TWO VARIABLE ORTHOGONAL POLYNOMIALS AND STRUCTURED MATRICES[*]

ANTONIA M. DELGADO[†], JEFFREY S. GERONIMO[‡], PLAMEN ILIEV[‡], AND FRANCISCO MARCELLÁN[†]

Abstract. We consider bivariate real valued polynomials orthogonal with respect to a positive linear functional. The lexicographical and reverse lexicographical orderings are used to order the monomials. Recurrence formulas are derived between polynomials of different degrees. These formulas link the orthogonal polynomials constructed using the lexicographical ordering with those constructed using the reverse lexicographical ordering. Relations between the coefficients in the recurrence formulas are established and used to give necessary and sufficient conditions for the existence of a positive linear functional. Links to the theory of matrix orthogonal polynomials are developed as well the consequences of a zero assumption on one of the coefficients in the the recurrence formulas.

Key words. bivariate orthogonal polynomials, positive linear functionals, moment problem, Hankel matrices.

AMS subject classifications. 42C05, 30E05, 47A57

DOI. 10.1137/05062946X

1. Introduction. Bivariate orthogonal polynomials have been investigated by many authors. Special examples of these types of polynomials have arisen in studies related to symmetry groups [3], [14], [20], as extensions of one variable polynomials [5], [13], and as eigenfunctions of partial differential equations [12], [17], [11], [19] (see also the references in [4]). The general theory of these polynomials can trace its origins back to [10] and an excellent review of the theory can be found in [4] (see also [21]). A major difficulty encountered in the theory of orthogonal polynomials of more than one variable is which monomial ordering to use. Except for the special cases that have arisen from the subject mentioned above, the preferred ordering is the total degree ordering which is the one set by Jackson [4]. For polynomials with the same total degree the ordering is lexicographical. There is a good reason to use this ordering, which is that if new orthogonal polynomials of higher degree are to be constructed, then their orthogonality relations will not affect the relations governing the lower degree polynomials. This can be seen especially in Xu's vector formulation of the problem [22] (see also [2], [6], and [15], [16]). However, in their work on the Fejér–Riesz factorization problem, Geronimo and Woerdeman [8], [9] noticed that the most useful ordering was the lexicographical ordering and reverse lexicographical ordering. Important in their work were the relations of the orthogonal polynomials in these orderings. The reason for this is that in these orderings the moment matrix is a structured matrix, i.e., it is a block Toeplitz matrix where the blocks are themselves Toeplitz matrices. In the

---

[†]Departamento de Matemáticas, Universidad Carlos III de Madrid, 28911 Leganés, Spain (adelgado@math.uc3m.es, pacomar@ing.uc3m.es). These authors were partially supported by grant BFM2003-06335-C03-02 from the Dirección General de Investigación, Ministerio de Educación y Ciencia of Spain.

[‡]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332–0160 (geronimo@math.gatech.edu, iliev@math.gatech.edu).

one variable case the connection between orthogonal polynomials and the Hankel or Toeplitz matrices associated with them plays an important role in the theory. The coefficients in the recurrence formulas for the orthogonal polynomials give a parameterization of positive definite Hankel or Toeplitz matrices. Furthermore, structured matrices come up in a variety of engineering and physics problems and so the orthogonal polynomials associated with them need to be investigated. The aim of this paper is to study orthogonal polynomials associated with positive definite block Hankel matrices whose entries are also Hankel and to develop methods for constructing such matrices. We proceed as follows. In section 2 we consider finite subspaces of monomials of the form $x^i y^j$, $0 \leq i \leq 2n$, $0 \leq j \leq 2m$, and show the connection between positive linear functionals defined on this space and positive doubly Hankel matrices, i.e., block Hankel matrices whose blocks are Hankel matrices. These structured matrices arise when using the lexicographical or reverse lexicographical ordering on the monomials. We then introduce certain matrix orthogonal polynomials and show how they give the Cholesky factors for the doubly Hankel matrix considered above. These polynomials may be thought of as arising from a parameterized moment problem. In section 3 we construct two variable orthogonal polynomials, where the monomials are ordered according to the lexicographical ordering. When these polynomials are organized into vector orthogonal polynomials they can be related to the matrix orthogonal polynomials constructed previously. From this relation it is shown that these vector polynomials are the minimizers of a certain quadratic functional. Using the orthogonality relation, recurrence relations satisfied by the vector polynomials and their counterparts in the reverse lexicographical ordering are derived and some elementary properties of the matrices entering these recurrence relations are deduced. Because of the size and shape of the coefficients in the recurrence formulas they must be related. In section 4 we derive and examine these relations, and in section 5 a number of Christoffel–Darboux-like formulas are derived. In section 6 we use the relations between the coefficients derived in section 4 to develop an algorithm to construct the coefficients in the recurrence formulas at a particular level, (n, m), say, in terms of the coefficients at the previous levels plus a certain number of unknowns. The collection of these unknowns is in one-to-one correspondence with the number of moments needed to construct the vector polynomials up to level (n, m). This is used in section 7 to construct a positive linear functional from the recurrence coefficients. The construction allows us to find necessary and sufficient conditions on the recurrence coefficients for the existence of a positive linear functional which is in one to one correspondence with the set of positive definite "doubly" Hankel matrices. In the above construction an important role is played by a set of matrices that must be contractions. In section 8 we explore the consequences of setting these contractive matrices equal to zero and show that this condition characterizes product measures. Finally in section 9 we give a numerical example for the case $n = 2$, $m = 2$, which illustrates the above algorithm. We also present an example for which the moment problem is not extendable.

2. Positive linear functionals and Hankel matrices. In this section we consider moment matrices associated with the lexicographical ordering which is defined by

$$(k, \ell) <_{\text{lex}} (k_1, \ell_1) \Leftrightarrow k < k_1 \text{ or } (k = k_1 \text{ and } \ell < \ell_1)$$

and the reverse lexicographical ordering defined by

$$(k, \ell) <_{\text{revlex}} (k_1, \ell_1) \Leftrightarrow (\ell, k) <_{\text{lex}} (\ell_1, k_1).$$

Both of these orderings are linear orders and in addition satisfy

$$(k, \ell) < (m, n) \Rightarrow (k + p, \ell + q) < (m + p, n + q).$$

Note that none of these orderings respects the total degree. Denote $\prod^{n,m}(x, y)$ as the span$\{x^i y^j, 0 \le i \le n, 0 \le j \le m\}$. Let $\mathcal{L}_{n,m}$ be a linear functional defined on $\prod^{2n,2m}(x, y)$ by

$$\mathcal{L}_{n,m}(x^i y^j) = h_{i,j}.$$

We will call $h_{i,j}$ the (i, j) moment of $\mathcal{L}_{n,m}$ and $\mathcal{L}_{n,m}$ a moment functional. If we form the $(n+1)(m+1) \times (n+1)(m+1)$ matrix $H_{n,m}$ for $\mathcal{L}_{n,m}$ in the lexicographical ordering then, as noted in the introduction, it has the special form

$$(2.1) \qquad H_{n,m} = \begin{bmatrix} H_0 & H_1 & \cdot^{\cdot^\cdot} & H_n \\ H_1 & H_2 & & H_{n+1} \\ \cdot^{\cdot^\cdot} & & \cdot^{\cdot^\cdot} & \cdot^{\cdot^\cdot} \\ H_n & H_{n+1} & \cdot^{\cdot^\cdot} & H_{2n} \end{bmatrix},$$

where each H_i is a $(m+1) \times (m+1)$ matrix of the form

$$(2.2) \qquad H_i = \begin{bmatrix} h_{i,0} & h_{i1} & \cdot^{\cdot^\cdot} & h_{i,m} \\ h_{i1} & h_{i2} & \cdot^{\cdot^\cdot} & \\ \cdot^{\cdot^\cdot} & \cdot^{\cdot^\cdot} & & \cdot^{\cdot^\cdot} \\ h_{i,m} & & \cdot^{\cdot^\cdot} & h_{i,2m} \end{bmatrix}, \qquad i = 0, \ldots, 2n.$$

Thus $H_{n,m}$ is a block Hankel matrix where each block is a Hankel matrix so it has a doubly Hankel structure. If the reverse lexicographical ordering is used in place of the lexicographical ordering we obtain another moment matrix $\tilde{H}_{n,m}$ where the roles of n and m are interchanged. We have the following useful lemmas, which characterize doubly Hankel matrices. An analogous characterization of doubly Toeplitz matrices was given in [9].

LEMMA 2.1. *Let H be a real square $k \times k$ matrix and H^1 be the same as H except with the first row and last column removed. Then H is a Hankel matrix if and only if $H = H^\top$ and $H^1 = H^{1\top}$.*

Proof. Recall that $H = (h_{i,j}) = (h_{i+j})$ characterizes a Hankel matrix. Thus the necessary conditions of the lemma follow immediately. To prove the converse, note that $H = H^\top$ implies that $h_{i,j} = h_{j,i}$. Since $H^1 = (h^1_{i,j}) = (h_{i+1,j})$, $i = 1, \ldots, k-1$, $j = 1, \ldots, k-1$, the second condition implies that

$$h_{i+1,j} = h_{j+1,i}.$$

Thus $h_{i+1,j} = h_{i,j+1}$, which completes the result. □

LEMMA 2.2. *Let $H = (H_{i,j})$, $i = 1, \ldots, k$, $j = 1, \ldots, k$, where each $H_{i,j}$ is a real $m \times m$ matrix. Then H is a doubly Hankel matrix if and only if $H = H^\top$, $H^1 = H^{1\top}$, and $H^2 = H^{2\top}$. Here H^1 is obtained from H by deleting the first block row and last block column and H^2 is obtained from H by removing the first row and last column of each $H_{i,j}$.*

Proof. Again the necessary conditions follow from the structure of H. To see the converse, note that $H^\top = H$ implies that $H_{i,j} = H_{j,i}^\top$ so that $H_{i,i} = H_{i,i}^\top$. The

condition on H^1 shows that $H_{i+1,j} = H_{i,j+1}$. Thus H is block Hankel with each entry being symmetric. The result now follows from Lemma 2.1.   □

Remark 2.3. The above results are true if the roles of the rows and columns are interchanged.

We say that the moment functional $\mathcal{L}_{n,m} : \prod^{2n,2m} \to \mathbb{R}$ is positive definite if

$$(2.3) \qquad \mathcal{L}_{n,m}(p^2) > 0$$

for all nonzero $p \in \prod^{n,m}$. Likewise, the moment functional $\mathcal{L}_{n,m} : \prod^{2n,2m} \to \mathbb{R}$ is nonnegative definite if $\mathcal{L}_{n,m}(p^2) \geq 0$ for every $p \in \prod^{n,m}$. Note that it follows from a simple quadratic form argument that $\mathcal{L}_{n,m}$ is positive definite or nonnegative definite if and only if its moment matrix $H_{n,m}$ is positive definite or nonnegative definite, respectively.

We will say that $\mathcal{L}$ is positive definite or nonnegative definite if

$$\mathcal{L}(p^2) > 0 \qquad \text{or} \qquad \mathcal{L}(p^2) \geq 0,$$

respectively, for all nonzero polynomials. Again these conditions are equivalent to the moment matrices $H_{n,m}$ being positive definite or nonnegative definite for all positive integers n and m.

From the above remark we easily find the next lemma.

LEMMA 2.4. *Let $H_{n,m}$ be a positive (nonnegative) definite $(n+1)(m+1) \times (n+1)(m+1)$ matrix given by (2.1) and (2.2). Then there is a positive (nonnegative) moment functional $\mathcal{L}_{n,m} : \prod^{2n,2m}(x,y) \to \mathbb{R}$ associated with $H_{n,m}$ given by*

$$h_{i,j} = \mathcal{L}_{n,m}(x^i y^j), \qquad 0 \leq i \leq 2n, \qquad 0 \leq j \leq 2m.$$

If the positive moment functional $\mathcal{L}_{n,m} : \prod^{2n,2m} \to \mathbb{R}$ is extended to two variable polynomials with matrix coefficients in the obvious way, we can associate to it a positive matrix function $\mathcal{L}_m : \prod^n_{m+1}(x) \times \prod^n_{m+1}(x) \to M^{m+1,m+1}$ defined by

$$(2.4) \qquad \mathcal{L}_m(P(x), Q(x)) = \mathcal{L}_{n,m}(P(x,y)\, Q^\top(x,y)),$$

where

$$P(x,y) = P(x) \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix} \text{ and } Q(x,y) = Q(x) \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix}.$$

Here, $\prod^n_{m+1}(x)$ is the set of all $(m+1) \times (m+1)$ real valued matrix polynomials of degree n or less and $M^{m,n}$ is the space of $m \times n$ matrices. Because of the structure of $H_{n,m}$ we can associate to $\mathcal{L}_m$ matrix valued orthogonal polynomials in the following manner. Let $\{R_i(x)\}^n_{i=0}$ and $\{L_i(x)\}^n_{i=0}$ be $(m+1) \times (m+1)$ real valued matrix polynomials given by

$$(2.5) \qquad R_i(x) = R_{i,i}x^i + R_{i,i-1}x^{i-1} + \cdots, \qquad i = 0, \ldots, n,$$

and

$$(2.6) \qquad L_i(x) = L_{i,i}x^i + L_{i,i-1}x^{i-1} + \cdots, \qquad i = 0, \ldots, n,$$

satisfying

$$(2.7) \qquad \mathcal{L}_m(R_i^\top, R_j^\top) = \delta_{ij} I_{m+1}$$

and

$$(2.8) \qquad \mathcal{L}_m(L_i, L_j) = \delta_{ij} I_{m+1},$$

respectively, where I_{m+1} denotes the $(m+1) \times (m+1)$ identity matrix. The above relations uniquely determine the sequences $\{R_i\}_{i=0}^n$ and $\{L_i\}_{i=0}^n$ up to a unitary factor and we fix this factor by requiring $R_{i,i}$ to be an upper triangular matrix with positive diagonal entries and $L_{i,i}$ to be a lower triangular matrix also with positive diagonal entries. From the defining equations (2.7) and (2.8) it follows that $R_i^\top = L_i$ hence we will concentrate on L_i. We write

$$(2.9) \qquad L_i(x) = [L_{i,0} \ L_{i,1} \ \cdots \ L_{i,i} \ 0 \ \cdots \ 0] \begin{bmatrix} I_{m+1} \\ x I_{m+1} \\ \vdots \\ x^n I_{m+1} \end{bmatrix}$$

and

$$(2.10) \qquad L^n(x) = \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_n(x) \end{bmatrix} = L \begin{bmatrix} I_{m+1} \\ x I_{m+1} \\ \vdots \\ x^n I_{m+1} \end{bmatrix},$$

where

$$(2.11) \qquad L = \begin{bmatrix} L_{0,0} & 0 & \cdots & 0 \\ L_{1,0} & L_{1,1} & \cdots & 0 \\ \vdots & & \ddots & \\ L_{n,0} & & \cdots & L_{n,n} \end{bmatrix}.$$

By lower A (respectively, upper B) Cholesky factor of a positive definite matrix M we mean

$$(2.12) \qquad M = AA^\top = BB^\top,$$

where A is a lower triangular matrix with positive diagonal elements and B is an upper triangular matrix with positive diagonal elements. With the above we have the following lemma.

LEMMA 2.5. *With the above normalization, $L^\top$ is the upper Cholesky factor of $H_{n,m}^{-1}$.*

Proof. Note that (2.8) implies that

$$I = \mathcal{L}_m(L^n, L^n) = L \mathcal{L}_m \left(\begin{bmatrix} I_{m+1} \\ x I_{m+1} \\ \vdots \\ x^n I_{m+1} \end{bmatrix}, \begin{bmatrix} I_{m+1} \\ x I_{m+1} \\ \vdots \\ x^n I_{m+1} \end{bmatrix} \right) L^\top = L H_{n,m} L^\top,$$

where I is the $(n+1)(m+1) \times (n+1)(m+1)$ identity matrix. Thus

$$H_{n,m}^{-1} = L^\top L. \qquad \square$$

From this formula and the fact that L is upper triangular we see that $L_{n,n}^\top$ is the upper Cholesky factor of $[0, \dots, I_{m+1}]H_{n,m}^{-1}[0, \dots, I_{m+1}]^\top$. Hence from (2.11) we find

$$(2.13) \qquad L_n(x) = \left[0, 0, \dots, 0, (L_{n,n}^\top)^{-1}\right] H_{n,m}^{-1} [I_{m+1}, xI_{m+1}, \dots, x^n I_{m+1}]^\top.$$

The theory of matrix orthogonal polynomials [1], [7], [18] can be applied to obtain the recurrence formula

$$(2.14) \quad \begin{aligned} xL_i(x) &= A_{i+1,m}L_{i+1}(x) + B_{i,m}L_i(x) + A_{i,m}^\top L_{i-1}(x), \quad i = 0, \dots, n-1, \\ L_{-1} &= 0, \end{aligned}$$

where

$$(2.15) \qquad\qquad A_{i+1,m} = \mathcal{L}_m(xL_i, L_{i+1}) = L_{i,i}L_{i+1,i+1}^{-1}$$

and

$$(2.16) \qquad\qquad B_{i,m} = \mathcal{L}_m(xL_i, L_i).$$

The above equations show that $B_{i,m}$ is an $(m+1) \times (m+1)$ real symmetric matrix and $A_{i,m}$ is an $(m+1) \times (m+1)$ real lower triangular matrix.

Routine manipulations of (2.14) using the fact that $B_{n,m}$ is self-adjoint give

$$(2.17) \quad \begin{aligned} & L_i^\top(x_1)A_{i+1,m}L_{i+1}(x) - L_{i+1}^\top(x_1)A_{i+1,m}^\top L_i(x) \\ & = (x - x_1)L_i^\top(x_1)L_i(x) + L_{i-1}^\top(x_1)A_{i,m}L_i(x) - L_i^\top(x_1)A_{i,m}^\top L_{i-1}(x), \end{aligned}$$

and iteration of this formula to $i = 0$ yields the important Christoffel–Darboux formula.

We note that the same results hold for the reverse lexicographical ordering with x replaced by y and the roles of n and m interchanged.

As in the scalar case, matrix orthogonal polynomials satisfy a minimization principle [7]. Let sym $\mathbb{R}_{m+1}$ be the space of $(m+1) \times (m+1)$ real symmetric matrices and let $\mathfrak{L} : \prod_{m+1}^n \to \text{sym } \mathbb{R}_{m+1}$ be given by

$$(2.18) \qquad\qquad \mathfrak{L}(Y) = \mathcal{L}_m(Y, Y) - 2\text{sym } Y_n.$$

Here

$$Y(x) = Y_n x^n + \cdots + Y_0 = [Y_0, Y_1, \dots, Y_n][I_{m+1}, xI_{m+1}, \dots, x^n I_{m+1}]^\top$$

and

$$\text{sym } Y_n = \frac{Y_n + Y_n^\top}{2}.$$

The equation (2.18) can be evaluated as

$$\mathfrak{L}(Y(x)) = [Y_0, Y_1, \dots, Y_n]H_{n,m}[Y_0, Y_1, \dots, Y_n]^\top - 2\text{sym } Y_n.$$

Set

$$X = [Y_0, Y_1, \ldots, Y_n] H_{n,m}^{1/2} - V_n H_{n,m}^{-1/2},$$

where any square root of $H_{n,m}$ may be used and where $V_n = [0, 0, \ldots, I_{m+1}]$. Then (2.18) becomes

$$(2.19) \qquad \mathfrak{L}(Y) = XX^\top - V_n H_{n,m}^{-1} V_n^\top.$$

Thus there is a unique $W \in \prod_{m+1}^n$, corresponding to $X = 0$ given by

$$(2.20) \qquad W(x) = V_n H_{n,m}^{-1} [I_{m+1}, x I_{m+1}, \ldots, x^n I_{m+1}]^\top,$$

that minimizes $\mathfrak{L}$ in the sense that

$$(2.21) \qquad \mathfrak{L}(W) \le \mathfrak{L}(Y)$$

for all $Y \in \prod_{m+1}^n$. From formula (2.13) we find

$$(2.22) \qquad L_n(x) = (L_{n,n}^\top)^{-1} W(x).$$

3. Lexicographic order and orthogonal polynomials. In this section we examine the properties of two variable orthogonal polynomials where the monomial ordering is either the lexicographical or reverse lexicographical. Given a positive definite linear functional $\mathcal{L}_{N,M} : \prod^{2N,2M} \to \mathbb{R}$ we perform the Gram–Schmidt procedure using the lexicographical ordering and define the orthonormal polynomials $p_{n,m}^l(x, y)$, $0 \le n \le N$, $0 \le m \le M$, $0 \le l \le m$, by the equations

$$(3.1) \quad \begin{aligned} &\mathcal{L}_{N,M}(p_{n,m}^l, x^i y^j) = 0, \quad 0 \le i < n \text{ and } 0 \le j \le m \text{ or } i = n \text{ and } 0 \le j < l, \\ &\mathcal{L}_{N,M}(p_{n,m}^l, p_{n,m}^l) = 1, \end{aligned}$$

and

$$(3.2) \qquad p_{n,m}^l(x, y) = k_{n,m,l}^{n,l} x^n y^l + \sum_{(i,j) <_{\text{lex}} (n,l)} k_{n,m,l}^{i,j} x^i y^j.$$

With the convention $k_{n,m,l}^{n,l} > 0$, the above equations uniquely specify $p_{n,m}^l$. Polynomials orthonormal with respect to $\mathcal{L}_{N,M}$ but using the reverse lexicographical ordering will be denoted by $\tilde{p}_{n,m}^l$. They are uniquely determined by the above relations with the roles of n and m interchanged.

Set

$$(3.3) \qquad \mathbb{P}_{n,m} = \begin{bmatrix} p_{n,m}^0 \\ p_{n,m}^1 \\ \vdots \\ p_{n,m}^m \end{bmatrix} = K_{n,m} \begin{bmatrix} 1 \\ y \\ \vdots \\ x^n y^m \end{bmatrix},$$

where the $(m+1) \times [(n+1)(m+1)]$ matrix $K_{n,m}$ is given by

$$(3.4) \qquad K_{n,m} = \begin{bmatrix} k_{n,m,0}^{0,0} & k_{n,m,0}^{0,1} & \cdots & k_{n,m,0}^{n,0} & 0 & \cdots \\ k_{n,m,1}^{0,0} & k_{n,m,1}^{0,1} & \cdots & k_{n,m,1}^{n,0} & k_{n,m,1}^{n,1} & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ k_{n,m,m}^{0,0} & k_{n,m}^{0,1} & \cdots & \cdots & \cdots & k_{n,m,m}^{n,m} \end{bmatrix}.$$

As indicated above denote

$$(3.5) \qquad \tilde{\mathbb{P}}_{n,m} = \begin{bmatrix} \tilde{p}_{n,m}^0 \\ \tilde{p}_{n,m}^1 \\ \vdots \\ \tilde{p}_{n,m}^n \end{bmatrix} = \tilde{K}_{n,m} \begin{bmatrix} 1 \\ x \\ \vdots \\ x^n y^m \end{bmatrix},$$

where the $(n+1) \times [(n+1)(m+1)]$ matrix $\tilde{K}_{n,m}$ is given similarly to (3.4) with the roles of n and m interchanged. In order to find recurrence formulas for the vector polynomials $\mathbb{P}_{n,m}$ we introduce the inner product

$$(3.6) \qquad \langle X, Y \rangle = \mathcal{L}_{N,M}(XY^\top).$$

Let $\prod_{(k)}^{(n,m)}$ be the vector space of k dimensional vectors with entries in $\prod^{n,m}(x,y)$. Utilizing the orthogonality relations (3.1) we see as in the next lemma.

LEMMA 3.1. *Suppose* $\mathbb{P} \in \prod_{(k)}^{(n,m)}$. *If* $\mathbb{P}$ *satisfies the orthogonality relations,*

$$(3.7) \qquad \langle \mathbb{P}, x^i y^j \rangle = 0, \quad 0 \le i < n, \quad 0 \le j \le m,$$

then $\mathbb{P} = C\mathbb{P}_{n,m}$, *where* C *is an* $k \times (m+1)$ *matrix. If* $k = m+1$, C *is lower triangular with positive diagonal entries, and* $\langle \mathbb{P}, \mathbb{P} \rangle = I_{m+1}$, *then* $C = I_{m+1}$.

Likewise we have the next lemma.

LEMMA 3.2. *Suppose* $\tilde{\mathbb{P}} \in \prod_{(k)}^{(n,m)}$. *If* $\tilde{\mathbb{P}}$ *satisfies the orthogonality relations,*

$$(3.8) \qquad \langle \tilde{\mathbb{P}}, x^i y^j \rangle = 0, \quad 0 \le i \le n, \quad 0 \le j < m,$$

then $\tilde{\mathbb{P}} = C\tilde{\mathbb{P}}_{n,m}$, *where* C *is an* $k \times (n+1)$ *matrix. If* $k = n+1$, C *is lower triangular with positive diagonal entries, and* $\langle \tilde{\mathbb{P}}, \tilde{\mathbb{P}} \rangle = I_{n+1}$, *then* $C = I_{n+1}$.

The discussion above allows us to make contact with the matrix orthogonal polynomials introduced in section 2.

LEMMA 3.3. *Let* $\mathbb{P}_{n,m}$ *be given by* (3.3). *Then*

$$(3.9) \qquad \mathbb{P}_{n,m} = L_n(x)[1, y, y^2, \dots, y^m]^\top$$

and

$$(3.10) \qquad \begin{bmatrix} \mathbb{P}_{0,m}(x,y) \\ \mathbb{P}_{1,m}(x,y) \\ \vdots \\ \mathbb{P}_{n,m}(x,y) \end{bmatrix} = \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_n(x) \end{bmatrix} [1, y, \dots, y^m]^\top = L \begin{bmatrix} I_{m+1} \\ xI_{m+1} \\ \vdots \\ x^n I_{m+1} \end{bmatrix} [1, y, \dots, y^m]^\top.$$

Proof. If we substitute the equation

$$\mathbb{P}_{n,m} = \hat{L}_n(x)[1 \ \cdots \ y^m]^\top = \sum_i \hat{L}_{n,i} x^i [1 \ \cdots \ y^m]^\top$$

into (3.7), where $\hat{L}_n(x)$ is some $(m+1) \times (m+1)$ matrix polynomial of degree n, we

find, for $j = 0, \ldots, n-1$,

$$
0 = \left\langle \mathbb{P}_{n,m}, x^j \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix} \right\rangle = \sum_{i=0}^{n} \hat{L}_{n,i} \left\langle x^i \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix}, x^j \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix} \right\rangle
$$

$$
= \sum_{i=1}^{n} \hat{L}_{n,i} \begin{bmatrix} \mathcal{L}_{NM}(x^{i+j}) & \cdots & \mathcal{L}_{NM}(x^{i+j}y^m) \\ \vdots & & \vdots \\ \mathcal{L}_{NM}(x^{i+j}y^m) & \cdots & \mathcal{L}_{NM}(x^{i+j}y^{2m}) \end{bmatrix}
$$

$$
= \sum_{i=1}^{n} \hat{L}_{n,i} \mathcal{L}_m(x^i, \ x^j) = \mathcal{L}_m(\hat{L}_n(x), \ x^j).
$$

Similarly,

$$
\langle \mathbb{P}_{n,m}, \mathbb{P}_{n,m} \rangle = I_{m+1} = \mathcal{L}_m \langle \hat{L}_n(x), \hat{L}_n(x) \rangle.
$$

This coupled with (2.8) and the fact that (3.3) implies that $\hat{L}_{n,m}$ is lower triangular with positive diagonal entries, gives the result.   □

As mentioned earlier, analogous formulas exist for orthogonal polynomials in the reverse lexicographical ordering with the roles of n and m interchanged.

THEOREM 3.4. *Given* $\{\mathbb{P}_{n,m}\}$ *and* $\{\tilde{\mathbb{P}}_{n,m}\}$, $0 \le n \le N$, $0 \le m \le M$, *the following recurrence formulas hold:*

$$
(3.11) \qquad x\mathbb{P}_{n,m} = A_{n+1,m}\mathbb{P}_{n+1,m} + B_{n,m}\mathbb{P}_{n,m} + A_{n,m}^{\top}\mathbb{P}_{n-1,m},
$$

$$
(3.12) \qquad \Gamma_{n,m}\mathbb{P}_{n,m} = \mathbb{P}_{n,m-1} - \mathcal{K}_{n,m}\tilde{\mathbb{P}}_{n-1,m},
$$

$$
(3.13) \qquad J_{n,m}^1\mathbb{P}_{n,m} = y\mathbb{P}_{n,m-1} + J_{n,m}^2\tilde{\mathbb{P}}_{n-1,m} + J_{n,m}^3\tilde{\mathbb{P}}_{n-1,m-1},
$$

$$
(3.14) \qquad \mathbb{P}_{n,m} = I_{n,m}\tilde{\mathbb{P}}_{n,m} + \Gamma_{n,m}^{\top}\mathbb{P}_{n,m-1},
$$

where

$$
(3.15) \qquad A_{n,m} = \langle x\mathbb{P}_{n-1,m}, \mathbb{P}_{n,m} \rangle \in M^{m+1,m+1},
$$

$$
(3.16) \qquad B_{n,m} = \langle x\mathbb{P}_{n,m}, \mathbb{P}_{n,m} \rangle \in M^{m+1,m+1},
$$

$$
(3.17) \qquad J_{n,m}^1 = \langle y\mathbb{P}_{n,m-1}, \mathbb{P}_{n,m} \rangle \in M^{m,m+1},
$$

$$
(3.18) \qquad J_{n,m}^2 = -\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m} \rangle \in M^{m,n},
$$

$$
(3.19) \qquad J_{n,m}^3 = -\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m-1} \rangle \in M^{m,n},
$$

$$
(3.20) \qquad \Gamma_{n,m} = \langle \mathbb{P}_{n,m-1}, \mathbb{P}_{n,m} \rangle \in M^{m,m+1},
$$

$$
(3.21) \qquad \mathcal{K}_{n,m} = \langle \mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m} \rangle \in M^{m,n},
$$

$$
(3.22) \qquad I_{n,m} = \langle \mathbb{P}_{n,m}, \tilde{\mathbb{P}}_{n,m} \rangle \in M^{m+1,n+1}.
$$

Similar formulas hold for $\tilde{\mathbb{P}}_{n,m}(x,y)$ *and will be denoted by* $(\tilde{3}.11)$, $(\tilde{3}.12)$, *etc.*

Proof. (3.11) follows from Lemma 3.3 and (2.14). To prove (3.12) note that because of the linear independence of the entries of $\mathbb{P}_{n,m}$, there is an $m \times (m+1)$ matrix $\Gamma_{n,m}$ such that $\Gamma_{n,m}\mathbb{P}_{n,m} - \mathbb{P}_{n,m-1} \in \prod_{(m)}^{(n-1,m)}(x,y)$. Furthermore,

$$
\langle \Gamma_{n,m}\mathbb{P}_{n,m} - \mathbb{P}_{n,m-1}, x^iy^j \rangle = 0, \qquad 0 \le i \le n-1 \quad 0 \le j \le m-1.
$$

Thus Lemma 3.2 implies that

$$\Gamma_{nm}\mathbb{P}_{n,m} - \mathbb{P}_{n,m-1} = \mathcal{K}_{n,m}\tilde{\mathbb{P}}_{n-1,m}.$$

The remaining recurrence formulas follow in a similar manner.   □

Remark 3.5. As indicated in the proof, formula (3.11) follows from the theory of matrix orthogonal polynomials and so allows us to move along a strip of size $m + 1$. This formula does not mix the polynomials in the two orderings. However, to increase m by one for polynomials constructed in the lexicographical ordering, the remaining relations show that orthogonal polynomials in the reverse lexicographical ordering must be used.

Remark 3.6. We saw in the previous section that $A_{n,m}$ is a lower triangular matrix with positive entries on the main diagonal, and $B_{n,m}$ is a symmetric matrix. From the orthogonality relations it follows immediately that $(J^1_{n,m})_{ij} = \langle yp^{i-1}_{n,m-1}, p^{j-1}_{n,m}\rangle = 0$ if $i + 1 < j$, and $(J^1_{n,m})_{i,i+1} > 0$. Thus $J^1_{n,m}$ has the form

$$J^1_{n,m} = \begin{bmatrix} (J^1_{n,m})_{1,1} & (J^1_{n,m})_{1,2} & 0 & 0 & 0 \\ (J^1_{n,m})_{2,1} & (J^1_{n,m})_{2,2} & (J^1_{n,m})_{2,3} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (J^1_{n,m})_{m,1} & (J^1_{n,m})_{m,2} & (J^1_{n,m})_{m,3} & \cdots & (J^1_{n,m})_{m,m+1} \end{bmatrix}.$$

Similarly, $(\Gamma_{n,m})_{ij} = \langle p^{i-1}_{n,m-1}, p^{j-1}_{n,m}\rangle = 0$ if $i < j$, and $(\Gamma_{n,m})_{i,i} > 0$, i.e., $\Gamma_{n,m}$ has the form

$$\Gamma_{n,m} = \begin{bmatrix} (\Gamma_{n,m})_{11} & 0 & \cdots & 0 & 0 \\ (\Gamma_{n,m})_{21} & (\Gamma_{n,m})_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\Gamma_{n,m})_{m1} & (\Gamma_{n,m})_{m2} & \cdots & (\Gamma_{n,m})_{mm} & 0 \end{bmatrix}.$$

Finally notice that $p^m_{n,m} = \tilde{p}^n_{n,m}$ and therefore $(I_{n,m})_{m+1,n+1} = 1$, $(I_{n,m})_{m+1,j} = 0$ for $j \le n$ and $(I_{n,m})_{i,n+1} = 0$ for $i \le m$, i.e.,

$$I_{n,m} = \begin{bmatrix} * & * & \cdots & * & 0 \\ \vdots & \vdots & & \vdots & 0 \\ * & * & \cdots & * & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Using the orthogonality relations and Theorem 3.4 one can verify the following.

PROPOSITION 3.7. *The following relations hold:*

$$\tag{3.23} \tilde{\mathcal{K}}_{n,m} = \mathcal{K}^\top_{n,m},$$

$$\tag{3.24} \mathcal{K}_{n,m}\mathcal{K}^\top_{n,m} + \Gamma_{n,m}\Gamma^\top_{n,m} = I_m,$$

$$\tag{3.25} J^3_{n,m} = -\mathcal{K}_{n,m}\tilde{A}^\top_{n-1,m},$$

$$\tag{3.26} \tilde{I}_{n,m} = I^\top_{n,m},$$

$$\tag{3.27} I_{n,m}I^\top_{n,m} + \Gamma^\top_{n,m}\Gamma_{n,m} = I_{m+1}.$$

Proof. We prove only (3.25) since the others are obvious from their defining relations. Beginning with (3.19) we find

$$J^3_{n,m} = -\langle\mathbb{P}_{n,m-1}, y\tilde{\mathbb{P}}_{n-1,m-1}\rangle.$$

The result now follows by using the counterpart of (3.11) for $\tilde{\mathbb{P}}$ and the defining equation for $\mathcal{K}_{n,m}$.   □

The function $\mathfrak{L}$ given by (2.18) can be used to show that $\mathbb{P}_{n,m}$ satisfies a certain minimization condition. Define $\hat{\mathfrak{L}} : \prod_{(m+1)}^{(n,m)} \to \text{sym } \mathbb{R}_{m+1}$ by

$$(3.28) \qquad \hat{\mathfrak{L}}(P) = \langle P, P \rangle - 2\text{sym } K_n,$$

where $P \in \prod_{(m+1)}^{(n,m)}$ is written as

$$(3.29) \qquad P(x,y) = K_n x^n \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix} + R(x,y)$$

with $R \in \prod_{(m+1)}^{(n-1,m)}$. We find,

LEMMA 3.8. *The polynomial $\hat{\mathbb{P}}_{n,m}(x,y) = L_{n,n}^{-1}\mathbb{P}_{n,m}(x,y)$ is the unique minimizer of $\hat{\mathfrak{L}}$ in the sense of (2.21).*

Proof. Write P as

$$P(x,y) = \sum_{i=0}^{n} K_i x^i \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix}.$$

Since

$$\left\langle x^i \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix}, x^j \begin{bmatrix} 1 \\ \vdots \\ y^m \end{bmatrix} \right\rangle = \begin{bmatrix} \mathcal{L}_{N,M}(x^i, x^j) & \cdots & \mathcal{L}_{N,M}(x^i, x^j y^m) \\ \vdots & & \vdots \\ \mathcal{L}_{N,M}(x^i y^m, x^j) & \cdots & \mathcal{L}_{N,M}(x^i y^m, x^j y^m) \end{bmatrix} = \mathcal{L}_m(x^i, x^j),$$

we see that $\hat{\mathfrak{L}}(P)$ can be written as $\hat{\mathfrak{L}}(P) = \mathfrak{L}(K)$, where $K(x) = K_n x^n + \cdots \in \prod_{m+1}^{n}$. The result now follows from (2.22) and Lemma 3.3.   □

4. Relations. As is evident from the previous section, there are relations between the various coefficients in (3.11)–(3.14) and their ($\tilde{3}$.11)–($\tilde{3}$.14) analogues. In this section we exhibit these relations.

LEMMA 4.1 (relations for $\mathcal{K}_{n,m}$).

$$(4.1) \qquad \Gamma_{n,m-1}\mathcal{K}_{n,m}\tilde{A}_{n-1,m}^{\top} = -J_{n,m-1}^2 - \mathcal{K}_{n,m-1}\tilde{B}_{n-1,m-1},$$

$$(4.2) \qquad A_{n,m-1}\mathcal{K}_{n,m}\tilde{\Gamma}_{n-1,m}^{\top} = -\tilde{J}_{n-1,m}^{2\top} - B_{n-1,m-1}\mathcal{K}_{n-1,m}.$$

Proof. We have

$$\mathcal{K}_{n,m}\tilde{A}_{n-1,m}^{\top} = \langle \mathbb{P}_{n,m-1}, \tilde{A}_{n-1,m}\tilde{\mathbb{P}}_{n-1,m} \rangle$$

$$= \langle \mathbb{P}_{n,m-1}, y\tilde{\mathbb{P}}_{n-1,m-1} - \tilde{B}_{n-1,m-1}\tilde{\mathbb{P}}_{n-1,m-1} - \tilde{A}_{n-1,m-1}^{\top}\tilde{\mathbb{P}}_{n-1,m-2} \rangle$$

$$= \langle \mathbb{P}_{n,m-1}, y\tilde{\mathbb{P}}_{n-1,m-1} \rangle.$$

Thus,

$$\Gamma_{n,m-1}\mathcal{K}_{n,m}\tilde{A}_{n-1,m}^{\top} = \langle \Gamma_{n,m-1}\mathbb{P}_{n,m-1}, y\tilde{\mathbb{P}}_{n-1,m-1} \rangle$$

$$= \langle \mathbb{P}_{n,m-2} - \mathcal{K}_{n,m-1}\tilde{\mathbb{P}}_{n-1,m-1}, y\tilde{\mathbb{P}}_{n-1,m-1} \rangle$$

$$= -J_{n,m-1}^2 - \mathcal{K}_{n,m-1}\tilde{B}_{n-1,m-1},$$

which completes the proof of (4.1). Writing (4.1) for $\tilde{\mathcal{K}}_{n,m}$ and using (3.23) we obtain (4.2).   □

LEMMA 4.2 (relations for $J^2_{n,m}$).

$$(4.3) \qquad \Gamma_{n,m-1}J^2_{n,m} = -J^1_{n,m-1}\mathcal{K}_{n,m} + \mathcal{K}_{n,m-1}\tilde{A}_{n-1,m},$$

$$-A_{n,m-1}J^2_{n,m}\tilde{\Gamma}^\top_{n-1,m} = J^1_{n-1,m}A^\top_{n-1,m}I_{n-2,m} - J^2_{n-1,m}\tilde{\Gamma}_{n-1,m}\tilde{J}^{1\,\top}_{n-1,m}$$

$$+ J^2_{n-1,m}\mathcal{K}^\top_{n-1,m}\tilde{J}^{2\,\top}_{n-1,m} + J^3_{n-1,m}I^\top_{n-2,m-1}\tilde{J}^{3\,\top}_{n-1,m}$$

$$+ B_{n-1,m-1}J^2_{n-1,m} - A^\top_{n-1,m-1}I_{n-2,m-1}\tilde{A}_{n-2,m}$$

$$(4.4) \qquad + A_{n,m-1}J^3_{n,m}I^\top_{n-1,m-1}\mathcal{K}_{n-1,m}.$$

Proof. The first equation can be derived by multiplying (3.18) on the left by $\Gamma_{n,m-1}$ then using (3̃.15) to obtain

$$\Gamma_{n,m-1}J^2_{n,m} = -\langle y\mathbb{P}_{n,m-2}, \tilde{\mathbb{P}}_{n-1,m}\rangle + \mathcal{K}_{n,m-1}\tilde{A}_{n-1,m}.$$

Eliminating $y\mathbb{P}_{n,m-2}$ using (3.13), then using the orthogonality of the polynomials and (3.19) yields (4.3).

To derive (4.4) notice that

$$(4.5) \qquad \begin{aligned} -A_{n,m-1}J^2_{n,m}\tilde{\Gamma}^\top_{n-1,m} &= A_{n,m-1}\langle y\mathbb{P}_{n,m-1}, \tilde{\Gamma}_{n-1,m}\tilde{\mathbb{P}}_{n-1,m}\rangle \\ &= A_{n,m-1}\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-2,m}\rangle \\ &\quad - A_{n,m-1}\langle y\mathbb{P}_{n,m-1}, \mathbb{P}_{n-1,m-1}\rangle\mathcal{K}_{n-1,m}. \end{aligned}$$

Using (3.11) in the first term on the right-hand side of (4.5) gives

$$(4.6) \qquad \begin{aligned} A_{n,m-1}\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-2,m}\rangle &= \langle x\mathbb{P}_{n-1,m-1}, y\tilde{\mathbb{P}}_{n-2,m}\rangle - B_{n-1,m-1}J^2_{n-1,m} \\ &\quad - A^\top_{n-1,m-1}\langle \mathbb{P}_{n-2,m-1}, y\tilde{\mathbb{P}}_{n-2,m}\rangle. \end{aligned}$$

Interchanging the positions of x and y in the first term on the right-hand side of (4.6), then using (3.13) and its (3̃.13) analogue yield

$$\begin{aligned} \langle x\mathbb{P}_{n-1,m-1}, y\tilde{\mathbb{P}}_{n-2,m}\rangle &= J^1_{n-1,m}(I_{n-1,m}\tilde{J}^{1\,\top}_{n-1,m} - \Gamma^\top_{n-1,m}\tilde{J}^{2\,\top}_{n-1,m}) \\ &\quad - J^2_{n-1,m}\tilde{\Gamma}_{n-1,m}\tilde{J}^{1\,\top}_{n-1,m} + J^2_{n-1,m}\mathcal{K}^\top_{n-1,m}\tilde{J}^{2\,\top}_{n-1,m} + J^3_{n-1,m}I^\top_{n-2,m-1}\tilde{J}^{3\,\top}_{n-1,m}, \end{aligned}$$

where (3.21), (3.22), and their (3̃.21), (3̃.22) analogues have been used.

Substituting (3.26) as well as the transpose of (4̃.18) into the above equation yields

$$(4.7) \qquad \begin{aligned} \langle x\mathbb{P}_{n-1,m-1}, y\tilde{\mathbb{P}}_{n-2,m}\rangle &= J^1_{n-1,m}A^\top_{n-1,m}I_{n-2,m} - J^2_{n-1,m}\tilde{\Gamma}_{n-1,m}\tilde{J}^{1\,\top}_{n-1,m} \\ &\quad + J^2_{n-1,m}\mathcal{K}^\top_{n-1,m}\tilde{J}^{2\,\top}_{n-1,m} + J^3_{n-1,m}I^\top_{n-2,m-1}\tilde{J}^{3\,\top}_{n-1,m}. \end{aligned}$$

The last term in (4.6) can be computed using (3̃.11) and (3.22), which gives

$$(4.8) \qquad \langle \mathbb{P}_{n-2,m-1}, y\tilde{\mathbb{P}}_{n-2,m}\rangle = I_{n-2,m-1}\tilde{A}_{n-2,m}.$$

Substituting (4.7) and (4.8) into (4.6) we see that the first term of (4.5) is

$$(4.9) \qquad \begin{aligned} A_{n,m-1}\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-2,m}\rangle &= J^1_{n-1,m}A^\top_{n-1,m}I_{n-2,m} - J^2_{n-1,m}\tilde{\Gamma}_{n-1,m}\tilde{J}^{1\,\top}_{n-1,m} \\ &\quad + J^2_{n-1,m}\mathcal{K}^\top_{n-1,m}\tilde{J}^{2\,\top}_{n-1,m} + J^3_{n-1,m}I^\top_{n-2,m-1}\tilde{J}^{3\,\top}_{n-1,m} \\ &\quad + B_{n-1,m-1}J^2_{n-1,m} - A^\top_{n-1,m-1}I_{n-2,m-1}\tilde{A}_{n-2,m}. \end{aligned}$$

Substituting (3.13) in the second term on the right-hand side of (4.5) and using the equations

$$\langle \mathbb{P}_{n,m-1}, \mathbb{P}_{n-1,m} \rangle = \mathcal{K}_{n,m} I_{n-1,m}^{\top} \tag{4.10}$$

and

$$\langle \mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-2,m} \rangle = \mathcal{K}_{n,m} \tilde{\Gamma}_{n-1,m}^{\top}, \tag{4.11}$$

which follow easily from (3.12), yields

$$
\begin{aligned}
A_{n,m-1} & \langle y\mathbb{P}_{n,m-1}, \mathbb{P}_{n-1,m-1} \rangle \mathcal{K}_{n-1,m} \\
&= A_{n,m-1}\mathcal{K}_{n,m} I_{n-1,m}^{\top} J_{n-1,m}^{1\ \top} \mathcal{K}_{n-1,m} - A_{n,m-1}\mathcal{K}_{n,m} \tilde{\Gamma}_{n-1,m}^{\top} J_{n-1,m}^{2\ \top} \mathcal{K}_{n-1,m} \\
&= A_{n,m-1}\mathcal{K}_{n,m} (I_{n-1,m}^{\top} J_{n-1,m}^{1\ \top} - \tilde{\Gamma}_{n-1,m}^{\top} J_{n-1,m}^{2\ \top}) \mathcal{K}_{n-1,m} \\
&= -A_{n,m-1} J_{n,m}^{3} I_{n-1,m-1}^{\top} \mathcal{K}_{n-1,m}.
\end{aligned}
\tag{4.12}
$$

In the last equality we used (3.25) and (4.18). Finally, combining (4.9) and (4.12) we obtain (4.4).   □

LEMMA 4.3 (relations for $J_{n,m}^{1}$).

$$\Gamma_{n,m-1} J_{n,m}^{1} = J_{n,m-1}^{1} \Gamma_{n,m}, \tag{4.13}$$

$$J_{n,m}^{1} \Gamma_{n,m}^{\top} \Gamma_{n,m-1}^{\top} = J_{n,m-1}^{1\ \top} + J_{n,m}^{3} \mathcal{K}_{n,m-1}^{\top} + J_{n,m}^{2} \mathcal{K}_{n,m}^{\top} \Gamma_{n,m-1}^{\top}. \tag{4.14}$$

Proof. (4.13) can be derived by multiplying (3.17) by $\Gamma_{n,m-1}$ then using (3.13). For the second equality we multiply (3.17) by $\Gamma_{n,m}^{\top}$ then use (3.12) to obtain

$$J_{n,m}^{1} \Gamma_{n,m}^{\top} = \langle y\mathbb{P}_{n,m-1}, \Gamma_{n,m}\mathbb{P}_{n,m} \rangle = \langle y\mathbb{P}_{n,m-1}, \mathbb{P}_{n,m-1} \rangle + J_{n,m}^{2} \mathcal{K}_{n,m}^{\top}.$$

Multiplying on the right of the above formula by $\Gamma_{n,m-1}^{\top}$, then using (3.12) followed by (3.13) twice, leads to the result.   □

LEMMA 4.4 (relations for $A_{n,m}$).

$$\Gamma_{n-1,m} A_{n,m} = A_{n,m-1} \Gamma_{n,m}, \tag{4.15}$$

$$J_{n-1,m}^{1} A_{n,m} = A_{n,m-1} J_{n,m}^{1}. \tag{4.16}$$

Proof. First we compute

$$\langle \mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n,m} \rangle = \langle \mathbb{P}_{n-1,m-1}, \mathbb{P}_{n-1,m} \rangle A_{n,m} = \Gamma_{n-1,m} A_{n,m}.$$

On the other hand,

$$\langle \mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n,m} \rangle = A_{n,m-1} \Gamma_{n,m}.$$

This gives (4.15). If we use (3.17) with n changed to $n-1$, then multiply on the right by $A_{n,m}$ and use (3.11), we find

$$J_{n-1,m}^{1} A_{n,m} = \langle x\mathbb{P}_{n-1,m-1}, y\mathbb{P}_{n,m} \rangle.$$

Now eliminating $x\mathbb{P}_{n-1,m-1}$ using (3.11), then applying (3.17), yields (4.16).   □

LEMMA 4.5 (relations for $I_{n,m}$).

$$(4.17) \qquad \Gamma_{n,m} I_{n,m} = -\mathcal{K}_{n,m} \tilde{\Gamma}_{n,m},$$

$$(4.18) \qquad J^1_{n,m} I_{n,m} = I_{n,m-1} \tilde{A}_{n,m} + J^2_{n,m} \tilde{\Gamma}_{n,m}.$$

Proof. The first relation follows in a straightforward manner by multiplying (3.22) on the left by $\Gamma_{n,m}$ and then using (3.12).

For (4.18) we first compute

$$I_{n,m-1} \tilde{A}_{n,m} = \langle \mathbb{P}_{n,m-1}, y\tilde{\mathbb{P}}_{n,m} \rangle,$$

which follows by using ($\tilde{3}$.11) to eliminate $y\tilde{\mathbb{P}}_{n,m}$. Next, in the above equation use (3.17) to obtain

$$\langle y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n,m} \rangle = J^1_{n,m} I_{n,m} - J^2_{n,m} \tilde{\Gamma}_{n,m},$$

which gives (4.18).   □

LEMMA 4.6 (relations for $B_{n-1,m}$).

$$\begin{aligned}
\Gamma_{n-1,m} B_{n-1,m} = &-\mathcal{K}_{n-1,m} I^\top_{n-2,m} A_{n-1,m} + B_{n-1,m-1} \Gamma_{n-1,m} \\
(4.19) \qquad & + A_{n,m-1} \mathcal{K}_{n,m} I^\top_{n-1,m},
\end{aligned}$$

$$\begin{aligned}
J^1_{n-1,m} B_{n-1,m} = &B_{n-1,m-1} J^1_{n-1,m} + J^2_{n-1,m} I^\top_{n-2,m} A_{n-1,m} \\
& + J^3_{n-1,m} I^\top_{n-2,m-1} A_{n-1,m-1} \Gamma_{n-1,m} - A_{n,m-1} J^2_{n,m} I^\top_{n-1,m} \\
(4.20) \qquad & - A_{n,m-1} J^3_{n,m} I^\top_{n-1,m-1} \Gamma_{n-1,m}.
\end{aligned}$$

Proof. We begin by multiplying (3.20) on the left by $B_{n-1,m}$ and then using (3.11) to obtain

$$\Gamma_{n-1,m} B_{n-1,m} = \langle \mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n-1,m} \rangle - \langle \mathbb{P}_{n-1,m-1}, \mathbb{P}_{n-2,m} \rangle A_{n-1,m}.$$

We see from (4.10) that the second term on the right-hand side of the above formula gives the first term on the right-hand side in (4.19). We can compute the first term on the right-hand side of the above formula by eliminating $x\mathbb{P}_{n-1,m}$ using (3.11) to find

$$\begin{aligned}
\langle \mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n-1,m} \rangle &= A_{n,m-1} \langle \mathbb{P}_{n,m-1}, \mathbb{P}_{n-1,m} \rangle + B_{n-1,m-1} \langle \mathbb{P}_{n-1,m-1}, \mathbb{P}_{n-1,m} \rangle \\
&= A_{n,m-1} \mathcal{K}_{n,m} I^\top_{n-1,m} + B_{n-1,m-1} \Gamma_{n-1,m},
\end{aligned}$$

where in the last equality we used again (4.10). This completes the proof of (4.19). Relation (4.20) can be derived as follows. First we multiply (3.16) with n reduced by one on the left by $J^1_{n-1,m}$ to obtain

$$\begin{aligned}
(4.21) \qquad J^1_{n-1,m} B_{n-1,m} = &\langle y\mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n-1,m} \rangle + J^2_{n-1,m} \langle \tilde{\mathbb{P}}_{n-2,m}, x\mathbb{P}_{n-1,m} \rangle \\
& + J^3_{n-1,m} \langle \tilde{\mathbb{P}}_{n-2,m-1}, xP_{n-1,m} \rangle.
\end{aligned}$$

Next we compute the three terms on the right-hand side of the above formula. For the first term we eliminate $x\mathbb{P}_{n-1,m-1}$ using (3.11) to find

$$\langle y\mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n-1,m} \rangle = A_{n,m-1} \langle \mathbb{P}_{n,m-1}, y\mathbb{P}_{n-1,m} \rangle + B_{n-1,m-1} J^1_{n-1,m}.$$

The term $\langle \mathbb{P}_{n,m-1}, y\mathbb{P}_{n-1,m}\rangle$ can be computed using (3.13) so that

$$\langle y\mathbb{P}_{n-1,m-1}, x\mathbb{P}_{n-1,m}\rangle = - A_{n,m-1}J_{n,m}^2 I_{n-1,m}^\top - A_{n,m-1}J_{n,m}^3 \tilde{I}_{n-1,m-1}\Gamma_{n-1,m}$$

(4.22)
$$+ B_{n-1,m-1}J_{n-1,m}^1.$$

Next we compute the second term in (4.21) using (3.13) to obtain

$$\langle \tilde{\mathbb{P}}_{n-2,m}, x\mathbb{P}_{n-1,m}\rangle = \tilde{J}_{n-1,m}^1 I_{n-1,m}^\top - \tilde{J}_{n-1,m}^2 \Gamma_{n-1,m}.$$

Using (4.18) for $\tilde{I}_{n-1,m}$ and (3.26) we obtain the following formula for the second term

(4.23) $$J_{n-1,m}^2\langle \tilde{\mathbb{P}}_{n-2,m}, x\mathbb{P}_{n-1,m}\rangle = J_{n-1,m}^2 I_{n-2,m}^\top A_{n-1,m}.$$

Finally, the third term in (4.21) can be computed with the help of (3.13),

$$\langle \tilde{\mathbb{P}}_{n-2,m-1}, x\mathbb{P}_{n-1,m}\rangle = \tilde{J}_{n-1,m-1}^1\langle \tilde{\mathbb{P}}_{n-1,m-1}, \mathbb{P}_{n-1,m}\rangle - \tilde{J}_{n-1,m-1}^2\langle \mathbb{P}_{n-1,m-2}, \mathbb{P}_{n-1,m}\rangle.$$

Notice that

$$\langle \tilde{\mathbb{P}}_{n-1,m-1}, \mathbb{P}_{n-1,m}\rangle = \tilde{I}_{n-1,m-1}\Gamma_{n-1,m},$$

and using (3.14) for $\mathbb{P}_{n-1,m}$, we get

$$\langle \mathbb{P}_{n-1,m-2}, \mathbb{P}_{n-1,m}\rangle = \Gamma_{n-1,m-1}\Gamma_{n-1,m}.$$

Thus we have

$$J_{n-1,m}^3\langle \tilde{\mathbb{P}}_{n-2,m-1}, x\mathbb{P}_{n-1,m}\rangle = J_{n-1,m}^3 \tilde{I}_{n-2,m-1}A_{n-1,m-1}\Gamma_{n-1,m},$$

where in the last equality we used again (4.18). Combining the above equation, (4.22), and (4.23), we obtain (4.20).   □

LEMMA 4.7 (relations for $\tilde{J}_{n,m}^1$).

(4.24)
$$\tilde{J}_{n,m}^1 = I_{n-1,m}^\top A_{n,m}I_{n,m} + I_{n-1,m}^\top B_{n-1,m}I_{n-1,m}\tilde{\Gamma}_{n,m}$$
$$+ I_{n-1,m}^\top A_{n-1,m}^\top I_{n-2,m}\tilde{\Gamma}_{n-1,m}\tilde{\Gamma}_{n,m} + \tilde{\Gamma}_{n-1,m}^\top \tilde{J}_{n-1,m}^1\tilde{\Gamma}_{n,m}.$$

Proof. We begin by eliminating $\tilde{\mathbb{P}}_{n-1,m}$ and $\tilde{\mathbb{P}}_{n,m}$ in (3.17) using (3.14) and (3.13),

$$\tilde{J}_{n,m}^1 = I_{n-1,m}^\top\langle x\mathbb{P}_{n-1,m}, I_{n,m}^\top\mathbb{P}_{n,m} + \tilde{\Gamma}_{n,m}^\top\tilde{\mathbb{P}}_{n-1,m}\rangle$$
$$+ \tilde{\Gamma}_{n-1,m}^\top\langle \tilde{J}_{n-1,m}^1\tilde{\mathbb{P}}_{n-1,m} - \tilde{J}_{n-1,m}^2\mathbb{P}_{n-1,m-1} - \tilde{J}_{n-1,m}^3\mathbb{P}_{n-2,m-1}, \tilde{\mathbb{P}}_{n,m}\rangle,$$

which simplifies to

$$\tilde{J}_{n,m}^1 = I_{n-1,m}^\top A_{n,m}I_{n,m} + I_{n-1,m}^\top\langle x\mathbb{P}_{n-1,m}, \tilde{\mathbb{P}}_{n-1,m}\rangle\tilde{\Gamma}_{n,m} + \tilde{\Gamma}_{n-1,m}^\top\tilde{J}_{n-1,m}^1\tilde{\Gamma}_{n,m}.$$

In the above equation, (3.15) and (3.20) have been used. Next use (3.11) to eliminate $x\mathbb{P}_{n-1,m}$. All terms can be evaluated using (3.15)–(3.22) except the term $\langle \mathbb{P}_{n-2,m}, \tilde{\mathbb{P}}_{n-1,m}\rangle$, which can be evaluated by applying (3.14) to $\tilde{\mathbb{P}}_{n-1,m}$.   □

Similar arguments show the next lemma.

LEMMA 4.8 (relations for $\tilde{J}_{n,m}^2$).

$$(4.25) \quad \begin{aligned} \tilde{J}_{n,m}^2 = &- (I_{n-1,m}^\top A_{n,m} \Gamma_{n,m}^\top + I_{n-1,m}^\top B_{n-1,m} I_{n-1,m} \mathcal{K}_{n,m}^\top \\ &+ I_{n-1,m}^\top A_{n-1,m}^\top I_{n-2,m} \tilde{\Gamma}_{n-1,m} \mathcal{K}_{n,m}^\top + \tilde{\Gamma}_{n-1,m}^\top \tilde{J}_{n-1,m}^1 \mathcal{K}_{n,m}^\top). \end{aligned}$$

LEMMA 4.9 (relations for $\tilde{A}_{n,m}$).

$$(4.26) \quad \tilde{A}_{n,m} = I_{n,m-1}^\top J_{n,m}^1 I_{n,m} - I_{n,m-1}^\top J_{n,m}^2 \tilde{\Gamma}_{n,m} + \tilde{\Gamma}_{n,m-1}^\top \tilde{A}_{n-1,m} \tilde{\Gamma}_{n,m}.$$

Proof. From $(\tilde{3}.15)$ we obtain

$$\tilde{A}_{n,m} = I_{n,m-1}^\top \langle y \mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n,m} \rangle + \tilde{\Gamma}_{n,m-1}^\top \langle y \tilde{\mathbb{P}}_{n-1,m-1}, \tilde{\mathbb{P}}_{n,m} \rangle.$$

Eliminate $y \mathbb{P}_{n,m-1}$ using (3.13) and $y \tilde{\mathbb{P}}_{n-1,m-1}$ using $(\tilde{3}.11)$ leads to the result.   □

Again in an analogous fashion, we have the next lemma.

LEMMA 4.10 (relations for $\tilde{B}_{n,m}$).

$$(4.27)$$
$$\begin{aligned} \tilde{B}_{n,m-1} = &\, I_{n,m-1}^\top \Gamma_{n,m} J_{n,m}^{1\,\top} I_{n,m-1} - I_{n,m-1}^\top \mathcal{K}_{n,m} J_{n,m}^{2\,\top} I_{n,m-1} \\ &- \tilde{\Gamma}_{n,m-1}^\top J_{n,m}^{3\,\top} I_{n,m-1} - I_{n,m-1}^\top J_{n,m}^3 \tilde{\Gamma}_{n,m-1} + \tilde{\Gamma}_{n,m-1}^\top \tilde{B}_{n-1,m-1} \tilde{\Gamma}_{n,m-1}. \end{aligned}$$

5. Christoffel–Darboux-like formulas. It is well known that the Christoffel–Darboux formula plays an important role in the theory of orthogonal polynomials of one variable. Using the connection between matrix orthogonal polynomials and two variable orthogonal polynomials developed in section 3 we will present two variable analogues of this celebrated formula.

THEOREM 5.1 (Christoffel–Darboux formula).

$$\frac{\mathbb{P}_{n,m}^\top(x_1, y_1) A_{n+1,m} \mathbb{P}_{n+1,m}(x, y) - \mathbb{P}_{n+1,m}^\top(x_1, y_1) A_{n+1,m}^\top \mathbb{P}_{n,m}(x, y)}{x - x_1}$$
$$= \sum_{k=0}^n \mathbb{P}_{k,m}^\top(x_1, y_1) \mathbb{P}_{k,m}(x, y)$$
$$= \sum_{j=0}^m \tilde{\mathbb{P}}_{n,j}^\top(x_1, y_1) \tilde{\mathbb{P}}_{n,j}(x, y).$$

Proof. The first equality follows from (3.11) and standard manipulations. The second equality follows since both sums are reproducing kernels for the same space.   □

An analogous result holds for the reverse lexicographical ordering. The above Theorem also implies the next lemma.

LEMMA 5.2.

$$(5.1)$$

$$\begin{aligned} &\mathbb{P}_{n,m}^\top(x_1, y_1) A_{n+1,m} \mathbb{P}_{n+1,m}(x, y) - \mathbb{P}_{n+1,m}^\top(x_1, y_1) A_{n+1,m}^\top \mathbb{P}_{n,m}(x, y) \\ &= (x - x_1) \tilde{\mathbb{P}}_{n,m}^\top(x_1, y_1) \tilde{\mathbb{P}}_{n,m}(x, y) \\ &+ \mathbb{P}_{n,m-1}^\top(x_1, y_1) A_{n+1,m-1} \mathbb{P}_{n+1,m-1}(x, y) - \mathbb{P}_{n+1,m-1}^\top(x_1, y_1) A_{n+1,m-1}^\top \mathbb{P}_{n,m-1}(x, y), \end{aligned}$$

and

$$\mathbb{P}_{n+1,m+1}^{\top}(x_1,y_1)\mathbb{P}_{n+1,m+1}(x,y) - \mathbb{P}_{n+1,m}^{\top}(x_1,y_1)\mathbb{P}_{n+1,m}(x,y)$$

$$(5.2)\qquad = \tilde{\mathbb{P}}_{n+1,m+1}^{\top}(x_1,y_1)\tilde{\mathbb{P}}_{n+1,m+1}(x,y) - \tilde{\mathbb{P}}_{n,m+1}^{\top}(x_1,y_1)\tilde{\mathbb{P}}_{n,m+1}(x,y).$$

Proof. To prove the first formula let

$$Z_{n,m}(x,y) = [1,y,\ldots,y^m][I_{m+1},xI_{m+1},\ldots,x^nI_{m+1}],$$

and let $\tilde{Z}_{n,m}(x,y)$ be given by a similar formula with the roles of x and y, and n and m interchanged. Then from the Christoffel–Darboux formula, Lemma 2.5, and (3.10) we find

$$\frac{\mathbb{P}_{n,m}^{\top}(x_1,y_1)A_{n+1,m}\mathbb{P}_{n+1,m}(x,y) - \mathbb{P}_{n+1,m}^{\top}(x_1,y_1)A_{n+1,m}^{\top}\mathbb{P}_{n,m}(x,y)}{x-x_1}$$

$$= Z_{n,m}(x_1,y_1)H_{n,m}^{-1}Z_{n,m}(x,y)^{\top} = \tilde{Z}_{n,m}(x_1,y_1)\tilde{H}_{n,m}^{-1}\tilde{Z}_{n,m}(x,y)^{\top}$$

$$= \tilde{\mathbb{P}}_{n,m}^{\top}(x_1,y_1)\tilde{\mathbb{P}}_{n,m}(x,y) + \tilde{Z}_{n,m-1}(x_1,y_1)\tilde{H}_{n,m-1}^{-1}\tilde{Z}_{n,m-1}(x,y)^{\top}.$$

Switching back to the lexicographical ordering in the second term in the last equation implies the result. (5.2) can be obtained by using the equality of the sums in Theorem 5.1 to find

$$\mathbb{P}_{n+1,m+1}^{\top}(x_1,y_1)\mathbb{P}_{n+1,m+1}(x,y) - \sum_{j=0}^{m}\tilde{\mathbb{P}}_{n+1,j}^{\top}(x_1,y_1)\tilde{\mathbb{P}}_{n+1,j}(x,y)$$

$$= \tilde{\mathbb{P}}_{n+1,m+1}^{\top}(x_1,y_1)\tilde{\mathbb{P}}_{n+1,m+1}(x,y) - \sum_{j=0}^{n}\mathbb{P}_{j,m+1}^{\top}(x_1,y_1)\mathbb{P}_{j,m+1}(x,y).$$

Switching to the lexicographical ordering in the sum on the left-hand side of the above equation and reverse lexicographical ordering in the sum on the right-hand side then extracting out the highest terms and using Theorem 5.1 gives the result. □

Remark 5.3. The above equations can be derived from the recurrence formulas in the previous sections. (5.2) follows easily from (3.12) and Proposition 3.7. However, the derivation of (5.1) is rather tedious.

6. Algorithm. With the use of the relations derived in the previous section we develop an algorithm that allows us to compute the coefficients in the recurrence formulas at higher levels in terms of those at lower levels plus some indeterminates that are equivalent to the moments (see Theorem 7.1). More precisely, at each level (n,m) we construct the matrices $\mathcal{K}_{n,m}$, $\Gamma_{n,m}$, $J_{n,m}^2$, $J_{n,m}^1$, $A_{n,m}$, $I_{n,m}$, $B_{n-1,m}$, $\tilde{A}_{n,m}$, $\tilde{J}_{n,m}^1$, $\tilde{J}_{n,m}^2$, $\tilde{B}_{n,m-1}$ and the polynomials $\mathbb{P}_{n,m}(x,y)$ and $\tilde{\mathbb{P}}_{n,m}(x,y)$ recursively, using the matrices at levels $(n-1,m)$ and $(n,m-1)$. In order to construct the above matrices we will have need of the $m \times (m+1)$ matrix U_m given by

$$U_m = [I_m|0] = [\delta_{i,j}] = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

and the $m \times m$ elementary matrix $E_{m,m}$ having just one nonzero entry at (m,m). The matrix norm used in that and the remaining sections is the l_2 norm.

At level $(0,0)$ we have just one free parameter (corresponding to $h_{0,0} = \langle 1 \rangle$), at level $(n,0)$ (resp., $(0,m)$) we have two new parameters corresponding to $h_{2n-1,0} = \langle x^{2n-1} \rangle$ and $h_{2n,0} = \langle x^{2n} \rangle$ (resp., $h_{0,2m-1} = \langle y^{2m-1} \rangle$, and $h_{0,2m} = \langle y^{2m} \rangle$), and, if $n > 0$ and $m > 0$, we have four new parameters corresponding to the moments $h_{2n-1,2m-1} = \langle x^{2n-1} y^{2m-1} \rangle$, $h_{2n-1,2m} = \langle x^{2n-1} y^{2m} \rangle$, $h_{2n,2m-1} = \langle x^{2n} y^{2m-1} \rangle$, $h_{2n,2m} = \langle x^{2n} y^{2m} \rangle$.

Level(0,0). When $n = m = 0$ we simply put

(6.1)
$$\mathbb{P}_{0,0}(x,y) = \tilde{\mathbb{P}}_{0,0}(x,y) = s_{0,0},$$

where $s_{0,0}$ is the new parameter corresponding to the moment $\langle 1 \rangle$.

Level(n,0). When $m = 0$, $\mathbb{P}_{n,0} = (p_{n,0}^0)$ is just a scalar valued function in x and clearly

(6.2)
$$\tilde{\mathbb{P}}_{n,0} = \begin{bmatrix} p_{0,0}^0 \\ p_{1,0}^0 \\ \vdots \\ p_{n,0}^0 \end{bmatrix}, \text{ i.e., } \tilde{p}_{n,0}^k = p_{k,0}^0.$$

Thus $I_{n,0} = (0,0,\ldots,0,1)$ $(1 \times (n+1)$ matrix$)$.

To construct all other matrices at level $(n,0)$ from $(n-1,0)$ we have two new parameters, $s_{2n-1,0}$ and $s_{2n,0}$, corresponding to the moments $\langle x^{2n-1} \rangle$ and $\langle x^{2n} \rangle$.

We take $A_{n,0} = s_{2n,0} > 0$ and $B_{n-1,0} = s_{2n-1,0}$. Then,

(6.3)
$$p_{n,0}^0 = A_{n,0}^{-1}(x p_{n-1,0}^0 - B_{n-1,0} p_{n-1,0}^0 - A_{n-1,0} p_{n-2,0}^0),$$

and $\tilde{J}_{n,0}^1$ is the tridiagonal matrix

$$\tilde{J}_{n,0}^1 = \begin{bmatrix} B_{0,0} & A_{1,0} & & & & \\ A_{1,0} & B_{1,0} & A_{2,0} & & & \\ & A_{2,0} & B_{2,0} & A_{3,0} & & \\ & & \ddots & \ddots & \ddots & \\ & & & A_{n-1,0} & B_{n-1,0} & A_{n,0} \end{bmatrix}.$$

Level(0,m). In this case $\tilde{\mathbb{P}}_{0,m} = (\tilde{p}_{0,m}^0)$ is a scalar function in y and

(6.4)
$$\mathbb{P}_{0,m} = \begin{bmatrix} \tilde{p}_{0,0}^0 \\ \tilde{p}_{0,1}^0 \\ \vdots \\ \tilde{p}_{0,m}^0 \end{bmatrix}, \text{ i.e., } p_{0,m}^k = \tilde{p}_{0,k}^0.$$

We have two new parameters, $s_{0,2m-1}$ and $s_{0,2m}$, corresponding to the moments $\langle y^{2m-1} \rangle$ and $\langle y^{2m} \rangle$.

Clearly $\Gamma_{0,m} = U_m$, $I_{0,m}$ is the $(m+1) \times 1$ matrix $(0,0,\ldots,0,1)^\top$.

We take $\tilde{A}_{0,m} = s_{0,2m} > 0$ and $\tilde{B}_{0,m-1} = s_{2m-1,0}$. Then

(6.5)
$$\tilde{p}_{0,m}^0 = \tilde{A}_{0,m}^{-1}(y \tilde{p}_{0,m-1}^0 - \tilde{B}_{0,m-1} \tilde{p}_{0,m-1}^0 - \tilde{A}_{0,m-1} \tilde{p}_{0,m-2}^0)$$

and

$$
J_{0,m}^1 =
\begin{bmatrix}
\tilde{B}_{0,0} & \tilde{A}_{0,1} & & & & \\
\tilde{A}_{0,1} & \tilde{B}_{0,1} & \tilde{A}_{0,2} & & & \\
& \tilde{A}_{0,2} & \tilde{B}_{0,2} & \tilde{A}_{0,3} & & \\
& & \ddots & \ddots & \ddots & \\
& & & \tilde{A}_{0,m-1} & \tilde{B}_{0,m-1} & \tilde{A}_{0,m}
\end{bmatrix}.
$$

Level(n,m). If $n \geq 1$ and $m \geq 1$ we have four new parameters $s_{2n-1,2m-1}$, $s_{2n-1,2m}$, $s_{2n,2m-1}$, and $s_{2n,2m}$, corresponding to the moments $\langle x^{2n-1}y^{2m-1}\rangle$, $\langle x^{2n-1}y^{2m}\rangle$, $\langle x^{2n}y^{2m-1}\rangle$, and $\langle x^{2n}y^{2m}\rangle$.

Computation of $\mathcal{K}_{n,m}$. In $\mathcal{K}_{n,m}$ the only new entry is $(\mathcal{K}_{n,m})_{m,n}$. (Everything else can be recovered from the previous levels.) Indeed, if $m > 1$ we show below that all rows except the last one can be obtained from (4.1). Notice that $\Gamma_{n,m-1}E_{m,m} = 0$ and therefore

$$
\Gamma_{n,m-1} = \Gamma_{n,m-1}(U_{m-1}^\top U_{m-1} + E_{m,m}) = \Gamma_{n,m-1}U_{m-1}^\top U_{m-1}.
$$

But we know that $\Gamma_{n,m-1}U_{m-1}^\top$ is an invertible matrix, which allows us to rewrite (4.1) as follows:

$$
U_{m-1}\mathcal{K}_{n,m} = -(\Gamma_{n,m-1}U_{m-1}^\top)^{-1}(J_{n,m-1}^2 + \mathcal{K}_{n,m-1}\tilde{B}_{n-1,m-1})\tilde{A}_{n-1,m}^{\top\,-1}.
$$

The matrix $U_{m-1}\mathcal{K}_{n,m}$ is the $(m-1) \times n$ matrix obtained from $\mathcal{K}_{n,m}$ by deleting the last row.

Similarly, if $n > 1$ we can write

$$
\tilde{\Gamma}_{n-1,m} = \tilde{\Gamma}_{n-1,m}(U_{n-1}^\top U_{n-1} + E_{n,n}) = \tilde{\Gamma}_{n-1,m}U_{n-1}^\top U_{n-1},
$$

i.e., $\tilde{\Gamma}_{n-1,m}^\top = U_{n-1}^\top(U_{n-1}\tilde{\Gamma}_{n-1,m}^\top)$, and (4.2) can be rewritten as

$$
\mathcal{K}_{n,m}U_{n-1}^\top = -A_{n,m-1}^{-1}(\tilde{J}_{n-1,m}^{2\,\top} + B_{n-1,m-1}\mathcal{K}_{n-1,m})(U_{n-1}\tilde{\Gamma}_{n-1,m}^\top)^{-1}.
$$

Thus the $m \times (n-1)$ matrix $\mathcal{K}_{n,m}U_{n-1}^\top$, which is obtained from $\mathcal{K}_{n,m}$ by deleting the last column, is known from the previous levels. This allows us to compute all entries in the last row of $\mathcal{K}_{n,m}$ except $(\mathcal{K}_{n,m})_{m,n}$. Finally we put $(\mathcal{K}_{n,m})_{m,n} = s_{2n-1,2m-1}$. From the computation of $\Gamma_{n,m}$ below we see that the parameters must be chosen so that $I - \mathcal{K}_{n,m}\mathcal{K}_{n,m}^\top$ is positive definite.

Computation of $\Gamma_{n,m}$. If $||\mathcal{K}_{n,m}|| < 1$, the matrix $I - \mathcal{K}_{n,m}\mathcal{K}_{n,m}^\top$ is symmetric and positive definite. Rewriting (3.24) as

$$
I - \mathcal{K}_{n,m}\mathcal{K}_{n,m}^\top = (\Gamma_{n,m}U_m^\top)(\Gamma_{n,m}U_m^\top)^\top,
$$

we see that $\Gamma_{n,m}U_m^\top$ (which is $\Gamma_{n,m}$ except the last zero column) is the lower-triangular factor in the Cholesky factorization of the matrix $I - \mathcal{K}_{n,m}\mathcal{K}_{n,m}^\top$.

Computation of $J_{n,m}^2$. The computation of $J_{n,m}^2$ is similar to the computation of $\mathcal{K}_{n,m}$. First, if $m > 1$ we can write (4.3) in the form

$$
U_{m-1}J_{n,m}^2 = (\Gamma_{n,m-1}U_{m-1}^\top)^{-1}(-J_{n,m-1}^1\mathcal{K}_{n,m} + \mathcal{K}_{n,m-1}\tilde{A}_{n-1,m}),
$$

which gives all entries of $J_{n,m}^2$ except the entries in the last row. Rewriting equation (4.4) as

$$
\begin{aligned}
J_{n,m}^2 U_{n-1}^\top = & - A_{n,m-1}^{-1}(J_{n-1,m}^1 A_{n-1,m}^\top I_{n-2,m} - J_{n-1,m}^2 \tilde{\Gamma}_{n-1,m} \tilde{J}_{n-1,m}^{1\,\top} \\
& + J_{n-1,m}^2 \mathcal{K}_{n-1,m}^\top \tilde{J}_{n-1,m}^{2\,\top} + J_{n-1,m}^3 I_{n-2,m-1}^\top \tilde{J}_{n-1,m}^{3\,\top} \\
& + B_{n-1,m-1} J_{n-1,m}^2 - A_{n-1,m-1}^\top I_{n-2,m-1} \tilde{A}_{n-2,m} \\
& + A_{n,m-1} J_{n,m}^3 I_{n-1,m-1}^\top \mathcal{K}_{n-1,m})(U_{n-1}\tilde{\Gamma}_{n-1,m}^\top)^{-1}
\end{aligned}
$$

allows us to compute everything except the last column of $J_{n,m}^2$.

Finally, we put $(J_{n,m}^2)_{m,n} = s_{2n-1,2m}$.

Computation of $J_{n,m}^1$. We put $(J_{n,m}^1)_{m,m} = s_{2n,2m-1}$ and $(J_{n,m}^1)_{m,m+1} = s_{2n,2m}$ using the last two parameters. Everything else can be recovered from the previous levels. Rewriting (4.13) as

$$
U_{m-1}J_{n,m}^1 = (\Gamma_{n,m-1}U_{m-1}^\top)^{-1}J_{n,m-1}^1 \Gamma_{n,m},
$$

we get the matrix obtained from $J_{n,m}^1$ by deleting the last row.

Consider now the $(m+1) \times (m-1)$ matrix $\Gamma_{n,m}^\top \Gamma_{n,m-1}^\top$. It is easy to see that the last two rows of this matrix are zeros and deleting these two rows we obtain an $(m-1) \times (m-1)$ upper-triangular matrix with positive entries on the main diagonal. Therefore the matrix $U_{m-1}U_m\Gamma_{n,m}^\top \Gamma_{n,m-1}^\top$ is invertible. Since $U_m^\top U_{m-1}^\top U_{m-1}U_m = I_{m+1} - E_{m,m} - E_{m+1,m+1}$, we can write

$$
\Gamma_{n,m}^\top \Gamma_{n,m-1}^\top = U_m^\top U_{m-1}^\top(U_{m-1}U_m\Gamma_{n,m}^\top \Gamma_{n,m-1}^\top).
$$

Combining this with formula (4.14) we see that

$$
J_{n,m}^1 U_m^\top U_{m-1}^\top = (J_{n,m-1}^{1\,\top} + J_{n,m}^3 \mathcal{K}_{n,m-1}^\top + J_{n,m}^2 \mathcal{K}_{n,m}^\top \Gamma_{n,m-1}^\top)(U_{m-1}U_m\Gamma_{n,m}^\top \Gamma_{n,m-1}^\top)^{-1}.
$$

The matrix $J_{n,m}^1 U_m^\top U_{m-1}^\top$ is obtained from $J_{n,m}^1$ by deleting the last two columns. This completes the computation of $J_{n,m}^1$.

Computation of $A_{n,m}$. Let us denote by $\mathcal{M}_{n-1,m}$ the $(m+1) \times (m+1)$ matrix obtained by adding the last row of $J_{n-1,m}^1$ to the bottom of $\Gamma_{n-1,m}$. This is an upper-triangular invertible matrix. Using (4.15) and the last row of (4.16) we obtain a formula for $\mathcal{M}_{n-1,m}A_{n,m}$ in terms of known matrices, which allows us to compute $A_{n,m}$.

Computation of $I_{n,m}$. Writing (4.17) as

$$
U_m I_{n,m} = -(\Gamma_{n,m}U_m^\top)^{-1}\mathcal{K}_{n,m}\tilde{\Gamma}_{n,m},
$$

we can compute all entries of $I_{n,m}$ except the last row. But the last row is simply $(0,0,\ldots,0,1)$, which completes the computation of $I_{n,m}$.

Computation of $B_{n-1,m}$. Similarly to $A_{n,m}$, we can combine (4.19) and the last of (4.20) to obtain a formula for $\mathcal{M}_{n-1,m}B_{n-1,m}$ in terms of known matrices.

We can compute $\tilde{A}_{n,m}$, $\tilde{J}_{n,m}^1$, $\tilde{J}_{n,m}^2$, and $\tilde{B}_{n,m-1}$ from (4.26), (4.24), (4.25), and (4.27).

Finally, we compute $\mathbb{P}_{n,m}(x,y)$ using (3.13) and (3.12). Similar to the computation of $A_{n,m}$ we obtain a formula for $\mathcal{M}_{n,m}\mathbb{P}_{n,m}$ in terms of known expressions, where $\mathcal{M}_{n,m}$ denotes the $(m+1) \times (m+1)$ matrix obtained by adding the last row of $J_{n,m}^1$ to the bottom of $\Gamma_{n,m}$. $\tilde{\mathbb{P}}_{n,m}(x,y)$ can be computed from the relation $(\tilde{3}.14)$ analogous to (3.14) for $\tilde{\mathbb{P}}_{n,m}$.

7. Construction of the linear functional. The above algorithm allows us to find a linear functional given the coefficients in the recurrence formulas. More precisely, we have the next theorem.

THEOREM 7.1. *Given parameters $s_{0,0}, \ldots, s_{2n,2m} \in \mathbb{R}$, we construct*
- *scalars $A_{i+1,0}$, $B_{i,0}$, $i = 0, \ldots, n-1$, and $\tilde{A}_{0,j+1}$, $\tilde{B}_{0,j}$, $j = 0, \ldots, m-1$;*
- *$j \times i$ matrices $\mathcal{K}_{i,j}$ and $J^2_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$;*
- *$j \times (j+1)$ matrices $J^1_{i,j}$ for $i = 1, \ldots, n$, $j = 1, 2 \ldots, m$.*

If

$$(7.1) \qquad\qquad s_{2i,2j} > 0 \ and \ ||\mathcal{K}_{i,j}|| < 1,$$

then there exists a positive linear functional $\mathcal{L}$ such that

$$(7.2) \qquad \mathcal{L}(\mathbb{P}_{i,m}, \mathbb{P}_{j,m}) = \delta_{i,j} I_{m+1} \ and \ \mathcal{L}(\tilde{\mathbb{P}}_{n,i}, \tilde{\mathbb{P}}_{n,j}) = \delta_{i,j} I_{n+1}.$$

The conditions (7.1) *are also necessary.*

Remark 7.2. The condition $||\mathcal{K}_{i,j}|| < 1$ imposes restrictions on the parameters $s_{i,j}$. In particular, it forces $|s_{2i-1,2j-1}| < 1$ for $i = 1, \ldots, n$, and $j = 1, \ldots, m$.

Proof. We construct the linear functional by induction. First, if $n = m = 0$ we set

$$\mathcal{L}(1) = \frac{1}{s^2_{0,0}} \ and \ p_{0,0} = \tilde{p}_{0,0} = s_{0,0},$$

and thus $\mathcal{L}(p_{0,0}, p_{0,0}) = \mathcal{L}(\tilde{p}_{0,0}, \tilde{p}_{0,0}) = 1$.

If $m = 0$, we construct $\mathbb{P}_{n,0} = p^0_{n,0}$ using (6.3) and then we define

$$\mathcal{L}(\mathbb{P}_{i,0}, \mathbb{P}_{j,0}) = \delta_{i,j}.$$

This gives a well-defined positive linear functional on x^j for $j = 0, 1 \ldots, n$.

Likewise, if $n = 0$, we construct $\tilde{\mathbb{P}}_{0,k} = \tilde{p}^0_{0,k}$ using (6.5) and define

$$\mathcal{L}(\tilde{\mathbb{P}}_{0,i}, \tilde{\mathbb{P}}_{0,j}) = \delta_{i,j},$$

which defines the linear functional on y^j for $j = 0, 1 \ldots, m$. Thus formula (7.2) will hold if $m = 0$ or $n = 0$.

Assume now that the functional $\mathcal{L}$ is defined for all levels before (n, m). We first extend $\mathcal{L}$ so that

$$(7.3) \qquad\qquad \mathcal{L}(\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m}) = \mathcal{K}_{n,m}.$$

To check that the above equation is consistent with how $\mathcal{L}$ is defined on the previous levels, note that

$$(7.4) \qquad\qquad \mathcal{L}(\Gamma_{n,m-1} \mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m}) = \Gamma_{n,m-1} \mathcal{K}_{n,m},$$

which follows from the construction of $\mathcal{K}_{n,m}$ and the definition of $\mathcal{L}$ on the previous levels (see Lemma 4.1). Similarly, using the second defining relation of $\mathcal{K}_{n,m}$ (i.e., the last row of (4.2)) we see that

$$(7.5) \qquad \mathcal{L}(E_{m,m} \mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m} \tilde{\Gamma}^\top_{n-1,m}) = E_{m,m} \mathcal{K}_{n,m} \tilde{\Gamma}^\top_{n-1,m}.$$

Equations (7.4) and (7.5) show that (7.3) is automatically true except the equality of the entries at (m, n) place (i.e., the definition of the linear functional on the previous levels and the construction of $\mathcal{K}_{n,m}$ imply most of (7.3)). We use the (m, n) entry to extend the functional on $x^{2n-1}y^{2m-1}$, i.e., we define $\mathcal{L}(x^{2n-1}y^{2m-1})$ so that (7.3) holds.

Using the same arguments as in the proof of (3.25) we show that

$$(7.6) \qquad J_{n,m}^3 = -\mathcal{K}_{n,m}\tilde{A}_{n-1,m}^\top = -\mathcal{L}(y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m-1}).$$

Similar to $\mathcal{K}_{n,m}$ we can use the construction of $J_{n,m}^2$ to extend the functional on $x^{2n-1}y^{2m}$ so that

$$(7.7) \qquad J_{n,m}^2 = -\mathcal{L}(y\mathbb{P}_{n,m-1}, \tilde{\mathbb{P}}_{n-1,m}).$$

Finally we use $J_{n,m}^1$ to extend the functional on $x^{2n}y^{2m-1}$ and $x^{2n}y^{2m}$ in such a way that

$$(7.8) \qquad J_{n,m}^1 = \mathcal{L}(y\mathbb{P}_{n,m-1}, \mathbb{P}_{n,m}).$$

This completes the extension of the linear functional. It remains to show that the orthogonality relations (7.2) hold. Recall that $\mathbb{P}_{n,m}$ is constructed by using (3.12) and the last row of (3.13). The orthogonality relations in the previous levels and (7.3), (7.6), and (7.7) imply that

$$\mathcal{L}(\Gamma_{n,m}\mathbb{P}_{n,m}, \tilde{\mathbb{P}}_{n-1,k}) = \mathcal{L}(E_{m,m}J_{n,m}^1\mathbb{P}_{n,m}, \tilde{\mathbb{P}}_{n-1,k}) = 0 \text{ for } k = 0, 1, \ldots, m,$$

hence

$$\mathcal{L}(\mathbb{P}_{n,m}, \tilde{\mathbb{P}}_{n-1,k}) = 0 \text{ for } k = 0, 1, \ldots, m.$$

From the last equation it follows that

$$(7.9) \qquad \mathcal{L}(\mathbb{P}_{n,m}, \mathbb{P}_{k,m}) = 0 \text{ for } k = 0, 1, \ldots, n-1.$$

It remains to show that

$$(7.10) \qquad \mathcal{L}(\mathbb{P}_{n,m}, \mathbb{P}_{n,m}) = I_{m+1}.$$

This can be derived from the two equalities

$$\mathcal{L}(\Gamma_{n,m}\mathbb{P}_{n,m}, \Gamma_{n,m}\mathbb{P}_{n,m}) = \Gamma_{n,m}\Gamma_{n,m}^\top,$$
$$\mathcal{L}(E_{m,m}J_{n,m}^1\mathbb{P}_{n,m}, \mathbb{P}_{n,m}) = E_{m,m}J_{n,m}^1.$$

Conversely, one can easily show that conditions (7.1) are necessary. Indeed, $s_{2i,2j} > 0$ follows from the normalization in (3.2) that the coefficient of the highest term is positive and (3.13). (3.24) shows that $\mathcal{K}_{i,j}$ must be a contraction, i.e., $||\mathcal{K}_{i,j}|| < 1$. $\square$

Remark 7.3. The above construction gives simple criteria for the existence of a one-step extension of the functional. That is, given moments so that there exists a positive linear functional on $\prod^{2n-2,2m} \cup \prod^{2n,2m-2}$, any set

$$\{h_{2n-1,2m-1}, h_{2n-1,2m}, h_{2n,2m-1}, h_{2n,2m}\}$$

that satisfies (7.1) can be used to extend the functional to $\prod^{2n,2m}$.

8. Interpretation of the condition $\mathcal{K}_{n,m} = 0$. In this section we classify two variable orthogonal polynomials, which can be obtained as a tensor product of two sets of (one variable) orthogonal polynomials. In other words, we want to see when $\mathbb{P}_{i,m}(x, y)$ can be written as

$$(8.1) \qquad \mathbb{P}_{i,m}(x, y) = p_i(x) \begin{bmatrix} \tilde{p}_0(y) \\ \vdots \\ \tilde{p}_m(y) \end{bmatrix}$$

for some orthogonal polynomials $p_i(x)$ and $\tilde{p}_j(y)$. The next proposition lists simple implications of (8.1).

PROPOSITION 8.1. *Assume that* (8.1) *holds for* $i = 1, 2, \ldots, m$. *Then, for all* $i = 1, 2, \ldots, m$, *and* $j = 1, 2, \ldots, n$, *we have*

$$(8.2) \qquad \mathbb{P}_{i,j}(x, y) = p_i(x) \begin{bmatrix} \tilde{p}_0(y) \\ \vdots \\ \tilde{p}_j(y) \end{bmatrix}, \qquad \tilde{\mathbb{P}}_{i,j}(x, y) = \tilde{p}_j(y) \begin{bmatrix} p_0(x) \\ \vdots \\ p_i(x) \end{bmatrix},$$

and

(8.3)

$$\mathcal{K}_{ij} = 0, \quad J^2_{i,j} = 0, \quad A_{i,j} = a_i I_{j+1}, \quad \tilde{A}_{i,j} = \tilde{a}_j I_{i+1},$$

$$J^1_{i,j} = \begin{bmatrix} \tilde{b}_0 & \tilde{a}_1 & & & & \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{b}_2 & & & \\ & \tilde{a}_2 & \tilde{b}_2 & \tilde{a}_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \tilde{a}_{j-1} & \tilde{b}_{j-1} & \tilde{a}_j \end{bmatrix}, \quad \tilde{J}^1_{i,j} = \begin{bmatrix} b_0 & a_1 & & & & \\ a_1 & b_1 & b_2 & & & \\ & a_2 & b_2 & a_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{i-1} & b_{i-1} & a_i \end{bmatrix}.$$

The scalars a_i, b_{i-1} *for* $i = 1, 2, \ldots, n$ *and* $\tilde{a}_j, \tilde{b}_{j-1}$ *for* $j = 1, 2, \ldots, m$, *are the coefficients in the three term recurrence formulas for the orthogonal polynomials* $p_i(x)$ *and* $\tilde{p}_j(y)$ *respectively, i.e.,*

$$(8.4) \qquad x p_i(x) = a_{i+1} p_{i+1}(x) + b_i p_i(x) + a_i p_{i-1}(x),$$

$$(8.5) \qquad y \tilde{p}_j(x) = \tilde{a}_{j+1} \tilde{p}_{j+1}(y) + \tilde{b}_j \tilde{p}_j(y) + \tilde{a}_j \tilde{p}_{j-1}(y).$$

Proof. If (8.1) holds, then the orthogonality $\langle \mathbb{P}_{i_1,m}, \mathbb{P}_{i_2,m} \rangle = \delta_{i_1,i_2} I_{m+1}$ is equivalent to

$$(8.6) \qquad \langle p_{i_1}(x) \tilde{p}_{j_1}(y), \, p_{i_2}(x) \tilde{p}_{j_2}(y) \rangle = \delta_{i_1,i_2} \delta_{j_1,j_2}.$$

From this relation one can easily obtain (8.2) and (8.3). □

Remark 8.2. Notice that if $p_i(x)$ and $\tilde{p}_j(y)$ satisfy (8.2) and c is a nonzero constant, then the polynomials

$$(8.7) \qquad q_i(x) = c p_i(x) \text{ and } \tilde{q}_j(y) = \tilde{p}_j(y)/c$$

also satisfy (8.2). Conversely, if two pairs $p_i, \tilde{p}_j$ and $q_i, \tilde{q}_j$ of scalar orthogonal polynomials satisfy (8.2), then there is a nonzero constant such that (8.7) holds. Thus the polynomials $p_i, \tilde{p}_j$ are unique up to a multiplicative constant.

LEMMA 8.3. *If* $\mathcal{K}_{i,j} = 0$ *for* $i = 1, \ldots, (n+1)$, *and* $j = 1, \ldots, (m+1)$, *then* (8.2) *holds for* $i \leq n$ *and* $j \leq m$.

Proof. The proof is by induction. This is obvious if $n = 0$ or $m = 0$. Assume now that this is true for levels $(n-1, m)$ and $(n, m-1)$. We need to show that it is also true for level (n, m). The induction hypothesis and Proposition 8.1 show that (8.2) and (8.3) hold

- if $i \leq n-1$ and $j \leq m$, or
- if $i \leq n$ and $j \leq m-1$.

In particular, $A_{n,m-1} = a_n I_m$. (3.24), (3.25), (4.1), and (4.2), combined with $\mathcal{K}_{i,j} = 0$ for $i \leq n+1$ and $j \leq m+1$, imply that

$$(8.8) \qquad \Gamma_{i,j} = [I_j | 0], \quad J_{i,j}^3 = 0, \quad \text{and} \quad J_{i,j}^2 = 0 \quad \text{for} \quad i \leq n+1, \; j \leq m.$$

Using (8.8) and the defining relations (4.13)–(4.14) for $J_{n,m}^1$ we see that $J_{n,m}^1$ has the form

$$(8.9) \qquad J_{n,m}^1 = \begin{bmatrix} \tilde{b}_0 & \tilde{a}_1 & & & & & \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{b}_2 & & & & \\ & \tilde{a}_2 & \tilde{b}_2 & \tilde{a}_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \tilde{a}_{m-2} & \tilde{b}_{m-2} & \tilde{a}_{m-1} & \\ & & & & \tilde{a}_{m-1} & c_1 & c_2 \end{bmatrix}$$

for some scalars c_1, c_2. (4.15) combined with (8.3) and (8.8) shows that

$$(8.10) \qquad A_{n,m} = \begin{bmatrix} a_n & & & & & \\ & a_n & & & & \\ & & a_n & & & \\ & & & \ddots & & \\ & & & & a_n & \\ d_1 & d_2 & \ldots & d_{m-1} & d_m & d \end{bmatrix}$$

for some scalars $d_1, d_2, \ldots, d_m, d$. Plugging (8.9), (8.10), and $A_{n,m-1} = a_n I_m$ into (4.16) and comparing the entries on the last row, we get the equalities

$$
\begin{aligned}
\tilde{a}_m d_1 &= 0, \\
\tilde{a}_m d_2 &= 0, \\
&\;\;\vdots \\
\tilde{a}_m d_{m-2} &= 0, \\
\tilde{a}_{m-1} a_n + \tilde{a}_m d_{m-1} &= a_n \tilde{a}_{m-1}, \\
\tilde{b}_{m-1} a_n + \tilde{a}_m d_m &= a_n c_1, \\
\tilde{a}_m d &= a_n c_2.
\end{aligned}
$$

$$(8.11)$$

The first $(m-1)$ equations simply imply that $d_1 = d_2 = \cdots = d_{m-1} = 0$.

Formula (4.4) (for $J_{n+1,m}^2$) becomes

$$0 = J_{n,m}^1 A_{n,m}^\top I_{n-1,m} - A_{n,m-1}^\top I_{n-1,m-1} \tilde{A}_{n-1,m}.$$

Using (8.3) we can rewrite the last equation as

$$J^1_{n,m} A^\top_{n,m} I_{n-1,m} = a_n \tilde{a}_m I_{n-1,m-1}.$$

On the other hand, from (4.17) it follows that for $i \leq n$ and $j \leq m$ all entries, except $(j+1, i+1)$, of the matrix $I_{i,j}$ are equal to zero. Thus the last equality simply means that the last column of the matrix $J^1_{n,m} A^\top_{n,m}$ is equal to $(0, 0, \ldots, 0, a_n \tilde{a}_m)^\top$. Using (8.9) and (8.10) we see that the bottom two entries of the last column of the matrix $J^1_{n,m} A^\top_{n,m}$ are $\tilde{a}_{m-1} d_m$ and $c_2 d$, i.e., we have

(8.12)
$$\tilde{a}_{m-1} d_m = 0,$$
$$c_2 d = a_n \tilde{a}_m.$$

The first equation implies that $d_m = 0$, while the second one combined with (8.15) implies that $d^2 = a_n^2$ and $c_2^2 = \tilde{a}_m^2$. Since all the numbers d, c_2, a_n, and $\tilde{a}_m$ are positive, it follows that $d = a_n$ and $c_2 = \tilde{a}_m$. Finally notice that the $(m-1)$st equation in formula (8.11) gives $c_1 = \tilde{b}_{m-1}$. Thus we proved that

(8.13)
$$A_{n,m} = a_n I_{m+1} \text{ and } J^1_{n,m} = \begin{bmatrix} \tilde{b}_0 & \tilde{a}_1 & & & & \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{b}_2 & & & \\ & \tilde{a}_2 & \tilde{b}_2 & \tilde{a}_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \tilde{a}_{m-1} & \tilde{b}_{m-1} & \tilde{a}_m \end{bmatrix}.$$

From the last formula and (3.12), (3.13), and (3.14), one can easily see that (8.2)–(8.3) hold for $i = n$ and $j = m$, which completes the proof. □

As a corollary we get the following theorem.

THEOREM 8.4. *The following conditions are equivalent:*

(i) *For all n and m, $\mathbb{P}_{n,m}$ is a tensor product of scalar polynomials $p_k(x)$ and $\tilde{p}_j(y)$, i.e., we have*

(8.14)
$$\mathbb{P}_{n,m}(x, y) = p_n(x) \begin{bmatrix} \tilde{p}_0(y) \\ \vdots \\ \tilde{p}_m(y) \end{bmatrix}.$$

(ii) $\mathcal{K}_{n,m} = 0$ *for all $n, m = 1, 2, \ldots$.*

Next we want to prove a finite analogue of Theorem 8.4, i.e., to give necessary and sufficient conditions for (8.4) to hold up to a given level (n, m). We need the following lemma.

LEMMA 8.5. *If $\mathcal{K}_{i,j} = 0$ and $(J^2_{i,j})_{j,i} = 0$ for $i = 1, \ldots, n$, and $j = 1, \ldots, m$, then (8.2) holds if $i < n$ and $j \leq m$ or if $i \leq n$ and $j < m$.*

Proof. We prove the statement by induction. If $n = 0$ or $m = 0$ this is a trivial statement. Assume now that this is true for levels $(n-1, m)$ and $(n, m-1)$. We will show that it also holds for level (n, m). The induction hypothesis means that (8.2) and (8.3) hold

- if $i \leq n - 2$ and $j \leq m$, or
- if $i \leq n - 1$ and $j \leq m - 1$, or
- if $i \leq n$ and $j \leq m - 2$.

As before, the condition $\mathcal{K}_{i,j} = 0$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$ implies that

(8.15) $\qquad J_{i,j}^3 = 0$ and $\Gamma_{i,j} = [I_j | 0]$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

Formula (4.2) gives

$$\tilde{J}_{n-1,m}^2 = 0,$$

which combined with (4.25) for $\tilde{J}_{n-1,m}^2$ shows that

(8.16) $\qquad\qquad\qquad I_{n-2,m}^\top A_{n-1,m} \Gamma_{n-1,m}^\top = 0.$

Using the induction hypothesis and (4.3)–(4.4) one can prove that $J_{n-1,m}^2 = 0$. Indeed, the right-hand side of (4.3) is obviously zero, while on the right-hand side of (4.4) there are only two nonzero terms:

$$J_{n-2,m}^1 A_{n-2,m}^\top I_{n-3,m} - A_{n-2,m-1}^\top I_{n-3,m-1} \tilde{A}_{n-3,m}.$$

Using (8.3) one can easily see that the above expression vanishes. This shows that all entries of $J_{n-1,m}^2$ are equal to 0, except probably $(J_{n-1,m}^2)_{m,n-1}$, but we know that this entry is 0, because this is one of the assumptions in the lemma.

Formula (8.3) and the defining relations (4.13)–(4.14) for $J_{n-1,m}^1$ show that it has the form

(8.17) $\qquad J_{n-1,m}^1 = \begin{bmatrix} \tilde{b}_0 & \tilde{a}_1 & & & & & \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{b}_2 & & & & \\ & \tilde{a}_2 & \tilde{b}_2 & \tilde{a}_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \tilde{a}_{m-2} & \tilde{b}_{m-2} & \tilde{a}_{m-1} & \\ & & & & \tilde{a}_{m-1} & c_1 & c_2 \end{bmatrix}$

for some scalars c_1, c_2. (4.15) for $A_{n-1,m}$ together with (8.3) imply

(8.18) $\qquad A_{n-1,m} = \begin{bmatrix} a_{n-1} & & & & & \\ & a_{n-1} & & & & \\ & & a_{n-1} & & & \\ & & & \ddots & & \\ & & & & a_{n-1} & \\ d_1 & d_2 & \cdots & d_{m-1} & d_m & d \end{bmatrix}$

for some scalars $d_1, d_2, \ldots, d_m, d$. Plugging (8.18) into (8.16) and using (8.15) we see that $d_1 = d_2 = \cdots = d_m = 0$ and $A_{n-1,m}$ is a diagonal matrix of the form

(8.19) $\qquad A_{n-1,m} = \begin{bmatrix} a_{n-1} & & & & & \\ & a_{n-1} & & & & \\ & & a_{n-1} & & & \\ & & & \ddots & & \\ & & & & a_{n-1} & \\ & & & & & d \end{bmatrix}.$

Writing (4.16) for $n-1$, plugging (8.17) and (8.19), and comparing the last two entries on the last row we get the equalities

$$(8.20) \qquad \tilde{b}_{m-1}a_{n-1} = a_{n-1}c_1,$$

$$(8.21) \qquad \tilde{a}_m d = a_{n-1}c_2.$$

The first equality gives $c_1 = \tilde{b}_{m-1}$.

Finally, notice that the last column of the left-hand side of (4.4) must be 0, which, in particular, implies that the last entry in the last column on the right-hand side of (4.4) must be zero. Computing this entry we get

$$(8.22) \qquad c_2 d - a_{n-1}\tilde{a}_m = 0.$$

(8.21) and (8.22) show that $c_2 = \tilde{a}_m$ and $d = a_{n-1}$. Thus

$$A_{n-1,m} = a_{n-1}I_{m+1} \quad \text{and} \quad J^1_{n-1,m} = \begin{bmatrix} \tilde{b}_0 & \tilde{a}_1 & & & & & \\ \tilde{a}_1 & \tilde{b}_1 & \tilde{b}_2 & & & & \\ & \tilde{a}_2 & \tilde{b}_2 & \tilde{a}_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \tilde{a}_{m-2} & \tilde{b}_{m-2} & \tilde{a}_{m-1} & \\ & & & & \tilde{a}_{m-1} & \tilde{b}_{m-1} & \tilde{a}_m \end{bmatrix},$$

which combined with (3.13) shows that (8.2)–(8.3) hold also for $i = n-1$ and $j = m$. Similarly one can show that (8.2)–(8.3) hold for $i = n$ and $j = m-1$, completing the induction. $\quad\square$

COROLLARY 8.6. *For given* $n, m \geq 0$, *the following conditions are equivalent:*
(i) *For all* $i = 0, 1, 2, \ldots, n$, $\mathbb{P}_{i,m}$ *is a tensor product of scalar polynomials* $p_k(x)$ *and* $\tilde{p}_j(y)$, *i.e.,* (8.1) *holds.*
(ii) $\quad$ – $\mathcal{K}_{i,j} = 0$ *for all* $i \leq n$, $j \leq m$;
$\quad$ – $(J^2_{i,j})_{j,i} = 0$ *if* $i = n$ *or* $j = m$;
$\quad$ – $(J^1_{n,m})_{m,m+1} = \tilde{A}_{0,m}$ *and* $(J^1_{n,m})_{m,m} = \tilde{B}_{0,m-1}$.

9. Numerical examples. In this last section we give several numerical examples which illustrate the algorithm.

Example 9.1. Let us choose parameters $s_{i,j}$ as follows:

$s_{0,0} = 0.7543$	$s_{0,1} = 0$	$s_{0,2} = 0.4633$	$s_{0,3} = 0$	$s_{0,4} = 0.4997$
$s_{1,0} = 0$	$s_{1,1} = -0.1185$	$s_{1,2} = 0$	$s_{1,3} = -0.1128$	$s_{1,4} = 0$
$s_{2,0} = 0.4634$	$s_{2,1} = 0$	$s_{2,2} = 0.4681$	$s_{2,3} = 0$	$s_{2,4} = 0.4998$
$s_{3,0} = 0$	$s_{3,1} = -0.1128$	$s_{3,2} = 0$	$s_{3,3} = -0.1073$	$s_{3,4} = 0$
$s_{4,0} = 0.4997$	$s_{4,1} = 0$	$s_{4,2} = 0.4682$	$s_{4,3} = 0$	$s_{4,4} = 0.4997$

First we apply the algorithm at levels $(0,0)$, $(0,1)$, $(0,2)$, $(1,0)$, $(2,0)$. (There is no restriction on the parameters at these levels.) Next we compute $\mathcal{K}_{1,1} = [s_{1,1}]$:

$$\mathcal{K}_{1,1} = [-0.1185].$$

Since $|s_{1,1}| < 1$, $\mathcal{K}_{1,1}$ is a contraction and we can compute all other matrices and the polynomials at level $(1,1)$. Then we proceed to levels $(1,2)$ and $(2,1)$. We obtain

$$\mathcal{K}_{1,2} = \begin{bmatrix} 0 \\ -0.1128 \end{bmatrix}$$

and

$$\mathcal{K}_{2,1} = \begin{bmatrix} 0 & -0.1128 \end{bmatrix},$$

which are obviously contractions, hence we can compute everything else at these levels using the algorithm. Finally, at level $(2,2)$ we get

$$\mathcal{K}_{2,2} = \begin{bmatrix} 0.0174 & 0 \\ 0 & -0.1073 \end{bmatrix},$$

which again is a contraction and allows us to apply the algorithm.

The example described above corresponds to the following moment problem:

$h_{0,0} = 1.7575$	$h_{0,1} = 0$	$h_{0,2} = 0.3773$	$h_{0,3} = 0$	$h_{0,4} = 0.1752$
$h_{1,0} = 0$	$h_{1,1} = -0.0447$	$h_{1,2} = 0$	$h_{1,3} = -0.0196$	$h_{1,4} = 0$
$h_{2,0} = 0.3773$	$h_{2,1} = 0$	$h_{2,2} = 0.0838$	$h_{2,3} = 0$	$h_{2,4} = 0.0395$
$h_{3,0} = 0$	$h_{3,1} = -0.0196$	$h_{3,2} = 0$	$h_{3,3} = -0.0087$	$h_{3,4} = 0$
$h_{4,0} = 0.1752$	$h_{4,1} = 0$	$h_{4,2} = 0.0395$	$h_{4,3} = 0$	$h_{4,4} = 0.0188.$

Next we present an example of a moment problem which cannot be extended.

Example 9.2. Let us fix the free parameters at levels $(0,0)$, $(0,1)$, $(0,2)$, $(1,0)$, $(2,0)$, and $(1,1)$ as follows:

Level $(0,0)$: $s_{0,0} = 1$;
Level $(0,1)$: $s_{0,1} = 1/7$, $s_{0,2} = 2$;
Level $(0,2)$: $s_{0,3} = 1/3$, $s_{0,4} = 7$;
Level $(1,0)$: $s_{1,0} = 1$, $s_{2,0} = 3$;
Level $(2,0)$: $s_{3,0} = 0$, $s_{4,0} = 0.33$;
Level $(1,1)$: $s_{1,1} = 0.5$, $s_{1,2} = 0$, $s_{2,1} = 0$, $s_{2,2} = 1$.

Applying the algorithm described in section 6, we obtain a functional $\mathcal{L}$ defined on the space $\{x^i y^j : i+j \leq 2\}$ and the orthogonal polynomials $\mathbb{P}$ and $\tilde{\mathbb{P}}$ corresponding to this functional.

Computing $\mathcal{K}_{2,1}$ and $\mathcal{K}_{1,2}$ we obtain

$$\mathcal{K}_{2,1} = \begin{bmatrix} 0.9997407262 & s_{3,1} \end{bmatrix}, \qquad \mathcal{K}_{1,2} = \begin{bmatrix} -0.02749286996 \\ s_{1,3} \end{bmatrix},$$

which shows that if we pick $s_{3,1}$ and $s_{1,3}$ with absolute value less than 1 and such that $\mathcal{K}_{2,1}$ and $\mathcal{K}_{1,2}$ are contractions, we can extend the functional to the space $\{x^i y^j : i+j \leq 3\}$. All other parameters at levels $(0,3)$, $(1,2)$, $(2,1)$, and $(3,0)$ can be chosen arbitrary. (Of course $s_{0,6}$, $s_{2,4}$, $s_{4,2}$ and $s_{6,0}$ must be positive.)

Finally let us compute $\mathcal{K}_{2,2}$. Entry $(1,1)$ of this matrix is

$$(\mathcal{K}_{2,2})_{1,1}$$
$$= \frac{-59.47189 - 0.2717694 \times 10^{-10} s_{1,3} + 0.1087078 \times 10^{-11} s_{3,1} s_{1,3}^2 - 43.93372 s_{3,1} s_{1,3}}{\sqrt{1 - 1928.713 s_{3,1}^2}\sqrt{1 - 1.000756 s_{1,3}^2}(1 + 1.237179 \times 10^{-14} s_{1,3})}.$$

From the above formula it is clear that $|(\mathcal{K}_{2,2})_{1,1}| > 1$, which means that $\mathcal{K}_{2,2}$ is not a contraction. Thus, we see that the functional $\mathcal{L}$ cannot be extended to level

$(2, 2)$ no matter how we choose the parameters at levels $(1, 2)$ and $(2, 1)$. In particular, it follows that $\mathcal{L}$ cannot be extended to the space of polynomials of (total) degree 3.

Remark 9.3.   The above example shows that not every functional defined on levels $(n, m - 1)$ and $(n - 1, m)$ can be extended to level (n, m) even if we modify the parameters entering one step back in each direction, that is, at levels $(n, m - 1)$ and $(n - 1, m)$. Several numerical experiments indicate that deforming the parameters two steps back in each direction is enough to extend the functional. Whether this is true or false in general is an interesting open problem.

Remark 9.4. Example 9.2 shows the simplest possible case of a moment problem which cannot be extended to a level (n, m) even if we modify the parameters entering one step back in each direction. More precisely, one can easily show that if $n = 1$ or $m = 1$ the moment problem can always be extended by deforming just one parameter entering one step back. Indeed, if, for example, $m = 1$, then $\mathcal{K}_{n,1}$ is a $1 \times n$ matrix. The first $(n - 1)$ entries are computed from (4.2). Notice that in this equation the only matrix coming from level $(n, 0)$ is $A_{n,0} = (s_{2n,0})$. Thus, if we make $s_{2n,0}$ large enough, $\mathcal{K}_{n,1}$ will be a contraction.

REFERENCES

[1] Ju. M. Berezanskii, *Expansions in eigenfunctions of self-adjoint operators*, Trans. Math. Mono. Amer. Math. Soc., 17 (1968).

[2] Ju. M. Berezanskii, *Direct and inverse spectral problems for a Jacobi field*, Algebra i Analiz, 9 (1997), pp. 38–61; translation in St. Petersburg Math. J., 9 (1998), pp. 1053–1071.

[3] C. F. Dunkl, *Intertwining operators and polynomials associated with the symmetric group*, Monatsh. Math., 126 (1998), pp. 181–209.

[4] C. F. Dunkl and Y. Xu, *Orthogonal Polynomials of Several Variables*, Encyclopedia of Mathematics and Its Applications 81. Cambridge University Press, Cambridge, UK, 2001.

[5] L. Fernández, T. E. Pérez, and M. A. Piñar, *Weak classical orthogonal polynomials in two variables*, J. Comput. Appl. Math., 178 (2005), pp. 191–203.

[6] M. I. Gekhtman and A. A. Kalyuzhny, *On the orthogonal polynomials in several variables*, Integral Equations Operator Theory, 19 (1994), pp. 404–418.

[7] J. S. Geronimo, *Scattering theory and matrix orthogonal polynomials on the real line*, Circuits Systems Signals Process, 1 (1982), pp. 471–495.

[8] J. S. Geronimo and H. J. Woerdeman, *Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables*, Ann. of Math., 160 (2004), pp. 839–906.

[9] J. S. Geronimo and H. J. Woerdeman, *Two variable orthogonal polynomials on the bi-circle and structured matrices.*

[10] D. Jackson, *Formal properties of orthogonal polynomials in two variables*, Duke Math. J., 2 (1936), pp. 423–434.

[11] Y. J. Kim, K. H. Kwon, and J. K. Lee, *Multi-variate orthogonal polynomials and second order partial differential equations*, Commun. Appl. Anal., 6 (2002), pp. 479–504.

[12] T. H. Koornwinder, *Orthogonal polynomials in two variables which are eigenfunctions of two algebraically independent partial differential operators*, I, II, Indag. Math., 36 (1974), pp. 48–66.

[13] T. H. Koornwinder, *Two variable analogues of the classical orthogonal polynomials*, in Theory and Applications of Special Functions, R. Askey, ed., Academic Press, NY, 1975, pp. 435–495.

[14] T. H. Koornwinder, *Askey-Wilson polynomials for root systems of type BC*, in Hypergeometric Functions on Domains of Positivity, Jack Polynomials, and Applications, Contemp. Math., Amer. Math. Soc., 138 (1992), pp. 189–204.

[15] M. A. Kowalski, *The recursion formulas for orthogonal polynomials in n variables*, SIAM J. Math. Anal., 13 (1982), pp. 309–315.

[16] M. A. Kowalski, *Orthogonality and recursion formulas for polynomials in n variables*, SIAM J. Math. Anal., 13 (1982), pp. 316–323.

[17] H. L. Krall and I. M. Sheffer, *Orthogonal polynomials in two variables*, Ann. Mat. Pura Appl., 76 (1967), pp. 325–376.

[18] M. G. Krein, *Infinite J-matrices and the matrix moment problem*, Dokl. Akad. Nauk. SSSR, 69 (1949), pp. 125–128.

[19] K. H. Kwon, J. K. Lee, and L. L. Littlejohn, *Orthogonal polynomial eigenfunctions of second order partial differential equations*, Trans. Amer. Math. Soc., 353 (2001), pp. 3629–3647.

[20] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, New York, 1995.

[21] P. K. Suetin, *Orthogonal Polynomials in Two Variables*, Anal. Methods Spec. Funct., 3, Gordon and Breach Science Publishers, Amsterdam, 1999.

[22] Y. Xu, *On multivariate orthogonal polynomials*, SIAM J. Math. Anal., 24 (1993), pp. 783–794.