Journal of Statistical Physics, Vol. 116, No. 5/6, September 2004 (© 2004)

Displacement Convexity for the Generalized Orthogonal Ensemble

Gordon Blower[1]

Received December 5, 2002; accepted February 17, 2004

The generalized orthogonal ensemble of $n \times n$ real symmetric matrices X has probability measure $\nu_n(dX) = Z_n^{-1} \exp\{-n \operatorname{trace} v(X)\} dX$ where dX is the product of Lebesgue measure on the matrix entries and $v(x) \geqslant (2+\delta) \log |x|$ with $\delta > 0$. The eigenvalue distribution is concentrated on $[-A/2, A/2]$ for some $A < \infty$. This paper establishes concentration and transportation inequalities for the distribution of eigenvalues of X under ν_n when v is twice differentiable with $v''(x) \geqslant -\kappa$ where $3A^2\kappa < 1$. If $v''(x) \geqslant \kappa_0 > 0$, or if the variance of the trace is $O(1/n^2)$, then the empirical distribution of eigenvalues converges weakly almost surely to some non-random probability measure on $[-A/2, A/2]$ as $n \to \infty$. These conditions are satisfied for certain polynomial potentials. The logarithmic energy is displacement convex as a functional on charge distributions, with fixed mean, along the real line. When the trace distribution satisfies a logarithmic Sobolev inequality, or equivalently a quadratic transportation inequality, the joint eigenvalue distributions and the limiting equilibrium measure likewise satisfy quadratic transportation inequalities in the sense of Talagrand.[24]

KEY WORDS: Random matrices; transportation; statistical mechanics.

1. INTRODUCTION AND MAIN RESULTS

This paper is concerned with the distribution of the eigenvalues of random matrices under the generalized orthogonal ensemble, as studied by Dyson, Boutet de Monvel, Pastur and Shcherbina,[6] and in the text of Mehta,[18] page 56. Let X be a real symmetric $n \times n$ matrix, and let dX be the product of the usual Lebesgue measure on the entries that are on

[1]Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, England, United Kingdom; e-mail: g.blower@lancaster.ac.uk

0022-4715/04/0900-1359/0 © 2004 Springer Science+Business Media, Inc.

or above the leading diagonal. There is a natural action of the orthogonal group by conjugation on such matrices

$$\mathrm{O}(n) \times M_n^s(\mathbb{R}) \to M_n^s(\mathbb{R}) : (U, X) \mapsto UXU^\dagger,$$

and the orbit of each X contains a unique diagonal matrix with leading diagonal entries in increasing order, which we identify with an element of

$$\Delta^n = \{\lambda = \lambda^{(n)} = (\lambda_j)_{j=1}^n \in \mathbb{R}^n : \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n\}.$$

Throughout this paper, the potential function for the ensemble is a real function v that is twice continuously differentiable and that satisfies $v(x) \geqslant (2+\delta)\log|x|$ for some $\delta > 0$ and all sufficiently large $|x|$. Then one can form the normalized trace $\tau = \frac{1}{n}\sum_{j=1}^n \lambda_j$ and $V(X) = \mathrm{trace}_n v(X) = \frac{1}{n}\sum_{j=1}^n v(\lambda_j)$ by functional calculus, and there exists Z_n with $0 < Z_n < \infty$ such that

$$\nu_n(dX) = Z_n^{-1}\exp\{-n^2 V(X)\}dX \tag{1.1}$$

defines a probability measure on $M_n^s(\mathbb{R})$. This $\nu_n(dX)$ is invariant under the orthogonal conjugation action on $M_n^s(\mathbb{R})$; hence it is termed the *generalized orthogonal ensemble.*

Here we consider the empirical distribution of eigenvalues of X, namely

$$\mu_n^{(\lambda)} = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j}, \tag{1.2}$$

where δ_{λ_j} denotes the unit point mass at the eigenvalue λ_j; typically the eigenvalues of X will be random, but are unlikely to be large. Let

$$\Omega_n^A = \{X \in M_n^s(\mathbb{R}) : -A/2 < \lambda_1 < \cdots < \lambda_n < A/2\},$$

which is invariant under orthogonal conjugation. Then Lemma 1 of ref. 6 asserts that there exist $c > 0$ and $A < \infty$ such that $\nu_n(\Omega_n^A) \geqslant 1 - e^{-cn}$. Hence there is no loss in restricting the ensemble to Ω_n^A and working with $\tilde{\nu}_n(dX) = \nu_n(\Omega_n^A)^{-1}\nu_n(dX)$, so that $\mu_n^{(\lambda)}$ is supported on $[-A/2, A/2]$. We impose the measure $\otimes_{n=1}^\infty \tilde{\nu}_n$ on $\prod_{n=1}^\infty M_n^s(\mathbb{R})$. Let $q_n(\tau)\,d\tau$ be the distribution of the normalized trace τ of X, where X is random subject to ν_n. The main result of this paper is the following.

Theorem 1.1. (i) Suppose that v is twice differentiable with $v''(x) \geqslant \kappa_0$ for all $x \in [-A/2, A/2]$, where $\kappa_0 > 0$. Then under the laws $\tilde{\nu}_n$, the empirical distributions of eigenvalues converge weakly almost surely to some non-random probability measure ρ which is supported in $[-A/2, A/2]$; so that

$$\int_{-\infty}^{\infty} f(x)\, \mu_n^{(\lambda)}(dx) \to \int_{-A/2}^{A/2} f(x)\, \rho(dx) \qquad (n \to \infty) \tag{1.3}$$

for almost all λ and for each bounded and continuous real function f.

(ii) Suppose that $v''(x) \geqslant -\kappa$, where $\kappa < 1/(3A^2)$, and that the variance of q_n satisfies

$$\int_{-\infty}^{\infty} |\tau - \bar{\tau}(n)|^2\, q_n(\tau)\, d\tau \leqslant \frac{C_1}{n^2} \qquad (n \geqslant 1) \tag{1.4}$$

where $\bar{\tau}(n) = \int \tau q_n(\tau) d\tau$ and C_1 is a constant. Then the same conclusions as in (i) hold.

Wigner[18] considered the case of the Gaussian ensembles, which arise when $v(x) = x^2/2$; moreover, when the potential is uniformly convex as in (i), the generalized orthogonal resembles the Gaussian orthogonal ensemble in many respects, as discussed in ref. 2. The hypotheses of (ii) are weaker than those of (i). In section 2 of ref. 22, Pastur and Shcherbina assert that (1.4) holds for unitarily invariant ensembles of Hermitian matrices, under very mild conditions on the potential. Their proof depends upon the orthogonal polynomial technique, so does not apply to orthogonal ensembles; nevertheless, (1.4) has been verified for certain orthogonal ensembles discussed below.

The form of the equilibrium distribution ρ was determined by Boutet de Monvel *et al.*[6] for orthogonal ensembles with a wide class of Hölder–continuous potentials. They show that ρ is absolutely continuous and the equilibrium density of states $p(x) = d\rho/dx$ satisfies $v(x) = \int \log|x - y|\, p(y)dy + c_1$ on the support of p for some constant c_1. When v is non-convex, p may have several local minima and the support of p may consist of several disjoint intervals; see section 2 of ref. 21. Boutet de Monvel *et al.*[6] show in their equation (2.2) that $\int |\tau - \bar{\tau}(n)|^2 q_n(\tau) d\tau \leqslant C(\log n)/n$ holds for general potentials.

In certain cases, for instance if v is convex as in Theorem 1.1(i), then $p(x)$ satisfies the principal value integral equation

$$p(x) = PV \frac{1}{\pi^2} \big[(b-x)(x-a)\big]^{1/2} \int_a^b \frac{v'(u)}{(x-u)[(b-u)(u-a)]^{1/2}} du \qquad (x \in [a, b]), \tag{1.5}$$

with constants $-A/2 \leqslant a < b \leqslant A/2$ that are determined by

$$\int_a^b p(x)\,dx = 1 \quad \text{and} \quad \int_a^b \frac{v'(u)}{[(b-u)(u-a)]^{1/2}}\,du = 0; \tag{1.6}$$

see ref. 19. Further, when v is a polynomial, we can use Tricomi's method to solve (1.5). Let $t_{2\ell-1,2k-1} = 2^{2-2k}\binom{2k-1}{k-\ell}$ for integers $1 \leqslant \ell \leqslant k$ and let U_j be the Chebyshev polynomials of the second kind, so that $U_j(\cos\theta) = \sin(j+1)\theta/\sin\theta$ for $j \geqslant 0$ as in ref. 13, section 8.94. We summarize some known results.

Proposition 1.2. Let $v(x) = \sum_{k=1}^{m} a_{2k}x^{2k}/2k$ be an even polynomial with $v(0) = 0$ and $a_{2m} > 0$. Suppose that the equilibrium distribution has support $[-b, b]$ where $b \leqslant A/2$.

(1) Then b is a root of

$$\sum_{k=1}^{m} 2^{-1} a_{2k} t_{1,2k-1} b^{2k} = 1, \tag{1.7}$$

and the equilibrium distribution has density

$$p(x) = \pi^{-1}\sqrt{b^2 - x^2} \sum_{k=1}^{m} \sum_{\ell=1}^{k} a_{2k} t_{2\ell-1,2k-1} b^{2k-1} U_{2\ell-2}(x/b). \tag{1.8}$$

(2) Moreover, if the polynomial factor in (1.8) has no real roots, then the trace distribution satisfies (1.4); indeed, the Gaussian concentration limit holds as $n \to \infty$:

$$\int_{-\infty}^{\infty} e^{nt\tau}\, q_n(\tau)\,d\tau \to e^{b^2t^2/4} \qquad (t \in \mathbb{R}). \tag{1.9}$$

Proof. 1. To solve (1.5), we express v' in terms of Chebyshev polynomials of the first kind, which satisfy $T_j(\cos\theta) = \cos j\theta$. Since v is even, we can suppose that $a = -b$ and use the trigonometric substitution $x = b\cos\theta$ in (1.5). The identity (1.8) follows easily from the identities on page 180 of ref. 25. The identity (1.7) is equivalent to the condition in (1.6) that p have integral equal to 1.

2. The trace is an example of a linear statistic in the eigenvalues, for which there is a central limit theorem for suitable polynomial potentials; see ref. 18, p. 315. Hence (1.9) is a special case of Johansson's Theorem 2.4 of ref. 15. To deduce (1.4) of Theorem 1.1(ii) we can take $t=1$ and use Chebyshev's inequality to obtain n_0 such that

$$\int_{\{\tau:|\tau|\geqslant s\}} q_n(\tau)\,d\tau \leqslant \big(1+e^{b^2/4}\big)e^{-ns} \qquad (s\geqslant 0; \quad n\geqslant n_0). \tag{1.10}$$

Then by integrating this concentration inequality we see that $\int \tau^2 q_n(\tau) d\tau \leqslant C/n^2$. See Propositions 6.5 and 6.6 for other conditions related to (1.4) and (1.9). ∎

Examples 1.3. We re-visit some examples that are discussed in ref. 6.

(i) Proposition 1.2 applies to any convex polynomial potential, and to non-convex polynomials such that $p(x)$ given by (1.9) is positive on $[-b,b]$.

(ii) Quartic potentials are used by Brézin[8] *et al.* in the planar approximation to field theory with global invariance group SO(n). Let us take $v(x)=x^4/4+(-\kappa)x^2/2$, which defines a potential with two wells. For small positive κ the equilibrium distribution is nevertheless supported on the interval $[-b,b]$ where $b^2=2(\kappa+\sqrt{\kappa^2+6})/3$, and is given by

$$p(x)=\frac{1}{\pi}\sqrt{b^2-x^2}\Big(x^2-\kappa+\frac{b^2}{2}\Big). \tag{1.11}$$

When $\kappa<1/20$ holds, the condition $3(2b)^2\kappa<1$ of Theorem 1.1(ii) is satisfied.

(iii) For the sextic $v(x)=x^6/6+a_4x^4/4+a_2x^2/2$, it is again possible to solve (1.7), using Cardan's method or Maple, and thus to obtain the coefficients of

$$p(x)=\frac{\sqrt{b^2-x^2}}{\pi}\left(x^4+\left(a_4+\frac{b^2}{2}\right)b^2x^2+a_2+\frac{3b^4}{8}+\frac{a_4b^2}{2}\right) \tag{1.12}$$

explicitly from (1.8). With $\delta=120a_2^3-27a_2^2a_4^2+810a_2a_4+2025-162a_4^3$ and $S=135a_2a_4+675-27a_4^3+15\sqrt{\delta}$ we have

$$b^2=\frac{2S^{1/3}}{15}-\frac{20a_2-6a_4^2}{5S^{1/3}}-\frac{2a_4}{5}.$$

One can check that $v(x)=x^6/6-3x^4/8+x^2/2$ is a nonconvex polynomial to which Theorem 1.1(ii) applies, and in this case the equilibrium distribution is trimodal.

(iv) For the potential $v(x)=|x|^p/p$ with $p\geqslant 2$, Theorem 1.6 of ref. 2 shows that $\nu_n\{X:|\tau(X)|\geqslant s\}\leqslant e^{-c_p n^2 s^p}$ for some constant $c_p>0$ and all $s\geqslant 0$. The proof of Theorem 1.1 in section 5 below works with this estimate.

We have not succeeded in obtaining a concentration inequality for any orthogonal ensemble such that the equilibrium density is supported on more than one interval. The main obstacle is the lack of any known inequality such as (1.10) in this context.

Boutet de Monvel *et al.*[6] established weak convergence in probability for the $\mu_n^{(\lambda)}$ to ρ. To obtain *almost sure* weak convergence, we prove a concentration inequality which shows that $\mu_n^{(\lambda)}$ is unlikely to deviate much from ρ when n is large; it is then straightforward to pair the measures with functions and deduce (1.3). This approach was used by the author in ref. 2 for uniformly convex potentials, and here we introduce significant technical refinements which allow us to deal with a wider and more realistic class of potentials.

Dyson and previously Wigner considered an analogy in electrostatics to describe the ensembles; see ref. 18 p. 70. If unit positive charge is distributed along the real line according to a probability measure μ and is subject to a potential field v, then the total energy is

$$E(\mu)=\int_{-\infty}^{\infty} v(x)\,\mu(dx)-\frac{1}{2}\iint_{[x\neq y]}\log|x-y|\,\mu(dx)\mu(dy); \tag{1.13}$$

here as in subsequent double integrals we exclude the diagonal $D=\{(x,x)\in\mathbb{R}^2\}$ since a point charge does not repel itself. A Radon probability measure μ is of *finite logarithmic energy* when $\int\int_{[x\neq y]}|\log|x-y||\mu(dx)\mu(dy)<\infty$. The ρ of Theorem 1.1 is the unique minimizer of this E over all probability measures of finite logarithmic energy; see p.27 of Saff and Totik[23] and ref. 16.

Whereas E is viewed most naturally as a functional on the probability measures on the line, its convexity properties are best interpreted in the phase space Δ^n of eigenvalues with its linear structure. In Theorem 2.1 we show that, for v as in Theorem 1.1, E is displacement convex in the sense of McCann[17], and hence we obtain the quantitative effect on E of rearranging the equilibrium configuration. In section three we consider the effect of displacement on the potential energy. In the context of Theorem 1.1 (ii), it seems to be necessary to condition the distribution of λ on the values of τ when proving displacement convexity. Transportation inequalities bound from above the cost of changing a measure μ into another measure ν by the relative entropy of μ with respect to ν. The required transportation inequality Theorem 4.1 follows for the conditional distribution of eigenvalues by a procedure due to Bobkov and Ledoux.[5] This implies a concentration inequality Theorem 4.2, with which we conclude the proof of Theorem 1.1 in section five.

In section six we show that, if the tracial distribution satisfies a suitable transportation inequality, then the unconditional joint eigenvalue distribution σ_n also satisfies a transportation inequality. Further, we present sufficient conditions for q_n to satisfy (1.4) and for σ_n to satisfy a transportation inequality with constants that improve with increasing n.

Under these slightly stronger hypotheses we are able to deduce analogues of results known in the case of Gaussian ensembles from refs. 2 and 24. Section seven features a transportation inequality for ρ. Displacement convexity has also been considered by Otto and Villani,[20] in the context of the Fokker-Planck equation, and by Carrillo *et al.*[9] for the granular medium diffusion equation, to establish logarithmic Sobolev and transportation inequalities.

2. DISPLACEMENT CONVEXITY OF LOGARITHMIC ENERGY

We begin this section by introducing some fundamental functionals to describe the problem, before stating the convexity properties of the energy. The Hamiltonian for our system may be written as a function on the phase space Δ^n of ordered eigenvalues

$$H(\lambda)=n\sum_{j=1}^{n} v(\lambda_j)-\sum_{j,k:j<k}\log|\lambda_j-\lambda_k| \qquad (\lambda\in\Delta^n), \tag{2.1}$$

or equivalently in terms of the empirical distribution as $n^2E(\mu_n^{(\lambda)})$. By restricting the ensemble, we can assume that all eigenvalues lie in $[-A/2, A/2]$. On Δ^n we shall use the metric associated with the norm $\|(x_j)\|_{\ell^2(n)}=\big(\frac{1}{n}\sum_{j=1}^n x_j^2\big)^{1/2}$ on $\mathbb{R}^n$.

Let (Ω, d) be a complete and separable metric space, and let μ and ν be Radon probability measures on Ω. When μ is absolutely continuous with respect to ν, we can unambiguously define the *relative entropy* of μ with respect to ν by

$$\mathrm{Ent}(\mu\mid\nu)=\int_\Omega \frac{d\mu}{d\nu}\log\frac{d\mu}{d\nu}d\nu,$$

where by Jensen's inequality $0\leqslant \mathrm{Ent}(\mu\mid\nu)\leqslant\infty$.

When $\int_\Omega d(x_0,y)^p\mu(dy)<\infty$ and $\int_\Omega d(x_0,y)^p\nu(dy)<\infty$ hold for some $x_0\in\Omega$ and $p>0$, we define the Wasserstein *p-transportation cost* to be

$$W_p(\mu,\nu)^p = \inf_\pi \left\{ \iint_{\Omega^2} d(x,y)^p \pi(dxdy) \,\middle|\, \pi \quad \text{has marginals} \quad \mu,\nu \right\}, \tag{2.2}$$

where the infimum is taken over all those Radon probability measures π with the prescribed marginals. The topology associated with the metric W_p for $p \geqslant 1$ is weaker than the topology of weak convergence of the probabilities. We shall use the dual characterization of transportation cost, due to Kantorovich and Rubinstein,[11] that

$$W_p(\mu,\nu)^p = \sup_{f,g} \left\{ \int_\Omega f(x)\,\mu(dx) - \int_\Omega g(y)\,\nu(dy) \,\middle|\, f(x) - g(y) \leqslant d(x,y)^p \right\} \tag{2.3}$$

where the functions f and g are bounded and continuous.

A continuous map $\psi:(\Omega_1,d_1)\to(\Omega_2,d_2)$ between complete and separable metric spaces is said to *induce* the Radon probability measure σ on Ω_2 from ν on Ω_1 when

$$\int_{\Omega_2} g(y)\,\sigma(dy) = \int_{\Omega_1} g(\psi(x))\nu(dx) \qquad (g \in C_b(\Omega_2)). \tag{2.4}$$

For notational convenience, we shall sometimes identify an absolutely continuous probability distribution on the real line with its associated probability density function. Let f_0 and f_1 be probability density functions on $[-A/2, A/2]$ and let $\varphi:[-A/2,A/2]\to[-A/2,A/2]$ be the increasing and continuous function that satisfies

$$\int_{-A/2}^{x} f_0(u)du = \int_{-A/2}^{\varphi(x)} f_1(u)du \qquad (x \in [-A/2, A/2]); \tag{2.5}$$

then φ induces $f_1(u)du$ from $f_0(u)du$, or, more briefly, f_1 from f_0. We introduce the increasing and continuous functions

$$\varphi_s(x) = (1-s)x + s\varphi(x) \qquad (x \in [-A/2, A/2], s \in [0,1]) \tag{2.6}$$

and set

$$\int_{-A/2}^{\varphi_s(x)} f_s(u)du = \int_{-A/2}^{x} f_0(u)du \qquad (x \in [-A/2, A/2]), \tag{2.7}$$

so that (f_s) $(0 \leqslant s \leqslant 1)$ is a family of probability density functions which *interpolates* between f_0 and f_1 by displacement in the sense of McCann.[17]

Theorem 2.1. Suppose that the probability density functions f_0 and f_1 on $[-A/2, A/2]$ have finite logarithmic energy.

(i) If $v''(x) \geqslant \kappa_0 > 0$, then the energy E of (1.13) is uniformly displacement convex; that is, $E(f_s)$ is a convex function of s with

$$(1-s)E(f_0) + sE(f_1) - E(f_s) \geqslant \frac{1}{2}s(1-s)\kappa_0 W_2(f_0, f_1)^2 \qquad (0 \leqslant s \leqslant 1). \tag{2.8}$$

(ii) If $v''(x) \geqslant -\kappa$ where $\kappa < 1/(3A^2)$, and f_0 and f_1 have equal means, then the energy E is uniformly displacement convex with

$$(1-s)E(f_0) + sE(f_1) - E(f_s) \geqslant s(1-s)\left(\frac{1-3A^2\kappa}{6A^2}\right)W_2(f_0, f_1)^2 \qquad (0 \leqslant s \leqslant 1). \tag{2.9}$$

Proof of Theorem 2.1. The proof is contained in sections two and three. First we recall that the logarithmic energy associated with f_s is

$$\frac{1}{2}\int\int_{[-A/2,A/2]^2} \log \frac{1}{|\varphi_s(x) - \varphi_s(y)|} f_0(x) f_0(y)\, dx\, dy. \tag{2.10}$$

The logarithmic energy in E makes a contribution to (2.8) of

$$\begin{aligned}
&\frac{1-s}{2}\int\int_{[-A/2,A/2]^2} \log \frac{1}{|x-y|} f_0(x) f_0(y)\, dx\, dy \\
&\quad + \frac{s}{2}\int\int_{[-A/2,A/2]^2} \log \frac{1}{|x-y|} f_1(x) f_1(y)\, dx\, dy \\
&\quad - \frac{1}{2}\int\int_{[-A/2,A/2]^2} \log \frac{1}{|x-y|} f_s(x) f_s(y)\, dx\, dy \\
&\quad = \frac{1}{2}\iint_{[-A/2,A/2]^2} \log \left\{ \frac{|\varphi_s(x) - \varphi_s(y)|}{|x-y|^{1-s}|\varphi(x) - \varphi(y)|^s} \right\} f_0(x) f_0(y)\, dx\, dy.
\end{aligned} \tag{2.11}$$

Here φ is increasing, so it follows from the inequality of the means that

$$\begin{aligned}
|\varphi_s(x) - \varphi_s(y)| &= |(1-s)(x-y) + s(\varphi(x) - \varphi(y))| \\
&\geqslant |x-y|^{1-s}|\varphi(x) - \varphi(y)|^s;
\end{aligned} \tag{2.12}$$

hence the integrand of (2.11) is non-negative. This shows that the logarithmic energy is displacement convex, so (2.10) is convex in s; a fact which turns out to be adequate for the proof of Theorem 2.1(i) in section three.

To obtain uniform convexity, as required for Theorem 2.1(ii), we improve upon (2.12) by using the inequality

$$\log\left\{\frac{|\varphi_s(x)-\varphi_s(y)|}{|x-y|^{1-s}|\varphi(x)-\varphi(y)|^s}\right\} \geqslant \frac{2}{3}s(1-s)\left\{\frac{x-\varphi(x)-y+\varphi(y)}{|x-y|+|\varphi(x)-\varphi(y)|}\right\}^2. \tag{2.13}$$

To see this, we set $\theta = x-y$ and $\psi = \varphi(x)-\varphi(y)$; here we suppose without loss that $\theta > 0$, and then $\psi > 0$ holds since φ is increasing. The function

$$h(s) = \log((1-s)\theta + s\psi) - (1-s)\log\theta - s\log\psi \qquad (s \in [0,1]) \tag{2.14}$$

is concave with $h(0) = h(1) = 0$ and hence satisfies the simple estimates

$$\begin{aligned} h(s) &\geqslant 2s(1-s)h(1/2) = 2s(1-s)\log\frac{\theta+\psi}{2\sqrt{\theta\psi}} \\ &= 2s(1-s)\log\left\{1+\frac{(\sqrt{\theta}-\sqrt{\psi})^2}{2\sqrt{\theta\psi}}\right\} \qquad (s\in[0,1]). \end{aligned}$$

The mean value theorem shows that $\log(1+u/2) \geqslant 2u/3$ for $0 \leqslant u \leqslant 1$; hence

$$h(s) \geqslant \frac{2}{3}s(1-s)\left\{\frac{\theta-\psi}{\theta+\psi}\right\}^2 \tag{2.15}$$

holds, and this is equivalent to (2.13). We deduce that

$$\begin{aligned} &\frac{1}{2}\int\int_{[-A/2,A/2]^2}\log\left\{\frac{|\varphi_s(x)-\varphi_s(y)|}{|x-y|^{1-s}|\varphi(x)-\varphi(y)|^s}\right\}f_0(x)f_0(y)dxdy \\ &\quad\geqslant \frac{s(1-s)}{12A^2}\int\int_{[-A/2,A/2]^2}\left\{(x-\varphi(x))-(y-\varphi(y))\right\}^2 f_0(x)f_0(y)dxdy, \end{aligned} \tag{2.16}$$

and since f_0 is a probability density function this is

$$\begin{aligned} &= \frac{s(1-s)}{6A^2}\left\{\int_{-A/2}^{A/2}(x-\varphi(x))^2 f_0(x)dx \right.\\ &\qquad\left. -\left(\int_{-A/2}^{A/2}(y-\varphi(y))f_0(y)dy\right)^2\right\}. \end{aligned} \tag{2.17}$$

As φ induces f_1 from f_0, the final integral in (2.17) vanishes

$$\int_{-A/2}^{A/2} y f_0(y)dy - \int_{-A/2}^{A/2} y f_1(y)dy = 0,$$

since f_0 and f_1 are assumed to have equal means.

Sudakov and Brenier[7] have shown that the most economical way of transporting f_0 to f_1 is via the probability measure π that is induced on $[-A/2, A/2]^2$ from $f_0(x)dx$ by $y \mapsto (y, \varphi(y))$; clearly π has marginal densities f_0 and f_1. Hence we have

$$(2.17) = \frac{s(1-s)}{6A^2} \int_{-A/2}^{A/2} (x - \varphi(x))^2 f_0(x)dx = \frac{s(1-s)}{6A^2} W_2(f_0, f_1)^2, \quad (2.18)$$

which gives our basic estimate on the effect on the logarithmic energy of displacing the charge distribution. In the next section we shall conclude the proof of Theorem 2.1 by considering the potential energy. ∎

3. DISPLACEMENT CONVEXITY FOR GENERAL POTENTIALS

Proposition 3.1. Suppose that $v''(x) \geqslant \alpha$ for all $x \in [-A/2, A/2]$. Then the potential energy satisfies

$$\begin{aligned}(1-s)\int_{-A/2}^{A/2} v(x) f_0(x)dx \; + \; s\int_{-A/2}^{A/2} v(x) f_1(x)dx - \int_{-A/2}^{A/2} v(x) f_s(x)dx \\ \geqslant \frac{\alpha s(1-s)}{2} W_2(f_0, f_1)^2. \quad (3.1)\end{aligned}$$

The potential energy is displacement convex whenever v is convex; the potential energy is uniformly displacement convex whenever $\alpha = \kappa_0 > 0$.

This result is known to McCann and others,[26] but for the sake of completeness, we include the proof.

Proof of Proposition 3.1. It follows from the mean value theorem that for $x, y \in \mathbb{R}$ and $s \in [0, 1]$, there exists $\bar{s} \in (0, 1)$ such that

$$\begin{aligned}(1-s)v(x) \; &+ s v(y) - v((1-s)x + sy) \\ &= \frac{1}{2} s(1-s)(x-y)^2 v''((1-\bar{s})x + \bar{s}y). \quad (3.2)\end{aligned}$$

Since φ_s induces f_s from f_0, we can write

$$\begin{aligned}(1-s)\int_{-A/2}^{A/2} v(x) f_0(x)dx + s\int_{-A/2}^{A/2} v(x) f_1(x)dx - \int_{-A/2}^{A/2} v(x) f_s(x)dx \\ = \int_{-A/2}^{A/2} \big\{(1-s)v(x) + sv(\varphi(x)) - v(\varphi_s(x))\big\} f_0(x)dx,\end{aligned} \tag{3.3}$$

where we can apply (3.2) with $y=\varphi(x)$ to the latest integrand and thereby obtain the bound

$$(3.3) \geqslant \frac{\alpha}{2} s(1-s)\int_{-A/2}^{A/2} \big\{\varphi(x)-x\big\}^2 f_0(x)dx. \tag{3.4}$$

The statement of the Proposition follows as with (2.18). ∎

Conclusion of the proof of Theorem 2.1. (i) When the potential is uniformly convex, we can take $\alpha=\kappa_0>0$ in Proposition 3.1. On adding (2.11) and (3.4), we obtain the required result (2.8).

(ii) When the potential satisfies the weaker estimate $v''\geqslant -\kappa$, and the densities have equal means, we obtain (2.9) by adding (2.18) to (3.4). This concludes the proof. ∎

We now consider the consequences for the empirical distribution of eigenvalues, as required for Theorem 1.1.

Corollary 3.2. (i) For v as in Theorem 2.1(i), the Hamiltonian is uniformly convex with

$$\begin{aligned}(1-s)H(\lambda)+sH(\xi) \;-\; & H((1-s)\lambda+s\xi) \\ & \geqslant 2^{-1}n^2 s(1-s)\kappa_0\|\lambda-\xi\|^2_{\ell^2(n)} \qquad (\lambda,\xi\in\Delta^n).\end{aligned}$$

(ii) For v as in Theorem 2.1(ii), the Hamiltonian satisfies

$$\begin{aligned}(1-s)H(\lambda) \;+\; & sH(\xi)-H((1-s)\lambda+s\xi) \\ & \geqslant n^2 s(1-s)\left(\frac{1-3A^2\kappa}{6A^2}\right)\|\lambda-\xi\|^2_{\ell^2(n)}\end{aligned} \tag{3.5}$$

for all $\lambda,\xi\in\Delta^n\cap[-A/2,A/2]^n$ with $\sum_{j=1}^n\lambda_j=\sum_{j=1}^n\xi_j$.

Proof. This may be verified directly as in the proof of Theorem 2.1. ∎

4. CONDITIONAL TRANSPORTATION AND CONCENTRATION

In ref. 14, Its *et al.* consider matrix ensembles, in which the eigenvalues $\lambda = (\lambda_1, \dots, \lambda_n)$ are constrained to lie on a hyperplane, as a model for a random word problem; here we consider such an ensemble as a technical device. Let $\mathbb{I}_{[-A/2,A/2]}$ be the indicator function of $[-A/2, A/2]$, and let

$$\sigma_n(d\lambda) = Z_n^{-1} \exp\{-H(\lambda)\}\mathbb{I}_{[-A/2,A/2]}(\lambda_1)\cdots\mathbb{I}_{[-A/2,A/2]}(\lambda_n)d\lambda_1 d\lambda_2 \dots d\lambda_n \tag{4.1}$$

be the probability measure on Δ^n that is induced from $\tilde{\nu}_n(dX)$ by the eigenvalue map $\Lambda : X \mapsto \lambda$. Ref. 18 p. 56 features a discussion of this formula.

It is convenient to condition measures with respect to the values taken by the normalized trace $\tau = \frac{1}{n}\sum_{j=1}^n \lambda_j$. Let q_n be the probability density function, induced by $\lambda \mapsto \tau$, that satisfies

$$\int_{-A/2}^{A/2} f(\tau) q_n(\tau)\, d\tau = \int_{\Delta^n} f\Big(\frac{1}{n}\sum_{j=1}^n \lambda_j\Big)\,\sigma_n(d\lambda) \qquad (f \in C_b(\mathbb{R})). \tag{4.2}$$

It is important to distinguish q_n from the integrated density of states, which involves $\frac{1}{n}\sum_{j=1}^n f(\lambda_j)$ instead of $f(\frac{1}{n}\sum_{j=1}^n \lambda_j)$ in (4.2), since their limiting behaviour is quite different; in fact, $q_n(\tau)\,d\tau$ converges weakly to $\delta_{\bar{x}}$ as $n \to \infty$ where $\bar{x} = \int x p(x) dx$. We suppose that $\sigma_n(d\lambda) = \sigma_{n,\tau}(d\chi) q_n(\tau) d\tau$ where $\sigma_{n,\tau}$ is the conditional probability measure on

$$\Pi_\tau^n = \Big\{\lambda \in \mathbb{R}^n : \frac{1}{n}\sum_{j=1}^n \lambda_j = \tau\Big\} \cap \Delta^n \cap [-A/2, A/2]^n. \tag{4.3}$$

This set is the closure of the intersection of a hyperplane in $\mathbb{R}^n$ with a convex open set, so has a natural Lebesgue measure $d\chi$ upon it, and we take the metric on Π_τ^n to be the $\ell^2(n)$ metric restricted to the said hyperplane. We shall use χ to denote a vector parallel to Π_τ^n, so that $\lambda = (\tau, \chi) \in \Pi_\tau^n$.

In this section we consider the joint distribution of eigenvalues of a random matrix subject to $\tilde{\nu}_n(dX)$. We shall assume as in Theorem 1.1(ii) that $3A^2\kappa < 1$ holds, and derive a concentration-of-measure theorem, itself a consequence of the following transportation inequality. Under the stronger hypotheses of Theorem 1.1(i), similar results hold with better constants as in Theorem 2.1(i); see section 6 of ref. 2.

Theorem 4.1. Let $\omega_{n,\tau}$ be a probability measure on Π_τ^n that is absolutely continuous and of finite relative entropy with respect to $\sigma_{n,\tau}$. Then the quadratic transportation cost is bounded by the relative entropy and satisfies

$$W_2(\omega_{n,\tau}, \sigma_{n,\tau})^2 \leqslant \frac{6A^2}{n^2(1-3A^2\kappa)} \mathrm{Ent}(\omega_{n,\tau} \mid \sigma_{n,\tau}). \tag{4.4}$$

Proof. This follows from Corollary 3.2(ii) by Proposition 4.2 of Bobkov and Ledoux.[5] Ultimately, their proof depends upon the Prékopa–Leindler inequality. ∎

The following concentration inequality is the functional form of Theorem 4.1, and we shall use it in the proof of almost sure convergence in Theorem 1.1(ii).

Theorem 4.2. Let $F : (\Pi_\tau^n, \ell^2(n)) \to \mathbb{R}$ be an L-Lipschitz function, so that $|F(\lambda) - F(\xi)| \leqslant L\|\lambda - \xi\|_{\ell^2(n)}$ for some $L < \infty$ and all $\lambda, \xi \in \Pi_\tau^n$. Suppose further that $\int F(\lambda)\, \sigma_{n,\tau}(d\chi) = 0$. Then

$$\int_{\Pi_\tau^n} \exp\{tF(\lambda)\}\, \sigma_{n,\tau}(d\chi) \leqslant \exp\left\{\frac{3t^2L^2A^2}{2n^2(1-3A^2\kappa)}\right\} \qquad (t \in \mathbb{R}). \tag{4.5}$$

Proof. This follows from Theorem 1.3 of Bobkov and Götze;[4] see also p. 342 of Villani.[26] ∎

We now present a logarithmic Sobolev inequality which formally strengthens Theorem 4.1. The precise connection between logarithmic Sobolev inequalities and transportation inequalities is discussed on p. 297 of Villani[26] within the unifying context of HWI inequalities.

Theorem 4.3. Let $g : \Pi_\tau^n \to \mathbb{R}$ be an $L^2(\sigma_{n,\tau})$ function such that $\|\nabla g\|_{\ell^2(n)}$ also belongs to $L^2(\sigma_{n,\tau})$. Then

$$\int_{\Pi_\tau^n} g(\chi)^2 \log\Big(g(\chi)^2 / \|g\|^2_{L^2(\sigma_{n,\tau})}\Big) \sigma_{n,\tau}(d\chi) \leqslant \frac{6A^2}{1-3A^2\kappa} \int_{\Pi_\tau^n} \|\nabla g(\chi)\|^2_{\ell^2(n)} \sigma_{n,\tau}(d\chi). \tag{4.6}$$

Proof. This follows from Proposition 3.2 of Bobkov and Ledoux.[5] ∎

5. ALMOST SURE WEAK CONVERGENCE

In this section we conclude the proof of Theorem 1.1 by arguments which exploit Theorem 4.2; compare ref. 2. The empirical distribution $\mu_n^{(\lambda)}$ is defined in (1.2).

Proposition 5.1. Suppose that v is as in Theorem 1.1. Then the empirical eigenvalue distribution converges weakly almost surely to the equilibrium distribution ρ as $n \to \infty$.

Proof. We shall concentrate on part (ii) of Theorem 1.1, this being the more difficult. By the Weierstrass approximation theorem, it suffices to prove (1.3) for an arbitrary L-Lipschitz function $f:[-A,A]\to\mathbb{R}$. We set

$$F_n(\lambda)=\int_{-A}^{A} f(x)\mu_n^{(\lambda)}(dx) \tag{5.1}$$

so that $F_n:(\Delta^n,\ell^2(n))\to\mathbb{R}$ defines an L-Lipschitz function. We introduce the means

$$m_{n,\tau}=\int_{\Pi_\tau^n} F_n(\lambda)\,\sigma_{n,\tau}(d\chi) \quad \text{and} \quad m_n=\int_{\Delta^n} F_n(\lambda)\,\sigma_n(d\lambda), \tag{5.2}$$

so that $m_n=\int m_{n,\tau}q_n(\tau)d\tau$ holds by definition of $\sigma_{n,\tau}$ and the limit

$$m_n\to m:=\int_{-A}^{A} f(x)\rho(dx) \qquad (n\to\infty) \tag{5.3}$$

holds by the weak convergence theorem of ref. 6.

Given $\varepsilon>0$, we shall prove that the sequence of probability values

$$\sigma_n\{\lambda\in\Delta^n:|F_n(\lambda)-m|>\varepsilon\}=\int_{-\infty}^{\infty}\sigma_{n,\tau}\{\lambda\in\Pi_\tau^n:|F_n(\lambda)-m|>\varepsilon\}q_n(\tau)d\tau \tag{5.4}$$

is summable over $n\geqslant 1$. To this end, we bound the right–hand side from above by

$$\int_{-\infty}^{\infty}\Big(\sigma_{n,\tau}\{\lambda\in\Pi_\tau^n:|F_n(\lambda)-m_{n,\tau}|>\varepsilon/4\}+\sigma_{n,\tau}\left[|m_{n,\tau}-m_{n,\bar\tau(n)}|>\varepsilon/4\right]$$
$$+\sigma_{n,\tau}\left[|m_{n,\bar\tau(n)}-m_n|>\varepsilon/4\right]+\sigma_{n,\tau}\left[|m_n-m|>\varepsilon/4\right]\Big)\,q_n(\tau)\,d\tau. \tag{5.5}$$

The concentration Theorem 4.2 leads via Chebyshev's inequality to the bound

$$\sigma_{n,\tau}\left\{\lambda\in\Pi_\tau^n : |F_n(\lambda)-m_{n,\tau}|>\varepsilon/4\right\}\leqslant 2\exp\left\{\frac{-\varepsilon t}{4}+\frac{3A^2L^2t^2}{2(1-3A^2\kappa)n^2}\right\}$$

and hence

$$\sigma_{n,\tau}\left\{\lambda\in\Pi_\tau^n : |F_n(\lambda)-m_{n,\tau}|>\varepsilon/4\right\}\leqslant 2\exp\left\{\frac{-n^2(1-3A^2\kappa)\varepsilon^2}{96A^2L^2}\right\};\qquad (5.6)$$

the constants here are independent of τ. This gives a satisfactory bound on the first term in (5.5).

The final term in (5.5) contributes zero for all sufficiently large n, independently of τ, on account of (5.3).

So it remains to bound the second and third terms of (5.5), which we do by establishing Lipschitz continuity of $m_{n,\tau}$ with respect to τ. For notational correctness, we introduce the probability measure $\tilde\sigma_{n,\tau}$ that is induced on Π_τ^n from $\sigma_{n,\tau}$ by the isometric map $\Phi:(\lambda_j)\mapsto(\lambda_j+\bar\tau(n)-\tau)$, which changes the trace from τ to $\bar\tau(n)$. We have

$$\begin{aligned} m_{n,\bar\tau(n)}-m_{n,\tau} &= \int_{\Pi^n_{\bar\tau(n)}} F(\lambda)\,\sigma_{n,\bar\tau(n)}(d\lambda)-\int_{\Pi^n_{\bar\tau(n)}} F(\lambda)\,\tilde\sigma_{n,\tau}(d\lambda)\\ &\quad+\int_{\Pi^n_\tau}(F(\Phi(\lambda))-F(\lambda))\,\sigma_{n,\tau}(d\lambda) \end{aligned}\qquad (5.7)$$

and since F is L-Lipschitz, this latest integral is bounded in modulus by $L\|\Phi(\lambda)-\lambda\|_{\ell^2(n)}\leqslant L|\tau-\bar\tau(n)|$. It follows from the Kantorovich-Rubinstein duality theorem (2.3) and Theorem 4.1 that

$$\begin{aligned} |m_{n,\bar\tau(n)}-m_{n,\tau}| &\leqslant LW_1(\sigma_{n,\bar\tau(n)},\tilde\sigma_{n,\tau})+L|\tau-\bar\tau(n)|\\ &\leqslant L\left(\frac{c_A}{n^2}\mathrm{Ent}\left(\sigma_{n,\bar\tau(n)}\,|\,\tilde\sigma_{n,\tau}\right)\right)^{1/2}+L|\tau-\bar\tau(n)| \end{aligned}\qquad (5.8)$$

for some constant c_A. In analogy with (2.1) and (4.1), the measures $\sigma_{n,\bar\tau(n)}$ and $\tilde\sigma_{n,\tau}$ arise from Hamiltonians which have difference

$$H_{\sigma_{n,\bar\tau(n)}}(\lambda)-H_{\tilde\sigma_{n,\tau}}(\lambda)=n\sum_{j=1}^n\left\{v(\lambda_j)-v(\lambda_j+\tau-\bar\tau(n))\right\}$$

since Φ does not affect the logarithmic energy term in (2.1), and the ratio of the normalizing constants of these probability measures is

$$\frac{Z(\tilde{\sigma}_{n,\tau})}{Z(\sigma_{n,\bar{\tau}(n)})} = \frac{\int_{\Pi^n_{\bar{\tau}(n)}} \exp\{-H_{\tilde{\sigma}_{n,\tau}}(\lambda)\}d\chi}{\int_{\Pi^n_{\bar{\tau}(n)}} \exp\{-H_{\sigma_{n,\bar{\tau}(n)}}(\lambda)\}d\chi}. \tag{5.9}$$

When all the λ_j have $-A \leqslant \lambda_j \leqslant A$ and $|v'(x)| \leqslant K$ for all $x \in [-A, A]$, we have

$$\left|H_{\sigma_{n,\bar{\tau}(n)}}(\lambda) - H_{\tilde{\sigma}_{n,\tau}}(\lambda)\right| \leqslant n^2 K |\tau - \bar{\tau}(n)|$$

and hence

$$\left|\log \frac{d\tilde{\sigma}_{n,\tau}}{d\sigma_{n,\bar{\tau}(n)}}\right| \leqslant n^2 K |\tau - \bar{\tau}(n)|. \tag{5.10}$$

On substituting the consequent bound on the relative entropy into (5.8), we obtain

$$|m_{n,\bar{\tau}(n)} - m_{n,\tau}| \leqslant L\big(K|\tau - \bar{\tau}(n)|\big)^{1/2} + L|\tau - \bar{\tau}(n)|. \tag{5.11}$$

When $|\tau - \bar{\tau}(n)| < s_0 := \min\{\varepsilon^2/(64KL^2), \varepsilon/(8L)\}$ we have $|m_{n,\bar{\tau}(n)} - m_{n,\tau}| < \varepsilon/4$ and hence

$$\begin{aligned}\int_{-\infty}^{\infty} \sigma_{n,\tau}\big[|m_{n,\tau} - m_{n,\bar{\tau}(n)}| > \varepsilon/4\big] q_n(\tau)\, d\tau &\leqslant \int_{\{\tau : |\tau - \bar{\tau}(n)| > s_0\}} q_n(\tau)\, d\tau \\ &\leqslant \frac{C_1}{n^2 s_0^2},\end{aligned} \tag{5.12}$$

where the last step follows from the variance inequality (1.4) via Chebyshev's inequality. We also have from (5.11) the bound

$$\begin{aligned}|m_{n,\bar{\tau}(n)} - m_n| &\leqslant \int_{-\infty}^{\infty} |m_{n,\bar{\tau}(n)} - m_{n,\tau}| q_n(\tau) d\tau \\ &\leqslant LK^{1/2} \int_{-\infty}^{\infty} |\tau - \bar{\tau}(n)|^{1/2} q_n(\tau)\, d\tau + L \int_{-\infty}^{\infty} |\tau - \bar{\tau}(n)| q_n(\tau)\, d\tau \\ &\leqslant LK^{1/2} \left(\int_{-\infty}^{\infty} |\tau - \bar{\tau}(n)|\, q_n(\tau)\, d\tau\right)^{1/2} \\ &\quad + L \int_{-\infty}^{\infty} |\tau - \bar{\tau}(n)|\, q_n(\tau)\, d\tau;\end{aligned} \tag{5.13}$$

where the latest step follows from the Cauchy-Schwarz inequality. By hypothesis (1.4) the right–hand side of (5.13) converges to zero as $n \to \infty$, and consequently we have $|m_{n,\bar{\tau}(n)} - m_n| < \varepsilon/4$ whenever n is sufficiently large; hence $\sigma_{n,\tau}[|m_{n,\bar{\tau}(n)} - m_n| > \varepsilon/4] = 0$ for all τ and large n.

From the preceding estimates, it is clear that only the first and second terms in (5.5) contribute when n is large; moreover from (5.6) and (5.12) the upper bound

$$\sigma_n\{\lambda \in \Delta^n : |F_n(\lambda) - m| > \varepsilon\} \leqslant \frac{C_1}{n^2 s_0^2} + 2\exp\left\{-\frac{n^2(1-3\kappa A^2)\varepsilon^2}{96A^2L^2}\right\} \tag{5.14}$$

follows for all sufficiently large n. By the first Borel-Cantelli Lemma, the probability of the event

$$\left\{(\lambda^{(n)}) \in \prod_{n=1}^{\infty} \Delta^n : |F_n(\lambda^{(n)}) - m| > \varepsilon \quad \text{for infinitely many} \quad n\right\} \tag{5.15}$$

is zero with respect to the measure $\otimes_{n=1}^{\infty}\sigma_n$. This establishes that the $\mu_n^{(\lambda)}$ converge almost surely to ρ in the weak topology as $n \to \infty$. ∎

6. TRANSPORTATION INEQUALITIES

In this section we present sufficient conditions for σ_n to satisfy a quadratic transportation inequality similar to (4.4). Having already achieved such a result for the eigenvalue distributions conditioned on the trace, we need to consider the tracial distribution $q_n(\tau)\,d\tau$.

Theorem 6.1. Let v be as in Theorem 1.1(ii) and suppose that $|v''| \leqslant K_2$ on $[-A/2, A/2]$. Suppose further that the tracial density function satisfies the quadratic transportation inequality

$$W_2(p_n, q_n)^2 \leqslant (1/\alpha_n)\mathrm{Ent}(p_n \mid q_n) \tag{6.1}$$

for all probability density functions p_n that are of finite relative entropy with respect to q_n. Then σ_n also satisfies the quadratic transportation inequality

$$W_2(w_n, \sigma_n)^2 \leqslant \max\left\{\frac{9A^4K_2^2}{\alpha_n(1-3A^2\kappa)^2}, \frac{6A^2}{n^2(1-3A^2\kappa)}\right\}\mathrm{Ent}(w_n \mid \sigma_n) \tag{6.2}$$

for any probability measure w_n that is absolutely continuous and of finite relative entropy with respect to σ_n.

Proof of Theorem 6.1. We let $g_n(\tau,\chi)=dw_n/d\sigma_n$, and also introduce $g_{n,\tau}(\chi) = g_n(\tau,\chi)/h_n(\tau)$, where $h_n(\tau)=\int_{\Pi^n_\tau} g_n(\tau,\chi)\sigma_{n,\tau}(d\chi)$ is so chosen that $g_{n,\tau}$ is a probability density function with respect to $\sigma_{n,\tau}$. Further, h_n is a probability density function with respect to $q_n(\tau)d\tau$. We shall obtain the Theorem from two lemmas which describe the quadratic transportation cost and the relative entropy. Talagrand[24] considered the corresponding results for product measures. ∎

Lemma 6.2. *Let $\varphi:[-A/2,A/2]\to[-A/2,A/2]$ be the continuous function that induces $h_n(\tau)q_n(\tau)d\tau$ from $q_n(\tau)d\tau$. Then the Wasserstein distances for the quadratic transportation cost satisfy*

$$\begin{aligned}3^{-1}W_2(w_n,\sigma_n)^2 \leqslant &\int_{-A/2}^{A/2} W_2(g_{n,\tau}\sigma_{n,\tau},\sigma_{n,\tau})^2 h_n(\tau)q_n(\tau)d\tau\\ &+W_2(h_nq_n,q_n)^2\\ &+\int_{-A/2}^{A/2} W_2(\sigma_{n,\varphi(\tau)},\sigma_{n,\tau})^2 q_n(\tau)d\tau. \end{aligned}\tag{6.3}$$

Proof. By the triangle inequality we have

$$\begin{aligned}W_2(g_{n,\tau}\sigma_{n,\tau}h_nq_n,\sigma_{n,\tau}q_n) \leqslant\ & W_2(g_{n,\tau}\sigma_{n,\tau}h_nq_n,\sigma_{n,\tau}h_nq_n)\\ &+W_2(\sigma_{n,\tau}h_nq_n,\sigma_{n,\varphi(\tau)}q_n)\\ &+W_2(\sigma_{n,\varphi(\tau)}q_n,\sigma_{n,\tau}q_n)\end{aligned}\tag{6.4}$$

and we can square up this inequality if we introduce the constant 3. We let $\psi_\tau:\Pi^n_\tau\to\Pi^n_\tau$ be the optimal transportation map that induces $g_{n,\tau}(\chi)\sigma_{n,\tau}(d\chi)$ from $\sigma_{n,\tau}(d\chi)$ at minimal W_2-cost. Then $(\tau,\chi)\mapsto(\tau,\psi_\tau(\chi))$ induces $g_{n,\tau}\sigma_{n,\tau}h_nq_n$ from $\sigma_{n\tau}h_nq_n$ and hence

$$\begin{aligned}&W_2(g_{n,\tau}\sigma_{n,\tau}h_nq_n,\sigma_{n,\tau}h_nq_n)^2\\ &\quad\leqslant\int_{-A/2}^{A/2}\left\{\int_{\Pi^n_\tau}\|\psi_\tau(\chi)-\chi\|^2_{\ell^2(n)}\sigma_{n,\tau}(d\chi)\right\} h_n(\tau)q_n(\tau)d\tau\end{aligned}\tag{6.5}$$

since the τ distribution is unchanged. As ψ_τ is the optimal transportation map, we can identify the inner integral with the minimal transportation cost for the χ-distribution and obtain

$$W_2(g_{n,\tau}\sigma_{n,\tau}h_nq_n,\sigma_{n,\tau}h_nq_n)^2\leqslant\int_{-A/2}^{A/2} W_2(g_{n,\tau}\sigma_{n,\tau},\sigma_{n,\tau})^2h_n(\tau)q_n(\tau)d\tau.\tag{6.6}$$

Hence the first term on the right–hand side of (6.4) gives rise to the first term on the right–hand side of (6.3).

Likewise, the final term in (6.4) gives rise to the final term in (6.3); when σ_n is the product of $q_n(\tau)d\tau$ with another measure, this term is zero since then $\sigma_{n,\tau} = \sigma_{n,\varphi(\tau)}$.

It remains to deal with the middle term in (6.4). The map $(\tau, \chi) \mapsto (\varphi(\tau), \chi)$ induces $\sigma_{n,\tau} h_n q_n$ from $\sigma_{n,\varphi(\tau)} q_n$ since

$$\int\int G(\varphi(\tau), \chi)\sigma_{n,\varphi(\tau)}(d\chi)\, q_n(\tau)d\tau = \iint G(\tau, \chi)\sigma_{n,\tau}(d\chi)\, h_n(\tau)q_n(\tau)\, d\tau \tag{6.7}$$

holds by the choice of φ. Further, since the χ- co-ordinate is unchanged, the cost is

$$W_2(\sigma_{n,\tau} h_n q_n, \sigma_{n,\varphi(\tau)} q_n)^2 \leqslant \int_{-A/2}^{A/2} |\varphi(\tau) - \tau|^2 q_n(\tau)\, d\tau = W_2(h_n q_n, q_n)^2.$$

This accounts for the remaining term in (6.3) and concludes the proof of Lemma 6.2. ∎

Lemma 6.3. The relative entropy of w_n with respect to σ_n satisfies

$$\mathrm{Ent}(w_n \mid \sigma_n) = \mathrm{Ent}(h_n q_n \mid q_n) + \int_{-A/2}^{A/2} \mathrm{Ent}(g_{n,\tau}\sigma_{n,\tau} \mid \sigma_{n,\tau}) h_n(\tau) q_n(\tau)\, d\tau. \tag{6.8}$$

Proof. The latest integral involves

$$\begin{aligned}
&\int_{-A/2}^{A/2} \Big\{ \int_{\Pi_\tau^n} g_n(\tau, \chi) \log \frac{g_n(\tau, \chi)}{h_n(\tau)}\, \sigma_{n,\tau}(d\chi) \Big\} q_n(\tau) d\tau \\
&\quad = \int_{\Delta^n} g_n(\lambda) \log g_n(\lambda)\, \sigma_n(d\lambda) \\
&\qquad - \int_{-A/2}^{A/2} \int_{\Pi_\tau^n} g_n(\tau, \chi) \log h_n(\tau)\, \sigma_{n,\tau}(d\chi)\, q_n(\tau)\, d\tau.
\end{aligned} \tag{6.9}$$

We evaluate the inner integral in the final double integral and obtain

$$(6.9) = \int_{\Delta^n} g_n(\lambda) \log g_n(\lambda)\, \sigma_n(d\lambda) - \int_{-A/2}^{A/2} h_n(\tau) \log h_n(\tau)\, q_n(\tau)\, d\tau. \tag{6.10}$$

The identity (6.8) follows when we rearrange this, and the proof of the Lemma is complete. ∎

Conclusion of the proof of Theorem 6.1. We obtain in turn upper bounds on each of the transportation costs in Lemma 6.2 in terms of corresponding expressions from Lemma 6.3. First we have, by the hypothesis (6.1) and (6.8),

$$W_2(h_nq_n, q_n)^2 \leqslant (1/\alpha_n)\mathrm{Ent}(h_nq_n \mid q_n) \leqslant (1/\alpha_n)\mathrm{Ent}(w_n \mid \sigma_n). \tag{6.11}$$

For the other terms in (6.3) we exploit Theorem 4.1. We have the inequality

$$\int_{-A/2}^{A/2} W_2(g_{n,\tau}\sigma_{n,\tau}, \sigma_{n,\tau})^2 h_n(\tau)q_n(\tau)\,d\tau \\ \leqslant \frac{6A^2}{n^2(1-3A^2\kappa)} \int_{-A/2}^{A/2} \mathrm{Ent}(g_{n,\tau}\sigma_{n,\tau} \mid \sigma_{n,\tau})h_n(\tau)q_n(\tau)\,d\tau, \tag{6.12}$$

and by Lemma 6.3 this is

$$\leqslant \frac{6A^2}{n^2(1-3A^2\kappa)}\mathrm{Ent}(w_n \mid \sigma_n). \tag{6.13}$$

To deal with the remaining term we introduce the Radon-Nikodym derivative $S_{n,\tau} = d\sigma_{n,\varphi(\tau)}/d\sigma_{n,\tau}$ and apply Theorem 4.1 to obtain

$$W_2(S_{n,\tau}\sigma_{n,\tau}, \sigma_{n,\tau})^2 \leqslant \frac{6A^2}{n^2(1-3A^2\kappa)}\mathrm{Ent}(S_{n,\tau}\sigma_{n,\tau} \mid \sigma_{n,\tau}), \tag{6.14}$$

and by Theorem 4.3 we have the logarithmic Sobolev inequality

$$\mathrm{Ent}(S_{n,\tau}\sigma_{n,\tau} \mid \sigma_{n,\tau}) \leqslant \frac{6A^2}{(1-3A^2\kappa)} \int_{\Pi_\tau^n} \|\nabla_\chi \sqrt{S}_{n,\tau}\|^2_{\ell^2(n)}\sigma_{n,\tau}(d\chi) \tag{6.15}$$

where ∇_χ is the projection of the gradient onto Π_τ^n. Now we can write $\lambda = \chi + \tau a$ where $a = (1, \ldots, 1)$ so that $\chi = (\chi_j)$ has $\sum_{j=1}^n \chi_j = 0$. In terms of these co-ordinates, (2.1) gives, after we cancel the logarithmic terms,

$$\log S_{n,\tau} = n\sum_{j=1}^n v(\chi_j + \tau) - v(\chi_j + \varphi(\tau)) + C_n \tag{6.16}$$

for some constant C_n, and this has gradient with norm squared

$$\begin{aligned}\|\nabla_\chi \log S_{n,\tau}\|^2_{\ell^2(n)} &= n\sum_{j=1}^{n}\big|v'(\chi_j+\tau)-v'(\chi_j+\varphi(\tau))\big|^2\\ &\leqslant n^2K_2^2(\varphi(\tau)-\tau)^2 \qquad (6.17)\end{aligned}$$

since $|v''|\leqslant K_2$. Hence on combining (6.17), (6.15) and (6.14) we obtain

$$W_2(S_{n,\tau}\sigma_{n,\tau},\sigma_{n,\tau})^2\leqslant\frac{9A^4K_2^2}{(1-3A^2\kappa)^2}(\varphi(\tau)-\tau)^2 \qquad (6.18)$$

and so

$$\int_{-A/2}^{A/2} W_2(\sigma_{n,\varphi(\tau)},\sigma_{n,\tau})^2 q_n(\tau)\,d\tau\leqslant\frac{9A^4K_2^2}{(1-3A^2\kappa)^2}\int_{-A/2}^{A/2}(\varphi(\tau)-\tau)^2 q_n(\tau)\,d\tau. \qquad (6.19)$$

Since φ is the optimal transportation map that takes $q_n(\tau)\,d\tau$ to $h_n(\tau)$ $q_n(\tau)\,d\tau$, we recognise this as

$$\frac{9A^4K_2^2}{(1-3A^2\kappa)^2}W_2(h_nq_n,q_n)^2.$$

We can now bound this term by a multiple of $\mathrm{Ent}(w_n\mid\sigma_n)$ as in (6.11). Having bounded each of the transportation costs in Lemma 6.2, we have achieved a proof of Theorem 6.1. ∎

Examples 6.4. (i) Let v be as in Theorem 1.1(i). Then the potential is uniformly 2-convex and conclusion of Theorem 6.1 holds by Theorems 1.4 and Lemma 6.3 of ref. 2; see also Bobkov and Ledoux.[(5)] In particular when $v(x)=x^2/2$, we have the classical Wigner ensemble $\mathrm{GOE}(n,1/n)$ where τ has a Gaussian $N(0,n^{-2})$ distribution and the quadratic transportation constant satisfies $1/\alpha_n\leqslant 2/n^2$ by Talagrand's theorem.[(24)]

(ii) Let v be a polynomial potential as in Proposition 1.2. Then v satisfies the Gaussian concentration inequality (1.10). By results of ref. 3 and Otto and Villani[(20)], it is known that the slightly stronger Gaussian isoperimetric inequality implies the quadratic transportation inequality.

The hypotheses of Theorem 6.1 imply that $q_n(\tau)\approx\exp\{-cn^2(\tau-\bar\tau(n))\}$ holds for some $c>0$ in the sense of the theory of large deviations; see ref. 12. In the remainder of this section, we shall make this more precise. Since q_n is a one-dimensional distribution, we can express the value of the quadratic transportation constant in terms of computable quantities. Otto and Villani[20] show, under the general conditions of their Theorem 1, that the transportation constant α_n in (6.1) is equivalent to the constant in the logarithmic Sobolev inequality

$$\int_{-\infty}^{\infty} h(\tau)\log h(\tau)\, q_n(\tau)\, d\tau \leqslant \frac{1}{4\alpha_n}\int_{-\infty}^{\infty} |\nabla\sqrt{h}(\tau)|^2\, q_n(\tau)\, d\tau \tag{6.20}$$

which holds for all probability density functions h of finite relative information with respect to q_n. Let the cumulative distribution function of q_n be $Q_n(x)=\int_{-A/2}^{x} q_n(t)\,dt$ and the tail be $T_n(x)=\int_x^{A/2} q_n(t)\,dt$; the median a_n satisfies $Q_n(a_n)=1/2$. We introduce the constants

$$B_n^{(1)} = \sup_{x\leqslant a_n} Q_n(x)\left\{\log\frac{1}{Q_n(x)}\right\}\int_x^{a_n}\frac{dt}{q_n(t)} \tag{6.21}$$

and

$$B_n^{(2)} = \sup_{x\geqslant a_n} T_n(x)\left\{\log\frac{1}{T_n(x)}\right\}\int_{a_n}^{x}\frac{dt}{q_n(t)}. \tag{6.22}$$

Then by Theorem 5.3 of Bobkov[4] and Götze, there exist absolute positive constants $c^{(1)}$ and $c^{(2)}$ such that

$$c^{(1)}(B_n^{(1)}+B_n^{(2)}) \leqslant 1/\alpha_n \leqslant c^{(2)}(B_n^{(1)}+B_n^{(2)}). \tag{6.23}$$

Further, they show that the logarithmic Sobolev inequality implies a Gaussian concentration inequality, as in Theorem 4.2, and (1.10).

The following result shows that the distribution of $\mathrm{trace}_n v'(X)$ is tightly concentrated near to its mean value; the inequality resembles (1.9), which holds for special polynomial potentials. For notational convenience, we use the unconditioned ensemble ν_n on $M_n^s(\mathbb{R})$.

Proposition 6.5. Suppose that v is twice continuously differentiable with $v''(x)\leqslant K$, and that $|v(x)|\geqslant cx^2$ holds for some $c>0$ and all large $|x|$. Then $V'(X)=\frac{1}{n}\sum_{j=1}^{n} v'(\lambda_j)$ satisfies $\int V'(X)\,\nu_n(dX)=0$ and

$$\int_{M_n^s(\mathbb{R})} \exp\{t\,nV'(X)\}\nu_n(dX) \leqslant \exp\left\{\frac{Kt^2}{2}\right\} \qquad (t\in\mathbb{R}). \tag{6.24}$$

Proof. We investigate the effect of translating X to $X+tI$; compare Lemma 1 of ref. 22. Since X and I commute, by checking the case of polynomial potentials and using the Weierstrass approximation theorem, we can easily show that $\frac{d^k}{dt^k}V(X+tI)=\mathrm{trace}_n v^{(k)}(X+tI)$, when v is k-times continuously differentiable and hence we deduce the inequality

$$V(X+tI)\leqslant V(X)+t\,\mathrm{trace}_n v'(X)+Kt^2/2 \tag{6.25}$$

by the mean value theorem.

The growth condition on v and the dominated convergence theorem allow us to manipulate the following integrals. The measure dX is translation invariant, and hence

$$1=\int_{M_n^s(\mathbb{R})}\nu_n(dX)=Z_n^{-1}\int_{M_n^s(\mathbb{R})}\exp\{-n^2V(X+tI)\}dX; \tag{6.26}$$

so the t-derivative of the right–hand side is zero, and

$$1\geqslant Z_n^{-1}\int_{M_n^s(\mathbb{R})}\exp\{-n^2Kt^2/2-n^2t\,\mathrm{trace}_n v'(X)-n^2V(X)\}dX \tag{6.27}$$

holds on account of (6.25). After rearranging, we obtain the desired result by replacing t by $-t/n$. ∎

Due to Lemma 1 of ref. 6 as mentioned in the Introduction, it is the bounds of v'' on $[-A/2, A, 2]$ that are important in applications of Proposition 6.5. We can now compute the asymptotic form of the variance of $V'(X)$.

Proposition 6.6. Suppose that v is a real analytic function that satisfies the conditions of Proposition 6.5. Then $V''(X)=\frac{1}{n}\sum_{j=1}^{n}v''(\lambda_j)$ satisfies

$$\int_{M_n^s(\mathbb{R})}V'(X)^2\nu_n(dX)=\frac{1}{n^2}\int_{M_n^s(\mathbb{R})}V''(X)\nu_n(dX) \tag{6.28}$$

and

$$\int_{M_n^s(\mathbb{R})}V''(X)\nu_n(dX)\to\frac{1}{2\pi}\int_{-\infty}^{\infty}|\xi||\hat{p}(\xi)|^2\,d\xi \qquad (n\to\infty). \tag{6.29}$$

Proof. We obtain (6.28) from (6.26) by differentiating twice and setting $t=0$. By results from ref. 6, the empirical distribution is weakly convergent in probability to the equilibrium distribution, hence the limit in (6.29) exists and has value $\int v''(x)p(x)dx$. To express the limit as a Fourier transform, we need to justify some computations. For v as in the Proposition, the equilibrium measure is supported on a finite union of disjoint intervals $[a_j, b_j]$, and on each $[a_j, b_j]$ its density p has the form $p(x) = \sqrt{(b_j - x)(x - a_j)}r_j(x)$, where $r_j(x) \geqslant 0$ is real and analytic; see ref. 10. It follows that p is bounded and vanishes at the endpoints of its supporting intervals, and that p' belongs to $L^{4/3}(\mathbb{R})$ since p' has singularities no worse than $x^{-1/2}$. Further, we can integrate by parts to obtain

$$\int_{-A/2}^{A/2} v''(x)p(x)\,dx = -\int_{-A/2}^{A/2} v'(x)p'(x)\,dx \tag{6.30}$$

where, on the support of p, this v' equals the Hilbert transform of p by results mentioned in the introduction. It follows from the Hausdorff–Young inequality that $\xi\hat{p}(\xi)$ belongs to $L^4(\mathbb{R})$ and $\hat{p}(\xi)$ belongs to $L^{4/3}(\mathbb{R})$. Hence we can apply Plancherel's formula to (6.30) and thus obtain the stated value for the limit. ∎

7. TRANSPORTATION OF THE EQUILIBRIUM DISTRIBUTION

Theorem 7.1. Let v be as in Theorem 1.1(ii), and suppose that the tracial distribution satisfies the quadratic transportation inequality (6.1) with $1/\alpha_n \leqslant \beta/n^2$ for some $\beta > 0$; alternatively, suppose that v is as in Theorem 1.1(i). Suppose further that the equilibrium density p has support $[-A/2, A/2]$, and let q be a probability density function on $[-A/2, A/2]$ with $q \log q$ integrable. Then the quadratic transportation cost satisfies, for some $\gamma > 0$ independent of q,

$$W_2(q, p)^2 \leqslant \frac{1}{\gamma} \iint \log \frac{1}{|x-y|}(p(x) - q(x))(p(y) - q(y))\,dx\,dy. \tag{7.1}$$

Thus the quadratic transportation cost is bounded by the logarithmic energy of $q - p$. The right–hand side is equivalent to the squared norm of $p - q$ in the Sobolev space $\dot{H}^{-1/2}[-A, A]$; that is, to $\sum' \frac{1}{|k|}|\hat{p}(k) - \hat{q}(k)|^2$ where the Fourier coefficients are taken with respect to $(\exp\{i\pi kx/A\})_{k=-\infty}^{\infty}$ where $k \neq 0$.

Proof. We recall that

$$v(x) = \int \log|x - y|\, p(y)\,dy + c_1 \qquad (x \in [-A/2, A/2]) \tag{7.2}$$

where c_1 is a constant which cancels later. Let $u(x)=\int \log|x-y|q(y)dy$, so that u defines a bounded function on $[-A/2, A/2]$; indeed by Young's inequality

$$\begin{aligned} u(x) &\leqslant \int_{-A/2}^{A/2} q(y)\log q(y)\,dy + \frac{1}{e}\int_{-A/2}^{A/2} |x-y|\,dy \\ &= \int_{-A/2}^{A/2} q(y)\log q(y)\,dy + \frac{A^2}{2e}. \end{aligned} \tag{7.3}$$

The function u is absolutely continuous and its derivative $u'(x) = PV\int (1/(x-y))\,q(y)dy$ is integrable by Kolmogorov's Theorem on the Hilbert transform; in particular, u is Hölder continuous. Since q satisfies (1.8)–(1.12) of ref. 6, it is the unique probability density function on $[-A/2, A/2]$ that has u as its logarithmic potential.

Now we introduce

$$\begin{aligned} \omega_n(d\lambda) = Z_n(u)^{-1}\exp\Big\{-n\sum_{j=1}^{n} u(\lambda_j) + \sum_{j,k:j<k} \log|\lambda_j-\lambda_k|\Big\} \\ \times \prod_{j=1}^{n} \mathbb{I}_{[-A/2,A/2]}(\lambda_j)d\lambda_1\dots d\lambda_n \end{aligned} \tag{7.4}$$

where $Z_n(u)$ is so chosen as to make $\omega_n(d\lambda)$ be a probability measure on Δ^n; here $0<Z_n(u)<\infty$ holds since u is bounded. This $\omega_n(d\lambda)$ has finite relative entropy with respect to $\sigma_n(d\lambda)$ and

$$\frac{1}{n^2}\int_{\Delta^n}\frac{d\omega_n}{d\sigma_n}\log\frac{d\omega_n}{d\sigma_n}d\sigma_n = \frac{1}{n^2}\log\frac{Z_n}{Z_n(u)} + \int_{\Delta^n}\frac{1}{n}\sum_{j=1}^{n}\big(v(\lambda_j)-u(\lambda_j)\big)\,\omega_n(d\lambda). \tag{7.5}$$

By Theorem 1 of ref. 6 we have weak convergence of the empirical eigenvalue distributions under $\omega_n(d\lambda)$; hence

$$\begin{aligned} \int_{\Delta^n}\frac{1}{n}\sum_{j=1}^{n} v(\lambda_j)\,\omega_n(d\lambda) &\to \int_{\mathbb{R}} v(x)q(x)\,dx \\ &= \iint_{\mathbb{R}^2}\log|x-y|\,p(y)q(x)dydx + c_1 \end{aligned} \tag{7.6}$$

as $n \to \infty$, and similarly

$$\begin{aligned}\int_{\Delta^n} \frac{1}{n} \sum_{j=1}^{n} u(\lambda_j)\, \omega_n(d\lambda) &\to \int_{\mathbb{R}} u(x) q(x)\, dx \\ &= \int\int_{\mathbb{R}^2} \log|x-y|\, q(y) q(x)\, dx dy. \qquad (7.7)\end{aligned}$$

The latest iterated integral is finite since u is bounded and q is integrable.

By Jensen's inequality applied to (7.5), we have

$$\frac{1}{n^2} \log \frac{Z_n}{Z_n(u)} \leqslant \int_{\Delta^n} \frac{1}{n} \sum_{j=1}^{n} \big(u(\lambda_j) - v(\lambda_j)\big)\, \sigma_n(d\lambda) \qquad (7.8)$$

and hence by weak convergence

$$\limsup_{n\to\infty} \frac{1}{n^2} \log \frac{Z_n}{Z_n(u)} \leqslant \int_{\mathbb{R}} \big(u(x) - v(x)\big) p(x)\, dx. \qquad (7.9)$$

The right–hand side of (7.9) may be simplified using (7.2) and the definition of u to

$$\int\int_{\mathbb{R}^2} \log|x-y|\, (q(y) - p(y)) p(x)\, dx dy - c_1.$$

We deduce from (7.6) that

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n^2} \int_{\Delta^n} \frac{d\omega_n}{d\sigma_n} \log \frac{d\omega_n}{d\sigma_n} d\sigma_n \\ &\quad\leqslant \iint_{\mathbb{R}^2} \log \frac{1}{|x-y|} \big(p(x) - q(x)\big)\big(p(y) - q(y)\big) dx dy. \qquad (7.10)\end{aligned}$$

Moving attention to the left–hand side of (7.1), we can introduce, for each $\varepsilon > 0$, continuous functions f and g with $f(x) - g(y) \leqslant |x-y|^2$ on $[-A/2, A/2]$ and such that

$$\begin{aligned}W_2(q,p)^2 &\leqslant \int_{\mathbb{R}} f(x) q(x)\, dx - \int_{\mathbb{R}} g(y) p(y)\, dy + \varepsilon \\ &\leqslant \int_{\Delta^n} \frac{1}{n} \sum_{j=1}^{n} f(\lambda_j)\, \omega_n(d\lambda) \\ &\qquad - \int_{\Delta^n} \frac{1}{n} \sum_{j=1}^{n} g(\xi_j)\, \sigma_n(d\xi) + 2\varepsilon \qquad (7.11)\end{aligned}$$

holds for all sufficiently large n. It follows from the choice of f and g that

$$W_2(q,p)^2 \leqslant \iint_{\Delta^n\times\Delta^n} \|\lambda-\xi\|^2_{\ell^2(n)}\pi_n(d\lambda d\xi)+2\varepsilon \tag{7.12}$$

holds for all probability measures π_n on $\Pi^n_\tau\times\Pi^n_\tau$ with marginals $\omega_n(d\lambda)$ and $\sigma_n(d\xi)$. By Theorem 6.1, σ_n satisfies a quadratic transportation inequality with constant $1/(\gamma n^2)$ for some $\gamma>0$. We deduce from the definition of transportation cost and (7.12) that

$$W_2(q,p)^2 \leqslant \limsup_{n\to\infty} W_2(\omega_n,\sigma_n)^2 \leqslant \limsup_{n\to\infty}\frac{1}{\gamma n^2}\int_{\Delta^n}\frac{d\omega_n}{d\sigma_n}\log\frac{d\omega_n}{d\sigma_n}d\sigma_n. \tag{7.13}$$

The result follows when we combine (7.13) with (7.10).

Remark 1. When $v(x)=x^2$ and we have the semicircle law $p(x)=2\sqrt{1-x^2}/\pi$ for $-1<x<1$, the inequality (7.1) simplifies to Biane[1] and Voiculescu's Remark 2.9. The constants arising from our double integral are equivalent to those of ref. 1; checking this is a pleasant exercise in the theory of the Gamma function.

ACKNOWLEDGMENTS

I am grateful to G. Jameson and L. Pastur for helpful conversations, and to M. Ledoux for pointing out a mistake in a previous version of this paper. Referees also provided helpful advice on the exposition. This work was supported by the TMR grant 'Analysis and Operators.'

REFERENCES

1. P. Biane and D. Voiculescu, A free probability analogue of the Wasserstein metric on the trace-state space, *Geom. Funct. Anal.* **11**:1125–1138 (2001).
2. G. Blower, Almost sure weak convergence for the generalized orthogonal ensemble, *J. Statist. Phys.* **105**:309–355 (2001).
3. G. Blower, The Gaussian isoperimetric inequality and concentration, *Positivity* **7**:203–224 (2003).
4. S. G. Bobkov and F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.* **163**:1–28 (1999).
5. S. G. Bobkov and M. Ledoux, From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities, *Geom. Funct. Anal.* **10**:1028–1052 (2000).
6. A. Boutet de Monvel, L. Pastur, and M. Shcherbina, On the statistical mechanics approach in the random matrix theory: integrated density of states, *J. Statist. Phys.* **79**:585–611 (1995).
7. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* **44**:375–417 (1991).

8. E. Brézin, C. Itzykson, G. Parisi, and J.B. Zuber, Planar diagrams, *Commun. Math. Phys*. **59**:35–51 (1978).
9. J. A. Carrillo, R. J. McCann, and C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, To appear.
10. P. Deift, T. Kriecherbauer, T.-R. McLaughlin, S. Venakides, and X. Zhou, Asymptotics for polynomials orthogonal with respect to varying exponential weights. *Int. Math. Res. Not*. **16**:759–782 (1997).
11. R. M. Dudley, *Real Analysis and Probability* (Wadsworth & Brooks/Cole, Pacific Grove, 1989).
12. R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics* (Springer-Verlag, New York, 1985).
13. I. S. Gradsteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* 5th ed. (Academic Press, San Diego, 1994).
14. A. R. Its, C. A. Tracy, and H. Widom, Random words, Toeplitz determinants and integrable systems I, in *Random Matrix Models and Their Applications*, P.M. Pavel and A.R. Its, eds., MSRI Publications, Vol. 40, (Cambridge University Press, Cambridge, 2001), pp. 245–258.
15. K. Johansson, On fluctuations of eigenvalues of random hermitian matrices. Duke Math. J. **91**:151–204 (1998).
16. M. K.-H. Kiessling and H. Spohn, A note on the eigenvalue density of random matrices. Comm. Math. Phys. **199**:683–695 (1999).
17. R. J. McCann, A convexity principle for interacting gases. *Adv. Math.* **128**:153–179 (1997).
18. M. L. Mehta, Random Matrices, 2nd ed. (Academic Press, San Diego, 1991).
19. N. I. Muskhelishvili, *Singular Integral Equations: Boundary Problems of Function Theory and Their Applications to Mathematical Physics* (P. Noordhoff N.V., Groningen, 1953).
20. F. Otto and C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173**:361–400 (2000).
21. L. A. Pastur, Spectral and probabilistic aspects of matrix models, *Algebraic and Geometric Methods in Mathematical Physics*, A. Boutet de Monvel and V. A. Marchenko, eds. (Kluwer, Dordrecht, 1996) pp. 205–242.
22. L. Pastur and M. Shcherbina, Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. *J. Statist. Phys*. **86**:109–147 (1997).
23. E. B. Saff and V. Totik, *Logarithmic Potentials With External Fields* (Springer, Berlin, 1997).
24. M. Talagrand, Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.* **6**:587–600 (1996).
25. F. G. Tricomi, *Integral Equations* (Interscience, New York, 1957).
26. C. Villani, *Topics in Optimal Transportation* (American Mathematical Society, Rhode Island, 2003).