Small deviations for beta ensembles

Michel Ledoux   Brian Rider

July 15, 2010

Abstract

We establish various small deviation inequalities for the extremal (soft edge) eigenvalues in the β-Hermite and β-Laguerre ensembles. In both settings, upper bounds on the variance of the largest eigenvalue of the anticipated order follow immediately.

1 Introduction

In the context of their original discovery, the Tracy-Widom laws describe the fluctuations of the limiting largest eigenvalues in the Gaussian Orthogonal, Unitary, and Symplectic Ensembles (G{O/U/S}E) [23, 24]. These are random matrices of real, complex, or quaternion Gaussian entries, of mean zero and mean-square one, independent save for the condition that the matrix is symmetric (GOE), Hermitian (GUE), or appropriately self-dual (GSE). The corresponding Tracy-Widom distribution functions have shape

$$F_{TW}(t) \sim e^{\frac{1}{24}\beta t^3} \text{ as } t \to -\infty, \quad 1 - F_{TW}(t) \sim e^{-\frac{2}{3}\beta t^{3/2}} \text{ as } t \to \infty, \tag{1.1}$$

where $\beta = 1$ in the case of GOE, $\beta = 2$ for GUE, and $\beta = 4$ for GSE.

Since that time, it has become understood that the three Tracy-Widom laws arise in a wide range of models. First, the assumption of Gaussian entries may be relaxed significantly, see [21], [22] for instance. Outside of random matrices, these laws also describe the fluctuations in the longest increasing subsequence of a random permutation [2], the path weight in last passage percolation [11], and the current in simple exclusion [11, 25], among others.

It is natural to inquire as to the rate of concentration of these various objects about the limiting Tracy-Widom laws. Back in the random matrix setting, the limit theorem reads: with $\lambda_{\max}$ the largest eigenvalue in the $n \times n$ GOE, GUE or GSE, it is the normalized quantity $n^{1/6}(\lambda_{\max} - 2\sqrt{n})$ which converges to Tracy-Widom. Thus, one would optimally hope for estimates of the form:

$$\mathbb{P}\Big(\lambda_{\max} - 2\sqrt{n} \le -\varepsilon\sqrt{n}\Big) \le Ce^{-n^2\varepsilon^3/C}, \quad \mathbb{P}\Big(\lambda_{\max} - 2\sqrt{n} \ge \varepsilon\sqrt{n}\Big) \le Ce^{-n\varepsilon^{3/2}/C},$$

for all $n \geq 1$, all $\varepsilon \in (0, 1]$ say, and C a numerical constant. Such are "small deviation" inequalities, capturing exactly the finite n scaling and limit distribution shape (compare (1.1)). Taking ε beyond $O(1)$ in the above yields more typical large deviation behavior and different (Gaussian) tails (see below).

As discussed in [14, 15], the right-tail inequality for the GUE (as well as for the Laguerre Unitary Ensemble, again see below) may be shown to follow from results of Johansson [11] for a more general invariant model related to the geometric distribution that uses large deviation asymptotics and sub-additivity arguments. The left-tail inequality for the geometric model of Johansson (and thus by some suitable limiting procedure for the GUE and the Laguerre Unitary Ensemble) is established in [3] together with convergence of moments using delicate Riemann-Hilbert methods. We refer to [15] for a discussion and the relevant references, as well as for similar inequalities in the context of last passage percolation *etc.* By the superposition-decimation procedure of [10], the GUE bounds apply similarly to the GOE (see also [16]).

Our purpose here is to present unified proofs of these bounds which apply to all of the so-called beta ensembles. These are point-processes on $\mathbb{R}$ defined by the n-level joint density: for any $\beta > 0$,

$$\mathbb{P}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \frac{1}{Z_{n,\beta}} \prod_{j<k} |\lambda_j - \lambda_k|^\beta e^{-(\beta/4)\sum_{k=1}^n \lambda_k^2}. \tag{1.2}$$

At $\beta = 1, 2, 4$ this joint density is shared by the eigenvalues of G{O/U/S}E. Furthermore, these three values give rise to exactly solvable models. Specifically, all finite dimensional correlation functions may be described explicitly in terms of Hermite polynomials. For this reason, the measure (1.2) has come to be referred to the β-Hermite ensemble; we will denote it by H_β. Importantly, off of $\beta = 1, 2, 4$, despite considerable efforts (see [9], Chapter 13 for a comprehensive review), there appears to be no characterization of the correlation functions amenable to asymptotics. Still, Ramírez-Rider-Virág [19] have shown the existence of a general β Tracy-Widom law, TW_β, via the corresponding limit theorem: with self-evident notation,

$$n^{1/6}\big(\lambda_{\max}(H_\beta) - 2\sqrt{n}\,\big) \Rightarrow TW_\beta. \tag{1.3}$$

This result makes essential use of a (tridiagonal) matrix model valid at all beta due to Dumitriu-Edelman [5], and proves the conjecture of Edelman-Sutton [6]. As to finite n bounds, we have:

Theorem 1. *For all $n \geq 1$, $0 < \varepsilon \leq 1$ and $\beta \geq 1$:*

$$\mathbb{P}\Big(\lambda_{\max}(H_\beta) \geq 2\sqrt{n}(1+\varepsilon)\Big) \leq Ce^{-\beta n \varepsilon^{3/2}/C},$$

and

$$\mathbb{P}\Big(\lambda_{\max}(H_\beta) \leq 2\sqrt{n}(1-\varepsilon)\Big) \leq C^\beta e^{-\beta n^2 \varepsilon^3/C},$$

where C is a numerical constant.

The restriction to $\beta \geq 1$ is somewhat artificial, though note that bounds of this type cannot remain meaningful all the way down to $\beta = 0$. Our methods do extend, with some caveats, to $\beta < 1$.

Theorem 1$'$. *When $0 < \beta < 1$ upper bounds akin to those in Theorem 1 hold as soon as $n \geq 2\beta^{-1}$, with the right hand sides reading $(1 - e^{-\beta/C})^{-1} e^{-\beta n \varepsilon^{3/2}/C}$ for the right tail and $C e^{-\beta n^2 \varepsilon^3/C}$ for the left tail. A right tail upper bound is available without the restriction on n, but with the right hand side replaced by $(1 - e^{-\beta^2/C})^{-1} e^{-\beta^{3/2} n \varepsilon^{3/2}/C}$.*

Our remaining results will have like extensions to $\beta < 1$. We prefer though to restrict the statements to $\beta \geq 1$, which covers the cases of classical interest and allows for cleaner, more unified proofs.

At this point we should also mention that for ε beyond $O(1)$, the large-deviation right-tail inequality takes the form

$$\mathbb{P}\Big(\lambda_{\max}(H_\beta) \geq 2\sqrt{n}(1+\varepsilon)\Big) \leq C e^{-\beta n \varepsilon^2/C}. \tag{1.4}$$

For $\beta = 1$ and 2 this follows from standard net arguments on the corresponding Gaussian matrices (see e.g. [15]). For other values of β (this again for $\beta \geq 1$), crude bounds on the tridiagonal models discussed below immediately yield the claim.

Continuing, those well versed in random matrix theory will know that this style of small deviation questions are better motivated in the context of "null" Wishart matrices, given their application in multivariate statistics. Also known as the Laguerre Orthogonal or Unitary Ensembles (L{O/U}E), these are ensembles of type XX^* in which X is an $n \times \kappa$ matrix comprised of i.i.d. real or complex Gaussians.

By the obvious duality, we may assume here that $\kappa \geq n$. When $n \to \infty$ with the κ/n converging to a finite constant (necessarily larger than one), the appropriately centered and scaled largest eigenvalue was shown to converge to the natural Tracy-Widom distribution; first by Johansson [11] in the complex ($\beta = 2$) case, then by Johnstone [12] in the real ($\beta = 1$) case. Later, El Karoui [7] proved the same conclusion allowing $\kappa/n \to \infty$.

For $\beta = 2$ and κ a fixed multiple of n, a small deviation upper bound at the right-tail (as well as the corresponding statement for the minimal eigenvalue in the "soft-edge" scaling) was known earlier (see [14, 15]), extended recently to non-Gaussian matrices in [8].

Once again there is a general beta version. Consider a density of the form (1.2) in which the Gaussian weight $w(\lambda) = e^{-\beta\lambda^2/4}$ on $\mathbb{R}$ is replaced by $w(\lambda) = \lambda^{(\beta/2)(\kappa-n+1)+1}\, e^{-\beta\lambda/2}$, now restricted to $\mathbb{R}_+$. Here κ can be any real number strictly larger than $n-1$. It is when κ is an integer and $\beta = 1$ or 2 that one recovers the eigenvalue law for the real or complex Wishart matrices just described. For general κ and $\beta > 0$ the resulting law on positive points $\lambda_1, \dots, \lambda_n$ is referred to as the β-Laguerre ensemble, here L_β for short.

Using a tridiagonal model for L_β introduced in [5], it is proved in [19]: for $\kappa+1 > n \to \infty$ with $\kappa/n \to c \geq 1$,

$$\frac{(\sqrt{\kappa n})^{1/3}}{(\sqrt{\kappa}+\sqrt{n})^{4/3}}\Big(\lambda_{\max}(L_\beta) - (\sqrt{\kappa}+\sqrt{n})^2\Big) \Rightarrow TW_\beta. \tag{1.5}$$

This covers all previous results for real/complex null Wishart matrices. Comparing (1.3) and (1.5) one sees that $O(n^{2/3}\varepsilon)$ deviations in the Hermite case should correspond to deviations of order $(\kappa n)^{1/6}(\sqrt{\kappa}+\sqrt{n})^{2/3}\varepsilon = O(\kappa^{1/2}n^{1/6}\varepsilon)$ in the Laguerre case. That is, one might expect bounds exactly of the form found in Theorem 1 with appearances of n in each exponent replaced by $\kappa^{3/4}n^{1/4}$. What we have is the following.

Theorem 2. *For all $\kappa + 1 > n \geq 1$, $0 < \varepsilon \leq 1$ and $\beta \geq 1$:*

$$\mathbb{P}\Big(\lambda_{\max}(L_\beta) \geq (\sqrt{\kappa}+\sqrt{n})^2(1+\varepsilon)\Big) \leq Ce^{-\beta\sqrt{n\kappa}\varepsilon^{3/2}(\frac{1}{\sqrt{\varepsilon}}\wedge(\frac{\kappa}{n})^{1/4})/C},$$

and

$$\mathbb{P}\Big(\lambda_{\max}(L_\beta) \leq (\sqrt{\kappa}+\sqrt{n})^2(1-\varepsilon)\Big) \leq C^\beta e^{-\beta n\kappa\varepsilon^3(\frac{1}{\varepsilon}\wedge(\frac{\kappa}{n})^{1/2})/C}.$$

Again, C is some numerical constant.

The right-tail inequality is extended to non-Gaussian matrices in [8]. The rather cumbersome exponents in Theorem 2 do produce the anticipated decay, though only for $\varepsilon \leq \sqrt{n/\kappa}$. For $\varepsilon \geq \sqrt{n/\kappa}$, the right and left-tails become linear and quadratic in ε respectively. This is to say that the large deviation regime begins at the order $O(\sqrt{n/\kappa})$ rather than $O(1)$ as in the β-Hermite case. To understand this, we recall that, normalized by $1/\kappa$, the counting measure of the L_β points is asymptotically supported on the interval with endpoints $(1 \pm \sqrt{n/\kappa})^2$. This statement is precise with convergent n/κ, and the limiting measure that of Marčenko-Pastur. Either way, $\sqrt{n/\kappa}$ is identified as the spectral width, in contrast with the semi-circle law appearing in the β-Hermite case which is of width one (after similar normalization). Of course, in the more usual set-up when $c_1 n \leq \kappa \leq c_2 n$ ($c_1 \geq 1$ necessarily) all this is moot: the exponents above may then be replaced with $-\beta n\varepsilon^{3/2}/C$ and $-\beta n^2\varepsilon^3/C$ for ε in an $O(1)$ range with no loss of accuracy. And again, the large deviation tails were known in this setting for $\beta = 1, 2$.

An immediate consequence of the preceding is a finite n (and/or κ) bound on the variance of $\lambda_{\max}$ in line with the known limit theorems. This simple fact had only previously been available for GUE and LUE (see the discussion in [15]).

Corollary 3. *Take $\beta \geq 1$. Then,*

$$\mathrm{Var}\Big[\lambda_{\max}(H_\beta)\Big] \leq C_\beta\, n^{-1/3}, \quad \mathrm{Var}\Big[\lambda_{\max}(L_\beta)\Big] \leq C_\beta\, \kappa n^{-1/3} \tag{1.6}$$

with now constant(s) C_β dependent upon β. (By Theorem $1'$, the Hermite bound holds for $\beta < 1$ as well.)

The same computation behind Corollary 3 implies that

$$\limsup_{n\to\infty} n^{p/6}\mathbb{E}\big|\lambda_{\max}(H_\beta) - 2\sqrt{n}\big|^p < \infty$$

for any p, and similarly for $\lambda_{\max}(L_\beta)$. Hence, we also conclude that all moments of the (scaled) maximal H_β and L_β eigenvalues converge to those for the TW_β laws (see [3] for $\beta = 2$).

Finally, there is the matter of whether any of the above upper bounds are tight. We answer this in the affirmative in the Hermite setting.

Theorem 4. *There is a numerical constant C so that*

$$\mathbb{P}\Big(\lambda_{\max}(H_\beta) \geq 2\sqrt{n}(1+\varepsilon)\Big) \geq C^{-\beta} e^{-C\beta n\varepsilon^{3/2}},$$

and

$$\mathbb{P}\Big(\lambda_{\max}(H_\beta) \leq 2\sqrt{n}(1-\varepsilon)\Big) \geq C^{-\beta} e^{-C\beta n^2\varepsilon^3}.$$

The first inequality holds for all $n > 1, 0 < \varepsilon \leq 1$, and $\beta \geq 1$. For the second inequality, the range of ε must be kept sufficiently small, $0 < \varepsilon \leq 1/C$ say.

Our proof of the right-tail lower bound takes advantage of a certain independence in the β-Hermite tridiagonals not immediately shared by the Laguerre models, but the basic strategy also works in the Laguerre case. Contrariwise, our proof of the left-tail lower bound uses a fundamentally Gaussian argument that is not available in the Laguerre setting.

The next section introduces the tridiagonal matrix models and gives an indication of our approach. The upper bounds (Theorems 1, 1′, 2 and Corollary 3) are proved in Section 3; the H_β lower bounds in Section 4. Section 5 considers the analog of the right-tail upper bound for the minimal eigenvalue in the β-Laguerre ensemble, this case holding the potential for some novelty granted the existence of a different class of limit theorems (hard edge) depending on the limiting ratio n/κ. While our method does produce a bound, the conditions on the various parameters are far from optimal. For this reason we relegate the statement, along with the proof and further discussion, to a separate section.

2 Tridiagonals

The results of [19] identify the general $\beta > 0$ Tracy-Widom law through a random variational principle:

$$TW_\beta = \sup_{f\in L}\left\{ \frac{2}{\sqrt{\beta}}\int_0^\infty f^2(x)db(x) - \int_0^\infty \big[(f'(x))^2 + xf^2(x)\big]dx \right\}, \qquad (2.1)$$

in which $x \mapsto b(x)$ is a standard Brownian motion and L is the space of functions f which vanish at the origin and satisfy $\int_0^\infty f^2(x)dx = 1$, $\int_0^\infty [(f'(x))^2 + xf^2(x)]dx < \infty$. The equality here is in law, or you may view (2.1) as the definition of TW_β.

This variational point of view also guides the proof of the convergence of the centered and scaled $\lambda_{\max}$ (of H_β or L_β) to TW_β. In particular, given the random tridiagonals which we are about to introduce, one always has a characterization of $\lambda_{\max}$ through Raleigh-Ritz. In [19], the point is to show this "discrete" variational problem goes over to the continuum problem (2.1) in a suitable sense. Furthermore, an analysis of the continuum problem has been shown to give sharp estimates on the tails of the β Tracy-Widom law (again see [19]). Our idea here is therefore retool those arguments for the finite n, or discrete, setting.

We start with the Hermite case. Let $g_1, g_2, \dots g_n$ be independent Gaussians with mean 0 and variance 2. Let also χ_β, $\chi_{2\beta}$, $\dots$, $\chi_{(n-1)\beta}$ be independent χ random variables of the indicated parameter. Then, re-using notation, [5] proves that the n eigenvalues of the random tridiagonal matrix

$$H_\beta = \frac{1}{\sqrt{\beta}} \begin{bmatrix} g_1 & \chi_{\beta(n-1)} & & & \\ \chi_{\beta(n-1)} & g_2 & \chi_{\beta(n-2)} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{\beta 2} & g_{n-1} & \chi_\beta \\ & & & \chi_\beta & g_n \end{bmatrix}$$

have joint law (1.2).[1] Centering appropriately, we define: for $v = (v_1, \dots, v_n) \in \mathbb{R}^n$,

$$\begin{aligned} H(v) &= v^T[H_\beta - 2\sqrt{n} I_n]v \qquad (2.2) \\ &= \frac{1}{\sqrt{\beta}} \sum_{k=1}^n g_k v_k^2 + \frac{2}{\sqrt{\beta}} \sum_{k=1}^{n-1} \chi_{\beta(n-k)} v_k v_{k+1} - 2\sqrt{n} \sum_{k=1}^n v_k^2. \end{aligned}$$

The problem at hand (Theorem 1) then becomes that of estimating

$$\mathbb{P}\Big(\sup_{||v||_2 = 1} H(v) \ge \sqrt{n}\varepsilon \Big) \quad \text{and} \quad \mathbb{P}\Big(\sup_{||v||_2 = 1} H(v) \le -\sqrt{n}\varepsilon \Big), \qquad (2.3)$$

where we have introduced the usual Euclidean norm $||v||_2^2 = \sum_{k=1}^n v_k^2$. To make the connection between $H(v)$ and the continuum form (2.1) even more plain we have the following.

Lemma 5. *For any $c > 0$ define*

$$\begin{aligned} H_c(v) &= \frac{1}{\sqrt{\beta}} \sum_{k=1}^n g_k v_k^2 + \frac{2}{\sqrt{\beta}} \sum_{k=1}^{n-1} \big[\chi_{\beta(n-k)} - \mathbb{E}(\chi_{\beta(n-k)}) \big] v_k v_{k+1} \qquad (2.4) \\ & \quad - c\sqrt{n} \sum_{k=0}^n (v_{k+1} - v_k)^2 - \frac{c}{\sqrt{n}} \sum_{k=1}^n k\, v_k^2 \end{aligned}$$

[1] For $\beta = 1$ or 2 this can be seen by applying Householder transformation to the "full" GOE or GUE matrices, and appears to have been used first in a random matrix theory context by Trotter [26].

in which it is understood that $v_0 = v_{n+1} = 0$*. There exist numerical constants,* $a > b > 0$*, so that*

$$H_a(v) \le H(v) \le H_b(v) \quad \textit{for all } v \in \mathbb{R}^n, \tag{2.5}$$

granted $\beta \ge 1$.

We defer the proof until the end of the section, after a description of the allied L_β set-up. The point of Lemma 5 should be clear: for an upper bound on the first probability in (2.3) one may replace H by H_b with any sufficiently small $b > 0$, and so on. Lemma 5 also marks our first run-in with issues surrounding $\beta \ge 1$ versus $\beta < 1$. An available extension to $\beta < 1$ reads:

Lemma 5′. *If* $0 < \beta < 1$, *estimates of type* (2.5) *with numerical* a *and* b *hold whenever* $n \ge 2\beta^{-1}$. *This condition on* n *may be removed in the upper bound at the cost of choosing* $b = const.\, \beta^{1/2}$.

The model for L_β is as follows. For $\kappa > n-1$, introduce the random bidiagonal matrix

$$B_\beta = \frac{1}{\sqrt{\beta}} \begin{bmatrix} \chi_{\beta\kappa} & & & & \\ \widetilde{\chi}_{\beta(n-1)} & \chi_{\beta(\kappa-1)} & & & \\ & \ddots & \ddots & & \\ & & \widetilde{\chi}_{\beta 2} & \chi_{\beta(\kappa-n+2)} & \\ & & & \widetilde{\chi}_{\beta} & \chi_{\beta(\kappa-n+1)} \end{bmatrix},$$

with the same definition for the χ's and again all variables independent. (The use of $\widetilde{\chi}$ is meant to emphasize this independence between the diagonals.) Now [5] shows that it is the eigenvalues of $L_\beta = (B_\beta)(B_\beta)^T$ which have the required joint density.[2] Note that L_β does not have independent entries.

Similar to before, we define

$$\begin{aligned} \sqrt{\kappa} L(v) &= v^T \big(L_\beta - (\sqrt{\kappa} + \sqrt{n})^2 I_n \big) v \\ &= \frac{1}{\beta} \sum_{k=1}^{n} \chi^2_{\beta(\kappa-k+1)} v_k^2 + \frac{1}{\beta} \sum_{k=2}^{n} \widetilde{\chi}^2_{\beta(n-k+1)} v_k^2 \\ &\quad + \frac{2}{\beta} \sum_{k=1}^{n-1} \chi_{\beta(\kappa-k+1)} \widetilde{\chi}_{\beta(n-k)} v_k v_{k+1} - (\sqrt{\kappa} + \sqrt{n})^2 \sum_{k=1}^{n} v_k^2. \end{aligned}$$

The added normalization by $\sqrt{\kappa}$ makes for better comparison with the Hermite case. With this, and since $\kappa > n-1$, to prove Theorem 2 is to establish bounds on the following analogs of (2.3):

$$\mathbb{P}\Big(\sup_{||v||_2=1} L(v) \ge \sqrt{n}\varepsilon \Big) \quad \text{and} \quad \mathbb{P}\Big(\sup_{||v||_2=1} L(v) \le -\sqrt{n}\varepsilon \Big) \tag{2.6}$$

[2] Once again, at $\beta = 1, 2$ this connection had been noted previously (via Householder), see [20] for example.

Finally, we state the Laguerre version of Lemma 5. (We prove only the latter as they are much the same).

Lemma 6. *For $c > 0$ set*

$$\begin{aligned} L_c(v) &= \frac{1}{\sqrt{\beta}} \sum_{k=1}^{n} Z_k v_k^2 + \frac{1}{\sqrt{\beta}} \sum_{k=2}^{n} \widetilde{Z}_k v_k^2 + \frac{2}{\sqrt{\beta}} \sum_{k=1}^{n-1} Y_k v_k v_{k+1} \\ & -c\sqrt{n} \sum_{k=0}^{n} (v_{k+1} - v_k)^2 - \frac{c}{\sqrt{n}} \sum_{k=1}^{n} k v_k^2, \end{aligned}$$

where

$$\begin{aligned} & Z_k = \frac{1}{\sqrt{\beta\kappa}} \big(\chi^2_{\beta(\kappa-k+1)} - \beta(\kappa - k + 1)\big), \quad \widetilde{Z}_k = \frac{1}{\sqrt{\beta\kappa}} \big(\widetilde{\chi}^2_{\beta(n-k+1)} - \beta(n - k + 1)\big), \\ & \text{and } Y_k = \frac{1}{\sqrt{\beta\kappa}} \big(\chi_{\beta(\kappa-k+1)} \widetilde{\chi}_{\beta(n-k)} - \mathbb{E}[\chi_{\beta(\kappa-k+1)} \widetilde{\chi}_{\beta(n-k)}]\big). \end{aligned} \tag{2.7}$$

Then, for all $\beta \geq 1$ there are constants $a > b > 0$ so that $L_a(v) \leq L(v) \leq L_b(v)$ for all $v \in \mathbb{R}^n$.

Proof of Lemma 5. Writing,

$$\begin{aligned} H(v) &= \frac{1}{\sqrt{\beta}} \sum_{k=1}^{n} g_k v_k^2 + \frac{2}{\sqrt{\beta}} \sum_{k=1}^{n-1} \big(\chi_{\beta(n-k)} - \mathbb{E}[\chi_{\beta(n-k)}]\big) v_k v_{k+1} \\ & - \sum_{k=1}^{n-1} \mathbb{E}[\frac{\chi_{\beta(n-k)}}{\sqrt{\beta}}] (v_{k+1} - v_k)^2 \\ & - \sum_{k=1}^{n-1} \big(\sqrt{n} - \mathbb{E}[\frac{\chi_{\beta(n-k)}}{\sqrt{\beta}}]\big) (v_k^2 + v_{k+1}^2) - \sqrt{n}\,(v_1^2 + v_n^2) \end{aligned}$$

shows it is enough to compare, for every v,

$$I(v) = \sqrt{n} \sum_{k=1}^{n-1} (v_{k+1} - v_k)^2 + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} k v_k^2$$

and

$$J(v) = \sum_{k=1}^{n-1} \mathbb{E}[\frac{\chi_{\beta(n-k)}}{\sqrt{\beta}}] (v_{k+1} - v_k)^2 + \sum_{k=1}^{n-1} \big(\sqrt{n} - \mathbb{E}[\frac{\chi_{\beta(n-k)}}{\sqrt{\beta}}]\big) (v_k^2 + v_{k+1}^2).$$

(We implicitly assume here that $n > 1$; for $n = 1$ there is nothing to do.) For this, there is the formula $\mathbb{E}\chi_r = 2^{1/2} \frac{\Gamma(r/2+1/2)}{\Gamma(r/2)}$. By Jensen's inequality we have the upper bound $\mathbb{E}\chi_r \leq \sqrt{r}$ for any $r > 0$, while

$$\mathbb{E}\chi_r \geq \sqrt{r - 1/2}, \quad \text{for } r \geq 1, \tag{2.8}$$

see (2.8) of [17].

These bounds easily translate to

$$\frac{k}{2\sqrt{n}} \le \sqrt{n} - \frac{1}{\sqrt{\beta}}\,\mathbb{E}\chi_{\beta(n-k)} \le \frac{2k}{\sqrt{n}}\,, \tag{2.9}$$

for all $k \le n-1$ and $\beta \ge 1$. It is immediate from the second inequality that $J(v) \le 4I(v)$ for every v. Next, if $k \le n/2$, $\mathbb{E}[\chi_{\beta(n-k)}/\sqrt{\beta}] \ge \sqrt{n}/4$ while if $k \ge n/2$, $\sqrt{n} - \mathbb{E}[\chi_{\beta(n-k)}/\sqrt{\beta}] \ge \sqrt{n}/4$. By splitting $J(v)$ accordingly one can also see that $J(v) \ge I(v)/16$. $\square$

Proof of Lemma $5'$. The issue is the lower bound (2.8). For $r < 1$ this may be replaced by

$$\mathbb{E}\chi_r \ge \frac{r}{\sqrt{1+r}}, \tag{2.10}$$

valid for all $r > 0$ (this is due to Wendel, see now (2.2) of [17]).

In bounding $J(v)$ above, it is the second inequality of (2.9) that is affected for $\beta(n-k) < 1$. We still wish it to hold, with perhaps the 2 replaced by some other constant C. That is, making use of (2.10) we want a constant C so that

$$\sqrt{n} \le C\sqrt{n} + \beta(n-k)\left(\frac{1}{\sqrt{2\beta}} - \frac{C}{\sqrt{n\beta}}\right),$$

and we can take $C = 1$ if $n > 2\beta^{-1}$.

For the lower bound on $J(v)$, note that $\mathbb{E}[\chi_{\beta(n-k)}/\sqrt{\beta}] \ge \sqrt{n}/4$ for $k \le n/2$ still holds (*i.e.* we can still use (2.8)) when $n > 2\beta^{-1}$ and so everything is as before. On the other hand, (2.10) provides $\mathbb{E}[\chi_{\beta(n-k)}/\sqrt{\beta}] \ge (\sqrt{\beta}/2)\sqrt{n}$ on that same range, and so we always have $J(v) \ge bI(v)$ with $b \sim \sqrt{\beta}$. $\square$

3 Upper Bounds

Theorems 1 and 2 are proved, first for the β-Hermite case with all details present (and comments on Theorem $1'$ made along the way); a second subsection explains the modifications required for the β-Laguerre case. The proof of Corollary 3 appears at the end.

3.1 Hermite ensembles

Right-tail. This is the more elaborate of the two. The following is a streamlined version of what is needed.

Proposition 7. *Consider the model quadratic form,*

$$H_b(v,z) = \frac{1}{\sqrt{\beta}}\sum_{k=1}^{n} z_k v_k^2 - b\sqrt{n}\sum_{k=0}^{n}(v_{k+1}-v_k)^2 - \frac{b}{\sqrt{n}}\sum_{k=0}^{n} k v_k^2, \tag{3.1}$$

for fixed $b > 0$ and independent mean-zero random variables $\{z_k\}_{k=1,\dots,n}$ satisfying the uniform tail bound $\mathbb{E}[e^{\lambda z_k}] \le e^{c\lambda^2}$ for all $\lambda \in \mathbb{R}$ and some $c > 0$. There is a $C = C(b,c)$ so that

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_b(v,z) \ge \varepsilon\sqrt{n}\Big) \le (1 - e^{-\beta/C})^{-1} e^{-\beta n \varepsilon^{3/2}/C}$$

for all $\varepsilon \in (0,1]$ and $n \ge 1$.

The proof of the above hinges on the following version of integration by parts (as in fact does the basic convergence result in [19]).

Lemma 8. *Let $s_1, s_2, \dots, s_k, \dots$ be real numbers, and set $S_k = \sum_{\ell=1}^k s_\ell$, $S_0 = 0$. Let further $t_1, \dots, t_n$ be real numbers, $t_0 = t_{n+1} = 0$. Then, for every integer $m \ge 1$,*

$$\sum_{k=1}^n s_k t_k = \frac{1}{m} \sum_{k=1}^n [S_{k+m-1} - S_{k-1}] t_k + \sum_{k=0}^n \Bigg(\frac{1}{m} \sum_{\ell=k}^{k+m-1} [S_\ell - S_k] \Bigg) (t_{k+1} - t_k).$$

Proof. For any T_k, $k = 0, 1, \dots, n$, write

$$\begin{aligned} \sum_{k=1}^n s_k t_k &= \sum_{k=1}^n S_k (t_k - t_{k+1}) \\ &= \sum_{k=0}^n [T_k - S_k](t_{k+1} - t_k) - \sum_{k=0}^n T_k (t_{k+1} - t_k) \\ &= \sum_{k=0}^n [T_k - S_k](t_{k+1} - t_k) + \sum_{k=1}^n [T_k - T_{k-1}] t_k. \end{aligned}$$

Conclude by choosing $T_k = \frac{1}{m} \sum_{\ell=k}^{k+m-1} S_\ell$, $k = 0, 1, \dots, n$. □

Proof of Proposition 7. Applying Lemma 8 with $s_k = z_k$ and $t_k = v_k^2$ (bearing in mind that $v_0 = v_{n+1} = 0$, and we are free to set $s_k = 0$ for $k \ge n+1$) yields

$$\begin{aligned} \sum_{k=1}^n z_k v_k^2 &\le \frac{1}{m} \sum_{k=1}^n |S_{k+m-1} - S_{k-1}| v_k^2 + \sum_{k=0}^n \Big(\frac{1}{m} \sum_{\ell=k}^{k+m-1} |S_\ell - S_k| \Big) |v_{k+1}^2 - v_k^2| \\ &\le \frac{1}{m} \sum_{k=1}^n \Delta_m(k-1) v_k^2 + \sum_{k=0}^n \Delta_m(k) |v_{k+1} + v_k| |v_{k+1} - v_k| \end{aligned}$$

where

$$\Delta_m(k) = \max_{k+1 \le \ell \le k+m} |S_\ell - S_k|, \quad \text{for } k = 0, \dots, n. \tag{3.2}$$

Next, by the Cauchy-Schwarz inequality, for every $\lambda > 0$,

$$\frac{1}{\sqrt{\beta}} \sum_{k=1}^n z_k v_k^2 \le \frac{1}{m\sqrt{\beta}} \sum_{k=1}^n \Delta_m(k-1) v_k^2 + \lambda \sum_{k=0}^n (v_{k+1} - v_k)^2 + \frac{1}{4\lambda\beta} \sum_{k=0}^n \Delta_m(k)^2 (v_{k+1} + v_k)^2.$$

Choosing $\lambda = b\sqrt{n}$ we obtain

$$\sup_{||v||_2=1} H_b(z,v) \le \max_{1\le k\le n}\Big(\frac{1}{m\sqrt{\beta}}\,\Delta_m(k-1) + \frac{1}{2b\sqrt{n}\beta}\left[\Delta_m(k-1)^2 + \Delta_m(k)^2\right] - b\frac{k}{\sqrt{n}}\Big). \quad (3.3)$$

And since whenever $(j-1)m+1 \le k \le jm$, $1 \le j \le [n/m]+1$, it holds

$$\Delta_m(k) \vee \Delta_m(k-1) \le 2\Delta_{2m}\big((j-1)m\big),$$

we may recast (3.3) as in

$$\sup_{||v||_2=1} H_b(z,v)$$
$$\le \max_{1\le j\le [n/m]+1}\left(\frac{2}{m\sqrt{\beta}}\,\Delta_{2m}\big((j-1)m\big) + \frac{4}{b\sqrt{n}\beta}\,\Delta_{2m}\big((j-1)m\big)^2 - b\,\frac{(j-1)m+1}{\sqrt{n}}\right).$$

Continuing requires a tail bound on $\Delta_{2m}(J)$ for integer $J \ge 0$. By Doob's maximal inequality and our assumptions on z_k, for every $\lambda > 0$ and $t > 0$,

$$\mathbb{P}\Big(\max_{1\le \ell\le 2m} S_\ell \ge t\Big) \le e^{-\lambda t}\,\mathbb{E}\left[e^{\lambda S_{2m}}\right] \le e^{-\lambda t + 2cm\lambda^2}.$$

Optimizing in λ, and then applying the same reasoning to the sequence $-S_\ell$ produces

$$\mathbb{P}\Big(\max_{1\le \ell\le 2m} |S_\ell| \ge t\Big) \le 2\,e^{-t^2/8cm}.$$

Hence,

$$\mathbb{P}\Big(\Delta_{2m}(J) \ge t\Big) \le 2\,e^{-t^2/8cm}, \quad (3.4)$$

for all integers $m \ge 1$ and $J \ge 0$, and every $t > 0$.

From (3.4) it follows that

$$\mathbb{P}\Big(\max_{1\le j\le [n/m]+1}\Big(\frac{2}{m\sqrt{\beta}}\,\Delta_{2m}\big((j-1)m\big) - \frac{[b(j-1)m+1]}{2\sqrt{n}}\Big) \ge \frac{\varepsilon\sqrt{n}}{2}\Big) \quad (3.5)$$
$$\le \sum_{j=1}^{[n/m]+1}\mathbb{P}\Big(\frac{2}{m\sqrt{\beta}}\,\Delta_{2m}\big((j-1)m\big) \ge \frac{b[(j-1)m+1]}{2\sqrt{n}} + \frac{\varepsilon\sqrt{n}}{2}\Big)$$
$$\le 2\sum_{j=1}^{[n/m]+1}\exp\left(-\frac{\beta m}{128c}\Big[\frac{b[(j-1)m+1]}{\sqrt{n}} + \varepsilon\sqrt{n}\Big]^2\right),$$

and similarly

$$\mathbb{P}\Big(\max_{1\le j\le [n/m]+1}\Big(\frac{4}{b\sqrt{n}\beta}\,\Delta_{2m}\left((j-1)m\right)^2 - \frac{b[(j-1)m+1]}{2\sqrt{n}}\Big) \ge \frac{\varepsilon\sqrt{n}}{2}\Big) \quad (3.6)$$
$$\le 2\sum_{j=1}^{[n/m]+1}\exp\left(-\frac{\beta b\sqrt{n}}{64cm}\Big[\frac{b[(j-1)m+1]}{\sqrt{n}} + \varepsilon\sqrt{n}\Big]\right).$$

Combined, this reads

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_b(z,v) \geq \varepsilon\sqrt{n}\Big) \tag{3.7}$$
$$\leq \Big(\frac{2}{1-e^{-\beta\varepsilon b m^2/64c}}\Big) e^{-\beta m n \varepsilon^2/128c} + \Big(\frac{2}{1-e^{-\beta b^2/64c}}\Big) e^{-\beta b \varepsilon n/64cm},$$

which we have recorded in full for later use. In any case, the choice $m = [\varepsilon^{-1/2}]$ will now produce the claim. □

We may now dispense of the proof of Theorem 1 (Right-Tail). Before turning to the proof, we remark that if $\varepsilon > 1$, one may run through the above argument and simply choose $m = 1$ at the end to produce the classical form of the large deviation inequality (1.4) known previously for $\beta = 1, 2$.

We turn to the values $0 < \varepsilon \leq 1$. The form (2.4) is split into two pieces,

$$H_b(v) = H_{b/2}(v, g) + \tilde{H}_{b/2}(v, \chi),$$

Proposition 7 applying to each.

The first term on the right is precisely of the form (3.1) with each z_k an independent mean-zero Gaussian of variance 2, which obviously satisfies the tail assumption with $c = 1$. The second term, $\tilde{H}_{b/2}(v,\chi)$, is a bit different, having noise present through the quantity $\sum_{k=1}^{n-1}(\chi_{\beta(n-k)} - \mathbb{E}\chi_{\beta(n-k)})v_k v_{k+1}$. But carrying out the integration by parts on $t_k = v_k v_{k+1}$ (and $s_k = \chi_{\beta(n-k)} - \mathbb{E}\chi_{\beta(n-k)}$), will produce a bound identical to (3.4), with an additional factor of 2 before each appearance of Δ_{2m}. Thus, we will be finished granted the following bound.

Lemma 9. *For χ a χ random variable,*

$$\mathbb{E}[e^{\lambda\chi}] \leq e^{\lambda\mathbb{E}\chi + \lambda^2/2}, \ \textit{for all } \lambda \in \mathbb{R}. \tag{3.8}$$

Proof. When the parameter r is greater than one, this is a consequence of the Log-Sobolev estimate for general gamma random variables. The density function $f(x) = c_r x^{r-1} e^{-x^2/2}$ on $\mathbb{R}_+$ satisfies $(\log f(x))'' \leq -1$ (if $r \geq 1$) and so the standard convexity criterion (see *e.g.* [13]) applies to yield a Log-Sobolev inequality (with the same constant as in the Gaussian case). Then the well known Herbst argument gives the bound (3.8).

For $r < 1$ set $\phi(\lambda) = \mathbb{E}[e^{\lambda\chi}]$ and first consider $\lambda \geq 0$. Differentiating twice, then integrating by parts we have that

$$\phi''(\lambda) = \lambda\phi'(\lambda) + r\phi(\lambda),$$

subject to $\phi(0)=1$, $\phi'(0)=E\chi := \mathfrak{e}$, for short. Note of course that $\phi''(0)=r=E\chi^2$, and now integrating twice we also find that: with $\psi(\lambda)=e^{-\lambda^2/2}\phi(\lambda)$ and $\theta=\frac{r}{1+r}<1$,

$$\begin{aligned}\psi(\lambda) &= 1+\mathfrak{e}\lambda+\int_0^\lambda (r\lambda-(1+r)t)\psi(t)\,dt \\ &\le 1+\mathfrak{e}\lambda+r\theta\int_0^\lambda(\lambda-t)\psi(\theta t)\,dt.\end{aligned}$$

Next, by the inequality $\frac{r}{\sqrt{1+r}}\le\mathfrak{e}$ already used above (proof of Lemma 5′) we can continue the above as in $\psi(\lambda)\le 1+\mathfrak{e}\lambda+\mathfrak{e}^2\int_0^\lambda(\lambda-t)\psi(\theta t)dt$. Iterating, we get a next term which reads

$$\mathfrak{e}^2\int_0^t(t-s)(1+\mathfrak{e}\theta s)ds\le\mathfrak{e}^2\int_0^t(t-s)(1+\mathfrak{e}s)ds=\mathfrak{e}^2(t^2/2)+\mathfrak{e}^3(t^3/6),$$

and this easily propagates to complete the proof (which actually works for all r).

To prove (3.8) for $\lambda<0$ set $\phi(\lambda)=\mathbb{E}[e^{-\lambda\chi}]$ (viewing λ as nonnegative), and the basic differential equation becomes $\phi''(\lambda)=-\lambda\phi'(\lambda)+r\phi(\lambda)$. With $p(\lambda)=\phi'(\lambda)/\phi(\lambda)$ this transforms to $p'(\lambda)=-p^2(\lambda)-\lambda p(\lambda)+r$. Just using $p'(\lambda)\le-\lambda p(\lambda)+r$ we find that

$$p(\lambda)\le p(0)+re^{-\lambda^2/2}\int_0^\lambda e^{t^2/2}dt\le-\mathfrak{e}+r\lambda,\quad\text{or}\quad\phi(\lambda)\le e^{-\mathfrak{e}\lambda+r\lambda^2/2},$$

which is what we want (when of course $r\le1$). □

Remark. For Theorem 1′ simply examine (3.7) to note the new form of the constants, with b dependent on β for the second part of the statement.

Left-Tail. This demonstrates yet another advantage of the variational picture afforded by the tridiagonal models. Namely, the bound may be achieved by a suitable choice of test vector since

$$\mathbb{P}\Big(\sup_{||v||_2=1}H_a(v)\le-2C\sqrt{n}\varepsilon\Big)\le\mathbb{P}\Big(H_a(v)\le-2C\sqrt{n}\varepsilon\,||v||_2^2\Big)$$

for whatever $\{v_k\}_{k=1,\dots,n}$ on the right hand side. This same idea was used in the large deviation estimates for TW_β in [19]. (Here have thrown in the additional constant $2C$ for reasons that will be clear in a moment.) Simplifying, we write

$$\begin{aligned}&\mathbb{P}\Big(H_a(v)\le-2C\sqrt{n}\varepsilon\,||v||_2^2\Big) \\ &\le\ \mathbb{P}\Big(H_a(v,g)\le-C\sqrt{n}\varepsilon\,||v||_2^2\Big)+\mathbb{P}\Big(\chi(v)\le-C\sqrt{n}\varepsilon\,||v||_2^2\Big),\end{aligned}\tag{3.9}$$

where in $H_a(v,g)$ we borrow the notation of Proposition 7 and

$$\chi(v)=\frac{2}{\sqrt{\beta}}\sum_{k=1}^{n-1}\big(\chi_{\beta(n-k)}-\mathbb{E}\chi_{\beta(n-k)}\big)v_kv_{k+1}.$$

Focus on the first term on the right of (3.9), and note that

$$\begin{aligned}
&\mathbb{P}\Big(H_a(v,g) \le -C\sqrt{n}\varepsilon\, ||v||_2^2\Big) \qquad (3.10)\\
&= \mathbb{P}\bigg(\Big(\frac{2}{\beta}\sum_{k=1}^{n} v_k^4\Big)^{1/2}\mathfrak{g} \ge C\sqrt{n}\varepsilon\sum_{k=1}^{n} v_k^2 - a\sqrt{n}\sum_{k=0}^{n}(v_{k+1}-v_k)^2 - \frac{a}{\sqrt{n}}\sum_{k=1}^{n} k v_k^2\bigg)
\end{aligned}$$

with $\mathfrak{g}$ a single standard Gaussian. Our choice of v is motivated as follows. The event in question asks for a large eigenvalue (think of $\sqrt{n}\varepsilon$ as large for a moment) of an operator which mimics negative Laplacian plus potential. The easiest way to accomplish this would be for the potential to remain large on a relatively long interval, with a flat eigenvector taking advantage. We choose

$$v_k = \frac{k}{n\varepsilon} \wedge \Big(1 - \frac{k}{n\varepsilon}\Big) \text{ for } k \le n\varepsilon \text{ and zero otherwise,} \qquad (3.11)$$

for which

$$\sum_{k=1}^{n} v_k^2 \sim \sum_{k=1}^{n} v_k^4 \sim n\varepsilon, \quad \sum_{k=0}^{n}(v_{k+1}-v_k)^2 \sim \frac{1}{n\varepsilon}, \text{ and } \sum_{k=1}^{n} k v_k^2 \sim n^2\varepsilon^2. \qquad (3.12)$$

(Here $a \sim b$ indicates that the ratio a/b is bounded above and below by numerical constants.) Substitution into (3.10) produces, for choice of $C = C(a)$ large enough inside the probability on the left,

$$\mathbb{P}\Big(H_a(v,g) \le -C\sqrt{n}\varepsilon\, ||v||_2^2\Big) \le e^{-\beta n^2\varepsilon^3/C} \text{ for } n\varepsilon^{3/2} \ge 1.$$

The restriction of the range of ε stems from the gradient-squared term; it also ensures that $\varepsilon n \ge 1$ which is required for our test vector to be sensible in the first place.

Next, as a consequence of Proposition 9 (see (3.8)) we have the bound: for $c > 0$,

$$\mathbb{P}\Big(\chi(v) \le -c^2\, ||v||_2^2\Big) \le \exp\bigg(-\beta c\Big(\sum_{k=1}^{n} v_k^2\Big)^2 \Big/ 8\sum_{k=1}^{n-1} v_k^2 v_{k+1}^2\bigg). \qquad (3.13)$$

With $c = C\sqrt{n}\varepsilon$ and v as in (3.11), this may be further bounded by $e^{-\beta n^2\varepsilon^3/C}$. Here too we should assume that $n\varepsilon^{3/2} \ge 1$.

Introducing a multiplicative constant of the advertised form C^β extends the above bounds to the full range of ε in the most obvious way. Replacing ε with $\varepsilon/2C$ throughout completes the proof.

Remark. With the exception of the restriction on n, nothing changes for the related conclusion in Theorem $1'$.

3.2 Laguerre ensembles

Right-Tail. We wish to apply the same ideas from the Hermite case to the Laguerre form $L_b(v)$ (for small b). Recall:

$$\begin{aligned} L_b(v) \;=\; & \frac{1}{\sqrt{\beta}}\sum_{k=1}^{n} Z_k v_k^2 + \frac{1}{\sqrt{\beta}}\sum_{k=2}^{n} \tilde{Z}_k v_k^2 + \frac{2}{\sqrt{\beta}}\sum_{k=1}^{n-1} Y_k v_{k+1} v_k \\ & - b\sqrt{n}\sum_{k=0}^{n}(v_{k+1}-v_k)^2 - b\frac{1}{\sqrt{n}}\sum_{k=1}^{n} k v_k^2. \end{aligned} \tag{3.14}$$

Here, Z_k, $\widetilde{Z}_k$ and Y_k are as defined in (2.7), and the appropriate versions of the tail conditions for these variables (in order to apply Proposition 7) are contained in the next two lemmas.

Lemma 10. *For χ be a χ random variable of positive parameter,*

$$\mathbb{E}[e^{\lambda\chi^2}] \le e^{\mathbb{E}[\chi^2](\lambda+2\lambda^2)} \text{ for all real } \lambda < 1/4.$$

Proof. With $r = \mathbb{E}[\chi^2] > 0$ and $\lambda < \frac{1}{2}$,

$$\mathbb{E}[e^{\lambda\chi^2}] = \Big(\frac{1}{1-2\lambda}\Big)^{r/2}.$$

Now, since $x \ge -\frac{1}{2}$ implies $\log(1+x) \ge x - x^2$, for any $\lambda \le \frac{1}{4}$ the right hand side of the above is less $e^{r(\lambda+2\lambda^2)}$ as claimed. □

Lemma 11. *Let χ and $\widetilde{\chi}$ be independent χ random variables. Then, for every $\lambda \in \mathbb{R}$ such that $|\lambda| < 1$,*

$$\mathbb{E}\Big[e^{\lambda(\chi\widetilde{\chi}-\mathbb{E}[\chi\widetilde{\chi}])}\Big] \le \frac{1}{\sqrt{1-\lambda^2}}\exp\Big(\frac{\lambda^2}{2(1-\lambda^2)}\big[\mathbb{E}[\chi]^2 + \mathbb{E}[\widetilde{\chi}]^2 + 2\lambda\mathbb{E}[\chi]\mathbb{E}[\widetilde{\chi}]\big]\Big).$$

Proof. For $|\lambda| < 1$, using inequality (3.8) in the $\widetilde{\chi}$ variable,

$$\mathbb{E}[e^{\lambda\chi\widetilde{\chi}}] \le \mathbb{E}\big[e^{\lambda\mathbb{E}[\widetilde{\chi}]\chi + \lambda^2\chi^2/2}\big] = \int_{-\infty}^{\infty}\mathbb{E}\big[e^{\lambda[\mathbb{E}(\widetilde{\chi})+s]\chi}\big]d\gamma(s)$$

where γ is the standard normal distribution on $\mathbb{R}$. Now, for every s, with (3.8) in the χ variable,

$$\mathbb{E}\big[e^{\lambda(\mathbb{E}(\widetilde{\chi})+s)\chi}\big] \le e^{\lambda(\mathbb{E}[\widetilde{\chi}]+s)\mathbb{E}[\chi] + \lambda^2[\mathbb{E}[\widetilde{\chi}]+s]^2/2}.$$

The result follows by integration over s. □

What this means for the present application is that

$$\mathbb{E}[e^{\lambda Z_k}],\ \mathbb{E}[e^{\lambda\widetilde{Z}_k}] \le e^{2\lambda^2} \text{ for all real } \lambda \le \sqrt{\beta\kappa}/4, \tag{3.15}$$

and

$$\mathbb{E}[e^{\lambda Y_k}] \le 2e^{12\lambda^2} \text{ for all real } \lambda \text{ with } |\lambda| \le \sqrt{\beta\kappa}/2\sqrt{2}. \tag{3.16}$$

To proceed, we split $L_b(v)$ into three pieces now, isolating each of the noise components, and focus on the bound for $\sup_{||v||_2=1} L_{b/3}(v,Z)$ (the notation indicating (3.14) with only the Z noise term present). One must take some care when arriving at the analog of (3.4). In obtaining an inequality of the form $\mathbb{P}(\Delta_m(J,Z) > t) \le Ce^{-t^2/C}$ we must be able to apply (3.15) (and (3.16) when considering the Y noise term) with $\lambda = O(t/m)$. But, examining (3.5) and (3.6) shows we only need consider t's of order $\sqrt{\beta}(\sqrt{n}+\sqrt{\kappa}\varepsilon)m$. Thus we easily get by via

$$\begin{aligned}\mathbb{P}\Big(\sup_{||v||_2=1} L_b(v) \ge \sqrt{\kappa}\varepsilon\Big) &\le C\,\mathbb{P}\Big(\sup_{||v||_2=1} L_{b/3}(v,Z) \ge \sqrt{\kappa}\varepsilon/C\Big) \\ &\le \Big(\frac{C}{1-e^{-\beta\varepsilon\sqrt{\frac{\kappa}{n}}m^2/C}}\Big)e^{-\beta\kappa m\varepsilon^2/C} + \Big(\frac{C}{1-e^{-\beta/C}}\Big)e^{-\beta\sqrt{\kappa n}\varepsilon/Cm},\end{aligned} \tag{3.17}$$

with $C = C(b)$, compare (3.7). Setting m to be the nearest integer to $\frac{1}{\sqrt{\varepsilon}}\left(\frac{n}{\kappa}\right)^{1/4}$ puts both exponential factors on the same footing, namely on the order of $e^{-\beta\kappa^{3/4}n^{1/4}\varepsilon^{3/2}/C}$, and removes all ε, κ, and n dependence on the first prefactor. Certainly the best decay possible, but requires $\varepsilon \le \sqrt{n/\kappa}$. Otherwise, if $\varepsilon \ge \sqrt{n/\kappa}$, we simply choose $m=1$ in which case the second term of (3.17) is the larger and produces decay $e^{-\beta\sqrt{n\kappa}\varepsilon/C}$. Happily, both estimates agree at the common value $\varepsilon = \sqrt{n/\kappa}$.

Remark. That Lemma 11 holds for chi's of any parameter will allows an extension to $\beta < 1$ in the same spirit as the Hermite case, granted working out a Lemma 6′ standing in the same relationship as Lemma 5′ does to Lemma 5.

Left-Tail. It is enough to produce the bound for $\mathbb{P}(L_a(v,Z) \le -C\sqrt{\kappa}\varepsilon||v||_2^2)$ for large a, given $v \in \mathbb{R}^n$ and a $C = C(a)$ as in the Hermite case. Indeed, (3.15) and (3.16) show that $L_a(v,\widetilde{Z})$ and $L_a(v,Y)$ will follow suit.

We have the estimate

$$\begin{aligned}&\mathbb{P}\Big(L_a(v,Z) \le -C\sqrt{\kappa}\varepsilon||v||_2^2\Big) \\ &= \mathbb{P}\Big(\sum_{k=1}^n(-Z_k)v_k^2 \ge \sqrt{\beta}\,\Big[C\sqrt{\kappa}\varepsilon||v||_2^2 - a\sqrt{n}||\nabla v||_2^2 - (a/\sqrt{n})||\sqrt{k}v||_2^2\Big]\Big) \\ &\le \exp\Big(-\beta\,\frac{[C\sqrt{\kappa}\varepsilon||v||_2^2 - a\sqrt{n}||\nabla v||_2^2 - (a/\sqrt{n})||\sqrt{k}v||_2^2]^2}{8||v||_4^4}\Big).\end{aligned} \tag{3.18}$$

Here we have introduced the shorthand

$$||v||_4^4 = \sum_{k=1}^n v_k^4,\ ||\nabla v||_2^2 = \sum_{k=0}^n (v_{k+1}-v_k)^2,\ ||\sqrt{k}v||_2^2 = \sum_{k=1}^n kv_k^2, \tag{3.19}$$

and have also used the fact that (3.15) applies just as well to $-Z_k$. In fact, the sign precludes any concern over the required choice of λ. For the Y-noise term, care must be taken on this point, but one may check that all is fine given our selection of v below.

For the small deviation regime, we use a slight modification of the Hermite test vector (3.11), and set

$$v_k = \Big(\frac{\delta}{n\varepsilon}k\Big) \wedge \Big(1 - \frac{\delta}{n\varepsilon}k\Big) \text{ with } \delta = (n/\kappa)^{1/2}$$

for $k \le n\varepsilon/\delta$ and $v_k = 0$ otherwise. This requires $\varepsilon \le \delta = (n/\kappa)^{1/2}$ in order to be sensible, and produces the same appraisals for $||v||_2^2, ||v||_4^4, ||\nabla||_2^2$, and $||\sqrt{k}v||_2^2$ as in (3.12), with each appearance of ε replaced by ε/δ. Substitution into (3.18) yields

$$\mathbb{P}\Big(L_a(v, Z) \le -C\sqrt{\kappa}\varepsilon||v||_2^2\Big) \le \exp\Bigg(- (\beta/8)\kappa^{3/2}n^{1/2}\varepsilon^3\Big[C - O(1 \vee \frac{1}{\varepsilon^3\kappa^{3/2}n^{1/2}})\Big]^2\Bigg).$$

For $\varepsilon > \sqrt{n/\kappa}$, notice that the particularly simple choice of a constant v gives

$$\mathbb{P}\Big(L_a(1, Z) \le -C\sqrt{\kappa}\varepsilon||1||_2^2\Big) \le \exp\Big(-(\beta/8)\kappa n\varepsilon^2[C - (2a/n) - a]^2\Big).$$

Combined, these two bounds cover the claimed result, provided that $\kappa^{3/2}n^{1/2}\varepsilon^3$ is chosen larger than one in the former. Extending this to the full range of ε and all remaining considerations are the same as in the Hermite setting.

3.3 Variances

We provide details for $\lambda_{\max}(H_\beta)$, the Laguerre case is quite the same. (Neither is difficult.) Write

$$\mathrm{Var}\left[\lambda_{\max}(H_\beta)\right] \le n\int_0^\infty \mathbb{P}\left(|\lambda_{\max}(H_\beta) - 2\sqrt{n}| \ge \sqrt{n}\varepsilon\right) d\varepsilon^2,$$

and then split the integrand in two according whether $\lambda_{\max} \le 2\sqrt{n}$ or $\lambda_{\max} > 2\sqrt{n}$.

First note that our upper bound on the probability that $\lambda_{\max}(H_\beta) - 2\sqrt{n} \le -\sqrt{n}\varepsilon$ applies to any $\varepsilon = O(1)$. Further, from the tridiagonal model we see that $\lambda_{\max}$ stochastically dominates $(1/\sqrt{\beta})\max_{1\le k\le n} g_k$. Hence, for $\delta > 0$ we have the cheap estimate $\mathbb{P}(\lambda_{\max}(H_\beta) \le -\delta\sqrt{n}) \le e^{-\beta n^2\delta^2}$, and thus

$$\mathbb{P}\left(\lambda_{\max}(H_\beta) - 2\sqrt{n} \le -\sqrt{n}\varepsilon\right) \le C_\beta\, e^{-\beta n^2(\varepsilon^3\wedge\varepsilon^2)/C}$$

for all $\varepsilon > 0$. This easily produces

$$\Big(\int_0^2 + \int_2^\infty\Big)\,\mathbb{P}\left(\lambda_{\max}(H_\beta) - 2\sqrt{n} \le -\sqrt{n}\varepsilon\right) d\varepsilon^2 \le C_\beta\, n^{-4/3}.$$

For the other range, recall that we mentioned at the end of proof for the right-tail upper bound that the advertised estimate is easily extended to the large deviation regime (cf. (1.4)) to read

$$\mathbb{P}\left(\lambda_{\max}(H_\beta) - 2\sqrt{n} \geq \sqrt{n}\varepsilon\right) \leq C_\beta\, e^{-\beta n(\varepsilon^{3/2} \vee \varepsilon^2)/C}.$$

This results in

$$\left(\int_0^2 + \int_2^\infty\right) \mathbb{P}\left(\lambda_{\max}(H_\beta) - 2\sqrt{n} \geq \sqrt{n}\varepsilon\right) d\varepsilon^2 \leq C_\beta\, n^{-4/3}$$

and completes the proof.

4 (Hermite) Lower Bounds

Right-Tail. This follows from another appropriate choice of test vector v. To get started, write

$$\begin{aligned} \mathbb{P}\Big(\sup_{\|v\|_2=1} H(v) \geq \sqrt{n}\varepsilon\Big) &\geq \mathbb{P}\Big(H_a(v) \geq \sqrt{n}\varepsilon\,\|v\|_2^2\Big) \quad (4.1) \\ &\geq \mathbb{P}\Big(H_a(v,g) \geq 2\sqrt{n}\varepsilon\,\|v\|_2^2\Big)\,\mathbb{P}\Big(\chi(v) < \sqrt{n}\varepsilon\,\|v\|_2^2\Big). \end{aligned}$$

Here, as before, $\chi(v) = (2/\sqrt{\beta})\sum_{k=1}^{n-1}(\chi_{\beta(n-k)} - \mathbb{E}[\chi_{\beta(n-k)}])v_k v_{k+1}$.

Our choice of v is arrived at by examining the first factor above: with, as in the left-tail upper bound, a standard Gaussian $\mathfrak{g}$,

$$\begin{aligned} &\mathbb{P}\Big(H_a(v,g) \geq 2\varepsilon\sqrt{n}\,\|v\|_2^2\Big) \\ &= \mathbb{P}\bigg(\Big(\frac{2}{\beta}\sum_{k=1}^n v_k^2\Big)^{1/2}\mathfrak{g} \geq 2\sqrt{n}\varepsilon\sum_{k=1}^n v_k^2 + a\sqrt{n}\sum_{k=0}^n (v_{k+1}-v_k)^2 + \frac{a}{\sqrt{n}}\sum_{k=1}^n k v_k^2\bigg). \end{aligned}$$

Now the intuition is that the eigenvalue (of a discretized $-d^2/dx^2+$ potential) is being forced large positive, so the potential should localize with the eigenvector following suit.

Let then

$$v_k = \sqrt{\varepsilon}k \wedge \left(1 - \sqrt{\varepsilon}k\right) \quad \text{for } k \leq \varepsilon^{-1/2} \text{ and otherwise } 0,$$

where we will assume that $n \geq \varepsilon^{-3/2} \geq \varepsilon^{-1/2}$. With these choices we have

$$\|v\|_2^2 \sim \|v\|_4^4 \sim \frac{1}{\sqrt{\varepsilon}},\ \|\nabla v\|_2^2 \sim \sqrt{\varepsilon},\ \|\sqrt{k}v\|_2^2 \sim \frac{1}{\varepsilon},$$

(recall the notation from (3.19)) and thus the existence of a constant $C = C(a)$ so that

$$\mathbb{P}\Big(H_a(v,g) \geq 2\varepsilon\sqrt{n}\,\|v\|_2^2\Big) \geq \frac{1}{C}\,e^{-C\beta n\varepsilon^{3/2}}.$$

Similarly, returning to the second factor on the right hand side of (4.1) and invoking the estimate (3.13) we also have

$$\mathbb{P}\Big(\chi(v) \geq \sqrt{n}\varepsilon ||v||_2^2\Big) \leq e^{-\beta n \varepsilon^{3/2}/C}$$

for the same choice of v. And granted $n\varepsilon^{3/2} \geq 1$, it follows that $\mathbb{P}(\chi(v) < \sqrt{n}\varepsilon||v||_2^2) \geq 1 - e^{-1/C}$ throughout this regime. That is,

$$\mathbb{P}\Big(\sup_{||v||_2=1} H(v) \geq \sqrt{n}\varepsilon\Big) \geq \frac{1}{C}\, e^{-C\beta n\varepsilon^{3/2}} \text{ whenever } n\varepsilon^{3/2} \geq 1.$$

When $n\varepsilon^{3/2} \leq 1$, write

$$\mathbb{P}\Big(\sup_{||v||_2=1} H(v) \geq \sqrt{n}\varepsilon\Big) \geq \mathbb{P}\Big(\sup_{||v||_2=1} H(v) \geq \sqrt{n}\varepsilon_0\Big) \geq \frac{1}{Ce^{\beta C}} \geq \frac{1}{Ce^{\beta C}}\, e^{-C\beta n\varepsilon^{3/2}},$$

where $\varepsilon_0 = n^{-2/3} \leq 1$ to produce the advertised form of the bound for all n and ε.

Left-Tail. This relies heavily on the right-tail upper bound. The first step is to reduce to a Gaussian setting via independence: for whatever $b > 0$,

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_{2b}(v) \leq -\sqrt{n}\varepsilon\Big) \geq \mathbb{P}\Big(\sup_{||v||_2=1} H_b(v,g) \leq -2\sqrt{n}\varepsilon\Big)\,\mathbb{P}\Big(\sup_{||v||_2=1} \widetilde{H}_b(v,\chi) \leq \sqrt{n}\varepsilon\Big).$$

Here we also use the notation of the proof of Theorem 1 (right-tail), from which we know that

$$\mathbb{P}\Big(\sup_{||v||_2=1} \widetilde{H}_b(v,\chi) \geq \sqrt{n}\varepsilon\Big) \leq Ce^{-n\varepsilon^{3/2}/C}.$$

(As $\beta \geq 1$ we are simply dropping it from the exponent on the right at this stage.) Hence, if as we regularly have start with an assumption like $n\varepsilon^{3/2} \geq C^2 \geq 1$, it follows that

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_{2b}(v) \leq -\sqrt{n}\varepsilon\Big) \geq (1-e^{-1})\mathbb{P}\Big(\sup_{||v||_2=1} H_b(v,g) \leq -2\sqrt{n}\varepsilon\Big).$$

Turning to $H_b(v,g)$ we make yet another decomposition of the noise term. Let L be an integer ($1 \leq L \leq n$) to be specified. Set $S_L = \frac{1}{L}\sum_{k=1}^{L} g_k$, and

$$\eta_k = g_k - \frac{1}{L}S_L \text{ for } 1 \leq k \leq L,\ \eta_k = g_k \text{ for } L < k \leq n.$$

Note that the family $\{\eta_k\}_{k=1,\dots,n}$ is independent of S_L. If the procedure of Proposition 7 could be applied to $H_b(v,\eta)$, we would have an event of probability larger than $1-Ce^{-n\varepsilon^{3/2}/C}$ (again we simply drop the beta dependence at this intermediate stage) on which

$$\frac{1}{\sqrt{\beta}}\sum_{k=1}^{n}\eta_k v_k^2 - b\sqrt{n}\sum_{k=1}^{n}(v_{k+1}-v_k)^2 - \frac{b}{2\sqrt{n}}\sum_{k=1}^{n} k v_k^2 \leq \sqrt{n}\varepsilon\sum_{k=1}^{n} v_k^2. \tag{4.2}$$

Since we are still working under the condition $n\varepsilon^{3/2} \geq C^2$, this is to say that there is an event of probability a least $1 - 1/e$, depending only of the η_k's, and on which

$$H_b(v,g) \leq \frac{1}{\sqrt{\beta}} S_L \sum_{k=1}^{L} v_k^2 - \frac{b}{2\sqrt{n}} \sum_{k=1}^{n} k v_k^2 + \sqrt{n}\varepsilon \sum_{k=1}^{n} v_k^2,$$

for every $v \in \mathbb{R}^n$. If we now choose $L + 1 \geq 6n\varepsilon/b$, we have further

$$H_b(v,g) \leq \frac{1}{\sqrt{\beta}} S_L \sum_{k=1}^{L} v_k^2 + \sqrt{n}\,\varepsilon \sum_{k=1}^{L} v_k^2 - 2\sqrt{n}\,\varepsilon \sum_{k=L+1}^{n} v_k^2$$

on that same event. Note this choice requires $\varepsilon \leq b/6$; it is here that the range of valid epsilon gets cut down in our final statement. In any case, putting the last remarks together we have proved that

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_b(v,g) \leq -2\sqrt{n}\,\varepsilon\Big) \geq (1-e^{-1})\mathbb{P}\left(S_L \leq -3\sqrt{n\beta}\,\varepsilon\right)$$

and so also

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_{2b}(v) \leq -\sqrt{n}\,\varepsilon\Big) \geq \frac{1}{C}\, e^{-C\beta n^2 \varepsilon^3},$$

again under the constrains $n\varepsilon^{3/2} \geq C^2$ and $\varepsilon \leq b/6$. The last inequality follows as S_L is a mean-zero Gaussian with variance of order $(n\varepsilon)^{-1}$.

The range $n\varepsilon^{3/2} \leq C^2$ is handled as before,

$$\mathbb{P}\Big(\sup_{||v||_2=1} H_{2b}(v) \leq -\sqrt{n}\,\varepsilon\Big) \geq \mathbb{P}\Big(\sup_{||v||_2=1} H_{2b}(v) \leq -\sqrt{n}\,\varepsilon_0\Big) \geq \frac{1}{Ce^{\beta C^5}} \geq \frac{1}{Ce^{\beta C^5}}\, e^{-C\beta n^2\varepsilon^3},$$

where $\varepsilon_0 = (C^2/n)^{2/3}$. As ε_0 must lie under $b/6$, this last selection requires $n \geq (6/b)^{3/2} C^2$, but smaller values of n can now be covered by adjusting the constant.

It remains to go back and verify that $\mathbb{P}(\sup_{||v||_2=1} H_b(v,\eta) \geq \sqrt{n}\varepsilon) \leq Ce^{-n\varepsilon^{3/2}/C}$. The only reason that Proposition 7 cannot be followed verbatim is that the η_k's are not independent, the first L of them being tied together through S_L. We need the appropriate Gaussian tail inequality for the variables

$$\triangle_m(k,\eta) = \max_{k<\ell\leq k+m} \Big| \sum_{j=k}^{\ell} \eta_j \Big|,$$

and, comparing with (3.4), shows that an estimate of type $\mathbb{P}(\triangle_m(k,\eta) > t) \leq Ce^{-t^2/Cm}$ suffices. But

$$\sum_{j=k}^{\ell} \eta_j = \sum_{j=k}^{\ell} g_j + (\ell \wedge L - k \wedge L) S_L,$$

and so

$$\mathbb{P}\big(\triangle_m(k,\eta) > t\big) \leq \mathbb{P}\big(\triangle_m(k,g) > t/2\big) + \mathbb{P}(mS_L > t/2).$$

The first term we have already seen to be of the required order, and the second is less than $e^{-Lt^2/8m^2}$. Since we only apply this bound in the present setting when $L = Cn\varepsilon \geq C\varepsilon^{-1/2}$ and $m = [\varepsilon^{-1/2}]$ (the choice made in Proposition 7), we have that $\mathbb{P}(mS_L > t/2) \leq e^{-t^2/Cm}$, and the proof is complete.

5 Minimal Laguerre Eigenvalue

While not detailed there, the results of [19] will imply that

$$\frac{(\sqrt{\kappa n})^{1/3}}{(\sqrt{\kappa} - \sqrt{n})^{4/3}} \Big((\sqrt{\kappa} - \sqrt{n})^2 - \lambda_{\min}(L_\beta) \Big) \Rightarrow TW_\beta, \tag{5.1}$$

whenever $\kappa, n \to \infty, \kappa/n \to c > 1$. This appraisal was long understood for the minimal eigenvalue of L{O/U}E, and has recently been extended to non-Gaussian versions of those ensembles in [8]. The condition $\kappa/n \to c > 1$ keeps the limiting spectral density supported away from the origin, resulting in the same soft-edge behavior that one has for $\lambda_{\max}$. If instead $\kappa - n$ remains fixed in the limit, one has a different scaling and different limit law(s) for $\lambda_{\min}$, the so-called hard-edge distributions. Granted the existence of the "hard-to-soft transition" for all $\beta > 0$ (see [4] and [18]) it is believed that (5.1) holds as long as $\kappa - n \to \infty$, but (to the best of our knowledge) this has not been explicitly worked out in any setting.

We only consider the analogue of the right-tail upper bound for $\lambda_{\min}$ and have the following.

Theorem 12. *Let $\beta \geq 1$ and $\kappa \geq n + 1$. Then,*

$$\mathbb{P}\Big(\lambda_{\min}(L_\beta) \leq (\sqrt{\kappa} - \sqrt{n})^2 (1 - \varepsilon) \Big) \leq C e^{-\beta (\kappa n)^{1/4} (\sqrt{\kappa} - \sqrt{n}) \varepsilon^{3/2}/C}, \tag{5.2}$$

for a numerical constant C and all $0 < \varepsilon \leq \sqrt{\frac{n}{\kappa}} (\alpha^{14} \wedge \alpha^2 n^{-2/5})$ in which $\alpha = 1 - \sqrt{n/\kappa}$.

According to (5.1), the deviations are of the order of $(\sqrt{\kappa n})^{1/3} (\sqrt{\kappa} - \sqrt{n})^{2/3} \varepsilon$, which explains the exponent in (5.2). Our condition on ε is certainly not very satisfactory, although still sensible to the fluctuations in (5.1). One would hope for the range of ε to be understandable in terms of the soft/hard edge picture − what we have here arises from technicalities. On the other hand, if we place an additional, "soft-edge" type, restriction on κ and n, we obtain a more natural looking estimate.

Corollary 13. *Again take $\beta \geq 1$, but now assume that $\kappa > cn$ for $c > 1$. The right hand side of (5.2) may then be replaced by $Ce^{-\beta n \varepsilon^{3/2}/C}$ for a $C = C(c)$, with the resulting bound valid for all $0 < \varepsilon \leq 1$.*

The last statement should be compared with Corollary V.2.1(b) of [8], which applies to classes of non-Gaussian matrices.

As to the proof, we proceed in a by now familiar way. We first set

$$\sqrt{\kappa}L(v) = v^T\Big((\sqrt{\kappa}-\sqrt{n})^2 - L_\beta\Big)v.$$

Then, after a rescaling of ε, we will prove the equivalent

$$\mathbb{P}\Big(\sup_{||v||=1} L(v) \geq \alpha^{4/3}\sqrt{n}\varepsilon\Big) \leq Ce^{-\beta n\varepsilon^{3/2}/C} \text{ for } \varepsilon \leq \min(\alpha^{44/3}, \alpha^{8/3}n^{-2/5}).$$

Similar to the strategy employed above, a series of algebraic manipulations shows that we can work instead with the simplified quadratic form

$$\begin{aligned} L'(v) \;\; = \;\; & \frac{1}{\sqrt{\beta}}\sum_{k=1}^{n}(-Z_k)v_k^2 + \frac{1}{\sqrt{\beta}}\sum_{k=2}^{n}(-\widetilde{Z}_k)v_k^2 + \frac{2}{\sqrt{\beta}}\sum_{k=1}^{n-1}(-Y_k)v_kv_{k+1} \\ & -\sum_{k=1}^{n-1}\frac{1}{\beta\sqrt{\kappa}}\,\mathbb{E}[\chi_{\beta(\kappa-k+1)}\widetilde{\chi}_{\beta(n-k)}](v_{k+1}+v_k)^2 - \frac{\alpha^2}{\sqrt{n}}\sum_{k=1}^{n}kv_k^2. \end{aligned} \tag{5.3}$$

(The condition $\kappa \geq n+1$ in Theorem 12 is used in passing from L to L'.)

We remark that under the added condition $\kappa > cn$ for $c > 1$, α is bounded uniformly from below and $\frac{1}{\beta\sqrt{\kappa}}\,\mathbb{E}[\chi_{\beta(\kappa-k+1)}\widetilde{\chi}_{\beta(n-k)}]$ is bounded below by a constant multiple of $\sqrt{n-k}$. Hence, the deterministic part of L' is bounded above by a small negative multiple of $\sqrt{n}\sum_{k=1}^{n-1}(v_{k+1}+v_k)^2 + \frac{1}{\sqrt{n}}\sum_{k=1}^{n}kv_k^2$. The proof of Corollary 13 is then identical to that of the right-tail upper bound for $\lambda_{\max}(L_\beta)$.

Back to Theorem 12 and α's unbounded from below, we begin by rewriting the noise term in L' as $\frac{1}{\sqrt{\beta}}$ times

$$\begin{aligned} & \sum_{k=1}^{n}(-Z_k)v_k^2 + \sum_{k=2}^{n}(-\widetilde{Z}_k)v_k^2 + 2\sum_{k=1}^{n-1}(-Y_k)v_kv_{k+1} \\ & \quad = \; \sum_{k=1}^{n}(-U_k)v_k^2 + \sum_{k=2}^{n}(-\widetilde{Z}_k)(v_k^2 - v_{k-1}^2) + \sum_{k=1}^{n-1}(-Y_k)v_k(v_{k+1}+v_k), \end{aligned}$$

in which

$$U_k = \frac{1}{\sqrt{\beta\kappa}}\Big[\big(\chi_{\beta(\kappa-k+1)} - \widetilde{\chi}_{\beta(n-k)}\big)^2 - \mathbb{E}\big[\big(\chi_{\beta(\kappa-k+1)} - \widetilde{\chi}_{\beta(n-k)}\big)^2\big]\Big], \quad k = 1, \ldots, n,$$

(with the convention that $\widetilde{\chi}_0 = 0$). The idea is the following. For moderate k, $\mathrm{Var}(U_k) = O(\alpha^2)$, and thus it is this contribution to the noise which balances the drift term $\frac{\alpha^2}{\sqrt{n}}\sum_{k=1}^{n}kv_k^2$. Also, one may check that in the continuum limit the optimal v is such that $|v_k+v_{k+1}| = o(1)$, and so the $\widetilde{Z}$ and Y terms should "wash out".

We complete the argument in two steps. In step one, we simply drop the $\widetilde{Z}$ and Y terms and apply the method in Proposition 7 to the further simplified form

$$L(v,U)=\frac{1}{\sqrt{\beta}}\sum_{k=1}^{n}(-U_k)v_k^2-\sum_{k=1}^{n-1}\frac{\mathbb{E}[\chi_{\beta(\kappa-k+1)}\widetilde{\chi}_{\beta(n-k)}]}{\beta\sqrt{\kappa}}(v_{k+1}+v_k)^2-\frac{\alpha^2}{\sqrt{n}}\sum_{k=1}^{n}kv_k^2. \tag{5.4}$$

Even here we loose a fair bit in our estimates (resulting in non-optimal on ε) due to the variable coefficient in the energy term. Step two shows that, under yet additional restrictions on ε, the $\widetilde{Z}$ and Y noise terms may be absorbed into $L(v,U)$.

Step 1. We wish to prove $\mathbb{P}(\sup_{||v||=1}L(v,U)\geq\alpha^{4/3}\sqrt{n}\varepsilon)\leq Ce^{-\beta n\varepsilon^{3/2}/C}$ for some range of $\varepsilon>0$. (The optimal range being $0<\varepsilon\leq(\kappa/n)^{1/2}\alpha^{2/3}$.) A first ingredient is a tail bound on the U_k variables, for which we first bring in the following.

Lemma 14. *(Aida, Masuda, Shigekawa [1]) Given a measure η on the line which satisfies a logarithmic Sobolev inequality with constant $C>0$, there is the estimate*

$$\int e^{\lambda(F^2-\mathbb{E}[F^2])}d\eta\leq 2\,e^{8C\lambda^2\mathbb{E}[F]^2}\ \textit{whenever}\ |\lambda|\leq\frac{1}{16C},$$

for any 1-Lipschitz function F.

As a consequence, we have that:

Corollary 15. *Let χ and $\widetilde{\chi}$ be independent χ random variables (each of parameter larger than one) and set $U=(\chi-\widetilde{\chi})^2$ and $\sigma=\mathbb{E}[\chi-\widetilde{\chi}]$. There exists a numerical constant $C>0$ such that*

$$\mathbb{E}[e^{\lambda(U-\mathbb{E}U)}]\leq Ce^{C\sigma^2\lambda^2}$$

for all real $\lambda\in(-1/C,1/C)$.

Indeed, by the general theory (see Thm. 5.2 of [13] for example) the distribution of the pair $(\chi,\widetilde{\chi})$ on $\mathbb{R}_+\times\mathbb{R}_+$ satisfies a logarithmic Sobolev inequality. The lemma then applies with $F(x,y)=x-y$. In our setting, we record this bound as

$$\mathbb{E}[e^{\lambda U_k}]\leq Ce^{C\sigma_k^2\lambda^2}\text{ for }|\lambda|<\sqrt{\beta\kappa}/C\text{ and }\sigma_k^2=\mathbb{E}[U_k]^2.$$

We make no effort to extend matters to χ's of parameter less than one as would be needed to consider $\beta<1$.

Picking up the thread of Proposition 7, the variable coefficient in the energy term of $L(v,U)$ is dealt with by applying the Cauchy-Schwarz argument with $\lambda=\lambda_k$ defined by

$$\lambda_k=\frac{\mathbb{E}[\chi_{\beta(\kappa-k+1)}\widetilde{\chi}_{\beta(n-k)}]}{\beta\sqrt{\kappa}},\ k=1,\dots,n-1, \tag{5.5}$$

compare (3.3). Schematically, we are left to bound

$$\sum_{j=1}^{[n/m]} \mathbb{P}\Big(\frac{1}{m\sqrt{\beta}} \Delta_m(jm, U) \vee \frac{1}{\lambda_{jm}\beta} \Delta_m(jm, U)^2 \geq \frac{\alpha^2}{\sqrt{n}} jm + \varepsilon \alpha^{4/3} \sqrt{n} \Big) \tag{5.6}$$

for our choice of integer m. The $\Delta_m(\cdot, U)$ notation stands in analogy to that used in Section 3. Note we have taken the liberty to drop various constants and shifts of indices in the above display (which are irrelevant to the upshot).

Here the dependence of σ_k and λ_k on the relationship between n and κ comes into play. While at the top of the form everything works as anticipated, these quantities behave unfavorably for k near n. For this reason we deal with the sum (5.6) by dividing the range into $j \leq n/2m$ and $j > n/2m$ with the help of the appraisals:

$$\sigma_k^2 \leq \begin{cases} C\alpha^2, & 1 \leq k \leq n/2, \\ C, & n/2 < k \leq n. \end{cases} \qquad \lambda_k \geq \begin{cases} \sqrt{n}/C, & 1 \leq k \leq n/2, \\ \frac{\alpha}{C}\sqrt{n-k}, & n/2 < k < n. \end{cases} \tag{5.7}$$

Restricted to $j \leq n/2m$ (and hence substituting $\sigma_{jm}^2 = C\alpha^2$, $\lambda_{jm} = \sqrt{n}/C$), the sum (5.6) can be bounded by $Ce^{-\beta n \varepsilon^{3/2}/C}$ upon choosing $m = [\varepsilon^{-1/2}\alpha^{-2/3}]$. This holds for all values of ε so long as the choice of m is sensible, requiring that $\varepsilon \geq \alpha^{-4/3} n^{-2}$. But this is ensured if $\kappa \geq n+1$ and $\varepsilon^{3/2} n \geq 1$ (the former having been built into the hypotheses and the latter we may always assume).

On the range $j \geq n/2m$ the ε term on the right hand side within the probabilities is of no help, and we use, along with $\sigma_{jm}^2 \leq C$ and $\lambda_{jm} \leq \alpha\sqrt{n - jm}/C$, the crude estimates

$$\begin{aligned} \sum_{n/2m \leq j < n/m} \mathbb{P}\Big(\Delta_m(jm, U) \geq \sqrt{\beta}\alpha^2 m^2 \frac{j}{\sqrt{n}} \Big) &\leq C \sum_{j \geq n/2m} e^{-\beta m^3 \alpha^4 j^2 / Cn} \\ &\leq Cm^{-2}\alpha^{-4}\, e^{-\beta m \alpha^4 n / C}, \end{aligned}$$

and

$$\begin{aligned} \sum_{n/2m \leq j < n/m} \mathbb{P}\Big(\Delta_m(jm, U)^2 \geq \beta\alpha^2 m \lambda_{jm} \frac{j}{\sqrt{n}} \Big) &\leq C \sum_{1 \leq j \leq n/2m} e^{-\beta\alpha^3 (n/m)^{1/2} j^{1/2} / C} \\ &\leq C\big(1 + (m/n\alpha^6)\big)\, e^{-\beta\alpha^3 (n/m)^{1/2}/C}. \end{aligned}$$

The choice of $m = [\varepsilon^{-1/2}\alpha^{-2/3}]$ being fixed, we can bound each of the above by the desired $Ce^{-\beta n \varepsilon^{3/2}/C}$ only by restricting ε to be sufficiently small. The first estimate requires $\varepsilon \leq \alpha^{20/3}$, the second requires in addition that $\varepsilon \leq \alpha^{8/3} n^{-2/5}$ (and again uses $n\varepsilon^{3/2} \geq 1$).

In summary

$$\mathbb{P}\Big(\sup_{||v||=1} L(v, U) \geq \alpha^{4/3}\sqrt{n}\varepsilon \Big) \leq Ce^{-\beta n \varepsilon^{3/2}/C} \ \text{ if } \ 0 < \varepsilon \leq \min(\alpha^{20/3}, \alpha^{8/3} n^{-2/5}). \tag{5.8}$$

It is perhaps worth mentioning here that the bounds on λ_k and σ_k^2 for the range $k \ge n/2$ introduced in (5.7) may be improved slightly, though not apparently with great effect on the final result.

Step 2. To absorb the $\widetilde{Z}, Y$ noise terms, we show that $L'(v) \le \tilde{L}(v,U) + \mathcal{E}(\widetilde{Z}, Y, v)$ with a new form $\tilde{L}(v,U)$ comparable to $L(v,U)$, and an "error" term $\mathcal{E}$ for which we have $\mathbb{P}(\mathcal{E} \ge \alpha^{4/3}\sqrt{n}\varepsilon) \le Ce^{-\beta n \varepsilon^{3/2}/C}$, at least for some range of $\varepsilon > 0$. What follows could almost certainly be improved upon.

Define, for $k = 1, \dots, n-1$:

$$a_k = \frac{1}{4}\lambda_k \text{ for } k \le \alpha^4 n, \quad a_k = \frac{1}{16}\frac{\alpha^2 k}{\sqrt{n}} \text{ for } k \ge \alpha^4 n.$$

(Recall the definition of λ_k from (5.5).) Then, an application of the Cauchy-Schwarz inequality yields: for all v of length one,

$$\frac{1}{\sqrt{\beta}}\sum_{k=2}^{n}(-\widetilde{Z}_k)(v_k^2 - v_{k-1}^2) \le \frac{1}{4}\sum_{k=1}^{n-1}\lambda_k(v_{k+1}+v_k)^2 + \frac{\alpha^2}{4\sqrt{n}}\sum_{k=1}^{n} k v_k^2 + \max_{1\le k\le n-1}\frac{\widetilde{Z}_{k+1}^2}{\beta a_k}$$

A similar estimate applies to $\sum_{k=1}^{n-1} Y_k v_k (v_{k+1} - v_k)$. Accordingly,

$$L'(v) \le \tilde{L}(v,U) + \max_{1\le k\le n-1}\frac{\widetilde{Z}_{k+1}^2}{\beta a_k} + \max_{1\le k\le n-1}\frac{Y_k^2}{\beta a_k}$$

with

$$\tilde{L}(v,U) = \frac{1}{\sqrt{\beta}}\sum_{k=2}^{n}(-U_k)v_k^2 - \frac{1}{2}\sum_{k=1}^{n-1}\lambda_k(v_{k+1}+v_k)^2 - \frac{\alpha^2}{2\sqrt{n}}\sum_{k=1}^{n} k v_k^2.$$

Obviously, the arguments of step 1 apply to $\tilde{L}(v,U)$.

Finally, with W_k either $\widetilde{Z}_{k+1}$ or Y_k, Lemmas 10 and 11 imply that

$$\mathbb{P}\left(\max_{1\le k\le n-1}\frac{W_k^2}{\beta a_k} \ge \varepsilon\alpha^{4/3}\sqrt{n}\right) \le C\sum_{k=1}^{n} e^{-\beta\varepsilon\alpha^{4/3}a_k/C},$$

provided say $\varepsilon \le 1$. Since it may be assumed that $\alpha < 1/2$ (otherwise we are in the easy regime covered by Corollary 13), we have the bound $a_k = \lambda_k \le \sqrt{n}/C$ for $k \le \alpha^4 n \le n/2$ and so also

$$\sum_{k=1}^{n} e^{-\varepsilon\alpha^{4/3}\sqrt{n}a_k/C} \le \alpha^4 n\, e^{-\beta\varepsilon\alpha^{4/3}n/C} + \frac{C}{\varepsilon\alpha^{10/3}}\, e^{-\beta\varepsilon\alpha^{22/3}n/C},$$

by considering the sums over $k \le \alpha^4 n$ and $k > \alpha^4 n$ separately. If now $\varepsilon \le \alpha^{44/3}$ (still keeping in mind that $\varepsilon^{3/2}n \ge 1$), the right hand side is less than $Ce^{-\beta n\varepsilon^{3/2}/C}$. Adding this new constraint on ε to those stated in (5.8) completes the proof.

Acknowledgments The work of the first author was supported in part by the French ANR GRANDMA, that of the second author by NSF grant DMS-0645756. The second author also thanks the Institut de Mathématiques de Toulouse, during a visit to which much of the present work was completed, for its hospitality.

References

[1] Aida, S., Masuda, T., and Shigekawa, I. (1994) Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126**, no. 1, 83-101.

[2] Baik, J., Deift, P., Johansson, K. (1999) On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12**, no. 4, 1119–1178.

[3] Baik, J., Deift, P., McLaughlin, K., Miller, P. and Zhou, Z. (2001) Optimal tail estimates for directed last passage site percolation with geometric random variables. *Adv. Theor. Math. Phys.* **5**, 1207–1250.

[4] Borodin, A., Forrester, P. (2003) Increasing subsequences and the hard-to-soft transition in matrix ensembles. *J. Phys. A: Math and Gen.* **36**, no. 12, 2963-2982.

[5] Dumitriu, I. and Edelman, A. (2002) Matrix models for beta ensembles. *J. Math. Phys.* **43**, no. 11, 5830-5847.

[6] Edelman, A., Sutton, B. (2007) From random matrices to stochastic operators. *J. Stat. Phys.* **127**, no. 6, 1121-1165.

[7] El Karoui, N. (2003) On the largest eigenvalue of Wishart matrices with identity covariance when n, p, and $p/n \to \infty$. *To appear, Bernoulli.*

[8] Feldheim, O., Sodin, S. (2009) A universality result for the smallest eigenvalues of certain sample covariance matrices. *To appear, Geom. Funct. Anal.*

[9] Forrester, P. Log-gases and Random Matrices. *Textbook in press,* 2009.

[10] Forrester, P. and Rains, E. Interrelations between orthogonal, unitary and symplectic matrix ensembles (2001). In *Random Matrix Models and their Applications,* voume 40 of *Math. Sci. Inst. Res. Publ.*, 171–207. Cambridge University Press.

[11] Johansson, K. (2000) Shape fluctuations and random matrices. *Comm. Math. Phys.* **209**, no. 2, 437-476.

[12] JOHNSTONE, I. M. (2001) On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29**, no. 2, 295-327.

[13] LEDOUX, M. The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs **89**. Amer. Math. Soc. 2001.

[14] LEDOUX, M. (2004) Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials: The continuous case. *Elect. Journal Probab.* **9**, 177-208.

[15] LEDOUX, M. (2007) Deviation inequalities on largest eigenvalues. Lecture Notes in Math. **1910**, 167-219. Springer.

[16] LEDOUX, M. (2009) A recursion formula for the moments of the Gaussian Orthogonal Ensemble. *Annales Inst. H. Poincaré* **45**, 754-769.

[17] QI, F. (2010) Bounds for the Ratio of Two Gamma Functions *Journal of Inequalities Appl.* **2010**, Article ID 493058.

[18] RAMÍREZ, J., RIDER, B. (2009) Diffusion at the random matrix hard edge. *Comm. Math. Phys.* **288**, no. 3, 887-906.

[19] RAMÍREZ, J., RIDER, B., AND VIRÁG, B. (2007) Beta ensembles, stochastic Airy spectrum and a diffusion. *Preprint, arXiv:math.PR/0607331.*

[20] SILVERSTEIN, J. (1985) The smallest eigenvalue of a large dimensional Wishart matrix. *Ann Probab.* **13**, no. 4, 1364-1368.

[21] SOSHNIKOV, A. (1999) Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207**, 697-733.

[22] TAO, T., AND VU, V. (2009) Random matrices: Universality of local eigenvalue statistics up to the edge. *Preprint, arXiv:0908.1982.*

[23] TRACY, C., AND WIDOM, H. (1994) Level spacing distributions and the Airy kernel. *Comm. Math. Phys.* **159** no. 1, 151-174.

[24] TRACY, C., AND WIDOM, H. (1996) On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177** no. 3, 727-754.

[25] TRACY, C., AND WIDOM, H. (2009) Asymptotics in ASEP with Step Initial Condition. *Comm. Math. Phys.* **290** no. 1, 129-154.

[26] Trotter, H. F. (1984) Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegő. *Adv. in Math.* **54** no. 1, 67-82

Michel Ledoux
Institut de Mathématiques de Toulouse,
Université de Toulouse, F-31062 Toulouse, France.
ledoux@math.univ-toulouse.fr

Brian Rider
Department of Mathematics,
University of Colorado at Boulder, Boulder, CO 80309.
brian.rider@colorado.edu