

Average characteristic polynomials for multiple orthogonal polynomial ensembles

Steven Delvaux¹

Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium

Received 3 July 2009; accepted 9 November 2009

Available online 27 November 2009

Communicated by Guillermo López Lagomasino

Abstract

Multiple orthogonal polynomials (MOP) are a non-definite version of matrix orthogonal polynomials. They are described by a Riemann–Hilbert matrix Y consisting of four blocks $Y_{1,1}$, $Y_{1,2}$, $Y_{2,1}$ and $Y_{2,2}$. In this paper, we show that $\det Y_{1,1}$ ($\det Y_{2,2}$) equals the average characteristic polynomial (average inverse characteristic polynomial, respectively) over the probabilistic ensemble that is associated to the MOP. In this way we generalize the classical results for orthogonal polynomials, and also some recent results for MOP of type I and type II. We then extend our results to arbitrary products and ratios of characteristic polynomials. In the latter case an important role is played by a matrix-valued version of the Christoffel–Darboux kernel. Our proofs use determinantal identities involving Schur complements, and adaptations of the classical results by Heine, Christoffel and Uvarov.

© 2009 Elsevier Inc. All rights reserved.

Keywords: Multiple/matrix orthogonal polynomials; Christoffel–Darboux kernel; Riemann–Hilbert problem; Determinantal point process; Average characteristic polynomial; Schur complement; (Block) Hankel determinant

1. Introduction and statement of results

1.1. Random matrix ensembles

On the space \mathcal{H}_n of Hermitian n by n matrices consider the random matrix ensemble defined by the probability distribution

$$\frac{1}{Z_n} e^{-\text{Tr}V(M)} dM, \quad M \in \mathcal{H}_n, \quad (1.1)$$

E-mail address: steven.delvaux@wis.kuleuven.be.

¹ The author is a Postdoctoral Fellow of the Fund for Scientific Research - Flanders (Belgium).

for some given polynomial V of even degree. Here Z_n is a normalization constant, Tr denotes the trace and dM is the Lebesgue measure on \mathcal{H}_n .

The random matrix ensemble (1.1) leads to a probability distribution on the space

$$\mathbb{R}_{\leq}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 \leq \dots \leq x_n\}$$

of ordered eigenvalue tuples. It is well known that this joint probability distribution has the form

$$\frac{1}{Z_n} \prod_{i < j} (x_j - x_i)^2 \prod_{j=1}^n e^{-V(x_j)} \prod_{j=1}^n dx_j, \tag{1.2}$$

for some normalization constant \tilde{Z}_n . Thus the probability that the ordered eigenvalues of the matrix (1.1) lie in an infinitesimal box $[x_1, x_1 + dx_1] \times \dots \times [x_n, x_n + dx_n] \subset \mathbb{R}_{\leq}^n$ is given by (1.2). Note in particular that the density (1.2) is small if two eigenvalues x_i and x_j are close to each other. This means that the eigenvalues tend to ‘repel’ each other.

One can write (1.2) alternatively as

$$\frac{1}{\tilde{Z}_n} \det (f_i(x_j))_{i,j=1}^n \det (g_i(x_j))_{i,j=1}^n \prod_{j=1}^n dx_j \tag{1.3}$$

where

$$f_i(x) = x^{i-1}, \quad g_i(x) = x^{i-1} e^{-V(x)}. \tag{1.4}$$

Indeed, this follows upon recognizing (1.2) to be basically a product of two Vandermonde determinants.

Rather than the space \mathbb{R}_{\leq}^n of ordered eigenvalue tuples, we will find it convenient to consider the probability distribution (1.3)–(1.4) on the full space \mathbb{R}^n . To maintain a probability distribution we should then multiply the normalization constant \tilde{Z}_n by a factor $n!$.

To the probability distribution (1.3)–(1.4) one can associate the *average characteristic polynomial*

$$P_n(z) = \frac{1}{\tilde{Z}_n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\prod_{j=1}^n (z - x_j) \right) \det (f_i(x_j))_{i,j=1}^n \times \det (g_i(x_j))_{i,j=1}^n \prod_{j=1}^n dx_j. \tag{1.5}$$

It follows from a classical calculation of Heine, see e.g. [12,27], that P_n can be characterized as the n th monic *orthogonal polynomial* with respect to the weight function $e^{-V(x)}$ on \mathbb{R} . Thus the polynomials P_n satisfy the conditions

$$P_n(x) = x^n + O(x^{n-1})$$

for all $n \in \mathbb{N}$ and

$$\int_{-\infty}^{\infty} P_n(x) P_m(x) e^{-V(x)} dx = c_n c_m \delta_{m,n} \tag{1.6}$$

for all $n, m \in \mathbb{N}$, for certain $c_n \in \mathbb{R}$.

Intuitively, the above result states that the zeros of the monic orthogonal polynomial $P_n(z)$ determine the ‘typical’ eigenvalue configuration of the random matrix ensemble (1.1). This holds in particular in the Gaussian case $V(x) = x^2/2$, where the ensemble (1.1) reduces to the well-known *Gaussian unitary ensemble* (GUE) while the corresponding orthogonal polynomials are (up to scaling factors) the classical *Hermite polynomials*.

In the literature, many more results on average characteristic polynomials for random matrix ensembles can be found, see e.g. [4,7,8,21,26] and the references therein. We mention the following result of Fyodorov–Strahov [16]. Define the *average inverse characteristic polynomial* corresponding to the probability distribution (1.3)–(1.4) as

$$\begin{aligned}
 Q_n(z) &= \frac{1}{Z_n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\prod_{j=1}^n (z - x_j)^{-1} \right) \det (f_i(x_j))_{i,j=1}^n \\
 &\quad \times \det (g_i(x_j))_{i,j=1}^n \prod_{j=1}^n dx_j,
 \end{aligned}
 \tag{1.7}$$

for $z \in \mathbb{C} \setminus \mathbb{R}$. Then it holds that

$$Q_n(z) = \frac{1}{c_{n-1}^2} \int_{-\infty}^{\infty} \frac{P_{n-1}(x)e^{-V(x)}}{z - x} dx,
 \tag{1.8}$$

with c_n as in (1.6). Fyodorov–Strahov [16,26] also observed that both $P_n(z)$ and the right hand side of (1.8) have a natural interpretation in terms of the associated *Riemann–Hilbert problem* (briefly *RH problem*), and they obtained determinantal formulas for averages of more general products and ratios of characteristic polynomials; see further.

In recent years there has been interest in the following generalization of (1.1),

$$\frac{1}{Z_n} e^{-\text{Tr}(V(M)-AM)} dM.
 \tag{1.9}$$

Here A is a fixed diagonal matrix which is called the *external source*. Typically A has only a small number of distinct eigenvalues a_1, \dots, a_p , say with corresponding multiplicities n_1, \dots, n_p .

The model (1.9) was first studied by Brézin–Hikami [9,10] and Zinn–Justin [31] who showed that the eigenvalue correlations are determinantal. In [5] it was shown that the eigenvalues x_1, \dots, x_n have a joint probability distribution of the form (1.3), where now $f_i(x) = x^{i-1}$ (Vandermonde factor) while $g_i(x)$ are given by several Vandermonde-like series:

$$\begin{aligned}
 g_i(x) &= x^{i-1} e^{-(V(x)-a_1x)}, \quad i = 1, \dots, n_1, \\
 g_{n_1+i}(x) &= x^{i-1} e^{-(V(x)-a_2x)}, \quad i = 1, \dots, n_2, \\
 &\vdots \\
 g_{n_1+\dots+n_{p-1}+i}(x) &= x^{i-1} e^{-(V(x)-a_px)}, \quad i = 1, \dots, n_p.
 \end{aligned}$$

To this eigenvalue ensemble one can associate the average characteristic polynomial $P_n(z)$ in exactly the same way as before, see (1.5). Bleher–Kuijlaars [5] showed that $P_n(z)$ satisfies the multiple orthogonality relations

$$\int_{-\infty}^{\infty} P_n(x)x^i e^{-(V(x)-a_kx)} dx = 0,$$

for $i = 0, \dots, n_k - 1$ and $k = 1, \dots, p$. For more background on this kind of orthogonality relations, see e.g. [2,29]. Desrosiers–Forrester [13] showed that the result (1.8) on the average inverse characteristic polynomial (1.7) can also be generalized to the random matrix ensemble with external source (1.9).

The goal of this paper is to generalize the above results to the more general context of *multiple orthogonal polynomial ensembles* (briefly *MOP ensembles*) in the sense of Daems–Kuijlaars [11]. We will show that in general, (1.5) and (1.7) can be expressed as *Riemann–Hilbert minors*, i.e., determinants of certain submatrices of the Riemann–Hilbert matrix. Using these results, we will then obtain determinantal formulas for averages of arbitrary products and ratios of characteristic polynomials, by adapting the method of Baik–Deift–Strahov [4]. Our results will be stated in Section 1.4. In Section 1.2–1.3 we first recall the basic definitions concerning MOP ensembles.

1.2. MOP ensembles

We consider a stochastic model in the following way [11]. Let $p, q \in \mathbb{N}$ be two positive integers. Let there be given

- A (finite) sequence of positive integers $n_1, n_2, \dots, n_p \in \mathbb{N}$;
- A sequence of weight functions $w_{1,1}(x), w_{1,2}(x), \dots, w_{1,p}(x) : \mathbb{R} \rightarrow \mathbb{R}$;
- A sequence of positive integers $m_1, m_2, \dots, m_q \in \mathbb{N}$;
- A sequence of weight functions $w_{2,1}(x), w_{2,2}(x), \dots, w_{2,q}(x) : \mathbb{R} \rightarrow \mathbb{R}$.

We will use the vector notations $\mathbf{n} := (n_1, \dots, n_p)$, $|\mathbf{n}| := \sum_{k=1}^p n_k$ and similarly for \mathbf{m} and $|\mathbf{m}|$. Occasionally we will also write $\mathbf{w}_1(x) := (w_{1,1}(x), \dots, w_{1,p}(x))$ and similarly for $\mathbf{w}_2(x)$. Assume in what follows that $|\mathbf{n}| = |\mathbf{m}| = n$.

To the above data one can associate a stochastic model, called a *multiple orthogonal polynomial ensemble* or briefly *MOP ensemble*. This model consists of n random points x_1, \dots, x_n on the real line whose joint p.d.f. can be written as a product of two determinants:

$$\frac{1}{Z_n} \det (f_i(x_j))_{i,j=1}^n \cdot \det (g_i(x_j))_{i,j=1}^n, \tag{1.10}$$

where Z_n is a normalization factor, and with functions

$$\begin{aligned} f_i(x) &= x^{i-1} w_{1,1}(x), & i = 1, \dots, n_1, \\ f_{n_1+i}(x) &= x^{i-1} w_{1,2}(x), & i = 1, \dots, n_2, \\ &\vdots \\ f_{n_1+\dots+n_{p-1}+i}(x) &= x^{i-1} w_{1,p}(x), & i = 1, \dots, n_p, \end{aligned} \tag{1.11}$$

and

$$\begin{aligned} g_i(x) &= x^{i-1} w_{2,1}(x), & i = 1, \dots, m_1, \\ g_{m_1+i}(x) &= x^{i-1} w_{2,2}(x), & i = 1, \dots, m_2, \\ &\vdots \\ g_{m_1+\dots+m_{q-1}+i}(x) &= x^{i-1} w_{2,q}(x), & i = 1, \dots, m_q. \end{aligned} \tag{1.12}$$

Note that this is a special case of a biorthogonal ensemble [6]. The motivation for the name ‘MOP ensemble’ comes from the multiple orthogonal polynomials that are related to it, see Section 2. We note that polynomials of this kind were already studied in [24].

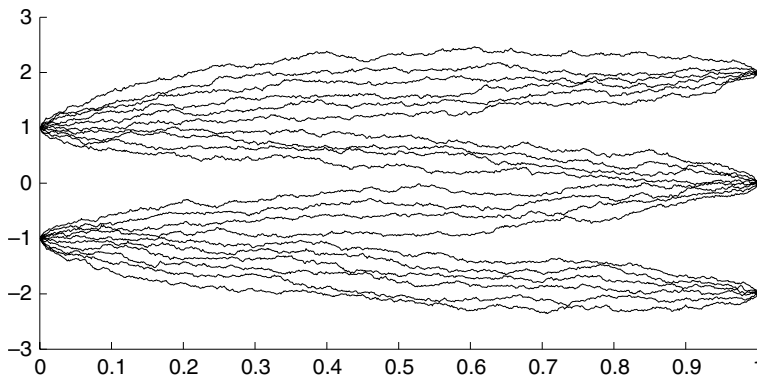


Fig. 1. The figure shows $n = 20$ non-intersecting Brownian motions with $p = 2$ starting points and $q = 3$ ending points. The horizontal axis denotes the time $t \in [0, 1]$ and the vertical axis denotes the space variable x . For each $t \in (0, 1)$ the positions x_1, \dots, x_n of the paths at time t form a MOP ensemble (1.10)–(1.12).

In order for the above model to be a valid probability distribution on \mathbb{R}^n one should have that (1.10) is positive on \mathbb{R}^n . We will not address this topic here since our algebraic results will be valid irrespective of this positivity condition.

The normalization constant Z_n in (1.10) serves to make the total probability of the MOP ensemble on \mathbb{R}^n equal to 1. We will further obtain an expression for Z_n as a block Hankel determinant; see the remark at the end of Section 3.1.

Of course, the above model contains the eigenvalue ensembles in Section 1.1 as special cases. Further motivation comes from the theory of non-intersecting one-dimensional Brownian paths with p distinct starting positions and q distinct ending positions, see Fig. 1. See [11,20] for details. Another type of application can be found in [14,24].

To facilitate comparison with the literature, we may note the following terminology which is often used: the MOP ensemble is said to be of *type I* if $q = 1$, of *type II* if $p = 1$ and of *mixed type* if $p, q \geq 2$. Thus the results of our paper will be valid for general MOP ensembles of mixed type, whereas the results of [5,13,22] only apply to MOP ensembles of type I or type II.

1.3. Riemann–Hilbert problem and kernel

The MOP ensemble in Section 1.2 and its associated multiple orthogonal polynomials are intimately related to the following Riemann–Hilbert problem (RH problem) introduced in [11]. The RH problem generalizes the well-known RH problem for orthogonal polynomials due to Fokas–Its–Kitaev [15] as well as its generalization in [30].

RH Problem 1.1. Consider the data $\mathbf{n}, \mathbf{w}_1(x), \mathbf{m}$ and $\mathbf{w}_2(x)$ as above and assume that $|\mathbf{n}| = |\mathbf{m}| =: n$. The RH problem consists in finding a matrix-valued function $Y(z) = Y_{\mathbf{n}, \mathbf{w}_1, \mathbf{m}, \mathbf{w}_2}(z)$ of size $p + q$ by $p + q$ such that

- (1) $Y(z)$ is analytic in $\mathbb{C} \setminus \mathbb{R}$;
- (2) For $x \in \mathbb{R}$, it holds that

$$Y_+(x) = Y_-(x) \begin{pmatrix} I_p & W(x) \\ 0 & I_q \end{pmatrix}, \tag{1.13}$$

where I_k denotes the identity matrix of size k ; where

$$\begin{aligned}
 W(x) &= \begin{pmatrix} w_{1,1}(x)w_{2,1}(x) & \dots & w_{1,1}(x)w_{2,q}(x) \\ \vdots & & \vdots \\ w_{1,p}(x)w_{2,1}(x) & \dots & w_{1,p}(x)w_{2,q}(x) \end{pmatrix} \\
 &= \begin{pmatrix} w_{1,1}(x) \\ \vdots \\ w_{1,p}(x) \end{pmatrix} \begin{pmatrix} w_{2,1}(x) & \dots & w_{2,q}(x) \end{pmatrix}, \tag{1.14}
 \end{aligned}$$

and where the notation $Y_+(x)$, $Y_-(x)$ denotes the limit of $Y(z)$ with z approaching $x \in \mathbb{R}$ from the upper or lower half plane in \mathbb{C} , respectively;

(3) As $z \rightarrow \infty$, we have that

$$Y(z) = (I_{p+q} + O(1/z))\text{diag}(z^{n_1}, \dots, z^{n_p}, z^{-m_1}, \dots, z^{-m_q}). \tag{1.15}$$

The solution $Y(z)$ to the RH problem is unique, if it exists. Partition this matrix as

$$Y(z) = \begin{matrix} p & q \\ \begin{pmatrix} Y_{1,1}(z) & Y_{1,2}(z) \\ Y_{2,1}(z) & Y_{2,2}(z) \end{pmatrix}, \end{matrix} \tag{1.16}$$

where the partition is such that $Y_{1,1}(z)$ has size $p \times p$, and so on. Then the entries of $Y_{1,1}$ and $Y_{2,1}$ can be described in terms of multiple orthogonal polynomials, while $Y_{1,2}$ and $Y_{2,2}$ contain certain Cauchy transforms thereof; see Section 2 for details.

The RH problem can be used to define the *Christoffel–Darboux kernel*. There are actually two kernels into play. The first is a $q \times p$ matrix-valued kernel $K_n(x, y)$ defined as

$$K_n(x, y) = \frac{1}{2\pi i(y-x)} \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(x)Y(y) \begin{pmatrix} I_p \\ 0 \end{pmatrix}, \tag{1.17}$$

for $x, y \in \mathbb{C}$, $x \neq y$. This definition may seem rather unmotivated; but see Section 2 for the connection with the classical definition of the Christoffel–Darboux kernel.

Second, there is also a scalar-valued kernel $\hat{K}_n(x, y)$ defined by

$$\hat{K}_n(x, y) = \begin{pmatrix} w_{2,1}(x) & \dots & w_{2,q}(x) \end{pmatrix} K_n(x, y) \begin{pmatrix} w_{1,1}(y) \\ \vdots \\ w_{1,p}(y) \end{pmatrix}. \tag{1.18}$$

It is known that the MOP ensemble in Section 1.2 is *determinantal* with *correlation kernel* $\hat{K}_n(x, y)$ [11]. In this paper, however, the quantity of interest will be the matrix-valued kernel $K_n(x, y)$.

1.4. Statement of results

Now we are ready to state our main results. Throughout this section we assume fixed data \mathbf{n} , $\mathbf{w}_1(x)$, \mathbf{m} and $\mathbf{w}_2(x)$ as before, and we consider the corresponding MOP ensemble (1.10)–(1.12).

For any $K, L \in \mathbb{N}$ we define $P_n^{[K,L]}$ as

$$\begin{aligned}
 &P_n^{[K,L]}(y_1, \dots, y_K; z_1, \dots, z_L) \\
 &:= \frac{1}{Z_n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{\prod_{k=1}^K \left(\prod_{j=1}^n (y_k - x_j) \right)}{\prod_{k=1}^L \left(\prod_{j=1}^n (z_k - x_j) \right)} \det (f_i(x_j))_{i,j=1}^n \\
 &\quad \times \det (g_i(x_j))_{i,j=1}^n dx_1 \dots dx_n,
 \end{aligned} \tag{1.19}$$

i.e., $P_n^{[K,L]}$ is the average with respect to the MOP ensemble (1.10)–(1.12) of the ratio of products of characteristic polynomials. Here we assume that $y_1, \dots, y_K \in \mathbb{C}, z_1, \dots, z_L \in \mathbb{C} \setminus \mathbb{R}$ and the numbers in the set $(y_1, \dots, y_K, z_1, \dots, z_L)$ are pairwise distinct.

Note that for $K = 1$ and $L = 0$, (1.19) reduces to the average characteristic polynomial (1.5), while for $K = 0$ and $L = 1$ it reduces to the average inverse characteristic polynomial (1.7). These are the cases under consideration in our first two main theorems.

Theorem 1.2 (Average Characteristic Polynomial). *We have that*

$$P_n^{[1,0]}(y) = \det Y_{1,1}(y), \tag{1.20}$$

according to the partition (1.16).

Theorem 1.3 (Average Inverse Characteristic Polynomial). *We have that*

$$P_n^{[0,1]}(z) = \det Y_{2,2}(z), \tag{1.21}$$

according to the partition (1.16).

The proofs of Theorem 1.2–1.3 will be given in Section 3.2.

Remark 1.4. (a) One may understand Theorem 1.2 as follows. The quantity $\det Y_{1,1}(y)$ is a monic polynomial of degree n , and its zeros determine the ‘typical’ point configuration of the MOP ensemble in Section 1.2. By specializing this to the MOP ensembles in the introduction, one obtains the typical eigenvalue configuration of random matrices with external source, the typical positions of non-intersecting Brownian motions, etc. However, one should be careful with these interpretations. For one thing, we do not even now if all the zeros of $\det Y_{1,1}(y)$ are real. This is an open problem.

(b) The functions $P_n^{[K,L]}$ in (1.19) are clearly invariant under the involution $(\mathbf{n}, \mathbf{w}_1) \leftrightarrow (\mathbf{m}, \mathbf{w}_2)$. The similar fact for the right hand sides of (1.20) and (1.21) is discussed in Section 2, see (2.12)–(2.13).

As we mentioned before, Theorem 1.2 generalizes a classical result for orthogonal polynomials, see e.g. [12], as well as its generalization to type II MOP due to Bleher–Kuijlaars [5]. On the other hand, Theorem 1.3 generalizes a result for orthogonal polynomials due to Fyodorov–Strahov [16], and also its generalization to type I MOP due to Desrosiers–Forrester [13], see also [22].

Now we state our third main theorem.

Theorem 1.5 (Average Ratio of Characteristic Polynomials). *For the case where $K = L = 1$ in (1.19), we have that*

$$P_n^{[1,1]}(y, z) = \det L_n(y, z) \tag{1.22}$$

where L_n is the q by q matrix

$$\begin{aligned} L_n(y, z) &:= I_q - (z - y) \int_{-\infty}^{\infty} K_n(y, x)W(x)/(z - x) \, dx \\ &= \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(y)Y(z) \begin{pmatrix} 0 \\ I_q \end{pmatrix}. \end{aligned} \tag{1.23}$$

Alternatively, we also have

$$P_n^{[1,1]}(y, z) = \det R_n(z, y) \tag{1.24}$$

where R_n is the p by p matrix

$$\begin{aligned} R_n(z, y) &:= I_p - (z - y) \int_{-\infty}^{\infty} W(x)K_n(x, y)/(z - x) \, dx \\ &= \begin{pmatrix} I_p & 0 \end{pmatrix} Y^{-1}(z)Y(y) \begin{pmatrix} I_p \\ 0 \end{pmatrix}. \end{aligned} \tag{1.25}$$

Here $W(x)$ is the weight matrix (1.14) and $K_n(x, y)$ is the Christoffel–Darboux kernel (1.17).

The proof of Theorem 1.5 will be given in Section 3.1.

Note that the matrices $W(x)$ and $K_n(x, y)$ are of size $p \times q$ and $q \times p$, respectively, so each of the matrix expressions in (1.23) and (1.25) is well defined. (The integrals are taken entrywise.) The equalities between these expressions are discussed in Section 2.5.

Corollary 1.6. *Assume that $q = 1$ and $w_{2,1}(x) \equiv 1$. Then*

$$P_n^{[1,1]}(y, z) = 1 - (z - y) \int_{-\infty}^{\infty} \hat{K}_n(y, x)/(z - x) \, dx, \tag{1.26}$$

where \hat{K}_n is the scalar kernel in (1.18). In particular,

$$\hat{K}_n(y, z) = \frac{1}{z - y} \lim_{\epsilon \rightarrow 0} \frac{1}{2\pi i} \left(P_n^{[1,1]}(y, z + \epsilon i) - P_n^{[1,1]}(y, z - \epsilon i) \right), \quad z \in \mathbb{R}. \tag{1.27}$$

Similar results hold when $p = 1$ and $w_{1,1}(x) \equiv 1$.

Formula (1.27) retrieves a result in [13], see also [7] for the scalar case $p = q = 1$.

Finally, let us consider the case of arbitrary K and L in (1.19). To this end we need some extra notation.

Definition 1.7. Assume that $|\mathbf{n}| = |\mathbf{m}|$. We set $(\mathbf{n}_0, \mathbf{m}_0) := (\mathbf{n}, \mathbf{m})$ and we define the sequence of multi-indices $(\mathbf{n}_k, \mathbf{m}_k)$, each \mathbf{n}_k being a vector of length p and each \mathbf{m}_k a vector of length q , for $k \in \mathbb{Z}$ inductively as follows

- For $k = 1, 2, \dots$, we set $\mathbf{n}_k := \mathbf{n}_{k-1} + (1, 1, \dots, 1)$ and we fix \mathbf{m}_k arbitrarily such that $\mathbf{m}_k \geq \mathbf{m}_{k-1}$ componentwise and $|\mathbf{m}_k - \mathbf{m}_{k-1}| = p$.
- For $k = -1, -2, \dots$, we set $\mathbf{m}_k := \mathbf{m}_{k+1} - (1, 1, \dots, 1)$ and we fix \mathbf{n}_k arbitrarily such that $\mathbf{n}_k \leq \mathbf{n}_{k+1}$ componentwise and $|\mathbf{n}_{k+1} - \mathbf{n}_k| = q$.

Definition 1.7 implies that $|\mathbf{n}_k| = |\mathbf{m}_k|$ for each k , so we can consider the RH problem with respect to the pair of multi-indices $(\mathbf{n}_k, \mathbf{m}_k)$. We will assume that the RH problem is solvable for all involved pairs of multi-indices.

Note that in Definition 1.7 the multi-indices $(\mathbf{n}_k, \mathbf{m}_k)$ for negative k are only meaningful provided $|k| \leq \min_{l=1, \dots, q} m_l$. Otherwise we would obtain multi-indices with negative components.

Now we are ready to consider (1.19) for arbitrary K and L . It turns out that the results can be expressed as determinants of large matrices constructed from the ‘building blocks’ in Theorem 1.2–1.5. To get the idea we first describe the three typical cases.

Theorem 1.8. For the expression $P_n^{[K,L]}$ in (1.19), we have

(a) If $L = 0$ then

$$\begin{aligned}
 P_n^{[K,0]}(y_1, \dots, y_K) &= \frac{1}{\prod_{1 \leq i < j \leq K} (y_j - y_i)^p} \\
 &\times \det \begin{pmatrix} (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_K) \\ \vdots & & \vdots \\ (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_K) \end{pmatrix}. \tag{1.28}
 \end{aligned}$$

(b) If $K = 0$ then

$$\begin{aligned}
 P_n^{[0,L]}(z_1, \dots, z_L) &= \frac{1}{\prod_{1 \leq i < j \leq L} (z_j - z_i)^q} \\
 &\times \det \begin{pmatrix} (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_L) \\ \vdots & & \vdots \\ (Y_{2,2})_{\mathbf{n}_{-(L-1)}, \mathbf{m}_{-(L-1)}}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_{-(L-1)}, \mathbf{m}_{-(L-1)}}(z_L) \end{pmatrix}. \tag{1.29}
 \end{aligned}$$

Here we assume that $L \leq 1 + \min_{l=1}^q m_l$.

(c) If $K = L$ then

$$P_n^{[K,K]}(y_1, \dots, y_K, z_1, \dots, z_K) = \left(\frac{\prod_{i,j=1}^K (z_j - y_i)}{\prod_{1 \leq i < j \leq K} (y_j - y_i)(z_i - z_j)} \right)^p \times \det \begin{pmatrix} \frac{1}{z_1 - y_1} R_n(z_1, y_1) & \dots & \frac{1}{z_1 - y_K} R_n(z_1, y_K) \\ \vdots & & \vdots \\ \frac{1}{z_K - y_1} R_n(z_K, y_1) & \dots & \frac{1}{z_K - y_K} R_n(z_K, y_K) \end{pmatrix}, \tag{1.30}$$

where R_n is the matrix in (1.25). A similar result holds with the matrix L_n in (1.23).

Theorem 1.8 generalizes results for the scalar case in [16]. Our proof will be based on the methods of [4].

We note that Parts (a)–(b) of **Theorem 1.8** are of a different flavor than Part (c). In fact, it is shown in [4,16] that in the scalar case, one can obtain ‘mixtures’ of the formulas in Parts (a)–(b); but these formulas do not seem to have an analogue in the present setting. This is the reason why we need to work with the formula of ‘two-point’ type in Part (c).

The case of arbitrary K, L in (1.19) is obtained from mixtures of either Parts (a) and (c), or Parts (b) and (c) of **Theorem 1.8**.

Theorem 1.9. For the expression $P_n^{[K,L]}$ in (1.19), we have

(a) If $K \geq L$ then

$$P_n^{[K,L]}(y_1, \dots, y_K, z_1, \dots, z_L) = \left(\frac{(-1)^{L(K-L)} \prod_{i,j} (z_j - y_i)}{\prod_{1 \leq i < j \leq K} (y_j - y_i) \prod_{1 \leq i < j \leq L} (z_i - z_j)} \right)^p \times \det \begin{pmatrix} \frac{1}{z_1 - y_1} R_n(z_1, y_1) & \dots & \frac{1}{z_1 - y_K} R_n(z_1, y_K) \\ \vdots & & \vdots \\ \frac{1}{z_L - y_1} R_n(z_L, y_1) & \dots & \frac{1}{z_L - y_K} R_n(z_L, y_K) \\ (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_K) \\ \vdots & & \vdots \\ (Y_{1,1})_{\mathbf{n}_{K-L-1}, \mathbf{m}_{K-L-1}}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_{K-L-1}, \mathbf{m}_{K-L-1}}(y_K) \end{pmatrix}. \tag{1.31}$$

Here all R_n matrices are taken with respect to the pair of multi-indices $\mathbf{n}_0, \mathbf{m}_0$.

(b) If $L \geq K$ then

$$\begin{aligned}
 P_n^{[K,L]}(y_1, \dots, y_K, z_1, \dots, z_L) &= \left(\frac{\prod_{i,j} (z_j - y_i)}{\prod_{1 \leq i < j \leq K} (y_i - y_j) \prod_{1 \leq i < j \leq L} (z_j - z_i)} \right)^q \\
 &\times \det \begin{pmatrix} \frac{1}{z_1 - y_1} L_n(y_1, z_1) & \dots & \frac{1}{z_L - y_1} L_n(y_1, z_L) \\ \vdots & & \vdots \\ \frac{1}{z_1 - y_K} L_n(y_K, z_1) & \dots & \frac{1}{z_L - y_K} L_n(y_K, z_L) \\ (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_L) \\ \vdots & & \vdots \\ (Y_{2,2})_{\mathbf{n}_{-(L-K-1)}, \mathbf{m}_{-(L-K-1)}}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_{-(L-K-1)}, \mathbf{m}_{-(L-K-1)}}(z_L) \end{pmatrix}. \tag{1.32}
 \end{aligned}$$

Here all L_n matrices are taken with respect to the pair of multi-indices $\mathbf{n}_0, \mathbf{m}_0$, and we assume that $L - K \leq 1 + \min_{l=1}^q m_l$.

1.5. About the proofs of Theorems 1.5 and 1.8(c)

Theorems 1.5 and 1.8(c) will be the key results from which all other main theorems will follow. The proofs of these two theorems will be based on a Schur complement formula for the kernel $K_n(x, y)$. This formula is valid for an arbitrary weight matrix

$$W(x) = (W_{k,l}(x))_{k=1, \dots, p, l=1, \dots, q} \tag{1.33}$$

of size p by q . Thus $W(x)$ does not need to have the rank-one factorization (1.14). For any $k \in \{1, \dots, p\}$ and $l \in \{1, \dots, q\}$ let

$$\mu_j^{(k,l)} = \int_{-\infty}^{\infty} x^j W_{k,l}(x) dx \in \mathbb{R}$$

denote the j th moment of the weight function $W_{k,l}(x)$, and for any $r, s \in \mathbb{N}$ let

$$H_{r,s}^{(k,l)} = (\mu_{i+j}^{(k,l)})_{i=0, \dots, r-1, j=0, \dots, s-1}$$

be the $r \times s$ Hankel matrix formed out of these moments. Stack these matrices in the block Hankel matrix

$$H_{\mathbf{n}, \mathbf{m}} = \begin{pmatrix} H_{n_1, m_1}^{(1,1)} & \dots & H_{n_1, m_q}^{(1,q)} \\ \vdots & & \vdots \\ H_{n_p, m_1}^{(p,1)} & \dots & H_{n_p, m_q}^{(p,q)} \end{pmatrix}. \tag{1.34}$$

Note that this matrix is of size $n \times n$.

Recall that for a block matrix

$$M := \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{1.35}$$

with A square of size $k \times k$ (say), the *Schur complement* of M with respect to the submatrix D is defined as the matrix

$$S_{M,D} := D - CA^{-1}B, \tag{1.36}$$

provided A is invertible. Schur complements are also known as *quasi-determinants* in the literature [18]. For more information on Schur complements see Section 1.6.

Proposition 1.10. *The kernel $K_n(x, y)$ equals the Schur complement of the matrix*

$$- \left(\begin{array}{ccc|cc} H_{n_1,m_1}^{(1,1)} & \dots & H_{n_1,m_q}^{(1,q)} & \mathbf{y}_{n_1} & \mathbf{0} \\ \vdots & & \vdots & & \ddots \\ H_{n_p,m_1}^{(p,1)} & \dots & H_{n_p,m_q}^{(p,q)} & \mathbf{0} & \mathbf{y}_{n_p} \\ \hline \mathbf{x}_{m_1}^T & & \mathbf{0} & 0 & \dots & 0 \\ & & \ddots & \vdots & & \vdots \\ \mathbf{0} & & \mathbf{x}_{m_q}^T & 0 & \dots & 0 \end{array} \right) \tag{1.37}$$

with respect to its bottom right $q \times p$ submatrix. Here the superscript T denotes the matrix transpose, and we use the column vector notations $\mathbf{y}_m = (1 \ y \ \dots \ y^{m-1})^T$ and $\mathbf{x}_m = (1 \ x \ \dots \ x^{m-1})^T$.

Proposition 1.10 will be proved in Section 2.5. Variants of this proposition in different contexts can be found in [6,23], see also [3].

Proposition 1.10 will allow the quantities $\det L_n(y, z)$ and $\det R_n(z, y)$ in Theorem 1.5 to be written as a ratio of two determinants. In Section 3.1 we will expand these determinants using an adaptation of the classical argument of Heine, and this will lead us to the proof of Theorem 1.5. In Section 3.3 we will give a similar argument proving Theorem 1.8(c).

1.6. Background on Schur complements

Since Schur complements frequently occur in this paper, we find it convenient to list here some basic properties.

Recall the setting in (1.35) and (1.36). Schur complements are related to Gaussian elimination as follows:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \sim \begin{pmatrix} A & 0 \\ 0 & S_{M,D} \end{pmatrix}, \tag{1.38}$$

where the \sim symbol relates matrices that can be obtained from one another by multiplying on the left and on the right with square transformation matrices of the form $\begin{pmatrix} I_k & 0 \\ * & I \end{pmatrix}$ and $\begin{pmatrix} I_k & * \\ 0 & I \end{pmatrix}$, respectively. The procedure to move from the left to the right hand side of (1.38) is sometimes called *Gaussian elimination with pivot block A* , see [19].

In the special case where D and (hence) $S_{M,D}$ are square matrices, one obtains from (1.38) the determinant relation

$$\det S_{M,D} = \frac{\det M}{\det A}. \tag{1.39}$$

Let $s_{i,j}$ denote the (i, j) th entry of $S = S_{M,D}$. By (1.36) one has

$$s_{i,j} = d_{i,j} - \mathbf{c}_i A^{-1} \mathbf{b}_j, \quad (1.40)$$

where \mathbf{c}_i denotes the i th row of C and \mathbf{b}_j denotes the j th column of B . Now (1.40) is nothing but the Schur complement of the matrix

$$\begin{pmatrix} A & \mathbf{b}_j \\ \mathbf{c}_i & d_{i,j} \end{pmatrix} \quad (1.41)$$

with respect to the entry $d_{i,j}$. From (1.39) we then obtain the determinantal expression

$$s_{i,j} = \frac{\det \begin{pmatrix} A & \mathbf{b}_j \\ \mathbf{c}_i & d_{i,j} \end{pmatrix}}{\det A}. \quad (1.42)$$

This shows that the entries of the Schur complement are ratios of determinants.

More generally than (1.41), one may observe that Schur complements are well behaved with respect to taking submatrices, in the sense that any submatrix of the Schur complement (1.36) is itself a Schur complement, of an appropriate submatrix of (1.35).

Finally, observe that for any matrix U of appropriate size, the matrix product $US_{M,D} = U(D - CA^{-1}B)$ can again be written as a Schur complement, of the matrix

$$\begin{pmatrix} A & B \\ UC & UD \end{pmatrix}. \quad (1.43)$$

A similar fact holds for matrix products of the form $S_{M,D}V$.

1.7. Outline of the rest of the paper

The rest of this paper is organized as follows. In Section 2 we discuss auxiliary results and notations that are used in the proofs of the main theorems, and we prove Proposition 1.10. Theorems 1.5 and 1.8(c) are then proved in Section 3, by an adaptation of the classical argument of Heine. Theorems 1.2 and 1.3 are obtained as limiting cases of Theorem 1.5. Finally, the generalizations to arbitrary products and ratios of characteristic polynomials are proved in Section 4.

2. Preliminaries for the proofs

In this section we collect some auxiliary results and notations that are used in the proofs of the main theorems. This section is organized as follows. Vector orthogonal polynomials (which include multiple orthogonal polynomials as a special case) are defined in Section 2.1 and their relation to the RH problem in Section 2.2. The connection with block Hankel determinants and the duality relations are discussed in Sections 2.3 and 2.4. In Section 2.5 we discuss the Christoffel–Darboux kernel, leading to the proof of Proposition 1.10. Finally, the related quantities L_n and R_n are investigated in Section 2.6.

Remark 2.1. Practical note: the reader who is mainly interested in Theorems 1.2, 1.3 and 1.5 and who wants to take Proposition 1.10 for granted may skip this entire section and move directly to Section 3.

2.1. Vector orthogonal polynomials

Let $p, q \in \mathbb{N}$ and consider the multi-indices \mathbf{n}, \mathbf{m} as in the beginning of Section 1.2, but now with $|\mathbf{n}| = |\mathbf{m}| + 1$. Let $W(x)$ be a p by q weight matrix as in (1.33). Our definition of vector orthogonal polynomials will be the same as in Sorokin–Van Iseghem [25]; this definition includes multiple orthogonal polynomials as a special case, when the weight matrix $W(x)$ has the rank-one factorization (1.14).

Let $\mathcal{P}_{\mathbf{n}}$ be the following space of polynomial vectors

$$\mathcal{P}_{\mathbf{n}} = \left\{ \mathbf{P}(x) = \begin{pmatrix} P_1(x) \\ \vdots \\ P_p(x) \end{pmatrix} \text{ with } P_k(x) \text{ a polynomial of degree at most } n_k - 1 \right\}. \tag{2.1}$$

The standard basis $(\mathbf{A}_1(x), \dots, \mathbf{A}_{|\mathbf{n}|}(x))$ of the vector space $\mathcal{P}_{\mathbf{n}}$ is defined by the column vectors of the p by $|\mathbf{n}|$ matrix

$$\begin{pmatrix} \mathbf{A}_1(x) & \dots & \mathbf{A}_{|\mathbf{n}|}(x) \end{pmatrix} = \begin{pmatrix} 1 & x & \dots & x^{n_1-1} & | & 0 & 0 & \dots & 0 & | & \dots & 0 \\ 0 & 0 & \dots & 0 & | & 1 & x & \dots & x^{n_2-1} & | & \dots & 0 \\ \vdots & \vdots & & \vdots & | & \vdots & \vdots & & \vdots & | & \dots & \vdots \\ 0 & 0 & \dots & 0 & | & 0 & 0 & \dots & 0 & | & \dots & x^{n_p-1} \end{pmatrix}. \tag{2.2}$$

Similarly, let $\mathcal{P}_{\mathbf{m}}$ denote the vector space

$$\mathcal{P}_{\mathbf{m}} = \left\{ \mathbf{Q}(x) = \begin{pmatrix} Q_1(x) \\ \vdots \\ Q_q(x) \end{pmatrix} \text{ with } Q_l(x) \text{ a polynomial of degree at most } m_l - 1 \right\}, \tag{2.3}$$

and define the standard basis $(\mathbf{B}_1(x), \dots, \mathbf{B}_{|\mathbf{m}|}(x))$ for $\mathcal{P}_{\mathbf{m}}$ by the columns of the q by $|\mathbf{m}|$ matrix

$$\begin{pmatrix} \mathbf{B}_1(x) & \dots & \mathbf{B}_{|\mathbf{m}|}(x) \end{pmatrix} = \begin{pmatrix} 1 & x & \dots & x^{m_1-1} & | & 0 & 0 & \dots & 0 & | & \dots & 0 \\ 0 & 0 & \dots & 0 & | & 1 & x & \dots & x^{m_2-1} & | & \dots & 0 \\ \vdots & \vdots & & \vdots & | & \vdots & \vdots & & \vdots & | & \dots & \vdots \\ 0 & 0 & \dots & 0 & | & 0 & 0 & \dots & 0 & | & \dots & x^{m_q-1} \end{pmatrix}. \tag{2.4}$$

Note that we use boldface notation to denote vector-valued objects.

We say that $\mathbf{P}(x) = \mathbf{P}_{\mathbf{n},\mathbf{m}}(x) \in \mathcal{P}_{\mathbf{n}}$ is a vector orthogonal polynomial with respect to the multi-indices \mathbf{n}, \mathbf{m} and the weight matrix $W(x)$ if

$$\int_{-\infty}^{\infty} \mathbf{P}_{\mathbf{n},\mathbf{m}}^T(x) W(x) \mathbf{Q}(x) dx = 0$$

for every $\mathbf{Q} \in \mathcal{P}_{\mathbf{m}}$. To stress the dependence on the weight matrix we will sometimes write $\mathbf{P}_{\mathbf{n},\mathbf{m}}(x) = \mathbf{P}_{\mathbf{n},\mathbf{m},W}(x)$.

The coefficients of the vector orthogonal polynomials can be found from a homogeneous linear system with $|\mathbf{n}|$ unknowns (polynomial coefficients) and $|\mathbf{m}|$ equations (orthogonality conditions). The restriction $|\mathbf{n}| = |\mathbf{m}| + 1$ guarantees that this system has a nontrivial solution.

If $\mathbf{P}_{\mathbf{n},\mathbf{m}}(x)$ is unique up to a multiplicative factor then the pair of multi-indices \mathbf{n}, \mathbf{m} is called *normal*.

Assume that \mathbf{n}, \mathbf{m} is a normal pair of indices. Then $\mathbf{P}_{\mathbf{n},\mathbf{m}}(x)$ is said to satisfy

- the *normalization of type I with respect to the l th index*, $l \in \{1, \dots, q\}$, if

$$\int_{-\infty}^{\infty} \mathbf{P}_{\mathbf{n},\mathbf{m}}^T(x) W(x) \mathbf{Q}(x) dx = 1$$

where $\mathbf{Q}(x)$ is the vector consisting of zeros except for the l th entry which is x^{m_l} .

- the *normalization of type II with respect to the k th index*, $k \in \{1, \dots, p\}$, if the k th component $(P_k)_{\mathbf{n},\mathbf{m}}(x)$ is monic, i.e., if

$$(P_k)_{\mathbf{n},\mathbf{m}}(x) = x^{n_k-1} + O(x^{n_k-2}).$$

The vector orthogonal polynomials $\mathbf{P}_{\mathbf{n},\mathbf{m}}(x)$ corresponding to the above normalizations, if they exist, will be denoted as $\mathbf{P}_{\mathbf{n},\mathbf{m}}^{(I,l)}(x)$ and $\mathbf{P}_{\mathbf{n},\mathbf{m}}^{(II,k)}(x)$, respectively.

2.2. Solution to the RH problem

Consider again the RH problem in Section 1.3 with $|\mathbf{n}| = |\mathbf{m}|$ but now with $W(x)$ in (1.14) replaced by an arbitrary matrix (1.33). The solution to the RH problem, if it exists, is uniquely described by vector orthogonal polynomials and their Cauchy transforms [11]. More precisely, consider again the partition of $Y(z)$ as in (1.16). Then one has that

- $Y_{1,1}(z)$ has its k th row given by the row vector

$$(\mathbf{P}_{\mathbf{n}+\mathbf{e}_k,\mathbf{m}}^{(II,k)})^T(z), \tag{2.5}$$

$k = 1, \dots, p$.

- $Y_{2,1}(z)$ has its l th row given by

$$-2\pi i (\mathbf{P}_{\mathbf{n},\mathbf{m}-\mathbf{e}_l}^{(I,l)})^T(z), \tag{2.6}$$

$l = 1, \dots, q$.

- $Y_{1,2}(z)$ has its k th row given by

$$-\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{(\mathbf{P}_{\mathbf{n}+\mathbf{e}_k,\mathbf{m}}^{(II,k)})^T(x) W(x)}{z-x} dx, \tag{2.7}$$

$k = 1, \dots, p$.

- $Y_{2,2}(z)$ has its l th row given by

$$\int_{-\infty}^{\infty} \frac{(\mathbf{P}_{\mathbf{n},\mathbf{m}-\mathbf{e}_l}^{(I,l)})^T(x) W(x)}{z-x} dx, \tag{2.8}$$

$l = 1, \dots, q$.

Here we are using the notation \mathbf{e}_l to denote the standard basis vector which is zero except for its l th entry which is one. The length of this vector will always be clear from the context.

For example, in the special case where $p = q = 2$, the solution $Y(z)$ is given by the 4×4 matrix

$$Y(z) = D \times \begin{pmatrix} (P_1)_{\mathbf{n}+\mathbf{e}_1, \mathbf{m}}^{(II,1)} & (P_2)_{\mathbf{n}+\mathbf{e}_1, \mathbf{m}}^{(II,1)} & * & * \\ (P_1)_{\mathbf{n}+\mathbf{e}_2, \mathbf{m}}^{(II,2)} & (P_2)_{\mathbf{n}+\mathbf{e}_2, \mathbf{m}}^{(II,2)} & * & * \\ (P_1)_{\mathbf{n}, \mathbf{m}-\mathbf{e}_1}^{(I,1)} & (P_2)_{\mathbf{n}, \mathbf{m}-\mathbf{e}_1}^{(I,1)} & * & * \\ (P_1)_{\mathbf{n}, \mathbf{m}-\mathbf{e}_2}^{(I,2)} & (P_2)_{\mathbf{n}, \mathbf{m}-\mathbf{e}_2}^{(I,2)} & * & * \end{pmatrix},$$

where $D := \text{diag}(1, 1, -2\pi i, -2\pi i)$, and where the entries denoted with $*$ are Cauchy transforms as in (2.7) and (2.8).

2.3. Biorthogonality and moment matrix

Recall the standard bases $\mathbf{A}_i(x)$ and $\mathbf{B}_j(x)$ in (2.2) and (2.4), where now $|\mathbf{n}| = |\mathbf{m}| =: n$. The moment matrix M with respect to these bases is of size n by n and has entries

$$M_{i,j} = \int_{-\infty}^{\infty} \mathbf{A}_i^T(x) W(x) \mathbf{B}_j(x) dx.$$

The moment matrix coincides with the block Hankel matrix $H_{\mathbf{n}, \mathbf{m}}$ in (1.34).

Consider ordered bases $(\mathbf{P}_1(x), \dots, \mathbf{P}_n(x))$ and $(\mathbf{Q}_1(x), \dots, \mathbf{Q}_n(x))$ of the vector spaces $\mathcal{P}_{\mathbf{n}}$ and $\mathcal{P}_{\mathbf{m}}$, respectively. We say that these bases are *biorthogonal* if

$$\int_{-\infty}^{\infty} \mathbf{P}_i^T(x) W(x) \mathbf{Q}_j(x) dx = \delta_{i,j}, \tag{2.9}$$

the Kronecker delta.

Lemma 2.2. *Assume that $|\mathbf{n}| = |\mathbf{m}|$. Then the following statements are equivalent*

1. *The RH problem for $Y(z)$ in Section 1.3 is solvable.*
2. *There is no non-zero vector in $\mathcal{P}_{\mathbf{n}}$ which is biorthogonal to the entire space $\mathcal{P}_{\mathbf{m}}$.*
3. *$\det H_{\mathbf{n}, \mathbf{m}} \neq 0$.*
4. *There exist biorthogonal bases for the spaces $\mathcal{P}_{\mathbf{n}}$ and $\mathcal{P}_{\mathbf{m}}$.*

Proof. First consider the equivalence between Statements 1 and 2. Suppose that statement 2 holds true. Then each of the polynomial vectors in (2.5)–(2.6) exists [11] and therefore the RH problem for $Y(z)$ is solvable. Conversely, suppose that statement 2 does not hold. Then the polynomial vectors in (2.5)–(2.6), if they exist, are not unique and therefore the RH problem for $Y(z)$ cannot be solvable, since that would contradict the uniqueness of $Y(z)$. Finally, the equivalence between Statements 2–4 follows from standard linear algebra arguments whose description we omit. \square

2.4. Duality

The role of the biorthogonal bases \mathbf{P}_i and \mathbf{Q}_j is interchanged by swapping the multi-indices \mathbf{n} and \mathbf{m} and by transposing the weight matrix $W(x)$:

$$(\mathbf{n}, \mathbf{m}, W(x)) \leftrightarrow (\mathbf{m}, \mathbf{n}, W^T(x)). \tag{2.10}$$

This duality has a nice form in terms of the RH problem [1,11]. If we partition the RH matrices corresponding to the original and dual data in (2.10) as

$$Y_{\mathbf{n},\mathbf{m},W} = \begin{matrix} p & q \\ \left(\begin{array}{cc} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{array} \right), \end{matrix} \quad Y_{\mathbf{m},\mathbf{n},W^T} = \begin{matrix} q & p \\ \left(\begin{array}{cc} \tilde{Y}_{1,1} & \tilde{Y}_{1,2} \\ \tilde{Y}_{2,1} & \tilde{Y}_{2,2} \end{array} \right), \end{matrix}$$

respectively, then it holds that

$$\begin{pmatrix} \tilde{Y}_{1,1} & \tilde{Y}_{1,2} \\ \tilde{Y}_{2,1} & \tilde{Y}_{2,2} \end{pmatrix} = \begin{pmatrix} Y_{2,2} & -Y_{2,1} \\ -Y_{1,2} & Y_{1,1} \end{pmatrix}^{-T}, \tag{2.11}$$

where the superscript $-T$ denotes the inverse transpose. From a well-known theorem on minors of the inverse matrix, see e.g. [17, page. 21, Eq. (33)], and the general fact that $\det Y(z) \equiv 1$ it then follows that

$$\det \tilde{Y}_{1,1}(y) = \det Y_{1,1}(y) \tag{2.12}$$

and

$$\det \tilde{Y}_{2,2}(z) = \det Y_{2,2}(z). \tag{2.13}$$

The duality relations (2.12)–(2.13) take care of a point made earlier, see the remark after the statement of Theorem 1.3.

From another application of [17, page. 21, Eq. (33)] it also follows that

$$\det \left(\begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix} Y^{-1}(z)Y(y) \begin{pmatrix} I_p \\ 0 \end{pmatrix} \right) = \det \left(\begin{pmatrix} 0 & I_q \\ I_p & 0 \end{pmatrix} Y^{-1}(y)Y(z) \begin{pmatrix} 0 \\ I_q \end{pmatrix} \right). \tag{2.14}$$

Comparing this with (1.23) and (1.25), this establishes the fact that $\det R_n(z, y) = \det L_n(y, z)$, which is of course implicit in the statement of Theorem 1.5.

2.5. The Christoffel–Darboux kernel

In this section we prove the Schur complement formula for the kernel $K_n(x, y)$ in Proposition 1.10. To this end we identify $K_n(x, y)$ as a reproducing kernel. In this way we will show that (1.17) corresponds to the kernels familiar from the theory of matrix orthogonal polynomials [23] and vector orthogonal polynomials [25].

We will assume throughout that the RH problem for $Y(z)$ is solvable, which is of course necessary in order for (1.17) to make sense.

Lemma 2.3. *The kernel $K_n(x, y)$ is a $q \times p$ matrix whose (i, j) entry is a bivariate polynomial of degree at most $m_i - 1$ in x and $n_j - 1$ in y . Equivalently,*

$$K_n(x, y) = \sum_{i,j=1}^n c_{i,j} \mathbf{B}_i(x) \mathbf{A}_j^T(y), \tag{2.15}$$

for certain $c_{i,j}$, where \mathbf{A}_j and \mathbf{B}_i denote the basis vectors in (2.2) and (2.4).

Proof. The expression $\begin{pmatrix} 0 & I_q \\ I_p & 0 \end{pmatrix} Y^{-1}(x)Y(y) \begin{pmatrix} I_p \\ 0 \end{pmatrix}$ in the numerator of (1.17) is made out of the polynomial entries of the matrices $Y^{-1}(x)$ and $Y(y)$, cf. (2.5)–(2.6) and (2.11). Since this

expression vanishes if $x = y$, each of these polynomials is divisible by $y - x$, so the denominator of (1.17) can be divided out. \square

It was shown in [11] that for a rank-one weight matrix $W(x)$ as in (1.14), the scalar-valued kernel $\hat{K}_n(x, y)$ in (1.18) has a certain reproducing property. We now obtain a similar property for the matrix-valued kernel $K_n(x, y)$ (without restrictions on $W(x)$).

Proposition 2.4. *The kernel K_n satisfies the reproducing property*

$$\int_{-\infty}^{\infty} K_n(x, y)W(y)\mathbf{Q}(y) \, dy = \mathbf{Q}(x) \tag{2.16}$$

for any $x \in \mathbb{C}$ and $\mathbf{Q} \in \mathcal{P}_m$. Moreover, this reproducing property uniquely characterizes $K_n(x, y)$ in the class of bivariate polynomial matrices of the form (2.15).

Proof. First we establish (2.16). This can be done by adapting the proof in [11]. However, we give a more streamlined proof. We have

$$\begin{aligned} \int K_n(x, y)W(y)\mathbf{Q}(y) \, dy &= \int K_n(x, y)W(y)(\mathbf{Q}(y) - \mathbf{Q}(x)) \, dy \\ &\quad + \int K_n(x, y)W(y)\mathbf{Q}(x) \, dy. \end{aligned}$$

By inserting (1.17) this becomes

$$\begin{aligned} &= \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(x) \left(\begin{aligned} &\frac{1}{2\pi i} \int Y_{1,1}(y)W(y) \frac{\mathbf{Q}(y) - \mathbf{Q}(x)}{y - x} \, dy \\ &\frac{1}{2\pi i} \int Y_{2,1}(y)W(y) \frac{\mathbf{Q}(y) - \mathbf{Q}(x)}{y - x} \, dy \end{aligned} \right) \\ &\quad + \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(x) \left(\begin{aligned} &\frac{1}{2\pi i} \int \frac{Y_{1,1}(y)W(y)}{y - x} \, dy \\ &\frac{1}{2\pi i} \int \frac{Y_{2,1}(y)W(y)}{y - x} \, dy \end{aligned} \right) \mathbf{Q}(x). \end{aligned} \tag{2.17}$$

The first term in (2.17) is zero since $\frac{\mathbf{Q}(y) - \mathbf{Q}(x)}{y - x}$ is a polynomial vector in y whose l th entry has degree at most $m_l - 2$, $l = 1, \dots, q$, and by invoking the orthogonality relations of the vector orthogonal polynomials. For the second term in (2.17), we recognize the defining relations for the Cauchy transforms in the second block column of the RH problem, see (2.7)–(2.8), so we get

$$= \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(x) \begin{pmatrix} Y_{1,2}(x) \\ Y_{2,2}(x) \end{pmatrix} \mathbf{Q}(x) = \begin{pmatrix} 0 & I_q \end{pmatrix} \begin{pmatrix} 0 \\ I_q \end{pmatrix} \mathbf{Q}(x) = \mathbf{Q}(x).$$

This establishes (2.16) when $x \in \mathbb{C} \setminus \mathbb{R}$; the case where $x \in \mathbb{R}$ follows by continuity.

Next we show that the reproducing property (2.16) uniquely characterizes $K_n(x, y)$ in the class of bivariate polynomial matrices (2.15). By Lemma 2.2, $1 \Rightarrow 4$, we can choose biorthogonal bases for the polynomial vector spaces \mathcal{P}_n and \mathcal{P}_m ; let $(\mathbf{P}_k)_{k=1}^n$ and $(\mathbf{Q}_k)_{k=1}^m$ be such bases. Any q by p matrix $M_n(x, y)$ of the form (2.15) can be rewritten as

$$M_n(x, y) = \sum_{i,j=1}^n \tilde{c}_{i,j} \mathbf{Q}_i(x) \mathbf{P}_j^T(y)$$

for suitable constants $\tilde{c}_{i,j}$. But then the reproducing property (2.16) and the biorthogonality relations (2.9) imply that $\tilde{c}_{i,j} = \delta_{i,j}$, the Kronecker delta. This ends the proof. \square

Proposition 2.4 has the following dual version.

Proposition 2.5. *The kernel K_n satisfies the reproducing property*

$$\int_{-\infty}^{\infty} \mathbf{P}^T(x)W(x)K_n(x, y) dx = \mathbf{P}^T(y) \tag{2.18}$$

for any $y \in \mathbb{C}$ and $\mathbf{P} \in \mathcal{P}_n$. Moreover, this reproducing property uniquely characterizes $K_n(x, y)$ in the class of bivariate polynomial matrices of the form (2.15).

Now we are ready for the

Proof of Proposition 1.10. Denote with $M_n(x, y)$ the Schur complement of (1.37). We check the reproducing kernel property

$$\int \mathbf{A}^T(x)W(x)M_n(x, y) dx = \mathbf{A}^T(y), \tag{2.19}$$

for each column vector of the form $\mathbf{A}(x) = (0 \dots x^{i_k} \dots 0)^T$ with $i_k \in \{0, \dots, n_k - 1\}$, $k \in \{1, \dots, p\}$. (These vectors $\mathbf{A}(x)$ are the standard basis of the space \mathcal{P}_n .) By using the multiplication property (1.43) for Schur complements, the left hand side of (2.19) equals the Schur complement of

$$- \left(\begin{array}{ccc|cc} H_{n_1, m_1}^{(1,1)} & \dots & H_{n_1, m_q}^{(1,q)} & \mathbf{y}_{n_1} & \mathbf{0} \\ \vdots & & \vdots & & \\ H_{n_p, m_1}^{(p,1)} & \dots & H_{n_p, m_q}^{(p,q)} & \mathbf{0} & \mathbf{y}_{n_p} \\ \hline \mathbf{h}_{i_k, m_1}^{(k,1)} & & \mathbf{h}_{i_k, m_q}^{(k,q)} & 0 & \dots & 0 \end{array} \right) \tag{2.20}$$

with respect to its bottom right $1 \times p$ submatrix, where we use the row vector notation $\mathbf{h}_{i_k, m_l}^{(k,l)} = (\mu_{i_k+j}^{(k,l)})_{j=0, \dots, m_l-1}$. Subtracting from the last row of (2.20) the i_k th row from the k th block row, we are led to the Schur complement of

$$- \left(\begin{array}{ccc|cc} H_{n_1, m_1}^{(1,1)} & \dots & H_{n_1, m_q}^{(1,q)} & \mathbf{y}_{n_1} & \mathbf{0} \\ \vdots & & \vdots & & \\ H_{n_p, m_1}^{(p,1)} & \dots & H_{n_p, m_q}^{(p,q)} & \mathbf{0} & \mathbf{y}_{n_p} \\ \hline \mathbf{0} & & \mathbf{0} & 0 & \dots & -y^{i_k} & \dots & 0 \end{array} \right), \tag{2.21}$$

where now the last row is zero except for the entry $-y^{i_k}$ in the k th column of the second block column. But clearly, the Schur complement of (2.21) is just the row vector $\mathbf{A}^T(y) = (0 \dots y^{i_k} \dots 0)$. This establishes the reproducing kernel property (2.19).

To finish the proof of the proposition, we note that the degree structure of (1.37) implies that $M_n(x, y)$ has the form in (2.15). Thus the proof is ended by invoking the uniqueness part of Proposition 2.5. \square

2.6. The matrices L_n and R_n

In this section we consider in more detail the quantities L_n and R_n which are derived from the kernel K_n . Let us first establish the equivalence between the two different formulas in the

definition of L_n in (1.23). From the identity $y - x = (y - z) + (z - x)$ we obtain

$$\begin{aligned} \int \frac{y-x}{z-x} K_n(y, x) W(x) dx &= \int K_n(y, x) W(x) I_q dx + \int \frac{y-z}{z-x} K_n(y, x) W(x) dx \\ &= I_q - (z-y) \int K_n(y, x) W(x)/(z-x) dx \end{aligned} \tag{2.22}$$

by virtue of the reproducing property in Proposition 2.4. On the other hand, the equality

$$\int \frac{y-x}{z-x} K_n(y, x) W(x) dx = \begin{pmatrix} 0 & I_q \end{pmatrix} Y^{-1}(y) Y(z) \begin{pmatrix} 0 \\ I_q \end{pmatrix}$$

follows from (1.17) and (2.7)–(2.8). This establishes the required equality in (1.23).

The equivalence between the different formulas in (1.25) can be established similarly.

In Section 4, we will need the following property of L_n .

Proposition 2.6. *The matrix L_n in (1.23) satisfies the ‘vanishing property’*

$$\int_{-\infty}^{\infty} \mathbf{P}^T(y) W(y) \frac{L_n(y, z)}{z-y} dy = 0 \tag{2.23}$$

for any $z \in \mathbb{C} \setminus \mathbb{R}$ and $\mathbf{P} \in \mathcal{P}_n$.

Proof. From the expression of L_n in (2.22), we obtain

$$\begin{aligned} &\int \mathbf{P}^T(y) W(y) \frac{L_n(y, z)}{z-y} dy \\ &= \int \mathbf{P}^T(y) \frac{W(y)}{z-y} dy - \int \int \frac{\mathbf{P}^T(y) W(y) K_n(y, x) W(x)}{z-x} dx dy \end{aligned}$$

which is

$$= \int \mathbf{P}^T(y) \frac{W(y)}{z-y} dy - \int \left(\int \mathbf{P}^T(y) W(y) K_n(y, x) dy \right) \frac{W(x)}{z-x} dx.$$

By inserting the reproducing property this becomes

$$= \int \mathbf{P}^T(y) \frac{W(y)}{z-y} dy - \int \mathbf{P}^T(x) \frac{W(x)}{z-x} dx = 0. \quad \square$$

Here is the dual version of Proposition 2.6.

Proposition 2.7. *The matrix R_n in (1.25) satisfies the ‘vanishing property’*

$$\int_{-\infty}^{\infty} \frac{R(z, y)}{z-y} W(y) \mathbf{Q}(y) dy = 0 \tag{2.24}$$

for any $z \in \mathbb{C} \setminus \mathbb{R}$ and $\mathbf{Q} \in \mathcal{P}_m$.

3. Proofs, part 1

In this section we prove Theorems 1.2, 1.3, 1.5 and 1.8(c).

3.1. Proof of Theorem 1.5

First we prove Theorem 1.5. The proof of this theorem will follow by expanding the moment determinant in Proposition 1.10 in a similar way as in the classical argument of Heine.

We will restrict ourselves to the proof of (1.24); the proof of (1.22) can then be devised in a similar way, or by simply invoking (2.14). For notational convenience and to keep things readable, we give the proof for the case $p = q = 2$. The case of general p, q is completely similar except that it requires more notational burden.

First we expand the left hand side of (1.24). By definition, the average ratio of characteristic polynomials $P_n^{[1,1]}(y, z)$ is given by

$$P_n^{[1,1]}(y, z) = \frac{1}{Z_n} \int \dots \int \prod_{j=1}^n \frac{y - x_j}{z - x_j} \left| \begin{array}{ccc} w_{1,1}(x_1) & \dots & w_{1,1}(x_n) \\ \vdots & & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \dots & x_n^{n_1-1} w_{1,1}(x_n) \\ w_{1,2}(x_1) & \dots & w_{1,2}(x_n) \\ \vdots & & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \dots & x_n^{n_2-1} w_{1,2}(x_n) \end{array} \right| \\
 \times \left| \begin{array}{ccc} w_{2,1}(x_1) & \dots & w_{2,1}(x_n) \\ \vdots & & \vdots \\ x_1^{m_1-1} w_{2,1}(x_1) & \dots & x_n^{m_1-1} w_{2,1}(x_n) \\ w_{2,2}(x_1) & \dots & w_{2,2}(x_n) \\ \vdots & & \vdots \\ x_1^{m_2-1} w_{2,2}(x_1) & \dots & x_n^{m_2-1} w_{2,2}(x_n) \end{array} \right| \prod_{j=1}^n dx_j. \tag{3.1}$$

We can expand the second determinant according to the Lagrange expansion with respect to the first block row:

$$\left| \begin{array}{ccc} w_{2,1}(x_1) & \dots & w_{2,1}(x_n) \\ \vdots & & \vdots \\ x_1^{m_1-1} w_{2,1}(x_1) & \dots & x_n^{m_1-1} w_{2,1}(x_n) \\ w_{2,2}(x_1) & \dots & w_{2,2}(x_n) \\ \vdots & & \vdots \\ x_1^{m_2-1} w_{2,2}(x_1) & \dots & x_n^{m_2-1} w_{2,2}(x_n) \end{array} \right| \\
 = \sum_S (-1)^S \left| \begin{array}{ccc} w_{2,1}(x_{s_1}) & \dots & w_{2,1}(x_{s_{m_1}}) \\ \vdots & & \vdots \\ x_{s_1}^{m_1-1} w_{2,1}(x_{s_1}) & \dots & x_{s_{m_1}}^{m_1-1} w_{2,1}(x_{s_{m_1}}) \end{array} \right| \\
 \times \left| \begin{array}{ccc} w_{2,2}(x_{\bar{s}_1}) & \dots & w_{2,2}(x_{\bar{s}_{m_2}}) \\ \vdots & & \vdots \\ x_{\bar{s}_1}^{m_2-1} w_{2,2}(x_{\bar{s}_1}) & \dots & x_{\bar{s}_{m_2}}^{m_2-1} w_{2,2}(x_{\bar{s}_{m_2}}) \end{array} \right| \tag{3.2}$$

where the sum is over all $\binom{n}{m_1}$ subsets $S \subset \{1, \dots, n\}$ with $|S| = m_1$, where we write the elements of S in increasing order: $s_1 < \dots < s_{m_1}$, similarly for those of the complement $\bar{S} = \{1, \dots, n\} \setminus S$, $\bar{s}_1 < \dots < \bar{s}_{m_2}$, and where we denote by $(-1)^S$ the sign of the permutation $s_1, \dots, s_{m_1}, \bar{s}_1, \dots, \bar{s}_{m_2}$.

For each term of (3.2) we can incorporate the factor $(-1)^S$ into the other determinant of (3.1) by permuting its columns. Next we can relabel the variables x_1, \dots, x_n back in their original form, and so (3.1) reduces to

$$\begin{aligned}
 P_n^{[1,1]}(y, z) &= \frac{\binom{n}{m_1}}{Z_n} \int \dots \int \prod_{j=1}^n \frac{y - x_j}{z - x_j} \left| \begin{array}{ccc} w_{1,1}(x_1) & \dots & w_{1,1}(x_n) \\ \vdots & & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \dots & x_n^{n_1-1} w_{1,1}(x_n) \\ \hline w_{1,2}(x_1) & \dots & w_{1,2}(x_n) \\ \vdots & & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \dots & x_n^{n_2-1} w_{1,2}(x_n) \end{array} \right| \\
 &\times \left| \begin{array}{ccc} w_{2,1}(x_1) & \dots & w_{2,1}(x_{m_1}) \\ \vdots & & \vdots \\ x_1^{m_1-1} w_{2,1}(x_1) & \dots & x_{m_1}^{m_1-1} w_{2,1}(x_{m_1}) \end{array} \right| \\
 &\times \left| \begin{array}{ccc} w_{2,2}(x_{m_1+1}) & \dots & w_{2,2}(x_n) \\ \vdots & & \vdots \\ x_{m_1+1}^{m_2-1} w_{2,2}(x_{m_1+1}) & \dots & x_n^{m_2-1} w_{2,2}(x_n) \end{array} \right| \prod_{j=1}^n dx_j
 \end{aligned}$$

which on evaluating the Vandermonde determinants becomes

$$\begin{aligned}
 P_n^{[1,1]}(y, z) &= \frac{\binom{n}{m_1}}{Z_n} \int \dots \int \prod_{j=1}^n \frac{y - x_j}{z - x_j} \left| \begin{array}{ccc} w_{1,1}(x_1) & \dots & w_{1,1}(x_n) \\ \vdots & & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \dots & x_n^{n_1-1} w_{1,1}(x_n) \\ \hline w_{1,2}(x_1) & \dots & w_{1,2}(x_n) \\ \vdots & & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \dots & x_n^{n_2-1} w_{1,2}(x_n) \end{array} \right| \\
 &\times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{m_1+1 \leq i < j \leq n} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \prod_{j=m_1+1}^n w_{2,2}(x_j) \prod_{j=1}^n dx_j. \quad (3.3)
 \end{aligned}$$

Next we expand the right hand side of (1.24). We recall from Proposition 1.10 that $-K_n(x, y)$ is the Schur complement of the matrix

$$\left(\begin{array}{cc|cc} H_{n_1, m_1}^{(1,1)} & H_{n_1, m_2}^{(1,2)} & \mathbf{y}_{n_1} & \mathbf{0} \\ H_{n_2, m_1}^{(2,1)} & H_{n_2, m_2}^{(2,2)} & \mathbf{0} & \mathbf{y}_{n_2} \\ \hline \mathbf{x}_{m_1}^T & \mathbf{0} & 0 & 0 \\ \mathbf{0} & \mathbf{x}_{m_2}^T & 0 & 0 \end{array} \right). \quad (3.4)$$

This implies that $\frac{1}{z-y} I_2 - \int \frac{W(x)K_n(x,y)}{z-x} dx$ is the Schur complement of the matrix

$$\left(\begin{array}{cc|cc} H_{n_1,m_1}^{(1,1)} & H_{n_1,m_2}^{(1,2)} & \mathbf{y}_{n_1} & \mathbf{0} \\ H_{n_2,m_1}^{(2,1)} & H_{n_2,m_2}^{(2,2)} & \mathbf{0} & \mathbf{y}_{n_2} \\ \hline \int \mathbf{x}_{m_1}^T W_{1,1}(x)/(z-x) dx & \int \mathbf{x}_{m_2}^T W_{1,2}(x)/(z-x) dx & 1/(z-y) & 0 \\ \int \mathbf{x}_{m_1}^T W_{2,1}(x)/(z-x) dx & \int \mathbf{x}_{m_2}^T W_{2,2}(x)/(z-x) dx & 0 & 1/(z-y) \end{array} \right). \tag{3.5}$$

From (1.39) this implies the determinant representation

$$\det \left(\frac{1}{z-y} I_2 - \int \frac{W(x)K_n(x,y)}{z-x} dx \right) = \frac{1}{\kappa} \left| \begin{array}{cccc} H_{n_1,m_1}^{(1,1)} & H_{n_1,m_2}^{(1,2)} & \mathbf{y}_{n_1} & \mathbf{0} \\ H_{n_2,m_1}^{(2,1)} & H_{n_2,m_2}^{(2,2)} & \mathbf{0} & \mathbf{y}_{n_2} \\ \int \mathbf{x}_{m_1}^T W_{1,1}(x)/(z-x) dx & \int \mathbf{x}_{m_2}^T W_{1,2}(x)/(z-x) dx & 1/(z-y) & 0 \\ \int \mathbf{x}_{m_1}^T W_{2,1}(x)/(z-x) dx & \int \mathbf{x}_{m_2}^T W_{2,2}(x)/(z-x) dx & 0 & 1/(z-y) \end{array} \right|, \tag{3.6}$$

where $\kappa = \det H_{\mathbf{n},\mathbf{m}}$ is the determinant of the top left block of (3.5).

Now we expand the integral in the right hand side of (3.6). By definition, each $H_{n_k,m_l}^{(k,l)}$ is a Hankel matrix whose (i, j) th entry is the $(i + j)$ th moment with respect to the weight function $W_{k,l}(x) = w_{1,k}(x)w_{2,l}(x)$, for $k, l = 1, 2$:

$$(H_{n_k,m_l}^{(k,l)})_{i,j} = \int x^{i+j} W_{k,l}(x) dx. \tag{3.7}$$

Note that the integration variable x in (3.7) is just a ‘dummy variable’ which we can give any name we want. By choosing the same integration variable x_1 for all entries in the first column of (3.6), we can rewrite (3.6) as follows:

$$= \frac{1}{\kappa} \left| \begin{array}{ccc|ccc} \int W_{1,1}(x_1) dx_1 & * & \dots & * & * & \dots & * & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \int x_1^{n_1-1} W_{1,1}(x_1) dx_1 & * & \dots & * & * & \dots & * & y^{n_1-1} & 0 \\ \hline \int W_{2,1}(x_1) dx_1 & * & \dots & * & * & \dots & * & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \int x_1^{n_2-1} W_{2,1}(x_1) dx_1 & * & \dots & * & * & \dots & * & 0 & y^{n_2-1} \\ \hline \int \frac{W_{1,1}(x_1)}{z-x_1} dx_1 & * & \dots & * & * & \dots & * & \frac{1}{z-y} & 0 \\ \int \frac{W_{2,1}(x_1)}{z-x_1} dx_1 & * & \dots & * & * & \dots & * & 0 & \frac{1}{z-y} \end{array} \right|$$

where the entries denoted with * are not important at this stage of the proof. Using the multilinearity of the determinant with respect to its first column, we can take out the integration with respect to x_1 :

$$= \frac{1}{\kappa} \int \begin{vmatrix} w_{1,1}(x_1) & * & \dots & * & * & \dots & * & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & * & \dots & * & * & \dots & * & y^{n_1-1} & 0 \\ \hline w_{1,2}(x_1) & * & \dots & * & * & \dots & * & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & * & \dots & * & * & \dots & * & 0 & y^{n_2-1} \\ \hline \frac{w_{1,1}(x_1)}{z-x_1} & * & \dots & * & * & \dots & * & \frac{1}{z-y} & 0 \\ \frac{w_{1,2}(x_1)}{z-x_1} & * & \dots & * & * & \dots & * & 0 & \frac{1}{z-y} \end{vmatrix} w_{2,1}(x_1) dx_1.$$

Here we also used that $W_{1,1}(x) = w_{1,1}(x)w_{2,1}(x)$ and $W_{2,1}(x) = w_{1,2}(x)w_{2,1}(x)$, allowing us to extract the common factor $w_{2,1}(x_1)$ from the first column. Applying the same technique to the other columns of the first block column we obtain

$$= \frac{1}{\kappa} \int \dots \int \begin{vmatrix} w_{1,1}(x_1) & \dots & w_{1,1}(x_{m_1}) & * & \dots & * & 1 & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \dots & x_{m_1}^{n_1-1} w_{1,1}(x_{m_1}) & * & \dots & * & y^{n_1-1} & 0 \\ \hline w_{1,2}(x_1) & \dots & w_{1,2}(x_{m_1}) & * & \dots & * & 0 & 1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \dots & x_{m_1}^{n_2-1} w_{1,2}(x_{m_1}) & * & \dots & * & 0 & y^{n_2-1} \\ \hline \frac{w_{1,1}(x_1)}{z-x_1} & \dots & \frac{w_{1,1}(x_{m_1})}{z-x_{m_1}} & * & \dots & * & \frac{1}{z-y} & 0 \\ \frac{w_{1,2}(x_1)}{z-x_1} & \dots & \frac{w_{1,2}(x_{m_1})}{z-x_{m_1}} & * & \dots & * & 0 & \frac{1}{z-y} \end{vmatrix} \\ \times \prod_{j=1}^{m_1} (x_j^{j-1} w_{2,1}(x_j) dx_j).$$

To get rid of the factor $\prod_{j=1}^{m_1} x_j^{j-1}$ one can apply the following well-known symmetrization trick. Sum the above expression over all permutations $x_{\sigma_1}, \dots, x_{\sigma_{m_1}}$ of x_1, \dots, x_{m_1} . For each term in this sum, apply a column permutation to arrange the columns of the determinant back in their original form; this releases a factor $(-1)^\sigma$. But then we get the expansion of a Vandermonde determinant:

$$\sum_{\sigma} (-1)^\sigma \prod_{j=1}^{m_1} x_{\sigma_j}^{j-1} = \prod_{1 \leq i < j \leq m_1} (x_j - x_i).$$

So we get

$$\begin{aligned}
 &= \frac{1}{\kappa(m_1)!} \int \cdots \int \left| \begin{array}{ccc|ccc}
 w_{1,1}(x_1) & \cdots & w_{1,1}(x_{m_1}) & * & \cdots & * & 1 & 0 \\
 \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\
 x_1^{n_1-1} w_{1,1}(x_1) & \cdots & x_{m_1}^{n_1-1} w_{1,1}(x_{m_1}) & * & \cdots & * & y^{n_1-1} & 0 \\
 w_{1,2}(x_1) & \cdots & w_{1,2}(x_{m_1}) & * & \cdots & * & 0 & 1 \\
 \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots \\
 x_1^{n_2-1} w_{1,2}(x_1) & \cdots & x_{m_1}^{n_2-1} w_{1,2}(x_{m_1}) & * & \cdots & * & 0 & y^{n_2-1} \\
 \hline
 \frac{w_{1,1}(x_1)}{z-x_1} & \cdots & \frac{w_{1,1}(x_{m_1})}{z-x_{m_1}} & * & \cdots & * & \frac{1}{z-y} & 0 \\
 \frac{w_{1,2}(x_1)}{z-x_1} & \cdots & \frac{w_{1,2}(x_{m_1})}{z-x_{m_1}} & * & \cdots & * & 0 & \frac{1}{z-y}
 \end{array} \right| \\
 &\times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \prod_{j=1}^{m_1} dx_j.
 \end{aligned}$$

Applying the same techniques to the second block column, we find

$$\begin{aligned}
 &= \frac{1}{\kappa(m_1)!(m_2)!} \int \cdots \int \left| \begin{array}{ccc|ccc}
 w_{1,1}(x_1) & \cdots & w_{1,1}(x_n) & 1 & 0 \\
 \vdots & & \vdots & \vdots & \vdots \\
 x_1^{n_1-1} w_{1,1}(x_1) & \cdots & x_n^{n_1-1} w_{1,1}(x_n) & y^{n_1-1} & 0 \\
 w_{1,2}(x_1) & \cdots & w_{1,2}(x_n) & 0 & 1 \\
 \vdots & & \vdots & \vdots & \vdots \\
 x_1^{n_2-1} w_{1,2}(x_1) & \cdots & x_n^{n_2-1} w_{1,2}(x_n) & 0 & y^{n_2-1} \\
 \hline
 \frac{w_{1,1}(x_1)}{z-x_1} & \cdots & \frac{w_{1,1}(x_n)}{z-x_n} & \frac{1}{z-y} & 0 \\
 \frac{w_{1,2}(x_1)}{z-x_1} & \cdots & \frac{w_{1,2}(x_n)}{z-x_n} & 0 & \frac{1}{z-y}
 \end{array} \right| \\
 &\times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{m_1+1 \leq i < j \leq n} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \prod_{j=m_1+1}^n w_{2,2}(x_j) \prod_{j=1}^n dx_j. \tag{3.8}
 \end{aligned}$$

We claim that this can be rewritten as

$$\begin{aligned}
 &= \frac{1}{(z-y)^2} \frac{1}{\kappa(m_1)!(m_2)!} \int \cdots \int \frac{\prod_{j=1}^n (y-x_j)}{\prod_{j=1}^n (z-x_j)} \\
 &\times \left| \begin{array}{ccc}
 w_{1,1}(x_1) & \cdots & w_{1,1}(x_n) \\
 \vdots & & \vdots \\
 x_1^{n_1-1} w_{1,1}(x_1) & \cdots & x_n^{n_1-1} w_{1,1}(x_n) \\
 \hline
 w_{1,2}(x_1) & \cdots & w_{1,2}(x_n) \\
 \vdots & & \vdots \\
 x_1^{n_2-1} w_{1,2}(x_1) & \cdots & x_n^{n_2-1} w_{1,2}(x_n)
 \end{array} \right| \\
 &\times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{m_1+1 \leq i < j \leq n} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \prod_{j=m_1+1}^n w_{2,2}(x_j) \prod_{j=1}^n dx_j. \tag{3.9}
 \end{aligned}$$

This follows from standard determinant manipulations and the exact formula for a Cauchy–Vandermonde determinant; see the [Appendix](#).

Now up to a constant factor times $1/(z - y)^2$, (3.9) equals (3.3). This proves (1.24), up to a constant factor; but this constant factor must be 1 as can be seen by matching the leading terms on both sides of (1.24).

Remark 3.1. By matching the leading constants in (3.3) and (3.9), we obtain the following expression for the normalization constant Z_n in (1.10):

$$Z_n = \kappa n!$$

where we recall that κ is the block Hankel determinant $\kappa = \det H_{n,m}$. This identity also follows directly from a ‘generalized Cauchy–Binet identity’, see e.g. [22].

3.2. Proofs of Theorems 1.2 and 1.3

Having established Theorem 1.5, Theorems 1.2 and 1.3 now follow as simple limiting cases. For Theorem 1.2 we let $z \rightarrow \infty$ in (1.24) and observe from (1.25) that

$$\begin{aligned} \det R_n(z, y) &= \det \left(\begin{pmatrix} I_p & 0 \end{pmatrix} Y^{-1}(z) Y(y) \begin{pmatrix} I_p \\ 0 \end{pmatrix} \right) \\ &= \det \left(\begin{pmatrix} (Y^{-1})_{1,1}(z) & (Y^{-1})_{1,2}(z) \end{pmatrix} \begin{pmatrix} Y_{1,1}(y) \\ Y_{2,1}(y) \end{pmatrix} \right) \\ &= z^{-n} \det Y_{1,1}(y) + O(z^{-n-1}), \end{aligned}$$

where in the last step we used the Cauchy–Binet identity and the asymptotics (1.15). This establishes Theorem 1.2. Theorem 1.3 can be obtained in a similar way by letting $y \rightarrow \infty$ in (1.22).

3.3. Proof of Theorem 1.8(c)

The proof that we have given for Theorem 1.5 can be extended to prove Theorem 1.8(c) as well. The key observation is that the matrix in the right hand side of (1.30) can be written in Schur complement form as well. For example, for $K = L = 2$ and $p = q = 2$ one has that

$$\begin{pmatrix} \frac{1}{z_1 - y_1} R_n(z_1, y_1) & \frac{1}{z_1 - y_2} R_n(z_1, y_2) \\ \frac{1}{z_2 - y_1} R_n(z_2, y_1) & \frac{1}{z_2 - y_2} R_n(z_2, y_2) \end{pmatrix}$$

is the Schur complement of the matrix.

$$\left(\begin{array}{cc|cccc} H_{n_1, m_1}^{(1,1)} & H_{n_1, m_2}^{(1,2)} & (y_1)_{n_1} & \mathbf{0} & (y_2)_{n_1} & \mathbf{0} \\ H_{n_2, m_1}^{(2,1)} & H_{n_2, m_2}^{(2,2)} & \mathbf{0} & (y_1)_{n_2} & \mathbf{0} & (y_2)_{n_2} \\ \hline \int_{\mathbf{x}_{m_1}^T} W_{1,1}(x)/(z_1 - x) dx & \int_{\mathbf{x}_{m_2}^T} W_{1,2}(x)/(z_1 - x) dx & 1/(z_1 - y_1) & 0 & 1/(z_1 - y_2) & 0 \\ \int_{\mathbf{x}_{m_1}^T} W_{2,1}(x)/(z_1 - x) dx & \int_{\mathbf{x}_{m_2}^T} W_{2,2}(x)/(z_1 - x) dx & 0 & 1/(z_1 - y_1) & 0 & 1/(z_1 - y_2) \\ \int_{\mathbf{x}_{m_1}^T} W_{1,1}(x)/(z_2 - x) dx & \int_{\mathbf{x}_{m_2}^T} W_{1,2}(x)/(z_2 - x) dx & 1/(z_2 - y_1) & 0 & 1/(z_2 - y_2) & 0 \\ \int_{\mathbf{x}_{m_1}^T} W_{2,1}(x)/(z_2 - x) dx & \int_{\mathbf{x}_{m_2}^T} W_{2,2}(x)/(z_2 - x) dx & 0 & 1/(z_2 - y_1) & 0 & 1/(z_2 - y_2) \end{array} \right). \quad (3.10)$$

Compare with (3.5).

The determinant of this Schur complement can be manipulated in the same way as in Section 3.1. This leads to the following analogue of (3.8):

$$\frac{1}{\kappa(m_1)!(m_2)!} \int \cdots \int \begin{vmatrix} w_{1,1}(x_1) & \cdots & w_{1,1}(x_n) & 1 & 0 & 1 & 0 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \cdots & x_n^{n_1-1} w_{1,1}(x_n) & y_1^{n_1-1} & 0 & y_2^{n_1-1} & 0 \\ w_{1,2}(x_1) & \cdots & w_{1,2}(x_n) & 0 & 1 & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \cdots & x_n^{n_2-1} w_{1,2}(x_n) & 0 & y_1^{n_2-1} & 0 & y_2^{n_2-1} \\ \frac{w_{1,1}(x_1)}{z_1-x_1} & \cdots & \frac{w_{1,1}(x_n)}{z_1-x_n} & \frac{1}{z_1-y_1} & 0 & \frac{1}{z_1-y_2} & 0 \\ \frac{w_{1,2}(x_1)}{z_1-x_1} & \cdots & \frac{w_{1,2}(x_n)}{z_1-x_n} & 0 & \frac{1}{z_1-y_1} & 0 & \frac{1}{z_1-y_2} \\ \frac{w_{1,1}(x_1)}{z_2-x_1} & \cdots & \frac{w_{1,1}(x_n)}{z_2-x_n} & \frac{1}{z_2-y_1} & 0 & \frac{1}{z_2-y_2} & 0 \\ \frac{w_{1,2}(x_1)}{z_2-x_1} & \cdots & \frac{w_{1,2}(x_n)}{z_2-x_n} & 0 & \frac{1}{z_2-y_1} & 0 & \frac{1}{z_2-y_2} \end{vmatrix} \\ \times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{m_1+1 \leq i < j \leq n} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \prod_{j=m_1+1}^n w_{2,2}(x_j) \prod_{j=1}^n dx_j, \tag{3.11}$$

and the following analogue of (3.9):

$$\left(\frac{(z_1 - z_2)(y_2 - y_1)}{\prod_{i,j=1}^2 (z_i - y_j)} \right)^2 \frac{1}{\kappa(m_1)!(m_2)!} \int \cdots \int \frac{\prod_{i=1}^2 \prod_{j=1}^n (y_i - x_j)}{\prod_{i=1}^2 \prod_{j=1}^n (z_i - x_j)} \\ \times \begin{vmatrix} w_{1,1}(x_1) & \cdots & w_{1,1}(x_n) \\ \vdots & & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \cdots & x_n^{n_1-1} w_{1,1}(x_n) \\ w_{1,2}(x_1) & \cdots & w_{1,2}(x_n) \\ \vdots & & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \cdots & x_n^{n_2-1} w_{1,2}(x_n) \end{vmatrix} \\ \times \prod_{1 \leq i < j \leq m_1} (x_j - x_i) \prod_{m_1+1 \leq i < j \leq n} (x_j - x_i) \prod_{j=1}^{m_1} w_{2,1}(x_j) \\ \times \prod_{j=m_1+1}^n w_{2,2}(x_j) \prod_{j=1}^n dx_j. \tag{3.12}$$

Here the transition from (3.11) to (3.12) follows again from standard determinant manipulations and the exact formula for a Cauchy–Vandermonde determinant, see the [Appendix](#).

Formula (3.12) gives an expression for the determinant in the right hand side of (1.30). Here the leftmost factor

$$\left(\frac{(z_1 - z_2)(y_2 - y_1)}{\prod_{i,j=1}^2 (z_i - y_j)} \right)^2 \tag{3.13}$$

is canceled by multiplying with the first factor in the right hand side in (1.30). Finally, by also expanding the left hand side of (1.30) in the same way as before, cf. (3.3), Theorem 1.8(c) follows.

4. Proofs, part 2

In this section we prove Theorems 1.8(a)–(b) and 1.9 about averages of general products and ratios of characteristic polynomials. To this end we use a similar strategy as in Baik–Deift–Strahov [4], i.e., we make use of the already established Theorems 1.2, 1.3 and 1.5, together with adaptations of the classical results by Christoffel and Uvarov. We mention that these Christoffel–Uvarov type results will be valid for arbitrary vector orthogonal polynomials, i.e., the weight matrix $W(x)$ does not need to have the rank-one factorization (1.14).

Throughout this section we assume that $|\mathbf{n}| = |\mathbf{m}|$ and we let $W(x)$ be an arbitrary weight matrix of size $p \times q$. We let $Y(z)$ be the solution of the corresponding RH problem and $K_n(x, y)$ the kernel (1.17). We also consider the modified weight matrix

$$\tilde{W}(x) = \frac{\prod_{k=1}^K (x - y_k)}{\prod_{l=1}^L (x - z_l)} W(x), \tag{4.1}$$

and we denote the corresponding RH matrix and kernel with $\tilde{Y}(z)$ and $\tilde{K}_n(x, y)$. Here we assume that $y_1, \dots, y_K \in \mathbb{C}, z_1, \dots, z_L \in \mathbb{C} \setminus \mathbb{R}$ and the numbers in the set $(y_1, \dots, y_K, z_1, \dots, z_L)$ are pairwise distinct. Our aim will be to relate the matrices \tilde{Y} and Y , and similarly for the kernels \tilde{K}_n and K_n .

Although Theorem 1.8 is just a special case of Theorem 1.9, for reasons of comprehension we discuss the former first, see Sections 4.1 and 4.2. The proof of Theorem 1.9 is then only briefly sketched in Section 4.3.

4.1. Proof of Theorem 1.8(a)

Recall that Christoffel’s formula expresses the orthogonal polynomials with respect to a polynomially modified weight function $\tilde{w}(x) = \prod_{k=1}^K (x - y_k)w(x)$ in terms of those with respect to the original weight function $w(x)$ [4,27]. We now state a generalization of this formula to the context of vector orthogonal polynomials. Throughout, we will use the notations in Definition 1.7.

Proposition 4.1. Assume that $L = 0$ in (4.1), so that

$$\tilde{W}(x) = \prod_{k=1}^K (x - y_k) W(x). \tag{4.2}$$

Then $\tilde{Y}_{1,1}(y)$ is the Schur complement of the matrix

$$\frac{1}{\prod_{k=1}^K (y - y_k)} \left(\begin{array}{ccc|c} (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_K) & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y) \\ \vdots & & \vdots & \vdots \\ (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_K) & (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y) \\ \hline (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y_K) & (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y) \end{array} \right) \tag{4.3}$$

with respect to its bottom right p by p submatrix.

Proof. Denote with $S(y)$ the Schur complement of the matrix (4.3). It is clear that $S(y)$ is a matrix of size p by p . We claim that the entries of $S(y)$ are polynomials in y , and moreover that

$$S(y) = (I_p + O(1/y))\text{diag}(y^{n_1}, \dots, y^{n_p}), \quad y \rightarrow \infty. \tag{4.4}$$

To prove these claims, first note that all entries in the last block column of (4.3) are polynomials in y . Now we use the expression (1.42) for the entries of the Schur complement of the matrix (1.35). It is clear that if we apply this formula to the input matrix (4.3), then the determinant in the numerator of (1.42) is zero whenever $y = y_k$ for certain $k = 1, \dots, K$. This means that all entries of $S(y)$ are polynomials divisible by $\prod_{k=1}^K (y - y_k)$, which on multiplying with the prefactor $1/\prod_{k=1}^K (y - y_k)$ in (4.3) indeed shows that $S(y)$ is a polynomial matrix.

To establish (4.4), observe that the leading term of $S(y)$ comes from the bottom right p by p block in (4.3), which has the asymptotics

$$(Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y) = (I_p + O(1/y))\text{diag}(y^{n_1+K}, \dots, y^{n_p+K}), \quad y \rightarrow \infty.$$

Thus on dividing by the prefactor $1/\prod_{k=1}^K (y - y_k)$ in (4.3), we obtain (4.4).

Next, we check the orthogonality conditions

$$\int S(x)\tilde{W}(x)\mathbf{Q}(x) \, dx = \mathbf{0} \tag{4.5}$$

for all $\mathbf{Q} \in \mathcal{P}_m$, where we recall the definition of $\tilde{W}(x)$ in (4.2). To prove this, note that the factor $\prod_{k=1}^K (x - y_k)$ in (4.2) cancels with the prefactor in (4.3), and by the multi-linearity of determinants the remaining integral can be taken into the last block column of (4.3). So the left hand side of (4.5) is the Schur complement of the matrix

$$\left(\begin{array}{ccc|c} (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_K) & \int (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(x)W(x)\mathbf{Q}(x) \, dx \\ \vdots & & \vdots & \vdots \\ (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(y_K) & \int (Y_{1,1})_{\mathbf{n}_{K-1}, \mathbf{m}_{K-1}}(x)W(x)\mathbf{Q}(x) \, dx \\ \hline (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(y_K) & \int (Y_{1,1})_{\mathbf{n}_K, \mathbf{m}_K}(x)W(x)\mathbf{Q}(x) \, dx \end{array} \right). \tag{4.6}$$

But each of the matrices $Y_{1,1}$ in the last column of (4.6) satisfies the orthogonality relations with respect to $\mathbf{Q} \in \mathcal{P}_m$. So the last column of (4.6) is zero and therefore the Schur complement of this matrix is zero as well. This establishes (4.5).

The proof of the proposition now follows from (4.4) and (4.5), see (2.5). \square

Having established Theorem 1.2 and Proposition 4.1, and taking into account the property (1.39), Theorem 1.8(a) can now be proved by induction on K , in exactly the same way as in [4, Proof of Th. 4.3]. We refer to the latter reference for details.

4.2. Proof of Theorem 1.8(b)

Recall that Uvarov’s formula expresses the orthogonal polynomials with respect to an inverse-polynomially modified weight function $\tilde{w}(x) = w(x)/\prod_{l=1}^L (x - z_l)$ in terms of those with respect to the original weight function $w(x)$ [4,28]. We now state a generalization of this formula to the context of vector orthogonal polynomials.

Proposition 4.2. Assume that $K = 0$ in (4.1), so that

$$\tilde{W}(x) = \frac{1}{\prod_{l=1}^L (x - z_l)} W(x). \tag{4.7}$$

Then $\tilde{Y}_{2,1}(z)$ is the Schur complement of the matrix

$$\left(\begin{array}{ccc|c} (Y_{2,2})_{n_0, m_0}(z_1) & \cdots & (Y_{2,2})_{n_0, m_0}(z_L) & (Y_{2,1})_{n_0, m_0}(z) \\ \vdots & & \vdots & \vdots \\ (Y_{2,2})_{n_{-L+1}, m_{-L+1}}(z_1) & \cdots & (Y_{2,2})_{n_{-L+1}, m_{-L+1}}(z_L) & (Y_{2,1})_{n_{-L+1}, m_{-L+1}}(z) \\ \hline (Y_{2,2})_{n_{-L}, m_{-L}}(z_1) & \cdots & (Y_{2,2})_{n_{-L}, m_{-L}}(z_L) & (Y_{2,1})_{n_{-L}, m_{-L}}(z) \end{array} \right) \quad (4.8)$$

with respect to its bottom right q by p submatrix, and $\tilde{Y}_{2,2}(z)$ is the Schur complement of the matrix

$$\frac{1}{\prod_{l=1}^L (z - z_l)} \left(\begin{array}{ccc|c} (Y_{2,2})_{n_0, m_0}(z_1) & \cdots & (Y_{2,2})_{n_0, m_0}(z_L) & (Y_{2,2})_{n_0, m_0}(z) \\ \vdots & & \vdots & \vdots \\ (Y_{2,2})_{n_{-L+1}, m_{-L+1}}(z_1) & \cdots & (Y_{2,2})_{n_{-L+1}, m_{-L+1}}(z_L) & (Y_{2,2})_{n_{-L+1}, m_{-L+1}}(z) \\ \hline (Y_{2,2})_{n_{-L}, m_{-L}}(z_1) & \cdots & (Y_{2,2})_{n_{-L}, m_{-L}}(z_L) & (Y_{2,2})_{n_{-L}, m_{-L}}(z) \end{array} \right) \quad (4.9)$$

with respect to its bottom right q by q submatrix.

Proof. Denote with $S(z)$ the Schur complement of the matrix (4.8). It is clear that $S(z)$ is a polynomial matrix of size q by p , which has the required degree structure

$$S(z) = (O(1/z))\text{diag}(z^{n_1}, \dots, z^{n_p}), \quad z \rightarrow \infty. \quad (4.10)$$

On the other hand, denote with $T(z)$ the Schur complement of the matrix (4.9). From the asymptotic condition

$$(Y_{2,2})_{n_{-L}, m_{-L}}(z) = (I_q + O(1/z))\text{diag}(z^{L-m_1}, \dots, z^{L-m_q}), \quad z \rightarrow \infty,$$

it is clear that $T(z)$ has the required asymptotics

$$T(z) = (I_q + O(1/z))\text{diag}(z^{-m_1}, \dots, z^{-m_q}), \quad z \rightarrow \infty. \quad (4.11)$$

Taking into account (4.10) and (4.11), the proposition will follow if we can show that

$$T(z) = -\frac{1}{2\pi i} \int \frac{S(x)\tilde{W}(x)}{z-x} dx. \quad (4.12)$$

Consider the left hand side of (4.12). From the formula $Y_{2,2}(z) = -\frac{1}{2\pi i} \int \frac{Y_{2,1}(x)W(x)}{z-x} dx$ we see that all entries of matrix (4.9) are integrals. Choose a common integration variable x_i for all entries in the i th row. Consider the coefficients in the partial fraction decomposition

$$\sum_{l=1}^L \frac{c_l}{x - z_l} + \frac{1}{x - z} = \frac{c}{(x - z) \prod_{l=1}^L (x - z_l)}, \quad (4.13)$$

where c_l, c are functions of $\{z_1, \dots, z_L, z\}$ but not of x . Consider the column operation where to the last block column of matrix (4.9) we add c_l times the l th block column, $l = 1, \dots, L$. Performing these operations inside the integrals, this causes the integrand in the last block column to be multiplied with $c / \prod_{l=1}^L (x - z_l) = \prod_{l=1}^L (z - z_l) / \prod_{l=1}^L (x - z_l)$. Here the factor $\prod_{l=1}^L (z - z_l)$ can be taken out of the integrand, canceling the prefactor in (4.9), while the factor $W(x) / \prod_{l=1}^L (x - z_l)$ in the integrand is nothing but $\tilde{W}(x)$. So we obtain precisely the right hand side of (4.12). This establishes (4.12), and thereby the proposition is proved. \square

Having established Theorem 1.3 and Proposition 4.2, and taking into account the property (1.39), Theorem 1.8(b) can now be proved by induction on L , in exactly the same way as in [4, Proof of Th. 4.10]. We refer to the latter reference for details.

4.3. Proof of Theorem 1.9

Finally we prove Theorem 1.9. The key to the proof of Theorem 1.9(a) is the following result.

Proposition 4.3. Assume that $K \geq L$ in (4.1). Then $\tilde{Y}_{1,1}(y)$ is the Schur complement of the matrix

$$\frac{\prod_{l=1}^L (y - z_l)}{\prod_{k=1}^K (y - y_k)} \left(\begin{array}{ccc|ccc} \frac{1}{z_1 - y_1} R_n(z_1, y_1) & \dots & \frac{1}{z_1 - y_K} R_n(z_1, y_K) & \frac{1}{z_1 - y} R_n(z_1, y) & & \\ \vdots & & \vdots & \vdots & & \\ \frac{1}{z_L - y_1} R_n(z_L, y_1) & \dots & \frac{1}{z_L - y_K} R_n(z_L, y_K) & \frac{1}{z_L - y} R_n(z_L, y) & & \\ (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y_K) & (Y_{1,1})_{\mathbf{n}_0, \mathbf{m}_0}(y) & & \\ \vdots & & \vdots & \vdots & & \\ \hline (Y_{1,1})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(y_1) & \dots & (Y_{1,1})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(y_K) & (Y_{1,1})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(y) & & \end{array} \right) \quad (4.14)$$

with respect to its bottom right p by p submatrix. Here all R_n matrices are taken with respect to the pair of multi-indices $\mathbf{n}_0, \mathbf{m}_0$.

The proof follows the same lines as the proof of Proposition 4.1, and we omit the details. We suffice to mention that the proof makes use of the vanishing property for the matrix R_n in Proposition 2.7.

From Proposition 4.3, Theorem 1.9(a) can now be proved by induction on K , $K \geq L$ (for L fixed), with induction basis $K = L$ corresponding to Theorem 1.8(c).

Finally, we note that the proof of Theorem 1.9(b) hinges on the following fact.

Proposition 4.4. Assume that $L \geq K$ in (4.1). Then $\tilde{Y}_{2,1}(z)$ is the Schur complement of the matrix

$$\left(\begin{array}{ccc|ccc} \frac{1}{z_1 - y_1} L_n(y_1, z_1) & \dots & \frac{1}{z_L - y_1} L_n(y_1, z_L) & 2\pi i K_n(y_1, z) & & \\ \vdots & & \vdots & \vdots & & \\ \frac{1}{z_1 - y_K} L_n(y_K, z_1) & \dots & \frac{1}{z_L - y_K} L_n(y_K, z_L) & 2\pi i K_n(y_K, z) & & \\ (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_L) & (Y_{2,1})_{\mathbf{n}_0, \mathbf{m}_0}(z) & & \\ \vdots & & \vdots & \vdots & & \\ \hline (Y_{2,2})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z_1) & \dots & (Y_{2,2})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z_L) & (Y_{2,1})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z) & & \end{array} \right) \quad (4.15)$$

with respect to its bottom right q by p submatrix, and $\tilde{Y}_{2,2}(z)$ is the Schur complement of the matrix

$$\frac{\prod_{k=1}^K (z - y_k)}{\prod_{l=1}^L (z - z_l)} \left(\begin{array}{ccc|c} \frac{1}{z_1 - y_1} L_n(y_1, z_1) & \cdots & \frac{1}{z_L - y_1} L_n(y_1, z_L) & \frac{1}{z - y_1} L_n(y_1, z) \\ \vdots & & \vdots & \vdots \\ \frac{1}{z_1 - y_K} L_n(y_K, z_1) & \cdots & \frac{1}{z_L - y_K} L_n(y_K, z_L) & \frac{1}{z - y_K} L_n(y_K, z) \\ (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_1) & \cdots & (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z_L) & (Y_{2,2})_{\mathbf{n}_0, \mathbf{m}_0}(z) \\ \vdots & & \vdots & \vdots \\ (Y_{2,2})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z_1) & \cdots & (Y_{2,2})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z_L) & (Y_{2,2})_{\mathbf{n}_{K-L}, \mathbf{m}_{K-L}}(z) \end{array} \right) \quad (4.16)$$

with respect to its bottom right q by q submatrix. Here all L_n matrices are taken with respect to the pair of multi-indices $\mathbf{n}_0, \mathbf{m}_0$.

Proof. The proof follows the same lines as the proof of Proposition 4.2. The properties (4.10) and (4.11) are easily extended to the present situation. The main difficulty lies in proving (4.12). To this end one can use the following analogue of the partial fraction decomposition (4.13),

$$\sum_{l=1}^L \frac{c_l}{x - z_l} + \frac{1}{x - z} = \frac{c \prod_{k=1}^K (x - y_k)}{(x - z) \prod_{l=1}^L (x - z_l)}, \quad (4.17)$$

where c_l, c are functions of $\{y_1, \dots, y_K, z_1, \dots, z_L, z\}$ but not of x . Observe that these very same coefficients c_l, c also appear in the more complicated partial fraction decomposition

$$\sum_{l=1}^L \frac{c_l(x - y_k)}{(z_l - y_k)(x - z_l)} + \frac{x - y_k}{(z - y_k)(x - z)} = \frac{c \prod_{k=1}^K (x - y_k)}{(x - z) \prod_{l=1}^L (x - z_l)}, \quad (4.18)$$

for any k . The proof of (4.12) then follows as in the proof of Proposition 4.2, using the integral representations $Y_{2,2}(z) = -\frac{1}{2\pi i} \int \frac{Y_{2,1}(x)W(x)}{z-x} dx$ (as before) and

$$L_n(y, z) = \int \frac{x - y}{x - z} K_n(y, x)W(x) dx,$$

and applying the relations, (4.17) for the integrals for $Y_{2,2}$ and (4.18) for those for L_n . \square

Appendix. Calculations with Cauchy–Vandermonde determinants

Define the functions

$$f_i(x) = \begin{cases} x^{i-1}, & i = 1, \dots, n, \\ 1/(z_{i-n} - x), & i = n + 1, \dots, n + m, \end{cases}$$

where z_1, \dots, z_m are a given set of numbers. It is well known that

$$\det (f_i(x_j))_{i,j=1}^{n+m} = \frac{\prod_{1 \leq i < j \leq m} (z_i - z_j) \prod_{1 \leq i < j \leq n+m} (x_j - x_i)}{\prod_{i,j} (z_i - x_j)}. \quad (4.19)$$

The expression in the left hand side of (4.19) is called a *Cauchy–Vandermonde determinant*.

Proof that (3.8) implies (3.9)

By applying a suitable row permutation, the determinant inside (3.8) takes the form

$$(-1)^{n_2} \begin{vmatrix} w_{1,1}(x_1) & \dots & w_{1,1}(x_n) & 1 & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_1-1} w_{1,1}(x_1) & \dots & x_n^{n_1-1} w_{1,1}(x_n) & y^{n_1-1} & 0 \\ \frac{w_{1,1}(x_1)}{z-x_1} & \dots & \frac{w_{1,1}(x_n)}{z-x_n} & \frac{1}{z-y} & 0 \\ \hline w_{1,2}(x_1) & \dots & w_{1,2}(x_n) & 0 & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ x_1^{n_2-1} w_{1,2}(x_1) & \dots & x_n^{n_2-1} w_{1,2}(x_n) & 0 & y^{n_2-1} \\ \frac{w_{1,2}(x_1)}{z-x_1} & \dots & \frac{w_{1,2}(x_n)}{z-x_n} & 0 & \frac{1}{z-y} \end{vmatrix}.$$

By applying Lagrange expansion with respect to the first block row, this equals

$$\begin{aligned} &= \sum_S (-1)^S \begin{vmatrix} w_{1,1}(x_{s_1}) & \dots & w_{1,1}(x_{s_{n_1}}) & 1 \\ \vdots & & \vdots & \vdots \\ x_{s_1}^{n_1-1} w_{1,1}(x_{s_1}) & \dots & x_{s_{n_1}}^{n_1-1} w_{1,1}(x_{s_{n_1}}) & y^{n_1-1} \\ \frac{w_{1,1}(x_{s_1})}{z-x_{s_1}} & \dots & \frac{w_{1,1}(x_{s_{n_1}})}{z-x_{s_{n_1}}} & \frac{1}{z-y} \end{vmatrix} \\ &\times \begin{vmatrix} w_{1,2}(x_{\bar{s}_1}) & \dots & w_{1,2}(x_{\bar{s}_{n_2}}) & 1 \\ \vdots & & \vdots & \vdots \\ x_{\bar{s}_1}^{n_2-1} w_{1,2}(x_{\bar{s}_1}) & \dots & x_{\bar{s}_{n_2}}^{n_2-1} w_{1,2}(x_{\bar{s}_{n_2}}) & y^{n_2-1} \\ \frac{w_{1,2}(x_{\bar{s}_1})}{z-x_{\bar{s}_1}} & \dots & \frac{w_{1,2}(x_{\bar{s}_{n_2}})}{z-x_{\bar{s}_{n_2}}} & \frac{1}{z-y} \end{vmatrix} \end{aligned} \tag{4.20}$$

where the sum is over all $\binom{n}{n_1}$ subsets $S \subset \{1, \dots, n\}$ with $|S| = n_1$, where we write the elements of S in increasing order: $s_1 < \dots < s_{n_1}$, similarly for those of the complement $\bar{S} = \{1, \dots, n\} \setminus S$, $\bar{s}_1 < \dots < \bar{s}_{n_2}$, and where we denote by $(-1)^S$ the sign of the permutation $s_1, \dots, s_{n_1}, \bar{s}_1, \dots, \bar{s}_{n_2}$. Note that the factor $(-1)^{n_2}$ of the previous formula has disappeared.

Up to a diagonal scaling of the columns, both determinants in (4.20) are Cauchy–Vandermonde determinants. By virtue of (4.19), this yields

$$\begin{aligned} &= \frac{1}{(z-y)^2} \left(\prod_{j=1}^n \frac{y-x_j}{z-x_j} \right) \sum_S (-1)^S \begin{vmatrix} w_{1,1}(x_{s_1}) & \dots & w_{1,1}(x_{s_{n_1}}) \\ \vdots & & \vdots \\ x_{s_1}^{n_1-1} w_{1,1}(x_{s_1}) & \dots & x_{s_{n_1}}^{n_1-1} w_{1,1}(x_{s_{n_1}}) \end{vmatrix} \\ &\times \begin{vmatrix} w_{1,2}(x_{\bar{s}_1}) & \dots & w_{1,2}(x_{\bar{s}_{n_2}}) \\ \vdots & & \vdots \\ x_{\bar{s}_1}^{n_2-1} w_{1,2}(x_{\bar{s}_1}) & \dots & x_{\bar{s}_{n_2}}^{n_2-1} w_{1,2}(x_{\bar{s}_{n_2}}) \end{vmatrix}. \end{aligned} \tag{4.21}$$

But this is precisely what one gets from Lagrange expansion of the determinant in (3.9).

Proof that (3.11) implies (3.12)

The proof that (3.11) implies (3.12) follows in the same way. The only difference is that there is an extra row and column in each Cauchy–Vandermonde determinant, thereby yielding the prefactor (3.13).

References

- [1] M. Adler, P. van Moerbeke, P. Vanhaecke, Moment matrices and multi-component KP, with applications to random matrix theory, *Comm. Math. Phys.* 286 (2009) 1–38.
- [2] A.I. Aptekarev, Multiple orthogonal polynomials, *J. Comput. Appl. Math.* 99 (1998) 423–447.
- [3] J. Baik, On the Christoffel–Darboux kernel for random Hermitian matrices with external source, manuscript 2009 (arXiv:0809.3970).
- [4] J. Baik, P. Deift, E. Strahov, Products and ratios of characteristic polynomials of random hermitian matrices, *J. Math. Phys.* 44 (2003) 3657–3670.
- [5] P.M. Bleher, A.B.J. Kuijlaars, Random matrices with external source and multiple orthogonal polynomials, *Int. Math. Res. Not.* 2004 (3) (2004) 109–129.
- [6] A. Borodin, Biorthogonal ensembles, *Nuclear Phys.* B536 (1999) 704–732.
- [7] A. Borodin, E. Strahov, Averages of characteristic polynomials in random matrix theory, *Comm. Pure Appl. Math.* 59 (2006) 161–253.
- [8] E. Brézin, S. Hikami, Characteristic polynomials of random matrices, *Comm. Math. Phys.* 214 (2000) 111–135. 4140–4149.
- [9] E. Brézin, S. Hikami, Universal singularity at the closure of the gap in a random matrix theory, *Phys. Rev. E* 57 (1998) 4140–4149.
- [10] E. Brézin, S. Hikami, Level spacing of random matrices in an external source, *Phys. Rev. E* 58 (1998) 7176–7185.
- [11] E. Daems, A.B.J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions, *J. Approx. Theory* 146 (2007) 91–114.
- [12] P. Deift, Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach, in: *Courant Lecture Notes in Mathematics*, vol. 3, Amer. Math. Soc., Providence RI, 1999.
- [13] P. Desrosiers, P.J. Forrester, A note on biorthogonal ensembles, *J. Approx. Theory* 152 (2008) 167–187.
- [14] U. Fidalgo Prieto, A. López García, G. López Lagomasino, V.N. Sorokin, Mixed type multiple orthogonal polynomials for two Nikishin systems, manuscript 2008 (arXiv:0812.1219).
- [15] A.S. Fokas, A.R. Its, A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, *Comm. Math. Phys.* 147 (1992) 395–430.
- [16] Y.V. Fyodorov, E. Strahov, An exact formula for general spectral correlation function of random Hermitian matrices, *J. Phys. A: Math. Gen.* 36 (2003) 3203–3213.
- [17] F.R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea Publishing Company, New York, 1959.
- [18] I. Gelfand, S. Gelfand, V. Retakh, R. Wilson, Quasideterminants, *Adv. Math.* 193 (2005) 56–141.
- [19] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third edition, The Johns Hopkins University Press, 1996.
- [20] S. Karlin, J. McGregor, Coincidence probabilities, *Pacific J. Math.* 9 (1959) 1141–1164.
- [21] J.P. Keating, N.C. Snaith, Random Matrix Theory and $\zeta(1/2 + it)$, *Comm. Math. Phys.* 214 (2000) 57–89.
- [22] A.B.J. Kuijlaars, Multiple orthogonal polynomial ensembles, manuscript 2009 (arXiv:0902.1058).
- [23] L. Miranian, Matrix-valued orthogonal polynomials on the real line: Some extensions of the classical theory, *J. Phys. A: Math. Gen.* 38 (2005) 5731–5749.
- [24] V.N. Sorokin, On simultaneous approximation of several linear forms, in: *Vestnik Mosk. Univ., Ser Matem No. 1*, 1983, pp. 44–47. English translation: *Moscow Univ. Math. Bull.* 38 (1) (1983) 53–56.
- [25] V.N. Sorokin, J. Van Iseghem, Algebraic aspects of matrix orthogonality for vector polynomials, *J. Approx. Theory* 90 (1997) 97–116.
- [26] E. Strahov, Y.V. Fyodorov, Universal results for correlations of characteristic polynomials: Riemann–Hilbert approach, *Comm. Math. Phys.* 241 (2003) 343–382.
- [27] G. Szegő, *Orthogonal polynomials*, in: *Amer. Math. Soc. Coll. Publ.*, Vol 23, Amer. Math. Soc., Providence, RI, 1975.
- [28] V.B. Uvarov, The connection between systems of polynomials orthogonal with respect to different distribution functions, *USSR Comput. Math. and Math. Phys.* 9 (Part 2) (1969) 25–36.

- [29] W. Van Assche, Multiple orthogonal polynomials, irrationality and transcendence, in: *Continued Fractions: From Analytic Number Theory to Constructive Approximation*, *Contemp. Math.* 236 (1999) 325–342.
- [30] W. Van Assche, J.S. Geronimo, A.B.J. Kuijlaars, Riemann–Hilbert problems for multiple orthogonal polynomials, in: J. Bustoz, et al. (Eds.), *Special Functions 2000: Current Perspectives and Future Directions*, Kluwer, Dordrecht, 2001, pp. 23–59.
- [31] P. Zinn-Justin, Universality of correlation functions of Hermitian random matrices in an external field, *Comm. Math. Phys.* 194 (1998) 631–650.