mazza, and S. Serra-Capizzano. Rectangular GLT sequences. *Electron. Trans. Numer. Anal.*, 55:585–617, 2022.

[16] G. Barbarino, C. Garoni, and S. Serra-Capizzano. Block generalized locally Toeplitz sequences: theory and applications in the multidimensional case. *Electron. Trans. Numer. Anal.*, 53:113–216, 2020.

[17] G. Barbarino, C. Garoni, and S. Serra-Capizzano. Block generalized locally Toeplitz sequences: theory and applications in the unidimensional case. *Electron. Trans. Numer. Anal.*, 53:28–112, 2020.

[18] P. Benedusi, P. Ferrari, M. E. Rognes, and S. Serra-Capizzano. Modeling excitable cells with the EMI equations: spectral analysis and iterative solution strategy. *J. Sci. Comput.*, 98(3):58, 2024.

[19] T. B. Benjamin. Internal waves of permanent form in fluids of great depth. *J. Fluid Mech.*, 29(3):559–592, 1967.

[20] J. P. Berrut and L. N. Trefethen. Barycentric Lagrange interpolation. *SIAM Rev.*, 46(3):501–517, 2004.

[21] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, second edition, 1999.

[22] D. Boffi, L. Gastaldi, and M. Ruggeri. Mixed formulation for interface problems with distributed Lagrange multiplier. *Comput. Math. Appl.*, 68(12):2151–2166, 2014.

[23] J. M. Bogoya, A. Böttcher, and E. A. Maximenko. From convergence in distribution to uniform convergence. *Bol. Soc. Mat. Mex.*, 22(2):695–710, 2016.

[24] M. Bogoya, S. M. Grudsky, and S. Serra-Capizzano. Fast non-Hermitian Toeplitz eigenvalue computations, joining matrixless algorithms and FDE approximation matrices. *SIAM J. Matrix Anal. Appl.*, 45(1):284–305, 2024.

[25] M. Bolten, M. Donatelli, P. Ferrari, and I. Furci. A symbol-based analysis for multigrid methods for block-circulant and block-Toeplitz systems. *SIAM J. Matrix Anal. Appl.*, 43(1):405–438, 2022.

[26] A. Böttcher and S. M. Grudsky. On the condition numbers of large semi-definite Toeplitz matrices. *Linear Algebra Appl.*, 279(1–3):285–301, 1998.

[27] A. Böttcher and B. Silbermann. *Introduction to Large Truncated Toeplitz Matrices*. Springer-Verlag, New York, 1999.

[28] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, 2007.

[29] H. Brass and K. Petras. *Quadrature Theory: The Theory of Numerical Integration on a Compact Interval*, volume 178 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[30] A. Bultheel, L. Daruis, and P. Gonzalez-Vera. A connection between quadrature formulas on the unit circle and the interval $[-1, 1]$. *J. Comput. Appl. Math.*, 132(1):1–14, 2001.

[31] A. Bultheel, L. Daruis, and P. González-Vera. Quadrature formulas on the unit circle with prescribed nodes and maximal domain of validity. *J. Comput. Appl. Math.*, 231(2):948–963, 2009.

[32] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. Orthogonality and quadrature on the unit circle. In C. Brezinski, L. Gori, and A. Ronveaux, editors, *Orthogonal Polynomials and Their Applications*, volume 9 of *IMACS Annals on Computing and Applied Mathematics*, pages 205–210. J. C. Baltzer AG, Basel, 1991.

[33] J. F. Cai, R. H. Chan, L. Shen, and Z. Shen. Convergence analysis of tight framelet approach for missing data recovery. *Adv. Comput. Math.*, 31(1–3):87–113, 2009.

[34] J. F. Cai, R. H. Chan, and Z. Shen. A framelet-based image inpainting algorithm. *Appl. Comput. Harmon. Anal.*, 24(2):131–149, 2008.

[35] R. H. Chan and M. K. Ng. Conjugate gradient methods for Toeplitz systems. *SIAM Rev.*, 38(3):427–482, 1996.

[36] T. F. Chan. An optimal circulant preconditioner for Toeplitz systems. *SIAM J. Sci. Stat. Comput.*, 9(4):766–771, 1988.

[37] M. M. Chawla and T. R. Ramakrishnan. Numerical evaluation of integrals of periodic functions with Cauchy and Poisson type kernels. *Numer. Math.*, 22(4):317–323, 1974.

[38] J. Chen and M. K. Ng. Color image inpainting via robust pure quaternion matrix completion: error bound and weighted loss. *SIAM J. Imaging Sci.*, 15(3):1469–1498, 2022.

[39] E. W. Cheney. *Introduction to Approximation Theory*. International Series in Pure and Applied Mathematics. McGraw–Hill Book Company, New York, 1966.

[40] G. Criscuolo and B. Della Vecchia. Extended Lagrange interpolation by zeros of Jacobi polynomials of second kind. *Rend. Circ. Mat. Palermo (2) Suppl.*, 33:281–287, 1993.

[41] G. Criscuolo, G. Mastroianni, and D. Occorsio. Convergence of extended Lagrange interpolation. *Math. Comp.*, 55(191):197–212, 1990.

[42] P. J. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Computer Science and Applied Mathematics. Academic Press, Orlando, FL, second edition, 1984.

[43] V. Del Prete, F. Di Benedetto, M. Donatelli, and S. Serra-Capizzano. Symbol approach in a signal-restoration problem involving block Toeplitz matrices. *J. Comput. Appl. Math.*, 272:399–416, 2014.

[44] R. A. DeVore and G. G. Lorentz. *Constructive Approximation*, volume 303 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1993.

[45] F. di Benedetto and S. Serra-Capizzano. Optimal multilevel matrix algebra operators. *Linear Multilinear Algebra*, 48(1):35–66, 2000.

[46] P. Díaz de Alba, L. Fermo, and V. Loi. Anti-Gauss Lagrange interpolation: Christoffel-Darboux form, barycentric representation, and orthogonal expansion, 2026. Posted 16 June 2026; arXiv:2606.18018v1 [math.NA].

[47] P. Díaz de Alba, L. Fermo, V. Loi, and D. Occorsio. Lagrange interpolation processes based on the zeros of anti-Gauss Jacobi polynomials, 2026. Posted 23 July 2026; arXiv:2607.21714v1 [math.NA].

[48] P. Díaz de Alba, L. Fermo, and G. Rodriguez. Solution of second kind Fredholm integral equations by means of Gauss and anti-Gauss quadrature rules. *Numer. Math.*, 146(4):699–728, 2020.

[49] P. Díaz de Alba, L. Fermo, and G. Rodriguez. AGquad: a MATLAB package for 1D and 2D anti-Gauss type rules. *J. Approx. Softw.*, 2(1):16–32, 2025.

[50] P. Díaz de Alba, L. Fermo, and G. Rodriguez. Anti-Gauss cubature rules with applications to Fredholm integral equations on the square. *SIAM J. Sci. Comput.*, 47(2):A689–A712, 2025.

[51] Z. Ditzian and V. Totik. *Moduli of Smoothness*, volume 9 of *Springer Series in Computational Mathematics*. Springer, New York, 1987.

[52] D. L. Djukić, L. Fermo, and R. M. Mutavdžić Djukić. Averaged cubature schemes on the real positive semiaxis. *Numer. Algorithms*, 92(1):545–569, 2023.

[53] D. L. Djukić, R. M. Mutavdžić Djukić, L. Reichel, and M. M. Spalević. Generalized averaged Gauss quadrature rules: a survey. *Mathematics*, 13(19):3145, 2025.

[54] M. Donatelli, A. Dorostkar, M. Mazza, M. Neytcheva, and S. Serra-Capizzano. Function-based block multigrid strategy for a two-dimensional linear elasticity-type problem. *Comput. Math. Appl.*, 74(5):1015–1028, 2017.

[55] M. Donatelli, P. Ferrari, I. Furci, S. Serra-Capizzano, and D. Sesana. Multigrid methods for block-Toeplitz linear systems: convergence analysis and applications. *Numer. Linear Algebra Appl.*, 28(4):e2356, 2021.

[56] M. Donatelli, M. Mazza, and S. Serra-Capizzano. Spectral analysis and structure preserving preconditioners for fractional diffusion equations. *J. Comput. Phys.*, 307:262–279, 2016.

[57] M. Donatelli, M. Neytcheva, and S. Serra-Capizzano. Canonical eigenvalue distribution of multilevel block Toeplitz sequences with non-Hermitian symbols. In W. Arendt, J. A. Ball, J. Behrndt, K.-H. Förster, V. Mehrmann, and C. Trunk, editors, *Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equations*, pages 269–291. Springer Basel, Basel, 2012.

[58] A. Dorostkar, M. Neytcheva, and S. Serra-Capizzano. Spectral analysis of coupled PDEs and of their Schur complements via generalized locally Toeplitz sequences in 2D. *Comput. Methods Appl. Mech. Eng.*, 309:74–105, 2016.

[59] J. Du. Quadrature formulas of quasi-interpolation type for singular integrals with Hilbert kernel. *J. Approx. Theory*, 93(2):231–257, 1998.

[60] J.-Y. Du. Quadrature formulas for singular integrals with Hilbert kernel. *J. Comput. Math.*, 6(3):205–225, 1988.

[61] M. Dumbser, F. Fambri, I. Furci, M. Mazza, S. Serra-Capizzano, and M. Tavelli. Staggered discontinuous Galerkin methods for the incompressible Navier-Stokes equations: spectral analysis and computational results. *Numer. Linear Algebra Appl.*, 25(5):e2151, 2018.

[62] S.-E. Ekström, I. Furci, and S. Serra-Capizzano. Exact formulae and matrix-less eigensolvers for block banded symmetric Toeplitz matrices. *BIT Numer. Math.*, 58(4):937–968, 2018.

[63] S.-E. Ekström, C. Garoni, and S. Serra-Capizzano. Are the eigenvalues of banded symmetric Toeplitz matrices known in almost closed form? *Exp. Math.*, 27(4):478–487, 2018.

[64] T. A. Ell, N. Le Bihan, and S. J. Sangwine. *Quaternion Fourier Transforms for Signal and Image Processing*. ISTE Ltd and John Wiley & Sons, Inc., London and Hoboken, NJ, 2014.

[65] G. Faber. Über die interpolatorische Darstellung stetiger Funktionen. *Jahresber. Dtsch. Math.-Ver.*, 23:192–210, 1914.

[66] D. Fasino and P. Tilli. Spectral clustering properties of block multilevel Hankel matrices. *Linear Algebra Appl.*, 306(1–3):155–163, 2000.

[67] L. Fermo. Embedding theorems for functions with inner singularities. *Acta Sci. Math. (Szeged)*, 75(3–4):547–573, 2009.

[68] L. Fermo and V. Loi. Approximation of the Hilbert transform on the unit circle. *J. Sci. Comput.*, 105(1), 2025. Article 23; published online 10 September 2025; arXiv:2409.07810v1 [math.NA].

[69] L. Fermo and D. Occorsio. Anti-Gauss Laguerre polynomials: some properties and a new interpolation process. *J. Comput. Appl. Math.*, 475:117034, 2026.

[70] L. Fermo, L. Reichel, G. Rodriguez, and M. M. Spalević. Averaged Nyström interpolants for the solution of Fredholm integral equations of the second kind. *Appl. Math. Comput.*, 467:128482, 2024.

[71] P. Ferrari, R. I. Rahla, C. Tablino-Possio, S. Belhaj, and S. Serra-Capizzano. Multigrid for $\mathbb{Q}_k$ finite element matrices using a (block) Toeplitz symbol approach. *Mathematics*, 8(1):5, 2020.

[72] S. Ferri, C. Giraudo, V. Loi, M. Kuchta, and S. Serra-Capizzano. Spectral analysis of the stiffness matrix sequence in the approximated Stokes equation. *Bol. Soc. Mat. Mex.*, 32(3):102, 2026. Article 102; published online 30 July 2026; arXiv:2510.25252v2 [math.NA].

[73] I. Furci, A. Adriani, and S. Serra-Capizzano. Block structured matrix-sequences and their spectral and singular value canonical distributions: a general theory, 2024. arXiv:2409.06465v2 [math.NA]; last revised 4 February 2025.

[74] A. L. Garkavi. Simultaneous approximation to a periodic function and its derivatives by trigonometric polynomials. *Izv. Akad. Nauk SSSR Ser. Mat.*, 24(1):103–128, 1960.

[75] C. Garoni. Estimates for the minimum eigenvalue and the condition number of Hermitian (block) Toeplitz matrices. *Linear Algebra Appl.*, 439(3):707–728, 2013.

[76] C. Garoni. Topological foundations of an asymptotic approximation theory for sequences of matrices with increasing size. *Linear Algebra Appl.*, 513:324–341, 2017.

[77] C. Garoni and S. Serra-Capizzano. *Generalized Locally Toeplitz Sequences: Theory and Applications: Volume I.* Springer, Cham, 2017.

[78] C. Garoni and S. Serra-Capizzano. Generalized locally Toeplitz sequences: a spectral analysis tool for discretized differential equations. In T. Lyche, C. Manni, and H. Speleers, editors, *Splines and PDEs: From Approximation Theory to Numerical Linear Algebra: Cetraro, Italy 2017*, volume 2219 of *Lecture Notes in Mathematics*, pages 161–236. Springer, Cham, 2018.

[79] C. Garoni and S. Serra-Capizzano. *Generalized Locally Toeplitz Sequences: Theory and Applications: Volume II.* Springer, Cham, 2018.

[80] C. Garoni and S. Serra-Capizzano. Block Jacobi/Gauss–Seidel preconditioning for GLT sequences, and GLH sequences, 2026. arXiv:2606.01888v1 [math.NA].

[81] C. Garoni, S. Serra-Capizzano, and D. Sesana. Spectral analysis and spectral symbol of d-variate $\mathbb{Q}_p$ Lagrangian FEM stiffness matrices. *SIAM J. Matrix Anal. Appl.*, 36(3):1100–1128, 2015.

[82] C. Garoni, H. Speleers, S.-E. Ekström, A. Reali, S. Serra-Capizzano, and T. J. R. Hughes. Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: exposition and review. *Arch. Comput. Methods Eng.*, 26(5):1639–1690, 2019.

[83] C. F. Gauss. Methodus nova integralium valores per approximationem inveniendi. *Comm. Soc. R. Sci. Göttingen Recens.*, 3:39–76, 1814. Werke **3**, 163–196, (1866).

[84] W. Gautschi. A survey of Gauss-Christoffel quadrature formulae. In P. L. Butzer and F. Fehér, editors, *E. B. Christoffel. The Influence of His Work on Mathematics and the Physical Sciences*, pages 72–147. Springer, 1981.

[85] W. Gautschi. *Orthogonal Polynomials: Computation and Approximation.* Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2004.

[86] W. Gautschi. Interpolation before and after Lagrange. *Rend. Semin. Mat. Univ. Politec. Torino*, 70(4):347–368, 2012.

[87] G. Gentili, C. Stoppato, and D. C. Struppa. *Regular Functions of a Quaternionic Variable.* Springer Monographs in Mathematics. Springer Berlin Heidelberg, 2013.

[88] L. Golinskii and S. Serra-Capizzano. The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences. *J. Approx. Theory*, 144(1):84–102, 2007.

[89] G. H. Golub and C. F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, MD, 4 edition, 2013.

[90] G. H. Golub and J. H. Welsch. Calculation of Gauss quadrature rules. *Math. Comp.*, 23(106):221–230, 1969.

[91] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products.* Academic Press, Amsterdam, seventh edition, 2007. Edited by Alan Jeffrey and Daniel Zwillinger.

[92] W. B. Gragg. Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle. *J. Comput. Appl. Math.*, 46(1–2):183–198, 1993.

[93] U. Grenander and G. Szegő. *Toeplitz Forms and Their Applications.* Chelsea Publishing Company, New York, second edition, 1984.

[94] G. Han, J. Yu, and Z. Sun. The Jordan–Chevalley decomposition and Jordan canonical form of a quaternionic linear operator, 2019. arXiv:1906.01918v1 [math.RA].

[95] N. J. Higham. The numerical stability of barycentric Lagrange interpolation. *IMA J. Numer. Anal.*, 24(4):547–556, 2004.

[96] E. Hitzer. Quaternion Fourier transform on quaternion fields and generalizations. *Adv. Appl. Clifford Algebras*, 17(3):497–517, 2007.

[97] C. Hu, W. Ding, and Y. Li. Preconditioned generalized minimal residual method for quaternion linear systems with structured coefficient matrices and its applications. *Numer. Algorithms*, 2026. Published online 17 March 2026.

[98] B. Huang, Z. Jia, and W. Li. Conjugate gradient normal residual method for solving quaternion Toeplitz-type linear systems with application to color image processing. *J. Sci. Comput.*, 103(1), 2025. Article 12.

[99] T. K. Huckle. Compact Fourier analysis for designing multigrid methods. *SIAM J. Sci. Comput.*, 31(1):644–666, 2008.

[100] A. Ilyas, M. F. Khan, V. Loi, and S. Serra-Capizzano. A maximal distribution result and numerical tests for geometric means of HPD GLT matrix-sequences with degenerate commuting/non-commuting GLT symbols. *Linear Algebra Appl.*, 2025. In press, corrected proof; available online 8 August 2025; arXiv:2505.03256v1 [math.NA].

[101] C. Jagels and L. Reichel. Szegö-Lobatto quadrature rules. *J. Comput. Appl. Math.*, 200(1):116–126, 2007.

[102] C. Jagels, L. Reichel, and T. Tang. Generalized averaged Szegő quadrature rules. *J. Comput. Appl. Math.*, 311:645–654, 2017.

[103] Z. Jia and M. K. Ng. Structure preserving quaternion generalized minimal residual method. *SIAM J. Matrix Anal. Appl.*, 42(2):616–634, 2021.

[104] Z. Jia, M. K. Ng, and G.-J. Song. Robust quaternion matrix completion with applications to image inpainting. *Numer. Linear Algebra Appl.*, 26(4):e2245, 2019.

[105] W. B. Jones, O. Njåstad, and W. J. Thron. Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle. *Bull. Lond. Math. Soc.*, 21(2):113–152, 1989.

[106] M. A. Khan and S. Khan. Quaternion Toeplitz matrices and their fundamental properties. *Filomat*, 40(4):1287–1300, 2026. arXiv:2510.04297v1 [math.FA].

[107] M. A. Khan and A. Yagoub. Algebras of Toeplitz matrices with quaternion entries. *J. Math. Ext.*, 18(10):1–19, 2024. Published online 5 February 2025; arXiv:2402.11332v1 [math.RA].

[108] S.-M. Kim and L. Reichel. Anti-Szegö quadrature rules. *Math. Comp.*, 76(258):795–810, 2007.

[109] R. Kress. *Linear Integral Equations*. Springer, 1999.

[110] A. B. J. Kuijlaars and S. Serra-Capizzano. Asymptotic zero distribution of orthogonal polynomials with discontinuously varying recurrence coefficients. *J. Approx. Theory*, 113(1):142–155, 2001.

[111] K. Kumar, M. N. N. Namboodiri, and S. Serra-Capizzano. Preconditioners and Korovkin-type theorems for infinite-dimensional bounded linear operators via completely positive maps. *Stud. Math.*, 218(2):95–118, 2013.

[112] A. Lailoune, V. Loi, and S. Serra-Capizzano. Weyl distributions, spectral properties, and circulant approximation results for quaternion block multilevel Toeplitz matrix sequences. *Boletín de la Sociedad Matemática Mexicana*, forthcoming. Accepted for publication; preprint arXiv:2511.22422v3 [math.NA], first posted 27 November 2025 and last revised 7 August 2026.

[113] N. Lamsahel, C. Manni, A. Ratnani, S. Serra-Capizzano, and H. Speleers. Outlier-free isogeometric discretizations for Laplace eigenvalue problems: closed-form eigenvalue and eigenvector expressions. *Numer. Math.*, 157(4):1397–1448, 2025.

[114] D. P. Laurie. Anti-Gaussian quadrature formulas. *Math. Comp.*, 65(214):739–747, 1996.

[115] H. C. Lee. Eigenvalues and canonical forms of matrices with quaternion coefficients. *Proc. R. Ir. Acad. Sect. A*, 52:253–260, 1949.

[116] T. Li and Q.-W. Wang. Structure preserving quaternion full orthogonalization method with applications. *Numer. Linear Algebra Appl.*, 30(5):e2495, 2023.

[117] T. Li and Q.-W. Wang. Structure preserving quaternion biconjugate gradient method. *SIAM J. Matrix Anal. Appl.*, 45(1):306–326, 2024.

[118] X.-L. Lin, M. K. Ng, and J. Pan. Hermitian quaternion Toeplitz matrices by quaternion-valued generating functions, 2025. arXiv:2504.15073v2 [math.NA].

[119] Q. Liu, S. Ling, and Z. Jia. Randomized quaternion singular value decomposition for low-rank matrix approximation. *SIAM J. Sci. Comput.*, 44(2):A870–A900, 2022.

[120] T. A. Loring. Factorization of matrices of quaternions. *Expo. Math.*, 30(3):250–267, 2012.

[121] L. Marple. Computing the discrete-time analytic signal via FFT. *IEEE Trans. Signal Process.*, 47(9):2600–2603, 1999.

[122] G. Mastroianni and G. V. Milovanović. *Interpolation Processes: Basic Theory and Applications*. Springer Monographs in Mathematics. Springer Berlin Heidelberg, Berlin, 2008.

[123] G. Mastroianni and D. Occorsio. Optimal systems of nodes for Lagrange interpolation on bounded intervals. a survey. *J. Comput. Appl. Math.*, 134(1–2):325–341, 2001.

[124] G. Mastroianni and M. G. Russo. Lagrange interpolation in some weighted uniform spaces. *Facta Univ. Ser. Math. Inform.*, 12:185–201, 1997.

[125] M. Mazza, A. Ratnani, and S. Serra-Capizzano. Spectral analysis and spectral symbol for the 2D curl-curl (stabilized) operator with applications to the related iterative solutions. *Math. Comp.*, 88(317):1155–1188, 2019.

[126] C. A. Micchelli, Y. Xu, and B. Yu. On computing with the Hilbert spline transform. *Adv. Comput. Math.*, 38(3):623–646, 2013.

[127] S. Miron, J. Flamant, N. Le Bihan, P. Chainais, and D. Brie. Quaternions in signal and image processing: a comprehensive and objective overview. *IEEE Signal Process. Mag.*, 40(6):26–40, 2023.

[128] H. Moghaderi, M. Dehghan, M. Donatelli, and M. Mazza. Spectral analysis and multigrid preconditioners for two-dimensional space-fractional diffusion equations. *J. Comput. Phys.*, 350:992–1011, 2017.

[129] P. Nevai. *Orthogonal Polynomials*, volume 18 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, RI, 1979. No. 213.

[130] P. Nevai. Géza Freud, orthogonal polynomials and Christoffel functions. a case study. *J. Approx. Theory*, 48(1):3–167, 1986.

[131] M. K. Ng. *Iterative methods for Toeplitz systems*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2004.

[132] S. E. Notaris. Anti-Gaussian quadrature formulae based on the zeros of Stieltjes polynomials. *BIT Numer. Math.*, 58(1):179–198, 2018.

[133] D. Noutsos, S. Serra-Capizzano, and P. Vassalos. Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate. *Theor. Comput. Sci.*, 315(2–3):557–579, 2004.

[134] S. Olver. Computing the Hilbert transform and its inverse. *Math. Comp.*, 80(275):1745–1767, 2011.

[135] H. Ono. Algebraic solitary waves in stratified fluids. *J. Phys. Soc. Jpn.*, 39(4):1082–1091, 1975.

[136] J. Pan and M. K. Ng. Separable quaternion matrix factorization for polarization images. *SIAM J. Imaging Sci.*, 16(3):1281–1307, 2023.

[137] J. Pan and M. K. Ng. Block-diagonalization of quaternion circulant matrices with applications. *SIAM J. Matrix Anal. Appl.*, 45(3):1429–1454, 2024.

[138] S. V. Parter. On the distribution of the singular values of Toeplitz matrices. *Linear Algebra Appl.*, 80:115–130, 1986.

[139] M. S. Pranić and L. Reichel. Generalized anti-Gauss quadrature rules. *J. Comput. Appl. Math.*, 284:235–243, 2015.

[140] L. Reichel and M. M. Spalević. A new representation of generalized averaged Gauss quadrature rules. *Appl. Numer. Math.*, 165:614–619, 2021.

[141] L. Reichel and M. M. Spalević. Averaged Gauss quadrature formulas: properties and applications. *J. Comput. Appl. Math.*, 410:114232, 2022.

[142] L. Reichel and M. M. Spalević. On the accuracy of averaged Gauss quadrature rules. *J. Comput. Appl. Math.*, 485:117611, 2026.

[143] A. Rifqui, A. Ratnani, and S. Serra-Capizzano. Block Schwarz methods and preconditioning strategies using generalized locally Toeplitz tools—part I: analysis of the preconditioners and numerical validation, 2026. arXiv:2602.04603v2 [math.NA]; last revised 17 August 2026.

[144] L. Rodman. *Topics in Quaternion Linear Algebra*. Princeton University Press, 2014.

[145] S. Serra-Capizzano. Multi-iterative methods. *Comput. Math. Appl.*, 26(4):65–87, 1993.

[146] S. Serra-Capizzano. Asymptotic results on the spectra of block Toeplitz preconditioned matrices. *SIAM J. Matrix Anal. Appl.*, 20(1):31–44, 1998.

[147] S. Serra-Capizzano. On the extreme eigenvalues of Hermitian (block) Toeplitz matrices. *Linear Algebra Appl.*, 270:109–129, 1998.

[148] S. Serra-Capizzano. Toeplitz preconditioners constructed from linear approximation processes. *SIAM J. Matrix Anal. Appl.*, 20(2):446–465, 1998.

[149] S. Serra-Capizzano. A Korovkin-based approximation of multilevel Toeplitz matrices (with rectangular unstructured blocks) via multilevel trigonometric matrix spaces. *SIAM J. Numer. Anal.*, 36(6):1831–1857, 1999.

[150] S. Serra-Capizzano. A Korovkin-type theory for finite Toeplitz operators via matrix algebras. *Numer. Math.*, 82(1):117–142, 1999.

[151] S. Serra-Capizzano. Spectral and computational analysis of block Toeplitz matrices having nonnegative definite matrix-valued generating functions. *BIT Numer. Math.*, 39(1):152–175, 1999.

[152] S. Serra-Capizzano. Distribution results on the algebra generated by Toeplitz sequences: a finite-dimensional approach. *Linear Algebra Appl.*, 328(1–3):121–130, 2001.

[153] S. Serra-Capizzano. Spectral behavior of matrix-sequences and discretized boundary value problems. *Linear Algebra Appl.*, 337:37–78, 2001.

[154] S. Serra-Capizzano. Matrix algebra preconditioners for multilevel Toeplitz matrices are not superlinear. *Linear Algebra Appl.*, 343–344:303–319, 2002.

[155] S. Serra-Capizzano. More inequalities and asymptotics for matrix-valued linear positive operators: the noncommutative case. In A. Böttcher, I. Gohberg, and P. Junghanns, editors, *Toeplitz Matrices and Singular Integral Equations*, volume 135 of *Operator Theory: Advances and Applications*, pages 293–315. Birkhäuser, Basel, 2002.

[156] S. Serra-Capizzano. Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations. *Linear Algebra Appl.*, 366:371–402, 2003.

[157] S. Serra-Capizzano. The GLT class as a generalized Fourier analysis and applications. *Linear Algebra Appl.*, 419(1):180–233, 2006.

[158] S. Serra-Capizzano. The spectral approximation of multiplication operators via asymptotic (structured) linear algebra. *Linear Algebra Appl.*, 424(1):154–176, 2007.

[159] S. Serra-Capizzano and C. Tablino-Possio. Spectral and structural analysis of high precision finite difference matrices for elliptic operators. *Linear Algebra Appl.*, 293(1–3):85–131, 1999.

[160] S. Serra-Capizzano and P. Tilli. On unitarily invariant norms of matrix-valued linear positive operators. *J. Inequal. Appl.*, 7(3):309–330, 2002.

[161] B. Simon. *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, volume 54 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2005.

[162] B. Simon. Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle. *J. Math. Anal. Appl.*, 329(1):376–382, 2007.

[163] M. M. Spalević. On generalized averaged Gaussian formulas. *Math. Comp.*, 76(259):1483–1492, 2007.

[164] M. M. Spalević. On generalized averaged Gaussian formulas. II. *Math. Comp.*, 86(306):1877–1885, 2017.

[165] X. Sun and P. Dang. Numerical stability of circular Hilbert transform and its application to signal decomposition. *Appl. Math. Comput.*, 359:357–373, 2019.

[166] G. Szegő. *Orthogonal Polynomials*, volume 23 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, fourth edition, 1975.

[167] P. Tilli. Locally Toeplitz sequences: spectral properties and applications. *Linear Algebra Appl.*, 278(1–3):91–120, 1998.

[168] P. Tilli. A note on the spectral distribution of Toeplitz matrices. *Linear Multilinear Algebra*, 45(2–3):147–159, 1998.

[169] E. E. Tyrtyshnikov. A unifying approach to some old and new theorems on distribution and clustering. *Linear Algebra Appl.*, 232:1–43, 1996.

[170] E. E. Tyrtyshnikov and N. Zamarashkin. Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships. *Linear Algebra Appl.*, 270(1–3):15–27, 1998.

[171] E. E. Tyrtyshnikov and N. L. Zamarashkin. A general equidistribution theorem for the roots of orthogonal polynomials. *Linear Algebra Appl.*, 366:433–439, 2003.

[172] H. A. van der Vorst. *Iterative Krylov Methods for Large Linear Systems*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2003.

[173] P. Vértesi. Optimal Lebesgue constant for Lagrange interpolation. *SIAM J. Numer. Anal.*, 27(5):1322–1331, 1990.

[174] H. S. Wilf. *Mathematics for the Physical Sciences*. John Wiley & Sons, New York, 1962.

[175] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra Appl.*, 251:21–57, 1997.

[176] M. M. Zheng and G. Ni. Block diagonalization of block circulant quaternion matrices and the fast calculation for T-product of quaternion tensors. *J. Sci. Comput.*, 100(3):69, 2024.