Communications in Mathematical Physics manuscript No.
(will be inserted by the editor)

On Universality for Orthogonal Ensembles of Random Matrices

M. Shcherbina

Mathematical Division of the Institute for Low Temperature Physics, Kharkov, Ukraine. E-mail: shcherbi@ilt.kharkov.ua

Abstract: We prove universality of local eigenvalue statistics in the bulk of the spectrum for orthogonal invariant matrix models with real analytic potentials with one interval limiting spectrum. Our starting point is the Tracy-Widom formula for the matrix reproducing kernel. The key idea of the proof is to represent the differentiation operator matrix written in the basis of orthogonal polynomials as a product of a positive Toeplitz matrix and a two diagonal skew symmetric Toeplitz matrix.

Key words. Random Matrices, Orthogonal Ensembles

1. Introduction and main results

In this paper we consider ensembles of $n \times n$ real symmetric (or Hermitian) matrices M with the probability distribution

$$P_n(M)dM = Z_{n,\beta}^{-1} \exp\{-\frac{n\beta}{2}\mathrm{Tr}V(M)\}dM, \qquad (1)$$

where $Z_{n,\beta}$ is a normalization constant, $V : \mathbb{R} \to \mathbb{R}_+$ is a Hölder function satisfying the condition

$$|V(\lambda)| \geq 2(1+\epsilon)\log(1+|\lambda|). \qquad (2)$$

A positive parameter β here assumes the values $\beta = 1$ (in the case of real symmetric matrices) or $\beta = 2$ (in the Hermitian case), and dM means the Lebesgue measure on the algebraically independent entries of M. Ensembles of random matrices (1) in the real symmetric case are usually called orthogonal, and in the Hermitian case - unitary ensembles. This terminology reflects the fact that the density of (1) is invariant with respect to orthogonal or unitary transformations of matrices M.

The joint eigenvalue distribution corresponding to (1) has the form (see [13])

$$p_{n,\beta}(\lambda_1,...,\lambda_n)=Q_{n,\beta}^{-1}\prod_{i=1}^{n}e^{-n\beta V(\lambda_i)/2}\prod_{1\le j<k\le n}|\lambda_i-\lambda_j|^{\beta}, \qquad (3)$$

where $Q_{n,\beta}$ is a normalization constant. The simplest question in both cases ($\beta = 1,2$) is the behavior of the Normalized Counting Measure (NCM) of eigenvalues. According to [3,12], NCM converges weakly in probability to the non random limiting measure $\mathcal{N}$ known as the Integrated Density of States (IDS) of the ensemble. The IDS is absolutely continuous, if V' satisfies the Lipshitz condition [18]. The non-negative density $\rho(\lambda)$ is called the Density of States (DOS) of the ensemble. The IDS can be found as a unique solution of a certain variational problem (see [3,5,18]).

To study local regimes for ensembles (1) means to study the behavior of marginal densities

$$p_{l,\beta}^{(n)}(\lambda_1,...,\lambda_l)=\int_{\mathbb{R}^{n-l}}p_{n,\beta}(\lambda_1,...\lambda_l,\lambda_{l+1},...,\lambda_n)d\lambda_{l+1}...d\lambda_n \qquad (4)$$

in the scaling limit, when $\lambda_i=\lambda_0+s_i/n^{\kappa}$ $(i=1,\dots,l)$, and κ is a constant, depending on the behavior of DOS $\rho(\lambda)$ in a small neighborhood of λ_0. If $\rho(\lambda_0)\neq 0$, then $\kappa=1$, if $\rho(\lambda_0)=0$ and $\rho(\lambda)\sim|\lambda-\lambda_0|^{\alpha}$, then $\kappa=1/(1+\alpha)$. The universality conjecture states that the scaling limits of all marginal densities are universal, i.e. do not depend on V and depend only on α and β.

For unitary ensembles all marginal densities can be represented (see [13]) as

$$p_{l,\beta}^{(n)}(\lambda_1,...,\lambda_l)=\frac{(n-l)!}{n!}\det\{K_{n,2}(\lambda_j,\lambda_k)\}_{j,k=1}^{l}, \qquad (5)$$

where

$$K_{n,2}(\lambda,\mu)=\sum_{l=0}^{n-1}\psi_l^{(n)}(\lambda)\psi_l^{(n)}(\mu). \qquad (6)$$

This function is known as a reproducing kernel of the orthonormalized system

$$\psi_l^{(n)}(\lambda)=\exp\{-nV(\lambda)/2\}p_l^{(n)}(\lambda),\ \ l=0,..., \qquad (7)$$

in which $\{p_l^{(n)}\}_{l=0}^{n}$ are orthogonal polynomials on $\mathbb{R}$ associated with the weight $w_n(\lambda)=e^{-nV(\lambda)}$, i.e.,

$$\int p_l^{(n)}(\lambda)p_m^{(n)}(\lambda)w_n(\lambda)d\lambda=\delta_{l,m}. \qquad (8)$$

Hence, the problem to study marginal distributions is replaced by the problem to study the reproducing kernel $K_n(\lambda,\mu)$ in the scaling limit.

This problem was solved in many cases. For example, in the bulk case ($\rho(\lambda_0)\neq 0$) it was shown in [14] (see also [16]) that for a general class of V (the third derivative is bounded in the some neighborhood of λ_0)

$$\lim_{n\to\infty}\frac{1}{n\rho(\lambda_0)}K_{n,2}(\lambda_0+s_1/n\rho(\lambda_0),\lambda_0+s_2/n\rho(\lambda_0))=\mathcal{K}_{\infty,2}^{(0)}(s_1,s_2),$$

where $\mathcal{K}_0(s_1, s_2)$ is a *sin*-kernel

$$\mathcal{K}^{(0)}_{\infty,2}(s_1 - s_2) = \frac{\sin \pi(s_1 - s_2)}{\pi(s_1 - s_2)}. \tag{9}$$

This result for the case of real analytic V was obtained also in [6]. Universality near the edge, i.e. the case when λ_0 is the edge point of the spectrum and $\rho(\lambda) \sim |\lambda - \lambda_0|^{1/2}$, as $\lambda \sim \lambda_0$, was studied in [6]. There are also results on universality near the extreme point, where $\rho(\lambda) \sim (\lambda - \lambda_0)^2$, as $\lambda \sim \lambda_0$ (see [4] for real analytic V and [19] for general V).

For orthogonal ensembles ($\beta = 1$) the situation is more complicated. Instead of (6) we have to use the matrix kernel

$$K_{n,1}(\lambda, \mu) = \begin{pmatrix} S_n(\lambda, \mu) & S_n d(\lambda, \mu) \\ IS_n(\lambda, \mu) - \epsilon(\lambda - \mu) & S_n(\mu, \lambda) \end{pmatrix}. \tag{10}$$

Here

$$S_n(\lambda, \mu) = -\sum_{i,j=0}^{n-1} \psi_i^{(n)}(\lambda)(\mathcal{M}^{(0,n)})^{-1}_{i,j}(n\epsilon\psi_j^{(n)})(\mu), \tag{11}$$

where $\psi_i^{(n)}$ are defined by (7)-(8) and the matrix $\mathcal{M}^{(0,n)}$ is defined as

$$M_{j,l} = n(\psi_j^{(n)}, \epsilon\psi_l^{(n)}); \quad \mathcal{M}^{(0,\infty)} = \{M_{j,l}\}_{j,l=0}^{\infty}; \quad \mathcal{M}^{(0,n)} = \{M_{j,l}\}_{j,l=0}^{n-1}, \tag{12}$$

where ϵ is the integral operator with the kernel

$$\epsilon(\lambda) = \frac{1}{2}\text{sign}(\lambda); \quad \epsilon f(\lambda) = \int \epsilon(\lambda - \mu) f(\mu) d\mu. \tag{13}$$

The symbol d in (10) denotes the differentiating, and $IS_n(\lambda, \mu)$ can be obtained from S_n by some integration procedure. Then similarly to the unitary case all marginal densities can be expressed in terms of the kernel K_{n1} (see [23]) The matrix kernel (10) was introduced first in [11] for circular ensemble and then in [13] for orthogonal ensembles. The scalar kernels of (10) could be defined in principle in terms of any family of polynomials complete in $L_2(\mathbb{R}, w_n)$ (see [23]), but usually the families of skew orthogonal polynomials were used (see [13] and references therein). Unfortunately, for general weights the properties of skew orthogonal polynomials are not studied enough. Hence, using of skew orthogonal polynomials for general V rises serious technical difficulties.

The main technical obstacle to study the kernel (11) defined in terms of orthogonal polynomials is that there is no uniform bound for $||(\mathcal{M}^{(0,n)})^{-1}||$. Widom observed (see [24]) that if the potential V is a rational function, then the question of bounded invertibility of $\mathcal{M}^{(0,n)}$ can be reduced by the question of invertibility of some matrix of fixed size, which depends on the degree of the numerator and denominator of V (e.g. if V is polynomial of degree $2m$, then we should study some $(2m-1) \times (2m-1)$ matrix). In the paper [8] it was shown that the Widom matrix is invertible in the case when (in our notations) $V(\lambda) = \lambda^{2m} + n^{-1/2m} a_{2m-1} \lambda^{2m-1} + \dots$. This allows to prove bulk universality. The same approach was used in [9] to prove edge universality and in [10] to prove bulk and edges universality (including the case of hard edge) for the Laguerre type ensembles. But since small terms $n^{-1/2m}a_{2m-1}\lambda^{2m-1}+\dots$ have no influence on the limiting behavior of $K_n(\lambda,\mu)$ (see [14], [16]), these results in fact prove universality for monomial $V(\lambda)=\lambda^{2m}$. In the papers [21,22] universality in the bulk and near the edges were studied for V being an even quatric polynomial.

In the present paper we prove universality in the bulk of the spectrum for real analytic V with one interval support. Differently from the Widom approach we show that the size of the matrix which we need to control depends only on the number of intervals of σ, and in the one interval case it is enough to control only some rank one matrix. This allow to generalize the method to a more wide class of potentials V and to simplify the proof.

Let us state our main conditions.

C1. *$V(\lambda)$ satisfies (2) and is an even analytic function in*

$$\Omega[d_1,d_2]=\{z:-2-d_1\le\Re z\le 2+d_1,\ |\Im z|\le d_2\},\quad d_1,d_2>0. \tag{14}$$

C2. *The support σ of IDS of the ensemble consists of a single interval:*

$$\sigma=[-2,2].$$

C3. *DOS $\rho(\lambda)$ is strictly positive in the internal points $\lambda\in(-2,2)$ and $\rho(\lambda)\sim|\lambda\mp 2|^{1/2}$, as $\lambda\sim\pm 2$.*

C4. *The function*

$$u(\lambda)=2\int\log|\mu-\lambda|\rho(\mu)d\mu-V(\lambda) \tag{15}$$

achieves its maximum if and only if $\lambda\in\sigma$.

Consider a semi infinite Jacoby matrix $\mathcal{J}^{(n)}$, generated by the recursion relations for orthogonal polynomials (8)

$$J_l^{(n)}\psi_{l+1}^{(n)}(\lambda)+q_l^{(n)}\psi_l^{(n)}(\lambda)+J_{l-1}^{(n)}\psi_{l-1}^{(n)}(\lambda)=\lambda\psi_l^{(n)}(\lambda),\quad J_{-1}^{(n)}=0,\quad l=0,\dots. \tag{16}$$

It is known (see [2]) that under conditions $C1-C4$ $q_l^{(n)}=0$ and there exists some fixed γ such that uniformly in $k:|k|\le 2n^{1/2}$

$$\left|J_{n+k}^{(n)}-1-\frac{k}{n}\gamma\right|\le C\frac{|k|^2+n^{2/3}}{n^2}. \tag{17}$$

Remark 1. The convergence $J_{n+k}^{(n)}\to 1$ $(n\to\infty)$ without uniform bounds for the remainder terms was shown in [1] under much more weak conditions ($V'(\lambda)$ is a Hölder function in some neighborhood of the limiting spectrum).

Note also (see [2]) that under conditions $C1-C4$ the limiting density of states (DOS) ρ has the form

$$\rho(\lambda)=\frac{1}{2\pi}P(\lambda)\sqrt{4-\lambda^2}\,\mathbf{1}_{|\lambda|<2}, \tag{18}$$

where the function P can be represented in the form

$$P(z)=\frac{1}{2\pi i}\oint_{\mathcal{L}}\frac{V'(z)-V'(\zeta)}{(z-\zeta)(\zeta^2-4)^{1/2}}d\zeta=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{V'(z)-V'(2\cos y)}{z-2\cos y}dy. \tag{19}$$

Here the contour $\mathcal{L}\subset\Omega[d_1/2,d_2/2]$, and $\mathcal{L}$ contains inside $[-2,2]$. If V is a polynomial of $2m$th degree, then it is evident that $P(z)$ is a polynomial of $(2m-2)$th degree, and conditions C3 guarantee that

$$|P(z)|\le C,\quad z\in\Omega[d_1/2,d_2/2],\quad P(\lambda)\ge\delta>0,\quad \lambda\in[-2,2]. \tag{20}$$

An important role below belongs to the following two operators:

$$P_{j,k}=\frac{1}{2\pi}\int_{-\pi}^{\pi}P(2\cos y)e^{i(j-k)y}dy=\frac{1}{2\pi i}\oint_{|\zeta|=1}P(\zeta+\zeta^{-1})\zeta^{j-k-1}d\zeta \tag{21}$$

and $\mathcal{R}=\mathcal{P}^{-1}$

$$R_{j,k}=R_{j-k}=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{i(j-k)x}dx}{P(2\cos x)}=\frac{1}{2\pi i}\oint\frac{\zeta^{-1}\zeta^{j-k}d\zeta}{P(\zeta+\zeta^{-1})}. \tag{22}$$

It is important that

$$\delta_1\le\mathcal{R}\le\delta_2,\quad \delta_1=\inf_{\sigma}P^{-1}(\lambda),\quad \delta_2=\sup_{\sigma}P^{-1}(\lambda). \tag{23}$$

Remark also that if we denote by $\mathcal{J}^*$ an infinite Jacobi matrix with constant coefficients

$$\mathcal{J}^*=\{J^*_{j,k}\}_{j,k=-\infty}^{\infty},\quad J^*_{j,k}=\delta_{j+1,k}+\delta_{j-1,k}, \tag{24}$$

then the spectral theorem yields that $\mathcal{P}=P(\mathcal{J}^*)$, $\mathcal{R}=P^{-1}(\mathcal{J}^*)$.

The main result of the paper is

Theorem 1. *Consider the orthogonal ensemble of random matrices defined by (1)-(3) with V satisfying conditions C1-C4 and even n. Then for λ_0 in the bulk ($\rho(\lambda_0)\ne 0$) there exist weak limits of the scaled correlation functions (4) and these limits are given in terms of the universal matrix kernel*

$$K^{(0)}_{\infty,1}(s_1,s_2)=\lim_{n\to\infty}\frac{1}{n\rho(\lambda_0)}K_{n,1}(\lambda_0+s_1/n\rho(\lambda_0),\lambda_0+s_2/n\rho(\lambda_0)), \tag{25}$$

where $K_{n,1}(\lambda,\mu)$ is defined by (10)-(11), and

$$K^{(0)}_{\infty,1}(s_1,s_2)=\begin{pmatrix} K^{(0)}_{\infty,2}(s_1-s_2) & \frac{\partial}{\partial s_1}K^{(0)}_{\infty,2}(s_1-s_2)\\ \int_0^{s_1-s_2}K^{(0)}_{\infty,2}(t)dt-\epsilon(s_1-s_2) & K^{(0)}_{\infty,2}(s_1-s_2)\end{pmatrix},$$

with $K^{(0)}_{\infty,2}(s_1-s_2)$ of the form (9).

The proof of the theorem is based on the following result

Theorem 2. *Under conditions of Theorem 1 for even n the matrix $(\mathcal{M}^{(0,n)})^{-1}$ defined in (12) is bounded uniformly in n, i.e. $||(\mathcal{M}^{(0,n)})^{-1}||\le C$ where C is independent of n and $||.||$ is a standard norm for $n\times n$ matrices.*

The main idea of the proof of Theorem 2 is to consider the matrix $\mathcal{V}^{(0,\infty)}_{j,k}$ of the operator $n^{-1}\frac{d}{dx}$ in the basis $\{\psi^{(n)}_k\}_{k=0}^\infty$ and to prove that (see Lemma 1)

$$\mathcal{V}^{(0,\infty)}_{j,k}=(\mathcal{P}\mathcal{D})_{j,k}+o(1),\quad \text{for}\quad |j-n|,|k-n|<<n,$$

where

$$\mathcal{D}=\{D_{j,k}\}_{j,k=-\infty}^{\infty},\quad D_{j,k}=\delta_{j+1,k}-\delta_{j-1,k}. \tag{26}$$

This allows to construct a local inverse of $\mathcal{M}^{(0,n)}$ up to the rank one perturbation that can be controlled (see Corollary 1).

The paper is organized as follows. In Section 2 we prove Theorems 1 and 2. The proofs of auxiliary results are given in Section 3.

2. Proof of the main results

Proof of Theorem 2. According to the results of [2] and [15], if we restrict the integration in (4) by $|\lambda_i|\le L=2+d_1/2$, consider the polynomials $\{p_k^{(n,L)}\}_{k=0}^\infty$ orthogonal on the interval $[-L,L]$ with the weight e^{-nV} and set $\psi_k^{(n,L)}=e^{-nV/2}p_k^{(n,L)}$, then for $k\le n(1+\varepsilon)$ with some $\varepsilon>0$

$$\sup_{|\lambda|\le L}|\psi_k^{(n,L)}(\lambda)-\psi_k^{(n)}(\lambda)|\le e^{-nC},\quad |\psi_k^{(n)}(\pm L)|\le e^{-nC} \tag{27}$$

with some absolute C. Therefore from the very beginning we can take all integrals in (4), (8), (13) and (12) over the interval $[-L,L]$. Besides, observe that since V is an analytic function in $\Omega[d_1,d_2]$ (see (14)), for any $m\in\mathbb{N}$ there exists a polynomial V_m of the $(2m)$th degree such that

$$|V_m(z)|\le C_0,\quad |V(z)-V_m(z)|\le e^{-Cm},\quad z\in\Omega[d_1/2,d_2/2]. \tag{28}$$

Here and everywhere below we denote by $C,C_0,C_1,...$ positive n,m-independent constants (different in different formulas).

Take

$$m=[\log^2 n] \tag{29}$$

and consider the system of polynomials $\{p_k^{(n,L,m)}\}_{k=0}^\infty$ orthogonal in the interval $[-L,L]$ with respect to the weight $e^{-nV_m(\lambda)}$. Set $\psi_k^{(n,L,m)}=p_k^{(n,L,m)}e^{-nV_m/2}$ and construct $\mathcal{M}_m^{(0,n)}$ by (12) with $\psi_k^{(n,L,m)}$. Then for any $k\le n+2n^{1/2}$ and uniformly in $\lambda\in[-L,L]$

$$\begin{array}{l}|\psi_k^{(n,L)}(\lambda)-\psi_k^{(n,L,m)}(\lambda)|\le e^{-C\log^2 n},\quad |\varepsilon\psi_k^{(n,L)}(\lambda)-\varepsilon\psi_k^{(n,L,m)}(\lambda)|\le e^{-C\log^2 n}\\ ||\mathcal{M}_m^{(0,n)}-\mathcal{M}^{(0,n)}||\le e^{-C\log^2 n}.\end{array} \tag{30}$$

The proof of the first bound here is identical to the proof of (27) (see [15]). The second bound follows from the first one because the operator $\varepsilon: L_2[-L,L]\to C[-L,L]$ is bounded by L. The last bound in (30) follows from the first one and the inequality valid for the norm of an arbitrary matrix $\mathcal{A}$

$$||\mathcal{A}||^2\le\max_i\sum_j|A_{i,j}|\cdot\max_j\sum_i|A_{i,j}|. \tag{31}$$

Remark also that if for arbitrary matrices $\mathcal{A}, \mathcal{B}$ $||\mathcal{A}^{-1}|| \le C$ and $||\mathcal{A}-\mathcal{B}|| \le qC^{-1}$ with some $0 < q < 1$, then we can write $\mathcal{B} = \mathcal{A}(I - \mathcal{A}^{-1}(\mathcal{A}-\mathcal{B}))$. Since $||\mathcal{A}^{-1}(\mathcal{A}-\mathcal{B})|| \le q < 1$, $||(I - \mathcal{A}^{-1}(\mathcal{A}-\mathcal{B}))^{-1}|| \le (1-q)^{-1}$, (see any textbook on linear algebra). Thus $\mathcal{B}$ is invertible and $||\mathcal{B}^{-1}|| \le C(1-q)^{-1}$. Moreover, $||\mathcal{A}^{-1} - \mathcal{B}^{-1}|| \le q(1-q)^{-1}C^2$. Due to this simple observation and (30), we obtain that if $||(\mathcal{M}_m^{(0,n)})^{-1}|| \le C_1$, then $||(\mathcal{M}^{(0,n)})^{-1}|| \le C_1(1 - C_1 e^{-C\log^2 n})^{-1} \le 2C_1$ and

$$||(\mathcal{M}^{(0,n)})^{-1} - (\mathcal{M}_m^{(0,n)})^{-1}|| \le C_2 e^{-C\log^2 n}.$$

Using this bound combined with the first and the second bound of (30) we can compare each term of the kernel $S_{n,m}(\lambda,\mu)$ constructed by formula (11) with new orthogonal polynomials $\{p_k^{(n,L,m)}\}_{k=0}^{\infty}$ with the corresponding term of $S_n(\lambda,\mu)$. Then, since by the result of [14]

$$|\psi_k^{(n)}(\lambda)|^2 \le K_{n,2}(\lambda,\lambda) \le nC, \quad \lambda \in [-L, L],$$

and by the Schwarz inequality

$$|\epsilon\psi_k^{(n)}(\lambda)| \le (2L)^{1/2}||\psi_k^{(n)}||_2 \le (2L)^{1/2}, \quad \lambda \in [-L, L], \tag{32}$$

where $||.||_2$ is a standard norm in $L_2[-L, L]$, we obtain that uniformly in $\lambda, \mu \in [-L, L]$

$$|S_{n,m}(\lambda,\mu) - S_n(\lambda,\mu)| \le Cn^4 e^{-C\log^2 n} \le e^{-C'\log^2 n}. \tag{33}$$

Therefore below we will study $\mathcal{M}_m^{(0,n)}$ and $S_{n,m}(\lambda,\mu)$ instead of $\mathcal{M}^{(0,n)}$ and $S_n(\lambda,\mu)$. To simplify notations we omit the indexes m, L, but keep the dependence on m in the estimates.

Let us set our main notations. We denote by $\mathcal{H} = l_2(-\infty,\infty)$ the Hilbert space of all infinite sequences $\{x_i\}_{i=-\infty}^{\infty}$ with the standard scalar product $(.,.)$ and a norm $||.||$. Let also $\{e_i\}_{i=-\infty}^{\infty}$ be a standard basis in $\mathcal{H}$, and $\mathcal{I}^{(n_1,n_2)}$ with $-\infty \le n_1 < n_2 \le \infty$ be an orthogonal projection operator defined as

$$\mathcal{I}^{(n_1,n_2)} e_i = \begin{cases} e_i, & n_1 \le i < n_2, \\ 0, & \text{otherwise}. \end{cases} \tag{34}$$

For any infinite or semi infinite matrix $\mathcal{A} = \{A_{i,j}\}$ we denote by

$$\begin{aligned} &\mathcal{A}^{(n_1,n_2)} = \mathcal{I}^{(n_1,n_2)} \mathcal{A} \mathcal{I}^{(n_1,n_2)}, \\ &(\mathcal{A}^{(n_1,n_2)})^{-1} = \mathcal{I}^{(n_1,n_2)} \Big(I - \mathcal{I}^{(n_1,n_2)} + \mathcal{A}^{(n_1,n_2)} \Big)^{-1} \mathcal{I}^{(n_1,n_2)}, \end{aligned} \tag{35}$$

so that $(\mathcal{A}^{(n_1,n_2)})^{-1}$ is a block operator which is inverse to $\mathcal{A}^{(n_1,n_2)}$ in the space $I^{(n_1,n_2)}\mathcal{H}$ and zero on the $(I - I^{(n_1,n_2)})\mathcal{H}$. We denote also by $(.,.)_2$ a standard scalar product in $L_2[-L, L]$.

Set $\mathcal{V}^{(0,\infty)} = \{\mathcal{V}_{j,l}\}_{j,l=0}^{\infty}$, where

$$\mathcal{V}_{j,l} = \operatorname{sign}(l-j)(\psi_j^{(n)}, V'\psi_l^{(n)})_2 = \frac{2}{n} \begin{cases} (\psi_j^{(n)}, (\psi_l^{(n)})')_2, & j > l, \\ (\psi_j^{(n)}, (\psi_l^{(n)})')_2 + O(e^{-C\log^2 n}), & j \le l. \end{cases} \tag{36}$$

Here $O(e^{-C\log^2 n})$ appears because of the integration by parts and bounds (27), (30). Since $(\psi_k^{(n)})' = q_k e^{-nV/2}$, where q_k is a polynomial of the $(k+2m-1)$th degree, its Fourier expansion in the basis $\{\psi_k^{(n)}\}_{k=0}^{\infty}$ contains not more than $(k+2m-1)$ terms and for $|j-k| > 2m-1$ the jth coefficient is $O(e^{-C\log^2 n})$. Therefore for $k \le n + 2n^{1/2}$

$$n^{-1}(\psi_k^{(n)})' = \frac{1}{2}\sum_j \mathcal{V}_{j,k}\psi_j^{(n)} + O_2(e^{-C\log^2 n}). \tag{37}$$

Here and below we write $\phi(\lambda) = O_2(\varepsilon_n)$, if $||\phi||_2 \le C\varepsilon_n$. The above relation implies

$$\frac{1}{2}\epsilon\Big(\sum_j \mathcal{V}_{j,k}^{(0,\infty)}\psi_j^{(n)}\Big) = n^{-1}\psi_k^{(n)} + O_2(e^{-C\log^2 n}). \tag{38}$$

Hence, by (12), for $0 \le j, k \le n + 2n^{1/2}$

$$\frac{1}{2}\Big(\mathcal{M}^{(0,\infty)}\mathcal{V}^{(0,\infty)}\Big)_{j,k} = \delta_{j,k} + O(e^{-C\log^2 n}). \tag{39}$$

Thus,

$$\frac{1}{2}\mathcal{M}^{(0,n)}\mathcal{V}^{(0,n)} = I^{(0,n)} - \mu^{(0,n)}\nu^{(0,n)} + \mathcal{E}^{(0,n)}, \quad ||\mathcal{E}^{(0,n)}|| = O(e^{-C\log^2 n}), \tag{40}$$

where $\nu^{(0,n)}$ is a matrix with entries equal to zero except the block $(2m-1)\times(2m-1)$ in the right bottom corner and $\mu^{(0,n)}$ in (40) has $(n-2m+1)$ first columns equal to zero and the last $(2m-1)$ ones of the form

$$\mu_{l,n-2m-1+k}^{(0,n)} = M_{l,n-1+k}, \quad k = 1,\dots,2m-1, \quad l = 0,\dots,n-1.$$

The relation (40) was obtained in [8].

If we transpose the matrices in (40) we get

$$\frac{1}{2}\Big(\mathcal{V}^{(0,n)}\mathcal{M}^{(0,n)}\Big)_{j,k} = \delta_{j,k} - \sum_{l=1}^{2m-1} f_k^{(l)}\delta_{n-l,j} + \mathcal{E}^{(0,n)T}, \tag{41}$$

where $f^{(1)},\dots,f^{(2m-1)} \in \mathcal{H}^{(0,n)}$ are some vectors, whose form is not important for us.

The idea of the proof is to show that for $|j-n| \le N := 4[\log^2 n]$

$$\begin{aligned} &M_{k,j-1} - M_{k,j+1} = M_{j+1,k} - M_{j-1,k} = 2R_{k-j} + \varepsilon'_{j,k},\\ &\sum_{k=0}^{\infty}|\varepsilon'_{j,k}|^2 \le Cm^2N^2n^{-1}, \end{aligned} \tag{42}$$

where R_k is defined by (22).

Remark 2. To prove Theorems 1, 2 we need not to know $M_{j,k}$, but (42) allows to find the limiting expressions for $M_{j,k}$ up to some additive constant. Indeed, if we know, e.g., $M_{n-1,n}$, we can find $M_{n+2j,n+1+2k}$, going step by step from the point $(n-1,n)$ to $(n+2j+1,n+2k)$. Then, using the symmetry $M_{j,k}=-M_{k,j}$ we obtain $M_{n+2j,n+2k+1}$. Hence, since $M_{j,k}=0$ for even $j-k$ because of the evenness, we find in such a way all $M_{j,k}$ with $|j-n|,|k-n|\le[2\log^2 n]$. Thus, if we denote $C(n)=M_{n-1,n}-M_2$ (see (44) for the definition of M_2), then for odd $j-k$ we have

$$\begin{aligned} M_{j,k}&=M^*_{j,k}+\varepsilon_{j,k},\quad |\varepsilon_{2j-1,2k}|\le C_*Nmn^{-1/2}(1+|j-n|+|k-n|),\\ M^*_{j,k}&=M_{k-j+1}-\frac{1}{2}((1+(-1)^j)M_{-\infty}-(-1)^jC(n), \end{aligned}\tag{43}$$

where for odd k $M_k=0$ and for even k

$$\begin{aligned} M_k&=(1+(-1)^k)\sum_{j=k}^{\infty}R_j=P^{-1}(2)-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin(k-1)x\,dx}{P(2\cos x)\sin x},\\ M_{-\infty}&=2\sum_{j=-\infty}^{\infty}R_j=2P^{-1}(2), \end{aligned}\tag{44}$$

and P is defined in (19). It is possible to show also that $C(n)\to 0$, as $n\to\infty$ (see [20]). For the case $V(\lambda)=\lambda^{2p}+o(1)$ expressions for $M_{j,k}$ were found in [8].

Let us assume that we know (42) and obtain the assertion of Theorem 2. Define

$$\mathcal{Q}^{(0,n)}_{j,k}=\frac{1}{2}\begin{cases}\mathcal{V}^{(0,n)}_{j,k}, & \text{for}\quad 0\le j\le n-2m,\ 0\le k<n,\\ (\mathcal{R}^{(-\infty,n)})^{-1}\mathcal{D}^{(-\infty,n)})_{j,k}, & \text{for}\quad n-2m<j<n,\ 0\le k<n.\end{cases}\tag{45}$$

Remark that since $(\mathcal{R})^{-1}_{j,k}=\mathcal{P}_{j,k}=0$ for $|j-k|>2m-2$, the standard linear algebra yields that $(\mathcal{R}^{(0,n)})^{-1}$ possesses the same property, i.e.,

$$(\mathcal{R}^{(-\infty,n)})^{-1}_{j,k}=0,\ \text{for}\ |j-k|>2m-2\Rightarrow\mathcal{Q}^{(0,n)}_{j,k}=0,\ \text{for}\ |j-k|>2m-1.\tag{46}$$

It follows from (41) that for $0\le j\le n-2m$, $0\le k<n$

$$(\mathcal{Q}^{(0,n)}\mathcal{M}^{(0,n)})_{j,k}=\delta_{j,k}+O(e^{-C\log^2 n}).\tag{47}$$

For $n-2m<j\le n$, $0\le k<n$, using (42) and (46), we get

$$\begin{aligned}(\mathcal{Q}^{(0,n)}\mathcal{M}^{(0,n)})_{j,k}&=\sum_{l=n-4m}^{n-1}(\mathcal{R}^{(-\infty,n)})^{-1}_{j,l}(\mathcal{D}^{(-\infty,n)}\mathcal{M}^{(0,n)})_{l,k}\\ &=\sum_{l=n-4m}^{n-1}(\mathcal{R}^{(-\infty,n)})^{-1}_{j,l}(\mathcal{R}^{(-\infty,n)})_{l,k}+\sum_{l=n-4m}^{n-1}(\mathcal{R}^{(-\infty,n)})^{-1}_{j,l}\varepsilon'_{l,k}\\ &\quad-(\mathcal{R}^{(-\infty,n)})^{-1}_{j,n-1}M_{n,k}=\delta_{j,k}-(\mathcal{R}^{(-\infty,n)})^{-1}_{j,n-1}M_{n,k}+r_{j,k},\end{aligned}\tag{48}$$

where $\varepsilon'_{l,k}$ are defined in (42). According to (42),

$$\sum_{k=0}^{n-1}|r_{j,k}|^2=\sum_{k=0}^{n-1}\Big|\sum_{l=n-4m}^{n-1}(\mathcal{R}^{(-\infty,n)})^{-1}_{j,l}\varepsilon'_{l,k}\Big|^2\le CN^2m^3n^{-1}.$$

Hence,

$$\sum_{j=n-2m}^{n}\sum_{k=0}^{n-1}|r_{j,k}|^2\le CN^2m^4n^{-1}, \tag{49}$$

and we obtain from (47)-(49)

$$\mathcal{Q}^{(0,n)}\mathcal{M}^{(0,n)}=\mathcal{I}^{(0,n)}-\Pi+\widetilde{\mathcal{E}}^{(0,n)},\quad ||\widetilde{\mathcal{E}}^{(0,n)}||\le CNm^2n^{-1/2}, \tag{50}$$

where

$$\Pi x=\frac{1}{2}(x,\mu_1)a,\quad \text{with}\quad \mu_{1k}=M_{n,k},\quad a=(\mathcal{R}^{(-\infty,n)})^{-1}e_{n-1}. \tag{51}$$

Note that by (46) $a\in\mathcal{I}^{(n-2m,n)}\mathcal{H}$.

Now let x be an eigenvector of $\mathcal{M}^{(0,n)}$, corresponding to the eigenvalue $i\varepsilon_0$ ($||x||=1$). Since by definition (45)

$$||\mathcal{Q}^{(0,n)}||\le\max\{||\mathcal{V}^{(0,n)}||,2||(\mathcal{R}^{(0,n)})^{-1}||\}\le C_V,$$

we have from (50)

$$x=(x,\mu_1)a+y,\quad y=i\varepsilon_0\mathcal{Q}^{(0,n)}x-\widetilde{\mathcal{E}}^{(0,n)}x,\quad ||y||\le 2C_V|\varepsilon_0|. \tag{52}$$

But since $\mathcal{M}^{(0,n)}$ is a skew symmetric matrix of even dimensionality with real entries, if $i\varepsilon_0$ is its eigenvalue, $-i\varepsilon_0$ is its eigenvalue too (if $\varepsilon_0=0$, then this eigenvalue has multiplicity at least 2). Thus, there exists an eigenvector $x^{(1)}$ ($||x^{(1)}||=1$) such that $\mathcal{M}^{(0,n)}x^{(1)}=-i\varepsilon_0x^{(1)}$ and (50) implies

$$x^{(1)}=(x^{(1)},\mu_1)a+y^{(1)},\quad ||y^{(1)}||\le 2C_V|\varepsilon_0|. \tag{53}$$

Now it is easy to see that for $|\varepsilon_0|\le C_V^*$ with some C_V^*, depending only on C_V, relations (52) and (53) contradict to the condition

$$(x,x^{(1)})=0$$

valid for any eigenvectors of $\mathcal{M}^{(0,n)}$, corresponding to different eigenvalues. Thus, we conclude that $||(\mathcal{M}^{(0,n)})^{-1}||\le|C_V^*|^{-1}$.

Since Π from (50) satisfy the relation

$$\Pi^2=\lambda\Pi,\quad \text{with}\quad \lambda=(\mu_1,a), \tag{54}$$

(50) and the bound $||(\mathcal{M}^{(0,n)})^{-1}||\le|C_V^*|^{-1}$ imply $|1-\lambda|\ge C_V^*/2$. Thus,

$$(\mathcal{M}^{(0,n)})^{-1}=\mathcal{Q}^{(0,n)}+(1-\lambda)^{-1}\Pi\mathcal{Q}^{(0,n)}+\widetilde{\mathcal{E}}_1^{(0,n)}, \tag{55}$$
$$||\widetilde{\mathcal{E}}_1^{(0,n)}||\le CNm^{3/2}n^{-1/2}.$$

To finish the proof of Theorem 2 we are left to prove (42). Define $\mathcal{V}^* = \{\mathcal{V}^*_{j,l}\}_{j,l=-\infty}^{\infty}$ with $\mathcal{V}^*_{j,l} = \mathrm{sign}(l-j)V'(\mathcal{J}^*)_{j,l}$, where $\mathcal{J}^*$ is defined in (24). Then by the spectral theorem

$$\mathcal{V}^*_{j,l} = \mathcal{V}^*_{j-l} = \frac{\mathrm{sign}(l-j)}{2\pi}\int_{-\pi}^{\pi} dx V'(2\cos x)e^{i(j-l)x}. \tag{56}$$

The key point in the proof of (42) is the lemma:

Lemma 1. *Under conditions of Theorem 1*

$$\mathcal{V}^* = \mathcal{P}\mathcal{D} = \mathcal{D}\mathcal{P}, \tag{57}$$

where $\mathcal{P}$ and $\mathcal{D}$ are defined in (21) and (26) respectively.

Moreover, taking $N = 4[\log^2 n]$, we have

$$\mathcal{V}_{j,k} = \mathcal{D}\mathcal{P}_{j,k} + \widetilde{\varepsilon}_{j,k}, \quad |k-n| \le 2N, \ |j-n| \le 2N + 2m, \tag{58}$$

where $\widetilde{\varepsilon}_{j,k} = 0$, if $|j-k| > 2m-1$ and

$$|\widetilde{\varepsilon}_{j,k}| \le CNmn^{-1}, \quad if \quad |j-k| \le 2m-1. \tag{59}$$

We will use also

Proposition 1. *For any $j : |j-n| < 4N$*

$$||\epsilon\psi_j^{(n)}||_2^2 \le Cn^{-1}. \tag{60}$$

Since ε is a bounded operator in $L_2[-2-d/2, 2+d/2]$, by (38), (58) and (38), we have for $|k-n| < 2N$

$$\frac{1}{2}\sum_j \mathcal{P}^{(0,\infty)}_{j,k}\left(\epsilon\psi^{(n)}_{j-1} - \epsilon\psi^{(n)}_{j+1}\right) = n^{-1}\psi_k^{(n)} + r_k, \tag{61}$$

where $O_2(.)$ is defined in (37) and for $|k-n| < 2N$ we have by (60) and (59)

$$r_k := \frac{1}{2}\sum_{|j-k|\le 2m-1} \widetilde{\varepsilon}_{j,k}\epsilon\psi_j^{(n)} + O_2(e^{-C\log^2 n}) = O_2(Nmn^{-3/2}). \tag{62}$$

Let us extend (61) to all $0 \le k < \infty$, choosing r_k for $|k-n| \ge 2N$ in such a way to obtain for these k identical equalities:

$$r_k := \sum_{j>0} \mathcal{P}^{(0,\infty)}_{j,k}\left(\epsilon\psi^{(n)}_{j-1} - \epsilon\psi^{(n)}_{j+1}\right) - \frac{1}{n}\psi_k^{(n)} = O_2(1). \tag{63}$$

Applying $(\mathcal{P}^{(0,\infty)})^{-1}$ to both sides of (61), we get

$$\frac{1}{2}\left(\epsilon\psi^{(n)}_{j-1} - \epsilon\psi^{(n)}_{j+1}\right) = n^{-1}\sum_{k>0}(\mathcal{P}^{(0,\infty)})^{-1}_{k,j}\psi_k^{(n)} + \sum_{k>0}(\mathcal{P}^{(0,\infty)})^{-1}_{k,j} r_k = \Sigma_{1j} + \Sigma_{2j}. \tag{64}$$

Now we need some facts from the theory of Jacobi matrices.

Proposition 2. *Let J be a Jacobi matrix with entries $|J_{j,j+1}| \le 1 + d_1/4$ and Q be a bounded analytic function ($|Q(z)| \le C_0$) in $\Omega[d_1/2, d_2/2]$. Then:*
(i) for any j, k

$$|Q(\mathcal{J})_{j,k}| \le Ce^{-d|j-k|}; \tag{65}$$

(ii) if $\widetilde{\mathcal{J}}$ is another Jacobi matrix, satisfying the same conditions, then for any $j, k \in (n_1, n_2)$

$$|Q(\mathcal{J})_{j,k} - Q(\widetilde{\mathcal{J}})_{j,k}| \le Ce^{-d|j-k|} \sup_{i \in [n_1,n_2)} |J_{i,i+1} - \widetilde{J}_{i,i+1}| + C(e^{-d(|n_1-j|+|n_1-k|)} + e^{-d(|n_2-j|+|n_2-k|)}); \tag{66}$$

(iii) if $Q(\lambda) > \delta > 0$ for $\lambda \in [-2 - d_1/2, 2 + d_1/2]$, then for $i, j \in (n_1, n_2)$

$$|(Q(\mathcal{J})^{(n_1,n_2)})^{-1}_{j,k} - Q^{-1}(\mathcal{J})_{j,k}| \le C \min\left\{e^{-d|n_1-j|} + e^{-d|n_2-j|}, e^{-d|n_1-k|} + e^{-d|n_2-k|}\right\}, \tag{67}$$

where C and d depend only on d_1, d_2, C_0 and δ.

The proof of Proposition 2 is given at the end of Section 3.

Using (67) and (65) to estimate $(\mathcal{P}^{(0,\infty)})^{-1}_{j,k}$ and (62)-(63) to estimate r_k, it is easy to obtain that uniformly in $|j - n| \le N$

$$||\Sigma_{2j}||_2 \le C \sup_{|k-j|\le N} ||r_k||_2 + Ce^{-cN} \sup_k ||r_k||_2 \le CNmn^{-3/2}.$$

Besides, it follows from (67) that uniformly in $|j - n| \le N$

$$\Sigma_{1j} - \sum_{k>0} \mathcal{P}^{-1}_{j,k} \psi^{(n)}_k(\lambda) = O_2(e^{-cn}).$$

Hence, we obtain

$$\epsilon\psi^{(n)}_{j-1} - \epsilon\psi^{(n)}_{j+1} = 2n^{-1} \sum_{k>0} R_{j,k}\psi^{(n)}_k + O_2(Nmn^{-3/2}). \tag{68}$$

Multiplying the relation by $n\psi^{(n)}_k$, we get (42).
□

Corollary 1. *Under conditions of Theorem 1*

$$(\mathcal{M}^{(0,n)})^{-1}_{j,k} = \mathcal{Q}^{(0,n)}_{j,k} + \frac{1}{2}a_j b_k + O(n^{-1/2}\log^6 n), \tag{69}$$

where $\mathcal{Q}^{(0,n)}$, a are defined by (45) and (51) respectively, and

$$b_k = ((\mathcal{R}^{(-\infty,n)})^{-1} r^*)_k, \quad r^*_{n-i} = R_i \tag{70}$$

with R_i defined by (22).

Proof of Corollary 1. Using (55) we have for any $x \in \mathcal{I}^{(n-2m,n)}\mathcal{H}$, $||x|| \le 1$

$$2(\mathcal{M}^{(0,n)})^{-1}x = ((\mathcal{R}^{(-\infty,n)})^{-1}\mathcal{D}^{(-\infty,n)})x + (\nu,x)a + O(n^{-1/2}\log^6 n), \quad (71)$$

where a is defined by (51), ν is some unknown vector and we write $x = O(\varepsilon_n)$, if $||x|| = O(\varepsilon_n)$.

Making transposition of both sides of the last equation (recall that $\mathcal{M}^{(0,n)T} = -\mathcal{M}^{(0,n)}$ and $\mathcal{D}^{(-\infty,n)T} = -\mathcal{D}^{(-\infty,n)}$), we get for any $x \in \mathcal{I}^{(n-2m,n)}\mathcal{H}$

$$-2(\mathcal{M}^{(0,n)})^{-1}x = -(\mathcal{D}^{(-\infty,n)}(\mathcal{R}^{(-\infty,n)})^{-1})x + (x,a)\nu + O(n^{-1/2}\log^6 n). \quad (72)$$

Subtracting (71) from (72) we have

$$[(\mathcal{R}^{(-\infty,n)})^{-1}, \mathcal{D}^{(-\infty,n)}]x = -(a,x)\nu - (\nu,x)a + O(n^{-1/2}\log^6 n), \quad (73)$$

where the symbol $[.,.]$ means the commutator.

On the other hand, it is easy to see that

$$[\mathcal{D}^{(-\infty,n)}, \mathcal{R}^{(-\infty,n)}]x = (x,r^*)e_{n-1} + (x,e_{n-1})r^*,$$

where r^* is defined in (70). Hence,

$$[(\mathcal{R}^{(-\infty,n)})^{-1}, \mathcal{D}^{(-\infty,n)}]x = -(x,a)b - (x,b)a,$$

with a, b defined in (51) and (70). Using the last relation and (73), we obtain that for any $x \in \mathcal{I}^{(n-2m,n)}\mathcal{H}$

$$(x,a)b + (x,b)a = (a,x)\nu + (\nu,x)a + O(n^{-1/2}\log^6 n). \quad (74)$$

Taking an arbitrary x such that $(a,x) = (b,x) = 0$, we get that

$$\nu = \lambda_1 a + \lambda_2 b + O_2(n^{-1/2}\log^6 n)$$

Using this expression in (74), we obtain $\lambda_1 = O(n^{-1/2}\log^6 n)$, $\lambda_2 = 1 - O(n^{-1/2}\log^6 n)$. These relations combined with (71) prove (69).

□

Proof of Theorem 1. Substituting (69) in (11) and using (38), we obtain

$$S_n(\lambda,\mu) = K_{n,2}(\lambda,\mu) + nr_n(\lambda,\mu), \quad (75)$$

where $K_{n,2}(\lambda,\mu)$ is defined by (6) and

$$r_n(\lambda,\mu) = \sum_{j,k=-2m+1}^{2m-1} r_{j,k}\psi_{n-j}^{(n)}(\lambda)(\epsilon\psi_{n-j}^{(n)})(\mu), \quad |r_{j,k}| \le C. \quad (76)$$

According to the result of [23], to prove the weak convergence of all correlation functions it is enough to prove the weak convergence of cluster functions, which have the form

$$R_n(s_1,\ldots,s_k) = \frac{\mathrm{Tr}K_{n,1}(\lambda_0 + \frac{s_1}{n\rho(\lambda_0)}, \lambda_0 + \frac{s_2}{n\rho(\lambda_0)})\ldots K_{n,1}(\lambda_0 + \frac{s_1}{n\rho(\lambda_0)}, \lambda_0 + \frac{s_1}{n\rho(\lambda_0)})}{(n\rho(\lambda_0))^k}, \quad (77)$$

where the matrix kernel $K_{n,1}(\lambda,\mu)$ has the form (10) with

$$Sd_n(\lambda,\mu) = -n^{-1}\frac{\partial}{\partial\mu}S_n(\lambda,\mu), \quad IS_n(\lambda,\mu) = n\int \epsilon(\lambda-\lambda')S_n(\lambda',\mu)d\lambda'.$$

Define similarly

$$\begin{aligned} Kd_{n,2}(\lambda,\mu) &= n^{-1}\frac{\partial}{\partial\mu}K_{n,2}(\lambda,\mu), \quad IK_{n,2}(\lambda,\mu) = n\int \epsilon(\lambda-\lambda')K_{n,2}(\lambda',\mu)d\lambda' \\ rd_n(\lambda,\mu) &= -n^{-1}\frac{\partial}{\partial\mu}r_n(\lambda,\mu), \quad Ir_n(\lambda,\mu) = n\int \epsilon(\lambda-\lambda')r_n(\lambda',\mu)d\lambda'. \end{aligned} \tag{78}$$

Lemma 2. *Under conditions of Theorem 1 uniformly in* $|k-n|\le 2[\log^2 n]$ *and* $\lambda\in[-2+\delta,2-\delta]$

$$|\epsilon\psi_k^{(n)}(\lambda)| \le C\left[n^{-1} + (1-(-1)^k)n^{-1/2}\right]. \tag{79}$$

Moreover, for any compact $\mathbb{K}\subset\mathbb{R}$ *uniformly in* $s_1,s_2\in\mathbb{K}$

$$\left|\left(\frac{\partial}{\partial s_1}+\frac{\partial}{\partial s_2}\right)K_{n,2}(\lambda_0+s_1/n,\lambda_0+s_2/n)\right| \le C, \tag{80}$$

$$n^{-1}IK_{n,2}(\lambda_0+s_1/(n\rho(\lambda_0)),\lambda_0+s_2/(n\rho(\lambda_0))) \to \epsilon K^{(0)}_{\infty,2}(s_1-s_2). \tag{81}$$

Since in (75)-(76) $r_{j,k}=0$, if both j,k are odd, the bounds (79) and relations (75)-(76) yield that uniformly in $s_1,s_2\in\mathbb{K}$

$$\left\|\frac{1}{n}K_{n,1}(\lambda_0+\frac{s_1}{n\rho(\lambda_0)},\lambda_0+\frac{s_2}{n\rho(\lambda_0)}) - \frac{1}{n}\widetilde{K}_{n,1}(\lambda_0+\frac{s_1}{n\rho(\lambda_0)},\lambda_0+\frac{s_2}{n\rho(\lambda_0)})\right\| \le o(1), \tag{82}$$

where

$$\widetilde{K}_{n,1}(\lambda,\mu) = \begin{pmatrix} K_{n,2}(\lambda,\mu) & Kd_{n,2}S_n(\lambda,\mu) \\ IK_{n,2}(\lambda,\mu)-\epsilon(\lambda-\mu) & K_{n,2}(\mu,\lambda) \end{pmatrix}$$

Hence, we can replace $K_{n,1}$ by $\widetilde{K}_{n,1}$ in (77). Then, using integration by parts and (80), we obtain that the integral

$$I(a,b) = \int_a^b \dots \int_a^b R_n(s_1,\dots s_k)ds_1\dots ds_k$$

can be represented as a finite sum of the terms:

$$\begin{aligned} T(a,b;k_1,\dots,k_p;l_1,\dots,l_q) &= \int_a^b \dots \int_a^b ds_1\dots ds_k F_1(s_1,s_2)\dots F_k(s_k,s_1) \\ &\left(\delta(s_{k_1}-a)-\delta(s_{k_1}-b)+\dots+\delta(s_{k_p}-a)-\delta(s_{k_p}-b)\right), \end{aligned} \tag{83}$$

where

$$F_i(s,s') = \frac{1}{n\rho(\lambda_0)}\begin{cases} IK_{n,2}(\lambda_0+\frac{s}{n\rho(\lambda_0)},\lambda_0+\frac{s'}{n\rho(\lambda_0)})-\epsilon(s_1-s_2), & i=l_1.\dots,l_q, \\ K_{n,2}(\lambda_0+\frac{s}{n\rho(\lambda_0)},\lambda_0+\frac{s'}{n\rho(\lambda_0)}), & \text{otherwise}. \end{cases}$$

Using the result of [14] and (81) we can take the limit $n\to\infty$ in each of these term. Theorem 1 is proved.

3. Auxiliary results

Proof of Lemma 1. According to the standard theory of Toeplitz matrices

$$\mathcal{V}^*_{k,j} = \mathcal{V}^*_{k-j} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-j)x}\widetilde{\mathcal{V}}(x)dx,$$

where

$$\widetilde{\mathcal{V}}(x) = 2\sum_{k=1}^{\infty} V'_k \sin kx, \quad V'_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx}V'(2\cos x)dx, \tag{84}$$

and to prove (57) it is enough to prove that

$$\widetilde{\mathcal{V}}(x) = 2\sin x \cdot P(2\cos x). \tag{85}$$

Replacing in (19) $z \to 2\cos x$, $2\cos y \to (\zeta+\zeta^{-1})$, $dy \to (i\zeta)^{-1}d\zeta$ and using the Cauchy theorem, we get

$$P(2\cos x) = \frac{1}{2\pi i}\oint_{|\zeta|=1+\delta} \frac{V'(\zeta+\zeta^{-1}) - V'(2\cos x)}{\zeta+\zeta^{-1}-2\cos x}\zeta^{-1}d\zeta =$$
$$\frac{1}{2\pi i}\oint_{|\zeta|=1+\delta} \frac{\sum V'_k(\zeta^k+\zeta^{-k})d\zeta}{(\zeta-e^{ix})(\zeta-e^{-ix})} = \frac{1}{2\pi i}\oint_{|\zeta|=1+\delta} \frac{\sum V'_k\zeta^k d\zeta}{(\zeta-e^{ix})(\zeta-e^{-ix})}$$
$$= \sum_k V'_k \frac{\sin kx}{\sin x} = \frac{\widetilde{\mathcal{V}}(x)}{2\sin x}.$$

Since V' is a polynomial of $(2m-1)$th degree, $V'(\mathcal{J})_{j,k} = 0$ for $|j-k| \ge 2m$ and

$$|V'(\mathcal{J}^{(n)})_{j,k} - V'(\mathcal{J}^*)_{j,k}| \le Cm \max_{|l-k|\le 2m}\{|J_l^{(n)} - 1|\}.$$

This bound implies (59). □

Proof of Proposition 2. To prove (60) we use the result of [7], according to which for any $\delta > 0$ and any $\lambda \in (-2+\delta, 2-\delta)$ and $|k| \le 16[\log^2 n]$

$$\psi_{n+k}(\lambda) = \frac{2+\varepsilon_{n+k}}{\sqrt{2\pi}|4-\lambda^2|^{1/4}}\cos\left(n\pi\int_{\lambda}^{2}\rho(\mu)d\mu + k\gamma(\lambda) + \theta(\lambda) + o(1)\right) + O(n^{-1}), \tag{86}$$

where $\varepsilon_{n+k} \to 0$ does not depend on λ, $\rho(\lambda)$ is the limiting IDS, and $\gamma(\lambda)$, $\theta(\lambda)$ are smooth functions in $(-2,2)$.

Moreover, it follows from the result of [7] that there exists $\delta > 0$ such that for $|\lambda \mp 2| \le \delta$

$$\psi_{n+k}(\lambda) = n^{1/6}(B_1 + O(k/n))\mathrm{Ai}\left(n^{2/3}\left(\Phi_{\pm}(\lambda\mp 2) + k\gamma^{(1)}_{\pm}(\lambda)/n\right)\right) \tag{87}$$
$$+ n^{-1/6}(B_2 + O(k/n))\mathrm{Ai}'\left(n^{2/3}\left(\Phi_{\pm}(\lambda\mp 2) + k\gamma^{(2)}_{\pm}/n\right)\right) + O(n^{-1}),$$

where Ai is the Airy function, B_1, B_2 are some uniformly bounded constants, $\Phi_+, \gamma^{(1)}_+, \gamma^{(2)}_+$ are some functions analytic in $(2-\delta, 2+\delta)$, $\Phi_-, \gamma^{(1)}_-, \gamma^{(2)}_-$ are some functions analytic in $(-2-\delta, -2+\delta)$ and $\Phi'_+(2) \ne 0$, $\Phi'_-(-2) \ne 0$.

Integrating (86) and (87), we get

$$|\epsilon\psi_{n+k}(\lambda)| \le Cn^{-1/2}, \tag{88}$$

which implies (60).
□

Proof of Lemma 2. Integrating the first line of (86) between 0 and λ, we get

$$|\epsilon\psi^{(n)}_{n-k}(\lambda) - \epsilon\psi^{(n)}_{n-k}(0)| \le Cn^{-1}.$$

Then, using the fact that $(\epsilon f)(0) = 0$ for even f, we get (79) for even k (recall, that n is even). For odd k the above inequality imply

$$||\epsilon\psi^{(n)}_{n-k}||_2 \ge C|\epsilon\psi^{(n)}_{n-k}(0)| + Cm^2n^{-1}.$$

Combining the above bounds with (88), we get (79).

Inequality (80) follows from the result [14] (see Lemma 7), according to which

$$\Big|\Big(\frac{\partial}{\partial s_1} + \frac{\partial}{\partial s_2}\Big)K_{n,2}(\lambda_0 + s_1/n, \lambda_0 + s_2/n)\Big| \le C\Big(n^{-1}|s_1 - s_2|^2$$
$$+|\psi^{(n)}_n(\lambda_0+s_1/n)|^2+|\psi^{(n)}_{n-1}(\lambda_0+s_1/n)|^2+|\psi^{(n)}_n(\lambda_0+s_2/n)|^2+|\psi^{(n)}_{n-1}(\lambda_0+s_2/n)|^2\Big).$$

Since by(86) $\psi^{(n)}_n$, $\psi^{(n)}_{n-1}$ are uniformly bounded in each compact $\mathbb{K} \subset (-2, 2)$, we obtain (81).

To prove (81) we use the Christoffel-Darboux formula, which gives us

$$n^{-1}IK_{n,2}(\lambda, \mu) = \int_{|\lambda'-\lambda_0|\ge\delta'} \epsilon(\lambda - \lambda')\frac{\psi^{(n)}_n(\lambda')\psi^{(n)}_{n-1}(\mu) - \psi^{(n)}_{n-1}(\lambda')\psi^{(n)}_n(\mu)}{\lambda' - \mu}d\lambda'$$
$$+\int_{|\lambda'-\lambda_0|\le\delta'} \epsilon(\lambda - \lambda')\frac{\psi^{(n)}_n(\lambda')\psi^{(n)}_{n-1}(\mu) - \psi^{(n)}_{n-1}(\lambda')\psi^{(n)}_n(\mu)}{\lambda' - \mu}d\lambda' = I_1 + I_2 \tag{89}$$

Integrating by parts (we use again that $\psi^{(n)}_k = (\epsilon\psi^{(n)}_k)'$) and taking into account (88), we get

$$|I_1| \le C\delta^{-1}n^{-1/2} + \delta^{-2}\int_{-L}^{L}(|\epsilon\psi^{(n)}_n(\lambda')| + |\epsilon\psi^{(n)}_{n-1}(\lambda')|)d\lambda'$$
$$\le C\delta^{-1}n^{-1/2} + C\delta^{-2}(||\epsilon\psi^{(n)}_{n-1}||_2 + ||\epsilon\psi^{(n)}_n||_2) = O(n^{-1/2}).$$

To find I_2 observe that (86) yields for $\lambda, \mu \in (-2 + \varepsilon, 2 - \varepsilon)$

$$n^{-1}K_{n,2}(\lambda, \mu) = R(\lambda)\frac{\sin\Big(n\pi\int_\mu^\lambda \rho(\lambda')d\lambda'\Big)}{n(\lambda - \mu)}(1 + (\lambda - \mu)\phi_1(\lambda, \mu))$$
$$+n^{-1}\cos\Big(n\pi\int_\mu^\lambda \rho(\lambda')d\lambda'\Big)\phi_2(\lambda, \mu)+n^{-1}\cos\Big(n\pi(\int_2^\lambda + \int_2^\mu)\rho(\lambda')d\lambda'\Big)\phi_3(\lambda, \mu),$$

where R and ϕ_1, ϕ_2, ϕ_3 are smooth functions of λ. Hence, using the Riemann-Lebesgue theorem to estimate integrals with $\phi_i(\lambda,\mu)$, we obtain

$$I_2 = \int_{-n\delta'}^{n\delta'} ds' \epsilon(s_1 - s') R(\lambda_0 + s'/n) \frac{\sin\left(n\pi \int_{\lambda_0+s_2/n}^{\lambda_0+s'/n} \rho(\lambda')d\lambda'\right)}{s' - s_2} + o(1).$$

Now we split here the integration domain in two parts: $|s'| \le A$ and $|s'| \ge A$ and take the limits $n \to \infty$ and then $A \to \infty$. Relation (81) follows.
□

Proof of Proposition 2. Assertion (i) follows from the spectral theorem, according to which

$$Q(\mathcal{J})_{j,k} = \frac{1}{2\pi i} \oint_{d(z)=d} R_{j,k}(z) Q^{-1}(z) dz, \tag{90}$$

and the bound, valid for the resolvent $R(z) = (\mathcal{J} - z)^{-1}$ of any Jacobi matrix $\mathcal{J}$, satisfying conditions of the proposition (see [17])

$$|R_{j,k}| \le \frac{C}{d(z)} e^{-Cd(z)|j-k|}, \quad d(z) = \text{dist}\,\{z, [-2 - d_1/2, 2 + d_1/2]\}. \tag{91}$$

To prove assertion (ii) consider

$$\mathcal{J}(n_1,n_2) = \mathcal{J}^{(n_1,n_2)} + \mathcal{J}^{(-\infty,n_1)} + \mathcal{J}^{(n_2,\infty)}, \quad \widetilde{\mathcal{J}}(n_1,n_2) = \widetilde{\mathcal{J}}^{(n_1,n_2)} + \widetilde{\mathcal{J}}^{(-\infty,n_1)} + \widetilde{\mathcal{J}}^{(n_2,\infty)}$$

and denote

$$R^{(1)}(z) = (\mathcal{J}(n_1,n_2) - z)^{-1}, \quad R^{(2)}(z) = (\widetilde{\mathcal{J}}(n_1,n_2) - z)^{-1}, \quad \widetilde{R}(z) = (\widetilde{\mathcal{J}} - z)^{-1}.$$

It is evident that for $n_1 \le j,k \le n_2$ and $z \notin [-2,2]$

$$R^{(1)}_{j,k}(z) = (\mathcal{J}^{(n_1,n_2)} - z)^{-1}_{j,k}, \quad R^{(2)}_{j,k}(z) = (\widetilde{\mathcal{J}}^{(n_1,n_2)} - z)^{-1}_{j,k}.$$

Then, using the resolvent identity

$$H_1^{-1} - H_2^{-1} = H_1^{-1}(H_2 - H_1)H_2^{-1} \tag{92}$$

and (91), we get

$$|R^{(1)}_{j,k}(z) - R^{(2)}_{j,k}(z)| \le C \sup_{i \in [n_1,n_2)} |J_{i,i+1} - \widetilde{J}_{i,i+1}| \frac{e^{-d(z)|j-k|/2}}{d^2(z)}.$$

On the other hand, by (92) and (91), we obtain

$$\begin{aligned} |R_{j,k} - R^{(1)}_{j,k}| &\le |R_{j,n_1+1} R^{(1)}_{n_1,k}| + |R_{j,n_1} R^{(1)}_{n_1+1,k}| + |R_{j,n_2} R^{(1)}_{n_2-1,k}| + |R_{j,n_2-1} R^{(1)}_{n_2,k}| \\ &\le \frac{C}{d^2(z)} (e^{-d(z)(|n_1-j|+|n_1-k|)} + e^{-d(z)(|n_2-j|+|n_2-k|)}). \end{aligned}$$

Similar bound is valid for $|\widetilde{R}_{j,k} - R^{(2)}_{j,k}|$. Then (91) and (90) yield (66).

To prove assertion (iii) observe that $x_j = (Q(\mathcal{J})^{(n_1,n_2)})^{-1}_{j,k}$ is the solution of the infinite linear system:

$$\begin{aligned} &\sum Q(\mathcal{J})_{i,j} x_j = \delta_{i,k}, && i \in [n_1, n_2), \\ &\sum Q(\mathcal{J})_{i,j} x_j = r_i := \sum Q(\mathcal{J})_{i,j} (Q(\mathcal{J})^{(n_1,n_2)})^{-1}_{j,k}, && i \notin [n_1, n_2). \end{aligned}$$

Hence,

$$(Q(\mathcal{J})^{(n_1,n_2)})^{-1}_{j,k} = Q^{-1}(\mathcal{J})_{j,k} + \sum_{i \notin [n_1,n_2)} Q^{-1}(\mathcal{J})_{j,i} r_i$$

Now, using assertion (i), we obtain (67).
□

Acknowledgements. The author thanks P.Deift, T.Kriecherbauer and L.Pastur for the fruitful discussion. The author would like to thank also the unknown referee for the suggestion which helped to simplify the proof of Theorem 2. The author acknowledges the INTAS Research Network 03-51-6637 for the financial support.

References

1. Albeverio, S., Pastur, L., Shcherbina, M.: On Asymptotic Properties of the Jacobi Matrix Coefficients. Matem. Fizika, Analiz, Geometriya **4**, 263-277 (1997)
2. Albeverio, S., Pastur, L., Shcherbina, M.: On the $1/n$ expansion for some unitary invariant ensembles of random matrices. Commun. Math. Phys. **224**, 271-305 (2001)
3. Boutet de Monvel, A., Pastur L., Shcherbina M.: On the statistical mechanics approach in the random matrix theory. Integrated density of states. J. Stat. Phys. **79**, 585-611 (1995)
4. Claeys, T., Kuijalaars, A.B.J.: Universality of the double scaling limit in random matrix models, Comm. Pure Appl. Math. **59**, 1573-1603 (2006)
5. P. Deift, T. Kriecherbauer, K. T.- R. McLaughlin: New results on the equilibrium measure for logarithmic potentials in the presence of an external field, J. Approx. Theory **95**, 388–475 (1998)
6. Deift, P., Kriecherbauer, T., McLaughlin, K., Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. Commun. Pure Appl. Math. **52**, 1335-1425 (1999)
7. Deift, P., Kriecherbauer, T., McLaughlin, K., Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. Commun. Pure Appl. Math. **52**, 1491-1552 (1999)
8. Deift, P., Gioev, D.: Universality in random matrix theory for orthogonal and symplectic ensembles. Int. Math. Res. Papers.2007; 004-116
9. Deift, P., Gioev, D.: Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices Comm. Pure Appl. Math. **60**, 867-910 (2007)
10. Deift, P., Gioev, D., Kriecherbauer, T., Vanlessen, M.: Universality for orthogonal and symplectic Laguerre-type ensembles. J.Stat.Phys **129**, 949-1053 (2007)
11. Dyson, D.J.: A Class of Matrix Ensembles. J.Math.Phys.,**13**, 90-107 (1972)
12. Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. **91**, 151-204 (1998)
13. M.L.Mehta, M.L.: *Random Matrices.* New York: Academic Press, 1991
14. Pastur, L., Shcherbina, M.: Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Stat. Phys.* **86**, 109-147 (1997)
15. Pastur, L., Shcherbina, M.: On the edge universality of the local eigenvalue statistics of matrix models. Matematicheskaya fizika, analiz, geometriya **10**, N3, 335-365 (2003)
16. Pastur, L., Shcherbina, M.: Bulk universality and related properties of Hermitian matrix models **130**, 205-250 (2007)
17. Reed,M., Simon,B.:*Methods of Modern Mathematical Physics, Vol.IV*, Academic Press: New York, 1978

18. Saff, E., Totik, V.: *Logarithmic Potentials with External Fields*. Springer-Verlag, Berlin, 1997
19. Shcherbina, M.: Double scaling limit for matrix models with non analytic potentials. J. Math. Phys. **49**, 033501-033535 (2008)
20. M.Shcherbina. On Universality for Orthogonal Ensembles of Random Matrices Preprint arXiv:math-ph/0701046
21. Stojanovic, A.: Universality in orthogonal and symplectic invariant matrix models with quatric potentials. Math.Phys.Anal.Geom. **3**, 339-373 (2002)
22. Stojanovic, A.: Universalité pour des modéles orthogonale ou symplectiqua et a potentiel quartic. Math.Phys.Anal.Geom. Preprint Bibos 02-07-98
23. Tracy, C.A., Widom, H.: Correlation functions, cluster functions, and spacing distributions for random matrices. J.Stat.Phys. **92**, 809-835 (1998)
24. Widom, H.: On the relations between orthogonal, symplectic and unitary matrix models. J.Stat.Phys. **94**, 347-363 (1999)