QUANTITATIVE NORMAL APPROXIMATION OF LINEAR STATISTICS OF β-ENSEMBLES

BY GAULTIER LAMBERT[1], MICHEL LEDOUX AND CHRISTIAN WEBB[2]

Universität Zürich, Université de Toulouse and Aalto University

We present a new approach, inspired by Stein's method, to prove a central limit theorem (CLT) for linear statistics of β-ensembles in the one-cut regime. Compared with the previous proofs, our result requires less regularity on the potential and provides a rate of convergence in the quadratic Kantorovich or Wasserstein-2 distance. The rate depends both on the regularity of the potential and the test functions, and we prove that it is optimal in the case of the Gaussian Unitary Ensemble (GUE) for certain polynomial test functions.

The method relies on a general normal approximation result of independent interest which is valid for a large class of Gibbs-type distributions. In the context of β-ensembles, this leads to a multi-dimensional CLT for a sequence of linear statistics which are approximate eigenfunctions of the infinitesimal generator of *Dyson Brownian motion* once the various error terms are controlled using the rigidity results of Bourgade, Erdős and Yau.

CONTENTS

2620 G. LAMBERT, M. LEDOUX AND C. WEBB

1. Introduction and main results.

In this article, we study linear statistics of β-ensembles. By a β-ensemble, we mean a probability measure on $\mathbb{R}^N$ of the form

$$(1.1) \qquad \mathbb{P}_{V,\beta}^N(d\lambda_1, \ldots, d\lambda_N) = \frac{1}{Z_{V,\beta}^N} e^{-\beta \mathcal{H}_V^N(\lambda_1, \ldots, \lambda_N)} \prod_{j=1}^N d\lambda_j,$$

where

$$\mathcal{H}_V^N(\lambda) = \sum_{1 \leq i < j \leq N} \log \frac{1}{|\lambda_i - \lambda_j|} + N \sum_{j=1}^N V(\lambda_j).$$

The parameter $\beta > 0$ is interpreted as the inverse temperature, $V : \mathbb{R} \to \mathbb{R}$ is a suitable function known as the confining potential, and $Z_{V,\beta}^N$ is a normalization constant, known as the partition function. By a linear statistic, we mean a random variable of the form $\sum_{j=1}^N f(\lambda_j)$, where the configuration $(\lambda_j)_{j=1}^N$ is drawn from $\mathbb{P}_{V,\beta}^N$ and $f : \mathbb{R} \to \mathbb{R}$ is a test function. Our goal is to study the asymptotic fluctuations of these linear statistics in the large N limit.

A general reference about the asymptotic behavior of β-ensembles is [29], Section 11. It is a well-known fact that when N is large, a linear statistic is close to its mean or equilibrium value which is given by $N \int f(x) \mu_V(dx)$ where μ_V, the so-called equilibrium measure associated to V, is a probability measure with compact support on $\mathbb{R}$ and can be characterized as being the unique solution of a certain optimization problem that we recall below; see the discussion around (1.7). Moreover, under suitable conditions on V and f, it is known that the fluctuations of $\sum_{j=1}^N f(\lambda_j)$ around the equilibrium value are of order 1 and described by a certain Gaussian random variable. This CLT first appeared in the fundamental work of Johansson [16], Theorem 2.4, valid for suitable polynomial potentials V. The argument is inspired from statistical physics and based on the asymptotic expansion of the partition function $Z_{V,\beta}^N$ and relies on the analysis of the first *loop equation* by the methods of the perturbation theory. The method was further developed and simplified in the subsequent papers [17], Theorem 1, [5], Proposition 5.2 and [31], Theorem 1, but still under the assumption that the potential V is real-analytic in a neighborhood of the support of the equilibrium measure. We also point out the

article [4] which appeared nearly simultaneously with this article, and where the assumption of real-analyticity is relaxed by using a priori estimates from [19]. In particular, they obtain a CLT which is valid when the potential $V \in \mathscr{C}^5(\mathbb{R})$ for test functions $f \in \mathscr{C}_c^3(\mathbb{R})$, these assumptions being slightly weaker than Assumption 1.1, which covers the assumptions under which we prove our main result, Theorem 1.2. Their results are also stronger in the sense that they obtain the CLT by proving convergence of the Laplace transform of the law of a linear statistic and they also provide sufficient conditions to obtain Gaussian fluctuations in the multicut setting even when the potential is not regular; that is beyond the context of our Assumption 1.1. We shall come back to their method and compare it with ours in Section 1.2. In the special case $\beta = 2$, there are also other types of proofs which rely on the determinantal structure of the ensembles $\mathbb{P}_{V,2}^N$ and are valid in greater generality; see, for example, [8] or [18].

Our goal is to offer a new proof for this CLT with an argument inspired by Stein's method; see Theorem 1.2. In addition to novelty, the benefits of our approach are a rate of convergence (in the Kantorovich or Wasserstein distance W_2, see the discussion around (1.6) for a definition) as well as requiring less regularity of the potential V than the more classical proofs. We also demonstrate that at least in some special cases (e.g., the GUE with certain polynomial linear statistics), the rate of convergence is optimal. Compared to the references above, the drawback of our approach is that we are not able to provide estimates for exponential moments of linear statistics, or in other words, strong asymptotic estimates for the partition function of a β-ensemble, although we do get convergence of the first and second moment of a linear statistic of a nice enough function. This being said, it does not seem to be obvious how one would obtain our rate of convergence in the Kantorovich distance from, for example the Laplace transform estimates from [4].

In this section, we will first introduce some key concepts and then state our main results concerning the precise asymptotic behavior of $\sum_{j=1}^N f(\lambda_j)$. After this, we offer a brief comparison of our approach to other approaches to the CLT for linear statistics of β ensembles, and then conclude this Introduction with an outline of the rest of the article. Before proceeding, we describe some notation that we will use repeatedly throughout the article.

Notational conventions and basic concepts. For two sequences (A_n) and (B_n), we will write $A_n \ll B_n$ if there exists a constant $C > 0$ independent of n such that $A_n \leq C B_n$ for all n. When this constant depends on some further parameter, say ϵ, we write $A_n \ll_\epsilon B_n$. If $A_n \ll B_n$ and $B_n \ll A_n$, we write $A_n \asymp B_n$ (and $\asymp_\epsilon$ if the constants depend on the parameter ϵ). We will also use analogous notation for functions, for example, $f \ll g$ if there exists a constant C independent of x such that $f(x) \leq C g(x)$ for all x. In most of our estimates, these constants will depend on the parameters β and V in the definition of $\mathbb{P}_{V,\beta}^N$, but as these parameters remain fixed throughout the article, we shall not emphasize these dependencies and simply write $\ll$ instead of $\ll_{V,\beta}$, etc.

We write $\mathbb{N} = \{0, 1, \ldots\}$ and for $k \in \mathbb{N}$ as well as for any open interval $\mathrm{I} \subseteq \mathbb{R}$, we write $\mathscr{C}^k(\mathrm{I})$ for the space of functions which are k-times continuously differentiable on I. If I is a finite closed interval, we also write $\mathscr{C}^k(\mathrm{I})$ to indicate that the kth-derivative of the functions have finite limits at the end points of I. For any $k \in \mathbb{N}$, $\alpha \in (0, 1)$, and interval $\mathrm{I} \subset \mathbb{R}$, we write $\mathscr{C}^{k,\alpha}(\mathrm{I})$ for the space of functions $f \in \mathscr{C}^k(\mathrm{I})$ which satisfy

$$(1.2) \qquad \|f\|_{\mathscr{C}^{k,\alpha}(\mathrm{I})} = \sup_{x,y \in \mathrm{I}, x \neq y} \frac{|f^{(k)}(x) - f^{(k)}(y)|}{|x - y|^\alpha} < \infty.$$

We also use the notation $\mathscr{C}_c^k(\mathrm{I})$ for the subspace of $\mathscr{C}^k(\mathrm{I})$ whose elements have compact support in I. For a function $f : \mathbb{R} \to \mathbb{R}$, we also find it convenient to write

$$(1.3) \qquad \|f\|_{\infty,\mathrm{I}} = \sup_{x \in \mathrm{I}} |f(x)|.$$

We will impose several constraints on the potential V (see Assumption 1.1)—one being that it is normalized so that the support of the equilibrium measure μ_V is $[-1, 1]$ (the actual constraint here being that the support is a single interval—normalizing this interval to be $[-1, 1]$ can always be achieved by scaling and translating space). We thus find it convenient to introduce the following notation: let $\mathrm{J} = (-1, 1)$ and $\mathrm{J}_\epsilon = (-1 - \epsilon, 1 + \epsilon)$ for any $\epsilon > 0$. We also introduce notation for the semicircle law and arcsine law on J:

$$(1.4) \qquad \mu_{\mathrm{sc}}(dx) = \frac{2}{\pi}\sqrt{1 - x^2}\mathbf{1}_{\mathrm{J}}(x)\,dx \quad \text{and} \quad \varrho(dx) = \frac{\mathbf{1}_{\mathrm{J}}(x)}{\pi\sqrt{1 - x^2}}\,dx.$$

If $\lambda = (\lambda_1, \ldots, \lambda_N)$ is a configuration distributed according to (1.1), we define

$$(1.5) \qquad \nu_N(dx) = \sum_{j=1}^N \delta_{\lambda_j}(dx) - N\mu_V(dx).$$

This measure corresponds to the centered empirical measure when $\beta = 2$ and, in general, it describes the fluctuation of the configuration λ around equilibrium—we suppress the dependence on V here.

A further notion that is instrumental in the statement of our main results is the quadratic Kantorovich or Wasserstein-2 distance between two probability distributions. For μ and ν probability distributions on $\mathbb{R}^d$ with finite second moment (so $\int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty$ and similarly for ν), we write

$$(1.6) \qquad \mathrm{W}_2(\mu, \nu) = \inf_\pi \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) \right)^{\frac{1}{2}},$$

where the infimum is over all probability measures π on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν, that is, over all couplings π so that $\pi(dx, \mathbb{R}) = \mu(dx)$ and $\pi(\mathbb{R}, dy) = \nu(dy)$. Here, $|\cdot|$ is the Euclidean distance on $\mathbb{R}^d$. A basic fact about the quadratic Kantorovich distance is that convergence with respect to it implies convergence in

distribution along with convergence of the second absolute moment. We refer to [36], Chapter 6, for further information about Kantorovich distances. If μ is the distribution of a random variable X and ν the distribution of a random variable Y, we will occasionally find it convenient to write $W_2(X, Y)$ or $W_2(X, \nu)$ instead of $W_2(\mu, \nu)$.

Finally, for any $d \geq 1$ we will denote by γ_d the standard Gaussian law on $\mathbb{R}^d$:

$$\gamma_d(dx) = \frac{1}{(2\pi)^{\frac{d}{2}}} \prod_{j=1}^{d} e^{-\frac{1}{2}x_j^2} \, dx_j.$$

1.1. *Main results.* In addition to smoothness, we will need some further conditions on V. These are slightly indirect as they are statements about the equilibrium measure associated to V. We recall that if V is say continuous and grows sufficiently fast at infinity, its equilibrium measure μ_V is defined as the unique minimizer of the energy functional

$$\mu \mapsto \frac{1}{2} \int_{\mathbb{R} \times \mathbb{R}} \log \frac{1}{|x - y|} \mu(dx)\mu(dy) + \int_{\mathbb{R}} V(x)\mu(dx)$$

over all probability measures on $\mathbb{R}$. It turns out that μ_V has compact support and is characterized by the following two conditions:

- There exists a constant ℓ_V, such that for μ_V-almost every $x \in \text{supp}(\mu_V)$,

(1.7) $$Q(x) := V(x) - \int_{\mathbb{R}} \log |x - y|\mu_V(dy) = \ell_V.$$

- $Q(x) \geq \ell_V$ for all $x \notin \text{supp}(\mu_V)$.

We refer to [1], Section 2.6, for further details.

It is well known that the assumption that the support of μ_V is a single interval is necessary to observe Gaussian fluctuations for all test functions. In the multicut regime, namely when the support consists of several intervals, there is a subspace of smooth functions where the fluctuations are Gaussian, but for a general smooth test function the fluctuations can oscillate with the dimension N and there is no CLT in general. This is due to the possibility of eigenvalues tunnelling from one component of the support to another. See, for example, [28] for a heuristic description of this phenomenon, or [31], Theorem 2, and [6], Section 8.3, for precise results as well as [4] for a description of when one observes Gaussian fluctuations. When the potential V is sufficiently smooth, the behavior of the equilibrium measure is well understood. In fact, by [29], Theorem 11.2.4, if one normalizes V so that $\text{supp}(\mu_V) = \bar{J} = [-1, 1]$ (one can check that this is equivalent to the conditions (11.2.17) in [29], Theorem 11.2.4), we have

(1.8) $$\mu_V(dx) = S(x)\mu_{\text{sc}}(dx),$$

where $S : \overline{J} \to [0, \infty]$ is given by

$$(1.9) \qquad S(x) = \frac{1}{2} \int_J \frac{V'(x) - V'(y)}{x - y} \varrho(dy).$$

Note that this formula can be used to define S outside of $\overline{J}$. Moreover, one can check (see, e.g., Lemma A.1 in the Appendix) that this expression shows that if $V \in \mathscr{C}^{\kappa+3}(\mathbb{R})$, then the extension given by (1.9) satisfies $S \in \mathscr{C}^{\kappa+1}(\mathbb{R})$ for any $\kappa \in \mathbb{N}$.

Finally, we assume that $S(x) > 0$ for all $x \in \overline{J}$. In the random matrix theory literature, this is known as the *off-criticality condition*. Combining these remarks, we now state our standing assumptions about V.

ASSUMPTION 1.1. Let $V \in \mathscr{C}^{\kappa+3}(\mathbb{R})$ be such that $\liminf_{x \to \pm\infty} \frac{V(x)}{\log |x|} > 1$ and $\inf_{x \in \mathbb{R}} V''(x) > -\infty$. Moreover, assume that the support of the measure μ_V is $\overline{J} = [-1, 1]$ and $S(x) > 0$ for all $x \in \overline{J}$.

There is a rather large class of potentials which satisfy these assumptions. For instance, any function V that is sufficiently convex on $\mathbb{R}$. A concrete example being $V(x) = x^2$, for which (1.1) is the so-called Gaussian or Hermite β-ensemble. In particular, it is a well-known fact that $\mathbb{P}^N_{x^2, 2}$ is the distribution of the eigenvalues of a $N \times N$ GUE random matrix.

As a final step before stating our main result, we introduce some notation concerning the limiting mean and variance of a linear statistic. Let $f \in \mathscr{C}^1_c(\mathbb{R})$ and define

$$(1.10) \qquad \begin{aligned} \mathbf{m}(f) &= \frac{1}{2}(f(1) + f(-1)) - \int_J f(x)\varrho(dx) \\ &\quad - \frac{1}{2} \int_J \frac{S'(x)}{S(x)} \left[\int_J \frac{f(x) - f(y)}{x - y} \varrho(dy) \right] \mu_{sc}(dx) \end{aligned}$$

and

$$(1.11) \qquad \mathbf{\Sigma}(f) = \frac{1}{4} \iint_{J \times J} \left(\frac{f(x) - f(y)}{x - y} \right)^2 (1 - xy) \varrho(dx) \varrho(dy),$$

where the measures μ_{sc} and ϱ are as in (1.4). As we see below, $N \int_J f \, d\mu_V + (\frac{1}{2} - \frac{1}{\beta}) \mathbf{m}(f)$ describes the mean of a linear statistic (a fact first noted in [31]) and $\frac{2}{\beta} \mathbf{\Sigma}(f)$ its variance. There are different possible expressions for the variance $\mathbf{\Sigma}$. For example, one can check that for $f \in \mathscr{C}^1_c(\mathbb{R})$,

$$(1.12) \qquad \mathbf{\Sigma}(f) = \frac{1}{4} \sum_{k=1}^{\infty} k f_k^2,$$

where f_k, $k \geq 1$, denote the Fourier–Chebyshev coefficients of the function f, see (4.9). We also refer to (4.15) for yet another formula for $\mathbf{\Sigma}(f)$. We wish to

emphasize here the fact (well known since [16]) that while μ_V and $\mathbf{m}(f)$ depend on V, the variance $\mathbf{\Sigma}(f)$ is independent of V, or in other words, the fluctuations of the linear statistics are universal.

We now turn to the statement of our main result.

THEOREM 1.2. *Let $V : \mathbb{R} \to \mathbb{R}$ satisfy Assumption* 1.1 *with $\kappa \geq 5$ and let $f \in \mathscr{C}_c^{\kappa+4}(\mathbb{R})$ be a given function which is nonconstant on* J. *Moreover, let $(\lambda_1, \ldots, \lambda_N)$ be distributed according to* (1.1) *and*

(1.13)
$$\mathcal{X}_N(f) := \sqrt{\frac{\beta}{2\mathbf{\Sigma}(f)}}\left(\int_{\mathbb{R}} f(x)\nu_N(dx) - \left(\frac{1}{2} - \frac{1}{\beta}\right)\mathbf{m}(f)\right)$$
$$= \sqrt{\frac{\beta}{2\mathbf{\Sigma}(f)}}\left(\sum_{j=1}^{N} f(\lambda_j) - N\int_{J} f(x)\mu_V(dx) - \left(\frac{1}{2} - \frac{1}{\beta}\right)\mathbf{m}(f)\right).$$

Then the W_2-distance between the law of the random variable $\mathcal{X}_N(f)$ and the standard Gaussian measure on $\mathbb{R}$ satisfies for any $\epsilon > 0$,
$$W_2(\mathcal{X}_N(f), \gamma_1) \ll_f N^{-\theta+\epsilon},$$
where $\theta = \min\{\frac{2\kappa-9}{2\kappa+11}, \frac{2}{3}\}$. In particular, as $N \to \infty$, $\sum_{j=1}^{N} f(\lambda_j) - N\int_J f \, d\mu_V$ converges in law to a Gaussian random variable with mean $(\frac{1}{2} - \frac{1}{\beta})\mathbf{m}(f)$ and variance $\frac{2}{\beta}\mathbf{\Sigma}(f)$.

REMARK 1.3. The function $\theta(\kappa)$ is nondecreasing in κ and positive if $\kappa \geq 5$. This means that the rate of convergence in Theorem 1.2 is improving with the regularity of the potential V as well as the regularity of the test function f. Moreover, we point out that the condition $\theta \leq \frac{2}{3}$ comes only from the fluctuations of the eigenvalues near the edges of the spectrum. For instance, we deduce from the proof of Theorem 1.2 (see Section 6.3) that if $f \in \mathscr{C}_c^\infty(J)$, then we could take $\theta = 1$ which, according to Theorem 1.4 below, leads to the optimal rate (at least for such test functions f) up to a factor of N^ϵ for an arbitrary $\epsilon > 0$.

One interpretation of Theorem 1.2 is that, if centered and normalized to have a β-independent covariance, the empirical measure $\sum_{j=1}^{N} \delta_{\lambda_j}$, viewed as a random generalized function, converges say in the sense of finite dimensional distributions to a mean-zero Gaussian process $\mathcal{X}$ which is an element of a suitable space of generalized functions and has the covariance structure

$$\mathbb{E}[\mathcal{X}(f)\mathcal{X}(g)] = \mathbf{\Sigma}(f, g) := \frac{1}{4}\sum_{k \geq 1} k f_k g_k.$$

In fact, one can understand $\mathcal{X}$ as being the (distributional) derivative of the random generalized function

$$\mathcal{Y}(x) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{k+1}} X_k U_k(x) \frac{2}{\pi}\sqrt{1-x^2},$$

where X_k are i.i.d. standard Gaussians, $x \in \mathrm{J} = (-1, 1)$, U_k are Chebyshev polynomials of the second kind (see the Appendix for their definition) and the interpretation is that one should consider this as a random element of a suitable space of generalized functions. With some work (we omit the details), one can then check that the covariance kernel of the process $\mathcal{Y}$ is given by

$$\mathbb{E}[\mathcal{Y}(x)\mathcal{Y}(y)] = -\frac{4}{\pi^2} \log\left(\frac{|x-y|}{1-xy+\sqrt{1-x^2}\sqrt{1-y^2}}\right), \qquad x, y \in \mathrm{J}.$$

We also point out that one can check formally (say using Lemma A.3) that the Hilbert transform of the process $\mathcal{Y}$ is a centered Gaussian process whose covariance kernel is proportional to $-\log 2|x - y|$ on $\mathrm{J} \times \mathrm{J}$—which is the covariance kernel of the $2d$-Gaussian free field restricted to J.

Returning to the statement of Theorem 1.2, we note that the assumption of f having compact support is not essential. In view of the large deviation principle for the extreme values of the configuration $(\lambda_j)_{j=1}^N$ (see, e.g., [5], Proposition 2.1, or [1], Section 2.6.2), one could easily replace compact support by polynomial growth rate at infinity and the assumption of f being smooth on $\mathbb{R}$ by f being smooth on $(-1 - \epsilon, 1 + \epsilon)$ for some $\epsilon > 0$ and some very mild regularity assumptions outside of this interval. We choose not to focus on these classical generalizations here and be satisfied with the assumption that f is smooth and compactly supported.

We also wish to point out a variant of Theorem 1.2 for polynomial linear statistics of the GUE to emphasize that our approach can yield, at least in some cases, an optimal rate for normal approximation in the W_2-metric.

THEOREM 1.4. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a nonconstant polynomial* (*independent of* N) *and* ν_N *be as in* (1.5) *where* $(\lambda_1, \ldots, \lambda_N)$ *is distributed according to the ensemble* $\mathbb{P}_{x^2,2}^N$. *Then, as* $N \to \infty$,

$$\mathrm{W}_2\left(\frac{1}{\sqrt{\mathbf{\Sigma}(f)}} \int_{\mathbb{R}} f(x)\nu_N(dx), \gamma_1\right) \ll_f \frac{1}{N}.$$

Moreover, if $f(x) = x^2$, *the result is sharp*:

$$\mathrm{W}_2\left(\frac{1}{\sqrt{\mathbf{\Sigma}(f)}} \int_{\mathbb{R}} x^2 \nu_N(dx), \gamma_1\right) \asymp \frac{1}{N}.$$

Our method of proof of Theorem 1.2 (and Theorem 1.4) is to first use a Stein's method type argument to prove a general normal approximation result of independent interest—Proposition 2.1—giving a bound on the W_2-distance between the law of an essentially arbitrary function of some collection of random variables and a standard Gaussian distribution. As this bound is true in such generality, it is naturally not a very useful one in most cases. What is of utmost importance to our approach is that if this function is chosen to be *an approximate eigenfunction* of a certain differential operator associated to the distribution of the random

variables (e.g., in the setting of the Gaussian β-ensembles, this operator is the infinitesimal generator of Dyson–Ornstein–Uhlenbeck process), then this bound on the W_2-distance becomes quite precise. We then proceed to prove that in the setting of the β-ensembles, there are many such approximate eigenfunctions—in fact, enough that linear statistics of nice enough functions can be approximated in terms of linear combinations of these approximate eigenfunctions and a multidimensional CLT for the approximate eigenfunctions implies a CLT for linear statistics of smooth enough functions.

Related reasoning has appeared elsewhere in the literature as well. For example, in [20], a normal approximation result was found for functions that are exact eigenfunctions of the relevant differential operator. Also, we wish to point out that our normal approximation result can be viewed as a W_2-version of the multivariate normal approximation in the Kantorovich metric W_1 developed by Meckes in [24] and relying on exchangeable pairs. For the application to linear statistics associated to classical compact groups and circular β-ensembles, studied in [11–13, 37], the pairs are created through circular Dyson–Brownian motion. Moreover, while this is not explicitly emphasized in these works, the CLT is actually proven for certain approximate eigenfunctions of the infinitesimal generator of the circular Dyson–Brownian motion.

These functions and the multidimensional CLT for the corresponding linear statistics are of such central importance to our proof of Theorem 1.2 that despite their properties being slightly technical we wish to formulate general results about them in this introduction.

THEOREM 1.5. *Let V be as in Assumption 1.1 and $\eta > \frac{4(\kappa+1)}{2\kappa-1}$. For small enough $\delta > 0$ depending only on V, there exists a sequence of functions $(\phi_n)_{n=1}^{\infty}$ such that for $\epsilon_n = \delta n^{-\eta}$, $\phi_n \in \mathscr{C}_c^{\kappa}(J_{2\epsilon_n})$ for all $n \geq 1$, and they satisfy*

$$(1.14) \qquad V'(x)\phi_n'(x) - \int_J \frac{\phi_n'(x) - \phi_n'(y)}{x - y}\mu_V(dy) = \sigma_n \phi_n(x)$$

on J_{ϵ_n}, where $\sigma_n = \frac{1}{2\Sigma(\phi_n)} > 0$. Moreover, we have the estimate $\sigma_n \asymp n$ and the bounds for any $0 \leq k < \kappa$,

$$(1.15) \qquad \qquad \|\phi_n^{(k)}\|_{\infty,\mathbb{R}} \ll n^{k\eta}.$$

Finally, restricted to J, the functions ϕ_n form an orthonormal basis of $\mathscr{H}_{\mu_V}$ (see the discussion around (4.1) for its definition) and for any f in $\mathscr{C}^{\kappa+4}(\mathbb{R})$, we can expand $f(x) = \widehat{f_0} + \sum_{n=1}^{\infty} \widehat{f_n}\phi_n(x)$ for all $x \in \bar{J}$, where the Fourier-ϕ coefficients, given by $\widehat{f_0} = \int_J f\, d\varrho$ and $\widehat{f_n} = \langle f, \phi_n \rangle_{\mu_V}$ (see (4.2) for a definition), satisfy $|\widehat{f_n}| \ll_f n^{-\frac{\kappa+3}{2}}$.

The proof of this theorem begins by noting that equation (1.14) can be inverted on J yielding an eigenvalue equation for a suitable operator which turns out to be

self-adjoint, compact, and positive on a suitable weighted Sobolev space that we call $\mathscr{H}_{\mu_V}$. This provides the existence of the functions ϕ_n on J, which are then continued outside of J by (1.14), which becomes an ODE outside of J.

The fact that the functions ϕ_n satisfy (1.14) turns out to imply that functions of the form $\sum_{j=1}^N \phi_n(\lambda_j)$ are approximate eigenfunctions of the operator

$$\mathrm{L} = \mathrm{L}_{V,\beta}^N = \sum_{j=1}^N \frac{\partial^2}{\partial \lambda_j^2} - \beta N \sum_{j=1}^N V'(\lambda_j) \frac{\partial}{\partial \lambda_j} + \beta \sum_{i \neq j} \frac{1}{\lambda_j - \lambda_i} \frac{\partial}{\partial \lambda_j}$$

in the sense that the random variable

$$(1.16) \qquad\qquad \mathrm{L}\big(\mathcal{X}_N(\phi_n)\big) + \beta N \sigma_n \mathcal{X}_N(\phi_n)$$

turns out to be small compared to the eigenvalue $\beta N \sigma_n$. Here and elsewhere in this article, we use a slight abuse of notation and write, for example, $\mathrm{L}(\sum_{i=1}^N \phi(\lambda_i))$ for the random variable which is obtained by first calculating $\mathrm{L}(\sum_{j=1}^N \phi(\lambda_i))$ with deterministic λ, and then evaluating this function at a random λ drawn from (1.1). This approximate eigenfunction property is precisely what allows us making use of the normal approximation result Proposition 2.1. This leads to the following multidimensional CLT.

PROPOSITION 1.6.   *Using the notation of Theorem* 1.5 *and* (1.13), *we have, for any $\epsilon > 0$,*

$$\mathrm{W}_2\big((\mathcal{X}_N(\phi_n))_{n=1}^d, \gamma_d\big) \ll d^{\frac{16(\kappa+1)}{2\kappa-1}} N^{-1+\epsilon}.$$

Our proof of Theorem 1.2 then proceeds by approximating a general function f by a linear combination of the ϕ_n's in $[-1, 1]$, arguing that essentially what happens outside of this interval is irrelevant, and then using the CLT for the ϕ_n. Actually, the proof of Theorem 1.4 is quite a bit simpler due to the fact that the functions ϕ_n are explicit in this case: they are just suitably normalized Chebyshev polynomials of the first kind and much of the effort going into the proof of Theorem 1.5 is not needed.

Besides the existence of these eigenfunctions, another fundamental property that is required to control the error term coming from Proposition 2.1 is a kind of rigidity of the random configuration $(\lambda_j)_{j=1}^N$. In particular, in the proof of Proposition 1.6 and Theorem 1.2, we will make use of the following strong rigidity result of Bourgade, Erdős and Yau.

THEOREM 1.7 ([7], Theorem 2.4).   *Let V be as in Assumption* 1.1, $\rho_0 = -1$, *and for any $j \in \{1, \ldots, N\}$, define the classical locations $\rho_j \in [-1, 1]$ by*

$$(1.17) \qquad\qquad \int_{\rho_0}^{\rho_j} \mu_V(dx) = \frac{j}{N}.$$

Moreover, write $\widehat{j} = \min(j, N - j + 1)$, *assume that the eigenvalues are ordered* $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$, *and consider the event*

$$(1.18) \qquad \mathcal{B}_\epsilon = \big\{ \forall j \in \{1, \ldots, N\} : |\lambda_j - \rho_j| \leq \widehat{j}^{-\frac{1}{3}} N^{-\frac{2}{3}+\epsilon} \big\}.$$

Then, for any $\epsilon > 0$ *there exist* $c_\epsilon, N_\epsilon > 0$ *such that for all* $N \geq N_\epsilon$,

$$\mathbb{P}^N_{V,\beta}\big(\mathbb{R}^N \setminus \mathcal{B}_\epsilon\big) \leq e^{-N^{c_\epsilon}}.$$

1.2. *Connection with the literature.* First, we point out that the operator L also plays an important role in many other approaches to the CLT for linear statistics as the first loop equation can be written as $\mathbb{E}^N_{V,\beta}[\mathrm{L}(\mathcal{X}_N(f))] = 0$. In particular, if (1.16) is asymptotically small compared to $\beta N \sigma_n$ in a strong enough sense, the first loop equation implies that $\mathbb{E}^N_{V,\beta}[\mathcal{X}_N(\phi_n)] \to 0$ as $N \to \infty$ so that the mean of the linear statistic $\sum_{j=1}^{N} f(\lambda_j)$ behaves like $N \int_\mathbb{R} f(x)\mu_V(dx) + (\frac{1}{2} - \frac{1}{\beta})\mathbf{m}(f)$ for large N. Moreover, the eigenequation (1.14) also plays a role in the analysis of the loop equations as well as for the transport map approach that we briefly present below. The starting point of the methods of the previous works [5, 6, 16, 31, 32] and [4] consists of expressing the Laplace transform of the law of a linear statistic as

$$(1.19) \qquad \mathbb{E}^N_{V,\beta}\big[e^{s \sum_{j=1}^{N} f(\lambda_j)}\big] = \frac{Z^N_{V_t,\beta}}{Z^N_{V_0,\beta}}, \qquad t = -\frac{s}{\beta N},$$

where one defines the deformed potential $V_t(x) = V(x) + tf(x)$ for any $t \in \mathbb{R}$. In order to obtain the asymptotics this ratio of partition functions, the idea from transport theory introduced in [3, 32] (to establish local universality) and used in [4] (to obtain a CLT) consists of making a change of variables $\lambda_j \leftarrow \vartheta_t(\lambda_j)$ for all $j = 1, \ldots, N$ in the integral

$$Z^N_{V_t,\beta} = \int_{\mathbb{R}^N} e^{-\beta \mathcal{H}^N_{V_t}(\lambda_1,\ldots,\lambda_N)} \prod_{j=1}^{N} d\lambda_j$$

$$= \int_{\mathbb{R}^N} e^{-\beta \mathcal{H}^N_{V_t}(\vartheta_t(\lambda_1),\ldots,\vartheta_t(\lambda_N)) + \sum_{j=1}^{N} \log \vartheta'_t(\lambda_j)} \prod_{j=1}^{N} d\lambda_j.$$

This implies that

$$(1.20) \qquad \frac{Z^N_{V_t,\beta}}{Z^N_{V_0,\beta}} = \mathbb{E}^N_{V,\beta}\big[e^{-\beta(\mathcal{H}^N_{V_t}(\vartheta_t(\lambda_1),\ldots,\vartheta_t(\lambda_N)) - \mathcal{H}^N_{V_0}(\lambda_1,\ldots,\lambda_N)) + \sum_{j=1}^{N} \log \vartheta'_t(\lambda_j)}\big].$$

Then it turns out that to obtain the CLT, it suffices to consider a simple diffeormorphism of the form $\vartheta_t(x) = x + t\psi(x)$ where ψ is a solution of the equation

$$(1.21) \qquad \square^V_x(\psi) = -V'(x)\psi(x) + \int_J \frac{\psi(x) - \psi(y)}{x - y} \mu_V(dy) = f(x) + c_f$$

for some suitably chosen constant $c_f \in \mathbb{R}$. This equation is important because if we expand the exponent on the RHS of (1.20) up to order t^2, we can check that

$$-\beta\big[\mathcal{H}_{V_t}^N(\vartheta_t(\lambda_1), \ldots, \vartheta_t(\lambda_N)) - \mathcal{H}_{V_0}^N(\lambda_1, \ldots, \lambda_N)\big] + \sum_{j=1}^{N} \log \vartheta_t'(\lambda_j)$$

$$= -\beta t N^2 \int f(x)\mu_V(dx) - \beta t N \int (f(x) - \square_x^V(\psi))\nu_N(dx)$$

$$+ \beta N^2 t^2 \Sigma(f) + \left(\frac{\beta}{2} - 1\right) t N \mathbf{m}(f) + \epsilon_{N,t} + O_{N\to\infty}(Nt^2),$$

where the error term is deterministic, $\epsilon_{N,t}$ is a small random quantity and we have

$$(1.22) \quad \mathbf{m}(f) = -\int \psi'(x)\mu_V(dx) \quad \text{and} \quad \Sigma(f) = -\frac{1}{2}\int f'(x)\psi(x)\mu_V(dx).$$

These formulae are those of the asymptotic mean and variance of the random variable $\int f(x)\nu_N$ given in [4], Theorem 1. Therefore, since $\beta Nt = -s$, combining these asymptotics with (1.19) and (1.20), we obtain

$$(1.23) \quad \mathbb{E}_{V,\beta}^N\big[e^{s\int f(x)\nu_N(dx)}\big] = e^{s(\frac{1}{2} - \frac{1}{\beta})\mathbf{m}(f) + \frac{s^2}{\beta}\Sigma(f) + O_{N\to\infty}(N^{-1})}\mathbb{E}_{V,\beta}^N\big[e^{\epsilon_{N,t}}\big].$$

To complete the proof of the CLT, most of the technical challenges consist of showing that the Laplace transform on the RHS of (1.23) converges to 1 as $N \to \infty$—in particular that the so-called *anisotropy term* contained in $\epsilon_{N,t}$ is small. This step is performed by using the regularity of the function ψ and some a priori estimates on the partition function $Z_{V_t,\beta}^N$ from [19] that we will not detail here. The bottom line is that the CLT holds for all test functions f for which equation (1.21) has a sufficiently smooth solution. This leads to sufficient conditions on the test function f (see (1.14) and (1.15) in [4]), valid in the multicut and certain critical situations under which the asymptotics

$$\mathbb{E}_{V,\beta}^N\big[e^{s\int f(x)\nu_N(dx)}\big] = e^{s(\frac{1}{2} - \frac{1}{\beta})\mathbf{m}(f) + \frac{s^2}{\beta}\Sigma(f) + o_{N\to\infty}(1)}$$

hold. Note that the operator $\square^V$ is related to the operator (4.5) in the following way:

$$\Xi^{\mu_V}(f) = -\square^V(f') \qquad \text{on J.}$$

Thus, in order to solve the eigenequation (4.6) which is fundamental to our proof of Theorem 1.5, it suffices to prove that the operator $\mathcal{R}$ given by

$$\mathcal{R}(\phi)' = (-\square^V)^{-1}(\phi)$$

is compact when acting on a suitable space $\mathscr{F}$ of functions $\phi : J \to \mathbb{R}$. Then taking into account the conditions [4] (1.14) and (1.15), in the definition of $\mathscr{F}$, one should be able—by adapting the arguments of Section 4—to generalize Theorem 1.5 to

the multicut and critical situations treated in [4]. Hence, it should be possible to generalize Theorem 1.2 and to obtain a rate of convergence in the Kantorovich distance in these cases as well (with a rate of convergence depending on the regularity of the potential V and the equilibrium measure μ_V). Finally, let us comment that if $f \in \mathscr{H}$ (see (4.1)) we may use the eigenbasis $(\phi_n)_{n=1}^{\infty}$ of Theorem 1.5 to solve equation (1.21). Namely, according to equation (1.14), we have

$$\square^V(\phi_n') = -\sigma_n \phi_n$$

and, if we expand $f = \widehat{f_0} + \sum_{n=1}^{\infty} \widehat{f_n} \phi_n$, then the function $\psi = -\sum_{n=1}^{\infty} \frac{\widehat{f_n}}{\sigma_n} \phi_n'$ solves (1.21) with $c_f = -\widehat{f_0}$ (note that $\psi \in L^2(\mu_V)$). Then we deduce from the formulae (1.22) that

$$\mathbf{m}(f) = \int \psi'(x)\mu_V(dx) = \sum_{n=1}^{\infty} \frac{\widehat{f_n}}{\sigma_n} \int \phi_n''(x)\mu_V(dx),$$

and since the functions $(\phi_n')_{n=1}^{\infty}$ are orthonormal with respect to $L^2(\mu_V)$,

$$\mathbf{\Sigma}(f) = -\frac{1}{2} \int f'(x)\psi(x)\mu_V(dx) = \sum_{n=1}^{\infty} \frac{\widehat{f_n}^2}{2\sigma_n}.$$

Using (5.5) and Lemma 6.1 below, we conclude that, if f is sufficiently smooth, the formulae (1.22) from [4], Theorem 1, are consistent with (1.10) and (1.11).

1.3. *Outline of the article.* We now describe the structure of the rest of the article. In Section 2, we prove Proposition 2.1—our general normal approximation result. After this, in Section 3 we apply it to the GUE in order to prove Theorem 1.4. We next move on to Section 4, where we prove the existence and basic regularity properties of the functions ϕ_n, namely Theorem 1.5. Armed with this information about the functions ϕ_n, we prove Proposition 1.6 in Section 5 and then in Section 6 we apply Proposition 1.6 to prove Theorem 1.2. Finally in the Appendix, we recall some basic properties of Chebyshev polynomials and the Hilbert transform which will play a critical role in our analysis.

2. A general normal approximation result. In this section, we describe and prove some general normal approximation results for functions of random variables drawn from probability distributions of the form $\mu(dx) = \frac{1}{Z} e^{-\mathcal{H}(x)} dx$ on $\mathbb{R}^N$, where $\mathcal{H} : \mathbb{R}^N \to \mathbb{R}$ is a nice enough function and Z a normalization constant. These approximation results will of course be useful (in the sense that they say that this function of random variables is close to a Gaussian random variable in the $N \to \infty$ limit) only for very special functions of these random variables. Nevertheless, we will be able to apply these approximation results to a suitably large class of linear statistics of measures of this form with the choice of $\mathcal{H} = \beta \mathcal{H}_V^N$ from (1.1). While our main interest is linear statistics of β-ensembles, we choose

to keep the discussion on a more general level as such normal approximation results might be of use in other settings as well. Keeping this in mind, we offer some further discussion about the approximation results, and even touching on some facts and approaches that may not be of use in our application to linear statistics of β-ensembles.

Before going into the statement and proof of the approximation results, we introduce some notation. An object that is of critical importance to our approach is a second-order differential operator which is symmetric with respect to the inner product of $L^2(\mu)$. More precisely, we write for smooth $f : \mathbb{R}^N \to \mathbb{R}$,

$$(2.1) \qquad \mathrm{L}f = \Delta f - \nabla \mathcal{H} \cdot \nabla f = \sum_{i=1}^N \partial_{ii} f - \sum_{i=1}^N \partial_i \mathcal{H} \partial_i f.$$

We point out for later reference that for β-ensembles, one has

$$(2.2) \qquad \mathrm{L} = \mathrm{L}_{V,\beta}^N = \sum_{j=1}^N \partial_{\lambda_j}^2 - N\beta \sum_{j=1}^N V'(\lambda_j) \partial_{\lambda_j} + \beta \sum_{i \neq j} \frac{1}{\lambda_j - \lambda_i} \partial_{\lambda_j}$$

which is just the infinitesimal generator of the diffusion with invariant measure $\mathbb{P}_{V,\beta}^N$ which we refer as *Dyson–Brownian motion*.

As mentioned above, we will make use of the fact that L is symmetric with respect to the inner product of $L^2(\mu)$. More precisely, if we assume sufficient regularity of $\mathcal{H}$ (say that it is smooth) then integrating by parts shows that for say smooth $f, g : \mathbb{R}^N \to \mathbb{R}$ with nice enough behavior at infinity,

$$(2.3) \qquad \int_{\mathbb{R}^N} f(-\mathrm{L}g) \, d\mu = \int_{\mathbb{R}^N} \nabla f \cdot \nabla g \, d\mu =: \int_{\mathbb{R}^N} \Gamma(f,g) \, d\mu,$$

where we thus write $\Gamma(f, g)$ for $\nabla f \cdot \nabla g$. To ease notation, we also write $\Gamma(f) := \Gamma(f, f) = |\nabla f|^2$. Note that in the setting of β-ensembles, $\mathcal{H}$ is not smooth, and when considering the integral $\int_{\mathbb{R}^N} f(-\mathrm{L})g \, d\mu$, one encounters singularities of the form $\frac{1}{\lambda_j - \lambda_k} \partial_j g(\lambda) \prod_{i<j} |\lambda_i - \lambda_j|^\beta$. As $\beta > 0$, this is an integrable singularity, so integration by parts is justified and (2.3) is still true. Our argument will implicitly impose several regularity assumptions on $\mathcal{H}$ and we will not be explicit about what precisely one should assume about $\mathcal{H}$. Nevertheless, when there are issues of this type, we will point out why there are no problems in the setting of linear statistics of β-ensembles.

Below in Section 2.1, we state our general normal approximation results and offer some further discussion about them in some particular cases. Then in Section 2.2, we prove the approximation results.

2.1. *Statement and discussion of the normal approximation results.* To simplify the statement of our approximation results, we fix some further notation. Let $X = (X_1, \ldots, X_N)$ be a random vector in $\mathbb{R}^N$ with distribution μ and let

$F = (F_1, \ldots, F_d) : \mathbb{R}^N \to \mathbb{R}^d$ be a smooth function. What we mean by a normal approximation result is that we wish to estimate the Kantorovich distance W_2 (and in case $d = 1$ also in the total variation distance) between the law of $F(X)$ and that of the standard Gaussian measure γ_d on $\mathbb{R}^d$. Bounds on these distances will be stated in terms of the function F and its differentials; more precisely, in the notation introduced above, the bounds will be expressed in terms of (the vector-valued versions of) LF and $\Gamma(F)$.

For positive numbers $\kappa_k > 0$, $k = 1, \ldots, d$, denote by K the diagonal matrix $K = \operatorname{diag}(\kappa_1, \ldots, \kappa_d)$. Given $F = (F_1, \ldots, F_d)$, let $\Gamma(F) = (\Gamma(F_k, F_\ell))_{1 \le k, \ell \le d}$ and introduce the quantities A and B by

$$(2.4) \qquad A = \left(\int_{\mathbb{R}^N} |F + K^{-1} LF|^2 \, d\mu \right)^{\frac{1}{2}}$$

and

$$(2.5) \qquad B = \left(\int_{\mathbb{R}^N} |\operatorname{Id} - K^{-1} \Gamma(F)|^2 \, d\mu \right)^{\frac{1}{2}},$$

the norms $|\cdot|$ being understood in the Euclidean space $\mathbb{R}^d$ and in the space of $d \times d$ matrices (Hilbert–Schmidt norm). The expressions A and B thus depend on F and on $\kappa_k > 0$, $k = 1, \ldots, d$. In this notation, our first approximation result reads as follows.

PROPOSITION 2.1. *Let $F : \mathbb{R}^N \to \mathbb{R}^d$ be of class $\mathscr{C}^2$ and in $L^2(\mu)$, and denote by $\mu \circ F^{-1}$ the law of F under μ (that is the law of $F(X)$ on $\mathbb{R}^d$). Then, for any choice of $\kappa_k > 0$, $k = 1, \ldots, d$,*

$$W_2(\mu \circ F^{-1}, \gamma_d) \le A + B,$$

where again γ_d is the law of a standard d-dimensional Gaussian.

We postpone the proof of Proposition 2.1 (as well as the forthcoming Proposition 2.2) to Section 2.2.

Before discussing further normal approximation results, we point out that for Proposition 2.1 to be of any use, one of course will want $F + K^{-1} LF$ and $\operatorname{Id} - K^{-1} \Gamma(F)$ to be in $L^2(\mu)$ and small in some sense. This typically will not be true for arbitrary F and K, but only for very special choices of F and K. We return to the choice of F and K later on.

We next mention that Proposition 2.1 is already of interest in dimension one (i.e., when $d = 1$) in which case we also get a bound in the total variation distance

$$\|\nu - \nu'\|_{\mathrm{TV}} = \sup_{E \in \mathcal{B}(\mathbb{R})} [\nu(E) - \nu'(E)] = \frac{1}{2} \sup \left[\int_{\mathbb{R}} \varphi \, d\nu - \int_{\mathbb{R}} \varphi \, d\nu' \right],$$

where the supremum is taken over all bounded measurable $\varphi : \mathbb{R} \to \mathbb{R}$ with $\|\varphi\|_\infty \le 1$. The normal approximation result is the following.

PROPOSITION 2.2.   *Let $F : \mathbb{R}^N \to \mathbb{R}$ be of class $\mathscr{C}^2$ and in $L^2(\mu)$, and denote by $\mu \circ F^{-1}$ the law of F under μ (that is the law of $F(X)$ on $\mathbb{R}$). Then, for any $\kappa > 0$, $\|\mu \circ F^{-1} - \gamma_1\|_{\mathrm{TV}} \le 2A + 2B$, that is,*

$$
(2.6) \qquad
\begin{aligned}
\|\mu \circ F^{-1} - \gamma_1\|_{\mathrm{TV}} &\le 2\left(\int_{\mathbb{R}^N} \left[F + \frac{1}{\kappa} \mathrm{L}F\right]^2 d\mu\right)^{\frac{1}{2}} \\
&\quad + 2\left(\int_{\mathbb{R}^N} \left[1 - \frac{1}{\kappa}\Gamma(F)\right]^2 d\mu\right)^{\frac{1}{2}}.
\end{aligned}
$$

Moreover, if $F = (F_1, \ldots, F_d)$ and $G = \sum_{k=1}^{d} \theta_k F_k$ where $\sum_{k=1}^{d} \theta_k^2 = 1$, then

$$
(2.7) \qquad \|\mu \circ G^{-1} - \gamma_1\|_{\mathrm{TV}} \le 2A + 2B
$$

and

$$
(2.8) \qquad \mathrm{W}_2(\mu \circ G^{-1}, \gamma_1) \le A + B.
$$

REMARK 2.3.   To underline the difference between Proposition 2.1 and Proposition 2.2, we mention here that the proof of Proposition 2.2 is a rather classical one-dimensional Stein's method argument, relying on Stein's lemma. In this setting, the natural metric in which one obtains approximation results is the total variation metric. The situation in Proposition 2.1 is slightly different. While there are generalizations of Stein's method to multivariate normal approximation (see, e.g., [24]), these typically yield approximation results in the metric W_1. Using these results, one could indeed prove a (weaker) version of Theorem 1.2. While philosophically very similar to classical Stein's method arguments, our proof of Proposition 2.1 relies instead on semigroup interpolation techniques, partly following [21], which allow upgrading convergence in the W_1-metric to the W_2-metric.

To widen the spectrum of Propositions 2.1 and 2.2, it is sometimes convenient to deal with a random vector X given as an image $X = U(Y)$ of another random vector Y on $\mathbb{R}^m$ (typically Gaussian) where $U : \mathbb{R}^m \to \mathbb{R}^N$. Depending on specific properties of the derivatives of U, Proposition 2.1 may be used to also control the distance between the law of $F(X)$ and γ_d. This follows from the description of A and B for the new map $G = F \circ U$.

In the same way, after a (linear) change of variables, the previous statements may be formulated with the target distribution being that of the Gaussian distribution $\gamma_{m,\Sigma}$ on $\mathbb{R}^d$ with mean m and invertible covariance matrix $\Sigma = M^t M$. For example, one has the following.

COROLLARY 2.4.   *Let $F : \mathbb{R}^N \to \mathbb{R}^d$ be of class $\mathscr{C}^2$ and in $L^2(\mu)$, and denote by $\mu \circ F^{-1}$ the law of F under μ (i.e., the law of $F(X)$ on $\mathbb{R}^d$). Then*

$$
\mathrm{W}_2(\mu \circ F^{-1}, \gamma_{m,\Sigma}) \le A_m + B_{m,\Sigma},
$$

where

$$A_m = \left(\int_{\mathbb{R}^N} |F - m + K^{-1}\mathrm{L}F|^2 \, d\mu \right)^{\frac{1}{2}}$$

and

$$B_{m,\Sigma} = \|M\| \left(\int_{\mathbb{R}^N} |\mathrm{Id} - (KM)^{-1}\Gamma(F)|^2 \, d\mu \right)^{\frac{1}{2}}$$

with $\|M\|$ *being the operator norm of* M.

This is a direct application of Proposition 2.1 and we skip the proof. Let us now turn to the choice of the coefficients κ_k. In applications, the choice of the coefficients κ_k might depend on the context to some degree, but we point out that for B from (2.5) to be small, one should at least expect $\mathrm{Id} - K^{-1}\int \Gamma(F) \, d\mu$ to be small as well (say in the Hilbert–Schmidt norm). This would suggest that one natural choice, where the diagonal entries of this matrix vanish, would be that $\kappa_k = \int_{\mathbb{R}^N} \Gamma(F_k) \, d\mu, k = 1, \ldots, d$ (these are strictly positive if ∇F_k is not μ-almost surely zero). This is essentially the choice we make in our application of the normal approximation results to linear statistics of β-ensembles; this choice would correspond to $\mathbb{E}[\sum_{j=1}^N f'(\lambda_j)^2]$ while our choice will be $N \int f'(x)^2 \mu_V(dx)$.

Let us consider some consequences of this choice of K. First of all, we point out that in this case a direct calculation shows that

$$(2.9) \qquad B^2 = \sum_{k=1}^d \frac{1}{\kappa_k^2} \mathrm{Var}_\mu\big(\Gamma(F_k)\big) + \sum_{k \neq \ell} \int_{\mathbb{R}^N} \frac{1}{\kappa_k^2} \Gamma(F_k, F_\ell)^2 \, d\mu.$$

Here, $\mathrm{Var}_\mu(f) = \mathrm{Var}(f(X))$ is the variance of $f : \mathbb{R}^N \to \mathbb{R}$ with respect to μ, equivalently the variance of the random variable $f(X)$. Note that if $d = 1$, $B^2 = \frac{1}{\kappa_1^2} \mathrm{Var}_\mu(\Gamma(F_1))$.

To simplify the expression of A^2, we recall some notation and facts from [2]. If f, g are smooth functions on $\mathbb{R}^N$, set

$$\Gamma_2(f, g) = \mathrm{Hess}(f) \cdot \mathrm{Hess}(g) + \mathrm{Hess}(\mathcal{H})\nabla f \cdot \nabla g$$

$$= \sum_{i,j=1}^N \partial_{ij} f \partial_{ij} g + \sum_{i,j=1}^N \partial_{ij}\mathcal{H}\partial_j f \partial_i g.$$

As for Γ, write below $\Gamma_2(f) = \Gamma_2(f, f)$. By integration by parts (again assuming, e.g., that $\mathcal{H}$ is smooth), for smooth functions $f, g : \mathbb{R}^N \to \mathbb{R}$,

$$\int_{\mathbb{R}^N} \mathrm{L}f\mathrm{L}g \, d\mu = \int_{\mathbb{R}^N} \Gamma_2(f, g) \, d\mu.$$

Therefore, in this notation, another simple calculation using our particular choice of κ_k shows that

$$(2.10) \qquad A^2 = \sum_{k=1}^{d} \int_{\mathbb{R}^N} \left[\frac{1}{\kappa_k^2} \Gamma_2(F_k) + F_k^2 - 2 \right] d\mu.$$

We point out here that it is not obvious that this type of argument is valid for β-ensembles. Indeed, a priori, $\mathrm{L}f\mathrm{L}g$ and $\partial_{ij}\mathcal{H}$ will have terms of the form $(\lambda_i - \lambda_j)^{-2}$ and if we assume just that $\beta > 0$, this could result in a noninte-grable singularity. Nevertheless, we note that if we are interested in linear statistics, namely we have $f(\lambda) = \sum_{j=1}^{N} u(\lambda_j)$ and $g(\lambda) = \sum_{j=1}^{N} v(\lambda_j)$ for some smooth bounded functions $u, v : \mathbb{R} \to \mathbb{R}$, then by symmetry, one has, for example,

$$\mathrm{L}f(\lambda) = \sum_{j=1}^{N} u''(\lambda_j) - N\beta \sum_{j=1}^{N} V'(\lambda_j)u'(\lambda_j) + \beta \sum_{i \neq j} \frac{1}{\lambda_j - \lambda_i} u'(\lambda_j)$$

$$= \sum_{j=1}^{N} u''(\lambda_j) - N\beta \sum_{j=1}^{N} V'(\lambda_j)u'(\lambda_j) + \frac{\beta}{2} \sum_{i \neq j} \frac{u'(\lambda_j) - u'(\lambda_i)}{\lambda_j - \lambda_i},$$

which no longer has singularities. Similarly, one has in this case

$$\sum_{i,j=1}^{N} \partial_{ij}\big(\beta \mathcal{H}_V^N(\lambda)\big) \partial_i f(\lambda) \partial_j g(\lambda)$$

$$= N\beta \sum_{i=1}^{N} V''(\lambda_i)u'(\lambda_i)v'(\lambda_i) + \beta \sum_{i \neq j} \frac{u'(\lambda_i)v'(\lambda_i) - u'(\lambda_i)v'(\lambda_j)}{(\lambda_i - \lambda_j)^2},$$

which has only singularities of type $\frac{1}{\lambda_i - \lambda_j}$ so as we are integrating against $\prod_{i<j} |\lambda_i - \lambda_j|^\beta$ with $\beta > 0$, we have only integrable singularities. Thus integration by parts is again justified in the setting we are considering.

Turning back to more general $\mathcal{H}$, we mention some further simplifications or bounds one can make use of in some special cases. Let us still assume that $\kappa_k = \int_{\mathbb{R}^N} \Gamma(F_k) \, d\mu$, $k = 1, \ldots, d$. In case the measure μ satisfies a Poincaré inequality in the sense that for any (smooth) function $f : \mathbb{R}^N \to \mathbb{R}$,

$$(2.11) \qquad \mathrm{Var}_\mu(f) \leq C \int_{\mathbb{R}^N} \Gamma(f) \, d\mu,$$

(e.g., $C = 1$ for $\mu = \gamma_N$, cf. [2], Chapter 4), the quantity B, rather B^2 from (2.9), is advantageously upper bounded by

$$(2.12) \qquad B'^2 = \sum_{k=1}^{d} \frac{C}{\kappa_k^2} \int_{\mathbb{R}^N} \Gamma(\Gamma(F_k)) \, d\mu + \sum_{k \neq \ell} \int_{\mathbb{R}^N} \frac{1}{\kappa_k^2} \Gamma(F_k, F_\ell)^2 \, d\mu.$$

In particular, in the setting of Proposition 2.2,

$$\|\mu \circ F^{-1} - \gamma_1\|_{\mathrm{TV}} \le 2\left(\int_{\mathbb{R}^N}\left[F + \frac{1}{\kappa}\mathrm{L}F\right]^2 d\mu\right)^{\frac{1}{2}}$$
$$+ \frac{2}{\kappa}\left(C\int_{\mathbb{R}^N}\Gamma\big(\Gamma(F)\big)d\mu\right)^{\frac{1}{2}}.$$

Before turning to the proofs, we mention that Propositions 2.1 and 2.2 are related to several earlier and parallel investigations. They may first be viewed as a W_2-version of the multivariate normal approximation in the Kantorovich metric W_1 developed by Meckes in [24] and relying on exchangeable pairs. For the application to linear statistics of random matrices, the pairs are created through Dyson–Brownian motion, and also in this setting, the operator L appears naturally and plays an important role.

The quantities A and B arising in Propositions 2.1 and 2.2 are also connected to earlier bounds in the literature. As a first instance, assume that $F : \mathbb{R}^N \to \mathbb{R}$ is an eigenvector of L in the sense that $-\mathrm{L}F = \kappa F$, normalized in $L^2(\mu)$ so that

$$\int_{\mathbb{R}^N}\Gamma(F)\,d\mu = \int_{\mathbb{R}^N}F(-\mathrm{L}F)\,d\mu = \kappa.$$

In this case, the inequality of Proposition 2.2 amounts to

$$\|\mu \circ F^{-1} - \gamma_1\|_{\mathrm{TV}} \le \frac{2}{\kappa}\mathrm{Var}_\mu\big(\Gamma(F)\big)^{\frac{1}{2}},$$

a result already put forward in [20]. As such, Propositions 2.1 and 2.2 suggest a similar result provided that F is approximately an eigenvector in the sense that A is small. This indeed is a central theme in our approach to proving the CLT for linear statistics of β-ensembles, and the motivation for our choice of the function F in later sections.

Another source of comparison is the works [10] and [26]. To emphasize the comparison, let us deal with the one-dimensional case $F : \mathbb{R}^N \to \mathbb{R}$ corresponding to Proposition 2.2. In the present notation, provided that $\int_{\mathbb{R}^N} F^2\,d\mu = 1$, the methodology of [10, 26] develops toward the inequality

$$\|\mu \circ F^{-1} - \gamma_1\|_{\mathrm{TV}} \le \mathrm{Var}_\mu(T),$$

where $T = \Gamma((-\mathrm{L})^{-1}F, F)$ and $(-\mathrm{L})^{-1}$ is the formal inverse of the positive operator $-\mathrm{L}$ ($-\mathrm{L}$ being positive because of the integration by parts formula (2.3)). Then, provided the measure μ satisfies a Poincaré inequality, the preceding variance is bounded from above by moments of differentials of F, as in A and B of Propositions 2.1 and 2.2.

An alternative point of view on the latter results may be expressed in terms of the Stein discrepancy and the inequality

$$(2.13) \qquad\qquad \mathrm{W}_2\big(\mu \circ F^{-1}, \gamma_d\big) \le \mathrm{S}_2\big(\mu \circ F^{-1}|\gamma_d\big)$$

emphasized in [21] where

$$S_2\big(\mu \circ F^{-1} | \gamma_d\big) = \left(\int_{\mathbb{R}^d} |\tau_{\mu \circ F^{-1}} - \mathrm{Id}|^2 \, d\mu \circ F^{-1} \right)^{\frac{1}{2}}$$

with $\tau_{\mu \circ F^{-1}}$ a so-called Stein kernel of the distribution $\mu \circ F^{-1}$. In dimension $d = 1$, this Stein kernel is characterized by

$$\int_{\mathbb{R}^N} F\varphi(F) \, d\mu = \int_{\mathbb{R}} x\varphi \, d\mu \circ F^{-1} = \int_{\mathbb{R}} \tau_{\mu \circ F^{-1}} \varphi' \, d\mu \circ F^{-1}$$

$$= \int_{\mathbb{R}^N} \tau_{\mu \circ F^{-1}}(F)\varphi'(F) \, d\mu$$

for every smooth $\varphi : \mathbb{R} \to \mathbb{R}$. In the previous notation, the kernel $\tau_{\mu \circ F^{-1}}$ may be described as the conditional expectation of $T = \Gamma((-\mathrm{L})^{-1}F, F)$ given F, so that, again under the normalization $\int_{\mathbb{R}^N} F^2 \, d\mu = 1$,

$$W_2\big(\mu \circ F^{-1}, \gamma_1\big)^2 \le S_2\big(\mu \circ F^{-1} | \gamma_1\big)^2 \le \mathrm{Var}_\mu(T).$$

With respect to the analysis of these prior contributions [10, 21, 26], the approach developed in Propositions 2.1 and 2.2 is additive rather than multiplicative, and concentrates directly on the generator L rather than its (possibly cumbersome) inverse. The formulation of Propositions 2.1 and 2.2 allows us to recover, sometimes at a cheaper price, several of the conclusions and illustrations developed in [10].

2.2. *Proof of Proposition 2.1 and Proposition 2.2.* The main argument of the proof relies on standard semigroup interpolation (cf. [2]) together with steps from [21]. Denote by $(P_t)_{t \ge 0}$ the Ornstein–Uhlenbeck semigroup on $\mathbb{R}^d$, with invariant measure γ_d the standard Gaussian measure on $\mathbb{R}^d$ and associated infinitesimal generator $\mathcal{L}_d = \Delta - x \cdot \nabla$. The operator P_t admits the classical integral representation

$$(2.14) \qquad P_t f(x) = \int_{\mathbb{R}^d} f\big(e^{-t}x + \sqrt{1 - e^{-2t}} y\big) \gamma_d(dy), \qquad t \ge 0, x \in \mathbb{R}^d.$$

A basic property of this operator we shall make use of is that it is symmetric with respect to the inner product of $L^2(\gamma_d)$ in the sense that for, say, bounded continuous functions $f, g : \mathbb{R}^d \to \mathbb{R}$, $\int_{\mathbb{R}^d} f P_t g \, d\gamma_d = \int_{\mathbb{R}^d} g P_t f \, d\gamma_d$ for each $t > 0$ (cf. [2], Chapter 2, Section 2.7).

We start with the first approximation result.

PROOF OF PROPOSITION 2.1. Assume first that $F : \mathbb{R}^N \to \mathbb{R}^d$ is smooth and such that the law $\mu \circ F^{-1}$ of F admits a smooth and positive density h with respect to γ_d. Denote then by

$$I(P_t h) = \int_{\mathbb{R}^d} \frac{|\nabla P_t h|^2}{P_t h} \, d\gamma_d, \qquad t \ge 0,$$

the Fisher information of the density $P_t h$ with respect to γ_d, which is assumed to be finite. It is also assumed throughout the proof that A and B are finite otherwise there is nothing to show. With $v_t = \log P_t h$, $t \geq 0$, after integration by parts and symmetry of P_t with respect to γ_d (cf. the analysis in [21]),

$$\mathrm{I}(P_t h) = \int_{\mathbb{R}^d} \frac{|\nabla P_t h|^2}{P_t h} \, d\gamma_d = -\int_{\mathbb{R}^d} \mathcal{L}_d[v_t] P_t h \, d\gamma_d = -\int_{\mathbb{R}^d} \mathcal{L}_d[P_t v_t] \, d\mu \circ F^{-1}.$$

Now, for any $t > 0$, and any $\kappa_k > 0$, $k = 1, \ldots, d$,

$$\begin{aligned}
\mathrm{I}(P_t h) &= -\int_{\mathbb{R}^N} \mathcal{L}_d[P_t v_t](F) \, d\mu \\
&= -\int_{\mathbb{R}^N} \left[\sum_{k=1}^d \partial_{kk} P_t v_t(F) - \sum_{k=1}^d F_k \partial_k P_t v_t(F) \right] d\mu \\
&= \int_{\mathbb{R}^N} \sum_{k=1}^d \left[F_k + \frac{1}{\kappa_k} \mathrm{L} F_k \right] \partial_k P_t v_t(F) \, d\mu \\
&\quad + \int_{\mathbb{R}^N} \sum_{k,\ell=1}^d \left[\frac{1}{\kappa_k} \Gamma(F_k, F_\ell) - \delta_{k\ell} \right] \partial_{k\ell} P_t v_t(F) \, d\mu,
\end{aligned}$$

where the last step follows by adding and subtracting $\frac{1}{\kappa_k} \mathrm{L}[F_k] \partial_k P_t v_t(F)$ and integrating by parts with respect to L, that is, formula (2.3).

Next, by the Cauchy–Schwarz inequality,

$$\int_{\mathbb{R}^N} \sum_{k=1}^d \left[F_k + \frac{1}{\kappa_k} \mathrm{L} F_k \right] \partial_k P_t v_t(F) \, d\mu \leq A \left(\int_{\mathbb{R}^N} |\nabla P_t v_t(F)|^2 \, d\mu \right)^{\frac{1}{2}}$$

and

$$\begin{aligned}
\int_{\mathbb{R}^N} |\nabla P_t v_t(F)|^2 \, d\mu &= \int_{\mathbb{R}^d} |\nabla P_t v_t|^2 \, d\mu \circ F^{-1} \\
&\leq e^{-2t} \int_{\mathbb{R}^d} P_t (|\nabla v_t|^2) \, d\mu \circ F^{-1} \\
&= e^{-2t} \mathrm{I}(P_t h).
\end{aligned}$$

Now, for every $k, \ell = 1, \ldots, d$, $\partial_{k\ell} P_t v_t = e^{-2t} P_t(\partial_{k\ell} v_t)$ and, by integration by parts in the integral representation of P_t,

$$\partial_{k\ell} P_t v_t(x) = \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}^d} y_k \partial_\ell v_t \left(e^{-t} x + \sqrt{1 - e^{-2t}} \, y \right) \gamma_d(dy).$$

Then, by another application of the Cauchy–Schwarz inequality,

$$\int_{\mathbb{R}^N} \sum_{k,\ell=1}^{d} \left[\frac{1}{\kappa_k} \Gamma(F_k, F_\ell) - \delta_{k\ell} \right] \partial_{k\ell} P_t v_t(F) \, d\mu$$

$$\leq \frac{e^{-2t} B}{\sqrt{1 - e^{-2t}}}$$

$$\times \left(\int_{\mathbb{R}^d} \sum_{k,\ell=1}^{d} \left[\int_{\mathbb{R}^d} y_k \partial_\ell v_t \big(e^{-t} x + \sqrt{1 - e^{-2t}} \, y \big) \gamma_d(dy) \right]^2 \mu \circ F^{-1}(dx) \right)^{\frac{1}{2}}$$

$$\leq \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} B$$

$$\times \left(\int_{\mathbb{R}^d} \sum_{\ell=1}^{d} \int_{\mathbb{R}^d} \big[\partial_\ell v_t \big(e^{-t} x + \sqrt{1 - e^{-2t}} \, y \big) \big]^2 \gamma_d(dy) \mu \circ F^{-1}(dx) \right)^{\frac{1}{2}}$$

$$= \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} B \left(\int_{\mathbb{R}^d} P_t \big(|\nabla v_t|^2 \big) \, d\mu \circ F^{-1} \right)^{\frac{1}{2}}$$

$$= \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} B \sqrt{\mathrm{I}(P_t h)},$$

where we used at the second step that for a given function $g : \mathbb{R}^d \to \mathbb{R}$ in $L^2(\gamma_d)$,

$$\sum_{k=1}^{d} \left(\int_{\mathbb{R}^d} y_k g(y) \gamma_d(dy) \right)^2 \leq \int_{\mathbb{R}^d} g^2 \, d\gamma_d,$$

which follows from the remark that the operator $g \mapsto \sum_{k=1}^{d} x_k \int y_k g(y) \gamma_d(dy)$ is an orthogonal projection on $L^2(\gamma_d)$.

Altogether, it follows that for every $t > 0$,

$$\mathrm{I}(P_t h) \leq \left(e^{-t} A + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} B \right) \sqrt{\mathrm{I}(P_t h)},$$

hence

$$\sqrt{\mathrm{I}(P_t h)} \leq e^{-t} A + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} B.$$

From [27], Lemma 2 (cf. also [36], Theorem 24.2(iv)), it follows that

$$\mathrm{W}_2 \big(\mu \circ F^{-1}, \gamma \big) \leq \int_0^\infty \sqrt{\mathrm{I}(P_t h)} \, dt \leq A + B,$$

which is the announced result in this case.

The general case is obtained by a regularization procedure which we outline in dimension $d = 1$. Let therefore $F : \mathbb{R}^N \to \mathbb{R}$ be of class $\mathscr{C}^2$ and in $L^2(\mu)$. Fix $\varepsilon > 0$ and consider, on $\mathbb{R}^N \times \mathbb{R}$, the vector (X, Z) where Z is a standard normal independent of X and

$$F_\varepsilon(x, z) = e^{-\varepsilon} F(x) + \sqrt{1 - e^{-2\varepsilon}} z, \qquad x \in \mathbb{R}^N, z \in \mathbb{R}.$$

Denote by μ_{F_ε} the distribution of $F_\varepsilon(X, Z)$, or image of $\mu \otimes \gamma_1$ under F_ε. The probability measure μ_{F_ε} admits a smooth and positive density h_ε with respect to γ_1 given by

$$h_\varepsilon(x) = \int_{\mathbb{R}} p_\varepsilon(x, y) \mu \circ F^{-1}(dy), \qquad x \in \mathbb{R},$$

where $p_t(x, y)$, $t > 0$, $x, y \in \mathbb{R}$, is the Mehler kernel of the semigroup representation (2.14) (and coincides with $P_\varepsilon h$ whenever $\mu \circ F^{-1}$ admits a density h with respect to γ_1). Furthermore, by the explicit representation of $p_\varepsilon(x, y)$ (cf., e.g., [2], (2.7.4)),

$$h'_\varepsilon(x) = -\frac{e^{-2\varepsilon}}{1 - e^{-2\varepsilon}} \int_{\mathbb{R}} [x - e^\varepsilon y] p_\varepsilon(x, y) \mu \circ F^{-1}(dy),$$

and from the Cauchy–Schwarz inequality, we obtain for any $x \in \mathbb{R}$,

$$\frac{h'_\varepsilon(x)^2}{h_\varepsilon(x)} \leq \left(\frac{e^{-2\varepsilon}}{1 - e^{-2\varepsilon}} \right)^2 \int_{\mathbb{R}} [x - e^\varepsilon y]^2 p_\varepsilon(x, y) \mu \circ F^{-1}(dy).$$

Since $F \in L^2(\mu)$, it follows that $\int_{\mathbb{R}} \frac{h'^2_\varepsilon}{h_\varepsilon} \, d\gamma_1 < \infty$ for any $\varepsilon > 0$. Now, for every $t \geq 0$, $P_t h_\varepsilon = h_{t+\varepsilon}$, so that the Fisher information $\mathrm{I}(P_t h_\varepsilon)$, $t \geq 0$, are well defined and finite. The proof we have presented thus far then applies to F_ε on the product space $\mathbb{R}^N \times \mathbb{R}$ with respect to the generator $\mathrm{L} \oplus \mathcal{L}_1$. Next $F_\varepsilon \to F$ in $L^2(\mu \times \gamma_1)$ from which

$$\mathrm{W}_2\big(\mu \circ F_\varepsilon^{-1}, \mu \circ F^{-1}\big)^2 \leq \int_{\mathbb{R}^N \times \mathbb{R}} |F_\varepsilon - F|^2 \, d\mu \times \gamma_1 \to 0.$$

Hence, by the triangle inequality, $\mathrm{W}_2(\mu \circ F_\varepsilon^{-1}, \gamma_1) \to \mathrm{W}_2(\mu \circ F^{-1}, \gamma_1)$. On the other hand, the Γ-calculus of [2], Chapter 3, developed on $\mathbb{R}^N \times \mathbb{R}$ yields the quantities A and B in the limit as $\varepsilon \to 0$. The proof of Proposition 2.1 is complete. $\square$

We next turn to our second approximation result which is similar but stays at the first order on the basis of the standard Stein equation.

PROOF OF PROPOSITION 2.2. The classical Stein lemma states that given a bounded function $\varphi : \mathbb{R} \to \mathbb{R}$, the equation

$$(2.15) \qquad \psi' - x\psi = \varphi - \int_{\mathbb{R}} \varphi \, d\gamma_1$$

may be solved with a function ψ, which along with its derivative, is bounded. More precisely, ψ may be chosen so that $\|\psi\|_\infty \leq \sqrt{2\pi} \|\varphi\|_\infty$ and $\|\psi'\|_\infty \leq 4\|\varphi\|_\infty$. Stein's lemma may then be used to provide the basic approximation bound

$$(2.16) \qquad \|\lambda - \gamma_1\|_{\mathrm{TV}} \leq \sup \left| \int_{\mathbb{R}} \psi'(x)\, d\lambda(x) - \int_{\mathbb{R}} x\psi(x)\, d\lambda(x) \right|,$$

where the supremum runs over all continuously differentiable functions $\psi : \mathbb{R} \to \mathbb{R}$ such that $\|\psi\|_\infty \leq \sqrt{\frac{\pi}{2}}$ and $\|\psi'\|_\infty \leq 2$.

We thus investigate

$$\int_{\mathbb{R}} \psi'(x)\, d\mu \circ F^{-1}(x) - \int_{\mathbb{R}} x\psi(x)\, d\mu \circ F^{-1}(x)$$

$$= \int_{\mathbb{R}^N} \psi'(F)\, d\mu - \int_{\mathbb{R}^N} F\psi(F)\, d\mu$$

and proceed as in the proof of Proposition 2.1. Namely, with $\kappa > 0$, by the integration by parts formula (2.3),

$$\int_{\mathbb{R}^N} [\psi'(F) - F\psi(F)]\, d\mu = -\int_{\mathbb{R}^N} \psi(F) \left[F + \frac{1}{\kappa} \mathrm{L}F \right] d\mu$$

$$+ \int_{\mathbb{R}^N} \psi'(F) \left[1 - \frac{1}{\kappa} \Gamma(F) \right] d\mu.$$

As a consequence,

$$\int_{\mathbb{R}^N} [\psi'(F) - F\psi(F)]\, d\mu \leq \|\psi\|_\infty \left(\int_{\mathbb{R}^N} \left[F + \frac{1}{\kappa} \mathrm{L}F \right]^2 d\mu \right)^{\frac{1}{2}}$$

$$+ \|\psi'\|_\infty \left(\int_{\mathbb{R}^N} \left[1 - \frac{1}{\kappa} \Gamma(F) \right]^2 d\mu \right)^{\frac{1}{2}}.$$

By definition of the total variation distance and Stein's lemma, Proposition 2.2 follows. It remains to briefly analyze (2.7) and (2.8). Arguing as in the proof (2.6), we find that

$$\int_{\mathbb{R}^N} [\psi'(F) - F\psi(F)]\, d\mu = -\int_{\mathbb{R}^N} \psi(F) \sum_{k=1}^{d} \theta_k \left[F_k + \frac{1}{\kappa_k} \mathrm{L}F_K \right] d\mu$$

$$+ \int_{\mathbb{R}^N} \psi'(F) \sum_{k,\ell=1}^{d} \theta_k \theta_\ell \left[\delta_{k\ell} - \frac{1}{\kappa_k} \Gamma(F_k, F_\ell) \right] d\mu.$$

(2.7) then follows from the Cauchy–Schwarz inequality and the definition of A and B. We note that (2.8) also holds as a consequence of Proposition 2.1 since, as is easily checked,

$$\mathrm{W}_2(\mu \circ G^{-1}, \gamma_1) \leq \mathrm{W}_2(\mu \circ F^{-1}, \gamma_d).$$

The proof of Proposition 2.2 is complete. $\quad\square$

3. The CLT for the GUE—Proof of Theorem 1.4.
On the basis of the general normal approximation results put forward in Section 2, we address the proof Theorem 1.4. In this section, we will write simply $\mathbb{P}$ and L for $\mathbb{P}^N_{V,\beta}$ and $L^N_{V,\beta}$ with $\beta = 2$ and $V(x) = x^2$, as well as $\mathbb{E}$ for the expectation with respect to $\mathbb{P}$.

PROOF OF THEOREM 1.4. We proceed in several steps, starting with establishing the fact that linear statistics of Chebyshev polynomials of the first kind are approximate eigenvectors (in a sense that will be made precise in the course of the proof) of the operator L. Then we move on to controlling the error in this approximate eigenvector property in order to apply Proposition 2.1 to get a joint CLT for linear statistics of Chebyshev polynomials. Finally, we expand a general polynomial in terms of Chebyshev polynomials to finish the proof.

Step 1—approximate eigenvector property. Recall that T_k and U_k, $k \geq 0$, denote the degree k Chebyshev polynomial of the first kind, respectively of the second kind; we refer to the Appendix for further details about the definition and basic properties of Chebyshev polynomials. See also [9] and [22] for related discussions.

We begin by noting that a direct application of Lemma A.3 shows that when $x \in J$, $k \geq 0$,

$$(3.1) \qquad \int_J \frac{T'_k(x) - T'_k(y)}{x - y} \mu_{\mathrm{sc}}(dy) = 2x T'_k(x) - 2k T_k(x).$$

As both sides of this equation are polynomials, this identity is actually valid for all $x \in \mathbb{R}$. We also point out here that this is precisely (1.14) for $V(x) = x^2$ with $\sigma_k = 2k$. In fact, this part of the proof is completely independent of β, but since it is much simpler to control the various error terms when $\beta = 2$, for simplicity, we choose to stick to $\beta = 2$ throughout the proof. The general case is covered by Theorem 1.2.

Next, we note that integrating (A.10) with respect to $\mu_{\mathrm{sc}}(dx)$ and making use of (A.9) along with the facts that $T'_k = k U_{k-1}$, $2x = U_1(x)$, and $(1 - x^2) = -\frac{1}{2}(T_2(x) - T_0(x))$, we find, for every $k \geq 1$,

$$
\begin{aligned}
\iint_{J \times J} & \frac{T'_k(x) - T'_k(y)}{x - y} \mu_{\mathrm{sc}}(dx) \mu_{\mathrm{sc}}(dy) \\
&= k \int_J U_1(x) U_{k-1}(x) \mu_{\mathrm{sc}}(dx) \\
&\quad + 2k \int_J T_k(x)\bigl(T_2(x) - T_0(x)\bigr) \varrho(dx) \\
&= 2k \delta_{k,2} \\
&= -4k \int_J T_k(x) \mu_{\mathrm{sc}}(dx).
\end{aligned}
$$

(3.2)

Applying first (3.1) and then (3.2), we see that

$$L\left(\sum_{j=1}^{N} T_k(\lambda_j)\right) = \sum_{j=1}^{N} T_k''(\lambda_j) - 4N \sum_{j=1}^{N} \lambda_j T_k'(\lambda_j) + 2 \sum_{i \neq j} \frac{T_k'(\lambda_j)}{\lambda_j - \lambda_i}$$

$$= -4kN \sum_{j=1}^{N} T_k(\lambda_j) - 2N \sum_{i=1}^{N} \int_J \frac{T_k'(\lambda_i) - T_k'(y)}{\lambda_i - y} \mu_{sc}(dy)$$

$$(3.3) \qquad + \sum_{i,j=1}^{N} \frac{T_k'(\lambda_i) - T_k'(\lambda_j)}{\lambda_i - \lambda_j}$$

$$= -4kN \left(\sum_{j=1}^{N} T_k(\lambda_j) - N \int_J T_k(x) \mu_{sc}(dx)\right)$$

$$+ \int_{\mathbb{R} \times \mathbb{R}} \frac{T_k'(x) - T_k'(y)}{x - y} \nu_N(dx) \nu_N(dy),$$

where ν_N is given by formula (1.5) with the equilibrium measure $\mu_V = \mu_{sc}$. We also used the fact that for all $k \in \mathbb{N}$, by symmetry,

$$2 \sum_{i \neq j} \frac{T_k'(\lambda_j)}{\lambda_j - \lambda_i} = \sum_{i,j=1}^{N} \frac{T_k'(\lambda_j) - T_k'(\lambda_i)}{\lambda_j - \lambda_i} - \sum_{i=1}^{N} T_k''(\lambda_i)$$

with the interpretation that the diagonal terms in the double sum are $T_k''(\lambda_i)$. This shows that the *re-centered* linear statistics $\int T_k(x)\nu_N(dx)$ are approximate eigenfunctions in the sense that

$$(3.4) \qquad L\left(\int_{\mathbb{R}} T_k(x)\nu_N(dx)\right) = -4kN \int_{\mathbb{R}} T_k(x)\nu_N(dx) + \zeta_k(\lambda),$$

where for any configuration $\lambda \in \mathbb{R}^N$,

$$(3.5) \qquad \zeta_k(\lambda) = \int_{\mathbb{R} \times \mathbb{R}} \frac{T_k'(x) - T_k'(y)}{x - y} \nu_N(dx) \nu_N(dy)$$

is interpreted as an error term. In fact, knowing the CLT for linear statistics, one can check that for any $k \in \mathbb{N}$, ζ_k converges in law to a finite sum of products of Gaussian random variables as $N \to \infty$, so that its fluctuations are negligible in comparison to $4kN$ (though we make no use of this). In particular, we will choose the $d \times d$ diagonal matrix $K = K_N$ appearing in the definition of A and B in (2.4) and (2.5) to have entries $K_{kk} = 4kN$.

Step 2—a priori bound on linear statistics. To apply Proposition 2.1, we will make use of the fact that $\frac{T_k'(x) - T_k'(y)}{x - y}$ is a polynomial in x and y so that we can express each ζ_k as a sum of products of (centered) polynomial linear statistics. Thus to obtain a bound on A in Proposition 2.1, it will be enough to have some

bound on the second moment of such statistics. In the case of the GUE, we are able to apply rather soft arguments for this (as opposed to the fact that we rely on rigidity estimates for general V and β).

It is known that for any polynomial f, the error term in Wigner's classical limit theorem (namely in the statement that $\lim_{N \to \infty} \frac{1}{N} \mathbb{E}[\sum_{j=1}^{N} f(\lambda_j)] = \int f(x) \mu_{sc}(dx)$) is of order $\frac{1}{N^2}$ (see, e.g., [25]). That is, we have

$$(3.6) \qquad \left| \mathbb{E}\left[\int_{\mathbb{R}} f(x) \nu_N(dx) \right] \right| \ll_f N^{-1}.$$

In addition, it is known that for any $p > 0$,

$$(3.7) \qquad \sup_{N \geq 1} \mathbb{E}\left| \int_{\mathbb{R}} f(x) \nu_N(dx) \right|^p < \infty.$$

This property seems part of the folklore (cf. [1, 29]) but we provide here a proof for completeness. For $\beta = 2$ and $V(x) = x^2$, the Hamiltonian $\beta \mathcal{H}_V^N$ in (1.1) is more convex than the quadratic potential $N \sum_{j=1}^{N} 2\lambda_j^2$. In particular, the GUE eigenvalue measure $\mathbb{P}_{x^2,2}^N$ satisfies a logarithmic Sobolev inequality with constant N^{-1} ([1], Theorem 4.4.18, p. 290, or Exercise 4.4.33, p. 302, or [2], Corollary 5.7.2) and by the resulting moment bounds (cf. [2], Proposition 5.4.2), for every smooth $g : \mathbb{R}^N \to \mathbb{R}$,

$$(3.8) \qquad \mathbb{E}|g - \mathbb{E}g|^p \leq \frac{C_p}{N^{p/2}} \mathbb{E}|\nabla g|^p$$

for some constant $C_p > 0$ only depending on $p \geq 2$. We thus apply the latter to $g = \int f(x) \nu_N(dx)$. At this stage, we detail the argument for the value $p = 4$ (which will be used below) but the proof is the same for any (integer) p (at the price of repeating the step). First, we have

$$\mathbb{E}|\nabla g|^4 = \mathbb{E}\left(\sum_{i=1}^{N} f'(x_i)^2 \right)^2$$

$$= \mathbb{E}\left(\int_{\mathbb{R}} f'(x)^2 \nu_N(dx) + N \int_{J} f'(x)^2 \mu_{sc}(dx) \right)^2$$

and, by (3.8),

$$(3.9) \qquad \begin{aligned} \mathbb{E}\left| \int_{\mathbb{R}} f'(x)^2 \nu_N(dx) - \mathbb{E}\left[\int_{\mathbb{R}} f'(x)^2 \nu_N(dx) \right] \right|^2 \\ \leq \frac{4C_2}{N} \mathbb{E}\left[\sum_{i=1}^{N} f''(x_i)^2 f'(x_i)^2 \right]. \end{aligned}$$

By Wigner's law, the RHS of (3.9) is uniformly bounded in N and by (3.6), if f is a polynomial, we obtain by the triangle inequality that

$$\mathbb{E}\left| \int_{\mathbb{R}} f'(x)^2 \nu_N(dx) \right|^2 \ll_f 1,$$

which implies that $\mathbb{E}|\nabla g|^4 \ll_f N^2$. Then, using the estimate (3.8) once more, we obtain

$$\mathbb{E}|g - \mathbb{E}g|^4 \leq \frac{C_4}{N^2}\mathbb{E}|\nabla g|^4 \ll_f 1.$$

By (3.6), we also have that $|\mathbb{E}g| \ll N^{-1}$ and the claim (3.7) follows by the triangle inequality.

Step 3—controlling B. We now move on to applying Proposition 2.1. Noting that by (1.12), one has $\boldsymbol{\Sigma}(T_k) = \frac{k}{4}$, we define $F = (F_1, \ldots, F_d) : \mathbb{R}^N \to \mathbb{R}^d$ by

$$F_k = \frac{2}{\sqrt{k}} \int T_k(x) v_N(dx),$$

where d is independent of N (in contrast to the proof of Proposition 1.6 in Section 5). Then, according to (2.5), we have

$$
\begin{aligned}
B^2 &= \mathbb{E}\big|\mathrm{Id} - K^{-1}\Gamma(F)\big|^2 \\
&= \sum_{i,k=1}^{d} \mathbb{E}\Big(\delta_{ik} - \frac{\Gamma(F_i, F_k)}{4kN}\Big)^2 \\
&= \sum_{k=1}^{d} \mathbb{E}\Big(1 - \frac{\Gamma(F_k)}{4kN}\Big)^2 + \sum_{i \neq k} \frac{\mathbb{E}\Gamma(F_i, F_k)^2}{16k^2 N^2}.
\end{aligned}
$$

(3.10)

By definition of Γ from (2.3), we have

$$\Gamma(F_k) = \frac{4}{k}\sum_{j=1}^{N} T_k'(\lambda_j)^2 = \frac{4}{k}\int_{\mathbb{R}} T_k'(x)^2 v_N(dx) + \frac{4}{k}N\int_{\mathrm{J}} T_k'(x)^2 \mu_{\mathrm{sc}}(dx).$$

Now, recalling that the Chebyshev polynomial of the second kind are orthonormal with respect to the semicircle law (see formulae (A.8) and (A.9)), we see that

$$\frac{\Gamma(F_k)}{4kN} = \frac{1}{k^2 N}\int T_k'(x)^2 v_N(dx) + 1$$

so that using the estimate (3.7), we obtain

$$\sum_{k=1}^{d} \mathbb{E}\Big(1 - \frac{\Gamma(F_k)}{4kN}\Big)^2 \ll_d N^{-2}.$$

A similar argument shows that for $i \neq k$,

$$
\begin{aligned}
\Gamma(F_i, F_k)^2 &= \frac{4}{ik}\sum_{j=1}^{N} T_i'(\lambda_j)T_k'(\lambda_j) \\
&= \frac{4}{ik}\int_{\mathbb{R}} T_i'(x)T_k'(x)v_N(dx)
\end{aligned}
$$

and

$$\sum_{i \neq k} \frac{\mathbb{E}\Gamma(F_i, F_k)^2}{16k^2N^2} \ll_d N^{-2}.$$

Combining the two previous estimates, by (3.10), we conclude that $B \ll_d N^{-1}$.

Step 4—controlling A. Let us begin by noting that according to (2.4) and the approximate eigenequation (3.4), we have

$$(3.11) \qquad A^2 = \mathbb{E}|F + K^{-1}\mathrm{L}F|^2 = \sum_{k=1}^{d} \frac{\mathbb{E}|\zeta_k|^2}{4k^3N^2}.$$

Since we are dealing with polynomials, there exists real coefficients $a_{i,j}^{(k)}$ which are zero if $i + j > k - 2$ so that for any $k \in \mathbb{N}$,

$$\frac{T_k'(x) - T_k'(y)}{x - y} = \sum_{i,j \geq 0} a_{i,j}^{(k)} x^i y^j.$$

This implies that the error term (3.5) factorizes:

$$\zeta_k = \sum_{i,j \geq 0} a_{i,j}^{(k)} \int_{\mathbb{R}} x^i \nu_N(dx) \int_{\mathbb{R}} y^j \nu_N(dy).$$

Thus, by (3.7) and Cauchy–Schwarz, we obtain that for any $k \in \mathbb{N}$,

$$\mathbb{E}|\zeta_k|^2 \ll_k 1.$$

Hence, by (3.11), we conclude that $A \ll_d N^{-1}$. Combining this estimate with the one coming from Step 3, we see that Proposition 2.1 gives, for any fixed $d \geq 1$, a multidimensional CLT for the linear statistics associated to Chebyshev polynomials

$$(3.12) \qquad \mathrm{W}_2(\mu \circ F^{-1}, \gamma_d) \ll_d N^{-1},$$

where $\mu \circ F^{-1}$ refers to the law of the vector $F = (\frac{2}{\sqrt{k}} \int T_k(x)\nu_N(dx))_{k=1}^{d}$.

Step 5—extending to general polynomials. Consider now an arbitrary degree d polynomial f. We can always expand it in the basis of Chebyshev polynomials of the first kind: $f = \sum_{k=0}^{d} f_k T_k$. By definition, this implies that

$$(3.13) \qquad \int_{\mathbb{R}} f(x)\nu_N(dx) = \frac{1}{2} \sum_{k=0}^{d} \sqrt{k} f_k F_k.$$

Moreover, by formula (1.12), one has $\mathbf{\Sigma}(f) = \frac{1}{4} \sum_{k=0}^{d} k f_k^2$. Then, since $\frac{1}{2\sqrt{\mathbf{\Sigma}(f)}} \times \sum_{k=0}^{d} \sqrt{k} f_k X_k \sim \gamma_1$ if the vector $X \sim \gamma_d$, it follows from the definition of the Kantorovich distance and the representation (3.13) that

$$\mathrm{W}_2^2\left(\int_{\mathbb{R}} f(x)\nu_N(dx), \sqrt{\mathbf{\Sigma}(f)}\gamma_1\right) \leq \mathbf{\Sigma}(f)\mathrm{W}_2^2(\mu \circ F^{-1}, \gamma_d).$$

Then the first statement of Theorem 1.4 follows directly, after simplifying by $\Sigma(f) > 0$, from the multidimensional CLT (3.12).

Step 6—optimality. We conclude the proof of Theorem 1.4 by verifying the optimality claim. Consider now $f(x) = x^2$. In this case indeed, the sum $\sum_{j=1}^{N} \lambda_j^2$ corresponds to the trace of the square of a GUE matrix and is thus equals $\frac{1}{4N}$ times a random variable whose law is the χ^2-distribution with $\frac{1}{2}N(N+1)$ degrees of freedom. Now, by standard arguments, given $X \sim \gamma_d$, we have

$$\left| \mathbb{E} \sin\left(\sum_{k=1}^{d} \frac{X_k^2}{\sqrt{d}} - \sqrt{d} \right) - \frac{c}{\sqrt{d}} \right| \ll \frac{1}{d},$$

where $c \neq 0$. Comparing this to the Kantorovich–Rubinstein representation of the W_1-distance shows that the total variation and all Kantorovich distances W_p (with $p > 1$) between the distribution of $\sum_{k=1}^{d} \frac{X_k^2}{\sqrt{d}} - \sqrt{d}$ and the limiting Gaussian distribution cannot be better than $\frac{1}{\sqrt{d}}$ in the $d \to \infty$ limit. The optimality claim then follows from this with a simple argument. $\square$

4. Existence and properties of the functions ϕ_n—Proof of Theorem 1.5. A central object in our construction of the functions ϕ_n in Theorem 1.5 is the following Sobolev space:

$$(4.1) \qquad \mathcal{H} = \left\{ g \in L^2(\varrho) : g' \in L^2(\mu_{\mathrm{sc}}) \text{ and } \int_J g(x) \varrho(dx) = 0 \right\},$$

which is equipped with the inner product

$$\langle f, g \rangle = \int_J f'(x) g'(x) \mu_{\mathrm{sc}}(dx).$$

In the following, we will consider other inner products defined on $\mathcal{H}$ which are given by

$$(4.2) \qquad \langle f, g \rangle_{\mu_V} = \int_J f'(x) g'(x) \mu_V(dx).$$

Note that, since we assume that $d\mu_V = S \, d\mu_{\mathrm{sc}}$ and $S > 0$ on $\bar{J}$,

$$\|g\| \asymp \|g\|_{\mu_V}$$

for all $g \in \mathcal{H}$. Thus the norms of the Hilbert spaces $\mathcal{H}_{\mu_V} := (\mathcal{H}, \langle \cdot \rangle_{\mu_V})$ are equivalent to that of $\mathcal{H}$. Recall also that we suppose that $V \in \mathscr{C}^{\kappa+3}(\mathbb{R})$ for some $\kappa \in \mathbb{N}$ and that the Hilbert transform (see the Appendix for further information and our conventions for the Hilbert transform) of the equilibrium measure satisfies

$$(4.3) \qquad \begin{cases} \mathcal{H}_x(\mu_V) = V'(x) & \text{if } x \in \bar{J}, \\ \mathcal{H}_x(\mu_V) < V'(x) & \text{if } x \in J_\delta \setminus \bar{J}. \end{cases}$$

The aim of this section is to investigate the solutions of the equation

$$(4.4) \qquad V'(x)\phi'(x) - \int_J \frac{\phi'(x) - \phi'(t)}{x - t}\mu_V(dt) = \sigma\phi(x),$$

where $\sigma > 0$ is an unknown of the problem. In particular, an important part of the analysis will be devoted to control the regularity of a solution ϕ. Let us consider the operator

$$(4.5) \qquad \Xi_x^{\mu_V}(\phi) = \mathrm{pv}\int_J \frac{\phi'(t)}{x - t}\mu_V(dt).$$

This operator is well defined on $\mathscr{H}$ and, since the Hilbert transform is bounded on $L^2(\mathbb{R})$ (again, see the Appendix), we have $\Xi^{\mu_V}(\phi) \in L^2(\mathbb{R})$. Moreover, using the variational condition (4.3), equation (4.4) reduces to the following for all $x \in J$,

$$(4.6) \qquad \Xi_x^{\mu_V}(\phi) = \sigma\phi(x).$$

Our main goal is to give a proof of Theorem 1.5. The proof is divided into several parts. In Section 4.1, we present the general setup and give basic facts that we will need in the remainder of the proof. In Sections 4.2 and 4.3, we analyze the regularity of a solution of equation (4.4) (or actually, we will find it convenient to invert the equation (see (4.16)) and consider the regularity of the solutions to the inverted equation). In Section 4.2, by a bootstrap argument, we show that a solution ϕ of equation (4.6) is of class $\mathscr{C}^{\kappa+1,1}(\bar{J})$. In particular, note that the $\kappa + 1$ first derivatives of the eigenfunctions have finite values at the edge-points ± 1. In Section 4.3, by viewing (4.4) as on ODE on $\mathbb{R} \setminus \bar{J}$, we show how to extend the eigenfunctions outside of the cut in such a way that that the eigenfunctions ϕ satisfy the following conditions: $\phi \in \mathscr{C}^\kappa(\mathbb{R})$ and ϕ solves equation (4.4) in a small neighborhood J_ϵ. We also control uniformly the sup norms on $\mathbb{R}$ of the derivatives $\phi^{(k)}$ for all $k < \kappa$.

As mentioned, we will actually study an inverted version of (4.4). In Sections 4.2 and 4.3, we establish the following result for the inverted equation.

PROPOSITION 4.1. *Let $\kappa \geq 1$ and $\eta > \frac{4(\kappa+1)}{2\kappa-1}$. Suppose that equation (4.16) below has a solution $\phi \in \mathscr{H}$ with $\sigma > 0$. Then we may extend this solution in such a way that $\phi \in \mathscr{C}^\kappa(\mathbb{R})$ with compact support and it satisfies equation (4.4) on the interval J_ϵ where $\epsilon = \sigma^{-\eta} \wedge \delta$ and for all $k < \kappa$,*

$$\|\phi^{(k)}\|_{\infty,\mathbb{R}} \ll \sigma^{\eta k}.$$

In Section 4.4, we relate equation (4.6) to a Hilbert–Schmidt operator $\mathcal{R}^S$ acting on $\mathscr{H}_{\mu_V}$ in order to prove the existence of the eigenvalues σ_n and eigenfunctions ϕ_n. Then, in Section 4.5, we put together the proof of Theorem 1.5 and get an estimate for the Fourier coefficients $\widehat{f_n}$ of the decomposition of a smooth function f on J in the eigenbasis $(\phi_n)_{n=1}^\infty$.

4.1. *Properties of the Sobolev spaces* $\mathcal{H}_{\mu V}$. We now review some basic properties of the spaces $\mathcal{H}_{\mu V}$ along with how suitable variants of the Hilbert transform act on these spaces. We will find it convenient to formulate many of the basic properties of the space $\mathcal{H}_{\mu V}$ in terms of Chebyshev polynomials. For definitions and basic properties, we refer to the Appendix. We begin by pointing out that elements of $\mathcal{H}_{\mu V}$ are actually Hölder continuous on $\overline{J}$ and in particular, bounded on $\overline{J}$. This follows from the simple remark that for $g \in \mathcal{H}$ and $x, y \in J$,

$$
\begin{aligned}
\left|g(x) - g(y)\right| &\leq 2\|g\|\sqrt{\int_x^y \varrho(dt)} \\
&= 2\|g\|\left|\arcsin(x) - \arcsin(y)\right|^{\frac{1}{2}} \\
&\ll \|g\|\,|x - y|^{\frac{1}{4}}.
\end{aligned}
$$

(4.7)

A basic fact about Chebyshev polynomials of the first kind, that we recall in the Appendix, is that after a simple normalization, they form an orthonormal basis for $L^2(J, \varrho)$, so in particular, we can also write for $g \in \mathcal{H}$,

$$
(4.8) \qquad g = \sum_{k \geq 1} g_k T_k,
$$

where g_k, $k \geq 1$, are the Fourier–Chebyshev's coefficients of $g \in \mathcal{H}$ (see (A.9)). On the other hand, since the Chebyshev polynomials of the second kind form an orthonormal basis for $L^2(\mu_{sc})$, the normalized Chebyshev polynomials of the first kind, $k^{-1}T_k$ form an orthonormal basis of $\mathcal{H}$. Combining these remarks, we see that

$$
(4.9) \qquad g_k = \frac{1}{k}\langle g, k^{-1}T_k\rangle = 2\int_J g(x)T_k(x)\varrho(dx).
$$

Thus by Parseval's formula (applied to the inner product of $\mathcal{H}$), we also obtain

$$
(4.10) \qquad \|g\|^2 = \sum_{k \geq 1} k^2 g_k^2 \ll 1.
$$

From this, one can check that for $g \in \mathcal{H}$, we may differentiate formula (4.8) term by term:

$$
(4.11) \qquad g' = \sum_{k \geq 1} k g_k U_{k-1},
$$

where the sum converges in $L^2(\mu_{sc})$. We define *the finite Hilbert transform* $\mathcal{U}$, for any ϕ for which

$$
\int_J |\phi(x)|^p (1 - x^2)^{-\frac{p}{2}}\, dx < \infty
$$

for some $p > 1$ by

$$\mathcal{U}_x(\phi) = \mathrm{pv} \int_{\mathrm{J}} \frac{\phi(y)}{y - x} \varrho(dy). \tag{4.12}$$

Note that, by boundedness of the Hilbert transform and the fact that elements of $\mathcal{H}$ are Hölder continuous on J so they are bounded, we see that for any $p < 2$, $\mathcal{U}$ maps $\mathcal{H}$ into $L^p(\mathbb{R})$. Moreover, it follows from Lemma A.3 that if $g \in \mathcal{H}$, then

$$\mathcal{U}(g) = \sum_{k \geq 1} g_k U_{k-1}, \tag{4.13}$$

where the sum converges in $L^2(\mu_{\mathrm{sc}})$ and uniformly on compact subsets of J. Indeed, it is well known that $|U_{k-1}(x)| \leq k \wedge (1 - x^2)^{-\frac{1}{2}}$ so that by (4.10) and the Cauchy–Schwarz inequality, for $x \in \mathrm{J}$,

$$|\mathcal{U}_x(g)| \ll (1 - |x|)^{-\frac{1}{2}} \sum_{k \geq 1} |g_k| \ll (1 - |x|)^{-\frac{1}{2}} \|g\|.$$

Note that from (4.12), one can check that $\mathcal{U}$ is actually a bounded operator from $\mathcal{H}$ to $L^2(\mu_{\mathrm{sc}})$ (and thus also a bounded operator from $\mathcal{H}_{\mu_V}$ to $L^2(\mu_{\mathrm{sc}})$).

Finally, let us note that for any $f, g \in \mathcal{H}$, by (4.11) and (4.13), we obtain

$$\int_{\mathrm{J}} f'(x) \mathcal{U}_x(g) \mu_{\mathrm{sc}}(dx) = \int_{\mathrm{J}} \mathcal{U}_x(f) g'(x) \mu_{\mathrm{sc}}(dx) = \sum_{k \geq 1} k g_k f_k \tag{4.14}$$

implying that according to (1.12), we can also write

$$\boldsymbol{\Sigma}(g) = \frac{1}{4} \int_{\mathrm{J}} g'(x) \mathcal{U}_x(g) \mu_{\mathrm{sc}}(dx). \tag{4.15}$$

4.2. *Regularity of the eigenfunctions in the cut* J. In this section, we begin the proof of Proposition 4.1. By Tricomi's inversion formula, Lemma A.4, if $\phi \in \mathcal{H}$ is a solution of the equation

$$\phi'(x) S(x) = \frac{\sigma}{2} \mathcal{U}_x(\phi) \tag{4.16}$$

for all $x \in \mathrm{J}$, since $d\mu_V = S \, d\mu_{\mathrm{sc}}$ and $\Xi^{\mu_V}(\phi) = \mathcal{H}(\phi' \mu_V)$, then ϕ also solves the equation (4.6). Hence, for now, we may focus on the properties of equation (4.16). In particular, by (4.7), we already know that if $\phi \in \mathcal{H}$, then ϕ is $\frac{1}{4}$-Hölder continuous on $\bar{\mathrm{J}}$ and we may use (4.16) to improve on the regularity of the eigenfunction ϕ by a bootstrap procedure.

PROPOSITION 4.2. *Suppose that equation (4.16) has a solution ϕ belonging to* $\in \mathscr{C}^{0,\alpha}(\bar{\mathrm{J}})$ *for some $\alpha > 0$. If $S \in \mathscr{C}^{\kappa+1}$ and $S > 0$ in a neighborhood of* J, *then* $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{\mathrm{J}})$.

The proof of Proposition 4.2 is given below and relies on certain regularity properties of the finite Hilbert transform $\mathcal{U}$; see Lemmas 4.3 and 4.4. If $f \in \mathscr{C}^k$ in a neighborhood of $x \in \mathbb{R}$, we denote its Taylor polynomial (a polynomial in t) by

$$\Lambda_k[f](x, t) = \sum_{i=0}^{k} f^{(i)}(x) \frac{(t - x)^i}{i!}.$$

Let ϕ be a solution of equation (4.16) and set for all $x, t \in \mathrm{J}$,

$$(4.17) \qquad\qquad F_k(x, t) = \frac{\phi(t) - \Lambda_k[\phi](x, t)}{(t - x)^{k+1}}$$

and

$$(4.18) \qquad\qquad \mathcal{U}_x^k(\phi) = \int_{\mathrm{J}} F_k(x, t) \varrho(dt).$$

In particular, by (A.4) and Lemma A.1, if $\phi \in \mathscr{C}^{k,\alpha}(\bar{\mathrm{J}})$ with $\alpha > 0$, then for almost all $x \in \mathrm{J}$,

$$(4.19) \qquad\qquad \frac{1}{k!} \left(\frac{d}{dx} \right)^k \mathcal{U}_x(\phi) = \mathcal{U}_x^k(\phi).$$

The following two technical results are our key ingredients in the proof of Proposition 4.2, and we will present their proofs once we have proven Proposition 4.2.

LEMMA 4.3.   *Let, for all $x \in \mathrm{J}$,*

$$(4.20) \qquad \mathrm{K}_\alpha(x) = \begin{cases} (1 - x)^{\alpha - \frac{1}{2}} + (1 + x)^{\alpha - \frac{1}{2}} & \text{if } \alpha \neq \frac{1}{2}, \\ 1 + \log\left(\dfrac{1}{1 - x^2} \right) & \text{if } \alpha = \dfrac{1}{2}. \end{cases}$$

For any $k \geq 0$, if $f \in \mathscr{C}^{k,\alpha}(\bar{\mathrm{J}})$ with $0 < \alpha < 1$, then for all $x \in \mathrm{J}$,

$$(4.21) \qquad\qquad \left| \mathcal{U}_x^k(f) \right| \ll_{\alpha, f} \mathrm{K}_\alpha(x).$$

LEMMA 4.4.   *For any $k \geq 0$, if $f \in \mathscr{C}^{k,1}(\bar{\mathrm{J}})$, then $\mathcal{U}^k(f) \in \mathscr{C}^{0,\alpha}(\bar{\mathrm{J}})$ for any $\alpha \leq \frac{1}{2}$.*

PROOF OF PROPOSITION 4.2.   Without loss of generality, we may assume that $\sigma = 2$ in (4.16). Let $0 \leq k \leq \kappa$ and suppose that $\phi \in \mathscr{C}^{k,\alpha}(\bar{\mathrm{J}})$ where $\alpha > 0$ and that for almost all $x \in \mathrm{J}$,

$$(4.22) \qquad \mathcal{U}_x^k(\phi) - S(x)\phi^{(k+1)}(x) = \sum_{j=1}^{k} \binom{k}{j} S^{(j)}(x)\phi^{(k+1-j)}(x).$$

Note that by our assumptions, these conditions are satisfied when $k = 0$ in which case the RHS equals zero. Since $\phi \in \mathscr{C}^{k,\alpha}(\bar{J})$, the RHS of (4.22) is uniformly bounded, so by Lemma 4.3 and since $S > 0$ on $\bar{J}$, we see that for all $x \in J$,

$$\left|\phi^{(k+1)}(x)\right| \ll_\alpha K_\alpha(x).$$

By the definition of K_α, this bound implies that $\phi^{(k)} \in \mathscr{C}^{0,\alpha'}(\bar{J})$ where $\alpha' = (\alpha + \frac{1}{2}) \wedge 1$. By repeating this argument, we obtain that $\phi \in \mathscr{C}^{k,1}(\bar{J})$ and, by Lemma 4.4, this implies that $\phi^{(k+1)} \in \mathscr{C}^{0,\alpha}(\bar{J})$. Then (4.19) shows that we can differentiate formula (4.22). In particular, this shows that $\phi^{(k+2)}$ exists and that for almost all $x \in J$,

$$S(x)\phi^{(k+2)}(x) = \mathcal{U}_x^{k+1}(\phi) - \sum_{j=1}^{k+1} \binom{k+1}{j} S^{(j)}(x)\phi^{(k+2-j)}(x).$$

Thus, we obtain (4.22) with k replaced by $k+1$. Therefore, we can repeat this argument until $k = \kappa + 1$, in which case we have seen that the solution $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{J})$. Proposition 4.2 is established. $\quad\square$

The remainder of this subsection is devoted to the proof of Lemmas 4.3 and 4.4 as well as bounding norms of derivatives of solutions of (4.16) in terms of σ.

PROOF OF LEMMA 4.3. Let $\phi \in \mathscr{C}^{k,\alpha}(\bar{J})$. We claim that for any $x, y \in J$,

$$\left|\phi(t) - \Lambda_k[\phi](x,t)\right| \ll_\phi |t - x|^{k+\alpha}.$$

Indeed, this is just the definition of the class $\mathscr{C}^{0,\alpha}(\bar{J})$ when $k = 0$. Moreover, if $k \geq 1$ it can be checked that

$$\phi(t) - \Lambda_k[\phi](x,t) = (t - x)\int_0^1 \left[\phi'(x + (t - x)u) - \Lambda_{k-1}[\phi'](x, x + (t - x)u)\right] du$$

so that the estimate follows directly by induction. According to (4.17)–(4.18), this implies that

$$\tag{4.23} \left|F_k(x,t)\right| \ll_\phi |t - x|^{\alpha-1}$$

and, by splitting the integral,

$$\tag{4.24} \left|\mathcal{U}_x^k(f)\right| \ll_\phi \int_{-1}^x \frac{(x - t)^{\alpha-1}}{\sqrt{1 - t^2}}\, dt + \int_x^1 \frac{(t - x)^{\alpha-1}}{\sqrt{1 - t^2}}\, dt.$$

The RHS of (4.24) is finite for all $x \in J$ and, by symmetry, we may assume that $x \geq 0$. Then we have

$$\int_x^1 \frac{(t - x)^{\alpha-1}}{\sqrt{1 - t^2}}\, dt \leq \int_x^1 \frac{(t - x)^{\alpha-1}}{\sqrt{1 - t}}\, dt$$

$$\tag{4.25} = (1 - x)^{\alpha-\frac{1}{2}} \int_0^1 u^{\alpha-1}(1 - u)^{-\frac{1}{2}}\, dt$$

$$\ll_\alpha K_\alpha(x).$$

On the other hand, we have

$$\int_{-1}^{x} \frac{(x-t)^{\alpha-1}}{\sqrt{1-t^2}} \, dt \le \int_{0}^{1} \frac{t^{\alpha-1}}{\sqrt{1-t^2}} \, dt + \int_{0}^{x} \frac{(x-t)^{\alpha-1}}{\sqrt{1-t}} \, dt.$$

The first integral does not depend on $x \ge 0$, and by a change of variable

$$\int_{-1}^{x} \frac{(x-t)^{\alpha-1}}{\sqrt{1-t^2}} \, dt \le C + (1-x)^{\alpha-\frac{1}{2}} \int_{0}^{\frac{x}{1-x}} u^{\alpha-1}(1+u)^{-\frac{1}{2}} \, du.$$

It is easy to verify that

$$\int_{0}^{\frac{x}{1-x}} u^{\alpha-1}(1+u)^{-\frac{1}{2}} \, du = \underset{x \to 1}{O} \begin{pmatrix} 1 & \text{if } \alpha < \frac{1}{2} \\ \log\left(\frac{1}{1-x}\right) & \text{if } \alpha = \frac{1}{2} \\ (1-x)^{\frac{1}{2}-\alpha} & \text{if } \alpha > \frac{1}{2} \end{pmatrix}$$

which shows that

(4.26) $$\int_{-1}^{x} \frac{(x-t)^{\alpha-1}}{\sqrt{1-t^2}} \, dt \ll_{\alpha} \mathrm{K}_{\alpha}(x).$$

Combining the estimates (4.24)–(4.26), we obtain (4.21). $\square$

To complete the proof of Proposition 4.2, we conclude with the proof of Lemma 4.4.

PROOF OF LEMMA 4.4. We claim that, if $\phi \in \mathscr{C}^{k,1}(\bar{\mathrm{J}})$, then the function $x \mapsto F_{k-1}(x,t)$ is $\mathscr{C}^{0,1}(\bar{\mathrm{J}})$. Indeed, by Taylor's formula, we have

(4.27) $$F_{k-1}(x,t) = \int_{0}^{1} \phi^{(k)}(t(1-u) + xu) \frac{u^{k-1}}{(k-1)!} \, du,$$

so that

$$\left| F_{k-1}(t,y) - F_{k-1}(t,x) \right|$$

(4.28)
$$\le \int_{0}^{1} \left| \phi^{(k)}(t(1-u) + yu) - \phi^{(k)}(t(1-u) + xu) \right| \frac{u^{k-1}}{(k-1)!} \, du$$

$$\ll |y-x|,$$

where the underlying constant is independent of the parameter $t \in \mathrm{J}$. Since

$$F_k(x,t) = \frac{F_{k-1}(x,t) - \frac{1}{k!}\phi^{(k)}(x)}{t-x},$$

we have that

$$
F_k(y, t) - F_k(x, t)
$$
$$
= \left(F_{k-1}(y, t) - \frac{1}{k!} \phi^{(k)}(y) - F_{k-1}(x, t) \right.
$$
$$
\left. + \frac{1}{k!} \phi^{(k)}(x) + (y - x) F_k(x, t) \right) \frac{1}{t - y}.
$$

Then, using the estimates (4.23) with $\alpha = 1$ and (4.28), we obtain that uniformly for all $x, t \in \bar{J}$,

$$
\left| F_k(t, y) - F_k(t, x) \right| \ll 1 \wedge \left(\frac{|y - x|}{|t - y| \vee |t - x|} \right).
$$

This implies that for all $-1 < x < y < 1$,

$$
(4.29) \quad
\begin{aligned}
\left| \mathcal{U}_x^k(\phi) - \mathcal{U}_y^k(\phi) \right| &\ll (y - x) \left(\int_{t<x} \frac{\varrho(dt)}{y - t} + \int_{y<t} \frac{\varrho(dt)}{t - x} \right) + \int_x^y \varrho(dt) \\
&\ll \sqrt{y - x} (1 + \theta(x, y)),
\end{aligned}
$$

where

$$
(4.30) \quad
\begin{aligned}
\theta(x, y) &= \sqrt{y - x} \left(\int_{t<x} \frac{\varrho(dt)}{y - t} + \int_{y<t} \frac{\varrho(dt)}{t - x} \right) \\
&= \sqrt{y - x} \frac{\sqrt{1 - y^2} + \sqrt{1 - x^2}}{\sqrt{1 - y^2} \sqrt{1 - x^2}} \\
&\quad \times \log \left(\frac{1 - xy + \sqrt{1 - x^2} \sqrt{1 - y^2}}{y - x} \right).
\end{aligned}
$$

Formula (4.30) follows by explicitly evaluating the integrals. This function extends by continuity to the domain $\mathcal{D} = \{-1 < x \le y < 1\}$ with $\theta(x, x) = 0$. By (4.29), it just remains to prove that the function θ is bounded on $\overline{\mathcal{D}}$. First, it is easy to check that

$$
\lim_{y \to 1} \theta(x, y) = \sqrt{1 + x} \quad \text{and} \quad \lim_{x \to -1} \theta(x, y) = \sqrt{1 - y}.
$$

To study the boundary values of θ around the point $(1, 1)$, we make the change of variables

$$
\widetilde{\theta}(\epsilon, \tau) = \theta \left(\sqrt{1 - \tau^2 \epsilon^2}, \sqrt{1 - \epsilon^2} \right),
$$

where $0 < \epsilon < 1$ and $1 < \tau < \frac{1}{\epsilon}$. Then $y = \sqrt{1 - \epsilon^2} \underset{\epsilon \to 0}{\simeq} 1 - \frac{\epsilon^2}{2}$ and $x = \sqrt{1 - \tau^2 \epsilon^2} \underset{\epsilon \to 0}{\simeq} 1 - \frac{\tau^2 \epsilon^2}{2}$ so that

$$
\lim_{\epsilon \to 0} \widetilde{\theta}(\epsilon, \tau) = g(\tau) = \sqrt{\frac{\tau^2 - 1}{2}} \frac{\tau + 1}{\tau} \log \left(\frac{\tau + 1}{\tau - 1} \right).
$$

The function $g(\tau)$ is continuous on the interval $(1, \infty)$ with boundary values $\lim_{\tau \to 1} g(\tau) = 0$ and $\lim_{\tau \to \infty} g(\tau) = \sqrt{2}$. In particular, this shows that

$$\limsup_{x \to 1, y \to 1} \theta(x, y) = \max_{\tau \geq 1} g(\tau) < \infty.$$

A similar analysis in a neighborhood of the point $(-1, -1)$ allows us to conclude that the function θ is bounded on $\overline{\mathcal{D}}$. Lemma 4.4 is established. $\quad\square$

We are now in a position to provide some quantitative bounds on solutions to (4.16). Our implicit assumption is that σ is not small—an assumption that will be fulfilled in our applications of the results we prove.

PROPOSITION 4.5. *Suppose that equation* (4.16) *has a solution* $\phi \in \mathcal{H}$ *such that* $\|\phi\|_{\infty,\mathrm{J}} \ll 1$ (*this bound does not depend on the parameter* $\sigma > 0$). *Then we have, for any* $0 < \alpha < \frac{1}{2}$ *and for all* $k \leq \kappa + 1$,

$$(4.31) \qquad \|\phi^{(k)}\|_{\infty,\mathrm{J}} \ll_\alpha \sigma^{\frac{k}{\alpha}}.$$

PROOF. By (4.7), we already know that $\phi \in \mathscr{C}^{0,\frac{1}{4}}(\bar{\mathrm{J}})$. Then, by Proposition 4.2, we obtain that $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{\mathrm{J}})$. By our assumptions, we note that the bound (4.31) holds when $k = 0$ and we proceed by induction on $k \leq \kappa$ to prove the general case. If we differentiate k times equation (4.16), we obtain

$$(4.32) \qquad S(x)\phi^{(k+1)}(x) = \frac{\sigma}{2}\mathcal{U}_x^k(\phi) - \sum_{j=1}^{k} \binom{k}{j-1} S^{(k+1-j)}(x)\phi^{(j)}(x).$$

By Lemma 4.4, the RHS of (4.32) is continuous for all $x \in \bar{\mathrm{J}}$ and, since $S > 0$ on $\bar{\mathrm{J}}$, this shows that

$$(4.33) \qquad \|\phi^{(k+1)}\|_{\infty,\mathrm{J}} \ll \sigma \|\mathcal{U}^k(\phi)\|_{\infty,\mathrm{J}} + \sigma^{\frac{k}{\alpha}}.$$

Here, we assumed that the estimate (4.32) is valid for all $j \leq k$ and that the parameter σ is large. Recall that, according to (4.27), we have the bound $\sup_{x,t \in \mathrm{J}} |F_{k-1}(t, x)| \leq \|\phi^{(k)}\|_{\infty,\mathrm{J}}$. Since

$$F_k(x, t) = \frac{F_{k-1}(x, t) - \frac{1}{k!}\phi^{(k)}(x)}{t - x},$$

this shows that for any $\alpha \in [0, 1]$ and for $x, t \in \bar{\mathrm{J}}$,

$$|F_k(t, x)| \ll \frac{\|\phi^{(k+1)}\|_{\infty,\mathrm{J}}^{1-\alpha}\|\phi^{(k)}\|_{\infty,\mathrm{J}}^{\alpha}}{|x - t|^\alpha} \ll \sigma^k \frac{\|\phi^{(k+1)}\|_{\infty,\mathrm{J}}^{1-\alpha}}{|x - t|^\alpha},$$

where we used the estimate (4.31) once more. By (4.18), this shows that for any $\alpha < \frac{1}{2}$,

$$\|\mathcal{U}^k(\phi)\|_{\infty,\mathrm{J}} \ll \sigma^k \|\phi^{(k+1)}\|_{\infty,\mathrm{J}}^{1-\alpha}.$$

Combining this bound with the estimate (4.33), we conclude that

$$\|\phi^{(k+1)}\|_{\infty,\mathrm{J}} \ll \sigma^{k+1}\|\phi^{(k+1)}\|_{\infty,\mathrm{J}}^{1-\alpha} + \sigma^{\frac{k}{\alpha}}.$$

Then either $\|\phi^{(k+1)}\|_{\infty,\mathrm{J}} \ll \sigma^{\frac{k}{\alpha}}$ or $\|\phi^{(k+1)}\|_{\infty,\mathrm{J}}^{\alpha} \ll \sigma^{k+1}$, which completes the proof. $\square$

4.3. *Extension of the eigenfunctions on* $\mathbb{R} \setminus \bar{\mathrm{J}}$. The aim of this subsection is to complete the proof of Proposition 4.1 by extending the eigenfunction ϕ outside of the cut in such a way that the resulting function, still denoted by ϕ, satisfies equation (4.4) in a neighborhood J_ϵ of the cut and that $\phi \in \mathscr{C}^\kappa(\mathbb{R})$. We will focus on the extension of ϕ in a neighborhood of the edge-point 1, the construction being completely analogous in a neighborhood of -1. Suppose that $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{\mathrm{J}})$ and let for all $x \in \mathbb{R}$,

$$(4.34) \qquad g(x) = \mathrm{pv} \int \frac{\phi'(t)}{x - t}\mu_V(dt).$$

Note that g is smooth on $\mathbb{R} \setminus \bar{\mathrm{J}}$ and that it does not depend on the values of ϕ outside of cut. Thus, we may interpret equation (4.4) as an ODE satisfied by the function $\mathrm{y} = \phi_{|\mathbb{R}\setminus\bar{\mathrm{J}}}$:

$$(4.35) \qquad Q'\mathrm{y}' = \sigma\mathrm{y} - g,$$

where Q is as in (1.7). The condition (4.3) guarantees that $Q' > 0$ and equation (4.35) is well-posed for all $x \in \mathrm{J}_\delta \setminus \bar{\mathrm{J}}$. Moreover, by Proposition A.6, since $\frac{1}{Q'}$ is integrable in a neighborhood of 1, equation (4.35) has a unique solution such that $\mathrm{y}(1) = \varkappa_0$ for any $\varkappa_0 \in \mathbb{R}$. Choosing $\varkappa_0 = \phi(1)$, in order to extend outside of the cut the solution ϕ of equation (4.6) in a smooth way, we need to check that y satisfies for all $1 \leq k \leq \kappa$,

$$(4.36) \qquad \lim_{x \to 1^+} \mathrm{y}^{(k)}(x) = \lim_{x \to 1^-} \phi^{(k)}(x).$$

The conditions (4.36) are checked using the following result and Proposition 4.7 below. Again, in our quantitative regularity bounds in terms of σ, we will be assuming that σ is large.

PROPOSITION 4.6. *Let* $\kappa \geq 1$ *and suppose that* $\eta > \frac{4(\kappa+1)}{2\kappa-1}$. *Let* G *in* $\mathscr{C}^{\kappa+1}(1, \infty)$ *be such that* $G(x) = o_{x\to1^+}(x - 1)^\kappa$ *and let* Y *be the solution of the equation*

$$(4.37) \qquad Q'Y' = \sigma Y + G$$

on the interval $(1, \infty)$ *with boundary condition* $Y(1) = 0$. *Then* Y *belongs to* $\mathscr{C}^{\kappa+2}(1, 1 + \delta)$ *and satisfies for all* $k \leq \kappa$,

$$Y^{(k)}(1) = \lim_{x \to 1^+} Y^{(k)}(x) = 0.$$

Moreover, letting $\epsilon = \sigma^{-\eta} \wedge \delta$, if we assume that for any $\frac{1}{\eta} < \alpha < \frac{1}{2}$,

$$\tag{4.38} \left\| G^{(\kappa-1)} \right\|_{\infty,[1,1+\epsilon]} \ll_\alpha \sigma^{\frac{\kappa+1}{\alpha}},$$

then we have for all $k < \kappa$,

$$\tag{4.39} \left\| Y^{(k)} \right\|_{\infty,[1,1+\epsilon]} \ll \sigma^{\eta k}.$$

PROOF. If we let, for all $1 \leq x < 1 + \delta$,

$$\tag{4.40} \mathrm{h}(x) = \int_1^x \frac{\sigma}{Q'(t)} \, dt \quad \text{and} \quad \Upsilon(x) = \int_1^x \frac{G(t)}{Q'(t)} e^{-\mathrm{h}(x)} \, dt,$$

then the solution of equation (4.37) for which $Y(1) = 0$ is given by

$$\tag{4.41} Y(x) = \Upsilon(x) e^{\mathrm{h}(x)}.$$

Then, if $G \in \mathscr{C}^{\kappa+1}(1, \infty)$, we immediately check that $Y \in \mathscr{C}^{\kappa+2}(1, 1+2\delta)$. Moreover, by Proposition A.6, the condition $G(x) = O_{x \to 1^+}(x-1)^\kappa$ implies that

$$\tag{4.42} |\Upsilon(x)| \ll (x-1)^{\kappa+\frac{1}{2}}.$$

Since the function $\mathrm{h}(x)$ is continuous on the interval $[1, 1+\delta]$ with $\mathrm{h}(0) = 0$, we conclude that the estimate (4.42) is also satisfied by the solution Y which proves that it is of class $\mathscr{C}^\kappa$ at 1 and that $Y^{(k)}(1) = 0$ for all $k \leq \kappa$.

Now we turn to the proof of (4.39). By Taylor's theorem, since $G^{(k)}(1) = 0$ for all $k \leq \kappa$,

$$G^{(j)}(t) = (t-1)^{\kappa-1-j} \int_0^1 G^{(\kappa-1)}(1 - u(1-t)) \frac{u^{\kappa-2-j}}{(\kappa-2-j)!} \, du,$$

and the estimate (4.38) implies that for all $j \leq \kappa - 1$ and for all $1 < x < \epsilon$,

$$\left| G^{(j)}(t) \right| \ll \sigma^{\frac{\kappa+1}{\alpha}} (t-1)^{\kappa-1-j}.$$

On the other hand, by Proposition A.6, for all $j \leq \kappa$,

$$\frac{d^j}{dx^j} \left(\frac{1}{Q'(x)} \right) = \mathop{O}_{x \to 1^+} (x-1)^{-\frac{1}{2}-j},$$

so that we deduce from (4.40) that for any $k < \kappa$ and for all $1 \leq x < 1 + \delta$,

$$\tag{4.43} \begin{aligned} \mathrm{h}^{(k)}(x) &\ll \sigma (x-1)^{\frac{1}{2}-k} \\ \left| \Upsilon^{(k)}(x) \right| &\ll \sigma^{(\kappa+1)/\alpha} (x-1)^{\kappa-\frac{1}{2}-k}. \end{aligned}$$

Combined with (4.41), the estimates (4.43) imply that if $\epsilon = \sigma^{-\eta} \wedge \delta$, we have for all $k < \kappa$,

$$\left\| Y^{(k)} \right\|_{\infty,[1,1+\epsilon]} \ll \sigma^{(\kappa+1)/\alpha - \eta(\kappa-\frac{1}{2}-k)}.$$

Now it is easy to check that if $\eta > \frac{4(\kappa+1)}{2\kappa-1}$, we can choose $\frac{1}{\eta} < \alpha < \frac{1}{2}$ so that $\frac{\kappa+1}{\alpha} \leq \eta(\kappa - \frac{1}{2})$. Hence, we obtain the estimate (4.39), completing the proof of the proposition. $\square$

The second ingredient of our extension of the eigenfunctions is the following result.

PROPOSITION 4.7. *Suppose that* $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{\mathrm{J}})$ *is a solution of equation* (4.6) *on* J. *For all* $k \leq \kappa + 1$, *define* $\varkappa_k = \lim_{x\to 1^-} \phi^{(k)}(x)$, *and let* $\Omega(x) = \sum_{i=0}^{\kappa+1} \varkappa_i \frac{(x-1)^i}{i!}$. *Then the function* g *given by* (4.34) *satisfies*

$$(4.44) \qquad g(x) = -\Omega'(x) Q'(x) + \sigma \Omega(x) + \underset{x\to 1}{o}(x-1)^\kappa.$$

Moreover, if for any $0 < \alpha < \frac{1}{2}$, *we have* $\|\phi^{(k)}\|_{\infty,\mathrm{J}} \ll_\alpha \sigma^{\frac{k}{\alpha}}$ *for all* $k \leq \kappa + 1$, *then the function* $G(x) = g(x) + Q'(x)\Omega'(x) - \sigma\Omega(x)$ *satisfies* (4.38) (*and we can apply the previous proposition*).

PROOF. On the one hand, according to (4.6) and (4.34), we have $g(x) = \sigma\phi(x)$ for all $x \in$ J. This shows that $g \in \mathscr{C}^{\kappa+1}(\bar{\mathrm{J}})$ and

$$(4.45) \qquad g(x) = \sigma\Omega(x) + \underset{x\to 1^-}{O}(x-1)^{\kappa+1}.$$

From the variational condition (4.3), $Q' = V' - \mathcal{H}(\mu_V) = 0$ on $\bar{\mathrm{J}}$, and we see that (4.44) holds when the limit is taken from the inside of the cut. On the other hand, we have for all $x \in$ J,

$$(4.46) \quad g(x) = V'(x)\phi'(x) - \Theta(x) \qquad \text{where } \Theta(x) = \int_{\mathrm{J}} \frac{\phi'(t) - \phi'(x)}{t - x} \mu_V(dt).$$

By Lemma A.1, since $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{\mathrm{J}})$, the function $\Theta \in \mathscr{C}^\kappa(\bar{\mathrm{J}})$ and for all $k \leq \kappa$,

$$(4.47) \qquad \Theta^{(k)}(1) = \lim_{x\to 1^-} \Theta^{(k)}(x) = k! \int \frac{\phi'(t) - \Lambda_k[\phi'](1,t)}{(t-1)^{k+1}} \mu_V(dt).$$

Hence, by (4.45) and (4.46), we obtain that

$$(4.48) \qquad \Theta^{(k)}(1) = \left.\frac{d^k(V'\phi')}{dx^k}\right|_{x=1^-} - \sigma\varkappa_k.$$

Define for all $x \geq 1$,

$$\widetilde{\Theta}(x) = \int \frac{\phi'(t) - \Omega'(x)}{t - x} \mu_V(dt),$$

so that

$$(4.49) \qquad g(x) = \mathcal{H}_x(\mu_V)\Omega'(x) - \widetilde{\Theta}(x).$$

It is also easy to check that the function $\widetilde{\Theta} \in \mathscr{C}^\infty(1, \infty)$ and that for all $k \geq 1$ and $x > 1$,

$$(4.50) \qquad \widetilde{\Theta}^{(k)}(x) = k! \int \frac{\phi'(t) - \Lambda_k[\Omega'](x, t)}{(t - x)^{k+1}} \mu_V(dt).$$

Since $\Lambda_k[\Omega'](1, t) = \Lambda_k[\phi'](1, t)$, by (4.47), we obtain for all $k \leq \kappa$,

$$(4.51) \qquad \lim_{x \to 1^+} \widetilde{\Theta}^{(k)}(x) = \Theta^{(k)}(1).$$

According to (4.49), this shows that for all $x \geq 1$,

$$(4.52) \qquad \begin{aligned} G(x) &= g(x) + Q'(x)\Omega'(x) - \sigma\Omega(x) \\ &= -\widetilde{\Theta}(x) + V'(x)\Omega'(x) - \sigma\Omega(x). \end{aligned}$$

By (4.48) and (4.51), since $\Theta^{(k)}(1) = \frac{d^k(V'\Omega')}{dx^k}|_{x=1} - \sigma \varkappa_k$, we see that there is a cancellation on the RHS of (4.52) so that $\lim_{x \to 1^+} G^{(k)}(x) = 0$ for all $k \leq \kappa$. This implies that $G(x) = o_{x \to 1^+}(x - 1)^\kappa$ and we obtain the expansion (4.44).

Now we turn to the proof of the estimate (4.38). We let $\epsilon = \sigma^{-\eta}$ where $\eta > \frac{1}{\alpha}$ and the parameter σ is assumed to be large. By hypothesis, $|\varkappa_k| \ll_\alpha \sigma^{\frac{k}{\alpha}}$ and it is easy to verify that for all $k \leq \kappa + 1$,

$$(4.53) \qquad \|\Omega^{(k)}\|_{\infty, [1, 1+\epsilon]} \ll \sigma^{\frac{k}{\alpha}}.$$

Thus, by (4.52), it suffices to show that

$$(4.54) \qquad \|\widetilde{\Theta}^{(\kappa-1)}\|_{\infty, [1, 1+\epsilon]} \ll_\alpha \sigma^{\frac{\kappa+1}{\alpha}}.$$

Since $\Lambda_{\kappa-1}[\Omega'](1, t) = \Lambda_{\kappa-1}[\phi'](1, t)$, by (4.50), we obtain for all $x \geq 1$,

$$(4.55) \qquad \begin{aligned} &\left|\widetilde{\Theta}^{(\kappa-1)}(x)\right| \\ &\ll \sup_{t \in J} \left\{ \left| \frac{\phi'(t) - \Lambda_{\kappa-1}[\phi'](1, t)}{(1 - t)^\kappa} \right| \right. \\ &\qquad \left. + \left| \frac{\Lambda_{\kappa-1}[\Omega'](x, t) - \Lambda_{\kappa-1}[\Omega'](1, t)}{(x - t)^\kappa} \right| \right\}. \end{aligned}$$

On the one hand, by Taylor's theorem, we have

$$\sup_{t \in J} \left| \frac{\phi'(t) - \Lambda_{\kappa-1}[\phi'](1, t)}{(1 - t)^\kappa} \right| \leq \|\phi^{(\kappa+1)}\|_{\infty, J} \ll \sigma^{\frac{\kappa+1}{\alpha}}.$$

On the other hand, if f is a smooth function, for all $k \geq 0$,

$$\frac{d\Lambda_k[f](x, t)}{dx} = \frac{f^{(k+1)}(x)}{k!}(x - t)^{k+1},$$

and by the mean-value theorem, there exists $t < \xi < x$ such that

$$\frac{\Lambda_k[f](x,t) - \Lambda_k[f](1,t)}{(t-x)^{k+1}} = \frac{f^{(k+1)}(\xi)}{k!}\left(\frac{\xi - t}{t - x}\right)^{k+1}.$$

Therefore, we have for all $x > 1$,

$$\left|\frac{\Lambda_{\kappa-1}[\Omega'](x,t) - \Lambda_{\kappa-1}[\Omega'](1,t)}{(t-x)^\kappa}\right| \ll \|\Omega^{(\kappa+1)}\|_\infty = |\varkappa_{\kappa+1}|,$$

as we defined, $\Omega(x) = \sum_{i=0}^{\kappa+1} \varkappa_i \frac{(x-1)^i}{i!}$. Since $|\varkappa_{\kappa+1}| \ll \sigma^{\frac{\kappa+1}{\alpha}}$, we deduce from the estimate (4.55) that $\|\widetilde{\Theta}^{(\kappa-1)}\|_{\infty,[1,\infty)} \ll_\alpha \sigma^{\frac{\kappa+1}{\alpha}}$ which obviously implies (4.54) and completes the proof. $\square$

We are now ready for the proof of Proposition 4.1.

PROOF OF PROPOSITION 4.1. First, if equation (4.16) has a solution $\phi \in \mathscr{H}$ with $\sigma > 0$, by Propositions 4.2 and 4.5, this solution $\phi \in \mathscr{C}^{\kappa+1,1}(\bar{J})$ and satisfies, for all $k \leq \kappa + 1$,

$$(4.56) \qquad \|\phi^{(k)}\|_{\infty,J} \ll_\alpha \sigma^{\frac{k}{\alpha}}.$$

Then, let us write as in Proposition 4.7: $\varkappa_i = \lim_{x \to 1^-} \phi^{(i)}(x)$ and $\Omega(x) = \sum_{i=0}^{\kappa+1} \varkappa_i \frac{(x-1)^i}{i!}$, so that making the change of variable $y(x) = Y(x) - \Omega(x)$ in (4.35), we obtain

$$Q'Y' = \sigma Y + G,$$

where as before, $G(x) = g(x) + Q'(x)\Omega'(x) - \sigma\Omega(x)$. By Tricomi's formula, Lemma A.4, the function ϕ also solves equation (4.6) and according to Proposition 4.7, $G(x) = o_{x\to 1^+}(x-1)^\kappa$ and we deduce from Proposition 4.6 that for all $k \leq \kappa$,

$$\lim_{x\to 1^+} y^{(k)}(1) = \lim_{x\to 1} \Omega^{(k)}(x) = \varkappa_k.$$

This shows that the conditions (4.36) are satisfied, and that we may extend the solution ϕ outside of J by setting $\phi(x) = y(x)$ for all $x \in J_\epsilon \setminus J$ in such a way that it satisfies (4.4) on J_ϵ for some $\epsilon \leq \delta$ and $\phi \in \mathscr{C}^\kappa(\mathbb{R})$. Moreover, the estimate (4.56) and Proposition 4.7 imply that the function G satisfies (4.38) so that $\|Y^{(k)}\|_{\infty,[1,1+\epsilon]} \ll \sigma^{\eta k}$ for all $k < \kappa$, where $\epsilon = \sigma^{-\eta} \wedge \delta$ and $\eta > \frac{4(\kappa+1)}{2\kappa-1}$. Then, by choosing the parameter $\alpha > \frac{1}{\eta}$ in the estimates (4.53) and (4.56), we obtain that $\|\phi^{(k)}\|_{\infty,[-1,1+\epsilon]} \ll \sigma^{\eta k}$. Moreover, an analogous analysis in the neighborhood of the other edge-point -1 allows us to conclude that for all $k < \kappa$,

$$(4.57) \qquad \|\phi^{(k)}\|_{\infty,J_\epsilon} \ll \sigma^{\eta k}.$$

Now let $\chi : \mathbb{R} \to [0, 1]$ be a smooth even function with compact support in $J = (-1, 1)$ such that $\chi(0) = 1$ and $\chi^k(0) = 0$ for all $1 < k \leq \kappa$, and recall that $t \mapsto \Lambda_\kappa[\phi](x, t)$ denotes the Taylor polynomial of the function ϕ of degree κ at x. Then, if we set for all $t \in \mathbb{R} \setminus \overline{J_\epsilon}$,

$$\phi(t) = \chi\left(\frac{t - 1 - \epsilon}{\epsilon}\right) \Lambda_\kappa[\phi](1 + \epsilon, t) + \chi\left(\frac{t + 1 + \epsilon}{\epsilon}\right) \Lambda_\kappa[\phi](-1 - \epsilon, t),$$

then we indeed have $\phi \in \mathscr{C}^\kappa(\mathbb{R})$ and it is easy to check that the estimate (4.57) remains true for the norms $\|\phi^{(k)}\|_{\infty, \mathbb{R}}$ and for all $k < \kappa$. Finally, note that this extension procedure guarantees that the function ϕ has support in $J_{2\epsilon}$. Proposition 4.1 is proved. $\square$

4.4. *Existence of the eigenfunctions.* In order to prove that there exists a sequence of eigenfunctions $\phi_n \in \mathscr{H}$, equation (4.16) suggests to consider the operator $\mathcal{R}^S(\phi) = \psi$ where ψ is the (weak) solution in $\mathscr{H}$ of equation

$$(4.58) \qquad\qquad \psi' S = \mathcal{U}(\phi).$$

Note that if $\phi \in \mathscr{H}$, since $\mathcal{U}(\phi)$ is continuous on J, there exists a solution and the condition $\int_J \psi(x) \varrho(dx) = 0$ guarantees that it is unique.

THEOREM 4.8. *The operator $\mathcal{R}^S$ is compact, positive, and self-adjoint on $\mathscr{H}_{\mu_V}$. If we denote by $(\frac{2}{\sigma_n})_{n=1}^\infty$ the eigenvalues of $\mathcal{R}^S$ in nonincreasing order, then $\sigma_n \asymp n$ and the corresponding (normalized) eigenfunctions ϕ_n satisfy $\|\phi_n\|_{\infty, J} \ll 1$ for all $n \geq 1$.*

PROOF. Since the finite Hilbert transform $\mathcal{U} : \mathscr{H} \mapsto L^2(\mu_{\text{sc}})$ is bounded (as we remarked in Section 4.1) and $S > 0$ on $\bar{J}$, by (4.58) we obtain

$$(4.59) \qquad \|\mathcal{R}^S(\phi)\|_{\mu_V}^2 = \int_J \mathcal{U}_x(\phi)^2 \frac{2\sqrt{1 - x^2}}{\pi S(x)} \, dx \ll \|\mathcal{U}(\phi)\|_{L^2(\mu_{\text{sc}})}^2,$$

so that $\mathcal{R}^S : \mathscr{H}_{\mu_V} \mapsto \mathscr{H}_{\mu_V}$ is a bounded operator. Moreover, by (4.14), for any $\phi, g \in \mathscr{H}$,

$$\langle \mathcal{R}^S(\phi), g \rangle_{\mu_V} = \int_J \mathcal{U}_x(\phi) g'(x) \, d\mu_{\text{sc}}(x) = \sum_{k \geq 1} k \phi_k g_k.$$

This shows that $\mathcal{R}^S$ is self-adjoint and positive-definite with

$$(4.60) \qquad\qquad \langle \mathcal{R}^S(\phi), \phi \rangle_{\mu_V} = 4\Sigma(\phi).$$

To prove compactness, we rely on the results of [30], Section VI.5. Introduce the operator $\mathcal{R}^{S,N} : \phi \mapsto \psi_N$ where ψ_N solves the equation $\psi_N' S = \mathcal{U}^N(\phi)$ and

$\mathcal{U}^N(\phi) = \sum_{k=1}^{N} \phi_k U_{k-1}$. Plainly, the operators $\mathcal{R}_N^S$ have finite rank for all $N \geq 1$ and for any $\phi \in \mathcal{H}$, we have

$$\|(\mathcal{R}^S - \mathcal{R}^{S,N})(\phi)\|_{\mu_V} \asymp \|(\mathcal{U} - \mathcal{U}_N)(\phi)\|_{L^2(\mu_{sc})} = \sqrt{\sum_{k>N} |\phi_k|^2} \ll \frac{\|\phi\|}{\sqrt{N}}.$$

The first step is similar to (4.59), for the second step we used that by (4.13), $(\mathcal{U} - \mathcal{U}_N)(\phi) = \sum_{k>N} \phi_k U_{k-1}$ and at last, we used the Cauchy–Schwarz inequality and (4.10). This proves that $\mathcal{R}^S = \lim_{N \to \infty} \mathcal{R}^{S,N}$ in operator-norm so that $\mathcal{R}^S$ is compact. Therefore, by the spectral theorem, there exists an orthonormal basis $(\phi_n)_{n \geq 1}$ of $\mathcal{H}_{\mu_V}$ so that

$$(4.61) \qquad \mathcal{R}^S(\phi_n) = \frac{2}{\sigma_n} \phi_n$$

and $\sigma_n > 0$ for all $n \geq 1$. Moreover, each eigenvalue has finite multiplicity, and we may order them so that the sequence σ_n is nondecreasing and $\sigma_n \to \infty$ as $n \to \infty$. In fact, by the Min-Max theorem, we have

$$\frac{2}{\sigma_n} = \max_{\substack{\mathcal{S} \subset \mathcal{H} \\ \dim \mathcal{S} = n}} \min_{\phi \in \mathcal{S}} \frac{\langle \mathcal{R}^S \phi, \phi \rangle_{\mu_V}}{\|\phi\|_{\mu_V}^2},$$

where the maximum ranges over all n-dimensional subspaces of the Hilbert space $\mathcal{H}$. By (4.60), since the quantity $\Sigma(\phi)$ does not depend on the equilibrium measure μ_V and $\|\phi\|_{\mu_V} \asymp \|\phi\|$, this shows that

$$(4.62) \qquad \frac{2}{\sigma_n} \asymp \max_{\substack{\mathcal{S} \subset \mathcal{H} \\ \dim \mathcal{S} = n}} \min_{\phi \in \mathcal{S}} \frac{\langle \mathcal{R}^1 \phi, \phi \rangle}{\|\phi\|^2}.$$

In the Gaussian case ($S = 1$), the eigenfunctions of $\mathcal{R}^1$ are the Chebyshev polynomials $\frac{1}{n} T_n$ and the eigenvalues satisfy $\sigma_n = 2n$, therefore, we deduce from (4.62) that, in the general case, we also have $\sigma_n \asymp n$.

Finally, to prove that the eigenfunctions are bounded, we may expand ϕ_n in a Fourier–Chebyshev series: $\phi_n = \sum_{k=1}^{\infty} \phi_{n,k} T_k$ on J. Since $\|T_k\|_{\infty, J} = 1$ for all $k \geq 1$, by (4.10) and the Cauchy–Schwarz inequality, we obtain

$$\|\phi_n\|_{\infty, J} \leq \sum_{k \geq 1} |\phi_{n,k}| \ll \|\phi_n\|.$$

Finally, since $\|\phi_n\| \asymp \|\phi_n\|_{\mu_V} = 1$ for all $n \in \mathbb{N}$, we conclude that $\|\phi_n\|_{\infty, \bar{J}} \ll 1$. □

We split the proof of Theorem 1.5 into two parts—we now focus on existence and bounds on the functions ϕ_n and then in the next section discuss the Fourier expansion.

PROOF OF THEOREM 1.5—EXISTENCE AND BOUNDS ON EIGENFUNCTIONS. By Theorem 4.8, we have $\mathcal{R}^S(\phi_n) = \frac{2}{\sigma_n}\phi_n$ and if we differentiate this equation and use (4.58), we obtain that $S(x)\phi_n'(x) = \frac{\sigma_n}{2}\mathcal{U}_x(\phi_n)$ for all $x \in J$. Then, by Lemma A.4, this implies that for all $n \geq 1$,

$$\Xi^{\mu_V}(\phi_n) = \sigma_n \phi_n. \tag{4.63}$$

Hence, $\phi_n \in \mathcal{H}$ and $\sigma_n \asymp n$ solve equation (4.6). By Proposition 4.1, this means that for any $n \geq 1$, we can extend the eigenfunction (keeping the notation ϕ_n) in such a way that $\phi_n \in \mathscr{C}^\kappa(\mathbb{R})$ with compact support and it satisfies equation (4.4) on the interval J_{ϵ_n} where $\epsilon_n = \sigma_n^{-\eta} \wedge \delta$ and that for all $k < \kappa$,

$$\|\phi_n^{(k)}\|_{\infty,\mathbb{R}} \ll \sigma_n^{\eta k}.$$

Finally, combining formulae (4.60) and (4.61), since $\|\phi_n\|_{\mu_V} = 1$, we obtain the identity $2\boldsymbol{\Sigma}(\phi_n) = \frac{1}{\sigma_n}$ for all $n \geq 1$. The part of the theorem concerning existence and bounds on the functions is now proven. $\quad\square$

4.5. *Fourier expansion.* The proof of Theorem 1.5 will be complete once we prove the following result.

PROPOSITION 4.9. *If $f \in \mathscr{C}^{\kappa+4}(\bar{J})$, we can expand $f = \int_J f(x)\varrho(dx) + \sum_{n=1}^\infty \widehat{f}_n \phi_n$ where the Fourier coefficients $\widehat{f}_n = \langle f, \phi_n \rangle_{\mu_V}$ satisfy*

$$|\widehat{f}_n| \ll \sigma_n^{-\frac{\kappa+3}{2}}$$

for all $n \geq 1$.

To do this, we prove an integration by parts result.

LEMMA 4.10 (Integration by parts). *The operator Ξ^{μ_V} is symmetric in the sense that for any function $g, f \in \mathscr{C}^2(\mathbb{R})$, we have*

$$\langle \Xi^{\mu_V}(f), g \rangle_{\mu_V} = \langle f, \Xi^{\mu_V}(g) \rangle_{\mu_V}. \tag{4.64}$$

PROOF. We let $\tilde{g}(x) = g'(x)\frac{d\mu_V}{dx}$ for all $x \in \mathbb{R}$ and similarly for $\tilde{f}$. By assumption $\tilde{g}$ is continuous with support in $\bar{J}$, it is differentiable and $\tilde{g}' \in L^p(\mathbb{R})$ for any $p < 2$. Moreover, since $\Xi^{\mu_V}(f) = \mathcal{H}(\tilde{f})$, using the differentiation properties of the Hilbert transform, we have $\frac{d}{dx}\mathcal{H}(\tilde{f}) = \mathcal{H}(\tilde{f}')$. This proves that under our assumptions, the function $\mathcal{H}(\tilde{f})$ is continuous on $\mathbb{R}$. Then an integration by parts shows that

$$\langle \Xi^{\mu_V}(f), g \rangle_{\mu_V} = \int_J \frac{d\mathcal{H}_x(\tilde{f})}{dx}\tilde{g}(x)\,dx = -\int_J \mathcal{H}_x(\tilde{f})\tilde{g}'(x)\,dx.$$

Note that there is no boundary term because we have just seen that the function $\tilde{g}\mathcal{H}(\tilde{f})$ is continuous with support in $\bar{\mathrm{J}}$. Using the antiself-adjointness and differentiation properties of the Hilbert transform, this implies that

$$\langle \Xi^{\mu_V}(f), g \rangle_{\mu_V} = \int_{\mathrm{J}} \tilde{f}(x) \mathcal{H}_x(\tilde{g}') \, dx = \int_{\mathrm{J}} \tilde{f}(x) \frac{d\mathcal{H}_x(\tilde{g})}{dx} \, dx.$$

Finally, using that $\mathcal{H}(\tilde{g}) = \Xi^{\mu_V}(g)$ and $\tilde{f}(x)\,dx = f'(x)\mu_V(dx)$, we obtain (4.64) establishing Lemma 4.10. $\quad\square$

We now turn to the missing ingredient in the proof of Theorem 1.5

PROOF OF PROPOSITION 4.9. First of all, we claim that if $\phi \in \mathscr{C}^{k+2}(\bar{\mathrm{J}})$ for some $k \leq \kappa + 2$, then the function $\Xi^{\mu_V}(\phi) \in \mathscr{C}^k(\bar{\mathrm{J}})$. Indeed, by (4.5) and the variational condition (4.3), we see that for all $x \in \mathrm{J}$,

$$\Xi_x^{\mu_V}(\phi) = -\int_{\mathrm{J}} \frac{\phi'(t) - \phi'(x)}{t - x} \mu_V(dt) - V'(x)\phi'(x).$$

Using Lemma A.1, if $\phi' \in \mathscr{C}^{k+1}(\bar{\mathrm{J}})$ and the potential $V \in \mathscr{C}^{\kappa+3}(\mathbb{R})$, this implies that the function $\Xi_x^{\mu_V}(\phi) \in \mathscr{C}^k(\bar{\mathrm{J}})$. Let $f^{(0)} = f \in \mathscr{C}^{\kappa+4}(\bar{\mathrm{J}})$ and define $f^{(k+1)} = \Xi^{\mu_V}(f^{(k)})$ for all $0 \leq k \leq K$ (here $f^{(k)}$, is not to be confused with the kth derivative of f). The previous observation shows that $f^{(1)} \in \mathscr{C}^{\kappa+2}(\bar{\mathrm{J}}), \ldots, f^{(K)} \in \mathscr{C}^{\kappa-2(K-2)}(\bar{\mathrm{J}})$. Thus, choosing $K = \frac{\kappa+3}{2}$, we obtain that $f^{(K)} \in \mathscr{C}^1(\bar{\mathrm{J}})$. By definition, we have

$$\begin{aligned}
\widehat{f}_n = \langle f, \phi_n \rangle_{\mu_V} &= \sigma_n^{-1} \langle f, \Xi^{\mu_V}(\phi_n) \rangle_{\mu_V} \\
&= \sigma_n^{-1} \langle f^{(1)}, \phi_n \rangle_{\mu_V} \\
&= \sigma_n^{-K} \langle f^{(K)}, \phi_n \rangle_{\mu_V}.
\end{aligned}$$

At first, we used the eigenequation (4.63). Then we used Lemma 4.10 observing that the functions $f^{(k)} \in \mathscr{C}^2(\bar{\mathrm{J}})$ for all $0 \leq k < K$. The last step follows by induction. Hence, since $\|\phi_n\|_{\mu_V} = 1$ for all $n \geq 1$, we obtain

$$|\widehat{f}_n| \leq \sigma_n^{-K} \|f^{(K)}\|_{\mu_V} \ll \sigma_n^{-K}$$

since the function $f^{(K)\prime}$ is uniformly bounded on $\bar{\mathrm{J}}$. Proposition 4.9 is established. $\quad\square$

5. The CLT for the functions ϕ_n—Proof of Proposition 1.6.

In this section, we establish a multidimensional CLT for the linear statistics of the test functions ϕ_n constructed in the previous section, for a general one-cut regular potential V and $\beta > 0$. This will closely parallel the argument for the GUE from Section 3, but we do encounter some technical difficulties. We point out here that we will find it convenient to assume that the eigenvalues are ordered, $\lambda_1 \leq \cdots \leq \lambda_N$, allowing us to use the rigidity result of Theorem 1.7 to control the behavior of the random measure ν_N (see Section 5.2).

5.1. *Step* 1—*approximate eigenvector property.* In this section, we establish that linear statistics of ϕ_n are approximate eigenfunctions of the generator L as was the case for the Chebyshev polynomials when the potential V is quadratic.

PROPOSITION 5.1. *Let* $(\phi_n)_{n=1}^\infty$ *and* $(\sigma_n)_{n=1}^\infty$ *be as in Theorem* 1.5. *If* **m** *is given by* (1.10) *and* L *by* (2.2), *then for any* $n \geq 1$

$$
\text{L}\left(\sum_{j=1}^N \phi_n(\lambda_j)\right)
$$

(5.1)

$$
= -\beta \sigma_n N \left(\sum_{j=1}^N \phi_n(\lambda_j) - N \int_{\text{J}} \phi_n \, d\mu_V - \left(\frac{1}{\beta} - \frac{1}{2}\right)\mathbf{m}(\phi_n)\right) + \zeta_n(\lambda),
$$

where $\zeta_n : \mathbb{R}^N \to \mathbb{R}$ *satisfies the following two conditions: for any* $\eta > \frac{4(\kappa+1)}{2\kappa-1}$, $|\zeta_n(\lambda)| \ll N^2 n^{2\eta}$ *for all* $\lambda \in \mathbb{R}^N$ *and, if* $\lambda_1, \ldots, \lambda_N \in [-1 - \epsilon_n, 1 + \epsilon_n]$ *(where* ϵ_n *is as in Theorem* 1.5*), then*

$$
\zeta_n(\lambda) = \left(1 - \frac{\beta}{2}\right) \int_{\mathbb{R}} \phi_n''(x) \nu_N(dx)
$$

(5.2)

$$
+ \frac{\beta}{2} \int_{\mathbb{R} \times \mathbb{R}} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \nu_N(dx) \nu_N(dt).
$$

To prove Proposition 5.1, we will need the following lemma.

LEMMA 5.2. *Let* $(\phi_n)_{n=1}^\infty$ *and* $(\sigma_n)_{n=1}^\infty$ *be as in Theorem* 1.5. *We have for all* $n \geq 1$,

$$
\int_{\text{J} \times \text{J}} \frac{\phi_n'(t) - \phi_n'(x)}{t - x} \mu_V(dt) \mu_V(dx) = -2\sigma_n \int_{\text{J}} \phi_n(x) \mu_V(dx).
$$

PROOF. First of all, let us observe that using the variational condition (4.3), for any function $f \in \mathscr{C}_c^1(\mathbb{R})$,

$$
\int_{\text{J}} f(x) V'(x) \mu_V(dx) = \int_{\text{J}} f(x) \mathcal{H}_x(\mu_V) \mu_V(dx) = -\int_{\text{J}} \mathcal{H}_x(f\mu_V) \mu_V(dx),
$$

where we used the antiself-adjointness of the Hilbert transform, (A.3), and the fact that μ_V is absolutely continuous with a bounded density function. In fact, one has

$$
\mathcal{H}_x(f\mu_V) = -\int_{\text{J}} \frac{f(t) - f(x)}{t - x} \mu_V(dt) + f(x) V'(x)
$$

and, by symmetry, this implies that

(5.3) $$\int_{\text{J}} f(x) V'(x) \mu_V(dx) = \frac{1}{2} \int_{\text{J} \times \text{J}} \frac{f(t) - f(x)}{t - x} \mu_V(dt) \mu_V(dx).$$

On the other hand, integrating equation (1.14) with respect to μ_V, we see that

$$\int_J V'(x)\phi_n'(x)\mu_V(dx) - \int_{J \times J} \frac{\phi_n'(x) - \phi_n'(y)}{x - y} \mu_V(dx)\mu_V(dy)$$

$$= \sigma_n \int_J \phi_n(u)\mu_V(du),$$

which, together with (5.3) completes the proof of this lemma. $\quad\square$

PROOF OF PROPOSITION 5.1. Let us first prove the uniform bound for the error term ζ_n on $\mathbb{R}^N$. Since the density $S = \frac{d\mu_V}{d\mu_{sc}}$ is $\mathscr{C}^1$ and positive on $\bar{J} = [-1, 1]$, we have (in the notation of (1.3))

$$\left|\mathbf{m}(\phi_n)\right| \ll \|\phi_n\|_{\infty,J} + \|\phi_n'\|_{\infty,J}.$$

One has

$$(5.4)\quad \begin{aligned} L\left(\sum_{j=1}^N \phi_n(\lambda_j)\right) &= \sum_{j=1}^N \phi_n''(\lambda_j) - \beta N \sum_{j=1}^N V'(\lambda_j)\phi_n'(\lambda_j) \\ &\quad + \beta \sum_{i \neq j} \frac{1}{\lambda_j - \lambda_i}\phi_n'(\lambda_j) \\ &= \left(1 - \frac{\beta}{2}\right)\sum_{j=1}^N \phi_n''(\lambda_j) - \beta N \sum_{j=1}^N V'(\lambda_j)\phi_n'(\lambda_j) \\ &\quad + \frac{\beta}{2}\sum_{i,j=1}^N \frac{\phi_n'(\lambda_j) - \phi_n'(\lambda_i)}{\lambda_j - \lambda_i}. \end{aligned}$$

Hence, we obtain

$$\left|L\left(\sum_{j=1}^N \phi_n(\lambda_j)\right)\right| \ll N^2(\|\phi_n''\|_{\infty,\mathbb{R}} + \|V'\|_{\infty,J_{2\delta}}\|\phi_n'\|_{\infty,\mathbb{R}}),$$

where the last bound follows from the fact that, by construction, the functions ϕ_n have compact support in the interval $J_{2\delta}$. Combining these two estimates, using the upper bound (1.15) and the fact that $\sigma_n \asymp n$ from Theorem 1.5, we obtain

$$\begin{aligned} |\zeta_n(\lambda)| &= \left|L\left(\sum_{j=1}^N \phi_n(\lambda_j)\right) + \beta \sigma_n N\left(\sum_{j=1}^N \phi_n(\lambda_j) - N\int_J \phi_n(\lambda)\mu_V(d\lambda) - \mathbf{m}(\phi_n)\right)\right| \\ &\ll N^2 \sigma_n^{2\eta}, \end{aligned}$$

which yields $\|\zeta_n\|_{L^\infty(\mathbb{R}^N)} \ll N^2 n^{2\eta}$.

In order to prove (5.2), observe that

$$\sum_{i,j=1}^{N} \frac{\phi_n'(\lambda_j) - \phi_n'(\lambda_i)}{\lambda_j - \lambda_i} = 2N \iint_{\mathbb{R} \times J} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \nu_N(dx) \mu_V(dt)$$

$$+ \iint_{\mathbb{R} \times \mathbb{R}} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \nu_N(dx) \nu_N(dt)$$

$$+ N^2 \iint_{J \times J} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \mu_V(dx) \mu_V(dt).$$

Thus, if $\lambda_1, \ldots, \lambda_N \in J_{\epsilon_n}$, using equation (1.14), we obtain

$$N \sum_{j=1}^{N} V'(\lambda_j) \phi_n'(\lambda_j) - \frac{1}{2} \sum_{i,j=1}^{N} \frac{\phi_n'(\lambda_j) - \phi_n'(\lambda_i)}{\lambda_j - \lambda_i}$$

$$= N \sigma_n \sum_{j=1}^{N} \phi_n(\lambda_j) - \frac{1}{2} \iint_{\mathbb{R} \times \mathbb{R}} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \nu_N(dx) \nu_N(dt)$$

$$+ \frac{N^2}{2} \iint_{J \times J} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \mu_V(dx) \mu_V(dt)$$

$$= N \sigma_n \int_{\mathbb{R}} \phi_n(x) \nu_N(dx) - \frac{1}{2} \iint_{\mathbb{R} \times \mathbb{R}} \frac{\phi_n'(x) - \phi_n'(t)}{x - t} \nu_N(dx) \nu_N(dt),$$

where we used Lemma 5.2 for the last step. According to (5.4) and (5.2), it follows that

$$L\left(\sum_{j=1}^{N} \phi_n(\lambda_j)\right)$$

$$= -\beta \sigma_n N \left(\int_{\mathbb{R}} \phi_n(x) \nu_N(dx) + \left(\frac{1}{2} - \frac{1}{\beta}\right) \frac{1}{\sigma_n} \int_J \phi_n''(x) \mu_V(dx)\right) + \zeta_n(\lambda).$$

Therefore, to complete the proof, it just remains to check that according to (1.10), we have for all $n \geq 1$

$$(5.5) \qquad\qquad \mathbf{m}(\phi_n) = \frac{1}{\sigma_n} \int_J \phi_n''(x) \mu_V(dx).$$

Since the density $S = \frac{d\mu_V}{d\mu_{\mathrm{sc}}}$ is $\mathscr{C}^1$ and positive on $\bar{J} = [-1, 1]$, an integration by parts shows that

$$\int_J \phi_n''(x) \mu_V(dx) = -\frac{2}{\pi} \int_J \phi_n'(x) S(x) \left(\frac{S'(x)}{S(x)} \sqrt{1 - x^2} - \frac{x}{\sqrt{1 - x^2}}\right) dx.$$

Since for any $n \geq 1$ ϕ_n is a solution of equation (4.16), this implies that

$$\frac{1}{\sigma_n} \int_J \phi_n''(x) \mu_V(dx) = -\frac{1}{\pi} \int_J \mathcal{U}_x(\phi_n) \left(\frac{S'(x)}{S(x)} \sqrt{1 - x^2} - \frac{x}{\sqrt{1 - x^2}}\right) dx.$$

By Lemma A.5, we may rewrite this formula as

$$\frac{1}{\sigma_n} \int_J \phi_n''(x) \mu_V(dx) = \frac{1}{2} \Big(\phi_n(1) + \phi_n(-1) - \int_J \mathcal{U}_x(\phi_n) \frac{S'(x)}{S(x)} \mu_{sc}(dx) \Big),$$

which, according to (A.4), completes the proof of (5.5). The proof of Proposition 5.1 is thus complete. $\square$

5.2. *Step 2—a priori bound on linear statistics.* In this section, we use the rigidity estimates from Theorem 1.7 to establish a result corresponding to (3.7). The following lemma is a standard consequence of rigidity, but we record it for completeness.

LEMMA 5.3. *Let $\epsilon > 0$ be fixed and $\mathcal{B}_\epsilon$ be the event defined in Theorem 1.7. We have, for any configuration $\lambda \in \mathcal{B}_\epsilon$ and for any Lipschitz continuous function $f : \mathbb{R} \to \mathbb{R}$,*

$$\Big| \sum_{j=1}^N f(\lambda_j) - N \int_J f(x) \mu_V(dx) \Big| \ll \|f\|_{\mathscr{C}^{0,1}(\mathbb{R})} N^\epsilon.$$

In particular, if f is a bounded Lipschitz function, then for any $p > 0$,

$$\mathbb{E} \Big| \sum_{j=1}^N f(\lambda_j) - N \int_J f(x) \mu_V(dx) \Big|^p \ll_p \|f\|_{\mathscr{C}^{0,1}(\mathbb{R})}^p N^{p\epsilon} + \|f\|_{\infty,\mathbb{R}}^p e^{-N^{c\epsilon}}.$$

PROOF. Recalling the definition of the classical locations ρ_j from Theorem 1.7, we may write

$$\sum_{j=1}^N f(\lambda_j) - N \int_J f(x) \mu_V(dx) = N \sum_{j=1}^N \int_{\rho_{j-1}}^{\rho_j} (f(\lambda_j) - f(x)) \mu_V(dx)$$

and we see that if $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathcal{B}_\epsilon$,

$$\Big| \sum_{j=1}^N f(\lambda_j) - N \int_J f(x) \mu_V(dx) \Big|$$

(5.6)
$$\leq \|f\|_{\mathscr{C}^{0,1}(\mathbb{R})} \sum_{j=1}^N \max(|\lambda_j - \rho_j|, |\lambda_j - \rho_{j-1}|)$$

$$\ll \|f\|_{\mathscr{C}^{0,1}(\mathbb{R})} \sum_{j=1}^N \{\widehat{j}^{-\frac{1}{3}} N^{-\frac{2}{3}+\epsilon} + (\rho_j - \rho_{j-1})\}.$$

Obviously, $\sum_{j=1}^N \widehat{j}^{-\frac{1}{3}} \ll N^{\frac{2}{3}}$ and $\sum_{j=1}^N (\rho_j - \rho_{j-1}) = \rho_N - \rho_0 = 2$ due to the normalization of the support of the equilibrium measure, so that we have proved

the first estimate. The second estimate follows directly from the facts that f is bounded and, by Theorem 1.7, that the probability of the complement of $\mathcal{B}_\epsilon$ is exponentially small in some power of N. Lemma 5.3 is established. $\quad\square$

LEMMA 5.4.   *Let $\epsilon > 0$ be fixed and $f \in \mathscr{C}^4(\mathbb{R})$. We have, for any $\lambda \in \mathcal{B}_\epsilon$,*

$$\left| \int_{\mathbb{R}\times\mathbb{R}} \frac{f'(x) - f'(y)}{x - y} v_N(dx) v_N(dy) \right| \ll \|f^{(4)}\|_{\infty,\mathbb{R}} N^{2\epsilon}.$$

PROOF.   We begin by pointing out the following fact that is easily checked by repeatedly performing the integrals: for any $f \in \mathscr{C}^4(\mathbb{R})$ and $s, t, u, v \in \mathbb{R}$,

$$
\begin{aligned}
\frac{f'(s) - f'(t)}{s - t} &+ \frac{f'(u) - f'(v)}{u - v} - \frac{f'(s) - f'(v)}{s - v} - \frac{f'(u) - f'(t)}{u - t} \\
&= (s - u)(t - v) \\
&\quad \times \int_{[0,1]^3} f^{(4)}\big(ac(s - u) + (1 - a)b(t - v) + au + (1 - a)v\big) \\
&\quad \times a(1 - a)\, da\, db\, dc,
\end{aligned}
$$
(5.7)

where the LHS is interpreted as f'' when some of the parameters coincide. Noting that

$$
\begin{aligned}
\int_{\mathbb{R}\times\mathbb{R}} &\frac{f'(x) - f'(y)}{x - y} v_N(dx) v_N(dy) \\
&= \sum_{i,j=1}^{N} N^2 \int_{\rho_{i-1}}^{\rho_i} \int_{\rho_{j-1}}^{\rho_j} \left[\frac{f'(\lambda_i) - f'(\lambda_j)}{\lambda_i - \lambda_j} + \frac{f'(x) - f'(y)}{x - y} \right. \\
&\qquad \left. - \frac{f'(\lambda_i) - f'(x)}{\lambda_i - x} - \frac{f'(\lambda_j) - f'(y)}{\lambda_j - y} \right] \mu_V(dx) \mu_V(dy),
\end{aligned}
$$

formula (5.7) implies that

$$
\begin{aligned}
\left| \int_{\mathbb{R}\times\mathbb{R}} \right. &\left. \frac{f'(x) - f'(y)}{x - y} v_N(dx) v_N(dy) \right| \\
&\le \|f^{(4)}\|_{\infty,\mathbb{R}} N^2 \sum_{i,j=1}^{N} \int_{\rho_{i-1}}^{\rho_i} |\lambda_i - x| \mu_V(dx) \int_{\rho_{j-1}}^{\rho_j} |\lambda_j - y| \mu_V(dy) \\
&\le \|f^{(4)}\|_{\infty,\mathbb{R}} \left(\sum_{i=1}^{N} \max(|\lambda_i - \rho_i|, |\lambda_i - \rho_{i-1}|) \right)^2.
\end{aligned}
$$

Like in the proof of Lemma 5.3, conditionally on $\mathcal{B}_\epsilon$, the previous sum is asymptotically of order at most N^ϵ which completes the proof. $\quad\square$

5.3. *Step 3—controlling the error terms A and B*. As for the GUE, we now move on to estimate the "*A*-term" and "*B*-term" by using the results of Section 5.2 and Theorem 1.5 in order to apply Proposition 2.1 and deduce the multidimensional CLT in the next subsection.

LEMMA 5.5. *Let $(\phi_n)_{n=1}^{\infty}$ and $(\sigma_n)_{n=1}^{\infty}$ be given as in Theorem 1.5. Let $F : \mathbb{R}^N \to \mathbb{R}^d$ be defined by*

$$F(\lambda) = F_n(\lambda)$$

(5.8)
$$= \sqrt{\beta \sigma_n}\left(\sum_{j=1}^{N} \phi_n(\lambda_j) - N \int_J \phi_n(x)\mu_V(dx) - \left(\frac{1}{\beta} - \frac{1}{2}\right)\mathbf{m}(\phi_n)\right),$$

and let the matrix $K = K_{V,\beta}^{N,d} = \beta N \mathrm{diag}(\sigma_1, \ldots, \sigma_d)$. Then we have, for any $\eta > \frac{4(\kappa+1)}{2\kappa-1}$, any $\epsilon > 0$, and for any $d \leq N^{\frac{1}{2\eta}}$,

$$A^2 = \mathbb{E}\|F + K^{-1}\mathrm{L}(F)\|^2 \ll d^{8\eta}N^{-2+4\epsilon}.$$

PROOF. According to (5.8), we may rewrite (5.1) as

$$\mathrm{L}F = -KF + \sqrt{\beta \sigma_n}\zeta,$$

where we extend the definition of $\zeta : \mathbb{R}^N \to \mathbb{R}^d$. Then we have

$$A^2 = \frac{1}{\beta N^2}\mathbb{E}\left[\sum_{n=1}^{d} \frac{\zeta_n(\lambda)^2}{\sigma_n}\right].$$

Since $\|\zeta_n\|_{L^{\infty}(\mathbb{R}^N)} \ll N^2 n^{2\eta}$ for all $n \geq 1$ by Theorem 1.7, this implies that

$$\left|A^2 - \frac{1}{\beta N^2}\sum_{n=1}^{d}\frac{\mathbb{E}[\zeta_n(\lambda)^2\mathbf{1}_{\mathcal{B}_\epsilon}]}{\sigma_n}\right| \ll N^2 d^{4\eta}e^{-N^c}.$$

Recall that $\epsilon_d \asymp \sigma_d^{-\eta}$ when d is large, so that for any $\epsilon > 0$, we have $\mathcal{B}_\epsilon \subset [-1 - \epsilon_d, 1 + \epsilon_d]^N$ when the parameter N is sufficiently large. Then, by (5.2),

$$\mathbb{E}[\zeta_n(\lambda)^2\mathbf{1}_{\mathcal{B}_\epsilon}] \ll \mathbb{E}\left[\mathbf{1}_{\mathcal{B}_\epsilon}\left(\int_{\mathbb{R}}\phi_n''(x)\nu_N(dx)\right)^2\right]$$
$$+ \mathbb{E}\left[\mathbf{1}_{\mathcal{B}_\epsilon}\left(\int_{\mathbb{R}\times\mathbb{R}}\frac{\phi_n'(x) - \phi_n'(t)}{x - t}\nu_N(dx)\nu_N(dt)\right)^2\right].$$

By Lemma 5.3, the first term is bounded by $\|\phi_n^{(3)}\|_{\infty,\mathbb{R}}^2 N^{2\epsilon}$, while by Lemma 5.4, the second term is bounded by $\|\phi_n^{(4)}\|_{\infty,\mathbb{R}}^2 N^{4\epsilon}$. Now, using the estimates of Theorem 1.5, we conclude that

$$A^2 \ll N^{-2+4\epsilon}\sum_{n=1}^{d}\sigma_n^{8\eta-1} \asymp N^{-2+4\epsilon}\sigma_d^{8\eta}$$

which completes the proof. $\square$

LEMMA 5.6. *Let B be given by formula* (2.5) *where $F : \mathbb{R}^N \to \mathbb{R}^d$ and the matrix K are as in Lemma* 5.5. *Then we have, for any $\eta > \frac{4(\kappa+1)}{2\kappa-1}$ and any $\epsilon > 0$,*

$$(5.9) \qquad\qquad B^2 = \mathbb{E}\|\mathrm{Id} - K^{-1}\Gamma(F)\|^2 \ll d^{6\eta+2}N^{-2+2\epsilon},$$

as long as the RHS converges to 0 as $N \to \infty$. (In particular, this is the case when $d \leq N^{\frac{1}{4\eta}}$.)

PROOF.   By definition, we have

$$\Gamma(F_i, F_j) = \beta\sqrt{\sigma_i\sigma_j} \sum_{\ell=1}^{N} \phi_i'(\lambda_\ell)\phi_j'(\lambda_\ell)$$

$$= \beta\sqrt{\sigma_i\sigma_j}\left(\int_{\mathbb{R}} \phi_i'(x)\phi_j'(x)\nu_N(dx) + N\delta_{i,j}\right)$$

since, by Theorem 1.5, $(\phi_n)_{n=1}^{\infty}$ is an orthonormal basis of the Sobolev space $\mathscr{H}_{\mu_V}$. This implies that, for the Hilbert–Schmidt norm,

$$\|\mathrm{Id} - K^{-1}\Gamma(F)\|^2 = \frac{1}{N^2} \sum_{i,j=1}^{d} \frac{\sigma_j}{\sigma_i}\left(\int_{\mathbb{R}} \phi_i'(x)\phi_j'(x)\nu_N(dx)\right)^2$$

$$\leq \frac{2}{N^2} \sum_{1\leq i \leq j \leq d} \frac{\sigma_j}{\sigma_i}\left(\int_{\mathbb{R}} \phi_i'(x)\phi_j'(x)\nu_N(dx)\right)^2.$$

Then, by Lemma 5.3, since the functions $x \mapsto \phi_i'(x)\phi_j'(x)$ are uniformly bounded and Lipschitz continuous on $\mathbb{R}$ with norm at most

$$\|\phi_j'\|_{\infty,\mathbb{R}}\|\phi_i''\|_{\infty,\mathbb{R}} + \|\phi_i'\|_{\infty,\mathbb{R}}\|\phi_j''\|_{\infty,\mathbb{R}} \ll \sigma_i^{\eta}\sigma_j^{2\eta},$$

using the estimate (1.15) when $i \leq j$, we obtain

$$B^2 \ll N^{-2(1-\epsilon)} \sum_{1\leq i \leq j \leq d} \sigma_i^{2\eta-1}\sigma_j^{4\eta+1} + d^2 e^{-N^c}$$

which yields the estimate (5.9) since $\sigma_j \asymp j$ and the second term is asymptotically negligible.   □

5.4. *Proof of Proposition* 1.6.   By Theorem 1.5, $\sigma_n = \frac{1}{2\Sigma(\phi_n)}$, so that according to (5.8), we have for all $n \geq 1$

$$\mathcal{X}_N(\phi_n) = F_n(\lambda).$$

Then, by Proposition 2.1 and using the bounds of Lemmas 5.5 and 5.6, we obtain for any $\eta > \frac{4(\kappa+1)}{2\kappa-1}$,

$$\mathrm{W}_2\big((\mathcal{X}_N(\phi_n))_{n=1}^{d}, \gamma_d\big) \leq A + B \ll d^{4\eta}N^{-1+2\epsilon}.$$

Here, we used that the main error term is given by A when the parameter d is large. Optimizing over the parameter η thus completes the proof.   □

6. The CLT for general test functions. In this section, we fix a test function $f \in \mathscr{C}_c^{\kappa+4}(\mathbb{R})$ with $\kappa \geq 5$. Without loss of generality, we may also assume that $\widehat{f_0} = \int f(x)\varrho(dx) = 0$ (otherwise we might consider the function $\tilde{f} = f - \widehat{f_0}$ instead and note that by formulae (1.10) and (1.13), we have both $\mathbf{m}(\tilde{f}) = \mathbf{m}(f)$ and $\mathcal{X}_N(\tilde{f}) = \mathcal{X}_N(f)$) and we define for any $d \geq 1$

$$(6.1) \qquad g_d(x) = \sum_{n=1}^{d} \widehat{f_n}\phi_n(x),$$

where $\widehat{f_n} = \langle f, \phi_n \rangle_{\mu_V}$. Note that if $\kappa \geq 5$, by Theorem 1.5, the sum (6.1) converges as $d \to \infty$ and this defines a function $g_\infty \in \mathscr{C}_c^1(\mathbb{R})$ which satisfies $g_\infty(x) = f(x)$ for all $x \in J$. The proof of Theorem 1.2 is divided into two steps. First, we make use of the multidimensional CLT—Proposition 1.6— to prove a CLT for the function g_d in the regime as $d = d(N) \to \infty$. Then we establish that along a suitable sequence $d(N)$ the random variable $\mathcal{X}_N(f)$ given by (1.13), is well approximated with respect the Kantorovich distance by

$$(6.2) \qquad X_N^d = \sqrt{\frac{\beta}{2\boldsymbol{\Sigma}(f)}}\left(\int_{\mathbb{R}} g_d(x)\nu_N(dx) - \left(\frac{1}{2} - \frac{1}{\beta}\right)\mathbf{m}(f)\right).$$

6.1. *Consequence of the multidimensional CLT.* In this section, we shall use Proposition 1.6 to establish that, if N is sufficiently large (compared to the parameter d), then the law of the random variable X_N^d is close to the Gaussian measure γ_1. In particular, in order to control the error term with respect to the Kantorovich metric, we need to express the correction term to the mean $\mathbf{m}(f)$ and the asymptotic variance $\boldsymbol{\Sigma}(f)$ of a linear statistics in terms of the Fourier coefficients of the test function f. This is the goal of the next lemma.

LEMMA 6.1. *If $f \in \mathscr{C}^{\kappa+4}(\bar{J})$ and $\kappa \geq 5$, using the notation of Proposition 4.9, we have*

$$(6.3) \qquad \boldsymbol{\Sigma}(f) = \sum_{n=1}^{\infty} \frac{\widehat{f_n}^2}{2\sigma_n}$$

and

$$(6.4) \qquad \mathbf{m}(f) = \sum_{n=1}^{\infty} \widehat{f_n}\mathbf{m}(\phi_n).$$

PROOF. On the one hand, by (4.15) and (4.58), we have

$$\boldsymbol{\Sigma}(f) = \frac{1}{4}\int_J f'(x)\mathcal{U}_x(f)\mu_{\mathrm{sc}}(dx) = \frac{1}{4}\langle f, \mathcal{R}^S(f)\rangle_{\mu_V}.$$

On the other hand, by linearity and (4.61), we have

$$\mathcal{R}^S(f) = \sum_{n \in \mathbb{N}} \widehat{f}_n \mathcal{R}^S(\phi_n) = 2 \sum_{n \in \mathbb{N}} \frac{\widehat{f}_n}{\sigma_n} \phi_n,$$

where the sum converges uniformly. Since $(\phi_n)_{n \geq 1}$ is an orthonormal basis of $\mathscr{H}_{\mu_V}$, this yields (6.3). Formula (6.4) follows by linearity of the operator $\mathbf{m}$, (1.10). Moreover, note that by (5.5),

$$\mathbf{m}(\phi_n) = \frac{1}{\sigma_n} \int_J \phi_n''(x) \mu_V(dx) = -\frac{2}{\pi \sigma_n} \int_J \phi_n'(x) \frac{d}{dx}\left(S(x)\sqrt{1-x^2}\right) dx,$$

so that for any $\eta > \frac{4(\kappa+1)}{2\kappa-1}$,

(6.5) $$\left|\mathbf{m}(\phi_n)\right| \ll \sigma_n^{-1} \|\phi_n'\|_{\infty, J} \ll \sigma_n^{\eta-1}.$$

Using the estimate of Theorem 1.5, this proves that the series (6.4) converges absolutely as long as $\eta < 3$ and $\kappa \geq 5$. $\square$

LEMMA 6.2.   *Let X_N^d be as in (6.2) and assume that $d^{\frac{\kappa+1}{2}} \leq N$. Then, for any $\epsilon > 0$,*

$$W_2^2(X_N^d, \gamma_1) \ll \left(d^{8\eta_*} N^{-2} + d^{2\eta_* - \kappa - 3}\right) N^\epsilon,$$

where $\eta_* := \frac{4(\kappa+1)}{2\kappa-1}$.

PROOF.   Using the notation (5.8) and by linearity of $\mathbf{m}$, we may rewrite

$$X_N^d = \frac{1}{\sqrt{2\Sigma(f)}} \left(\sum_{n=1}^d \frac{\widehat{f}_n F_n}{\sqrt{\sigma_n}} + \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right) \mathbf{m}(g_d - f) \right).$$

Let $Z \sim \gamma_d$. It follows from the previous formula and the triangle inequality that

(6.6)
$$W_2^2(X_N^d, \gamma_1) \leq \frac{2}{\Sigma(f)} \left\{ W_2^2\left(\sum_{n=1}^d \frac{\widehat{f}_n F_n}{\sqrt{2\sigma_n}}, \sum_{n=1}^d \frac{\widehat{f}_n Z_n}{\sqrt{2\sigma_n}} \right) \right.$$
$$+ W_2^2\left(\sum_{n=1}^d \frac{\widehat{f}_n Z_n}{\sqrt{2\sigma_n}} + \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right) \mathbf{m}(g_d - f), \right.$$
$$\left. \left. \sqrt{\Sigma(f)}\gamma_1 \right) \right\}.$$

Our task is to prove that both term on the RHS of (6.6) converge to 0 as $N \to \infty$. The second term corresponds to the distance between two Gaussian measures on

$\mathbb{R}$ and equals

$$W_2^2\left(\sum_{n=1}^d \frac{\widehat{f_n} Z_n}{\sqrt{2\sigma_n}} + \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right)\mathbf{m}(g_d - f), \sqrt{\boldsymbol{\Sigma}(f)}\gamma_1\right)$$

$$= \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right)^2 \mathbf{m}(g_d - f)^2 + \left(\sqrt{\sum_{n=1}^d \frac{\widehat{f_n^2}}{2\sigma_n}} - \sqrt{\boldsymbol{\Sigma}(f)}\right)^2.$$

By Lemma 6.1, this implies that

$$W_2^2\left(\sum_{n=1}^d \frac{\widehat{f_n} Z_n}{\sqrt{2\sigma_n}} + \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right)\mathbf{m}(g_d - f), \sqrt{\boldsymbol{\Sigma}(f)}\gamma_1\right)$$

$$\ll \left(\sum_{n>d} \widehat{f_n}\mathbf{m}(\phi_n)\right)^2 + \sum_{n>d} \frac{\widehat{f_n^2}}{2\sigma_n}.$$

Using the estimate (6.5), we see that both sums converge if $\kappa \geq 5$ and we obtain, for any $\epsilon > 0$,

(6.7)
$$W_2^2\left(\sum_{n=1}^d \frac{\widehat{f_n} Z_n}{\sqrt{2\sigma_n}} + \left(\frac{\sqrt{\beta}}{2} - \frac{1}{\sqrt{\beta}}\right)\mathbf{m}(g_d - f), \sqrt{\boldsymbol{\Sigma}(f)}\gamma_1\right)$$

$$\ll d^{2\eta - \kappa - 3}$$

$$\ll d^{2\eta_* - \kappa - 3} N^\epsilon.$$

On the other hand, by definition of the Kantorovich distance and the Cauchy–Schwarz inequality, the first term on the RHS of in (6.6) satisfies the bound

$$W_2^2\left(\sum_{n=1}^d \frac{\widehat{f_n} F_n}{\sqrt{2\sigma_n}}, \sum_{n=1}^d \frac{\widehat{f_n} Z_n}{\sqrt{2\sigma_n}}\right) \leq \left(\sum_{n=1}^d \frac{\widehat{f_n^2}}{2\sigma_n}\right) W_2^2((F_n)_{n=1}^d, \gamma_d).$$

By definition, $F_n = \mathcal{X}_N(\phi_n)$, so that by Proposition 1.6, we obtain

(6.8)
$$W_2^2\left(\sum_{n=1}^d \frac{\widehat{f_n} F_n}{\sqrt{2\sigma_n}}, \sum_{n=1}^d \frac{\widehat{f_n} Z_n}{\sqrt{2\sigma_n}}\right) \ll \boldsymbol{\Sigma}(f) d^{8\eta_*} N^{-2+2\epsilon}.$$

Combining the estimates (6.6), (6.7) and (6.8) completes the proof of the lemma.
$\square$

6.2. *Truncation.* The goal of this section is to prove that the linear statistic $\mathcal{X}_N(f)$ is asymptotically close to the random variable X_N^d with respect to the Kantorovich distance, as $d \to \infty$ and $N \to \infty$ simultaneously in a suitable way. At first, we compare the laws of the linear statistics $\mathcal{X}_N(f)$ and X_N^∞ by using rigidity of the random configurations under the Gibbs measure $\mathbb{P}_{V,\beta}^N$. Then we compare the laws of X_N^d and X_N^∞ by using the decay of the Fourier coefficients of the test function f; cf. Theorem 1.5.

LEMMA 6.3. *According to the notation* (1.13) *and* (6.2), *we have, for any* $\epsilon > 0$,

$$W_2^2(X_N^\infty, \mathcal{X}_N(f)) \ll_f N^{-\frac{4}{3}+\epsilon}.$$

PROOF. By definition,

$$\mathcal{X}_N(f) - X_N^\infty = \sqrt{\frac{\beta}{2\Sigma(f)}} \int_{\mathbb{R}} (f - g_\infty) \, d\nu_N$$

so that

$$W_2^2(X_N^\infty, \mathcal{X}_N(f)) \leq \frac{\beta}{2\Sigma(f)} \mathbb{E}\left(\int_{\mathbb{R}} (f - g_\infty) \, d\nu_N\right)^2.$$

Since both $f, g_\infty \in \mathscr{C}^1(\mathbb{R})$, up to an exponentially small error term that we shall neglect, we see that for any $\epsilon > 0$,

$$(6.9) \qquad W_2^2(X_N^\infty, \mathcal{X}_N(f)) \ll \mathbb{E}\left(\mathbf{1}_{\mathcal{B}_\epsilon} \int_{\mathbb{R}} (f - g_\infty) \, d\nu_N\right)^2.$$

We claim that it follows from the square-root vanishing of the equilibrium measure at the edges that for all $j \in \{1, \ldots, N\}$,

$$(6.10) \qquad \rho_j - \rho_0 \gg \left(\frac{j}{N}\right)^{\frac{2}{3}} \quad \text{and} \quad \rho_N - \rho_{N-j} \gg \left(\frac{j}{N}\right)^{\frac{2}{3}}.$$

This shows that for any $\lambda \in \mathcal{B}_\epsilon$, $\lambda_j \in \mathsf{J}$ for all $\widehat{j} \geq N^{3\epsilon}$. Since $g_\infty(x) = f(x)$ for all $x \in \mathsf{J}$, this implies that for any $\lambda \in \mathcal{B}_\epsilon$,

$$\left|\int_{\mathbb{R}} (f - g_\infty) \, d\nu_N\right| \leq \sum_{\widehat{j} \leq N^{3\epsilon}} \int_{\rho_{j-1}}^{\rho_j} \left|(f - g_\infty)(\lambda_j) - (f - g_\infty)(x)\right| \mu_V(dx).$$

Like in the proof of Lemma 5.3 (see in particular the estimate (5.6)), we obtain

$$\left|\int_{\mathbb{R}} (f - g_\infty) \, d\nu_N\right| \ll \|f' - g_\infty'\|_{L^\infty(\mathbb{R})} \sum_{\widehat{j} \leq N^{3\epsilon}} \{\widehat{j}^{-\frac{1}{3}} N^{-\frac{2}{3}+\epsilon} + (\rho_j - \rho_{j-1})\}$$

$$\ll_f N^{-\frac{2}{3}+4\epsilon} + (\rho_{\lfloor N^{3\epsilon}\rfloor} - \rho_0) + (\rho_N - \rho_{N-\lceil N^{3\epsilon}\rceil}).$$

By (6.10), this error term is of order $N^{-\frac{2}{3}+5\epsilon}$ and by (6.9), we conclude that

$$W_2^2(X_N^\infty, \mathcal{X}_N(f)) \ll_f N^{-\frac{4}{3}+10\epsilon}.$$

Since the parameter $\epsilon > 0$ is arbitrary, this completes the proof. □

LEMMA 6.4. *If* $d^{\frac{\kappa+1}{2}} \leq N$, *then we have, for any* $\epsilon > 0$,

$$W_2^2(X_N^d, X_N^\infty) \ll d^{2\eta_* - \kappa - 1} N^\epsilon,$$

where as above $\eta_* := \frac{4(\kappa+1)}{2\kappa-1}$.

PROOF. By Theorem 1.5, since $\sum_{n \geq 1} |\widehat{f}_n| < \infty$ and $\|\phi_n\|_{\infty, \mathbb{R}} \ll 1$ where the implied constant is independent of $n \geq 1$ we have

$$X_N^\infty - X_N^d = \sqrt{\frac{\beta}{2\boldsymbol{\Sigma}(f)}} \sum_{n > d} \widehat{f}_n \int_{\mathbb{R}} \phi_n(x) v_N(dx)$$

so that

$$(6.11) \qquad W_2^2(X_N^d; X_N^\infty) \leq \frac{\beta}{2\boldsymbol{\Sigma}(f)} \mathbb{E}\left(\sum_{n > d} \widehat{f}_n \int \phi_n(x) v_N(dx)\right)^2.$$

Note that by Lemma 5.3 and the estimate (1.15), conditionally on the event $\mathcal{B}_\epsilon$, we have for all $n \geq 1$

$$\left|\int \phi_n(x) v_N(dx)\right| \ll N^\epsilon \sigma_n^\eta.$$

This bound turns out to be useful when $n \ll N^{\frac{1}{\eta}}$, otherwise it is better to use the trivial bound

$$\left|\int \phi_n(x) v_N(dx)\right| \ll N.$$

From these two estimates, we deduce that

$$\mathbb{E}\left(\sum_{n > d} \widehat{f}_n \int \phi_n \, dv_N\right)^2 \ll \left(N^\epsilon \sum_{n = d+1}^{N^{\frac{1}{\eta}}} \widehat{f}_n \sigma_n^\eta + N \sum_{n > N^{\frac{1}{\eta}}} \widehat{f}_n\right)^2,$$

up to an error term coming from the complement of the event $\mathcal{B}_\epsilon$ which is exponential in N and that we have neglected. Now, using the fact that $|\widehat{f}_n| \ll \sigma_n^{-\frac{\kappa+3}{2}}$ and $\sigma_n \asymp n$, we obtain

$$\mathbb{E}\left(\sum_{n > d} \widehat{f}_n \int \phi_n \, dv_N\right)^2 \ll N^{2\epsilon} d^{2\eta - \kappa - 1} + N^{2 - \frac{\kappa+1}{\eta}}.$$

Note that if $d \to \infty$ as $N \to \infty$ and $\eta < \frac{\kappa+1}{2}$, both error terms converge to zero and the first term is the largest in the regime $d^\eta \leq N$. Then the claim follows directly from the estimate (6.11) and the fact that $\epsilon > 0$ and $\eta_* = \frac{4(\kappa+1)}{2\kappa - 1} < \eta < \frac{\kappa+1}{2}$ are arbitrary (the previous inequalities are possible only if $\kappa \geq 5$). $\quad\square$

6.3. *Proof of Theorem* 1.2. By the triangle inequality we have, for any $d \in \mathbb{N}$,

$$W_2(\mathcal{X}_N(f), \gamma_1) \leq W_2(\mathcal{X}_N(f), X_N^\infty) + W_2(X_N^d(f), X_N^\infty) + W_2(X_N^d(f), \gamma_1).$$

Note that by choosing

$$d = \left\lfloor N^{\frac{2}{6\eta_* + \kappa + 1}} \right\rfloor,$$

the error terms in Lemmas 6.2 and 6.4 match and are of order $N^{-2\tilde{\theta}+\epsilon}$ for any $\epsilon > 0$ where

$$\tilde{\theta}(\kappa) = \frac{\kappa + 1 - 2\eta_*}{6\eta_* + \kappa + 1}$$

and $\eta_* = \frac{4(\kappa+1)}{2\kappa-1}$. In particular, we easily check that for any $\kappa \geq 5$, $\tilde{\theta}(\kappa) = \frac{2\kappa-9}{2\kappa+11} > 0$. Compared with the error term in Lemma 6.3, coming from the fluctuations of the particles near the edge of the spectrum, we conclude that

$$W_2(\mathcal{X}_N(f), \gamma_1) \ll N^{-\min\{\tilde{\theta}, \frac{2}{3}\}+\epsilon},$$

which completes the proof. Finally, observe that if $\operatorname{supp}(f) \subset J$, then $g_\infty = f$ so that $W_2(\mathcal{X}_N(f), X_N^\infty) = 0$ for all $N \geq 1$. This means that in this case, the rate of convergence is given by

$$W_2(\mathcal{X}_N(f), \gamma_1) \ll N^{-\tilde{\theta}+\epsilon}.$$

Since $\lim_{\kappa \to \infty} \tilde{\theta}(\kappa) = 1$, when f is smooth, this establishes the second claim of Remark 1.3. The proof of Theorem 1.2 is therefore complete. $\square$

APPENDIX: PROPERTIES OF THE HILBERT TRANSFORM, CHEBYSHEV POLYNOMIALS AND THE EQUILIBRIUM MEASURE

In this appendix, we recall some basic properties of the standard Hilbert transform on $\mathbb{R}$ as well as some basic facts about expanding functions on J in terms of Chebyshev polynomials. We also discuss some properties related to the equilibrium measure μ_V. We normalize the Hilbert transform of a function $f \in L^p(\mathbb{R})$ for some $p > 1$ in the following way:

$$(\mathcal{H}_x f) = \operatorname{pv} \int_\mathbb{R} \frac{f(y)}{x - y} \, dy, \tag{A.1}$$

where pv means that we consider the integral as a Cauchy principal value integral. A basic fact about the Hilbert transform is that it is a bounded operator on $L^p(\mathbb{R})$: for finite $p > 1$, there exists a finite constant $C_p > 0$ such that for any $f \in L^p(\mathbb{R})$, $\mathcal{H}f \in L^p(\mathbb{R})$ and

$$\|\mathcal{H}f\|_{L^p(\mathbb{R})} \leq C_p \|f\|_{L^p(\mathbb{R})}. \tag{A.2}$$

See, for example, [14], Theorem 4.1.7, or [34], Theorem 101, for a proof of this fact. With some abuse of notation, if μ is a measure whose Radon–Nikodym derivative with respect to the Lebesgue measure is in $L^p(\mathbb{R})$ for some $p > 1$, we write $\mathcal{H}(\mu)$ for the Hilbert transform of $\frac{d\mu}{dx}$.

Another basic property of the Hilbert transform that we will make use of is its antiself-adjointness: if $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, where $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_\mathbb{R} f(x)(\mathcal{H}_x g) \, dx = - \int_\mathbb{R} (\mathcal{H}_x f) g(x) \, dx. \tag{A.3}$$

For a proof see, for example, [34], Theorem 102.

One of the main objects we study in Section 4 is the (weighted) finite Hilbert transform

$$\mathcal{U}_x \phi = -\mathcal{H}_x(\phi \varrho).$$

Since one can check that $\mathcal{H}_x \varrho = (x^2 - 1)^{-\frac{1}{2}} \mathbf{1}\{|x| > 1\}$, one can actually write for $x \in \mathrm{J}$,

$$(A.4) \qquad \mathcal{U}_x(f) = \int_{\mathrm{J}} \frac{f(t) - f(x)}{t - x} \varrho(dt).$$

We will need to consider derivatives of $\mathcal{U}_x f$ and other similar functions, and for this, we record the following simple result.

LEMMA A.1.   *Let* $\mathrm{I} \subseteq \mathbb{R}$ *be a closed interval,* μ *be a probability measure with compact support,* $\operatorname{supp} \mu \subseteq \mathrm{I}$, *and* $\phi \in \mathscr{C}^{k,\alpha}(\mathrm{I})$ *for some* $0 < \alpha \le 1$. *Then for almost every* $x \in \mathbb{R}$,

$$(A.5) \qquad \left(\frac{d}{dx}\right)^k \int_{\mathrm{I}} F_0(x,t)\mu(dt) = k! \int_{\mathrm{I}} F_k(x,t)\mu(dt),$$

where $F_k(x,t) = \frac{\phi(t) - \Lambda_k[\phi](x,t)}{(t-x)^{k+1}}$; *see* (4.17). *In addition, if* $\alpha = 1$, (A.5) *holds for all* $x \in \mathrm{I}$ *and the RHS is continuous.*

PROOF.   There is nothing to prove when $k = 0$ and since μ has compact support the integral exists. By induction, it suffices to show that for almost every $x \in \mathrm{I}$,

$$(A.6) \qquad \frac{d}{dx} \int_{\mathrm{I}} F_{k-1}(x,t)\mu(dt) = k \int_{\mathrm{I}} F_k(x,t)\mu(dt).$$

Plainly, the function $x \mapsto F_{k-1}(x,t)$ is differentiable for any $t \in \mathrm{I}$ and it is easy to check that its derivative $F'_{k-1}(x,t) = k F_k(x,t)$. Note that for any $u, v \in \mathrm{I}$ with $u < v$, the function $(t,x) \mapsto |x-t|^{\alpha-1}$ is integrable on $\mathbb{R} \times [u,v]$ with respect to $\mu \times dx$, so that by (4.23) and Fubini's theorem,

$$\int_u^v dx \int_{\mathrm{I}} F_k(x,t)\mu(dt) = \frac{1}{k} \int_{\mathrm{I}} \mu(dt) \int_u^v F'_{k-1}(x,t)\,dx$$
$$= \frac{1}{k}\left(\int_{\mathrm{I}} F_{k-1}(v,t)\mu(dt) - \int_{\mathrm{I}} F_{k-1}(u,t)\mu(dt)\right).$$

By the Lebesgue differentiation theorem, we conclude that (A.6) holds. Finally, note that when $\alpha = 1$, $\sup_{x,t \in \mathrm{I}} |F_k(x,t)| \ll 1$, so that the function on the RHS of (A.6) is continuous on $\mathbb{R}$. $\square$

While it might be possible to prove many of the properties of $\mathcal{U}$ and $\mathcal{R}^S$ (see Section 4.4 for the definition) through properties of $\mathcal{H}$ directly, we found it more convenient to expand elements of $\mathscr{H}_{\mu_V}$ in terms of Chebyshev polynomials. We

recall now the definitions of these. For any $k \in \mathbb{N}$, the kth Chebyshev polynomial of the first kind, $T_k(x)$, and the kth Chebyshev polynomial of the second kind, $U_k(x)$, are the unique polynomials for which
(A.7)
$$T_k(\cos\theta) = \cos(k\theta) \quad \text{and} \quad U_k(\cos\theta) = \frac{\sin((k+1)\theta)}{\sin\theta} \qquad \text{for } \theta \in [0, 2\pi].$$

Note that in particular that $T_0' = 0$ and for $k \geq 1$,
(A.8)
$$T_k'(x) = kU_{k-1}(x).$$

These satisfy the basic orthogonality relations (easily checked using orthogonality of Fourier modes):

(A.9)
$$\int_J T_k(x) T_l(x) \varrho(dx) = \begin{cases} \delta_{k,l}, & k = 0, \\ \frac{1}{2}\delta_{k,l}, & k \neq 0, \end{cases} \quad \text{and}$$

$$\int_J U_k(x) U_l(x) \mu_{sc}(dx) = \delta_{k,l},$$

where we used the notation of (1.4). We consider now expanding functions in $L^p(J, \varrho)$ for $p > 1$ in terms of series of Chebyshev polynomials of the first kind. We define for $f \in L^p(J, \varrho)$,

(A.10)
$$f_k = \begin{cases} \int_J f(x) \varrho(dx), & k = 0, \\ 2 \int_J f(x) T_k(x) \varrho(dx), & k > 0. \end{cases}$$

The following result is a direct consequence of the standard L^p-theory of Fourier series; see, for example, [14], Section 3.5, for the basic results about L^p convergence of Fourier series.

LEMMA A.2. *Let $f \in L^p(J, \varrho)$ for some $p > 0$. Then as $N \to \infty$, $\sum_{k=0}^N f_k \times T_k \to f$, where the convergence is in $L^p(J, \varrho)$.*

We will also need to know how T_k and U_k behave under suitable Hilbert transforms.

LEMMA A.3. *For $k \geq 1$ and $x \in J = (-1, 1)$, $\mathcal{U}_x T_k = U_{k-1}(x)$ and $\mathcal{H}_x[U_{k-1}\mu_{sc}] = 2T_k$.*

For a proof see, for example, the discussion around [35], Section 4.3: equation (22). We also recall a closely related result, namely Tricomi's formula for inverting the finite Hilbert transform (i.e., the Hilbert transform of functions on J). As it is in a slightly different form than the standard one (see [35], Section 4.3), we provide a proof.

LEMMA A.4 (Tricomi's formula).   *For any $\phi \in \mathscr{H}$, we have $\phi = \frac{1}{2}\mathcal{H}(\mathcal{U}(\phi) \times \mu_{\mathrm{sc}})$.*

PROOF.   Recall that the Hilbert transform of the semicircle law is given by

(A.11) $$\mathcal{H}_x(\mu_{\mathrm{sc}}) = 2\big(x - \mathbf{1}_{|x|>1}\sqrt{x^2 - 1}\big).$$

For any function $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$ with $\frac{1}{p} + \frac{1}{q} < 1$, the Hilbert transform satisfies the following *convolution property*:

$$\mathcal{H}(f\mathcal{H}(g) + \mathcal{H}(f)g) = \mathcal{H}(f)\mathcal{H}(g) - \pi^2 fg;$$

see [35], formula (14) in Section 4.2. We can apply this formula with $f = \frac{d\mu_{\mathrm{sc}}}{dx}$ and $g = \phi \frac{d\varrho}{dx}$ when $\phi \in \mathscr{H}$. Indeed, since $\phi \in \mathscr{C}(\bar{\mathrm{J}})$ (cf., the estimate (4.7)), the functions $f \in L^p(\mathbb{R})$ for any $p > 1$ and $g \in L^q(\mathbb{R})$ for any $1 < q < 2$. Moreover, note that both f and g have support on $\bar{\mathrm{J}}$ and that $\pi^2 fg = 2\phi$. Now, since $\mathcal{H}(g) = -\mathcal{U}(\phi)$, by formula (A.11), this implies that for almost every $x \in \mathrm{J}$,

(A.12) $$-\mathcal{H}_x\big(f(t)\mathcal{U}_t(\phi) - 2tg(t)\big) = 2x\mathcal{H}_x(g) - 2\phi(x).$$

Since $g = \phi \frac{d\varrho}{dx}$ where $\phi \in \mathscr{H}$, we obtain

$$\mathcal{H}_x\big(tg(t)\big) = \mathrm{pv}\int_{\mathrm{J}} \frac{tg(t)}{x - t}\,dt = -\int_{\mathrm{J}} \phi(t)\varrho(dt) + x\mathcal{H}_x(g)$$

and the first term of the RHS vanishes by definition of the Sobolev space $\mathscr{H}$. We conclude that $\mathcal{H}_x(tg(t)) = x\mathcal{H}_x(g)$ and deduce from (A.12) that for all $x \in \mathrm{J}$,

$$\phi(x) = \frac{1}{2}\mathcal{H}_x\big(\mathcal{U}(\phi)\mu_{\mathrm{sc}}\big) = \mathrm{pv}\int_{\mathrm{J}} \frac{\mathcal{U}_t(\phi)}{x - t}\frac{\sqrt{1 - t^2}}{\pi}\,dt. \qquad \square$$

We will also make use of the following result, whose proof is likely to exist somewhere in the literature, but we give one for completeness.

LEMMA A.5.   *For any function $f \in \mathscr{C}^2(\bar{\mathrm{J}})$ for which $\int_{\mathrm{J}} f(x)\varrho(dx) = 0$, we have*

(A.13) $$\int_{\mathrm{J}} x\mathcal{U}_x(f)\varrho(dx) = \frac{f(1) + f(-1)}{2}.$$

PROOF.   We begin by showing that (A.13) holds when f is a Chebyshev polynomial of the first kind. By Lemma A.3, it suffices to show that for any $k \geq 1$,

(A.14) $$\int_{\mathrm{J}} xU_{k-1}(x)\varrho(dx) = \frac{T_k(1) + T_k(-1)}{2}.$$

Using the identity, $xU_{k-1}(x) = U_k(x) - T_k(x)$ and (A.9), we obtain for any $k \geq 1$,

$$\int_{\mathrm{J}} xU_{k-1}(x)\varrho(dx) = \int_{\mathrm{J}} U_k(x)\varrho(dx).$$

Another standard fact about the Chebyshev polynomials of the second kind is that

$$U_{2k}(x) = 2\sum_{j=0}^{k} T_{2j}(x) - 1 \quad \text{and} \quad U_{2k+1}(x) = 2\sum_{j=0}^{k} T_{2j+1}(x),$$

which combined with (A.9) implies that $\int_J U_k(x)\varrho(dx) = \frac{1}{2}(1 + (-1)^k) = \frac{1}{2}(T_k(1) + T_k(-1))$, concluding the proof of (A.14). The claim in case f is a Chebyshev polynomial of the first kind.

Now we shall prove that formula (A.13) holds for arbitrary functions $f \in \mathscr{C}^2(\bar{J})$ by approximating $\mathcal{U}(f)$ by $\mathcal{U}^N(f) = \sum_{k=1}^{N} f_k U_{k-1}$. By definition,

$$f_k = \frac{2}{\pi}\int_0^\pi f(\cos\theta)\cos(k\theta)\,d\theta = -\frac{2}{\pi k^2}\int_0^\pi f(\cos\theta)\frac{d^2\cos(k\theta)}{d\theta^2}\,d\theta$$

and, an integration by parts shows that $|f_k| \ll k^{-2}$ where the implied constant depends only on the function $f \in \mathscr{C}^2(\bar{J})$. In particular, this estimate shows that the series $f = \sum_{k\geq 1} f_k T_k$ converges uniformly on $\bar{J}$. Thus, we obtain

$$
\begin{aligned}
(A.15) \quad \lim_{N\to\infty}\int_J x\mathcal{U}_x^N(f)\varrho(dx) &= \lim_{N\to\infty}\sum_{k=1}^{N} f_k\frac{T_k(1) + T_k(-1)}{2} \\
&= \frac{f(1) + f(-1)}{2}.
\end{aligned}
$$

On the other hand, using the results of Section 4.1 and Section 4.4, we immediately see that formula (A.13) follows from (A.15) and the dominated convergence theorem. The lemma is established. □

We conclude this appendix with an estimate concerning the equilibrium measure (or its Hilbert transform) that we need in Section 5.1. Recall that we denote by S the density of the equilibrium measure μ_V with respect to the semicircle law μ_{sc}. Here, we assume that S exists and $S(x) > 0$ for all $x \in \bar{J}$. Then it is proved, for example ([29], Theorem 11.2.4) that, if V is sufficiently smooth, then $S = \frac{1}{2}\mathcal{U}(V')$, so that we may extend S to a continuous function which is given by

$$S(x) = \frac{1}{2}\int_J \frac{V'(t) - V'(x)}{t - x}\varrho(dt).$$

Moreover, if the potential $V \in \mathscr{C}^{\kappa+3}(\mathbb{R})$, we obtain by Lemma A.1 that $S \in \mathscr{C}^{\kappa+1}(\mathbb{R})$. Observe that by Tricomi's formula, for all $x \in J$,

$$\mathcal{H}_x(S\mu_{\mathrm{sc}}) = V'(x).$$

This establishes the first of the variational conditions (4.3) and the next proposition shows that the second condition holds as well.

PROPOSITION A.6. *The function $\mathcal{H}(\mu_V) \in \mathscr{C}^\kappa(\mathbb{R})$ and*

$$(A.16) \qquad \mathcal{H}_x(\mu_V) = V'(x) - 2S(x)\operatorname{Re}\left(\sqrt{x^2-1}\right) + \underset{x\to 1}{o}(x-1)^\kappa.$$

PROOF. We have just seen that $\mathcal{H}(\mu_V) = V'$ on J so that (A.16) holds (without the error term) for all $x \in J$. Define for all $x \in \mathbb{R}$,

$$\Phi(x) = \int_J \frac{S(t) - S(x)}{t - x} \, d\mu_{\mathrm{sc}}(t) - 2x\,S(x).$$

Since $d\mu_V = S\,d\mu_{\mathrm{sc}}$, by (A.11), we obtain for all $x \in \mathbb{R}$,

$$\mathcal{H}_x(\mu_V) = -\Phi(x) - 2\operatorname{Re}\left(\sqrt{x^2-1}\right)S(x).$$

Since the density $S \in \mathscr{C}^{\kappa+1}(\mathbb{R})$, by Lemma A.1 the function $\Phi \in \mathscr{C}^\kappa(\mathbb{R})$ and $\Phi = -V'$ on J. By continuity, this implies that

$$\Phi(x) = -V'(x) + \underset{x\to 1}{o}(x-1)^\kappa.$$

This completes the proof of (A.16). $\square$

REFERENCES

[1] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[2] BAKRY, D., GENTIL, I. and LEDOUX, M. (2014). *Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Springer, Cham. MR3155209

[3] BEKERMAN, F., FIGALLI, A. and GUIONNET, A. (2015). Transport maps for β-matrix models and universality. *Comm. Math. Phys.* **338** 589–619. MR3351052

[4] BEKERMAN, F., LEBLÉ, T. and SERFATY, S. (2017). CLT for fluctuations of β-ensembles with general potential. Preprint. Available at arXiv:1706.09663.

[5] BOROT, G. and GUIONNET, A. (2013). Asymptotic expansion of β matrix models in the one-cut regime. *Comm. Math. Phys.* **317** 447–483. MR3010191

[6] BOROT, G. and GUIONNET, A. (2013). Asymptotic expansion of beta matrix models in the multi-cut regime. Preprint. Available at arXiv:1303.1045.

[7] BOURGADE, P., ERDÖS, L. and YAU, H.-T. (2014). Edge universality of beta ensembles. *Comm. Math. Phys.* **332** 261–353. MR3253704

[8] BREUER, J. and DUITS, M. (2017). Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. *J. Amer. Math. Soc.* **30** 27–66. MR3556288

[9] CABANAL-DUVILLARD, T. (2001). Fluctuations de la loi empirique de grandes matrices aléatoires. *Ann. Inst. Henri Poincaré Probab. Stat.* **37** 373–402. MR1831988

[10] CHATTERJEE, S. (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* **143** 1–40. MR2449121

[11] DÖBLER, C. and STOLZ, M. (2011). Stein's method and the multivariate CLT for traces of powers on the classical compact groups. *Electron. J. Probab.* **16** 2375–2405. MR2861678

[12] DÖBLER, C. and STOLZ, M. (2014). A quantitative central limit theorem for linear statistics of random matrix eigenvalues. *J. Theoret. Probab.* **27** 945–953. MR3245992

[13] FULMAN, J. (2012). Stein's method, heat kernel, and traces of powers of elements of compact Lie groups. *Electron. J. Probab.* **17** no. 66, 16. MR2968673

[14] GRAFAKOS, L. (2008). *Classical Fourier Analysis*, 2nd ed. *Graduate Texts in Mathematics* **249**. Springer, New York. MR2445437

[15] JOHANSSON, K. (1997). On random matrices from the compact classical groups. *Ann. of Math.* (2) **145** 519–545. MR1454702

[16] JOHANSSON, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** 151–204. MR1487983

[17] KRIECHERBAUER, T. and SHCHERBINA, M. (2010). Fluctuations of eigenvalues of matrix models and their applications. Preprint. Available at arXiv:1003.6121.

[18] LAMBERT, G. (2018). Limit theorems for biorthogonal ensembles and related combinatorial identities. *Adv. Math.* **329** 590–648. MR3783424

[19] LEBLÉ, T. and SERFATY, S. (2017). Large deviation principle for empirical fields of log and Riesz gases. *Invent. Math.* **210** 645–757. MR3735628

[20] LEDOUX, M. (2012). Chaos of a Markov operator and the fourth moment condition. *Ann. Probab.* **40** 2439–2459. MR3050508

[21] LEDOUX, M., NOURDIN, I. and PECCATI, G. (2015). Stein's method, logarithmic Sobolev and transport inequalities. *Geom. Funct. Anal.* **25** 256–306. MR3320893

[22] LEDOUX, M. and POPESCU, I. (2013). The one dimensional free Poincaré inequality. *Trans. Amer. Math. Soc.* **365** 4811–4849. MR3066771

[23] MASON, J. C. and HANDSCOMB, D. C. (2003). *Chebyshev Polynomials*. Chapman & Hall/CRC Press, Boca Raton, FL. MR1937591

[24] MECKES, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V*: *The Luminy Volume*. *Inst. Math. Stat.* (*IMS*) *Collect.* **5** 153–178. IMS, Beachwood, OH. MR2797946

[25] MEHTA, M. L. (2004). *Random Matrices*, 3rd ed. *Pure and Applied Mathematics* (*Amsterdam*) **142**. Elsevier/Academic Press, Amsterdam. MR2129906

[26] NOURDIN, I. and PECCATI, G. (2012). *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge Univ. Press, Cambridge. MR2962301

[27] OTTO, F. and VILLANI, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620

[28] PASTUR, L. (2006). Limiting laws of linear eigenvalue statistics for Hermitian matrix models. *J. Math. Phys.* **47** 103303, 22. MR2268864

[29] PASTUR, L. and SHCHERBINA, M. (2011). *Eigenvalue Distribution of Large Random Matrices. Mathematical Surveys and Monographs* **171**. Amer. Math. Soc., Providence, RI. MR2808038

[30] REED, M. and SIMON, B. (1980). *Methods of Modern Mathematical Physics*. I. *Functional Analysis*, 2nd ed. Academic Press, New York. MR0751959

[31] SHCHERBINA, M. (2013). Fluctuations of linear eigenvalue statistics of β matrix models in the multi-cut regime. *J. Stat. Phys.* **151** 1004–1034. MR3063494

[32] SHCHERBINA, M. (2014). Change of variables as a method to study general β-models: Bulk universality. *J. Math. Phys.* **55** 043504, 23. MR3390602

[33] STEIN, C. (1995). The accuracy of the normal approximation to the distribution of the traces of powers of random orthogonal matrices. Technical Report 470, Dept. Statistics, Stanford Univ., Stanford, CA.

[34] TITCHMARSH, E. C. (1986). *Introduction to the Theory of Fourier Integrals*, 3rd ed. Chelsea, New York. MR0942661

[35] TRICOMI, F. G. (1985). *Integral Equations*. Dover, New York. MR0809184

[36] VILLANI, C. (2009). *Optimal Transport. Old and New. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Springer, Berlin. MR2459454

[37] WEBB, C. (2016). Linear statistics of the circular β-ensemble, Stein's method, and circular Dyson Brownian motion. *Electron. J. Probab.* **21** Paper No. 25, 16. MR3485367

G. LAMBERT
INSTITUT FÜR MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURERSTRASSE 190
CH-8057 ZÜRICH
SWITZERLAND
E-MAIL: gaultier.lambert@math.uzh.ch

M. LEDOUX
INSTITUT DE MATHÉMATIQUES DE TOULOUSE
UNIVERSITÉ DE TOULOUSE—PAUL-SABATIER
F-31062 TOULOUSE
FRANCE
AND
INSTITUT UNIVERSITAIRE DE FRANCE
E-MAIL: ledoux@math.univ-toulouse.fr

C. WEBB
DEPARTMENT OF MATHEMATICS
   AND SYSTEMS ANALYSIS
AALTO UNIVERSITY
P.O. BOX 11000
00076 AALTO
FINLAND
E-MAIL: christian.webb@aalto.fi