FAST STRUCTURED JACOBI-JACOBI TRANSFORMS

JIE SHEN, YINGWEI WANG, AND JIANLIN XIA

Abstract. Jacobi polynomials are frequently used in scientific and engineering applications, and often times, one needs to use the so called Jacobi-Jacobi transforms which are transforms between two Jacobi expansions with different indices. In this paper, we develop a fast structured algorithm for Jacobi-Jacobi transforms. The algorithm is based on two main ingredients. (i) Derive explicit formulas for connection matrices of two Jacobi expansions with arbitrary indices. In particular, if the indices have integer differences, the connection matrices are relatively sparse or highly structured. The benefit of simultaneous promotion or demotion of the indices is shown. (ii) If the indices have non-integer differences, we explore analytically or numerically a low-rank property hidden in the connection matrices. Combining these two ingredients, we develop a fast structured Jacobi-Jacobi transform with nearly linear complexity, after an one-time precomputation with quadratic complexity, between coefficients of two Jacobi expansions with arbitrary indices. An important byproduct of the fast Jacobi-Jacobi transform is the fast Jacobi transform between the function values at a set of Chebyshev-Gauss-type points and coefficients of the Jacobi expansion with arbitrary indices. Ample numerical results are presented to illustrate the computational efficiency and accuracy of our algorithm.

1. Introduction

Jacobi polynomials have been used in many areas of mathematics and applied sciences, e.g., approximation theory [15, 16], the resolution of Gibbs' phenomenon [13], electrocardiogram data compression [43], and spectral methods for numerical partial differential equations [8, 35, 18]. See also [44, 29] which include extended lists of related work.

Many applications of Jacobi polynomials require transforms between the coefficients of Jacobi expansions and the values at Jacobi-Gauss-type points, and/or between coefficients of Jacobi expansions with different indices. Some examples are as follows. In spectral/spectral-element methods with triangles or tetrahedrons [18, 23], one uses the Koornwinder polynomials [21] (often known as Dubiner's polynomials in the spectral community [11]) which involve Jacobi polynomials with varying indices. In solving prolate spheroidal equations with large zonal wave numbers [6] which arise, e.g., from Helmholtz equations with large wave numbers, one

2010 *Mathematics Subject Classification.* Primary 65T50, 65D05, 65N35, 65F30.

Key words and phrases. Jacobi-Jacobi transform, Chebyshev-Jacobi transform, recurrence formula, connection matrix, low-rank property, structured matrix.

The work of Jie Shen is partially supported by NSF grants DMS-1620262, DMS-1720442 and AFOSR FA9550-16-1-0102.

The research of Jianlin Xia is supported in part by NSF CAREER Award DMS-1255416.

is led to use Jacobi polynomials with very large indices [1]. In dealing with singular solutions such as corner singularities or fractional PDEs, one often needs to work with the approximation space spanned by $\{(1 \pm x)^{k\alpha}\}$ with $\alpha \in (0,1)$, which requires transformation between Jacobi polynomials with index $(0,\alpha)$ or $(\alpha,0)$ and Legendre polynomials [36]. The Jacobi polynomials with index $(10,10)$ outperform those with index $(0,0)$ (Legendre polynomials) for the approximation to the truncated standard Gaussian function (see Appendix C in [54]).

Hence, it is highly desirable to develop algorithms which can perform these transforms as quickly and accurately as possible. Moreover, fast Jacobi-Jacobi transforms may also be useful for the development of fast spherical harmonic transforms which are of critical importance in many applications [31, 47]. The purpose of this work is thus to design fast Jacobi-Jacobi transforms.

Let $\mathbb{P}_N$ be the set of polynomials with degrees less than or equal to N, and (α_1,β_1), (α_2,β_2) be two pairs of Jacobi indices with $\alpha_i,\beta_i > -1$ $(i=1,2)$. For $f \in \mathbb{P}_N$, we can expand $f(x)$ in Jacobi expansions with indices (α_1,β_1) and (α_2,β_2), namely:

$$f(x) = \sum_{n=0}^{N} f_n^{(\alpha_1,\beta_1)} J_n^{(\alpha_1,\beta_1)}(x) = \sum_{n=0}^{N} f_n^{(\alpha_2,\beta_2)} J_n^{(\alpha_2,\beta_2)}(x), \quad x \in [-1,1], \tag{1.1}$$

where $\{J_n^{(\alpha_i,\beta_i)}\}_{n=0}^N$ are Jacobi polynomials with indices (α_i,β_i), $i=1,2$.

By the orthogonal properties of Jacobi polynomials, one can easily determine the *connection matrices* $\boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}$ and $\boldsymbol{K}^{(\alpha_2,\beta_2)\to(\alpha_1,\beta_1)}$ such that

$$\mathbf{f}^{(\alpha_2,\beta_2)} = \boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}\mathbf{f}^{(\alpha_1,\beta_1)}, \ \mathbf{f}^{(\alpha_1,\beta_1)} = \boldsymbol{K}^{(\alpha_2,\beta_2)\to(\alpha_1,\beta_1)}\mathbf{f}^{(\alpha_2,\beta_2)} \tag{1.2}$$

where $\mathbf{f}^{(\alpha_i,\beta_i)} = \{f_n^{(\alpha_i,\beta_i)}\}_{n=0}^N$, $i=1,2$. However, the connection matrices are full upper triangular matrices so that a direct Jacobi-Jacobi transform will cost $O(N^2)$. The main question we want to address in this paper is how to quickly and accurately perform the Jacobi-Jacobi transforms (1.2).

A pioneering work in this direction is done by Alpert and Rokhlin in [2] where a fast transform between Legendre and Chebyshev coefficients is proposed based on the fast multipole method (FMM). Another approach to compute the connection between classical orthogonal polynomials or associated functions is based on the observation that the corresponding connection matrix $\boldsymbol{K}$ can be represented as a properly scaled semiseparable eigenvector matrix. Employing a divide-and-conquer algorithm [10] to compute this eigen-decomposition enables us to perform the matrix-vector product $\boldsymbol{K}\mathbf{f}$ efficiently for any column vector $\mathbf{f}$. This approach has been used for the transform between associated Legendre functions [31], Gegenbauer polynomials [19, 20] and any families of Hermite, Laguerre, and Gegenbauer polynomials with single parameters [4]. However, there is no efficient extension of this approach to Jacobi transforms which have two parameters. *

Another important category of methods for fast Chebyshev-Legendre transform is to use asymptotic expansions [30, 27, 17]. A significant advantage of this methodology is that it does not require an expensive initialization phase. However, the approach with asymptotic expansions is somewhat limited to the Legendre case, and it is generally difficult to extend to Jacobi polynomials with arbitrary indices.

*It was brought to our attention by one of the referees that the authors of [45] developed very recently a fast polynomial transforms based on Toeplitz and Hankel transforms (see also related work in [48]).

We note that the asymptotic approach was generalized to the Chebyshev-Jacobi transform in [37], which also yields fast evaluations of Jacobi expansions at Gauss-Chebyshev nodes.

The main goal of this paper is to develop fast algorithms, with nearly linear complexity after possibly a one-time nearly quadratic-complexity precomputation step, for Jacobi-Jacobi transforms with arbitrary Jacobi indices. Explicit formulas are derived for connection matrices of two Jacobi expansions with arbitrary indices. We show that, if the indices have integer differences, the connection matrices have banded forms in the promotion case (from lower indices to higher ones) or are related to certain highly structured forms in the demotion case. These banded or structured forms can be used to conveniently perform the Jacobi-Jacobi transforms in $O(N)$ operations. We also show that, when it needs to promote or demote both indices α and β, it is more desirable to perform the promotion or demotion simultaneously for both indices, instead of performing one by one. We show how the simultaneous promotion/demotion is done and demonstrate the saving in the cost.

If the indices have non-integer differences such as in the Chebyshev-Legendre transform and more general cases, the connection matrices are dense. We explore a more general structure that is data sparse. This is based on a so-called *low-rank property* of the connection matrices, i.e., their appropriate off-diagonal blocks have small (numerical) ranks. This property can be verified either analytically or numerically. For connection matrices in Chebyshev-Legendre transforms, we can rigorously show an off-diagonal rank bound by deriving certain expansions of the relevant generating functions.

A useful feature for matrices with the low-rank property is that they can be approximated by rank structured matrices such as *hierarchically semiseparable* (HSS) forms [9, 51]. The HSS approximations enable us to perform the desired transforms with $O(rN)$ memory and $O(rN)$ flops, where r is the maximum off-diagonal numerical rank.

The HSS form also has another benefit in the numerical stability. That is, the numerical errors propagate along a binary tree instead of sequentially. Thus, the backward error of the matrix-vector multiplication in the transforms is proportional to $\log^2 N$ and a low-degree term of r [49]. In contrast, direct dense transforms may suffer from large numerical errors since the backward stability depends on the condition number which could be very large in some cases.

In Jacobi-spectral methods, one often needs to transform between the coefficients of the Jacobi expansion to the values at a given set of collocation points. An important byproduct of the proposed fast Jacobi-Jacobi transform is a fast algorithm to perform the Jacobi transform between the function values at a set of Chebyshev-Gauss-type points and coefficients of the Jacobi expansion with arbitrary indices. More precisely, let $f \in \mathbb{P}_N$ and $\{x_j \in [-1,1]\}_{j=0}^N$ be a set of Chebyshev-Gauss-type points. We need to determine $\{f_n^{(\alpha,\beta)}\}_{n=0}^N$ from $\{f(x_j)\}_{j=0}^N$ or vice versa through

$$f(x_j) = \sum_{n=0}^{N} f_n^{(\alpha,\beta)} J_n^{(\alpha,\beta)}(x_j), \quad 0 \le j \le N. \tag{1.3}$$

To obtain the coefficients of the Jacobi expansion $\{f_n^{(\alpha,\beta)}\}_{n=0}^N$ from the function values $\{f(x_j)\}_{j=0}^N$, we proceed in two steps:

- First, obtain the coefficients $\{f_n^{\mathbf{t}}\}_{n=0}^N$ of the Chebyshev expansion:

$$f(x) = \sum_{n=0}^{N} f_n^{\mathbf{t}} T_n(x)$$

  using the Fast Fourier Transform (FFT) in $O(N \log N)$ operations;
- Second, use the proposed Jacobi-Jacobi transform to determine $\{f_n^{(\alpha,\beta)}\}_{n=0}^N$ from $\{f_n^{\mathbf{t}}\}_{n=0}^N$ through the identity

$$f(x) = \sum_{n=0}^{N} f_n^{\mathbf{t}} T_n(x) = \sum_{n=0}^{N} f_n^{(\alpha,\beta)} J_n^{(\alpha,\beta)}(x), \quad x \in [-1,1]. \tag{1.4}$$

Conversely, one can determine the function values $\{f(x_j)\}_{j=0}^N$ from the coefficients of the Jacobi expansion $\{f_n^{(\alpha,\beta)}\}_{n=0}^N$ by reversing the above steps. Thus, the Jacobi transforms with arbitrary indices can be performed also in nearly linear complexity.

The remaining sections are organized as follows. In Section 2, we derive explicit recurrence formulas of the connection matrices for Jacobi-Jacobi transforms with arbitrary indices, and for Jacobi-Jacobi transforms of indices with integer differences. In Section 3, we explore the low-rank property of the connection matrices. We describe the algorithms for the proposed fast structured Jacobi-Jacobi transforms in Section 4, and present several numerical experiments in Section 5. Finally, we present some conclusions and possible directions for future research in the last section.

We list below some notations used throughout the paper:

- $\mathbb{P}_n$: the space of polynomials of degree at most n.
- $\deg(\cdot)$: the degree of a polynomial.
- $\delta_{i,j}$: the Kronecker delta.
- $\Gamma(\cdot)$: Gamma function defined as $\Gamma(t) = \int_0^\infty x^{t-1} \mathrm{e}^{-x} \mathrm{d}x$.
- $\langle \cdot, \cdot \rangle_\omega$: the inner product with respect to the weight ω, e.g., $\langle f, g \rangle_\omega = \int_{-1}^1 f(x) g(x) \omega(x) \mathrm{d}x$.
- $J_n^{(\alpha,\beta)}(x)$: Jacobi polynomial of degree n with indices $\mathbf{j} = (\alpha, \beta)$.
- $L_n(x) = J_n^{(0,0)}(x)$: Legendre polynomial of degree n, or equivalently, Jacobi polynomial of degree n with indices $\mathbf{l} = (0,0)$.
- $T_n(x) = c_n J_n^{(-1/2,-1/2)}(x)$: Chebyshev polynomial (of the first kind) of degree n, which is proportional to Jacobi polynomial of degree n with indices $\mathbf{t} = (-1/2, -1/2)$ (see (2.14) for details).
- $\boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}$ or $\boldsymbol{K}^{\mathbf{j}_1 \to \mathbf{j}_2}$ with $\mathbf{j}_1 = (\alpha_1, \beta_1)$ and $\mathbf{j}_2 = (\alpha_2, \beta_2)$: the connection matrix from $\{J_n^{(\alpha_1,\beta_1)}\}$ to $\{J_n^{(\alpha_2,\beta_2)}\}$.
- $\mathbf{f}^{(\alpha,\beta)}$: the expansion coefficients of a polynomial $f(x)$ in terms of Jacobi polynomials $\{J_n^{(\alpha,\beta)}\}$; in particular, $\mathbf{f}^{\mathbf{l}}$ means the coefficients of the Legendre expansion and $\mathbf{f}^{\mathbf{t}}$ means the coefficients of the Chebyshev expansion.

2. Connection coefficients

First, let us recall some basic properties of Jacobi polynomials $\{J_n^{(\alpha,\beta)}(x)\}$ associated with real indices $\alpha, \beta > -1$.

(1) The three-term recurrence relations for Jacobi polynomials $\{J_n^{(\alpha,\beta)}(x)\}$ read:

$$J_0^{(\alpha,\beta)}(x)=1,\quad J_1^{(\alpha,\beta)}(x)=\frac{1}{2}(\alpha+\beta+2)x+\frac{1}{2}(\alpha-\beta),$$

$$J_{n+1}^{(\alpha,\beta)}(x)=\left(p_n^{(\alpha,\beta)}x-q_n^{(\alpha,\beta)}\right)J_n^{(\alpha,\beta)}(x)-r_n^{(\alpha,\beta)}J_{n-1}^{(\alpha,\beta)}(x),\quad n\geq 1, \tag{2.1}$$

where the constants are

$$p_n^{(\alpha,\beta)}=\frac{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}{2(n+1)(n+\alpha+\beta+1)}, \tag{2.2}$$

$$q_n^{(\alpha,\beta)}=\frac{(\beta^2-\alpha^2)(2n+\alpha+\beta+1)}{2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)}, \tag{2.3}$$

$$r_n^{(\alpha,\beta)}=\frac{(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)}{(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)}. \tag{2.4}$$

(2) Jacobi polynomials $\{J_n^{(\alpha,\beta)}(x)\}_{n=0}$ are orthogonal with respect to their weights $\omega^{(\alpha,\beta)}$:

$$\langle J_k^{(\alpha,\beta)},J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha,\beta)}}=\gamma_k^{(\alpha,\beta)}\delta_{kj}, \tag{2.5}$$

where

$$\omega^{(\alpha,\beta)}=(1-x)^\alpha(1+x)^\beta, \tag{2.6}$$

$$\gamma_k^{(\alpha,\beta)}=\frac{2^{\alpha+\beta+1}\Gamma(\alpha+k+1)\Gamma(\beta+k+1)}{(\alpha+\beta+2k+1)k!\Gamma(\alpha+\beta+k+1)}. \tag{2.7}$$

Definition 2.1. (Connection coefficients) Let $\mathcal{J}_1=\{J_n^{(\alpha_1,\beta_1)}(x)\}_{n=0}^N$ and $\mathcal{J}_2=\{J_n^{(\alpha_2,\beta_2)}(x)\}_{n=0}^N$ with $\deg(J_n^{(\alpha_1,\beta_1)})=\deg(J_n^{(\alpha_2,\beta_2)})=n$ be two sequences of Jacobi polynomials with respect to inner products $\langle\cdot,\cdot\rangle_{\omega^{(\alpha_1,\beta_1)}}$ and $\langle\cdot,\cdot\rangle_{\omega^{(\alpha_2,\beta_2)}}$ respectively. (The weight functions are defined following (2.6).) Then each $J_j^{(\alpha_1,\beta_1)}\in\mathcal{J}_1$ can be represented as a linear combination of the polynomials $\{J_i^{(\alpha_2,\beta_2)}\}_{i=0}^j\subset\mathcal{J}_2$, i.e.,

$$J_j^{(\alpha_1,\beta_1)}=\sum_{i=0}^{j}\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}J_i^{(\alpha_2,\beta_2)},\quad j=0,1,\ldots,N,$$

where $\mathbf{j}_1=(\alpha_1,\beta_1)$ and $\mathbf{j}_2=(\alpha_2,\beta_2)$. Here, the polynomials in $\mathcal{J}_1$ (and $\mathcal{J}_2$) are called *source* (and *target*) polynomials, and the following matrix is called the *matrix of connection coefficients* or *connection matrix* of degree N from $\mathcal{J}_1$ to $\mathcal{J}_2$:

$$\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}=\left(\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}\right)_{i,j=0}^N\in\mathbb{R}^{(N+1)\times(N+1)}.$$

The following lemma then immediately follows.

Lemma 2.2. *The connection matrix* $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ *defined in Definition 2.1 is given by*

$$\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}=\frac{\langle J_i^{(\alpha_1,\beta_1)},J_j^{(\alpha_2,\beta_2)}\rangle_{\omega^{(\alpha_2,\beta_2)}}}{\langle J_i^{(\alpha_2,\beta_2)},J_i^{(\alpha_2,\beta_2)}\rangle_{\omega^{(\alpha_2,\beta_2)}}}=\frac{\langle J_i^{(\alpha_2,\beta_2)},J_j^{(\alpha_1,\beta_1)}\rangle_{\omega^{(\alpha_2,\beta_2)}}}{\gamma_i^{(\alpha_2,\beta_2)}}.$$

Remark 1. By the orthogonality, we observe that for $i>j$, $\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}=0$, which means that $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ is an upper triangular matrix.

With the matrix $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ in hand, we can handle the so-called *connection problem*, i.e., given any polynomial expressed in the basis of one set of orthogonal polynomials, to compute the coefficients with respect to a different set of orthogonal polynomials. This can be rigorously stated as follows.

Lemma 2.3. *Let $f(x)\in\mathbb{P}_N$, and $\mathcal{J}_1,\mathcal{J}_2$ be the Jacobi polynomial sets given in Definition 2.1. Consider two expansions of $f(x)$ as follows*

$$f(x)=\sum_{n=0}^{N} f_n^{\mathbf{j}_1} J_n^{(\alpha_1,\beta_1)}(x)=\sum_{n=0}^{N} f_n^{\mathbf{j}_2} J_n^{(\alpha_2,\beta_2)}(x). \tag{2.8}$$

Then the matrix $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ given in Lemma 2.2 leads to the transform between two column vectors $\mathbf{f}^{\mathbf{j}_1}=(f_n^{\mathbf{j}_1})_{n=0}^N$ and $\mathbf{f}^{\mathbf{j}_2}=(f_n^{\mathbf{j}_2})_{n=0}^N$, i.e., $\mathbf{f}^{\mathbf{j}_2}=\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}\mathbf{f}^{\mathbf{j}_1}$.

Determining these connection coefficients has been studied extensively as a theoretical problem [22, 41, 42, 32, 25, 28, 26, 5]. We focus on below two special cases that are useful in our algorithms, namely, recurrence formulas of *Jacobi-Jacobi transforms with arbitrary indices*, and explicit formulas of *Jacobi-Jacobi transforms with indices that have integer differences*.

Generally speaking, computing the entries of the connection matrix $\boldsymbol{K}$ defined in Lemma 2.3 explicitly and then applying it to a vector both require $O(N^2)$ storage and $O(N^2)$ operations. In order to obtain fast transforms, we proceed with the following strategies:

- For the transforms between any two Jacobi expansions with indices close to each other, we find that the connection matrix $\boldsymbol{K}$ enjoy the low-rank property and thus can be approximated by a rank structured matrix. The rank structured approximation can be quickly applied to a vector in linear complexity.
- For the problem between Jacobi polynomials with integer differences, we find that either the matrix itself can be written as the product of banded matrices, or the off-diagonal blocks of the matrix are low rank, which implies that the Jacobi-Jacobi transforms can be done in linear complexity.

2.1. **Jacobi-Jacobi transforms with arbitrary indices.** In this subsection, we consider the transform between the coefficients of two Jacobi expansions with different indices, which is a generalization of the *forward* and *backward Chebyshev-Jacobi transforms* (FCJT and BCJT) in (1.4).

Consider two Jacobi expansions for any $f(x)\in\mathbb{P}_N$ shown in (2.8). The connection matrices $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ and $\boldsymbol{K}^{\mathbf{j}_2\to\mathbf{j}_1}$ satisfy the following relations:

$$\mathbf{f}^{\mathbf{j}_1}=\boldsymbol{K}^{\mathbf{j}_2\to\mathbf{j}_1}\mathbf{f}^{\mathbf{j}_2},\quad \mathbf{f}^{\mathbf{j}_2}=\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}\mathbf{f}^{\mathbf{j}_1}. \tag{2.9}$$

Note that explicit formulae for the connection matrices are available (cf. Lemma 7.1.1 in [3]). However, the formulae involve the generalized hypergeometric function which is not easy to deal with. On the other hand, there are also explicit formulae for the connection matrices with the same second index (cf. Theorem 7.1.3 in [3]), but it is not easy to extend it to the general case. Therefore, we shall derive recurrence relations for the general connection matrices using the recurrence relation (2.1) and the orthogonal property (2.5).

Note that some recurrence relations for the connection matrices have been derived in [33]. Since the recurrence relations are frequently used in the sequel, we provide them and their proofs below for the reader's convenience and completeness.

Theorem 2.4. *(Recurrence formulas for Jacobi-Jacobi transform) The nonzero entries of* $\boldsymbol{K}^{\mathbf{j}_2\to\mathbf{j}_1}$ *and* $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ *in* (2.9) *can be generated recursively as follows:*

$$\kappa_{i,j+1}^{\mathbf{j}_2\to\mathbf{j}_1} = \varepsilon_1^{\mathbf{j}_2\to\mathbf{j}_1}\kappa_{i,j-1}^{\mathbf{j}_2\to\mathbf{j}_1} + \varepsilon_2^{\mathbf{j}_2\to\mathbf{j}_1}\kappa_{i-1,j}^{\mathbf{j}_2\to\mathbf{j}_1} + \varepsilon_3^{\mathbf{j}_2\to\mathbf{j}_1}\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} + \varepsilon_4^{\mathbf{j}_2\to\mathbf{j}_1}\kappa_{i+1,j}^{\mathbf{j}_2\to\mathbf{j}_1}, \quad j \ge i,$$
$$\kappa_{i,j+1}^{\mathbf{j}_1\to\mathbf{j}_2} = \varepsilon_1^{\mathbf{j}_1\to\mathbf{j}_2}\kappa_{i,j-1}^{\mathbf{j}_1\to\mathbf{j}_2} + \varepsilon_2^{\mathbf{j}_1\to\mathbf{j}_2}\kappa_{i-1,j}^{\mathbf{j}_1\to\mathbf{j}_2} + \varepsilon_3^{\mathbf{j}_1\to\mathbf{j}_2}\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2} + \varepsilon_4^{\mathbf{j}_1\to\mathbf{j}_2}\kappa_{i+1,j}^{\mathbf{j}_1\to\mathbf{j}_2}, \quad j \ge i,$$

where the coefficients $\{\varepsilon_k^{\mathbf{j}_m\to\mathbf{j}_{\hat m}}\}_{k=1,2,3,4}$ *with* $(m,\hat m) = (1,2)$ *or* $(2,1)$ *are given by*

$$\varepsilon_1^{\mathbf{j}_m\to\mathbf{j}_{\hat m}} = -r_j^m, \qquad \varepsilon_2^{\mathbf{j}_m\to\mathbf{j}_{\hat m}} = \begin{cases} 0, & i=0, \\ \pi_2^{\hat m}(i)p_j^m, & i\ge 1, \end{cases}$$
$$\varepsilon_3^{\mathbf{j}_m\to\mathbf{j}_{\hat m}} = \pi_3^{\hat m}(i)p_j^m - q_j^m, \qquad \varepsilon_4^{\mathbf{j}_m\to\mathbf{j}_{\hat m}} = \pi_4^{\hat m}(i)p_j^m,$$

with the parameters $\{p_j, q_j, r_j\}$ *given in* (2.2)–(2.4)*, and*

$$\pi_2^m(k) = \frac{2k(k+\alpha_m+\beta_m)}{(2k+\alpha_m+\beta_m-1)(2k+\alpha_m+\beta_m)},$$
$$\pi_3^m(k) = \frac{\beta_m^2-\alpha_m^2}{(2k+\alpha_m+\beta_m)(2k+\alpha_m+\beta_m+2)},$$
$$\pi_4^m(k) = \frac{2(k+\alpha_m+1)(k+\beta_m+1)}{(2k+\alpha_m+\beta_m+2)(2k+\alpha_m+\beta_m+3)},$$

for $m = 1,2$ *and* $k = 0,1,2,\dots$.

Moreover, the starting points of the above recurrence formulas are

$$\kappa_{0,0}^{\mathbf{j}_2\to\mathbf{j}_1} = 1, \quad \kappa_{0,1}^{\mathbf{j}_2\to\mathbf{j}_1} = \frac{(\beta_1-\alpha_1)(\alpha_2+\beta_2+2)}{2(\alpha_1+\beta_1+2)} - \frac{\beta_2-\alpha_2}{2},$$
$$\kappa_{1,0}^{\mathbf{j}_2\to\mathbf{j}_1} = 0, \quad \kappa_{1,1}^{\mathbf{j}_2\to\mathbf{j}_1} = \frac{\alpha_2+\beta_2+2}{\alpha_1+\beta_1+2},$$
$$\kappa_{0,0}^{\mathbf{j}_1\to\mathbf{j}_2} = 1, \quad \kappa_{0,1}^{\mathbf{j}_1\to\mathbf{j}_2} = \frac{(\alpha_1+\beta_1+2)(\beta_2-\alpha_2)}{2(\alpha_2+\beta_2+2)} - \frac{\beta_1-\alpha_1}{2},$$
$$\kappa_{1,0}^{\mathbf{j}_1\to\mathbf{j}_2} = 0, \quad \kappa_{1,1}^{\mathbf{j}_1\to\mathbf{j}_2} = \frac{\alpha_1+\beta_1+2}{\alpha_2+\beta_2+2}.$$

Proof. To simplify notations in this proof, denote

$$J_n^m := J_n^{(\alpha_m,\beta_m)}(x), \qquad \omega^m := (1+x)^{\alpha_m}(1-x)^{\beta_m}, \qquad \gamma_n^m := \gamma_n^{\alpha_m,\beta_m},$$
$$p_n^m := p_n^{\alpha_m,\beta_m}, \qquad q_n^m := q_n^{\alpha_m,\beta_m}, \qquad r_n^m := r_n^{\alpha_m,\beta_m},$$

where $n = 0,1,2,\dots$ and $m = 1,2$. The entries of the matrices $\boldsymbol{K}^{\mathbf{j}_2\to\mathbf{j}_1} = \left(\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1}\right)_{i,j=0}^N$ and $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2} = \left(\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}\right)_{i,j=0}^N$ are

$$\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} = \frac{\langle J_i^1, J_j^2\rangle_{\omega^1}}{\|J_i^1\|_{\omega^1}^2}, \quad \kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2} = \frac{\langle J_i^2, J_j^1\rangle_{\omega^2}}{\|J_i^2\|_{\omega^2}^2}, \tag{2.10}$$

respectively.

We focus on $\{\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1}\}_{i,j=0}^N$ first. Denote $\tilde\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} = \langle J_i^1, J_j^2\rangle_{\omega^1}$. Note that the relation between $\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1}$ and $\tilde\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1}$ is

$$\tilde\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} = \|J_i^1\|_{\omega^1}^2 \kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1}. \tag{2.11}$$

Then for $j \geq i \geq 1$,

$$\begin{aligned}
&\tilde{\kappa}_{i,j+1}^{\mathbf{j}_2\to\mathbf{j}_1} = \langle J_i^1, J_{j+1}^2\rangle_{\omega^1}\\
&= \langle J_i^1, (p_j^2 x - q_j^2)J_j^2 - r_j^2 J_{j-1}^2\rangle_{\omega^1}\\
&= \langle xJ_i^1, p_j^2 J_j^2\rangle_{\omega^1} - \langle J_i^1, q_j^2 J_j^2\rangle_{\omega^1} - \langle J_i^1, r_j^2 J_{j-1}^2\rangle_{\omega^1}\\
&= \frac{p_j^2}{p_i^1}\left[\langle J_{i+1}^1, J_j^2\rangle_{\omega^1} + q_i^1\langle J_i^1, J_j^2\rangle_{\omega^1} + r_i^1\langle J_{i-1}^1, J_j^2\rangle_{\omega^1}\right] - q_j^2\langle J_i^1, J_j^2\rangle_{\omega^1} - r_j^2\langle J_i^1, J_{j-1}^2\rangle_{\omega^1}\\
&= \frac{1}{p_i^1}\left[p_j^2\tilde{\kappa}_{i+1,j}^{\mathbf{j}_2\to\mathbf{j}_1} + (p_j^2 q_i^1 - p_i^1 q_j^2)\tilde{\kappa}_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} + p_j^2 r_i^1\tilde{\kappa}_{i-1,j}^{\mathbf{j}_2\to\mathbf{j}_1}\right] - r_j^2\tilde{\kappa}_{i,j-1}^{\mathbf{j}_2\to\mathbf{j}_1}.
\end{aligned}$$

It follows that

$$\kappa_{i,j+1}^{\mathbf{j}_2\to\mathbf{j}_1} = \frac{p_j^2\gamma_{i+1}^1}{p_i^1\gamma_i^1}\kappa_{i+1,j}^{\mathbf{j}_2\to\mathbf{j}_1} + \frac{p_j^2 q_i^1 - p_i^1 q_j^2}{p_i^1}\kappa_{i,j}^{\mathbf{j}_2\to\mathbf{j}_1} + \frac{p_j^2 r_i^1\gamma_{i-1}^1}{p_i^1\gamma_i^1}\kappa_{i-1,j}^{\mathbf{j}_2\to\mathbf{j}_1} - r_j^2\kappa_{i,j-1}^{\mathbf{j}_2\to\mathbf{j}_1}. \tag{2.12}$$

Recall that $x = \frac{2J_1^1 - (\alpha_1-\beta_1)}{\alpha_1+\beta_1+2}$. For $i = 0$, the formula becomes

$$\kappa_{0,j+1}^{\mathbf{j}_2\to\mathbf{j}_1} = \frac{2p_j^2\gamma_1^1}{(\alpha_1+\beta_1+2)\gamma_0^1}\kappa_{1,j+1}^{\mathbf{j}_2\to\mathbf{j}_1} - \left(\frac{p_j^2(\alpha_1-\beta_1)}{\alpha_1+\beta_1+2} + q_j^2\right)\kappa_{0,j}^{\mathbf{j}_2\to\mathbf{j}_1} - r_j^2\kappa_{0,j-1}^{\mathbf{j}_2\to\mathbf{j}_1}, \quad j \geq 1. \tag{2.13}$$

Besides, for $\{\kappa_{i,j}^{\mathbf{j}_1\to\mathbf{j}_2}\}_{i,j=0}^N$, we just need to interchange the indices 1 and 2 in (2.12)–(2.13). □

Remark 2. Thanks to the symmetry property of Jacobi polynomials

$$J_n^{(\alpha,\beta)}(-x) = (-1)^n J_n^{(\beta,\alpha)}(x),$$

the Jacobi polynomial $J_n^{(\alpha,\alpha)}(x)$ (up to a constant, referred to as the Gegenbauer or ultra-spherical polynomial) is an odd function for odd n and an even function for even n. Therefore, for the Chebyshev-Jacobi transforms shown in (1.4), we have $\kappa_{ij}^{\mathbf{t}\to(\alpha,\alpha)} = \kappa_{ij}^{(\alpha,\alpha)\to\mathbf{t}} = 0$ for odd $i+j$.

Note that the Chebyshev polynomials and Jacobi polynomials with indices $\mathbf{t} = (-1/2,-1/2)$ are proportional to each other, i.e.

$$T_n(x) \equiv \frac{\Gamma(1/2)\Gamma(n+1)}{\Gamma(n+1/2)}J_n^{(-1/2,-1/2)}(x), \quad \forall n = 0,1,\ldots \tag{2.14}$$

Hence, we can derive, as a special case of Theorem 2.4, the recurrence relation for the connection matrices between Chebyshev and Jacobi expansions.

Now let us consider the Legendre polynomials $L_n(x) \equiv J_n^{(0,0)}(x)$ as a special case of Jacobi polynomials with indices $\mathbf{l} = (0,0)$.

Corollary 2.5. (Recurrence formulas for Chebyshev-Legendre transforms) The nonzero entries of the connection coefficients of the backward and forward Chebyshev-Legendre transforms, i.e. $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}} = (\kappa_{i,j}^{\mathbf{l}\to\mathbf{t}})$ and $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}} = (\kappa_{i,j}^{\mathbf{t}\to\mathbf{l}})$ can be obtained recursively by

$$\kappa_{i,j}^{\mathbf{l}\to\mathbf{t}} = \frac{2}{c_i\pi}\tilde{\kappa}_{i,j}^{\mathbf{l}\to\mathbf{t}}, \quad \kappa_{i,j}^{\mathbf{t}\to\mathbf{l}} = (i+1/2)\tilde{\kappa}_{i,j}^{\mathbf{t}\to\mathbf{l}},$$

where $c_0 = 2, c_i = 1$ for $i \geq 1$ and

$$\tilde{\kappa}^{\mathbf{l}\to\mathbf{t}}_{i,j+1} = \frac{2j+1}{2j+2}\left(\tilde{\kappa}^{\mathbf{l}\to\mathbf{t}}_{i+1,j} + \tilde{\kappa}^{\mathbf{l}\to\mathbf{t}}_{i-1,j}\right) - \frac{j}{j+1}\tilde{\kappa}^{\mathbf{l}\to\mathbf{t}}_{i,j-1},$$
$$\tilde{\kappa}^{\mathbf{t}\to\mathbf{l}}_{i,j+1} = \frac{2i+2}{2i+1}\tilde{\kappa}^{\mathbf{t}\to\mathbf{l}}_{i+1,j} + \frac{2i}{2i+1}\tilde{\kappa}^{\mathbf{t}\to\mathbf{l}}_{i-1,j} - \tilde{\kappa}^{\mathbf{t}\to\mathbf{l}}_{i,j-1}.$$

In particular, for Chebyshev-Legendre transforms, explicit formulas are given in [2].

Lemma 2.6. *(Explicit formulas for Chebyshev-Legendre transforms) The explicit formula for the entries of backward and forward Chebyshev-Legendre transforms, i.e.* $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}} = (\kappa^{\mathbf{l}\to\mathbf{t}}_{i,j})$ *and* $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}} = (\kappa^{\mathbf{t}\to\mathbf{l}}_{i,j})$*, respectively, are given by the following:*

$$\kappa^{\mathbf{l}\to\mathbf{t}}_{i,j} = \begin{cases} \frac{1}{\pi}\left[\Lambda\left(\frac{j}{2}\right)\right]^2, & \text{if } 0 = i \leq j < n \text{ and } j \text{ is even},\\ \frac{2}{\pi}\Lambda\left(\frac{j-i}{2}\right)\Lambda\left(\frac{j+i}{2}\right), & \text{if } 0 < i \leq j < n \text{ and } i+j \text{ is even},\\ 0, & \text{otherwise},\end{cases}$$
$$\kappa^{\mathbf{t}\to\mathbf{l}}_{i,j} = \begin{cases} 1, & \text{if } i = j = 0,\\ \frac{1}{2}\frac{\sqrt{\pi}}{\Lambda(i)}, & \text{if } 0 < i = j < n,\\ \frac{-j(i+1/2)}{(j+i+1)(j-i)}\Lambda\left(\frac{j-i-2}{2}\right)\Lambda\left(\frac{j+i-1}{2}\right), & \text{if } 0 \leq i < j < n \text{ and } i+j \text{ is even},\\ 0, & \text{otherwise},\end{cases}$$

where the function $\Lambda(\cdot)$ *is defined as*

$$\Lambda(z) = \frac{\Gamma(z+1/2)}{\Gamma(z+1)}. \tag{2.15}$$

Remark 3. The explicit formula for the entries of connection coefficients between Gegenbauer polynomials, which are Jacobi polynomials with indices $\alpha = \beta$, can be found in [20]. Since it is costly to numerically evaluate the function $\Lambda(\cdot)$ defined above, the recurrence formulas given in Corollary 2.5 and Theorem 2.4 are usually employed to generate the corresponding connection coefficients (see for instance, [34]). On the other hand, the explicit formulas given in Lemma 2.6 are useful in analyzing the low-rank property hidden in the matrices $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$ and $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}$ in Section 3.2.

2.2. Jacobi-Jacobi transforms for indices with integer differences.

In this subsection, we consider the connection coefficients between the following Jacobi expansions with indices of integer differences:

$$f(x) = \sum_{n=0}^{N} f_n^{(\alpha,\beta)} J_n^{(\alpha,\beta)}(x) = \sum_{n=0}^{N} f_n^{(\alpha+1,\beta)} J_n^{(\alpha+1,\beta)}(x)$$
$$= \sum_{n=0}^{N} f_n^{(\alpha,\beta+1)} J_n^{(\alpha,\beta+1)}(x) = \sum_{n=0}^{N} f_n^{(\alpha+1,\beta+1)} J_n^{(\alpha+1,\beta+1)}(x).$$

For simplicity, let us denote four cases of Jacobi indices with integer differences as follows:

$$\mathbf{u}_{00} = (\alpha,\beta), \quad \mathbf{u}_{10} = (\alpha+1,\beta), \quad \mathbf{u}_{01} = (\alpha,\beta+1), \quad \mathbf{u}_{11} = (\alpha+1,\beta+1).$$

Then the column vectors of the expansion coefficients above are

$$\mathbf{f}^{\mathbf{u}_{00}} = (f_n^{(\alpha,\beta)})_{n=0}^N, \qquad \mathbf{f}^{\mathbf{u}_{10}} = (f_n^{(\alpha+1,\beta)})_{n=0}^N,$$
$$\mathbf{f}^{\mathbf{u}_{01}} = (f_n^{(\alpha,\beta+1)})_{n=0}^N, \qquad \mathbf{f}^{\mathbf{u}_{11}} = (f_n^{(\alpha+1,\beta+1)})_{n=0}^N.$$

First, let us consider the promotion case, i.e., from coefficients of lower indices to those of higher ones.

Lemma 2.7. *(Promotion relation [40]) The Jacobi polynomials with $n \geq 0$ satisfy*

$$(1-x)J_n^{(\alpha+1,\beta)}(x) = \xi_1^{(\alpha,\beta,n)} J_n^{(\alpha,\beta)}(x) - \xi_0^{(\alpha,\beta,n)} J_{n+1}^{(\alpha,\beta)}(x),$$
$$(1+x)J_n^{(\alpha,\beta+1)}(x) = \xi_2^{(\alpha,\beta,n)} J_n^{(\alpha,\beta)}(x) + \xi_0^{(\alpha,\beta,n)} J_{n+1}^{(\alpha,\beta)}(x),$$

where

$$\xi_0^{(\alpha,\beta,n)} = \frac{2(n+1)}{2n+\alpha+\beta+2}, \quad \xi_1^{(\alpha,\beta,n)} = \frac{2(n+\alpha+1)}{2n+\alpha+\beta+2}, \quad \xi_2^{(\alpha,\beta,n)} = \frac{2(n+\beta+1)}{2n+\alpha+\beta+2}.$$

Note that all of the coefficients $\xi_0^{(\alpha,\beta,n)}$, $\xi_1^{(\alpha,\beta,n)}$, $\xi_0^{(\alpha,\beta,n)}$ are x-independent, allowing us to compute these connection coefficients explicitly.

Theorem 2.8. *The promotion coefficients $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{10}}$ and $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{01}}$ in $\mathbf{f}^{\mathbf{u}_{01}} = \boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{10}}\mathbf{f}^{\mathbf{u}_{00}}$ and $\mathbf{f}^{\mathbf{u}_{11}} = \boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{01}}\mathbf{f}^{\mathbf{u}_{00}}$, respectively, are bidiagonal matrices as follows:*

$$\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{10}} = \begin{pmatrix} \mu^{(\alpha,\beta,0)} & \nu_1^{(\alpha,\beta,0)} & & & \\ & \mu^{(\alpha,\beta,1)} & \nu_1^{(\alpha,\beta,1)} & & \\ & & \ddots & \ddots & \\ & & & \ddots & \nu_1^{(\alpha,\beta,N-1)} \\ & & & & \mu^{(\alpha,\beta,N)} \end{pmatrix}_{(N+1)\times(N+1)},$$

$$\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{01}} = \begin{pmatrix} \mu^{(\alpha,\beta,0)} & \nu_2^{(\alpha,\beta,0)} & & & \\ & \mu^{(\alpha,\beta,1)} & \nu_2^{(\alpha,\beta,1)} & & \\ & & \ddots & \ddots & \\ & & & \ddots & \nu_2^{(\alpha,\beta,N-1)} \\ & & & & \mu^{(\alpha,\beta,N)} \end{pmatrix}_{(N+1)\times(N+1)},$$

where

$$\mu^{(\alpha,\beta,i)} = \frac{\alpha+\beta+i+1}{\alpha+\beta+2i+1}, \; \nu_1^{(\alpha,\beta,i)} = -\frac{\beta+i+1}{\alpha+\beta+2i+3}, \; \nu_2^{(\alpha,\beta,i)} = \frac{\alpha+i+1}{\alpha+\beta+2i+3}.$$

Proof. Similar to the procedure in the proof of Theorem 2.4, we have

$$\begin{aligned} \kappa_{i,j}^{\mathbf{u}_{00}\to\mathbf{u}_{10}} &= \frac{\langle J_i^{(\alpha+1,\beta)},\ J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha+1,\beta)}}}{\|J_i^{\alpha+1,\beta}\|^2_{\omega^{(\alpha+1,\beta)}}} = \frac{\langle (1-x)J_i^{(\alpha+1,\beta)},\ J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha,\beta)}}}{\gamma_i^{\alpha+1,\beta}} \\ &= \frac{\langle \xi_1^{(\alpha,\beta,i)} J_i^{(\alpha,\beta)} - \xi_0^{(\alpha,\beta,i)} J_{i+1}^{(\alpha,\beta)},\ J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha,\beta)}}}{\gamma_i^{\alpha+1,\beta}} \\ &= \frac{\gamma_i^{(\alpha,\beta)}\xi_1^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha+1,\beta)}}\delta_{i,j} - \frac{\gamma_{i+1}^{(\alpha,\beta)}\xi_0^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha+1,\beta)}}\delta_{i+1,j} := \mu^{(\alpha,\beta,i)}\delta_{i,j} + \nu_1^{(\alpha,\beta,i)}\delta_{i+1,j}. \end{aligned}$$

Also, it is easy to see that

$$\kappa_{i,j}^{\mathbf{u}_{00}\to\mathbf{u}_{01}} = \frac{\gamma_i^{(\alpha,\beta)}\xi_2^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha,\beta+1)}}\delta_{i,j} + \frac{\gamma_{i+1}^{(\alpha,\beta)}\xi_0^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha,\beta+1)}}\delta_{i+1,j} := \mu^{(\alpha,\beta,i)}\delta_{i,j} + \nu_2^{(\alpha,\beta,i)}\delta_{i+1,j}.$$

Finally, the parameters $\mu^{(\alpha,\beta,i)}, \nu_1^{(\alpha,\beta,i)}$ and $\nu_2^{(\alpha,\beta,i)}$ can be easily computed thanks to the definitions of $\{\gamma_i^{(\alpha,\beta)}\}$ in (2.7) and $\{\xi_j^{(\alpha,\beta,n)}\}$ in Lemma 2.7. □

The result above implies that the flops to promote an index by one is about $3N$. Next, we consider promoting both indices (α,β) simultaneously.

Lemma 2.9. *(Simultaneous promotion) The relation between* $\{J_n^{(\alpha+1,\beta+1)}\}_{n=0}$ *and* $\{J_n^{(\alpha,\beta)}\}_{n=0}$ *is*

$$(1-x)(1+x)J_n^{(\alpha+1,\beta+1)}(x) = \lambda_1^{\alpha,\beta,n}J_n^{(\alpha,\beta)}(x) + \lambda_2^{\alpha,\beta,n}J_{n+1}^{(\alpha,\beta)}(x) + \lambda_3^{(\alpha,\beta,n)}J_{n+2}^{(\alpha,\beta)}(x),$$

where

$$\lambda_1^{(\alpha,\beta,n)} = \frac{4(n+\alpha+1)(n+\beta+1)}{(2n+\alpha+\beta+2)(2n+\alpha+\beta+3)},$$
$$\lambda_2^{(\alpha,\beta,n)} = \frac{4(n+1)(\alpha-\beta)}{(2n+\alpha+\beta+2)(2n+\alpha+\beta+4)},$$
$$\lambda_3^{(\alpha,\beta,n)} = -\frac{4(n+1)(n+2)}{(2n+\alpha+\beta+3)(2n+\alpha+\beta+4)}.$$

Proof. Applying Lemma 2.7 twice leads to

$$(1+x)(1-x)J_n^{(\alpha+1,\beta+1)} = (1+x)\left(\xi_1^{(\alpha,\beta+1,n)}J_n^{(\alpha,\beta+1)} - \xi_0^{(\alpha,\beta+1,n)}J_{n+1}^{(\alpha,\beta+1)}\right)$$
$$= \xi_1^{(\alpha,\beta+1,n)}\xi_2^{(\alpha,\beta,n)}J_n^{(\alpha,\beta)} + \left(\xi_1^{(\alpha,\beta+1,n)}\xi_0^{(\alpha,\beta,n)} - \xi_0^{(\alpha,\beta+1,n)}\xi_2^{(\alpha,\beta,n+1)}\right)J_{n+1}^{(\alpha,\beta)}$$
$$- \xi_0^{(\alpha,\beta+1,n)}\xi_0^{(\alpha,\beta,n+1)}J_{n+2}^{(\alpha,\beta)}.$$

It implies that

$$\lambda_1^{(\alpha,\beta,n)} = \xi_1^{(\alpha,\beta+1,n)}\xi_2^{(\alpha,\beta,n)},$$
$$\lambda_2^{(\alpha,\beta,n)} = \xi_1^{(\alpha,\beta+1,n)}\xi_0^{(\alpha,\beta,n)} - \xi_0^{(\alpha,\beta+1,n)}\xi_2^{(\alpha,\beta,n+1)},$$
$$\lambda_3^{(\alpha,\beta,n)} = -\xi_0^{(\alpha,\beta+1,n)}\xi_0^{(\alpha,\beta,n+1)}.$$

By algebraic computations, we can get the expressions for the parameters $\lambda_1^{(\alpha,\beta,n)}$, $\lambda_2^{(\alpha,\beta,n)}$, and $\lambda_3^{(\alpha,\beta,n)}$. □

Theorem 2.10. *The promotion coefficient matrix* $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{11}}$ *in the transform* $\mathbf{f}^{\mathbf{u}_{11}} = \boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{11}}\mathbf{f}^{\mathbf{u}_{00}}$ *is a banded upper-triangular matrix with upper bandwidth* 2 *of the form*

$$\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{11}} = \begin{pmatrix} \tilde{\lambda}_1^{(\alpha,\beta,0)} & \tilde{\lambda}_2^{(\alpha,\beta,0)} & \tilde{\lambda}_3^{(\alpha,\beta,0)} & & \\ & \tilde{\lambda}_1^{(\alpha,\beta,1)} & \tilde{\lambda}_2^{(\alpha,\beta,1)} & \ddots & \\ & & \ddots & \ddots & \tilde{\lambda}_3^{(\alpha,\beta,N-2)} \\ & & & \ddots & \tilde{\lambda}_2^{(\alpha,\beta,N-1)} \\ & & & & \tilde{\lambda}_1^{(\alpha,\beta,N)} \end{pmatrix}_{(N+1)\times(N+1)},$$

where the entries are

$$\tilde{\lambda}_1^{(\alpha,\beta,i)} = \frac{(\alpha+\beta+i+1)(\alpha+\beta+i+2)}{(\alpha+\beta+2i+1)(\alpha+\beta+2i+2)},$$
$$\tilde{\lambda}_2^{(\alpha,\beta,i)} = \frac{(\alpha-\beta)(\alpha+\beta+i+2)}{(\alpha+\beta+2i+2)(\alpha+\beta+2i+4)},$$
$$\tilde{\lambda}_3^{(\alpha,\beta,i)} = -\frac{(\alpha+i+2)(\beta+i+2)}{(\alpha+\beta+2i+4)(\alpha+\beta+2i+5)}.$$

Proof. Let us denote $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{11}} = (\kappa_{i,j}^{\mathbf{u}_{00}\to\mathbf{u}_{11}})$. Then it is easy to see that

$$\begin{aligned}
\kappa_{i,j}^{\mathbf{u}_{00}\to\mathbf{u}_{11}} &= \frac{\langle J_i^{(\alpha+1,\beta+1)},\ J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha+1,\beta+1)}}}{\|J_i^{(\alpha+1,\beta+1)}\|^2_{\omega^{(\alpha+1,\beta+1)}}} \\
&= \frac{\langle (1-x)(1+x)J_i^{(\alpha+1,\beta+1)}, J_j^{(\alpha,\beta)}\rangle_{\omega^{(\alpha,\beta)}}}{\gamma_i^{(\alpha+1,\beta+1)}} \\
&= \frac{\langle \lambda_1^{(\alpha,\beta,i)} J_i^{(\alpha,\beta)} + \lambda_2^{(\alpha,\beta,i)} J_{i+1}^{(\alpha,\beta)} + \lambda_3^{(\alpha,\beta,i)} J_{i+2}^{(\alpha,\beta)},\ J_j^{(\alpha,\beta,i)}\rangle_{\omega^{(\alpha,\beta)}}}{\gamma_i^{(\alpha+1,\beta+1)}} \\
&= \frac{\gamma_i^{(\alpha,\beta)}\lambda_1^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha+1,\beta+1)}}\delta_{i,j} + \frac{\gamma_{i+1}^{(\alpha,\beta)}\lambda_2^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha+1,\beta+1)}}\delta_{i+1,j} + \frac{\gamma_{i+2}^{(\alpha,\beta)}\lambda_3^{(\alpha,\beta,i)}}{\gamma_i^{(\alpha+1,\beta+1)}}\delta_{i+2,j} \\
&:= \tilde{\lambda}_1^{(\alpha,\beta,i)}\delta_{i,j} + \tilde{\lambda}_2^{(\alpha,\beta,i)}\delta_{i+1,j} + \tilde{\lambda}_3^{(\alpha,\beta,i)}\delta_{i+2,j},
\end{aligned}$$

where the parameters $\tilde{\lambda}_1^{(\alpha,\beta,i)}$, $\tilde{\lambda}_2^{(\alpha,\beta,i)}$, and $\tilde{\lambda}_3^{(\alpha,\beta,i)}$ can be calculated easily. □

Remark 4. In many cases, we may need the promotion transform from $\mathbf{f}^{(\alpha,\beta)}$ to $\mathbf{f}^{(\alpha+m,\beta+m)}$, where m is a positive integer. If we promote the indices α and β separately, then the total cost is about $6mN$ flops. However, if we promote them simultaneously, then the total cost is about $5mN$.

Now, let us consider the demotion case, i.e., from coefficients of higher indices to those of lower ones. There are two strategies. One is to treat the demotion coefficient matrices as inverses of appropriate promotion coefficient matrices. Another is to explicitly write the structured forms of the corresponding matrices.

Lemma 2.11. *(Demotion relation* [40]*) The Jacobi polynomials satisfy*

$$J_n^{(\alpha+1,\beta)}(x) = \sum_{j=0}^{n} \kappa_{nj}^{(\alpha+1,\beta)\to(\alpha,\beta)} J_j^{(\alpha,\beta)}(x),$$
$$J_n^{(\alpha,\beta+1)}(x) = \sum_{j=0}^{n} \kappa_{nj}^{(\alpha,\beta+1)\to(\alpha,\beta)} J_j^{(\alpha,\beta)}(x),$$

where

$$\kappa_{nj}^{(\alpha+1,\beta)\to(\alpha,\beta)} = \frac{\Gamma(n+\beta+1)}{\Gamma(n+\alpha+\beta+2)}\cdot\frac{(2j+\alpha+\beta+1)\Gamma(j+\alpha+\beta+1)}{\Gamma(j+\beta+1)},$$
$$\kappa_{nj}^{(\alpha,\beta+1)\to(\alpha,\beta)} = (-1)^{n+j}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+\alpha+\beta+2)}\cdot\frac{(2j+\alpha+\beta+1)\Gamma(j+\alpha+\beta+1)}{\Gamma(j+\alpha+1)}.$$

Corollary 2.12. The matrices $\boldsymbol{K}^{\mathbf{u}_{10}\to\mathbf{u}_{00}}$ and $\boldsymbol{K}^{\mathbf{u}_{01}\to\mathbf{u}_{00}}$ are upper triangular matrices with their upper triangular parts given by those of $\boldsymbol{\rho}_1\boldsymbol{\eta}_1^T$ and $\boldsymbol{\rho}_2\boldsymbol{\eta}_2^T$, respectively, where

$$\boldsymbol{\rho}_1 = \left(\frac{(2i+\alpha+\beta+1)\Gamma(i+\alpha+\beta+1)}{\Gamma(i+\beta+1)}\right)_{i=0}^{N}, \quad \boldsymbol{\eta}_1 = \left(\frac{\Gamma(i+\beta+1)}{\Gamma(i+\alpha+\beta+2)}\right)_{i=0}^{N},$$

$$\boldsymbol{\rho}_2 = \left((-1)^i\frac{(2i+\alpha+\beta+1)\Gamma(i+\alpha+\beta+1)}{\Gamma(i+\alpha+1)}\right)_{i=0}^{N}, \quad \boldsymbol{\eta}_2 = \left((-1)^i\frac{\Gamma(i+\alpha+1)}{\Gamma(i+\alpha+\beta+2)}\right)_{i=0}^{N}.$$

To compute the entries in the vectors $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ accurately and efficiently, we should not evaluate the Gamma function $\Gamma(\cdot)$ repeatedly. Instead, we could make use of the recurrence relations. For example, for the vector $\boldsymbol{\rho}_1 = (\rho_j)_{i=0}^N$, we have

$$\rho_0 = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\beta+1)}, \quad \rho_{i+1} = \frac{(2i+\alpha+\beta+3)(i+\alpha+\beta+1)}{(2i+\alpha+\beta+1)(i+\beta+1)}\rho_i, \quad i=0,\ldots,N-1.$$

We can compute $\boldsymbol{\rho}_2, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2$ in a similar manner.

Remark 5. Since $\boldsymbol{K}^{\mathbf{u}_{10}\to\mathbf{u}_{00}}$ and $\boldsymbol{K}^{\mathbf{u}_{01}\to\mathbf{u}_{00}}$ correspond to the upper triangular parts of two rank-1 matrices, they are the so-called *semiseparable* matrices and are highly structured. Simple HSS representations can be analytically written down for these matrices. The demotion case can also be considered in another way. That is, we can treat $\boldsymbol{K}^{\mathbf{u}_{10}\to\mathbf{u}_{00}}$ and $\boldsymbol{K}^{\mathbf{u}_{01}\to\mathbf{u}_{00}}$ as inverses of the bidiagonal promotion coefficient matrices $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{10}}$ and $\boldsymbol{K}^{\mathbf{u}_{00}\to\mathbf{u}_{01}}$, respectively. Thus, the matrix-vector multiplication involves bidiagonal solution. In either case, the computational cost for demoting from $\mathbf{f}^{\alpha+m_\alpha,\beta+m_\beta}$ to $\mathbf{f}^{\alpha,\beta}$ is then no more than $3(m_\alpha+m_\beta)N$ flops, for integers m_α and m_β.

3. Low-rank property of the connection matrices

In this section, we show that the Jacobi-Jacobi connection matrices given in Section 2.1 enjoy a *low-rank* property, which allows us to construct HSS representations or approximations for the matrices. The HSS forms are data sparse in the sense that the dense off-diagonal blocks are in compressed low-rank format. This helps to significantly reduce the algorithmic complexity and storage, and yields structured Jacobi-Jacobi transforms that are both efficient and stable. The matrices in Section 2.2 are banded or have small off-diagonal ranks. The related transforms can be conveniently done. The matrices can also be considered as special HSS forms.

Thus, we focus on the connection matrices in Section 2.1 for Jacobi-Jacobi transforms with arbitrary indices. We show that the maximum off-diagonal numerical ranks of those matrices grow polylogarithmically by numerical verification in Section 3.1 and by theoretical analysis in Section 3.2.

3.1. **Low-rank property of Jacobi-Jacobi connection matrices: numerical verification.**

In order to gain some intuition, let us present the results of a few numerical experiments on the off-diagonal numerical ranks of the connection matrices $\boldsymbol{K}$. Each matrix is hierarchically partitioned into $l_{\max}$ levels of HSS blocks [52], which are block rows or columns excluding the diagonal blocks. (See Figure 2 for an illustration.) At levels $l = 0, 1, \ldots, l_{\max}$ (from the root to the leaf levels), the HSS block rows have row sizes N_l and maximum numerical rank r_l. For convenience, suppose the partition is uniform and $N_l \approx N/2^l$. We will see that when

l decreases and N_l roughly doubles, r_l only increases slightly. Here r_l is said to be a *rank pattern* with respect to l in [51].

We consider (α, β) in two square regions:

- $\Omega_1 = [-1, 0]^2$ with center $\mathbf{t} = (-1/2, -1/2)$, which corresponds to the Chebyshev-Jacobi case;
- $\Omega_2 = [\alpha^* - 1/2, \alpha^* + 1/2] \times [\beta^* - 1/2, \beta^* + 1/2]$ with $\mathbf{j}^* := (\alpha^*, \beta^*) = (3\sqrt{3}, \pi)$.

In Figure 1, we show the rank patterns r_l (versus N_l) for the HSS block rows at level l of the HSS partition, where the relative tolerance for computing the numerical ranks is $\tau = 10^{-8}$, and $N_{l_{\max}} = 20$. Some randomly chosen points $\mathbf{j}$ in the regions are used for the tests. For comparison purposes, we also plot reference lines for $O(\log N_l)$ and $O(\log \log N_l)$. We can observe that the following.

1. In all of the three cases, the numerical ranks r_l for HSS blocks at level l increase very slowly, in fact, much slower than $O(\log N_l)$. Instead, it roughly follows the pattern of $O(\log \log N_l)$ initially in our computation, although not yet analytically justified. The numerical ranks stop to increase when N_l is sufficiently large.
2. The numerical rank patterns r_l related to two sets of indices (α, β) and (α^*, β^*) appear to depend only on the their relative locations, by the comparison of the results from Ω_1 in Figure 1(c-d) and Ω_2 in Figure 1(e-f).

3.2. **Low-rank property of Chebyshev-Legendre connection matrices: theoretical analysis.** We then focus on the Chebyshev-Legendre case and show the low-rank property analytically. The following result is well known and its variations are frequently used in the FMM (see, e.g., [14]). Here, we make it slightly more precise for our case.

Lemma 3.1. *For a given tolerance $\epsilon > 0$, suppose the entries of an $m \times n$ matrix $\boldsymbol{A}$ satisfy the following:*

$$\boldsymbol{A}_{ij} = \sum_{k=1}^{r} f_k(i) g_k(j) + b_{i,j}, \quad 1 \le i \le m, \ 1 \le j \le n, \tag{3.1}$$

where $r < \min(m, n)$, $|b_{i,j}| \le \epsilon$, and $\{f_k\}_{k=1}^r$ and $\{g_k\}_{k=1}^r$ are continuous functions defined on $[1, m]$ and $[1, n]$, respectively. That is, $\boldsymbol{A}_{ij}$ is defined by a function with a separable approximation. Then the numerical rank of $\boldsymbol{A}$ with respect to the tolerance $\epsilon\sqrt{mn}$ is bounded by r.

Proof. The proof is relatively obvious. Eq.(3.1) means that $\boldsymbol{A}$ can be written as

$$\boldsymbol{A} = \boldsymbol{F}\boldsymbol{G}^T + \boldsymbol{B},$$

where

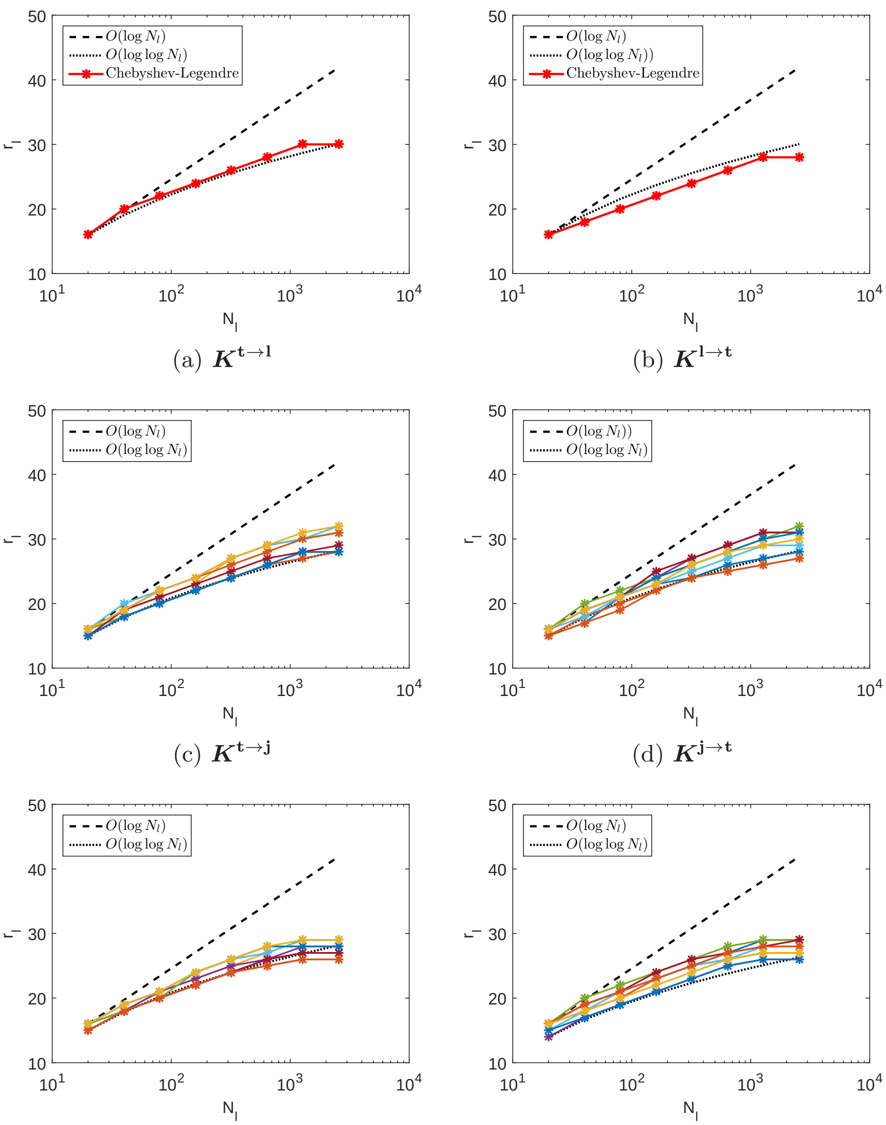
$$\boldsymbol{F} = (f_j(i))_{m \times r}, \quad \boldsymbol{G} = (g_j(i))_{n \times r}, \quad \boldsymbol{B} = (b_{i,j})_{m \times n}.$$

Clearly,

$$||\boldsymbol{B}||_2 \le \sqrt{mn} ||\boldsymbol{B}||_{\max} \le \epsilon\sqrt{mn},$$

where $\|\boldsymbol{B}\|_{\max}$ is the largest entry of the matrix $|\boldsymbol{B}|$. Thus, the numerical rank of $\boldsymbol{A}$ with respect to the tolerance $\epsilon\sqrt{mn}$ is at most r. □

Let us then consider some useful properties of the function $\Lambda(z)$ defined in (2.15).



(a) $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}$ (b) $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$

(c) $\boldsymbol{K}^{\mathbf{t}\to\mathbf{j}}$ (d) $\boldsymbol{K}^{\mathbf{j}\to\mathbf{t}}$

(e) $\boldsymbol{K}^{\mathbf{j}^*\to\mathbf{j}}$ (f) $\boldsymbol{K}^{\mathbf{j}\to\mathbf{j}^*}$

Figure 1: *Numerical ranks r_l of the HSS block rows (of row sizes N_l) at level l of the hierarchical partition of the (a-b) Chebyshev-Legendre, (c-d) Chebyshev-Jacobi, and (e-f) Jacobi-Jacobi connection matrices, where randomly selected indices $\mathbf{j} = (\alpha, \beta) \in \Omega_1$ and $\mathbf{j} = (\alpha, \beta) \in \Omega_2$ are used for the solid lines in (c-d) and (e-f), respectively.*

- $\Lambda(z)$ decreases with respect to z and is bounded, i.e., for any $z \geq 1$,

$$\Lambda(z+1) \leq \Lambda(z) \leq \Lambda(1) = \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}. \tag{3.2}$$

- It can be shown that for integer z,

$$\lim_{z\to\infty} \Lambda(z) = 0. \tag{3.3}$$

- $\Lambda(z)$ has an asymptotic expansion in the following form for any $z \geq 1$ [46, 12]:

$$\Lambda(z) = \sum_{m=0}^{n-1} a_m z^{-(m+1/2)} + \hat{E}_n(z), \tag{3.4}$$

  where

$$a_m = (-1)^m \frac{\Gamma(m+1/2)}{\Gamma(m+1)\Gamma(1/2)} B_m^{(1/2)}\left(\frac{1}{2}\right), \quad \hat{E}_n(z) = O(z^{-(n+1/2)}), \tag{3.5}$$

  with $B_m^{(\mu)}(x)$ being the generalized Bernoulli polynomials [29].
- The first order approximation to $\Lambda(z)$ is

$$\Lambda(z) = z^{-1/2} + O(z^{-3/2}), \tag{3.6}$$

  as $z \to \infty$. It implies that for any given $0 < \epsilon < 1$, there exists an integer

$$\tilde{N}_\epsilon = O(\epsilon^{-2}), \tag{3.7}$$

  such that

$$|\Lambda(z)| < \epsilon \quad \text{for any} \quad z > \tilde{N}_\epsilon. \tag{3.8}$$

The following lemma shows the asymptotic behavior of a_m for sufficiently large m.

Lemma 3.2. *The coefficients* $\{a_m\}$ *in* (3.4) *satisfy*

$$|a_m| = O(m^{m-1}(2\pi \mathrm{e})^{-m}), \quad \text{as } m \to \infty.$$

Proof. We observe from (3.2) that, for $m \geq 1$,

$$\frac{\Gamma(m+1/2)}{\Gamma(m+1)\Gamma(1/2)} = \frac{\Lambda(m)}{\Gamma(1/2)} \leq \frac{\sqrt{\pi}/2}{\sqrt{\pi}} = \frac{1}{2}. \tag{3.9}$$

In addition, the first order approximation to the $B_m^{(1/2)}(1/2)$ is [24]

$$B_m^{(1/2)}(1/2) = \frac{2(m!)}{m^{1/2}(2\pi)^m \Gamma(1/2)} \left[\cos\left(\frac{\pi}{2}(3-m)\right) + O(m^{-1})\right].$$

By Stirling's formula, we have

$$\left| B_m^{(1/2)}\left(\frac{1}{2}\right) \right| = O\left(\frac{m^m}{(2\pi e)^m}\right). \tag{3.10}$$

One can easily draw the conclusion from (3.9) and (3.10). □

Remark 6. The function $\Lambda(z)$ can be computed efficiently by the following way:

- For $z < 15$, one can compute $\Lambda(z)$ directly using the Γ function.
- For $z \geq 15$, one can compute $\Lambda(z)$ via the approximate expansion $\Lambda(z) \approx \sum_{m=0}^{5} a_m z^{-m-1/2}$, where the coefficients $\{a_m\}_{m=0}^{5}$ for z in different intervals are given in double precision in the Appendix of [2].

Lemma 3.3 ([7]). *The function $z^{-\theta}$ with $\theta > 0$, $z \in [1, R]$ can be approximated by a sum of exponentials with r terms, i.e.,*

$$z^{-\theta} = \sum_{i=1}^{r} s_i \exp(-t_i z) + E_{r,[1,R]}(z),$$

where $\{s_i\}_{i=1}^r$ and $\{t_i\}_{i=1}^r$ are non-negative real numbers independent of z, and the error bound is given by

$$|E_{r,[1,R]}(z)| \le 2^{3+\theta} \exp\left(-\frac{\pi^2 r}{\log(8R)}\right). \tag{3.11}$$

Moreover, for the same coefficients $\{s_k\}_{k=1}^r$ and $\{t_k\}_{k=1}^r$, the error bound for $z \in [1, \infty)$ is

$$|E_{r,[1,\infty)}(z)| = \max\left\{|E_{r,[1,R]}(z)|, R^{-\theta}\right\}. \tag{3.12}$$

Next, we consider the approximation of $\Lambda(z)$ by exponential sums.

Lemma 3.4. *The function $\Lambda(z)$ defined in (2.15) can be approximated by an r-term sum of exponentials with respect to a tolerance ϵ:*

$$\Lambda(z) = \sum_{k=1}^{r} w_k \mathrm{e}^{-v_k z} + \mathcal{E}_{r,[2,N]}(z), \quad \left|\mathcal{E}_{r,[2,N]}(z)\right| < \epsilon, \tag{3.13}$$

where $\{w_k\}_{k=1}^r$ and $\{v_k\}_{k=1}^r$ are non-negative real numbers independent of z, and for $2 \le z \le N$,

$$r = O\left(\log^2 \frac{1}{\epsilon} \log\log \frac{1}{\epsilon} \log N\right). \tag{3.14}$$

Moreover, the r-term approximation in (3.13) can be extended from $z \in [2, N]$ to $z \in [2, \infty)$ with the error bound

$$\left|\mathcal{E}_{r,[2,\infty)}(z)\right| < \max\{\epsilon, \Lambda(N)\}. \tag{3.15}$$

Besides, $\left|\mathcal{E}_{r,[2,\infty)}(z)\right| < \epsilon$ with

$$r = \tilde{r}_\epsilon := O\left(\log^3 \frac{1}{\epsilon} \log\log \frac{1}{\epsilon}\right). \tag{3.16}$$

Proof. By setting $\theta = m + 1/2$ in Lemma 3.3, we have for $z \in [2, N]$,

$$z^{-(m+1/2)} = \sum_{i=1}^{r_m} s_{m,i} \mathrm{e}^{-t_{m,i} z} + E_{r_m,[2,N]}(z), \tag{3.17}$$

where we use the subscript m in $r_m, s_{m,i}, E_{r_m,[2,N]}$ to indicate the dependence on m, and the error bound is

$$|E_{r_m,[2,N]}(z)| \le |E_{r_m,[1,N]}(z)| \le 2^{m+7/2} \exp\left(-\frac{\pi^2 r_m}{\log(8N)}\right). \tag{3.18}$$

By (3.4),

$$\begin{aligned}
\Lambda(z) &= \sum_{m=0}^{n-1} a_m \left(\sum_{i=1}^{r_m} s_{m,i} \mathrm{e}^{-t_{m,i} z} + E_{r_m,[2,N]}(z)\right) + \hat{E}_n(z), \\
&= \sum_{m=0}^{n-1} \sum_{i=1}^{r_m} a_m s_{m,i} \mathrm{e}^{-t_{m,i} z} + \sum_{m=0}^{n-1} a_m E_{r_m,[2,N]}(z) + \hat{E}_n(z).
\end{aligned}$$

Thus, $\Lambda(z)$ can be rewritten in the form of (3.13) with

$$\mathcal{E}_{r,[2,N]}(z) = \sum_{m=0}^{n-1} a_m E_{r_m,[2,N]}(z) + \hat{E}_n(z),$$

$$r = \sum_{m=0}^{n-1} r_m. \tag{3.19}$$

We choose n and r so that $|\mathcal{E}_{r,[2,N]}(z)| < \epsilon$ in (3.13) holds for any $z \geq 2$. According to (3.5), we can choose n so that

$$|\hat{E}_n(z)| \leq |\hat{E}_n(2)| = O(2^{-(n+1/2)}).$$

That is,

$$n = O\left(\log \frac{1}{\epsilon}\right). \tag{3.20}$$

According to (3.18), we can choose r_m so that

$$|a_m| 2^{m+7/2} \exp\left(-\frac{\pi^2 r_m}{\log(8N)}\right) = O\left(\frac{\epsilon}{n}\right)$$

That is,

$$r_m = O\left(\log N(\log|a_m| + \log \frac{n}{\epsilon} + m)\right) = O\left(\log N(\log|a_m| + \log\frac{1}{\epsilon} + \log n + m)\right)$$
$$= O\left(\log N(\log|a_m| + \log\frac{1}{\epsilon} + n)\right).$$

For sufficiently large m, the estimate in Lemma 3.2 yields

$$r_m = O\left((\log\frac{1}{\epsilon} + n\log n)\log N\right). \tag{3.21}$$

For m that is not very large, a smaller estimate $r_m = O((\log(1/\epsilon) + n)\log N)$ is obtained since n is the dominate term as compared with $\log|a_m|$.

Combining (3.19), (3.20), and (3.21), we have

$$r = \sum_{m=0}^{n-1} O\left((\log\frac{1}{\epsilon} + n\log n)\log N\right) = O\left(\log^2\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\log N\right).$$

Thus, by choosing r in (3.19) to be (3.14), we can ensure that (3.13) holds.

Moreover, by (3.4), (3.12), and the non-negativity of $\sum_{i=1}^{r} w_k \mathrm{e}^{-v_k z}$, we can conclude (3.15).

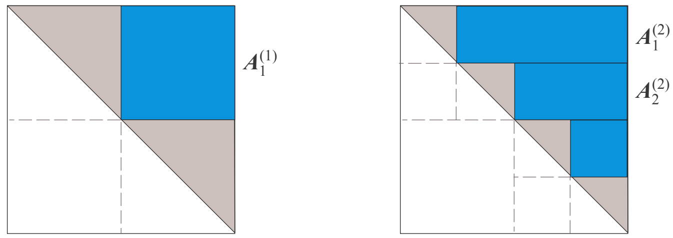
Finally, from (3.6) and $\tilde{N}_\epsilon$ defined in (3.7)–(3.8), if we replace N in (3.14) and (3.15) by $\tilde{N}_\epsilon$, we can obtain $\left|\mathcal{E}_{r,[2,\infty)}(z)\right| < \epsilon$ with the estimate of r in (3.16). □

We are now ready to present the following theorem regarding the Chebyshev-Legendre connection matrices.

Theorem 3.5. *The off-diagonal blocks of the backward and forward Chebyshev-Legendre transform matrices* $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}} = (\kappa_{i,j}^{\mathbf{l}\to\mathbf{t}})$ *and* $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}} = (\kappa_{i,j}^{\mathbf{t}\to\mathbf{l}})$ *are of low numerical ranks. More precisely, for a given tolerance* ϵ*, the numerical ranks of the HSS blocks are*

$$r = O\left(\log^6\frac{1}{\epsilon}\log^2(\log\frac{1}{\epsilon})\right)$$

with respect to the tolerance $\frac{1}{2}N\epsilon$.



(a) 1-level partition (b) 2-level partition

Figure 2: *Two levels of partition of* $\boldsymbol{A} := \boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$*, where the block rows without the diagonal subblocks are HSS blocks, as marked in blue color.*

Proof. Since $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$ is the inverse of $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}$, we only need to show the off-diagonal numerical rank bounds for one of them, and the same bounds also apply to the other case. We focus on $\boldsymbol{A} := \boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$. Note that $\boldsymbol{A}$ is upper triangular. We partition $\boldsymbol{A}$ as in Figure 2 following different levels of hierarchical partition. The HSS block rows at level l are denoted by $\boldsymbol{A}_m^{(l)}$, $m = 1, 2, \ldots, 2^l - 1$. For convenience, assume each $\boldsymbol{A}_m^{(l)}$ has row size N_l. We study the numerical ranks of all the blocks $\boldsymbol{A}_m^{(l)}$.

According to Lemma 2.6, each entry of $\boldsymbol{A}_m^{(l)}$ has the form

$$(\boldsymbol{A}_m^{(l)})_{ij} = \kappa^{\mathbf{l}\to\mathbf{t}}_{(m-1)N_l+i,\ mN_l+j},$$
$$i = 1, 2, \ldots, N_l, \quad j = 1, 2, \ldots, (2^l - m)N_l,$$

and $(\boldsymbol{A}_m^{(l)})_{ij}$ can be generated by a scalar multiple of the following function:

$$\phi_{l,m}(x, y) = \Lambda\left(\frac{1}{2}(N_l + y - x)\right)\Lambda\left(\frac{1}{2}((2m-1)N_l + y + x)\right), \quad (x, y) \in \Theta_{l,m} \cup \Omega_{l,m} \subset \mathbb{R}^2, \tag{3.22}$$

where

$$\Theta_{l,m} := [1, N_l] \times [1, 3], \quad \Omega_{l,m} := [1, N_l] \times [4, (2^l - m)N_l].$$

Here, $(\boldsymbol{A}_m^{(l)})_{ij} = \frac{1}{\pi}\phi_{l,m}(i, j)$ if $i = 1$, and $(\boldsymbol{A}_m^{(l)})_{ij} = \frac{2}{\pi}\phi_{l,m}(i, j)$ if $i > 1$. The difference in the scalar for generating the first row does not impact the study of the rank structure.

Besides, the points $(i, j) \in \Theta_{l,m}$ correspond to the first three columns in the block $(\boldsymbol{A}_m^{(l)})_{ij}$, the rank of which is at most 3. We will focus on the numerical rank of the $\boldsymbol{A}_m^{(l)}$ from the fourth column to the N-th column, where the indices (i, j) are located in the the subdomain $\Omega_{l,m}$.

Furthermore, we define $\tilde{\Omega}_{l,1} = [1, N_l] \times [4, 2N]$. It is easy to see that

$$\Omega_{l,2^l} \subset \cdots \subset \Omega_{l,m+1} \subset \Omega_{l,m} \subset \cdots \subset \Omega_{l,1} \subset \tilde{\Omega}_{l,1}.$$

For fixed l, on each domain $\Omega_{l,m}$, we consider the following two functions:

$$V_m^-(x, y) = \frac{1}{2}(N_l + y - x), \quad V_m^+(x, y) = \frac{1}{2}((2m-1)N_l + y + x), \quad \forall (x, y) \in \Omega_{l,m}.$$

It implies that $\phi_{l,m}(x, y) = \Lambda(V_m^-(x, y))\Lambda(V_m^+(x, y))$, $\forall (x, y) \in \Omega_{l,m}$.

(1) For any $(x,y)\in\Omega_{l,m}$, we can find $x'_-=x, y'_-=y$ such that $(x'_-,y'_-)\in\Omega_{l,1}$ and

$$V_m^-(x,y)=V_1^-(x'_-,y'_-). \tag{3.23}$$

(2) For any $(x,y)\in\Omega_{l,m}$, we can find $x'_+=x, y'_+=2(m-1)N_l+y$ such that $(x'_+,y'_+)\in\tilde{\Omega}_{l,1}$ and

$$V_m^+(x,y)=V_1^+(x'_+,y'_+). \tag{3.24}$$

This indicates that we can consider the case $m=1$ first. That is, we consider the approximation of $\Lambda(V_1^-(x,y))$ on $\Omega_{l,1}$ and $\Lambda(V_1^+(x,y))$ on $\tilde{\Omega}_{l,1}$ by sum of exponentials.

It is easy to know that

$$2\le V_1^-(x,y)\le\frac{1}{2}(N-1),\quad (x,y)\in\Omega_{l,1}, \tag{3.25}$$

$$\frac{1}{2}(N_l+5)\le V_1^+(x,y)\le\frac{1}{2}(N+N_l),\quad (x,y)\in\tilde{\Omega}_{l,1}, \tag{3.26}$$

For sufficient large N, we have $\min\left\{\frac{1}{2}(N-1),\frac{1}{2}(N+N_l)\right\}>\tilde{N}_\epsilon$ for ϵ in (3.7).

Setting $z=V_1^\pm(x,y)$ in Lemma 3.4 leads to the following approximations:

$$\Lambda\left(V_1^-(x,y)\right)=\mathcal{F}_l^-(x,y)+\mathcal{E}_{r^-,[2,\infty)}, \tag{3.27}$$

$$\Lambda\left(V_1^+(x,y)\right)=\mathcal{F}_l^+(x,y)+\mathcal{E}_{r^+,[2,\infty)}, \tag{3.28}$$

where

$$\mathcal{F}_l^-(x,y)=\sum_{k=1}^{r^-}w_k^-\mathrm{e}^{-\frac{1}{2}v_k^-N_l}\mathrm{e}^{\frac{1}{2}v_k^-x}\mathrm{e}^{-\frac{1}{2}v_k^-y}, \tag{3.29}$$

$$\mathcal{F}_l^+(x,y)=\sum_{k=1}^{r^+}w_k^+\mathrm{e}^{-\frac{1}{2}v_k^+N_l}\mathrm{e}^{-\frac{1}{2}v_k^+x}\mathrm{e}^{-\frac{1}{2}v_k^+y}, \tag{3.30}$$

and the superscripts $-$ and $+$ are added to the constants in (3.13) to distinguish the two cases. Moreover, for a given tolerance ϵ, by (3.16), we can choose $r^\pm$ such that

$$r^\pm=O\left(\log^3\frac{1}{\epsilon}\log\log\frac{1}{\epsilon}\right),\quad \max\left\{\mathcal{E}_{r^-,[2,\infty)},\ \mathcal{E}_{r^+,[2,\infty)}\right\}\le\epsilon. \tag{3.31}$$

It follows that

$$\begin{aligned}\phi_{l,1}(x,y)&=\Lambda\left(V_1^-(x,y))\right)\Lambda\left(V_1^+(x,y)\right)\\&=\left(\mathcal{F}_l^-(x,y)+\mathcal{E}_{r^-,[2,\infty)}\right)\left(\mathcal{F}_l^+(x,y)+\mathcal{E}_{r^+,[2,\infty)},\right)\\&=\sum_{k=1}^{r}\tilde{w}_{k,l}\mathrm{e}^{\tilde{v}_{k,l}^{(1)}x}\mathrm{e}^{\tilde{v}_{k,l}^{(2)}y}+\tilde{\mathcal{E}}_r,\end{aligned} \tag{3.32}$$

where the coefficients $\tilde{w}_{k,l}, \tilde{v}_{k,l}^{(1)}, \tilde{v}_{k,l}^{(2)}$ are obtained from the coefficients in (3.29)–(3.30), and

$$r = r^- r^+ = O\left(\log^6 \frac{1}{\epsilon} \log^2(\log \frac{1}{\epsilon})\right), \tag{3.33}$$

$$\begin{aligned} |\tilde{\mathcal{E}}_r| &\le \max\left\{\mathcal{E}_{r^-,[2,\infty)},\ \mathcal{E}_{r^+,[2,\infty)}\right\} \max\left\{\Lambda\left(V_1^-(x,y)\right), \Lambda\left(V_1^+(x,y)\right)\right\} \\ &\le \epsilon\Lambda(1) = \frac{\sqrt{\pi}}{2}\epsilon < \epsilon. \end{aligned} \tag{3.34}$$

We are done for the case with $m = 1$.

For the case with $m \ge 2$, by the relation (3.23), we know that the approximation (3.27) and (3.29) still hold for $\Lambda(V_m^-(x,y)), \forall (x,y) \in \Omega_{l,m}$, with the same estimates of r_l^- and the approximation error in (3.31). Besides, due to (3.24), (3.28), and (3.30), the function $\Lambda(V_m^+(x,y)), \forall (x,y) \in \Omega_{l,m}$, can be approximated by

$$\begin{aligned} \Lambda(V_m^+(x,y)) &= \Lambda(V_1^+(x, 2(m-1)N_l + y)) \\ &= \sum_{k=1}^{r^+} w_k^+ \mathrm{e}^{-\frac{1}{2}v_k^+(2m-1)N_l} \mathrm{e}^{-\frac{1}{2}v_k^+ x} \mathrm{e}^{-\frac{1}{2}v_k^+ y} + \mathcal{E}_{r^+,[2,\infty)}, \end{aligned} \tag{3.35}$$

where r^+ is still given in (3.31). The estimates of $r^\pm$ and approximation errors then do no depend on m. Thus, the rank estimate shown in (3.33) and error estimate shown in (3.34) hold for any $m \ge 1$.

Therefore, the function $\phi_{l,m}(x,y)$ that generates $\boldsymbol{A}_m^{(l)}$ has a separable approximation as in (3.32) with the error bounded by ϵ. According to Lemma 3.1, the numerical rank of the HSS block $\boldsymbol{A}_m^{(l)}$ is then r with respect to the tolerance $\epsilon\sqrt{N_l(N-N_l)} \le \frac{1}{2}N\epsilon$. □

The result above rigorously establishes the low-rank property for the transform matrices of the Chebyshev-Legendre case. The rank bound may be grossly overestimated since the numerical tests indicate much smaller ranks. For more general Chebyshev-Jacobi and Jacobi-Jacobi cases, it is difficult to rigorously prove their low-rank property due to the lack of explicit generating functions for the connection matrices. However, we showed numerically in Figure 1(c-f) that they also enjoy the low-rank property.

The detailed HSS matrix-vector multiplication algorithm can be found in [9, 49]. Since the upper bound of the numerical HSS rank is a constant for a fixed tolerance ϵ, we can conclude that the computational cost of HSS matrix-vector multiplication is linear $O(N)$.

4. Fast structured Jacobi-Jacobi transforms

In this section, we present our fast structured Jacobi-Jacobi transforms (FSJJT) and fast structured Jacobi transforms (FSJT). The studies of the low-rank in the previous section mean that HSS representations or approximations can be computed for the connection matrices.

For the case where the differences in the indices are integers, the connection matrices are banded or highly structured with analytical HSS forms. When the differences in the indices are not integers, a one-time HSS construction is needed. There are two popular ways to construct such an HSS form: direct block compression [52] and randomized sampling [53]. If the maximum off-diagonal numerical rank is r, the constructions cost $O(rN^2)$ flops. For cases like in Theorem 3.5, r is a constant for a fixed tolerance ϵ. Here, since our main purpose is to compute Jacobi-Jacobi transforms, to ensure stability, we spend a one-time cost in the precomputation stage to construct the HSS approximation based on the method in [52]. The method uses rank-revealing factorizations to compress the off-diagonal blocks, which is known to usually yield nearly optimal low-rank approximations. The resulting HSS form can be used to compute matrix-vector multiplications with superior efficiency and stability. The cost to multiply the HSS form and a vector is $O(rN)$.

4.1. **Fast structured Jacobi-Jacobi transforms (FSJJT).** We consider the following forward FSJJT: $\mathbf{f}^{(\alpha_2,\beta_2)} = \boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}\mathbf{f}^{(\alpha_1,\beta_1)}$, where $\mathbf{j}_1 = (\alpha_1,\beta_1)$, $\mathbf{j}_2 = (\alpha_2,\beta_2)$. Let $m_\alpha = \alpha_2 - \alpha_1$, $m_\beta = \beta_2 - \beta_1$ be the difference parameters. Depending on m_α and m_β, our algorithm works as follows.

- *Both m_α and m_β are positive integers.* We can quickly perform the integer promotion with linear cost based on Theorem 2.8. The algorithm has two steps:
  (1) **Initialization.** Compute the vectors as in Theorem 2.8 that define the bidiagonal matrices
  $$\boldsymbol{K}^{(\alpha_1+i,\beta_1)\to(\alpha_1+i+1,\beta)}, \quad \boldsymbol{K}^{(\alpha_2,\beta_1+i)\to(\alpha_2,\beta_1+i+1)}.$$
  (2) **Multiplication.** Promote the index from α_1 to α_2 while fixing β_1 via bidiagonal multiplication:
  $$\mathbf{f}^{(\alpha_2,\beta_1)} = \prod_{i=0}^{m_\alpha-1} \boldsymbol{K}^{(\alpha_1+i,\beta_1)\to(\alpha_1+i+1,\beta_1)}\mathbf{f}^{(\alpha_1,\beta_1)}.$$
  Promote the index from β_1 to β_2 while fixing α_2 via bidiagonal multiplication:
  $$\mathbf{f}^{(\alpha_2,\beta_2)} = \prod_{i=0}^{m_\beta-1} \boldsymbol{K}^{(\alpha_2,\beta_1+i)\to(\alpha_2,\beta_1+i+1)}\mathbf{f}^{(\alpha_2,\beta_1)}.$$
  The total cost is about $3(m_\alpha+m_\beta)N$ flops. This can be further reduced when simultaneous promotion as in Theorem 2.10 is used.
- *Both m_α and m_β are integers, with one or both negative.*
  (1) **Initialization.** Following Remark 5, get the HSS representations based on Corollary 2.12 (for multiplications), or the bidiagonal matrices based on Theorem 2.8 (for solutions).
  (2) **Multiplication.** Then perform the transform via fast HSS matrix-vector multiplications or bidiagonal solutions.
  The total cost is $O((m_\alpha + m_\beta)N)$ flops.
- m_α, $m_\beta \in (-1, 1)$.
  (1) **Initialization.** Compute an HSS approximation to $\boldsymbol{K}^{\mathbf{j}_1\to\mathbf{j}_2}$ via direct or randomized HSS construction in a precomputation.
  (2) **Multiplication.** Then perform fast HSS matrix-vector multiplications for the corresponding transform.
  After the one-time HSS construction, the multiplication cost is at most $O(rN)$, where r is the maximum numerical rank of the HSS blocks.

- *Other cases.* For a general difference parameter, we can split it into the sum of an integer and a number in $(-1, 1)$. This can then be handled by combining the procedures above.

4.2. **Fast structured Jacobi transforms (FSJT).** We now consider the fast structured Jacobi transform (FSJT) between function values $\mathbf{f}^{\mathbf{0}}$ at Chebyshev-Gauss-type points and expansion coefficients $\mathbf{f}^{(\alpha,\beta)}$ of Jacobi polynomials with any indices α, β:

$$\mathbf{f}^{\mathbf{0}} \overset{\text{DCT}}{\longleftrightarrow} \mathbf{f}^{\mathbf{t}} \overset{\text{FSCJT}}{\longleftrightarrow} \mathbf{f}^{(\alpha,\beta)}, \tag{4.1}$$

where $\mathbf{f}^{\mathbf{t}}$ is the vector of Chebyshev expansion coefficients as mentioned at the end of Section 1, DCT means discrete cosine transform which can be done in $O(\log N)$ operations, and FSCJT denotes the fast structured Chebyshev-Jacobi transform which is a special case of FSJJT. The algorithm also has two steps:

(1) **Initialization.** Given the function values $\mathbf{f}^{\mathbf{0}}$, perform the forward Chebyshev transform to obtain the expansion coefficients $\mathbf{f}^{\mathbf{t}}$, and compute an HSS representation/approximation for Chebyshev-Jacobi transform.
(2) **Multiplication.** Perform the forward FSCJT $\mathbf{f}^{(\alpha,\beta)} = \boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}\mathbf{f}^{\mathbf{t}}$ using fast HSS matrix-vector multiplications.

Some additional remarks are in order.

- The initialization stage of the above algorithms has $O((m_\alpha + m_\beta)N)$ complexity if m_α and m_β are integers. Otherwise, a one-time precomputation (HSS construction) of cost $O(rN^2)$ is used.
- According to Remark 4, for the integer parameters m_α and m_β satisfying, say, $m_\alpha \le m_\beta$, we should first promote the indices α, β simultaneously, and then promote the index β, i.e., $\mathbf{f}^{(\alpha,\beta)} \to \mathbf{f}^{(\alpha+m_\alpha,\beta+m_\alpha)} \to \mathbf{f}^{(\alpha+m_\alpha,\beta+m_\beta)}$.
- Our method works for more general Jacobi transforms than the method by Hale and Townsend in [17] for the Chebyshev-Legendre transform (see also [37] for the Chebyshev-Jacobi transform) based on an asymptotic formula. For the special case of Chebyshev-Legendre transform, the method in [17] costs $O(N\log^2 N/\log\log N)$, and our method costs $O(N)$ after the one-time HSS construction.
- In the direct Jacobi transform (1.1), the matrix $\boldsymbol{J}^{(\alpha,\beta)} = (J_n^{(\alpha,\beta)}(x_j))_{n,j=0}^{N-1}$ generally does not have the low-rank property. However, by employing the Chebyshev transform as the intermediate step as in (4.1), we can take advantage of the low-rank property of the connection matrices $\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}$ and $\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}$, and obtain FSJT together with FFT.
- HSS multiplication has superior stability, as shown in [49, 50]. In fact, the backward error only depends on $\log^2 N$ and a low-degree term of r. In comparison, in standard dense matrix-vector multiplication, the backward error depends on the condition number of the matrix.

5. Numerical experiments

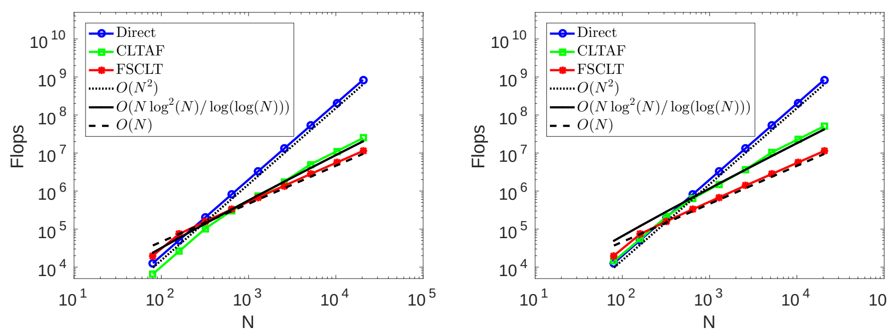
We now present some numerical experiments to illustrate the efficiency and accuracy of our fast structured transforms. When HSS constructions are needed for a matrix $\boldsymbol{A}$ in the tests, a relative tolerance $\tau = 10^{-12}$ is used, and the finest level HSS block row size is about 40.

First, let us start with the Chebyshev-Legendre transform. We test these methods:

- FSCLT: our proposed fast structured Chebyshev-Legendre transform;
- CLTAF: the $O(N \log^2 N / \log \log N)$ complexity method in [17] (which can apply in this special case);
- Direct: the direct Chebyshev-Legendre transform.

In FSCLT, a connection matrix $\boldsymbol{K}$ is approximated by an HSS form in a precomputation stage.

We perform the forward and backward Chebyshev-Legendre transforms with the three methods. The flop counts are given in Figure 3, together with reference lines for $O(N^2)$, $O(N \log^2 N / \log \log N)$, and $O(N)$. The results roughly follow the estimates. When N increases, both FSCLT and CLTAF are much faster than Direct. Besides, FSCLT also has a significant advantage over CLTAF for large N, and both methods achieve comparable accuracies. See Table 1.



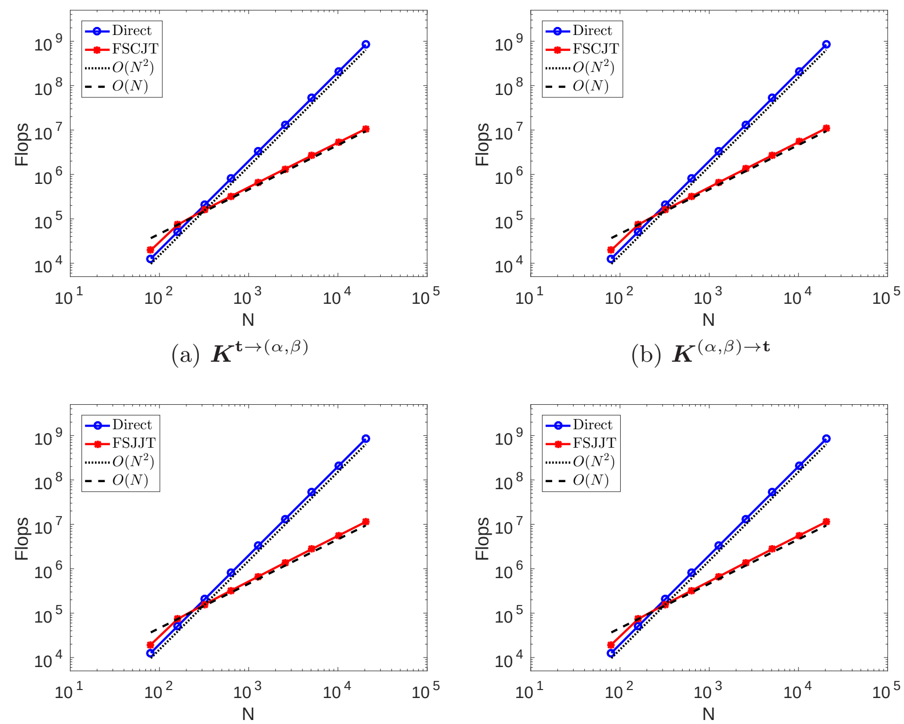
(a) $\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}$ (b) $\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$

Figure 3: *Flops of Chebyshev-Legendre transforms.*

Table 1: *Accuracies of FSCLT and CLTAF for Chebyshev-Legendre transforms of random vectors.*

N	$\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}$		$\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}$	
	FSCLT	CLTAF	FSCLT	CLTAF
160	6.1466e-14	1.5815e-14	4.3208e-14	2.1671e-14
320	2.7861e-13	3.1610e-14	3.0375e-13	1.5178e-13
640	2.2266e-13	2.7652e-14	3.0139e-13	8.5224e-13
1280	1.7845e-13	5.3438e-14	5.7378e-13	7.1993e-13
2560	2.2830e-13	9.9047e-14	1.5104e-12	1.3362e-11
5120	4.1844e-13	1.1875e-13	1.2649e-12	1.3817e-11

Our proposed FSCLT also applies to more general Jacobi-Jacobi transforms than CLTAF. For example, we show the numerical results for the Chebyshev-Jacobi transform with $(\alpha, \beta) = (-\frac{\sqrt{2}}{2}, \frac{\pi}{4})$ in Figure 4(a-b) and the Jacobi-Jacobi transform between $(\alpha_1, \beta_1) = (2, 1)$ and $(\alpha_2, \beta_2) = (3\sqrt{3}, \pi)$ in Figure 4(c-d). We can observe that the costs of our proposed fast structured transforms are nearly linear in N.



(a) $\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}$ (b) $\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}$

(c) $\boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}$ (d) $\boldsymbol{K}^{(\alpha_2,\beta_2)\to(\alpha_1,\beta_1)}$

Figure 4: *Flops of (a-b) Chebyshev-Jacobi transforms and (c-d) Jacobi-Jacobi transforms.*

Next, we demonstrate the benefit of numerical stability of the fast transform method. For relatively large α and/or β, a straightforward Jacobi transform based on the explicit evaluation of the connection matrices as in Lemma 2.2 may have numerical stability issues due to the large values of $J_n^{(\alpha,\beta)}(\pm 1)$ as n increases. On the other hand, our structured transform has nice stability as mentioned before. To illustrate this, we perform a forward transform followed by a backward one as in $\boldsymbol{K}^{(\alpha_2,\beta_2)\to(\alpha_1,\beta_1)}\boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}\mathbf{f}^{(\alpha_1,\beta_1)}$, and measure the error

$$\frac{\|\boldsymbol{K}^{(\alpha_2,\beta_2)\to(\alpha_1,\beta_1)}\boldsymbol{K}^{(\alpha_1,\beta_1)\to(\alpha_2,\beta_2)}\mathbf{f}^{(\alpha_1,\beta_1)} - \mathbf{f}^{(\alpha_1,\beta_1)}\|}{\|\mathbf{f}^{(\alpha_1,\beta_1)}\|},$$

where the norm is the Jacobi weighted norm

$$\|\mathbf{f}^{(\alpha,\beta)}\| = \sqrt{\sum_{n=0}^{N}(f_k^{(\alpha,\beta)})^2\gamma_k^{(\alpha,\beta)}}, \tag{5.1}$$

with $\gamma_k^{(\alpha,\beta)}$ defined in (2.7).

Since we are usually interested in Jacobi transforms of functions with certain smoothness, i.e., with decaying Jacobi coefficients, we take random vectors $\mathbf{f}^{(\alpha,\beta)}$

and scale their entries so that they decay like: n^{-1}, $n^{-1.5}$, $n^{-6.5}$, or $\exp(-36n/N)$. We perform three different transforms (i) Chebyshev-Legendre transform ($(\alpha,\beta)=(0,0)$); (ii) Chebyshev-Jacobi transform with small indices ($(\alpha,\beta)=(-\frac{\sqrt{2}}{2},\frac{\pi}{4})$); (iii) Chebyshev-Jacobi transform with large indices ($(\alpha,\beta)=(10\sqrt{3},10\pi)$). As discussed in Section 4.1, the case (iii) can be split into two steps: the first one is the Chebyshev-Jacobi transform with indices $(\bar{\alpha},\bar{\beta})=(10\sqrt{3}-17,10\pi-31)\approx(0.32,0.42)$, and the second one is the Jacobi-Jacobi transform with integer differences ($m_\alpha=17$, $m_\beta=31$) between (α,β) and $(\bar{\alpha},\bar{\beta})$.

The results are shown in Figure 5. Nice accuracies are observed. For some cases, the errors only slightly increase with the matrix size. This indicates the stability of our fast structured transform, even with respect to relatively large α and/or β.

6. Concluding remarks

In this paper, we developed efficient and robust algorithms for Jacobi-Jacobi transforms with arbitrary indices. To achieve this, we derived explicit formulas for the connection matrices between two Jacobi polynomials with different indices, and then showed that these matrices have the low-rank property. After an one-time precomputation with quadratic complexity, the Jacobi-Jacobi transforms can be accomplished in nearly linear complexity, which is verified ether analytically or numerically.

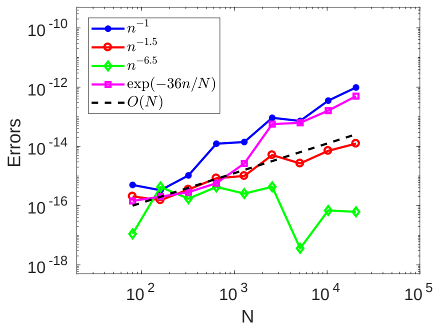
An important byproduct of the proposed fast Jacobi-Jacobi transform is a fast algorithm to perform the transform between the function values at a set of Chebyshev-Gauss-type points and coefficients of the Jacobi expansion with arbitrary indices. Numerical results indicate that our algorithm achieves the desired complexity, and is numerically stable for Jacobi transforms with relatively large indices.

It is apparent that the main techniques and strategies developed in this paper can be applied to many other situations. Indeed, we are currently working on the following situations:

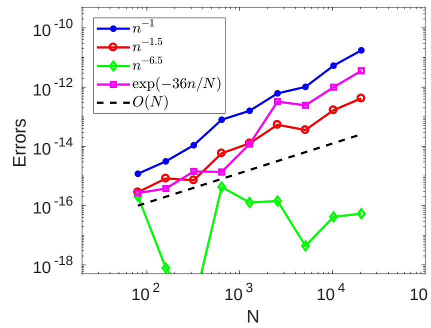
- The strategy of fast structured Jacobi-Jacobi transforms can be used to develop fast transforms between the family of generalized Laguerre polynomials.
- A more difficult problem is to construct a fast spherical harmonic transform. Many attempts have been made in this regard [31, 47, 39, 38], but they are still not fully satisfactory. The main difficulty, as compared with the Jacobi case, is that the spherical harmonic expansion involves associated Legendre polynomials with a full range of indices, rather than a fixed index. It is hopeful that, by exploring the relations between associated Legendre polynomials and Chebyshev polynomials, one can construct a robust and fast structured spherical harmonic transform.

References

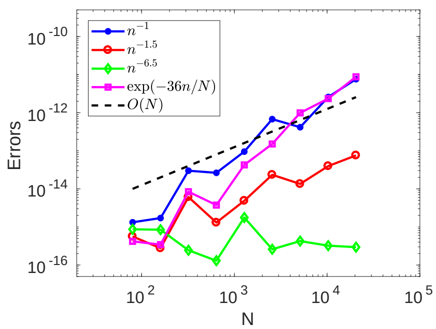
1. H. Alici and J. Shen, *Highly accurate pseudospectral approximations of the prolate spheroidal wave equation for any bandwidth parameter and zonal wavenumber*, J. Sci. Comput. **71** (2017), 804–821.
2. B. Alpert and V. Rokhlin, *A fast algorithm for the evaluation of Legendre expansions*, SIAM Journal on Scientific and Statistical Computing **12** (1991), no. 1, 158–179.
3. George E Andrews, Richard Askey, and Ranjan Roy, *Special functions, volume 71 of encyclopedia of mathematics and its applications*, 1999.



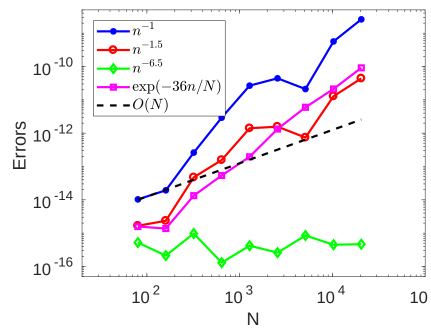
(a) $\frac{\|\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}\mathbf{f}^{\mathbf{l}}-\mathbf{f}^{\mathbf{l}}\|}{\|\mathbf{f}^{\mathbf{l}}\|}$



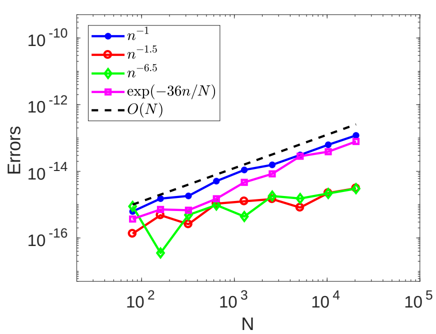
(b) $\frac{\|\boldsymbol{K}^{\mathbf{l}\to\mathbf{t}}\boldsymbol{K}^{\mathbf{t}\to\mathbf{l}}\mathbf{f}^{\mathbf{t}}-\mathbf{f}^{\mathbf{t}}\|}{\|\mathbf{f}^{\mathbf{t}}\|}$



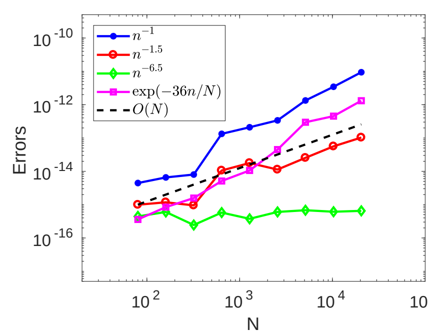
(c) $\frac{\|\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}\mathbf{f}^{(\alpha,\beta)}-\mathbf{f}^{(\alpha,\beta)}\|}{\|\mathbf{f}^{(\alpha,\beta)}\|}$



(d) $\frac{\|\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}\mathbf{f}^{\mathbf{t}}-\mathbf{f}^{\mathbf{t}}\|}{\|\mathbf{f}^{\mathbf{t}}\|}$



(e) $\frac{\|\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}\mathbf{f}^{(\alpha,\beta)}-\mathbf{f}^{(\alpha,\beta)}\|}{\|\mathbf{f}^{(\alpha,\beta)}\|}$



(f) $\frac{\|\boldsymbol{K}^{(\alpha,\beta)\to\mathbf{t}}\boldsymbol{K}^{\mathbf{t}\to(\alpha,\beta)}\mathbf{f}^{\mathbf{t}}-\mathbf{f}^{\mathbf{t}}\|}{\|\mathbf{f}^{\mathbf{t}}\|}$

Figure 5: *Stability test for (a-b) Chebyshev-Legendre transforms, (c-d) Chebyshev-Jacobi transforms with* $(\alpha,\beta)=(-\frac{\sqrt{2}}{2},\frac{\pi}{4})$, *and (e-f) Chebyshev-Jacobi transforms with* $(\alpha,\beta)=(10\sqrt{3},10\pi)$.

4. T. Bella and J. Reis, *The spectral connection matrix for classical orthogonal polynomials of a single parameter*, Linear Algebra and its Applications **458** (2014), 161–182.
5. T. Bella and J. Reis, *The spectral connection matrix for any change of basis within the classical real orthogonal polynomials*, Mathematics **3** (2015), no. 2, 382–397.
6. J. P. Boyd, *Algorithm 840: computation of grid points, quadrature weights and derivatives for spectral element methods using prolate spheroidal wave functions—prolate elements*, ACM Transactions on Mathematical Software (TOMS) **31** (2005), no. 1, 149–165.
7. D. Braess and W. Hackbusch, *On the efficient computation of high-dimensional integrals and the approximation by exponential sums*, Multiscale, nonlinear and adaptive approximation, Springer, 2009, pp. 39–74.
8. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral methods: Fundamentals in single domains*, Springer-Verlag, Berlin-Heidelberg, 2006.
9. S. Chandrasekaran, P. Dewilde, M. Gu, and T. Pals, *A fast decomposition solver for hierarchically semiseparable representations*, SIAM J. Matrix. Anal. Appl. **28** (2006), 603–622.
10. S. Chandrasekaran and M. Gu, *A divide-and-conquer algorithm for the eigendecomposition of symmetric block-diagonal plus semiseparable matrices*, Numerische Mathematik **96** (2004), no. 4, 723–731.
11. M. Dubiner, *Spectral methods on triangles and other domains*, Journal of Scientific Computing **6** (1991), no. 4, 345–390.
12. C. L. Frenzen, *Error bounds for asymptotic expansions of the ratio of two gamma functions*, SIAM journal on mathematical analysis **18** (1987), no. 3, 890–896.
13. D. Gottlieb and C. Shu, *On the Gibbs phenomenon and its resolution*, SIAM Review **39** (1997), no. 4, 644–668.
14. L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, Journal of Computational Physics **135** (1997), no. 2, 280–292.
15. B.-Y. Guo, *Jacobi approximations in certain Hilbert spaces and their applications to singular differential equations*, Journal of Mathematical Analysis and Applications **243** (2000), no. 2, 373–408.
16. B.-Y. Guo and L.-L. Wang, *Jacobi approximations in non-uniformly Jacobi-weighted Sobolev spaces*, Journal of Approximation Theory **128** (2004), no. 1, 1–41.
17. N. Hale and A. Townsend, *A fast, simple, and stable Chebyshev-Legendre transform using an asymptotic formula*, SIAM Journal on Scientific Computing **36** (2014), no. 1, A148–A167.
18. G. E. Karniadakis and S. J. Sherwin, *Spectral/hp element methods for CFD*, second ed., Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.
19. J. Keiner, *Gegenbauer polynomials and semiseparable matrices*, Electronic Transactions on Numerical Analysis **30** (2007), 26–53.
20. ______, *Computing with expansions in Gegenbauer polynomials*, SIAM Journal on Scientific Computing **31** (2009), no. 3, 2151–2171.
21. T. Koornwinder, *Two-variable analogues of the classical orthogonal polynomials*, Theory and applications of special functions (1975), 435–495.
22. S. Lewanowicz, *Quick construction of recurrence relations for the Jacobi coefficients*, Journal of Computational and Applied Mathematics **43** (1992), no. 3, 355–372.
23. H. Li and J. Shen, *Optimal error estimates in Jacobi-weighted Sobolev spaces for polynomial approximations on the triangle*, Mathematics of Computation **79** (2010), no. 271, 1621–1646.
24. J. L. López and N. M. Temme, *Large degree asymptotics of generalized Bernoulli and Euler polynomials*, Journal of Mathematical Analysis and Applications **363** (2010), no. 1, 197–208.
25. P. Maroni and Z. da Rocha, *Connection coefficients between orthogonal polynomials and the canonical sequence: an approach based on symbolic computation*, Numerical Algorithms **47** (2008), no. 3, 291–314.
26. ______, *Connection coefficients for orthogonal polynomials: symbolic computations, verifications and demonstrations in the Mathematica language*, Numerical Algorithms **63** (2013), no. 3, 507–520.
27. A. Mori, R. Suda, and M. Sugihara, *An improvement on Orszag's fast algorithm for Legendre polynomial transform*, Transactions of the Information Processing Society of Japan **40** (1999), no. 9, 3612–3615.
28. A. Narayan and J. S Hesthaven, *Computation of connection coefficients and measure modifications for orthogonal polynomials*, BIT Numerical Mathematics **52** (2012), no. 2, 457–483.

29. F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark (eds.), *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010.
30. S. A. Orszag, *Fast eigenfunction transforms*, Science and Computers, Adv. Math. Supplementary Studies (G. C. Rota, ed.), vol. 10, Academic Press, New York, USA, 1986, pp. 23–30.
31. V. Rokhlin and M. Tygert, *Fast algorithms for spherical harmonic expansions*, SIAM Journal on Scientific Computing **27** (2006), no. 6, 1903–1928.
32. A. Ronveaux, A. Zarzo, and E. Godoy, *Recurrence relations for connection coefficients between two families of orthogonal polynomials*, Journal of Computational and Applied Mathematics **62** (1995), no. 1, 67–73.
33. Herbert E Salzer, *A recurrence scheme for converting from one orthogonal expansion into another*, Communications of the ACM **16** (1973), no. 11, 705–707.
34. J. Shen, *Effcient Chebyshev-Legendre Galerkin methods for elliptic problems*, Proceedings of ICOSAHOM'95 **34** (1996), 233–239.
35. J. Shen, T. Tang, and L.-L. Wang, *Spectral methods: Algorithms, analysis and applications*, Springer Series in Computational Mathematics, no. 41, Springer-Verlag, Berlin-Heidelberg, 2011.
36. J. Shen and Y. Wang, *Müntz-Galerkin methods and applications to mixed Dirichlet-Neumann boundary value problems*, SIAM Journal on Scientific Computing **38** (2016), no. 4, A2357–A2381.
37. R. M. Slevinsky, *On the use of Hahn's asymptotic formula and stabilized recurrence for a fast, simple and stable Chebyshev–Jacobi transform*, IMA Journal of Numerical Analysis (2017), drw070.
38. R. Suda and M. Takami, *A fast spherical harmonics transform algorithm*, Mathematics of Computation **71** (2002), no. 238, 703–715.
39. P. N. Swarztrauber and W. F. Spotz, *Generalized discrete spherical harmonic transforms*, Journal of Computational Physics **159** (2000), no. 2, 213 – 230.
40. Gábor Szegö, *Orthogonal polynomials*, 4th ed., AMS Colloquium Publications 23, vol. 23, American Mathematical Society, Providence, RI, 1975.
41. R. Szwarc, *Connection coefficients of orthogonal polynomials*, Canad. Math. Bull. **35** (1992), no. 4, 548–556.
42. ———, *Connection coefficients of orthogonal polynomials with applications to classical orthogonal polynomials*, Contemp. Math. **183** (1995), 341–346.
43. D. Tchiotsop, D. Wolf, V. Louis-Dorr, and R. Husson, *ECG data compression using Jacobi polynomials*, Proceedings of the 29th Annual International Conference of the IEEE EMBS (Lyon France), 2007, pp. 1863–1867.
44. N. M. Temme, *Special functions: An introduction to the classical functions of mathematical physics*, John Wiley & Sons, Inc., New York, 1996.
45. Alex Townsend, Marcus Webb, and Sheehan Olver, *Fast polynomial transforms based on Toeplitz and Hankel matrices*, Math. Comp., to appear (2017).
46. F. G. Tricomi and A. Erdélyi, *The asymptotic expansion of a ratio of gamma functions*, Pacific J. Math **1** (1951), no. 1, 133–142.
47. M. Tygert, *Fast algorithms for spherical harmonic expansions II*, Journal of Computational Physics **227** (2008), no. 8, 4260–4279.
48. Haiyong Wang and Daan Huybrechs, *Fast and accurate computation of Jacobi expansion coefficients of analytic functions*, (2014).
49. Y. Xi and J. Xia, *On the stability of some hierarchical rank structured matrix algorithms*, SIAM J. Matrix Anal. Appl. **37** (2016), 1279–1303.
50. Y. Xi, J. Xia, S. Cauley, and V. Balakrishnan, *Superfast and stable structured solvers for Toeplitz least squares via randomized sampling*, SIAM J. Matrix. Anal. Appl. **35** (2014), 44–72.
51. J. Xia, *On the complexity of some hierarchical structured matrix algorithms*, SIAM J. Matrix Anal. Appl. **33** (2012), 388–410.
52. J. Xia, S. Chandrasekaran, M. Gu, and X. S. Li, *Fast algorithms for hierarchically semiseparable matrices*, Numer. Linear Algebra Appl. **17** (2010), 953–976.
53. J. Xia, Y. Xi, and M. Gu, *A superfast structured solver for Toeplitz linear systems via randomized sampling*, SIAM J. Matrix. Anal. Appl. **33** (2012), 837–858.
54. D. Xiu and G. E. Karniadakis, *Supersensitivity due to uncertain boundary conditions*, International Journal for Numerical Methods in Engineering **61** (2004), no. 12, 2114–2138.

Department of Mathematics, Purdue University, West Lafayette, IN 47907
E-mail address: shen7@purdue.edu

Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706
E-mail address: wywshtj@gmail.com

Department of Mathematics, Purdue University, West Lafayette, IN 47907
E-mail address: xiaj@purdue.edu