# A Stieltjes algorithm for generating multivariate orthogonal polynomials

Zexin Liu<sup>1</sup> and Akil Narayan<sup>2</sup>

<sup>1, 2</sup>Department of Mathematics, and Scientific Computing and Imaging (SCI) Institute, University of Utah

February 11, 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.

**Keywords.** Multivariate orthogonal polynomials, Recurrence coefficient matrices, Stieltjes procedure

# 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 [25, 9, 10].

It is well-known even in the multivariate setting that such families of polynomials satisfy threeterm recurrence relations [12, 15, 14, 13, 28, 29, 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, 18]. Such procedures are absent in the multivariate setting; this paper provides one algorithmic solution to fill this gap.

## 1.1 Challenges with computing multivariate 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  $\mu$  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  $\mu$  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 [22, 23, 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 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 theory 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 identify non-uniqueness issues for multivariate polynomials and propose a *canonical basis* that removes much of this ambiguity and identifies a computational strategy for direct evaluation of multivariate polynomials. We follow this by Section 4, which shows that if  $\mu$  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, 26], 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, 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  $\mu$  as the empirical measure associated to a large number of realizations that are independently and identically distributed according to  $\mu$ . Thus, sometimes our numerical examples compute orthogonal polynomials with respect to an approximate measure. 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 (via recurrence matrices).

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 dand 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  $\mu$  is non-degenerate with respect to *d*-variate polynomials. This is ensured if, for example,  $\mu$  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 is denoted by  $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ . For  $\alpha \in \mathbb{N}_0^d$ , and  $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ , a monomial in variables  $x_1, \ldots, x_d$  of an index  $\alpha$  is defined by  $x^{\alpha} = x_1^{\alpha_1}, \ldots, x_d^{\alpha_d}$ . The number  $|\alpha| \coloneqq \alpha_1 + \cdots + \alpha_d$  is the degree of  $x^{\alpha}$ . We denote the space of *d*-variate polynomials of exactly degree  $n \in \mathbb{N}_0$ , and up to degree *n*, respectively, by

$$\mathbb{P}_n^d \coloneqq \operatorname{span} \{ x^\alpha : |\alpha| = n, \alpha \in \mathbb{N}_0^d \}, \qquad \qquad \Pi_n^d \coloneqq \operatorname{span} \{ x^\alpha : |\alpha| \le n, \alpha \in \mathbb{N}_0^d \},$$

The dimensions of these spaces are, respectively,

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

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,

$$\Delta r_n \coloneqq r_n - r_{n-1} = \binom{n+d-2}{n}, \qquad n \ge 0, \qquad d \ge 2,$$

where we define  $r_{-1} \coloneqq 0$ . We will only be concerned with  $\Delta r_n$  when d = 3.

Throughout we assume that  $\mu$  is a given probability measure on  $\mathbb{R}^d$ , with  $d \in \mathbb{N}$ . The support of  $\mu$  may be unbounded. We assume that for any nontrivial polynomial p,

$$0 < \int_{\mathbb{R}^d} p^2(x) \mathrm{d}\mu(x) < \infty, \tag{1}$$

and we implicitly assume that  $\int p(x)d\mu(x)$  is computationally available. Our main goal is to compute an orthonormal basis for  $L^2_{\mu}(\mathbb{R}^d;\mathbb{R})$ , where orthonormality is defined through the inner product,

$$\langle f,g \rangle = \int f(x)g(x)\mathrm{d}\mu(x)$$

### 2.1 Univariate orthogonal polynomials

In the univariate case, let  $\mu$  be a positive Borel measure on  $\mathbb{R}$  with finite moments,  $m_n \coloneqq \int_{\mathbb{R}} x^n d\mu(x)$ . The standard Gram-Schmidt process applied to the moment sequence  $\{m_n\}$  yields a sequence of orthonormal polynomials  $\{p_n(x)\}_{n=0}^{\infty}$  satisfying  $\langle p_n, p_m \rangle = \delta_{m,n}$ , where  $\delta_{m,n}$  is the Kronecker delta. In practice, 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

$$xp_n(x) = b_{n+1}p_{n+1}(x) + a_{n+1}p_n(x) + b_n p_{n-1}(x), \qquad n \ge 0, \qquad (2)$$

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 [24], and in cases when they are not, several algorithms exist to accurately compute approximations [4, 19, 20, 27, 7, 11, 9, 18]. The procedure we present in this paper is generalization of the (univariate) Stieltjes procedure [22, 23, 8].

# 2.2 Multivariate orthogonal polynomials

Now let  $\mu$  be a probability measure on  $\mathbb{R}^d$  with finite moments. For  $\alpha \in \mathbb{N}_0^d$ , let  $\mu_\alpha = \int_{\mathbb{R}^d} x^\alpha d\mu(x)$  be the moments of  $\mu$ . Again a Gram-Schmidt process applied to the multivariate monomials  $x^\alpha$  produces a sequence of orthogonal polynomials in several variables. In order to directly apply a Gram-Schmidt process, a linear order of the moments  $\mu_\alpha$  is required, translating to a required order of multi-indices  $\alpha$  in  $\mathbb{N}_0^d$ . No such ordering is unique, but we will proceed in a graded total degree fashion. We shall say that  $p \in \prod_n^d$  is an orthogonal polynomial of degree n > 0 with respect to  $d\mu$  if

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

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  $\mu$  implies that dim  $\mathcal{V}_n^d = \dim \Pi_n^d = r_n$ . This allows us to state results in terms of size- $r_n$  vector functions containing orthonormal bases of  $\mathcal{V}_n^d$  for  $n \ge 0$ .

Let  $\{p_{\alpha} : \alpha \in \mathbb{N}_0^d\}$  be a sequence of orthogonal polynomials in d variables identified through the Gram-Schmidt procedure with the partial order  $|\alpha|$  on multi-indices. Then  $\{p_{\alpha} : |\alpha| = n\}$  is a basis of  $\mathcal{V}_n^d$ , motivating the vector notation:

$$\mathbf{p}_n = (p_\alpha)_{|\alpha|=n} = (p_{\alpha^{(n,1)}}, \dots, p_{\alpha^{(n,r_n)}})^T$$
(3)

where  $\alpha^{(n,1)}, \ldots, \alpha^{(n,r_n)}$  is any arrangement of elements in the set  $J_n \coloneqq \{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$ , 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

$$x_i \mathbb{p}_n(x) = B_{n+1,i} \mathbb{p}_{n+1}(x) + A_{n+1,i} \mathbb{p}_n(x) + B_{n,i}^T \mathbb{p}_{n-1}(x), \qquad 1 \le i \le d.$$
(4)

where we define  $\mathbb{p}_{-1} = 0$  and  $\mathbb{p}_0(x) = 1$ . Note that if d = 1, then  $\mathbb{p}_n = p_n$ , and  $(A_{n,1}, B_{n,1}) = (a_n, b_n)$ .

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 conditions on the recurrence matrices that define the appropriate feasible set corresponding to multivariate polynomials can be described through a multivariate version of Favard's Theorem. The first Favard-like condition is a set of necessary conditions called the *commuting conditions*.

**Theorem 2.1** ([30, 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  $i, j \in [d]$ :

$$B_{n+1,i}B_{n+1,j}^T + A_{n+1,i}A_{n+1,j} + B_{n,i}^T B_{n,j} = B_{n+1,j}B_{n+1,i}^T + A_{n+1,j}A_{n+1,i} + B_{n,j}^T B_{n,i}, \quad n \ge 0,$$
(5a)

$$B_{n,i}A_{n+1,j} + A_{n,i}B_{n,j} = B_{n,j}A_{n+1,i} + A_{n,j}B_{n,i}, \qquad n \ge 1,$$

$$B_{n,i}B_{n+1,j} = B_{n,j}B_{n+1,i},$$
  $n \ge 1.$ 

(5c)

(5b)

One way to informally motivate these conditions is as follows: Note that (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. Mechanically, these constraints (5) are identified through two different computations of identical matrix moments. Note that if d = 1 then the commuting conditions (5) are vacuous since the recurrence matrices reduce to scalars.

The multivariate Favard Theorem combines the commuting condition with a rank condition on the  $B_{n,i}$  matrices.

**Theorem 2.2** ([30, Theorem 2.3]). If

rank
$$B_n = r_n$$
,  $B_n \coloneqq \begin{pmatrix} B_{n,1} \\ B_{n,2} \\ \vdots \\ B_{n,d} \end{pmatrix} \in \mathbb{R}^{dr_{n-1} \times r_n}$ , (6)

and the recurrence matrices satisfy the commuting conditions (5), then (4) generates a set of polynomials orthonormal with respect to some positive-definite functional. Furthermore, recurrence matrices corresponding to any orthogonal family satisfy the rank condition above.

The 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:

$$A_{n+1,i} = \int_{\mathbb{R}^d} x_i \mathbb{p}_n(x) \mathbb{p}_n^T(x) \mathrm{d}\mu, \qquad B_{n+1,i} = \int_{\mathbb{R}^d} x_i \mathbb{p}_n(x) \mathbb{p}_{n+1}^T(x) \mathrm{d}\mu.$$
(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* polynomials.

# 3 Non-uniqueness and evaluation of polynomials

There are two computational hurdles we must address before discussing algorithms for computing recurrence matrices. The first hurdle is the non-uniqueness of elements in the basis  $\mathbb{p}_n$ , even though  $\mathcal{V}_n^d$  is unique. In one dimension, the only non-uniqueness corresponds to a multiplicative sign, but in more than one dimension we have non-uniqueness corresponding to arbitrary orthogonal transformations. The second challenge is that (4) does not provide a direct strategy for evaluating orthogonal polynomials since  $B_{n+1,i}$  is a rectangular, possibly rank-deficient matrix. We address these challenges simultaneously by introducing a canonical form for the basis vectors.

### 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 the equations (4) for all  $i \in [d]$  simultaneously. To aid in this type of procedure, we make a special choice of orthonormal basis that amounts to choosing a particular sequence of unitary transformations in section 3.

**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

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

where  $B_n$  is as defined in (6), and the matrices  $\{\Lambda_n\}_{n \in \mathbb{N}}$  are a sequence of diagonal matrices with elements of each matrices 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 Theorem 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.** 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  $n \ge 0$ , 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 (6), ensuring that  $r_{n+1} = \operatorname{rank}(B_{n+1}) = \operatorname{rank}(\Lambda_{n+1})$ .

The theorem above, in particular (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  $\Lambda_{n+1}$ . An alternative procedure in [1] explicitly transforms known recurrence matrices into a relation that can be used to evaluate monic polynomials. The main requirement for our evaluation technique above 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

Although the recurrence matrices  $A_{n,i}$  and  $B_{n,i}$  are unique if we fix the basis  $p_n$ , there is substantial freedom in how the basis is chosen. In particular, let  $\{U_n\}_{n\geq 0}$  be an arbitrary family of unitary matrices with sizes,

$$U_n \in \mathbb{R}^{r_n \times r_n}, \qquad \qquad U_n^T U_n = I.$$

Then a new orthonormal basis  $q_n$  can be defined,

$$\mathbf{q}_n(x) \coloneqq U_n \mathbf{p}_n(x). \tag{10}$$

The bases  $\{p_n\}_{n\geq 0}$  and  $\{q_n\}_{n\geq 0}$  satisfy a type of (total degree-)graded biorthogonality property:

- $\langle p_{\alpha}, p_{\beta} \rangle_{\mu} = \delta_{\alpha,\beta}$  for every  $\alpha, \beta \in \mathbb{N}_0^d$
- $\langle q_{\alpha}, q_{\beta} \rangle_{\mu} = \delta_{\alpha,\beta}$  for every  $\alpha, \beta \in \mathbb{N}_0^d$
- $\langle p_{\alpha}, q_{\beta} \rangle_{\mu} = 0$  if  $|\alpha| \neq |\beta|$ .

In addition, a manipulation of (4) shows that the basis elements  $q_n$  satisfy a three-term recurrence,

$$x_{i}\mathbf{q}_{n}(x) = D_{n+1,i}\mathbf{q}_{n+1}(x) + C_{n+1,i}\mathbf{q}_{n}(x) + D_{n,i}^{T}\mathbf{q}_{n-1}(x), \quad 1 \le i \le d,$$

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

$$C_{n,i} = U_{n-1}A_{n,i}U_{n-1}^T, D_{n,i} = U_{n-1}B_{n,i}U_n^T, (11)$$

for every  $i \in [d]$  and  $n \in \mathbb{N}_0$ . 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,

$$B_n^T B_n = V_n \Lambda_n V_n^T, \tag{12}$$

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

$$U_n = V_n^T, \qquad n \ge 1,$$

which identifies  $C_{n,i}$  and  $D_{n,i}$  through (11), then 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  $q_n$  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  $\Lambda_n$  have distinct values, then each of the functions in the vector  $\mathbf{q}_n$  is unique up to a multiplicative sign.

# 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  $\mu$ . If  $\mu$  is a tensorial measure, i.e.,

$$\mu = \bigotimes_{j=1}^{a} \mu_j,$$

where each  $\mu_j$  is a probability 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,

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

with starting conditions  $p_{j,-1} \equiv 0$  and  $p_{j,0}(x_j) = 1/b_{j,0}$ , are  $L^2_{\mu_j}(\mathbb{R})$ -orthonormal,

$$\int_{\mathbb{R}} p_{j,n}(x_j) p_{j,m}(x_j) d\mu_j(x_j) = \delta_{m,n}, \qquad n, m \in \mathbb{N}_0,$$

with deg  $p_{j,n} = n$ . With these univariate polynomials, we can directly construct multivariate orthonormal polynomials,

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

with deg  $p_{\alpha} = |\alpha|$ . We now construct vectors containing polynomials of a specific degree as in (3). Specifically, let  $J_n$  denote any ordered set of the multi-indices  $\alpha$  satisfying  $|\alpha| = n$ ,

$$J_n \coloneqq \left(\alpha^{(n,1)}, \dots, \alpha^{(n,r_n)}\right), \qquad n \ge 0.$$

We have  $|\alpha^{(n,k)}| = n$  for each  $k \in [r_n]$ . Then we define  $\mathbb{p}_n(x)$  as a vector of size  $r_n$ ,

$$\mathbf{p}_n = \left(p_{\alpha^{(n,1)}}, \dots, p_{\alpha^{(n,r_n)}}\right)^T.$$
(14)

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  $c: \mathbb{N}_0^d \times [d] \to \mathbb{N}$  defined as

$$c\left(\alpha^{(n,k)},i\right)$$
 = The index q 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  $(p_n)_{n=0}^{\infty}$  defined by (14) and (13), then the recurrence matrices are given by,

$$A_{n+1,i} = \text{diag}\left(a_{i,\alpha_i^{(n,k)}+1}\right)_{k=1}^{r_n}, \qquad i \in [d], \qquad (15a)$$

where  $\alpha_i^{(n,k)}$  is the *i*th component of  $\alpha^{(n,k)}$ , and

$$B_{n+1,i} = \sum_{k=1}^{r_n} b_{i,\alpha_i^{(n,k)}+1} e_{r_n,k} e_{r_{n+1},c(\alpha^{(n,k)},i)}^T,$$
(15b)

where  $e_{m,j}$  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 kth component of the vector  $\mathbb{p}_n$  satisfies,

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

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).

The result above is not necessarily new (see a specialized case in, e.g., [6, Section 3.3.4]), but we have not seen this formula in its generality appearing in the literature. When  $\mu$  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  $\mu_j$  is well-studied and has several solutions (cf. section 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):

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

In order for  $\mathbb{p}_n$  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  $\mathbb{p}_n$ , 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  $P_n : [r_{n-1}] \rightarrow [r_{n-1}]$  given by,

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

for each  $k \in [r_n]$ . Defining a new basis,

$$\mathbf{q}_n \coloneqq \left( p_{\alpha^{(n,P(k))}} \right)_{k=1}^{r_n},$$

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  $\mu$ , allowing us to compute errors relative to exact matrices in some numerical examples.

# 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) that directly uses moments 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. The MS algorithm is substantially more complicated than its univariate counterpart.

### 5.1 The moment method (MM)

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. We describe the analogous multivariate procedure here.

First, we suppose that a(ny) polynomial basis  $\{\phi_j\}_{j \in \mathbb{N}}$  is given with the properties,

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

Again, a simple example is that  $\phi_j(x)$  is a multivariate monomial  $x^{\alpha}$  for some total ordering of the multi-indices  $\alpha$  that respects the partial order induced by the  $\ell^1(\mathbb{N}_0^d)$  norm<sup>1</sup>. We assume that

<sup>&</sup>lt;sup>1</sup>Degree-graded lexicographic ordering over multi-indices is an example of one such ordering.

quadratic  $\phi_i$  moments are available, i.e., we have access to a function such that

$$m(i,j) = \int \phi_i(x)\phi_j(x)d\mu(x), \qquad \alpha \in \mathbb{N}_0^d.$$
(16)

This implies in particular that we can compute the Gram matrices,

$$G_n \in \mathbb{R}^{R_n \times R_n}, \qquad (G_n)_{i,j} = \int \phi_i(x)\phi_j(x)d\mu(x) = m(i,j). \qquad (17)$$

Given this matrix, we can compute monomial expansion coefficients for the orthonormal basis  $\mathbb{p}_n$ . Algebraically, this can be accomplished by orthogonalizing the columns of an identity matrix with respect to the quadratic form  $\langle v, w \rangle_n = w^T G_n v$ . As a consequence, the monomial expansion coefficients are given by the inverse Cholesky factor of  $G_n$ ,

$$G_n = L_n L_n^T, \qquad (\mathbb{p}_0, \ldots, \mathbb{p}_n)^T = L_n^{-1} \Phi_n, \qquad \Phi_n = (\phi_1, \ldots, \phi_{R_n})^T \in \mathbb{R}^{R_n}, \qquad (18)$$

where  $p_j$  is an  $r_j$ -vector containing degree-j orthonormal polynomials. Combining this with (7), the recurrence matrices can be directly computed,

$$A_{n+1,i} = \tilde{L}_n^{-1} G_{n,i} \tilde{L}_n^{-T}, \qquad B_{n+1,i} = \tilde{L}_n^{-1} \tilde{G}_{n+1,i} \tilde{L}_{n+1}^{-T}, \qquad i \in [d], \qquad (19)$$

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

$$(G_{n,i})_{i,j} = \int x_i \phi_i(x) \phi_j(x) \mathrm{d}\mu(x), \qquad i, j \in [R_n], \qquad (20)$$

and  $\widetilde{G}_{n+1,i} \in \mathbb{R}^{R_n \times R_{n+1}}$  equals the first  $R_n$  rows of  $G_{n+1,i}$ . Thus, so long as the polynomial moments m(i, j) and (20) 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 (18).

The main drawbacks to the procedures above stem from accuracy concerns due to ill-conditioning of the  $G_n$  matrices. If the  $\phi_i$  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 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  $i \in [d]$ .

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 Theorem 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  $B_{n+1,i}$ ,

$$B_{n+1,i} = U_{n+1,i} \Sigma_{n+1,i} V_{n+1,i}^T, \qquad i \in [d], \qquad (21)$$

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

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

and  $\Sigma_{n+1,i}$  is diagonal with non-negative entries in non-increasing order. We furthermore decompose the right-singular  $V_{n+1,i}$  matrices into blocks, with  $\hat{V}$  the first  $r_n$  rows, and  $\tilde{V}$  the remaining rows:

$$V_{n+1,i} = \begin{pmatrix} \widehat{V}_{n+1,i} \\ \widetilde{V}_{n+1,i} \end{pmatrix}, \qquad \qquad \widehat{V}_{n+1,i} \in \mathbb{R}^{r_n \times r_n} \qquad \qquad \widetilde{V}_{n+1,i} \in \mathbb{R}^{\Delta r_{n+1} \times r_n}$$
(22)

Thus, all together we have rewritten (21) as,

$$B_{n+1,i} = U_{n+1,i} \Sigma_{n+1,i} \left( \hat{V}_{n+1,i}^T \ \tilde{V}_{n+1,i}^T \right).$$
(23)

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 the blocks of  $V_{n+1,1}$ ,

$$\widehat{V}_{n+1,1} = I_{r_n}, \qquad \qquad \widetilde{V}_{n+1,1} = \mathbf{0}.$$
 (24)

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 more moment matrices provides a way to compute the U and  $\Sigma$  SVD matrices of B. We introduce "mixed" moments in Section 5.2.3 that allow us to obtain the  $\hat{V}$  block of the Vmatrices. The remaining block  $\tilde{V}$  is computed using different strategies depending on the dimension d of the problem. For d = 2, Section 5.2.4 shows that  $\tilde{V}$  can be computed almost directly. For  $d \geq 3$  dimensions, we must enforce the commuting conditions (5a), which is easily done in d = 3dimensions, 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,

$$S_{n,i} \coloneqq \int x_i \mathbb{p}_n \mathbb{p}_n^T \mathrm{d}\mu(x) \in \mathbb{R}^{r_n \times r_n}, \qquad T_{n,i,j} \coloneqq \int \widetilde{\mathbb{p}}_{n+1,i} \widetilde{\mathbb{p}}_{n+1,j}^T \mathrm{d}\mu(x) \in \mathbb{R}^{r_n \times r_n},$$
(25)

where the modified polynomial basis  $\tilde{p}$  is defined as,

$$\widetilde{\mathbf{p}}_{n+1,i}(x) \coloneqq x_i \mathbf{p}_n - A_{n+1,i} \mathbf{p}_n - B_{n,i}^T \mathbf{p}_{n-1}.$$
(26)

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  $S_{n,i}$  in (25) can be approximated via quadrature, just as is frequently done for the  $\phi_j$  moments of Section 5.1.

Inspection of (7) immediately reveals that,

$$A_{n+1,i} = S_{n,i},\tag{27}$$

and hence  $A_{n+1,i}$  is directly computable. This then allows  $\tilde{p}_{n+1,i}$  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  $\Sigma$  matrices in the SVD of B. A direct computation with the three-term recurrence (4) and the definition (25) reveals that

$$T_{n,i,i} = B_{n+1,i} B_{n+1,i}^T = U_{n+1,i} \Sigma_{n+1,i}^2 U_{n+1,i}^T.$$
(28)

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  $\Sigma_{n+1,i}$ . Thus, we are left only to compute the rectangular  $V_{n+1,i}$  matrices.

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

Using mixed moments  $T_{n,i,j}$  with  $i \neq j$ , we can compute the square matrices  $\widehat{V}_{n+1,i}$ , which are subblocks of  $V_{n+1,i}$ . (Recall from (24) that  $V_{n+1,j}$  is already known for j = 1, so we can consider j > 2 without loss.) A similar computation as the Stieltjes procedure in (28) 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 (23), we conclude that

$$\widehat{V}_{n+1,i}^T \widehat{V}_{n+1,j} + \widetilde{V}_{n+1,i}^T \widetilde{V}_{n+1,j} = \Sigma_{n+1,i}^{-1} U_{n+1,i}^T T_{n,i,j} U_{n+1,j} \Sigma_{n+1,j}^{-1}, \qquad i, j \in [d].$$
(29)

Letting i = 1 and utilizing (24), we have,

$$\widehat{V}_{n+1,j} = \Sigma_{n+1,1}^{-1} U_{n+1,1}^T T_{n,1,j} U_{n+1,j} \Sigma_{n+1,j}^{-1}, \qquad j \ge 2, \qquad (30)$$

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  $\widetilde{V}$  can already be computed. When  $d \geq 3$ , we do require  $1 \neq i \neq j$  to identify  $\widetilde{V}$ .

# **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  $V_{n+1,2}$  is a  $1 \times r_n$  vector. To reduce notational clutter, we consider fixed *n* and use the following notation,

$$\boldsymbol{y}\coloneqq \widetilde{V}_{n+1,2}^T\in \mathbb{R}^{r_n}$$

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,

$$yy^{T} = I_{r_{n}} - \hat{V}_{n+1,2}^{T} \hat{V}_{n+1,2}, \qquad (31)$$

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

$$\boldsymbol{y} = \pm \boldsymbol{z},\tag{32}$$

where z is computed either as a rank-1 Cholesky factor or as the positive semi-definite square root of the right-hand side of (31). Although it appears the multiplicative sign needs to be chosen, but if we choose (i, j) = (1, 2) in (5c), then we have

$$\left( B_{n,1} U_{n+1,2} \Sigma_{n+1,2} \widehat{V}_{n+1,2} \pm B_{n,1} U_{n+1,2} \Sigma_{n+1,2} \boldsymbol{z} \right) = B_{n,2} B_{n+1,1}$$

Such a choice of signs makes no difference because the last column of  $B_{n+1,1}$  is a zero vector.

#### **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  $\widetilde{V}_{n+1,i} \in \mathbb{R}^{1 \times d-1}$ . We have two conditions to impose on this vector of length  $d - 1 \ge 2$ :

- The scalar-valued commuting condition (5a) (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 (19) to compute the  $B_{n+1}$  matrix.

## 5.2.6 d > 2, n > 0: The commuting conditions – computing $V_{n+1,i}$

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 (5c) with i = 1 implies:

$$B_{n,1}U_{n+1,j}\Sigma_{n+1,j}\left(\widehat{V}_{n+1,j}^T \ \widetilde{V}_{n+1,j}^T\right) = B_{n,j}B_{n+1,1}, \qquad j \ge 2, \qquad (33)$$

where only  $\widetilde{V}_{n+1,j}$  is unknown. Note that we have made the choice (24) for  $\widetilde{V}_{n+1,1}$ , which implies that the last  $\Delta r_{n+1}$  columns of  $B_{n,j}B_{n+1,1}$  vanish, i.e., the last  $\Delta r_{n+1}$  columns of (33) read,

$$K_{n+1,j}\widetilde{V}_{n+1,j} = \mathbf{0}, \qquad \qquad K_{n+1,j} \coloneqq B_{n,1}U_{n+1,j}\Sigma_{n+1,j} \in \mathbb{R}^{r_{n-1} \times r_n}.$$
(34)

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

$$\widetilde{V}_{n+1,j}^T = \Psi_j C_j, \qquad \qquad \Psi_j \in \mathbb{R}^{r_n \times \Delta r_n}, \qquad \qquad C_j \in \mathbb{R}^{\Delta r_n \times \Delta r_{n+1}}, \tag{35}$$

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

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

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

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

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

$$C_j = E_j W_j, \qquad E_j \coloneqq \left(\sqrt{D_j} \quad \mathbf{0}_{\Delta r_n \times (\Delta r_{n+1} - \Delta r_n)}\right), \qquad W_j \in \mathbb{R}^{\Delta r_{n+1} \times \Delta r_{n+1}}, \tag{36}$$

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 (29) and writing in terms of the unknown  $W_i, W_j$  yields,

$$E_i W_i W_j^T E_j^T = H_{n,i,j} \qquad 2 \le i, j \ge d, \quad i \ne j \qquad (37)$$
  
subject to  $W_j^T W_j = I_{\Delta r_{n+1}} \qquad 2 \le j \le d.$ 

where,

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

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  $\widetilde{V}_{n+1,i}$  is determined through (35) and (36). This is a non-convex optimization, appearing as 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. 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 (37) 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 (37), 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 (35) and (36).

The determination of  $W_3$  in (37) now reduces to instance of a WOPP:

$$E_2 W_3^T E_3^T = H_{n,2,3}, \quad \text{subject to} \quad W_3 \in O(\Delta r_{n+1})$$
(38)

We now notice that  $E_j$  defined in (36) 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

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

the reduced singular value decompositions of  $E_2$  and  $E_3$ , then (38) can be rewritten as

$$I_{(n+1)\times(n+2)}WI_{(n+2)\times(n+1)} = Y_2^{-1}X_2^TH_{n,2,3}X_3Y_3^{-1}, \qquad W = Z_2^TW_3Z_3,$$

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  $Y_2^{-1}X_2^T H_{n,2,3}X_3Y_3^{-1}$ . To determine the last row and column of W, we write,

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

and since W is an orthogonal matrix and must have orthonormal columns, then the first n + 1 entries of  $\boldsymbol{w}$  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  $\boldsymbol{v}$  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 (37) 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. All the algorithms are implemented in Python and all the computations are carried out on a MacBook Pro laptop with a 3.1 GHz Intel(R) Core(TM) i5 processor and 8 GB of RAM.

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  $\mathbb{p}_n$ . We then use quadrature (approximate or exact as described) to evaluate the Gram matrices,

$$M_{m,n} = \int \mathbb{p}_m(x) \mathbb{p}_n^T(x) \mathrm{d}\mu(x), \qquad (39)$$

Example	Abbreviation	Section
tensorized Jacobi weight measure	JAC	6.1.1
Measure on an annulus	ANN	6.1.2
Measure within polar curves	CUR	6.1.3
Measure on a rectangle with a hole	HOL	6.2.1
Measure on map of Madagascar	MAP	6.2.2
Measure on a torus	TOR	6.1.4

 Table 1: Abbreviation, subsection, and algorithm for each method. Also included is a modern citation that explains each algorithm

and given a maximum degree N, subsequently construct a Gramian error matrix,

$$E = \begin{pmatrix} M_{0,0} & \cdots & M_{0,N} \\ \vdots & \ddots & \vdots \\ M_{N,0} & \cdots & M_{N,N} \end{pmatrix} - I_{R_N} \in \mathbb{R}^{R_N \times R_N}$$

$$\tag{40}$$

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  $\mu$  is a tensor-product measure, we use the procedure in section 4 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 1.
- (MM-M) The moment method of section 5.1, involving direct orthogonalization of the monomial basis, i.e., the functions  $\phi_j$  are monomials  $x^{\alpha}$ .
- (MM-L) The moment method of section 5.1, involving direct orthogonalization of the tensorial Legendre polynomial basis, i.e., the functions  $\phi_j$  are tensorial Legendre polynomials with respect to the uniform measure  $\nu$  whose support is a bounding box of the support of  $\mu$ .

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  $\mu$ ) are given in Table 1

### 6.1 Experiments with moments via tensorized Gaussian quadrature

In this subsection we compute polynomial moments with respect to the measure  $\mu$  via tensorized Gaussian quadrature. In all cases except subsection 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,

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

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  $\alpha_i, \beta_i$  parameters by uniformly sampling over the interval (-1, 10), resulting in the following choices:



$$(\alpha_1, \alpha_2) = (3.80, 0.78), \qquad (\beta_1, \beta_2) = (7.34, 8.26), \qquad d = 2$$
  
$$(\alpha_1, \alpha_2, \alpha_3) = (1.61, 0.32, 3.01), \qquad (\beta_1, \beta_2, \beta_3) = (-0.89, 9.83, 7.67), \qquad d = 3$$

Figure 1: JAC results, 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 O(1) error; we continue to impose this saturation value for all subsequent error plots.

Figure 1 visualizes the Gramian error matrix E in (40) 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  $\mu$  with support in an annular region in d = 2 dimensions centered at the origin. In polar coordinates  $(r, \theta)$ , this is represented as,

$$\operatorname{supp}(\mu) = \left\{ (r,\theta) \mid \theta_1 \le \theta \le \theta_2, \ r_1(\theta) \le r \le r_2(\theta) \right\},\tag{41}$$

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  $(r, \theta)$  space using Fourier quadrature in  $\theta$  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.



Figure 2: ANN results, 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.

### 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 (41), but this time it is a region bounded between two Archimedean spirals. In particular, we set,

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

Again we choose  $\mu$  as the uniform measure over the region defined in (41).

We write integrals as iterated, with the inner integral over r exactly computable using Legendre-Gauss quadrature. But the outer integral in  $\theta$  involves terms both polynomial and trigonometric polynomial in  $\theta$ , and we approximately integrate these values with a 10<sup>6</sup>-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,

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

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  $(r, \theta, \phi)$  space using Fourier quadrature in  $\theta$  and  $\phi$  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.



**Figure 3:** CUR results, 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  $\mathbb{P}_n$  for degree n = 1, 3, 5 using the MS algorithm.



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

## 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 repsect to a uniform measure over a domain D via,

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

where D is a two-dimensional domain and  $\{x_m\}_{m=1}^M$  are iid random samples from  $\mu$ . 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 rule. In all simulations in this section we take  $M = 10^8$ .

### 6.2.1 HOL: Square with a hole

We consider the uniform measure  $\mu$  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 the three procedures MS, MM-M, and MM-L perform. The MSprocedure is still the best among these three, 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 algorithm. We investigate this in more detail in section 6.3.



Figure 5: HOL results, 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 [21]. 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 6.3.



Figure 6: MAP results, 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  $p_n$  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 (17). For MS, we investigate the condition number of the moment matrix  $T_{n,i,i}$  in (25) 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  $\mu$  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  $\lambda_N$ , the Christoffel function,

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

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



Figure 7: Top row: Condition numbers of the moment matrices  $T_{n,i,i}$  in (25) used in the MS algorithm (averaged over  $i \in [d]$ ). Middle row: Condition numbers for the Gram matrices  $G_n$  in (17) used in the MM-M algorithm. Bottom row: Same as the middle row but for the MM-L algorithm.

approximations to functions in the  $L^2_{\mu}$  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 supp( $\mu$ ) [2, 17], which is a fundamental quantity of theoretical interest in polynomial approximation in several variables [16, 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 Nbehavior of  $K_N$ , which is not possible directly without the ability to stably compute an orthonormal basis. Figure 8 plots both  $K_N$  and  $\lambda_N$  for four of our two-dimensional domains with 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 multi-



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

variate 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 C<sup>d</sup>, Annales Polonici Mathematici, 106 (2012), pp. 53–81.
- [4] P. L. CHEBYSHEV, Sur l'interpolation par la méthode des moindres carrés, Mémoires de lÁcadémie 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] W. B. GRAGG AND W. J. HARROD, The numerically stable reconstruction of jacobi matrices from spectral data, Numerische Mathematik, 44 (1984), pp. 317–335.
- [12] D. JACKSON, Formal properties of orthogonal polynomials in two variables, Duke Mathematical Journal, 2 (1936), pp. 423–434.
- [13] M. KOWALSKI, Orthogonality and recursion formulas for polynomials in n variables, SIAM Journal on Mathematical Analysis, 13 (1982), pp. 316–323.
- [14] —, The recursion formulas for orthogonal polynomials in n variables, SIAM Journal on Mathematical Analysis, 13 (1982), pp. 309–315.
- [15] H. KRALL AND I. SHEFFER, Orthogonal polynomials in two variables, Annali di Matematica Pura ed Applicata, 76 (1967), pp. 325–376.
- [16] N. LEVENBERG, Approximation in C<sup>N</sup>, Surveys in Approximation Theory, 2 (2006), pp. 92– 140. arXiv: math/0611249.
- [17] N. LEVENBERG, Weighted Pluripotential Theory Results of Berman-Boucksom, arXiv:1010.4035, (2010).
- [18] Z. LIU AND A. NARAYAN, On the computation of recurrence coefficients for univariate orthogonal polynomials, Journal of Scientific Computing, 88 (2021), p. 53.
- [19] H. RUTISHAUSER, On jacobi rotation patterns, in Proceedings of Symposia in Applied Mathematics, vol. 15, 1963, pp. 219–239.
- [20] R. A. SACK AND A. F. DONOVAN, An algorithm for Gaussian quadrature given modified moments, Numerische Mathematik, 18 (1971), pp. 465–478.
- [21] W. H. SMITH AND D. T. SANDWELL, Global sea floor topography from satellite altimetry and ship depth soundings, Science, 277 (1997), pp. 1956–1962.
- [22] 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.
- [23] —, Some research on the theory of so-called mechanical quadratures, in Scientific annals of the 'E cole Normale Sup é rieure, vol. 1, 1884, pp. 409–426.
- [24] G. SZEGÖ, Orthogonal Polynomials, American Mathematical Soc., 4th ed., 1975.
- [25] 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.
- [26] S. WALDRON, Recursive Three-Term Recurrence Relations for the Jacobi Polynomials on a Triangle, Constructive Approximation, 33 (2011), pp. 405–424.
- [27] J. C. WHEELER, Modified moments and Gaussian quadratures, Rocky Mountain Journal of Mathematics, 4 (1974), pp. 287–296.
- [28] Y. XU, On multivariate orthogonal polynomials, SIAM journal on mathematical analysis, 24 (1993), pp. 783–794.
- [29] —, Multivariate orthogonal polynomials and operator theory, Transactions of the American Mathematical Society, 343 (1994), pp. 193–202.
- [30] W. ZU CASTELL, F. FILBIR, AND B. FORSTER, Inzell lectures on orthogonal polynomials, Nova Publishers, 2005.

# A Algorithms

Algorithm 1: Stieltjes procedure **Input:** dimension d, max degree N, measure  $d\mu$  on support  $\Omega$ . **Initialization:**  $A_{0,i} = [0]$  and  $B_{0,i} = [1]$  for  $1 \le i \le d$ 1 for n = 0 to N - 1 do for i = 1 to d do  $\mathbf{2}$ evaluate orthonormal polynomials  $p_n$  by (9); 3  $\mathbf{4}$ compute  $A_{n+1,i}$  using moment matrices (27); evaluate the modified polynomials basis  $\widetilde{p}_{n,i}$  by (26);  $\mathbf{5}$ compute symmetric moments  $T_{n,i,i}$  by (25); 6 compute  $U_{n+1,i}$  and  $\Sigma_{n+1,i}$  in the SVD of  $B_{n+1,i}$  from (28);  $\mathbf{7}$ for j = i + 1 to d do 8 compute the mixed moment  $T_{n,i,j}$  by (25); 9 determine  $V_{n+1,1}$  and  $V_{n+1,i}$  by (24), and assemble the matrix  $V_{n+1,1}$  by (22); 10 compute  $B_{n+1,1}$  by (21); 11 for j = 2 to d do 12 compute  $\widehat{V}_{n+1,j}$  by (30); 13 compute  $\boldsymbol{y}\boldsymbol{y}^T$  by (31);  $\mathbf{14}$ if d=2 then 15compute  $\boldsymbol{y}$  by (32)  $\mathbf{16}$ else if d=3 then 17 if n=0 then 18 compute  $B_{n+1,j}$  following the strategy in Section 5.2.5 19  $\mathbf{20}$ else compute  $K_{n+1,j}$  from (34) and its kernel  $\Psi_j$ ;  $\mathbf{21}$ compute  $E_j$  from (36); 22 if j=2 then  $\mathbf{23}$ Set  $W_2 = I_{\Delta_{r_n+1}}$  $\mathbf{24}$  $\mathbf{else}$  $\mathbf{25}$ compute  $W_3$  from (38) following the strategy in Section 5.2.7  $\mathbf{26}$ assemble the matrix  $V_{n+1,j}$  by (22);  $\mathbf{27}$ else 28  $\mathbf{29}$ solve  $W_i$  from (37) determine  $C_i$  by (36), and thus  $\widetilde{V}_{n+1,i}by(35)$ 30 compute  $B_{n+1,j}$  by (21); 31 **Output:** coefficient matrices  $\{A_{n,i}, B_{n,i}\}_{n=0}^N$