

A Stieltjes algorithm for generating multivariate orthogonal polynomials

Zexin Liu¹ and Akil Narayan²

^{1, 2}Department of Mathematics, and Scientific Computing and Imaging (SCI)
Institute, University of Utah

February 16, 2022

Abstract

Orthogonal polynomials of several variables have a vector-valued three-term recurrence relation, much like the corresponding one-dimensional relation. This relation requires only knowledge of certain recurrence matrices, and allows simple and stable evaluation of multivariate orthogonal polynomials. In the univariate case, various algorithms can evaluate the recurrence coefficients given the ability to compute polynomial moments, but such a procedure is absent in multiple dimensions. We present a new Multivariate Stieltjes (**MS**) algorithm that fills this gap in the multivariate case, allowing computation of recurrence matrices assuming moments are available. The algorithm is essentially explicit in two and three dimensions, but requires the numerical solution to a non-convex problem in more than three dimensions. Compared to direct Gram-Schmidt-type orthogonalization, we demonstrate on several examples in up to three dimensions that the **MS** algorithm is far more stable, and allows accurate computation of orthogonal bases in the multivariate setting, in contrast to direct orthogonalization approaches.

1 Introduction

Orthogonal polynomials are a mainstay tool in numerical analysis and scientific computing, and serve as theoretical and computational foundations for numerical algorithms involving approximation and quadrature [26, 9, 10].

It is well-known even in the multivariate setting that such families of polynomials satisfy three-term recurrence relations [13, 16, 15, 14, 29, 30, 6], which are commonly exploited for stable evaluation and manipulation of such polynomials. Identification or numerical approximation of the coefficients in such relations is therefore of great importance, and in the univariate setting many algorithms for accomplishing such approximations exist [9, 19]. Such procedures are absent in the multivariate setting; this paper provides one algorithmic solution to fill this gap.

1.1 Challenges with computing orthogonal polynomials

Throughout, we assume the ability to compute generalized polynomial moments, i.e., there is some algorithm available to us that evaluates $p \mapsto \int_{\mathbb{R}^d} p(x) d\mu(x)$ for a given positive measure μ on \mathbb{R}^d . This assumption is required for univariate algorithms as well.

With moment information, one could devise a linear algebraic scheme that orthogonalizes some known basis (say monomials) into an orthonormal basis, seemingly providing a solution to the evaluation of orthogonal polynomials. But in finite precision, even stable orthonormalization algorithms can be ineffective due to the high condition number of the map from integral moments to the orthonormal basis; we demonstrate this in Figure 7. Thus, even computational identification of an orthonormal polynomial basis is challenging, let alone computation of recurrence matrices.

Although computing moments with respect to fairly general multivariate measures μ is certainly an open challenge, it is not the focus of this article: We focus on the separate, open challenge of computing recurrence coefficients (allowing stable evaluation of multivariate polynomials) *given* the ability to compute moments.

1.2 Contributions

The main contribution of this paper is to extend existing methods for computing recurrence coefficients from the univariate case to the multivariate case. Recognizing that the *Stieltjes* algorithm for computing recurrence coefficients in univariate problems has shown tremendous success [23, 24, 8], we devise a new, *Multivariate Stieltjes* (**MS**) algorithm for computing recurrence coefficients (matrices in the multivariate setting), and hence also for computing a multivariate orthonormal basis. Thus, our contribution, the Multivariate Stieltjes algorithm, is a new method for tackling the challenge identified in Section 1.1. We demonstrate with several numerical examples the (substantially) improved stability of the **MS** algorithm compared to alternative Gram-Schmidt-type approaches for computing an orthonormal basis.

The tools we employ are, qualitatively, direct multivariate generalizations of existing univariate ideas. However, the technical details in the multivariate case are so disparate from the univariate case that we must employ somewhat different theories and develop new algorithms. Our **MS** algorithm has explicit steps in two and three dimensions, but requires non-convex optimization in four or more dimensions. We first review some existing methods to compute univariate recurrence coefficients (see Section 2.1) and introduce notation, properties, and the three-term relation for multivariate polynomials in Section 2.2. In Section 3, we propose a *canonical basis* that identifies a computational strategy for direct evaluation of multivariate polynomials. We follow this by Section 4, which shows that if μ is tensorial, then a tensor-product basis is in fact a canonical basis. Algorithms are discussed in Section 5; Section 5.1 describes a direct procedure using orthonormalization given polynomial moments. The new multivariate Stieltjes procedure is described in Section 5.2. Finally, we present a wide range of numerical examples in Section 6, which compares these approaches, and demonstrates the improved accuracy of the **MS** procedure.

We mention that our goals are similar to the results in [1, 27], which produce explicit recurrence relations. However, these results are either specialized to certain domains, or use recurrence matrices as known ingredients. Our procedures compute recurrence matrices for general measures, and hence are quite different.

1.3 Assumptions and caveats

Throughout this manuscript, we assume that integral moments of arbitrary polynomials are available/computable. For “simple” domains, we realize this through mapped/tensorized quadrature (which is sometimes exact and sometimes approximate). For more complicated domains, we discretize the measure μ as the empirical measure associated to a large number of realizations that are independently and identically distributed according to μ . Thus, sometimes our numerical examples compute orthogonal polynomials with respect to an approximate measure. However, we use a sufficiently dense grid that such approximation error is relatively small. We emphasize that this approximation error is not the focus of this article; our goal is to devise a scheme that, given the ability to compute moments, accurately computes an orthonormal polynomial basis.

The new **MS** algorithm we develop is effective compared to direct orthonormalization schemes when the condition number of the Gram moment matrix is large. For, e.g., small dimensions d and polynomial degree, this moment matrix typically is not too ill-conditioned, and so there is little benefit in the **MS** algorithm for such situations. However, when one requires polynomials of moderately large degree, or when the Gram matrix is extremely ill-conditioned, we show that the **MS** algorithm is effective.

Finally, we note that the technique we present leverages theories associated with total-degree

spaces of polynomials, and does not directly apply to more exotic spaces. In particular, we assume that μ is non-degenerate with respect to d -variate polynomials. This is ensured if, for example, μ has a positive Lebesgue density over any open ball in \mathbb{R}^d .

2 Background and notation

We use the standard multi-index notation in $d \in \mathbb{N}$ dimensions. With \mathbb{N}_0 the set of nonnegative integers, a multi-index in d dimensions is denoted by $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$. For $\alpha \in \mathbb{N}_0^d$, and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, we write monomials as $x^\alpha = x_1^{\alpha_1}, \dots, x_d^{\alpha_d}$. The number $|\alpha| := \alpha_1 + \dots + \alpha_d$ is the degree of x^α . We denote the space of d -variate polynomials of exactly degree $n \in \mathbb{N}_0$, and up to degree n , respectively, by

$$\mathbb{P}_n^d := \text{span} \{x^\alpha : |\alpha| = n, \alpha \in \mathbb{N}_0^d\}, \quad \Pi_n^d := \text{span} \{x^\alpha : |\alpha| \leq n, \alpha \in \mathbb{N}_0^d\},$$

The dimensions of these spaces are, respectively,

$$r_n = r_n^d := \dim \mathbb{P}_n^d = \binom{n+d-1}{n}, \quad R_n = R_n^d := \dim \Pi_n^d = \binom{n+d}{n} = \sum_{j=0}^n r_j,$$

Since the dimension d will be arbitrary but fixed in our discussion, we will frequently suppress notational dependence on d and write r_n, R_n . We will also require differences between dimensions of subspaces,

$$\Delta r_n := r_n - r_{n-1} = \binom{n+d-2}{n}, \quad n \geq 0, \quad d \geq 2,$$

where we define $r_{-1} := 0$. We will not be concerned with Δr_n when $d = 1$, although one could define $\Delta r_n = 0$ in this case.

Throughout we assume that μ is a given positive measure on \mathbb{R}^d , with $d \in \mathbb{N}$. The support of μ may be unbounded. We assume that, given any non-trivial polynomial p , we have,

$$0 < \langle p, p \rangle < \infty, \quad \langle f, g \rangle := \int f(x)g(x)d\mu(x). \quad (1)$$

and we implicitly assume that $\int q(x)d\mu(x)$ is computationally available for arbitrary polynomials q . Our main goal is to compute an orthonormal polynomial basis, where orthonormality is defined through the inner product above.

2.1 Univariate orthogonal polynomials

In the univariate case, let μ be a positive Borel measure on \mathbb{R} with finite moments. A standard Gram-Schmidt process applied to monomials with the inner product $\langle \cdot, \cdot \rangle$ yields a sequence of orthonormal polynomials $\{p_n(x)\}_{n=0}^\infty$, which satisfies $\langle p_n, p_m \rangle = \delta_{m,n}$, where $\delta_{m,n}$ is the Kronecker delta. In practical settings, it is well-known that utilizing the three-term recurrence formula for evaluation is more computationally stable compared to direct orthonormalization techniques. There exist coefficients b_0 and $\{a_n, b_n\}_{n \in \mathbb{N}}$, with $a_n = a_n(\mu)$ and $b_n = b_n(\mu)$, such that

$$xp_n(x) = b_{n+1}p_{n+1}(x) + a_{n+1}p_n(x) + b_np_{n-1}(x), \quad n \geq 0, \quad (2)$$

where $b_n > 0$ for all $n \in \mathbb{N}_0$, and $p_0 = 1/b_0$, $p_{-1} \equiv 0$. Availability of the coefficients (a_n, b_n) not only enables evaluation via the recurrence above, but also serve as necessary ingredients for various computational approximation algorithms, e.g., quadrature. Thus, knowledge of these coefficients is of great importance in the univariate case. In some cases these coefficients are explicitly known [25], and in cases when they are not, several algorithms exist to accurately compute approximations [4, 20, 21, 28, 7, 12, 9, 19]. The procedure we present in this paper is generalization of the (univariate) Stieltjes procedure [23, 24, 8].

2.2 Multivariate orthogonal polynomials

Now let μ be a measure on \mathbb{R}^d . Again a Gram-Schmidt process applied to the multivariate monomials x^α , $\alpha \in \mathbb{N}_0^d$ produces a sequence of orthogonal polynomials in several variables. However, in the multivariable setting it is more natural to proceed in a degree-graded fashion: We shall say that $p \in \Pi_n^d$ is an orthogonal polynomial of degree $n > 0$ with respect to $d\mu$ if

$$\langle p, q \rangle = 0, \quad \forall q \in \Pi_{n-1}^d.$$

Of course, $p \equiv 1$ is the unique unit-norm degree-0 polynomial with positive leading coefficient. The definition above allows us to introduce, \mathcal{V}_n^d , the space of orthogonal polynomials of degree of exactly n ; that is

$$\mathcal{V}_n^d = \{p \in \Pi_n^d : \langle p, q \rangle = 0, \forall q \in \Pi_{n-1}^d\}.$$

Our assumption (1) on non-degeneracy of μ implies that $\dim \mathcal{V}_n^d = \dim \mathbb{P}_n^d = r_n$. This allows us to state results in terms of size- r_n vector functions containing orthonormal bases of \mathcal{V}_n^d for $n \geq 0$. We fix an(y) orthonormal basis for each $n \in \mathbb{N}_0$,

$$\mathbb{P}_n = (p_{(n,j)})_{j=1}^{r_n} = (p_{(n,1)}, \dots, p_{(n,r_n)})^T, \quad \text{span}\{p_{(n,j)}\}_{j=1}^{r_n} = \mathcal{V}_n^d, \quad (3)$$

This fixed basis \mathbb{P}_n has a three-term recurrence relation analogous to (2): There exist unique matrices $A_{n+1,i} \in \mathbb{R}^{r_n \times r_n}$, $B_{n+1,i} \in \mathbb{R}^{r_n \times r_{n+1}}$, such that

$$x_i \mathbb{P}_n(x) = B_{n+1,i} \mathbb{P}_{n+1}(x) + A_{n+1,i} \mathbb{P}_n(x) + B_{n,i}^T \mathbb{P}_{n-1}(x), \quad i \in [d]. \quad (4)$$

where we define $\mathbb{P}_{-1} = 0$ and $\mathbb{P}_0(x) = 1$. These matrices must satisfy the conditions,

$$\text{rank}(B_{n,i}) = r_{n-1}, \quad \text{rank}(B_n) = r_n, \quad B_n := (B_{n,1}^T, \dots, B_{n,d}^T)^T \in \mathbb{R}^{dr_{n-1} \times r_n}, \quad (5)$$

see [6, Theorem 3.3.4].

Given an orthonormal basis $\{\mathbb{P}_n\}_{n \geq 0}$, there are certain recurrence matrices that make the relation (4) true. However, polynomials generated by (4) for an arbitrary set of matrices $A_{n,i}$ and $B_{n,i}$ need not be orthogonal polynomials. The following are a set of necessary conditions, the *commuting conditions*, that we will require.

Theorem 2.1 ([31, Theorem 2.4]). *If $A_{n,i}$ and $B_{n,i}$ for $i \in [d]$ and $n \in \mathbb{N}_0$ are recurrence matrices corresponding to an orthonormal polynomial sequence, then they satisfy the following conditions for every $i, j \in [d]$ and $n \in \mathbb{N}$:*

$$B_{n+1,i} B_{n+1,j}^T + A_{n+1,i} A_{n+1,j} + B_{n,i}^T B_{n,j} = B_{n+1,j} B_{n+1,i}^T + A_{n+1,j} A_{n+1,i} + B_{n,j}^T B_{n,i}, \quad (6a)$$

$$B_{n,i} A_{n+1,j} + A_{n,i} B_{n,j} = B_{n,j} A_{n+1,i} + A_{n,j} B_{n,i}, \quad (6b)$$

$$B_{n,i} B_{n+1,j} = B_{n,j} B_{n+1,i}, \quad (6c)$$

where (6a) holds for $n = 0$ also.

The conditions above can be derived by expressing certain matrix moments using two different applications of (4). More informally, (4) for a fixed n corresponds to d sets of conditions of size r_n that determine the r_{n+1} degrees of freedom in \mathbb{P}_{n+1} . In order for all these conditions to be consistent, the recurrence matrices must satisfy certain constraints, namely (6).

One starting point for our multivariate Stieltjes algorithm is the following direct observation: Inspection of the recurrence relation (4) reveals that the coefficient matrices can be computed via:

$$A_{n+1,i} = \int_{\mathbb{R}^d} x_i \mathbb{P}_n(x) \mathbb{P}_n^T(x) d\mu, \quad B_{n+1,i} = \int_{\mathbb{R}^d} x_i \mathbb{P}_n(x) \mathbb{P}_{n+1}^T(x) d\mu. \quad (7)$$

Note that $A_{n+1,i}$ is determined by (quadratic moments of) \mathbb{P}_n , but $B_{n+1,i}$ does not have a similar characterization using only degree- n orthogonal polynomials.

3 Evaluation of polynomials

The three-term relation (4) does not immediately yield an evaluation scheme. We discuss in this section one approach for such an evaluation, which prescribes a fixed orthonormal basis. (That is, we remove “most” of the unitary equivalence freedom for \mathbb{P}_n .) Our solution for this is introduction of a particular “canonical” form.

3.1 Canonical bases

The three-term recurrence (4) for a fixed $i \in [d]$ is an underdetermined set of equations for \mathbb{P}_{n+1} , and hence this cannot be used in isolation to evaluate polynomials. To make the system determined, one could consider (4) for all $i \in [d]$ simultaneously. To aid in this type of procedure, we make a special choice of orthonormal basis that we will see amounts to choosing a particular sequence of unitary transformations.

Definition 3.1. Let $\{\mathbb{P}_n\}_{n \in \mathbb{N}_0}$ be an orthonormal set of polynomials with recurrence matrices $A_{n,i}$ and $B_{n,i}$ for $i \in [d]$, $n \in \mathbb{N}_0$. We say that $\{\mathbb{P}_n\}_{n \in \mathbb{N}_0}$ is a canonical (orthonormal) basis, and that the matrices $A_{n,i}$ and $B_{n,i}$ are in canonical form if the following is true: For every $n \in \mathbb{N}$ we have the condition

$$B_n^T B_n = \sum_{i \in [d]} B_{n,i}^T B_{n,i} = \Lambda_n, \quad (8)$$

where B_n is as defined in (5), and the matrices $\{\Lambda_n\}_{n \in \mathbb{N}}$ are a sequence of diagonal matrices with the elements of each matrix appearing in non-decreasing order.

Although the condition for canonical form appears only explicitly through a condition on $B_{n,i}$, the matrices $A_{n,i}$ are coupled with $B_{n,i}$ through the commuting conditions in (2.1) so that canonical form is implicitly a condition on the $A_{n,i}$ matrices as well. The utility of a canonical basis is revealed by the following result.

Theorem 3.1 (See also [6, Theorem 3.3.5]). Let the orthonormal basis $\{\mathbb{P}_n\}_{n \in \mathbb{N}_0}$ be a canonical basis so that the associated matrices $A_{n,i}$ and $B_{n,i}$ satisfy (8). Then,

$$\Lambda_{n+1} \mathbb{P}_{n+1} = \left(\sum_{i \in [d]} x_i B_{n+1,i}^T \right) \mathbb{P}_n - \left(\sum_{i \in [d]} B_{n+1,i}^T A_{n+1,i} \right) \mathbb{P}_n - \left(\sum_{i \in [d]} B_{n+1,i}^T B_{n,i}^T \right) \mathbb{P}_{n-1}, \quad (9)$$

for each $n \geq 0$, where $\Lambda_{n+1} \in \mathbb{R}^{r_{n+1} \times r_{n+1}}$ is diagonal and positive-definite (and hence invertible).

Proof. The relation (9) is computed by (vertically) stacking the d relations (4) for $i \in [d]$. Although this stacked system is overdetermined, it is consistent since (4) must hold for all $i \in [d]$. Multiplying both sides of the stacked system by B_{n+1}^T yields (9), and we need only explain why $\Lambda_{n+1} = B_{n+1}^T B_{n+1}$ is diagonal and positive-definite. The diagonal property is immediate since the basis is canonical, and clearly is positive semi-definite. That it is in fact positive-definite is a consequence of the rank condition (5), ensuring that $r_{n+1} = \text{rank}(B_{n+1}) = \text{rank}(\Lambda_{n+1})$. \square

Equation (9) demonstrates how knowledge of $(A_{n+1,i}, B_{n+1,i})_{i \in [d]}$ translates into direct evaluation of \mathbb{P}_{n+1} : the right-hand side of (9) is computable, and need only be scaled elementwise by the inverse diagonal of Λ_{n+1} . The main requirement for this simple technique for evaluation is that the recurrence matrices are in canonical form. Fortunately, it is fairly simple to transform any valid recurrence matrices into canonical form.

3.2 Transformation to canonical form

There is substantial freedom in how the basis elements of the vectors \mathbf{p}_n are chosen. In particular, let $\{U_n\}_{n \geq 0}$ be an arbitrary family of orthogonal matrices defining a new basis \mathbf{q}_n ,

$$\mathbf{q}_n(x) := U_n \mathbf{p}_n(x), \quad U_n \in \mathbb{R}^{r_n \times r_n}, \quad U_n^T U_n = I. \quad (10)$$

A manipulation of (4) shows that the basis elements \mathbf{q}_n satisfy a three-term relation,

$$x_i \mathbf{q}_n(x) = D_{n+1,i} \mathbf{q}_{n+1}(x) + C_{n+1,i} \mathbf{q}_n(x) + D_{n,i}^T \mathbf{q}_{n-1}(x), \quad i \in [d],$$

where the new matrices $C_{n,i}$ and $D_{n,i}$ can be explicitly derived from the unitary matrices U_n and the recurrence matrices for \mathbf{p}_n ,

$$C_{n,i} = U_{n-1} A_{n,i} U_{n-1}^T, \quad D_{n,i} = U_{n-1} B_{n,i} U_n^T, \quad (i, n) \in [d] \times \mathbb{N}_0. \quad (11)$$

Our goal now is to take arbitrary valid recurrence matrices $(A_{n,i}, B_{n,i})$ and identify the unitary transform matrices $\{U_n\}_{n \in \mathbb{N}}$ so that $(C_{n,i}, D_{n,i})$ are in canonical form. Since $B_n^T B_n$ is symmetric, then it has an eigenvalue decomposition,

$$B_n^T B_n = V_n \Lambda_n V_n^T, \quad (12)$$

with diagonal eigenvalue matrix Λ_n and unitary matrix V_n , where we assume the diagonal elements of Λ_n are in non-increasing order. Then by defining,

$$U_n = V_n^T, \quad n \geq 1,$$

which identifies $C_{n,i}$ and $D_{n,i}$ through (11), we immediately have that

$$\sum_{i \in [d]} D_{n,i}^T D_{n,i} = \sum_{i \in [d]} (V_n^T B_{n,i}^T V_{n-1})(V_{n-1}^T B_{n,i} V_n) = V_n^T B_n^T B_n V_n = \Lambda_n,$$

and hence \mathbf{q}_n is a canonical basis, with $C_{n,i}$ and $D_{n,i}$ the associated recurrence matrices in canonical form. Thus, for each fixed n , a transformation to canonical form is accomplished via a single size- r_n symmetric eigenvalue decomposition.

The discussion above also reveals how much non-uniqueness there is in the choice of canonical form through non-uniqueness in the symmetric eigenvalue decomposition: each row of U_n is non-unique up to a sign, and arbitrary unitary transforms of sets of rows corresponding to invariant subspaces of $B_n^T B_n$ are allowed. If the non-increasing diagonal elements of Λ_n have distinct values, then each of the functions in the vector \mathbf{q}_n is unique up to a multiplicative sign.

3.3 Three-term relation conditions on matrices

Our computational strategy computes $A_{n,i}$, $B_{n,i}$ through manipulations of polynomials computed via (9), which are assumed to be orthonormal polynomials. One subtlety is that if we prescribe recurrence matrices (through a computational procedure), then usage of (9) requires additional conditions to be equivalent to (4). This equivalence is summarized as follows, and forms the basis for the conditions we aim to impose in our algorithm.

Theorem 3.2 ([6, Theorem 3.5.1]). *Given matrices $A_{n,i}, B_{n,i}$, let \mathbf{p}_n , $n \geq 0$ be generated through (9). Then $\{\mathbf{p}_n\}$ is orthonormal with respect to some positive-definite bilinear functional on \mathbb{R}^d if and only if:*

- $A_{n,i}$ is symmetric for all $(i, n) \in [d] \times \mathbb{N}_0$.
- $B_{n,i}$ satisfies the rank condition in (5): $\text{rank}(B_{n,i}) = r_{n-1}$.
- The matrices satisfy the commuting conditions (6).

Additional conditions are required to ensure that the stated positive-definite functional is integration with respect to a measure μ , but since our algorithm considers only finite n , such a distinction is not needed for us.

4 Recurrence matrices and canonical form for tensorial measures

We take a slight detour in this section to identify recurrence matrices and associated canonical forms for tensorial measures μ . If μ is a tensorial measure, i.e., $\mu = \otimes_{j=1}^d \mu_j$, where each μ_j is a positive measure on \mathbb{R} satisfying the non-degeneracy condition (1), then there exists a sequence of univariate orthonormal polynomials for each $j \in [d]$. In particular, fixing j , there exist coefficients $\{a_{j,n}\}_{n \in \mathbb{N}} \subset \mathbb{R}$ and $\{b_{j,n}\}_{n \in \mathbb{N}_0} \subset (0, \infty)$ such that the sequence of polynomials defined by the recurrence,

$$x_j p_{j,n}(x_j) = b_{j,n+1} p_{j,n+1}(x_j) + a_{j,n} p_{j,n}(x_j) + b_{j,n} p_{j,n-1}(x_j), \quad n \geq 0,$$

are $L^2_{\mu_j}(\mathbb{R})$ -orthonormal. With these univariate polynomials, we can directly construct multivariate orthonormal polynomials via tensorization,

$$p_\alpha(x) := \prod_{j=1}^d p_{j,\alpha_j}(x_j), \quad \alpha \in \mathbb{N}_0^d, \quad (13)$$

where $\deg p_\alpha = |\alpha|$. We now construct vectors containing polynomials of a specific degree as in (3). With J_n denote any ordered set of the multi-indices α satisfying $|\alpha| = n \geq 0$, define,

$$\mathbb{P}_n = (p_{\alpha^{(n,1)}}, \dots, p_{\alpha^{(n,r_n)}})^T, \quad J_n := (\alpha^{(n,1)}, \dots, \alpha^{(n,r_n)}), \quad (14)$$

where $|\alpha^{(n,k)}| = n$ for each $k \in [r_n]$. Clearly this set $\{\mathbb{P}_n\}_{n \in \mathbb{N}_0}$ is a sequence of multivariate orthogonal polynomials. The next subsection explicitly computes the associated recurrence matrices.

4.1 Recurrence matrices

Given the ordering of multi-indices defined by the sets $(J_n)_{n \in \mathbb{N}_0}$, we require a function that identifies the location of the index $\alpha^{(n,k)} + e_j$ in the set J_{n+1} . Fixing the sets $\{J_n\}_{n \geq 0}$, we define a function $c : \mathbb{N}_0^d \times [d] \rightarrow \mathbb{N}$ as,

$$c(\alpha^{(n,k)}, i) = \text{The index } q \in [r_{n+1}] \text{ such that } \alpha^{(n,k)} + e_i = \alpha^{(n+1,q)}.$$

Note that $c(\alpha, i)$ is well-defined for all $\alpha \in \mathbb{N}_0^d$ and $i \in [d]$. We can now identify recurrence matrices for the polynomials \mathbb{P}_n explicitly in terms of univariate recurrence coefficients.

Theorem 4.1. *With the set of polynomials $(\mathbb{P}_n)_{n=0}^\infty$ defined by (14) and (13), then the recurrence matrices are given by,*

$$A_{n+1,i} = \text{diag} \left(a_{i,\alpha_i^{(n,k)}+1} \right)_{k=1}^{r_n}, \quad B_{n+1,i} = \sum_{k=1}^{r_n} b_{i,\alpha_i^{(n,k)}+1} e_{r_n,k} e_{r_{n+1},c(\alpha^{(n,k)},i)}^T, \quad (15)$$

where $\alpha_i^{(n,k)}$ is the i th component of $\alpha^{(n,k)}$ and $e_{m,j}$ is an m -dimensional vector with entry 1 in location j and zeros elsewhere.

Proof. Fixing any $n \in \mathbb{N}_0$ and $i \in [d]$, and for each $k \in [r_n]$, the k th component of the vector \mathbb{P}_n satisfies,

$$\begin{aligned} x_i p_{\alpha^{(n,k)}}(x) &= \left(x_i p_{\alpha_i^{(n,k)}}(x_i) \right) \prod_{j \neq i} p_{\alpha_j^{(n,k)}}(x_j) \\ &= b_{i,n+1} p_{\alpha^{(n,k)}+e_{d,i}}(x) + a_{i,n+1} p_{\alpha^{(n,k)}}(x) + b_{i,n} p_{\alpha^{(n,k)}-e_{d,i}}(x). \end{aligned}$$

By comparing the final expression above with the three-term recurrence (4), the components of the matrices $A_{n+1,i}$ and $B_{n+1,i}$ can be identified and are given as in (15). \square

When μ is tensorial, then computational methods for evaluating p_n need not explicitly use recurrence matrices: the relations (14) and (13) show that one needs only knowledge of *univariate* recurrence coefficients $a_{j,n}, b_{j,n}$. Computation of these coefficients from the marginal measure μ_j is well-studied and has several solutions (cf. (2.1)). However, if one wanted the recurrence matrices for this situation, they are explicitly available through the result above. We next show that, up to a permutation, such tensorized basis elements form a canonical basis.

4.2 Canonical form for tensorial measures

The explicit formulas in (15) allow us to investigate the canonical condition (8):

$$B_n^T B_n = \sum_{i \in [d]} B_{n,i}^T B_{n,i} = \text{diag} \left(\sum_{i=1}^d b_{i,\alpha_i^{(n-1,1)}+1}^2, \dots, \sum_{i=1}^d b_{i,\alpha_i^{(n-1,r_n-1)}+1}^2 \right)$$

In order for p_n to be a canonical basis, the diagonal elements above must be non-increasing. While it is not transparent on how to accomplish this for generic univariate coefficients $(a_{i,n}, b_{i,n})_{i \times n \in [d] \times \mathbb{N}}$, one can always computationally achieve this by reordering the elements of p_n , equivalently by reordering the elements of J_n , according to the elements on the right-hand side above.

Corollary 4.1.1. *Define permutation operators $P_n : [r_{n-1}] \rightarrow [r_n]$ given by,*

$$P(k+1) \geq P(k) \implies \sum_{i=1}^d b_{i,\alpha_i^{(n-1,k+1)}+1}^2 \geq \sum_{i=1}^d b_{i,\alpha_i^{(n-1,k)}+1}^2,$$

for each $k \in [r_n]$. Defining a new basis, $q_n := (p_{\alpha^{(n,P(k))}})_{k=1}^{r_n}$, then $\{q_n\}_{n=0}^\infty$ is a canonical basis.

We re-emphasize that identifying a canonical basis through an algorithmic version of the procedure above is not computationally advantageous compared to direct usage of (14) and (13). However, the procedure above gives us a method for oracle evaluation of recurrence matrices associated to a tensorial measure μ .

5 Algorithms for computing recurrence coefficient matrices

We discuss two strategies for computing the recurrence matrices $(A_{n,i}, B_{n,i})$. The first is a straightforward Moment Method (MM-M) that directly uses moments (e.g., of monomials) to compute an orthonormal basis that can be used directly with (7) to compute the matrices. However, this approach is very ill-conditioned for moderately large degree and hence has limited utility. The second approach we present is the main novel advancement of this paper, a multivariate Stieltjes (MS) algorithm that computes the recurrence matrices through an alternative procedure.

5.1 The moment method (MM-M)

A straightforward way to compute recurrence coefficients in the univariate case is to first perform an orthogonalization step to numerically generate orthogonal polynomials as linear expansions in some specified basis (say monomials), and second to exploit the linear expansion expressions to compute the recurrence coefficients. This works in the multivariate setting as well. We suppose that a(ny) polynomial basis $\{\phi_j\}_{j \in \mathbb{N}}$ is given with the properties,

$$\begin{aligned} \text{span } \{\phi_j\}_{j=1}^{R_n} &= \Pi_n^d, & n \in \mathbb{N}_0. \\ r_n < j \leq r_{n+1} &\implies \deg \phi_j = n+1. \end{aligned}$$

Again, a simple example is that $\phi_j(x)$ is a multivariate monomial x^α for some total ordering of the multi-indices α that respects the partial order induced by the $\ell^1(\mathbb{N}_0^d)$ norm¹. We assume that quadratic ϕ_j moments are available, allowing us to compute a Gram matrix,

$$(G_n)_{i,j} = m(i,j), \quad G_n \in \mathbb{R}^{R_n \times R_n}, \quad m(i,j) = \int \phi_i(x)\phi_j(x)d\mu(x). \quad (16)$$

Given this matrix, we can compute monomial expansion coefficients for the orthonormal basis \mathbb{p}_n :

$$G_n = L_n L_n^T, \quad (\mathbb{p}_0, \dots, \mathbb{p}_n)^T = L_n^{-1} \Phi_n, \quad \Phi_n = (\phi_1, \dots, \phi_{R_n})^T \in \mathbb{R}^{R_n}, \quad (17)$$

where \mathbb{p}_j is an r_j -vector containing degree- j orthonormal polynomials. Combining this with (7), the recurrence matrices can be directly computed,

$$A_{n+1,i} = \tilde{L}_n^{-1} G_{n,i} \tilde{L}_n^{-T}, \quad B_{n+1,i} = \tilde{L}_n^{-1} \tilde{G}_{n+1,i} \tilde{L}_{n+1}^{-T}, \quad i \in [d], \quad (18)$$

where \tilde{L}_n^{-1} is the $r_n \times R_n$ matrix formed from the last r_n rows of L_n^{-1} , $G_{n,i} \in \mathbb{R}^{R_n \times r_n}$ has entries

$$(G_{n,i})_{i,j} = \int x_i \phi_i(x) \phi_j(x) d\mu(x), \quad i, j \in [R_n], \quad (19)$$

and $\tilde{G}_{n+1,i} \in \mathbb{R}^{R_n \times R_{n+1}}$ equals the first R_n rows of $G_{n+1,i}$. Thus, so long as the polynomial moments $m(i,j)$ and (19) can be evaluated, then this allows direct computation of the recurrence matrices. Of course, if only an orthonormal basis (without recurrence matrices) is desired, then one may stop at (17).

The main drawbacks to the procedures above stem from accuracy concerns due to ill-conditioning of the G_n matrices. If the ϕ_i are selected as “close” to an orthonormal basis, then the G_n matrices can be well-conditioned, but *a priori* knowledge of such a basis is not available in general scenarios. However, this method is flexible in the sense that with minor modifications one may compute an orthonormal basis (but not necessarily recurrence matrices) for very general, non-total-degree, polynomial spaces, such as hyperbolic cross spaces.

5.2 The Multivariate Stieltjes algorithm (MS)

In this section we describe a Stieltjes-like procedure for computing recurrence matrices, which partially overcomes ill-conditioning issues of the moment method. Like the univariate procedure, we directly compute the recurrence matrices instead of attempting to orthogonalize the basis, and the procedure is iterative on the degree n . Thus, throughout this section we assume that the recurrence matrices $\{A_{m,i}, B_{m,i}\}_{i \in [d], m \leq n}$ are available, and our goal is to compute $A_{n+1,i}$ and $B_{n+1,i}$ for $i \in [d]$. Our main strategy for accomplishing this will be to satisfy certain “degree- $(n+1)$ ” matrix moment conditions subject to the constraints identified in 3.2.

The availability of the matrices for index $m \leq n$ implies that \mathbb{p}_m for all $m \leq n$ can be evaluated through the procedure in 3.1. (In practice we transform $A_{m,i}, B_{m,i}$ for $m \leq n$ to be in canonical form.) We will compute $A_{n+1,i}$ directly, but compute factorized components and subblocks of $B_{n+1,i}$, and we identify those subblocks now. Consider the truncated singular value decomposition of $B_{n+1,i}$,

$$B_{n+1,i} = U_{n+1,i} \Sigma_{n+1,i} V_{n+1,i}^T = U_{n+1,i} \Sigma_{n+1,i} \begin{pmatrix} \hat{V}_{n+1,i}^T & \tilde{V}_{n+1,i}^T \end{pmatrix}, \quad i \in [d], \quad (20)$$

where $U_{n+1,i}, V_{n+1,i}$ have orthonormal columns and are of sizes,

$$U_{n+1,i} \in \mathbb{R}^{r_n \times r_n}, \quad \Sigma_{n+1,i} \in \mathbb{R}^{r_n \times r_n}, \quad V_{n+1,i} \in \mathbb{R}^{r_{n+1} \times r_n},$$

¹Degree-graded lexicographic ordering over multi-indices is an example of one such ordering.

and $\Sigma_{n+1,i}$ is diagonal with non-negative entries in non-increasing order. $\Sigma_{n+1,i}$ must be invertible due to (5). In (20) we have further decomposed the right-singular $V_{n+1,i}$ matrices into blocks, with \widehat{V} the first r_n rows, and \widetilde{V} the remaining rows:

$$V_{n+1,i} = \begin{pmatrix} \widehat{V}_{n+1,i} \\ \widetilde{V}_{n+1,i} \end{pmatrix}, \quad \widehat{V}_{n+1,i} \in \mathbb{R}^{r_n \times r_n} \quad \widetilde{V}_{n+1,i} \in \mathbb{R}^{\Delta r_{n+1} \times r_n} \quad (21)$$

We do not assume that the $B_{n+1,i}$ are in canonical form, and therefore have freedom to specify the unitary transform for degree $n+1$. Without loss we implicitly choose the unitary transform U_{n+1} in (11) so that $V_{n+1,1}$ is equal to the first r_n columns of the size- r_{n+1} identity matrix. This uniquely identifies $V_{n+1,1}$,

$$\widehat{V}_{n+1,1} = I_{r_n}, \quad \widetilde{V}_{n+1,1} = \mathbf{0}. \quad (22)$$

We therefore need only compute $\widehat{V}_{n+1,i}, \widetilde{V}_{n+1,i}$ for $i \geq 2$.

The remainder of this section is structured as follows: We define certain moment matrices through a modified polynomials basis in Section 5.2.1, which immediately yields the $A_{n+1,i}$ matrices. Section 5.2.2 shows how we compute the U and Σ SVD matrices of B . We introduce “mixed” moments in Section 5.2.3 that allow us to obtain the \widehat{V} block of the V matrices. The remaining block \widetilde{V} is computed using different strategies depending on the dimension d of the problem. For $d = 2$, Section 5.2.4 shows that \widetilde{V} can be computed almost directly. For $d \geq 3$ dimensions, we must enforce the commuting conditions (6), which is somewhat easily done in $d = 3$ dimensions, but requires nontrivial optimization for $d > 3$.

5.2.1 Moment matrices – computing $A_{n+1,i}$

The MS algorithm begins by considering moments of polynomials that are not explicitly known *a priori*, but are easily generated during an algorithmic procedure. In particular we introduce the moment matrices,

$$S_{n,i} := \int x_i \mathbb{P}_n \mathbb{P}_n^T d\mu(x), \quad T_{n,i,j} := \int \widetilde{\mathbb{P}}_{n+1,i} \widetilde{\mathbb{P}}_{n+1,j}^T d\mu(x), \quad (23)$$

both $r_n \times r_n$ matrices, where the modified polynomial basis $\widetilde{\mathbb{P}}$ is defined as,

$$\widetilde{\mathbb{P}}_{n+1,i}(x) := x_i \mathbb{P}_n - A_{n+1,i} \mathbb{P}_n - B_{n,i}^T \mathbb{P}_{n-1}. \quad (24)$$

Note that availability of $\{A_{m,i}, B_{m,i}\}_{i \in [d], m \leq n}$ along with the ability to evaluate \mathbb{P}_n imply that the moment matrices in (23) can be approximated via quadrature, just as is frequently done for the ϕ_j moments of Section 5.1.

Inspection of (7) immediately reveals that,

$$A_{n+1,i} = S_{n,i}, \quad (25)$$

and hence $A_{n+1,i}$ is directly computable. In addition, $A_{n+1,i}$ is symmetric (since $S_{n,i}$ is symmetric) and hence this satisfies the first condition in 3.2. This then allows $\widetilde{\mathbb{P}}_{n+1,i}$ to be evaluated, and hence allows $T_{n,i,j}$ to be computed. While evaluating $A_{n+1,i}$ is fairly straightforward from $S_{n,i}$, computing $B_{n+1,i}$ from $T_{n,i,j}$ is more involved.

5.2.2 Stieltjes symmetric moments – computing $U_{n+1,i}, \Sigma_{n+1,i}$

The matrices $T_{n,i,i}$ of symmetric moments allow us to compute the U and Σ matrices in the SVD of B . A direct computation with the three-term recurrence (4) and the definition (23) reveals that

$$T_{n,i,i} = B_{n+1,i} B_{n+1,i}^T = U_{n+1,i} \Sigma_{n+1,i}^2 U_{n+1,i}^T. \quad (26)$$

Therefore for each $i \in [d]$, we can first compute the square, symmetric matrix $T_{n,i,i}$, and subsequently its eigenvalue decomposition, ordering the eigenvalues in decreasing order. Then the eigenvector matrix of $T_{n,i,i}$ is $U_{n+1,i}$ and the square root of the eigenvalue matrix equals $\Sigma_{n+1,i}$. We can also conclude that $T_{n,i,i}$ is full-rank: Since p_n contains linearly independent polynomials, then $x_i p_n$ contains linearly independent polynomials, and hence $\tilde{P}_{n+1,i}$ contains linearly independent polynomials. Since $T_{n,i,i}$ is the Gram matrix for $\tilde{P}_{n+1,i}$ and we have assumed μ is non-degenerate, then it must be of full rank. This observation then implies that $\Sigma_{n+1,i}$ is invertible, and hence that our computed $B_{n+1,i}$ is also full rank. Therefore, we have satisfied the second condition in 3.2. Now we are left only to compute the rectangular $V_{n+1,i}$ matrices, which will involve enforcing the third condition (the commuting conditions).

5.2.3 Stieltjes mixed moments – computing $\hat{V}_{n+1,i}$

Using mixed moments $T_{n,i,j}$ with $i \neq j$, we can compute the square matrices $\hat{V}_{n+1,i}$, which are subblocks of $V_{n+1,i}$. (Recall from (22) that $V_{n+1,j}$ is already known for $j = 1$, so we consider only $j > 2$.) A similar computation as the Stieltjes procedure in (26) yields that $T_{n,i,j} = B_{n+1,i} B_{n+1,j}^T$. By using the decomposition of $B_{n+1,i}$ in (20), we conclude that

$$\hat{V}_{n+1,i}^T \hat{V}_{n+1,i} + \tilde{V}_{n+1,i}^T \tilde{V}_{n+1,i} = \Sigma_{n+1,i}^{-1} U_{n+1,i}^T T_{n,i,j} U_{n+1,j} \Sigma_{n+1,j}^{-1}, \quad i, j \in [d]. \quad (27)$$

Letting $i = 1$ and utilizing (22), we have,

$$\hat{V}_{n+1,j} = \Sigma_{n+1,1}^{-1} U_{n+1,1}^T T_{n,1,j} U_{n+1,j} \Sigma_{n+1,j}^{-1}, \quad j \geq 2, \quad (28)$$

where everything on the right-hand side is known and computable.

Note that here we have only utilized $T_{n,i,j}$ for $1 = i \neq j$. The case $1 \neq i \neq j$ is vacuous for $d = 2$, and we will see that the remaining block \tilde{V} can already be computed. When $d \geq 3$, we do require $1 \neq i \neq j$ to identify \tilde{V} .

5.2.4 $d = 2$: Orthonormality conditions for $\tilde{V}_{n+1,i}$

For $d = 2$, we now need only compute $\tilde{V}_{n+1,2}$. Since $\Delta r_n = 1$ for every n when $d = 2$, then $\tilde{V}_{n+1,2}$ is a $1 \times r_n$ vector. To reduce notational clutter, we consider fixed n and use the following notation,

$$\mathbf{y} := \tilde{V}_{n+1,2}^T \in \mathbb{R}^{r_n}.$$

Then since $V_{n+1,2}$ has orthonormal columns, i.e., $V_{n+1,2}^T V_{n+1,2} = I_{r_n}$, we have,

$$\mathbf{y} \mathbf{y}^T = I_{r_n} - \hat{V}_{n+1,2}^T \hat{V}_{n+1,2}, \quad (29)$$

which defines \mathbf{y} up to a sign. More precisely, we have

$$\mathbf{y} = \pm \mathbf{z}, \quad (30)$$

where \mathbf{z} is computed either as a rank-1 Cholesky factor or as the positive semi-definite square root of the right-hand side of (29). Although it appears the multiplicative sign needs to be chosen, if we choose $(i, j) = (1, 2)$ in (6c), then we have

$$\begin{pmatrix} B_{n,1} U_{n+1,2} \Sigma_{n+1,2} \hat{V}_{n+1,2} & \pm B_{n,1} U_{n+1,2} \Sigma_{n+1,2} \mathbf{z} \end{pmatrix} = B_{n,2} B_{n+1,1}.$$

Thus the choice of sign makes no difference because the last column of $B_{n+1,1}$ is a zero vector. Therefore, we arbitrarily choose the sign. This completes the computation of $B_{n+1,i}$ for $d = 2$; in this case we need not impose the commuting conditions, but they are needed for $d > 2$.

5.2.5 $d > 2, n = 0$: Falling back on moments

For $d > 2$ and $n = 0$, then $B_{n+1,i} \in \mathbb{R}^{1 \times d}$, and hence $\tilde{V}_{n+1,i} \in \mathbb{R}^{1 \times (d-1)}$. We have two conditions to impose on this vector of length $d - 1 \geq 2$:

- The scalar-valued commuting condition (6a) (the others do not apply for $n = 0$).
- A unit-norm condition on the column vector $V_{n+1,i}$ (as is used in the previous section)

This amounts to 2 conditions on this vector (although the second condition does not determine a multiplicative sign). However, as this approach combining all these conditions can be relatively cumbersome to simply determine a vector, in this case we fall back to using MM routines. When $n = 0$, the MM Gramians are typically well-conditioned. Therefore, when $d > 2$ and $n = 0$, we use (18) to compute the $B_{1,i}$ matrices.

5.2.6 $d > 2, n > 0$: The commuting conditions – computing $\tilde{V}_{n+1,i}$

We recall that our remaining task is to compute $\tilde{V}_{n+1,i} \in \mathbb{R}^{\Delta_{n+1} \times r_n}$ for $2 \leq i \leq d$. Our tools to accomplish this will be (i) the commuting conditions (ii) orthonormality conditions on the columns of $V_{n+1,i}$, and (iii) the mixed moments $T_{n,i,j}$ for $1 \neq i \neq j$. For $n > 0$, the commuting condition (6c) with $i = 1$ implies:

$$B_{n,1}U_{n+1,j}\Sigma_{n+1,j}\left(\hat{V}_{n+1,j}^T \quad \tilde{V}_{n+1,j}^T\right) = B_{n,j}B_{n+1,1}, \quad j \geq 2, \quad (31)$$

where only $\tilde{V}_{n+1,j}$ is unknown. Note that we have made the choice (22) for $\tilde{V}_{n+1,1}$, which implies that the last Δr_{n+1} columns of $B_{n,j}B_{n+1,1}$ vanish, i.e., the last Δr_{n+1} columns of (31) read,

$$K_{n+1,j}\tilde{V}_{n+1,j} = \mathbf{0}, \quad K_{n+1,j} := B_{n,1}U_{n+1,j}\Sigma_{n+1,j} \in \mathbb{R}^{r_{n-1} \times r_n}. \quad (32)$$

Thus, the columns of $\tilde{V}_{n+1,j}$ lie in the kernel of the known matrix $K_{n+1,j}$. I.e., we have,

$$\tilde{V}_{n+1,j}^T = \Psi_j C_j, \quad \Psi_j \in \mathbb{R}^{r_n \times \Delta r_n}, \quad C_j \in \mathbb{R}^{\Delta r_n \times \Delta r_{n+1}}, \quad (33)$$

where C_j is unknown and Ψ_j is known (computable), containing an orthonormal basis for $\ker(K_{n+1,j})$,

$$\text{range}(\Psi_j) = \ker(K_{n+1,j}), \quad \Psi_j^T \Psi_j = I_{\Delta r_n \times \Delta r_n}.$$

We now use orthonormality of the columns of $V_{n+1,j}$. In particular this implies,

$$C_j C_j^T = D_j := I_{\Delta r_n} - \Psi_j^T \hat{V}_{n+1,j}^T \hat{V}_{n+1,j} \Psi_j.$$

Since D_j is a symmetric, positive semi-definite matrix, then C_j must be given by,

$$C_j = E_j W_j, \quad E_j := \begin{pmatrix} \sqrt{D_j} & \mathbf{0}_{\Delta r_n \times (\Delta r_{n+1} - \Delta r_n)} \end{pmatrix}, \quad W_j \in \mathbb{R}^{\Delta r_{n+1} \times \Delta r_{n+1}}, \quad (34)$$

where W_j is a unitary matrix, and $\sqrt{D_j}$ is the symmetric positive semi-definite matrix square root of D_j . We therefore need only determine W_j . The final linear conditions we impose are the remaining mixed moment conditions from Section 5.2.3 involving $T_{n,i,j}$ for $1 \neq i \neq j \geq 2$. Using (27) and writing in terms of the unknown W_i, W_j yields,

$$\begin{aligned} E_i W_i W_j^T E_j^T &= H_{n,i,j} & 2 \leq i, j \leq d, \quad i \neq j \\ \text{subject to } W_j^T W_j &= I_{\Delta r_{n+1}} & 2 \leq j \leq d. \end{aligned} \quad (35)$$

where,

$$H_{n,i,j} = \Psi_i^T (\Sigma_{n+1,i}^{-1} U_{n+1,i}^T T_{n,i,j} U_{n+1,j} \Sigma_{n+1,j}^{-1} - \hat{V}_{n+1,i}^T \hat{V}_{n+1,j}) \Psi_j,$$

is a computable matrix. This optimization problem for $\{W_j\}_{j=2}^d$ must be solved to determine the W_j matrices. Once these are determined, then $\tilde{V}_{n+1,i}$ is determined through (33) and (34). This is a non-convex optimization, and is a generalization of a weighted orthogonal Procrustes problem (WOPP). Even the original WOPP has no known direct solution, so that numerical optimization must be employed to solve the above problem [11]. Fortunately, when $d = 3$ some extra manipulations do yield a direct solution.

5.2.7 $d = 3$: Circumventing the WOPP

When $d = 3$, the problem (35) simplifies substantially since we need only compute W_2, W_3 . First we note that if any pair of orthogonal matrices (W_2, W_3) satisfies (35), then so does the pair $(I_{\Delta r_{n+1}}, W_3 W_2^T)$. Therefore, we may choose $W_2 = I_{\Delta r_{n+1}}$ without loss. This then determines $\tilde{V}_{n+1,2}$ through (33) and (34).

The determination of W_3 in (35) now reduces to an instance of a WOPP:

$$E_2 W_3^T E_3^T = H_{n,2,3}, \quad \text{subject to} \quad W_3^T W_3 = I_{\Delta r_{n+1}} \quad (36)$$

We now notice that E_j defined in (34) is rectangular, but when $d = 3$ has only one more column than row, i.e., $\Delta r_n = n = 1$ and thus $\Delta r_{n+1} = 1 + \Delta r_n$. Then with

$$E_j = X_j (Y_j \ \mathbf{0}_{(n+1) \times 1}) Z_j^T,$$

the reduced singular value decompositions of E_2 and E_3 , then (36) can be rewritten as

$$I_{(n+1) \times (n+2)} W I_{(n+2) \times (n+1)} = Y_2^{-1} X_2^T H_{n,2,3} X_3 Y_3^{-1}, \quad W = Z_2^T W_3 Z_3,$$

where W is an orthogonal matrix of size $\Delta r_{n+1} = n + 2$. Determining W uniquely identifies W_3 , but the above relation shows that the $(n + 1) \times (n + 1)$ principal submatrix of W is given by $Y_2^{-1} X_2^T H_{n,2,3} X_3 Y_3^{-1}$. To determine the last row and column of W , we write,

$$W = \begin{pmatrix} Y_2^{-1} X_2^T H_{n,2,3} X_3 Y_3^{-1} & | \\ \mathbf{w}^T & | \end{pmatrix}, \quad \mathbf{w}, \mathbf{v} \text{ unknown},$$

and since W is an orthogonal matrix and must have orthonormal columns, then the first $n + 1$ entries of \mathbf{w} are determined up to a sign, and the signs can be determined by enforcing pairwise orthogonality conditions among the first $n + 1$ columns. The final column \mathbf{v} can be determined (up to an inconsequential sign) as a unit vector orthogonal to the first $n + 1$ columns. Thus, W is computable, which in turn determines W_3 , and completes the solution to (35) for $d = 3$.

6 Numerical test

We present numerical results showcasing the effectiveness of the MS algorithm for computing orthogonal polynomials on two and three-dimensional domains. We measure the efficacy of any particular method through a numerical orthogonality condition. Each method first generates computational approximations to the matrices $\{A_{n,i}, B_{n,i}\}$, and subsequently can generate computational approximations $\hat{\mathbb{P}}_n$ to an orthonormal basis \mathbb{P}_n . Fixing a maximum degree N , we then use quadrature (approximate or exact as described) to evaluate the error matrix,

$$E = \begin{pmatrix} M_{0,0} & \cdots & M_{0,N} \\ \vdots & \ddots & \vdots \\ M_{N,0} & \cdots & M_{N,N} \end{pmatrix} - I_{R_N}, \quad M_{m,n} = \int \mathbb{P}_m(x) \mathbb{P}_n^T(x) d\mu(x), \quad (37)$$

where E is an $R_N \times R_N$ matrix. We will both show plots of this error matrix and also report entrywise maximum values, $\|E\|_{\infty,\infty} = \max_{i,j \in [R_N]} |E_{i,j}|$. In all experiments, we set $N = 39$ for $d = 2$, and $N = 15$ for $d = 3$. (These choices result in $R_N = 820$ and $R_N = 816$, respectively.) We compare four methods:

- (Exact) When μ is a tensor-product measure, we use the procedure in Section 4 and summarized in Algorithm 1 to first explicitly compute the recurrence matrices, and subsequently to evaluate polynomials through the procedure in Section 3.
- (MS) The novel algorithm of this manuscript, the Multivariate Stieltjes algorithm, described in Section 5.2 and summarized in Algorithm 2.
- (MM-M) The moment method of Section 5.1, involving direct orthogonalization of the monomial basis, i.e., the functions ϕ_j are monomials x^α .
- (MM-L) The moment method of Section 5.1, involving direct orthogonalization of the tensorial Legendre polynomial basis, i.e., the functions ϕ_j are tensorial Legendre polynomials with respect to the uniform measure ν whose support is a bounding box of the support of μ .

All methods assume the ability to compute (general) polynomial moments in order to carry out computations, and we describe in subsections below how we approximately or exactly compute these moments via quadrature. Each experiment uses the same quadrature rule for all methods. A summary of the experiments (i.e., the various measures μ) is given in Table 1

Table 1: Abbreviation and subsection for the examples considered.

Example	Abbreviation	Dimension	Section
Jacobi weight function	JAC	2, 3	6.1.1
Uniform measure on an annulus	ANN	2	6.1.2
Uniform measure between polar curves	CUR	2	6.1.3
Uniform Measure within a torus	TOR	3	6.1.4
Uniform measure on a rectangle minus a ball	HOL	2	6.2.1
Uniform measure on Madagascar	MAP	2	6.2.2

6.1 Experiments with moments via tensorized Gaussian quadrature

In this subsection we compute polynomial moments with respect to the measure μ via tensorized Gaussian quadrature. In all cases except Section 6.1.3, this quadrature is exact (in infinite precision).

6.1.1 JAC: Tensorized Jacobi measure

For an initial investigation, we consider a tensorial measure,

$$d\mu(x) = \prod_{i=1}^d B(\alpha_i, \beta_i)(1-x_i)^{\alpha_i}(1+x_i)^{\beta_i}, \quad x \in [-1, 1]^d \quad \alpha_i, \beta_i > -1,$$

where $B(\cdot, \cdot)$ is the Beta function, which is a tensorized Jacobi (or Beta) measure. To compute moments, we utilize tensorized Gauss quadrature of sufficiently high order so that all integrals are exact (in exact arithmetic). We randomly generated the α_i, β_i parameters by uniformly sampling over the interval $(-1, 10)$, resulting in the choices:

$$\begin{aligned} (\alpha_1, \alpha_2) &= (3.80, 0.78), & (\beta_1, \beta_2) &= (7.34, 8.26), & d &= 2 \\ (\alpha_1, \alpha_2, \alpha_3) &= (1.61, 0.32, 3.01), & (\beta_1, \beta_2, \beta_3) &= (-0.89, 9.83, 7.67), & d &= 3 \end{aligned}$$

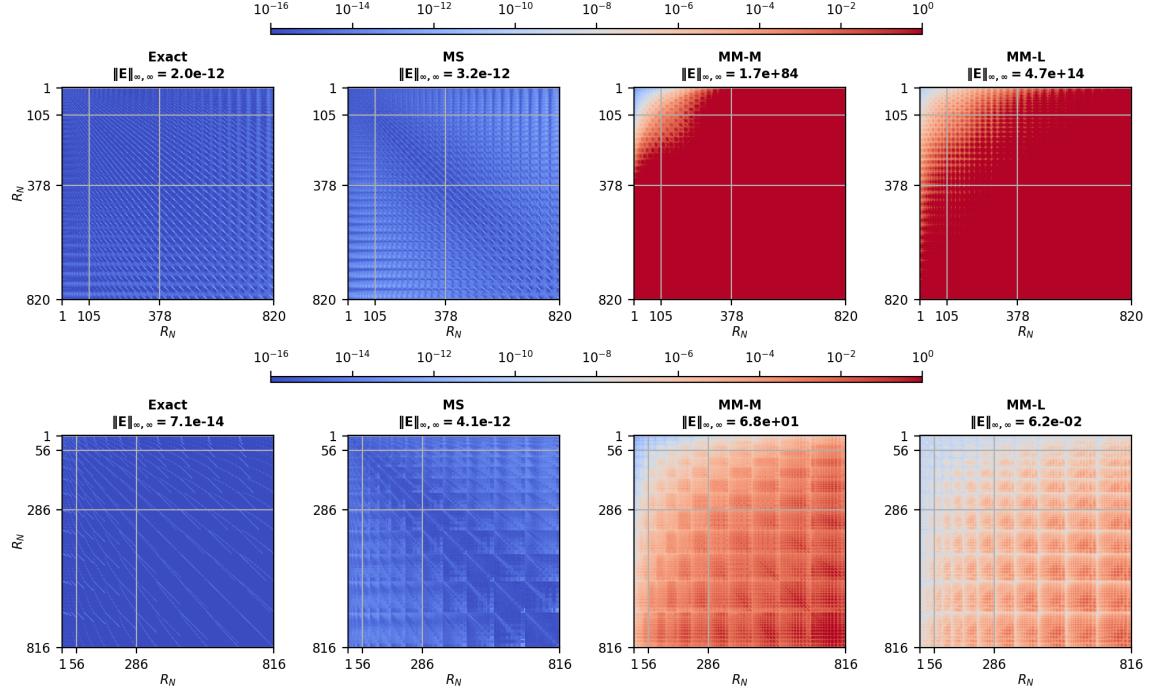


Figure 1: JAC results, Section 6.1.1: Visualization of the error matrix E in dimension $d = 2$ (top) and $d = 3$ (bottom). From left to right in each row: the Exact, MS, MM-M, and MM-L algorithms. We choose 10^0 as the upper saturation point for all colormaps as values beyond this indicate $\mathcal{O}(1)$ error; we continue to impose this saturation value for all subsequent error plots.

Figure 1 visualizes the Gramian error matrix E in (37) for $d = 2, 3$. One observes that the Exact and MS algorithms performs very well, but both the MM-M and MM-L algorithms suffer accumulation of roundoff error. Even in this case, when the MM-L algorithm uses a “reasonable” choice of basis for orthogonalization, instabilities develop quickly. In this simple case when quadrature is numerically exact, a standard orthogonalization routine produces unstable results.

6.1.2 ANN: Measure on an Annulus

Our second test is the uniform measure μ with support in an annular region in $d = 2$ dimensions centered at the origin. In polar coordinates (r, θ) , this is represented as,

$$\text{supp}(\mu) = \{(r, \theta) \mid \theta_1 \leq \theta \leq \theta_2, r_1(\theta) \leq r \leq r_2(\theta)\}, \quad (38)$$

where $(\theta_1, \theta_2) = (0, 2\pi)$ and $(r_1(\theta), r_2(\theta)) = (0.5, 1.0)$. Quadrature with respect to the uniform measure on this domain can exactly integrate polynomials (in the x variable) using a tensor-product quadrature over (r, θ) space using Fourier quadrature in θ and Legendre-Gauss quadrature in the r variable. We use a large enough quadrature rule so that all integrals are numerically exact.

In this case we do not know exact formulas for the recurrence matrices, so we rely on the metric E . Figure 2 shows again that the MS algorithm performs better than the MM-M and MM-L methods for large degrees.

6.1.3 CUR: Measure within polar curves

We now consider a more difficult example again in $d = 2$ dimensions. We again use a curve defined as (38), but this time it is a region bounded between two Archimedean spirals. In particular, we set,

$$(\theta_1, \theta_2) = (0, 6\pi), \quad (r_1(\theta), r_2(\theta)) = (0.8\theta, \theta).$$

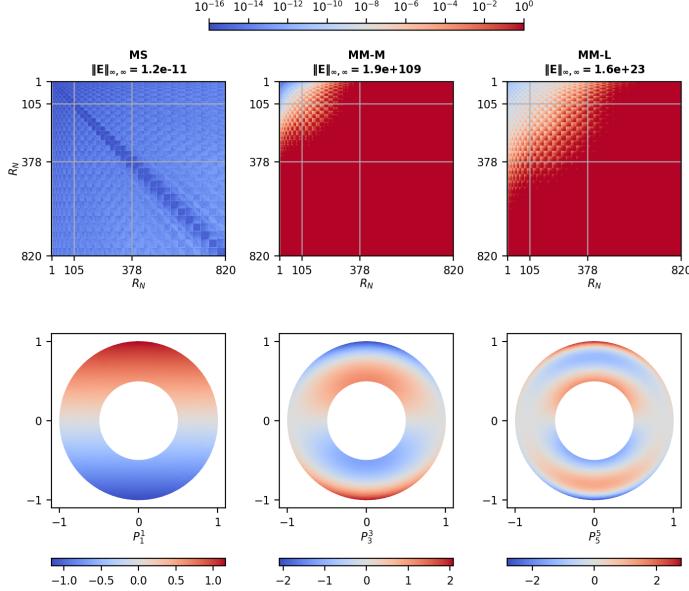


Figure 2: ANN results, Section 6.1.2. Top: Visualization of the error matrix E for the **MS**, **MM-M**, and **MM-L** algorithms from left to right. Bottom: Evaluations of the r_n 'th entry of p_n for degree $n = 1, 3, 5$ using the **MS** algorithm.

Again we choose μ as the uniform measure over the region defined in (38).

We write integrals as iterated, with the inner integral over r exactly computable using Legendre-Gauss quadrature. But the outer integral in θ involves terms both polynomial and trigonometric polynomial in θ , and we approximately integrate these values with a 10^6 -point Fourier quadrature rule. As can be seen in Figure 3, the novel **MS** procedure once again is much more stable than the **MM-M** and **MM-L** approaches.

6.1.4 TOR: Uniform measure inside a torus

We consider the uniform measure over the interior of a torus, whose parametric representation of the boundary is given by,

$$x_1(\theta, \phi) = (R + r \cos(\theta)) \cos \phi, \quad x_2(\theta, \phi) = (R + r \cos(\theta)) \sin \phi, \quad x_3(\theta, \phi) = r \sin(\theta),$$

for $\theta, \phi \in [0, 2\pi]$. The interior of the torus is defined by $(\sqrt{x_1^2 + x_2^2} - R^2)^2 + x_3^2 < r^2$. We choose $r = 1$ and $R = 2$.

Quadrature with respect to the uniform measure on this domain can exactly integrate polynomials using a tensor-product quadrature over (r, θ, ϕ) space using Fourier quadrature in θ and ϕ and an Legendre-Gauss quadrature in the r variable. We use a large enough quadrature rule so that all integrals are numerically exact. In Figure 4 we again observe that **MS** outperforms **MM-M** and **MM-L**. However, these two moment-based procedures give more reasonable results in this case since the polynomial degree is relatively low.

6.2 Moments via Monte Carlo techniques

We now consider two more complicated domains, where integration is performed approximately using a Monte Carlo quadrature rule. I.e., we approximate moments with respect to a uniform measure over a domain D via,

$$\int_D p(x) dx \approx \frac{1}{M} \sum_{m=1}^M p(x_m),$$

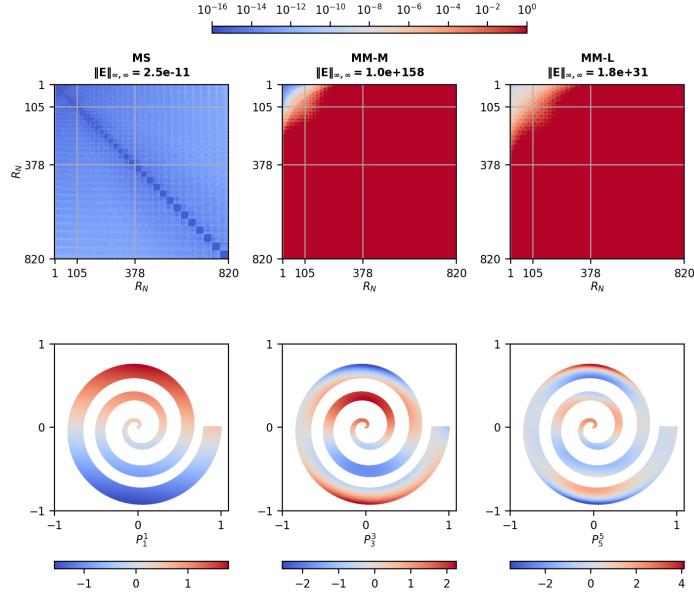


Figure 3: CUR results, Section 6.1.3. Top: Visualization of the error matrix E for the MS, MM-M, and MM-L algorithms from left to right. Bottom: Evaluations of the r_n 'th entry of p_n for degree $n = 1, 3, 5$ using the MS algorithm.

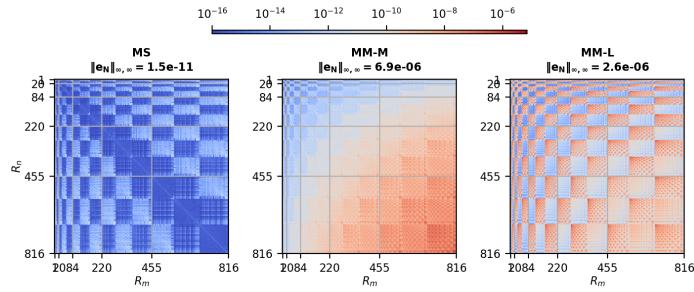


Figure 4: TOR results, Section 6.1.4. Visualization of the error matrix E for the MS, MM-M, and MM-L algorithms from left to right.

where D is a two-dimensional domain and $\{x_m\}_{m=1}^M$ are iid random samples from μ . In all examples we use a single, fixed instance of the Monte Carlo samples. Therefore one can consider this computing approximately orthogonal polynomials with respect to the uniform measure over D , or as computing (numerically) exactly orthogonal polynomials with respect to a size- M discrete measure. We emphasize again that our goal is not to construct accurate quadrature rules, but rather to construct orthogonal polynomials given some quadrature rules. In all simulations here we take $M = 10^8$.

6.2.1 HOL: Square with a hole

We consider the uniform measure μ over the two dimension domain $[-1, 1]^2 \setminus B_1(0)$, where $B_1(0)$ is the origin-centered unit ball of radius 1. Figure 5 shows results for this experiment. The MS algorithm again performs the best, but we see a notably increased error in the orthogonality metric compared with the previous examples. We attribute this increased error to an increase in the condition number of the associated matrices of the MS. We investigate this in more detail in Section 6.3.

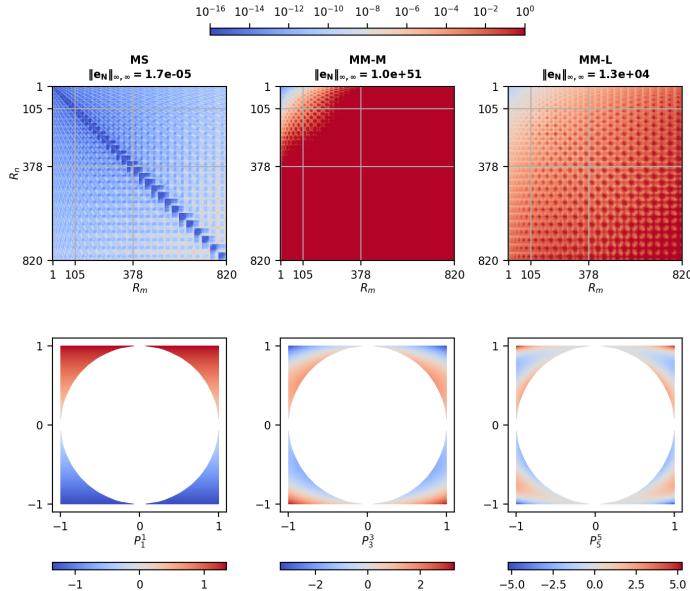


Figure 5: HOL results, Section 6.2.1: Visualization of the error matrix E for the MS, MM-M, and MM-L algorithms from left to right. Bottom: Evaluations of the r_n 'th entry of p_n for degree $n = 1, 3, 5$ using the MS algorithm.

6.2.2 MAP: Measure on map of Madagascar

Our final two-dimensional example is the region of the country of Madagascar. We draw random samples from this region via rejection sampling over a latitude-longitude bounding box, where the rule for inclusion in the domain is defined by positive elevation, which can be sampled via the data in [22]. We map the bounding box to $[-1, 1]^2$ for simpler plotting.

As can be seen in Figure 6, our orthogonality metric explodes very quickly for the MM-M and MM-L, even for relatively small polynomial degree. The MS succeeds to a much greater degree, but produces relatively large errors. Again we attribute this to increased ill-conditioning of associated matrices in our procedure, see the uptrend of condition number in Section 6.3.

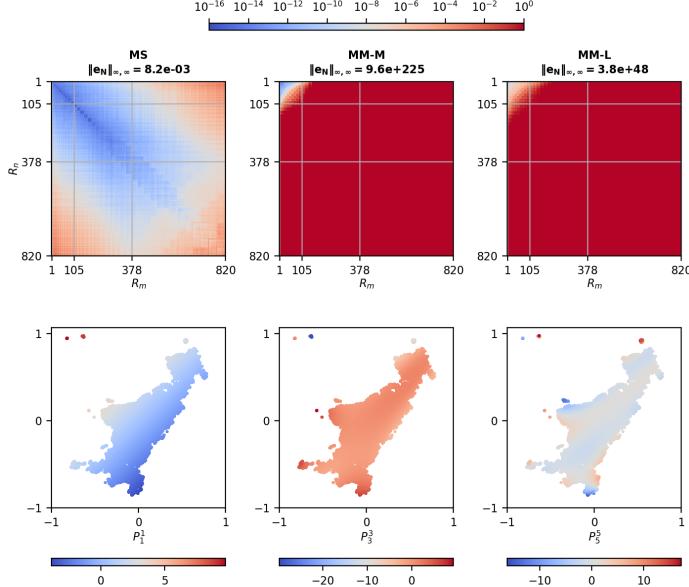


Figure 6: MAP results, Section 6.2.2: Top: Visualization of the error matrix E for the **MS**, **MM-M**, and **MM-L** algorithms from left to right. Bottom: Evaluations of the r_n 'th entry of \mathbb{p}_n for degree $n = 1, 3, 5$ using the **MS** algorithm.

6.3 Stability investigation via condition numbers

We justify the accuracy of the **MS**, **MM-M**, and **MM-L** algorithms by investigating the condition numbers of some of the associated matrices in each procedure. For the **MM-M** and **MM-L** algorithms, we investigate the condition number of the Gram matrix G_n in (16). For **MS**, we investigate the condition number of the moment matrix $T_{n,i,i}$ in (23) plotting the average of condition number of $T_{n,i,i}$ over all i .

Figure 7 shows these condition numbers for all our previous examples. We note that the G_n matrices are badly ill-conditioned for larger degrees, but the T_n matrices are much better conditioned. In addition, we see that for the **MAP** and **HOL** cases the condition number of T_n is larger than for other cases, which motivates why even the **MS** algorithm struggles for these domains.

6.4 The Christoffel function

The ability to stably compute an orthogonal basis in multiple dimensions allows us to investigate interesting phenomena. Let μ be uniform over a compact set in \mathbb{R}^d , and consider the diagonal K_N of the degree- N normalized reproducing kernel, and its inverse λ_N , the Christoffel function,

$$K_N(x) = \frac{1}{R_N} \sum_{n=0}^N \mathbb{p}_n^T(x) \mathbb{p}_n(x), \quad \lambda_N(x) = 1/K_N(x),$$

so that $K_N(x)d\mu(x)$ is a probability density. Random sampling from this probability measure is known to result in near-optimal sample complexity for constructing least-squares polynomial approximations to functions in the L_μ^2 norm [5]. Plotting such densities is itself interesting, but even more so is that fact that as $N \uparrow \infty$, such densities weakly converge to the Monge-Ampère measure over $\text{supp}(\mu)$ [2, 18], which is a fundamental quantity of theoretical interest in polynomial approximation in several variables [17, 3]. As analytic forms for such measures are unknown for general domains, it is interesting to use numerical algorithms to investigate the finite but large N behavior of K_N , which is not possible directly without the ability to stably compute an orthonormal basis. Figure 8 plots both K_N and λ_N for four of our two-dimensional domains with $N = 39$.

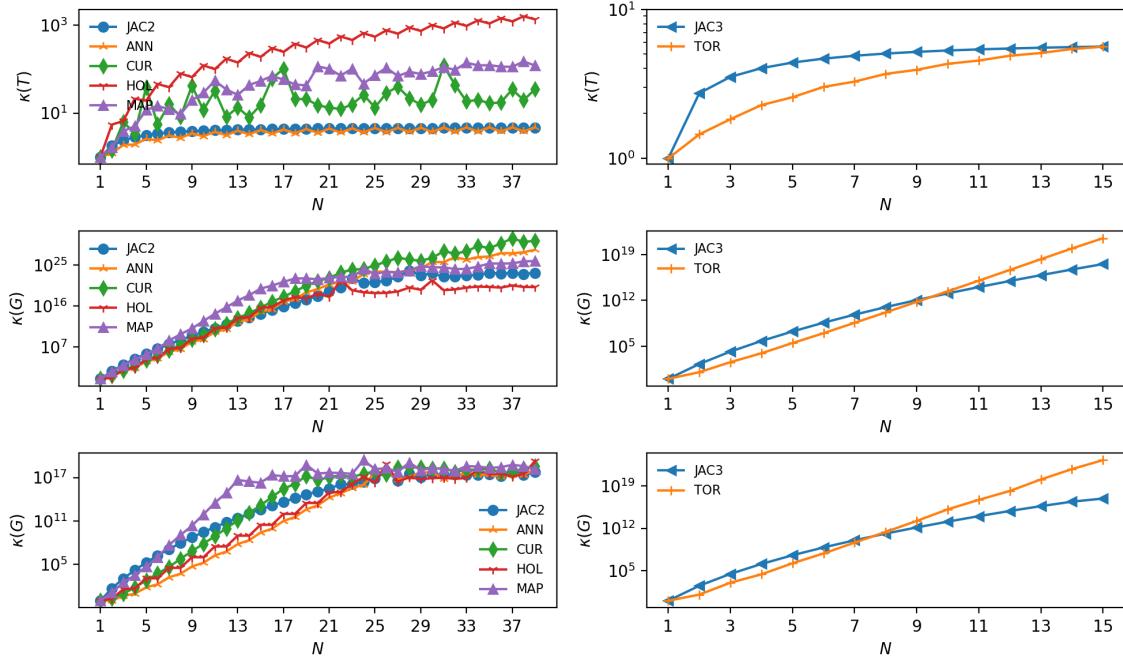


Figure 7: Top row: Condition numbers of the moment matrices $T_{n,i,i}$ in (23) used in the MS algorithm (averaged over $i \in [d]$). Middle row: Condition numbers for the Gram matrices G_n in (16) used in the MM-M algorithm. Bottom row: Same as the middle row but for the MM-L algorithm.

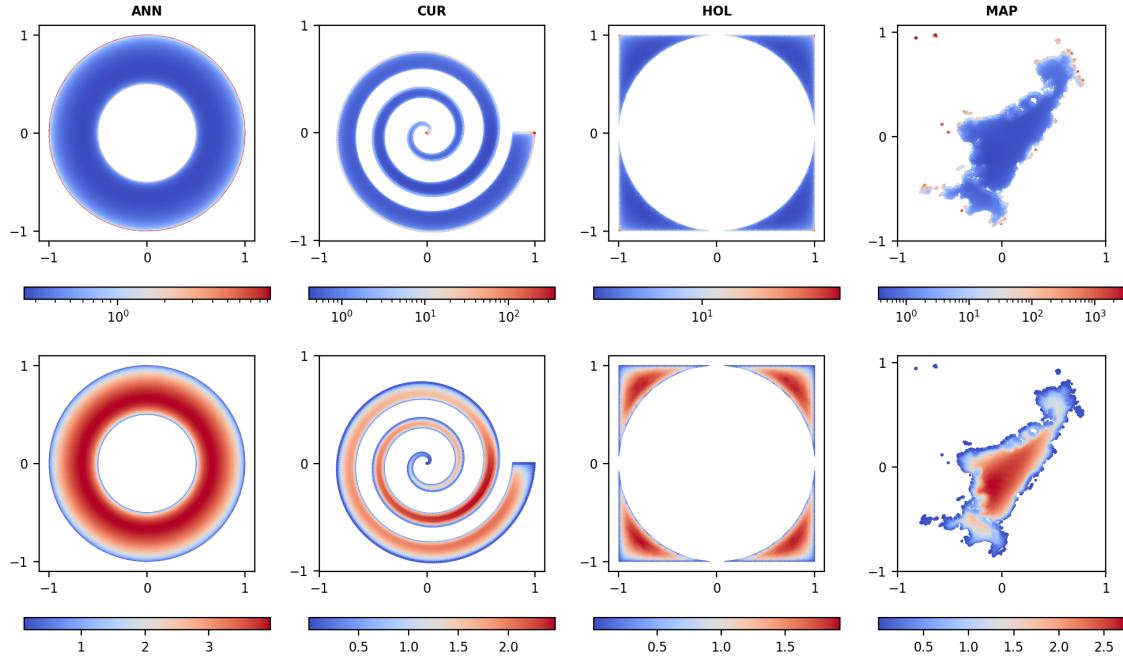


Figure 8: The normalized reproducing diagonal K_N (top row) and normalized Christoffel function λ_N (bottom row) for $N = 39$.

7 Conclusions

In this paper, we extend existing approaches for computing recurrence coefficients from the univariate case to the multivariate case. We propose a new, Multivariate Stieltjes (MS) algorithm for computing recurrence matrices that allows stable evaluation of multivariate orthonormal polynomials.

We demonstrate with several numerical examples the substantially improved stability of the new algorithm compared to direct orthogonalization approaches. For both small dimension and small polynomial degree, there is little benefit, but MS outperforms other methods when one requires polynomials of moderately large degree.

The algorithm is essentially explicit in two and three dimensions, but requires the numerical solution to a non-convex optimization problem in more than three dimensions, whose investigation would be a natural extension of this work.

References

- [1] R. BARRIO, J. M. PEÑA, AND T. SAUER, *Three term recurrence for the evaluation of multivariate orthogonal polynomials*, Journal of Approximation Theory, 162 (2010), pp. 407–420.
- [2] R. BERMAN AND S. BOUCKSOM, *Growth of balls of holomorphic sections and energy at equilibrium*, Inventiones mathematicae, 181 (2010), pp. 337–394.
- [3] T. BLOOM, L. P. BOS, J.-P. CALVI, AND N. LEVENBERG, *Polynomial interpolation and approximation in \mathbb{C}^d* , Annales Polonici Mathematici, 106 (2012), pp. 53–81.
- [4] P. L. CHEBYSHEV, *Sur l’interpolation par la méthode des moindres carrés*, Mémoires de l’Académie Impériale des sciences de St.-Pétersbourg, (1859), pp. 1–24.
- [5] A. COHEN AND G. MIGLIORATI, *Optimal weighted least-squares methods*, SMAI Journal of Computational Mathematics, 3 (2017), pp. 181–203. arxiv:1608.00512 [math.NA].
- [6] C. F. DUNKL AND Y. XU, *Orthogonal polynomials of several variables*, no. 155, Cambridge University Press, 2014.
- [7] W. GAUTSCHI, *A Survey of Gauss-Christoffel Quadrature Formulae*, in E. B. Christoffel, P. L. Butzer and F. Fehér, eds., Birkhäuser Basel, 1981, pp. 72–147.
- [8] ———, *On some orthogonal polynomials of interest in theoretical chemistry*, BIT Numerical Mathematics, 24 (1984), pp. 473–483.
- [9] ———, *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, USA, June 2004.
- [10] ———, *Orthogonal polynomials, quadrature, and approximation: computational methods and software (in matlab)*, in Orthogonal polynomials and special functions, Springer, 2006, pp. 1–77.
- [11] J. C. GOWER AND G. B. DIJKSTERHUIS, *Procrustes Problems*, Oxford Statistical Science Series, Oxford University Press, Oxford, 2004.
- [12] W. B. GRAGG AND W. J. HARROD, *The numerically stable reconstruction of jacobi matrices from spectral data*, Numerische Mathematik, 44 (1984), pp. 317–335.
- [13] D. JACKSON, *Formal properties of orthogonal polynomials in two variables*, Duke Mathematical Journal, 2 (1936), pp. 423–434.
- [14] M. KOWALSKI, *Orthogonality and recursion formulas for polynomials in n variables*, SIAM Journal on Mathematical Analysis, 13 (1982), pp. 316–323.

- [15] ———, *The recursion formulas for orthogonal polynomials in n variables*, SIAM Journal on Mathematical Analysis, 13 (1982), pp. 309–315.
- [16] H. KRALL AND I. SHEFFER, *Orthogonal polynomials in two variables*, Annali di Matematica Pura ed Applicata, 76 (1967), pp. 325–376.
- [17] N. LEVENBERG, *Approximation in C^N* , Surveys in Approximation Theory, 2 (2006), pp. 92–140. arXiv: math/0611249.
- [18] N. LEVENBERG, *Weighted Pluripotential Theory Results of Berman-Boucksom*, arXiv:1010.4035, (2010).
- [19] Z. LIU AND A. NARAYAN, *On the computation of recurrence coefficients for univariate orthogonal polynomials*, Journal of Scientific Computing, 88 (2021), p. 53.
- [20] H. RUTISHAUSER, *On jacobi rotation patterns*, in Proceedings of Symposia in Applied Mathematics, vol. 15, 1963, pp. 219–239.
- [21] R. A. SACK AND A. F. DONOVAN, *An algorithm for Gaussian quadrature given modified moments*, Numerische Mathematik, 18 (1971), pp. 465–478.
- [22] W. H. SMITH AND D. T. SANDWELL, *Global sea floor topography from satellite altimetry and ship depth soundings*, Science, 277 (1997), pp. 1956–1962.
- [23] T. J. STIELTJES, *Quelques recherches sur la théorie des quadratures dites mécaniques*, in Annales scientifiques de l’École Normale Supérieure, vol. 1, 1884, pp. 409–426.
- [24] ———, *Some research on the theory of so-called mechanical quadratures*, in Scientific annals of the ’Ecole Normale Supérieure, vol. 1, 1884, pp. 409–426.
- [25] G. SZEGÖ, *Orthogonal Polynomials*, American Mathematical Soc., 4th ed., 1975.
- [26] W. VAN ASSCHE, *Discrete painlevé equations for recurrence coefficients of orthogonal polynomials*, in Difference equations, special functions and orthogonal polynomials, World Scientific, 2007, pp. 687–725.
- [27] S. WALDRON, *Recursive Three-Term Recurrence Relations for the Jacobi Polynomials on a Triangle*, Constructive Approximation, 33 (2011), pp. 405–424.
- [28] J. C. WHEELER, *Modified moments and Gaussian quadratures*, Rocky Mountain Journal of Mathematics, 4 (1974), pp. 287–296.
- [29] Y. XU, *On multivariate orthogonal polynomials*, SIAM journal on mathematical analysis, 24 (1993), pp. 783–794.
- [30] ———, *Multivariate orthogonal polynomials and operator theory*, Transactions of the American Mathematical Society, 343 (1994), pp. 193–202.
- [31] W. ZU CASTELL, F. FILBIR, AND B. FORSTER, *Inzell lectures on orthogonal polynomials*, Nova Publishers, 2005.

A Algorithms

Algorithm 1: Algorithm via three-term relation

```

Input: dimension  $d$  and univariate recurrence coefficients  $\{a_{i,n}, b_{i,n}\}_{n=0}^N$ ,  $i \in [d]$ 
1 for  $n = 1$  to  $N$  do
2   for  $i = 1$  to  $d$  do
3     | compute  $c(\alpha^{(n,k)}, i)$  defined in Section 4.1 ;
4     | compute  $A_{n,i}, B_{n,i}$  by (15)
5   end
6 end
Output: coefficient matrices  $\{A_{n,i}, B_{n,i}\}_{n=0}^N$ 
```

Algorithm 2: Stieltjes procedure

```

Input: dimension  $d$ , max degree  $N$ , measure  $d\mu$  on support  $\Omega$ .
1 for  $n = 0$  to  $N - 1$  do
2   for  $i = 1$  to  $d$  do
3     | evaluate orthonormal polynomials  $p_n$  by (9) ;
4     | compute  $A_{n+1,i}$  using moment matrices ((25)) ;
5     | evaluate the modified polynomials basis  $\tilde{p}_{n,i}$  by (24) ;
6     | compute symmetric moments  $T_{n,i,i}$  by (23) ;
7     | compute  $U_{n+1,i}$  and  $\Sigma_{n+1,i}$  in the SVD of  $B_{n+1,i}$  from (26) ;
8     | for  $j = i + 1$  to  $d$  do
9       | | compute the mixed moment  $T_{n,i,j}$  by (23) ;
10    | determine  $\hat{V}_{n+1,1}$  and  $\tilde{V}_{n+1,i}$  by (22), and assemble the matrix  $V_{n+1,1}$  by (21) ;
11    | compute  $B_{n+1,1}$  by (20) ;
12    | for  $j = 2$  to  $d$  do
13      | | compute  $\hat{V}_{n+1,j}$  by (28) ;
14      | | compute  $\mathbf{y}\mathbf{y}^T$  by (29) ;
15      | | if  $d=2$  then
16        | | | compute  $\mathbf{y}$  by (30)
17      | | else if  $d=3$  then
18        | | | if  $n=0$  then
19          | | | | compute  $B_{n+1,j}$  following the strategy in Section 5.2.5
20        | | | else
21          | | | | compute  $K_{n+1,j}$  from (32) and its kernel  $\Psi_j$  ;
22          | | | | compute  $E_j$  from (34) ;
23          | | | if  $j=2$  then
24            | | | | Set  $W_2 = I_{\Delta_{r_n+1}}$ 
25          | | | else
26            | | | | compute  $W_3$  from (36) following the strategy in Section 5.2.7
27          | | | | assemble the matrix  $V_{n+1,j}$  by (21) ;
28        | | else
29          | | | solve  $W_j$  from (35)
30          | | | determine  $C_j$  by (34), and thus  $\tilde{V}_{n+1,j}$  by (33)
31    | | compute  $B_{n+1,j}$  by (20) ;
Output: coefficient matrices  $\{A_{n,i}, B_{n,i}\}_{n=0}^N$ 
```