

# Accuracy of the Tracy-Widom limit for the largest eigenvalue in white Wishart matrices

Zongming Ma  
*Stanford University*

October 8, 2008

## Abstract

Let  $A$  be a  $p$ -variate real Wishart matrix on  $n$  degrees of freedom with identity covariance. The distribution of the largest eigenvalue in  $A$  has important applications in multivariate statistics. Consider the asymptotics when  $p$  grows in proportion to  $n$ , it is known from Johnstone (2001) that after centering and scaling, these distributions approach the orthogonal Tracy-Widom law for real-valued data, which can be numerically evaluated and tabulated in software.

Under the same assumption, we show that more carefully chosen centering and scaling constants improve the accuracy of the distributional approximation by the Tracy-Widom limit to second order:  $O((n \wedge p)^{-2/3})$ . Together with the numerical simulation, it implies that the Tracy-Widom law is an attractive approximation to the distributions of these largest eigenvalues, which is important for using the asymptotic result in practice. We also provide a parallel accuracy result for the smallest eigenvalue of  $A$  when  $n > p$ .

**Key Words and Phrases.** Eigenvalues of random matrices, Laguerre orthogonal ensemble, Laguerre polynomial, Liouville-Green method, principal component analysis, rate of convergence, Tracy-Widom distribution, Wishart distribution.

## 1 Introduction

The central object of multivariate statistical analysis is an  $n \times p$  data matrix  $X$ , where each of the  $n$  rows corresponds to an observation of a random vector in a  $p$ -dimensional space. If we assume that the row vectors are i.i.d. samples from a multivariate Gaussian distribution  $N_p(\mu, \Sigma)$ , much of the classical theory in multivariate statistical analysis is reduced to study of the eigen-decomposition of a random matrix following a Wishart distribution. Typical examples include but are not limited to principal component analysis (PCA), factor analysis and multidimensional scaling (MDS). The fundamental setting is the determinantal equation

$$\det(A - \lambda I) = 0,$$

where  $A$  follows a central Wishart distribution with covariance matrix  $\Sigma$ .

In this setting, a common null hypothesis is  $H_0 : \Sigma = I$ . For instance, in PCA, this is the hypothesis of isotropic variation over all the principal components; see, for example, Mardia et al. (1979, Section 8.4.3). If  $H_0$  is true, we say that we are in the null case and call  $A$  a (real) white Wishart matrix. For testing this particular hypothesis, as for many others in multivariate statistics, there are two different systematic strategies: one is the likelihood ratio test (LRT),

which uses all the eigenvalues of  $A$ ; the other is the union intersection test (UIT) initiated by Roy (1953), which utilizes only the largest (or smallest) eigenvalue of  $A$  for the current problem.

An inconvenience of using UIT is that the exact evaluation of the marginal distribution of the extreme sample eigenvalues is not simply tractable, even in the null case considered here. Interested readers are referred to Muirhead (1982, Section 9.7) for the expressions of the marginal distributions in terms of hypergeometric function of matrix argument; see, in particular, Corollary 9.7.2 and 9.7.4 there. We remark that recent work of Koev and Edelman (2006) has developed efficient evaluations of hypergeometric functions of matrix argument and made the computation of the exact marginal distributions possible when both  $n$  and  $p$  are small.

An alternative approach is to approximate these exact finite sample distributions of the extreme eigenvalues by some other well-understood asymptotic distribution. This kind of approximation is ubiquitous in statistics: the normal approximation to the distribution of the Wald and score statistics, the Chi-square approximation to the Pearson statistic in fitting contingency tables, etc. For the problem studied here, Anderson (2003, Chapter 13) provides a complete summary of the established results in the conventional regime of asymptotics:

$$p \text{ is fixed and } n \rightarrow \infty.$$

However, many modern data (microarray data, stock prices, weather forecasting, etc.) we are now dealing with typically have the number of features  $p$  very large while the number of observations  $n$  much smaller than or just comparable to  $p$ . For these situations, the classical asymptotics is no longer always appropriate and new asymptotic results that could handle this type of data are desirable.

An advance in this direction was made in Johnstone (2001), where the asymptotic regime was switched to

$$p \rightarrow \infty, n = n(p) \rightarrow \infty \text{ and } n/p \rightarrow \gamma \in (0, \infty). \quad (1)$$

To state his result, let  $X$  be an  $n \times p$  data matrix with the  $n$  rows i.i.d. following  $N_p(0, I)$ . The  $p \times p$  matrix  $A = X'X$  has a standard Wishart distribution:  $A \sim W_p(I, n)$ . We denote the ordered eigenvalues of  $A$  by  $\lambda_1 \geq \dots \geq \lambda_p$ . Borrowing tools from the field of Random Matrix Theory (RMT), especially those established in Tracy and Widom (1994, 1996, 1998), Johnstone showed that if we define centering and scaling constants as

$$\mu_p = (\sqrt{n-1} + \sqrt{p})^2, \quad \sigma_p = (\sqrt{n-1} + \sqrt{p}) \left( \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{p}} \right)^{1/3}, \quad (2)$$

then under condition (1),

$$\frac{\lambda_1 - \mu_p}{\sigma_p} \xrightarrow{\mathcal{D}} W_1 \sim F_1, \quad (3)$$

where  $F_1$  is the orthogonal Tracy-Widom law, which was originally found by Tracy and Widom (1996) as the limiting law of the largest eigenvalue of a  $p \times p$  real Gaussian symmetric matrix. We remark that, prior to Johnstone (2001), as a byproduct of his analysis on random growth model, Johansson (2000) established the scaling limit for the largest eigenvalue in complex white Wishart matrix, which turns out to be the unitary Tracy-Widom law  $F_2$ . We'd also like to mention that for the weak limit (3) to hold, El Karoui (2006a) extended the asymptotic regime (1) to include the cases where  $n/p \rightarrow 0$  or  $\infty$ .

This type of asymptotic result, albeit emerging only recently in the statistics literature, has already found its relevance to applications with modern data. For instance, based on the weak limit (3), Patterson et al. (2006) developed a formal test for the presence of population heterogeneity

in a biallelic dataset and suggested a systematic way for assigning statistical significance to successive eigenvectors, which in turn has been used to correct population stratification (Price et al., 2006) and to perform genetic matching (Luca et al., 2008) in genome-wide association studies.

From a statistical point of view, to inform the use of any asymptotic result in practice, we need to have an understanding of the accuracy of the approximation to finite distributions by the limit, which usually appears in the form of a rate of convergence result. In the complex domain, El Karoui (2006b) established such a result for Johansson’s theorem with carefully chosen centering and scaling constants. With his choice, the error term in the Tracy-Widom approximation could be controlled at the order  $O((n \wedge p)^{-2/3})$ , as opposed to  $O((n \wedge p)^{-1/3})$  by using the original centering and scaling constants in Johansson (2000). For an up-to-date survey of higher order accuracy results of this fashion, we refer to Johnstone (2006, Section 3).

In statistics, we are typically more interested in real-valued data. However, for technical reasons, results for complex-valued data are usually easier to derive under the asymptotic regime (1) than in the real case. Recently, in analyzing the parallel problem for the greatest root statistics for pairs of Wishart matrices (Mardia et al., 1979, Definition 3.7.2), Johnstone (2007) figured out a way to connect the central object of study in the real case to that in the complex case. To be more specific, in both real and complex cases, the problem reduces to the study of operator convergence in appropriate metrics by using standard techniques from Random Matrix Theory. The key observation there is that the crucial element of the operator kernel in the real case could be represented in closed form as a rank one perturbation of the complex kernel; see Johnstone (2007, Eq.(50)), which is a consequence of Adler et al. (2000, Proposition 4.2).

Inspired by Johnstone (2007), we investigate in this paper the rate of convergence for the distributions of properly rescaled largest eigenvalues in real white Wishart matrices to the orthogonal Tracy-Widom law. We remark that, instead of using Adler et al. (2000, Proposition 4.2), the central formula (15) for the “complex to real” connection in our paper is derived from a slightly earlier result given in Widom (1999, Section 4) which is specific to white Wishart matrices. This new approach not only helps to avoid introducing a further nonlinear transformation after rescaling the largest eigenvalues as in Johnstone (2007) but also enables us to make direct use of the analysis done in El Karoui (2006b) for complex white Wishart matrices.

*Statement of the theoretical result.* It was suggested in Johnstone (2006) that if we modify the centering and scaling constants from (2) to

$$\begin{aligned}\tilde{\mu}_{np} &= \left( \sqrt{n - \frac{1}{2}} + \sqrt{p - \frac{1}{2}} \right)^2, \\ \tilde{\sigma}_{np} &= \left( \sqrt{n - \frac{1}{2}} + \sqrt{p - \frac{1}{2}} \right) \left( \frac{1}{\sqrt{n - \frac{1}{2}}} + \frac{1}{\sqrt{p - \frac{1}{2}}} \right)^{1/3},\end{aligned}\tag{4}$$

we might obtain second order accuracy in the Tracy-Widom approximation.

Indeed, the main theoretical result of the paper can be formulated as the following theorem, which establishes the above conjecture.

**Theorem 1.** *Let  $A \sim W_p(I, n)$  and  $\lambda_1$  be its largest eigenvalue. Define centering and scaling constants  $(\tilde{\mu}_{np}, \tilde{\sigma}_{np})$  as in (4), then under condition (1), there exists a continuous and nonincreasing function  $C(\cdot)$ , such that for all real  $s_0$ , there exists an integer  $N_0(s_0, \gamma)$  for which we have that for any  $s \geq s_0$  and  $n \wedge p \geq N_0(s_0, \gamma)$ ,*

$$|P\{\lambda_1 \leq \tilde{\mu}_{np} + \tilde{\sigma}_{np}s\} - F_1(s)| \leq C(s_0)(n \wedge p)^{-2/3} \exp(-s/2).$$

The theorem provides theoretical support for using the Tracy-Widom law  $F_1$  as approximate largest eigenvalue distribution in the null case. In addition, the numerical investigation pursued in Section 2.1 shows that the approximation yields reasonable accuracy even when  $n$  and  $p$  are as small as 2. Therefore, both theoretical and numerical results provide us with the confidence in using the Tracy-Widom approximation for nearly all finite  $n \times p$  distributions, at least under the Wishart assumption.

*Remark 1.* In fact, Theorem 1 will be proved only when  $p$  is even and  $n \neq p$  since our method relies on a determinant formula of de Bruijn (1955) which was only established for  $p$  even and the Laguerre polynomials which are essential for building the convergence rate are not well-defined when  $n = p$ . It would be of interest to have some theoretical support for the  $p$  odd and the square cases. However, numerical experiments suggest that the Tracy-Widom approximation works just as well for  $p$  odd as for  $p$  even and for  $n = p$  as for  $n \neq p$ .

*Organization of the paper.* In Section 2, we first investigate the numerical quality of the Tracy-Widom approximation for finite  $n \times p$  distributions, then review some important statistical settings to which our result is relevant and finally discuss several interesting issues involved in this study, including a parallel result for the smallest eigenvalue. The rest of the paper is dedicated to the proof of Theorem 1, mainly with tools from Random Matrix Theory. In Section 3, we start with the formulation of Theorem 1 in RMT terminology. After that, we derive the central formula (15) in this paper and reduce our problem to the study of operator convergence in some appropriate metric. We sketch our proof of the main result in Section 4 by assembling operator theoretic tools and asymptotic bounds on transformed Laguerre polynomials. Finally, Section 5 gives details of the Laguerre asymptotics required in the proof. Appendix A collects various necessary technical details not spelled out fully in the main text. Appendix B discusses the issues mentioned in Section 2.3 in a more concrete manner.

## 2 Statistical Implications and Discussion

### 2.1 Quality of the approximation

An important motivation for the current study is to promote practical use of the Tracy-Widom approximation. For example, one could tabulate the  $F_1$  table and use it to compute  $p$ -values. With such motivation, we investigate the quality of the approximation with numerical experiments.

**Distributional approximation.** First of all, we study the numerical accuracy of the approximation using our centering and scaling constants (4) and compare it with that of the original proposal (2) in Johnstone (2001), with results summarized in Table 1. We first look at the square cases with  $n = p = 2, 5, 20$  and 100 and then the cases with the same  $p$ 's but with the ratio  $n/p$  fixed at 4 : 1, and finally the cases where  $p = 5$  and 10 with  $n/p$  raised to as high as 100 : 1 and 1000 : 1, which, in some sense, fall into the situation  $n/p \rightarrow \infty$  as discussed in El Karoui (2006a). Finally, in all these cases, we use  $R = 40,000$  replications.

In terms of accuracy, from the last three columns of Table 1, the approximation seems reasonable at conventional significance levels of 10%, 5% and 1% (corresponding to right-hand tails of the distributions) even when  $p$  is as small as 2 or 5, and keeps improving as  $p$  grows large, regardless of the  $n/p$  ratio. When  $p$  is large, for instance, in the  $100 \times 100$  and  $400 \times 100$  cases, the Tracy-Widom law yields reasonable approximation over the whole range of interest and matches the finite distributions almost exactly on the right-hand tail.

Percentiles	-3.8954	-3.1804	-2.7824	-1.9104	-1.2686	-0.5923	0.4501	0.9793	2.0234
TW	.01	.05	.10	.30	.50	.70	<b>.90</b>	<b>.95</b>	<b>.99</b>
2 × 2	.000 (.000)	.000 (.000)	.000 (.000)	.034 (.015)	.379 (.345)	.690 (.669)	<b>.908</b> (.902)	<b>.953</b> (.950)	<b>.988</b> (.987)
5 × 5	.000 (.000)	.002 (.002)	.021 (.020)	.218 (.213)	.465 (.460)	.702 (.698)	<b>.908</b> (.907)	<b>.954</b> (.953)	<b>.989</b> (.989)
20 × 20	.003 (.003)	.029 (.029)	.071 (.071)	.275 (.274)	.490 (.489)	.700 (.699)	<b>.902</b> (.901)	<b>.952</b> (.952)	<b>.990</b> (.990)
100 × 100	.008 (.008)	.044 (.043)	.091 (.091)	.291 (.291)	.495 (.495)	.699 (.699)	<b>.901</b> (.901)	<b>.951</b> (.951)	<b>.990</b> (.990)
8 × 2	.000 (.000)	.000 (.004)	.013 (.031)	.200 (.274)	.458 (.534)	.704 (.755)	<b>.913</b> (.931)	<b>.956</b> (.966)	<b>.990</b> (.992)
20 × 5	.001 (.002)	.018 (.028)	.057 (.077)	.262 (.305)	.486 (.533)	.703 (.739)	<b>.905</b> (.919)	<b>.952</b> (.960)	<b>.990</b> (.992)
80 × 20	.005 (.006)	.035 (.043)	.082 (.096)	.287 (.312)	.497 (.524)	.700 (.723)	<b>.902</b> (.911)	<b>.951</b> (.956)	<b>.990</b> (.992)
400 × 100	.009 (.010)	.047 (.052)	.095 (.103)	.298 (.312)	.499 (.514)	.700 (.712)	<b>.899</b> (.905)	<b>.949</b> (.952)	<b>.989</b> (.990)
500 × 5	.010 (.022)	.050 (.084)	.100 (.154)	.303 (.387)	.502 (.589)	.705 (.770)	<b>.905</b> (.932)	<b>.953</b> (.968)	<b>.992</b> (.995)
5000 × 5	.012 (.027)	.056 (.098)	.108 (.169)	.311 (.406)	.511 (.606)	.711 (.783)	<b>.910</b> (.938)	<b>.957</b> (.971)	<b>.993</b> (.995)
1000 × 10	.010 (.017)	.049 (.073)	.099 (.136)	.296 (.363)	.500 (.567)	.701 (.754)	<b>.902</b> (.925)	<b>.952</b> (.964)	<b>.991</b> (.993)
10000 × 10	.012 (.022)	.054 (.084)	.104 (.148)	.306 (.381)	.506 (.579)	.707 (.764)	<b>.903</b> (.927)	<b>.950</b> (.964)	<b>.991</b> (.994)
2 × SE	.001	.002	.003	.005	.005	.005	.003	.002	.001

Table 1: Simulations for finite  $n \times p$  vs. Tracy-Widom approximation: accuracy comparison of the new centering and scaling constants (4) with that in Johnstone (2001). For each combination of  $n$  and  $p$ , we show in the first line the estimated cumulative probabilities for  $\lambda_1$ , rescaled using (4); and in the second line with parentheses, rescaled using Johnstone (2001, Eq.(1.3) and (1.4)), both computed from  $R = 40,000$  repeated draws using the Beta-ensemble sampling technique proposed by Dumitriu and Edelman (2002). The conventional significance levels are highlighted in bold font and the last line gives approximate standard errors based on binomial sampling. The orthogonal Tracy-Widom distribution  $F_1$  was computed using the method proposed in Edelman and Persson (2002) with percentiles obtained by inverse interpolation.

In terms of the comparison with the original centering and scaling constants, we could see from the first block of Table 1 that in the square cases, neither method seems superior to the other. However, when the ratio  $n/p$  is changed to 4 : 1 or larger (see the second and the third blocks of Table 1), the improvement by using the new constants is substantial. The new constants not only provide better absolute accuracy in most of the cases, but also seem to result in a faster convergence to the limiting distribution  $F_1$ .

Last but not least, the good performance on the right tail and the faster convergence by using the new constants, as reflected in Table 1, support our theoretical bound in Theorem 1.

**Approximate percentiles.** Except for computing  $p$ -values,  $F_1$  could also be used to compute approximate percentiles of finite  $n \times p$  distributions. To measure the accuracy of this approxima-

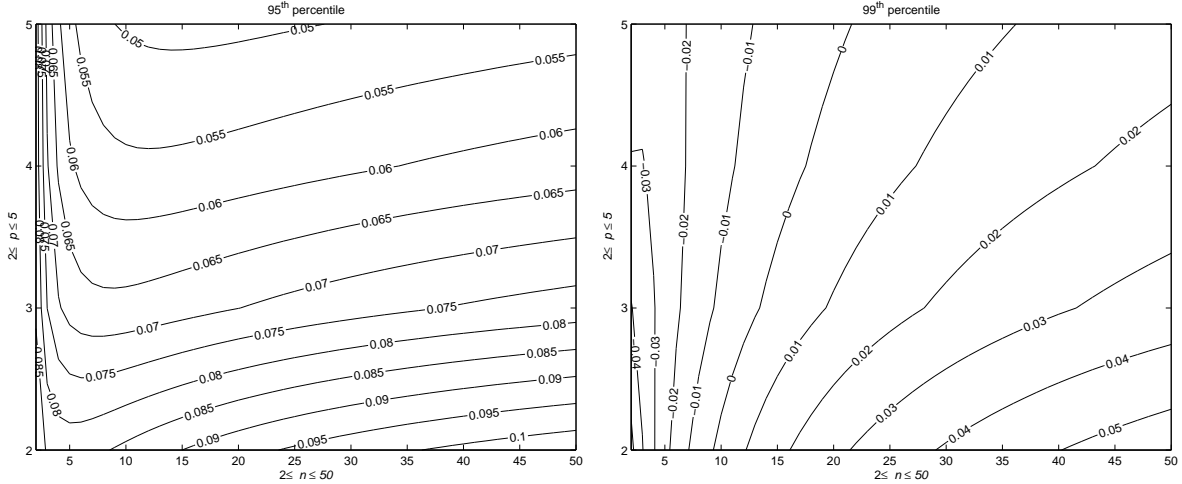


Figure 1: Plots of the relative error  $r_\alpha = \theta_\alpha^{TW}/\theta_\alpha - 1$  for the approximate percentiles computed from  $F_1$ : (a) 95-th percentile; (b) 99-th percentile. The exact finite  $n \times p$  largest eigenvalue distributions are computed using Plamen Koev's implementation in MATLAB of the recursive method proposed in Koev and Edelman (2006) and the orthogonal Tracy-Widom law  $F_1$  is computed using the method proposed in Edelman and Persson (2002). The percentiles are always obtained from inverse interpolation.

tion, we consider the relative error  $r_\alpha = \theta_\alpha^{TW}/\theta_\alpha - 1$ , where  $\theta_\alpha$  is the exact  $100\alpha$ -th percentile of the largest eigenvalue in finite  $n \times p$  model and  $\theta_\alpha^{TW}$  is its counterpart obtained from the Tracy-Widom law.

In Figure 1, we plot the relative error  $r_\alpha$  for  $\alpha = 0.95$  and  $0.99$ , with  $p$  ranging from 2 to 5 and  $n$  from 2 to 50. Although the minimum of  $n$  and  $p$  is no larger than 5, the numerical accuracy is reasonably satisfactory. For the 95-th percentile, the relative error ranges from 5% to 10% for most of the cases and slightly exceeds 10% only for the cases where  $p = 2$  and the  $n/p$  ratios are high. The approximation to the 99-th percentile is even better, with the absolute relative error  $|r_{.99}| \leq 5\%$  for most of the cases. Due to the computational limitation (Koev and Edelman, 2006), we could not compute the exact percentiles when  $n$  and  $p$  are large. However, we expect the approximate percentiles to become more accurate as the consequence of better distributional approximation.

## 2.2 Related statistical settings

In this part, we review several common settings in multivariate statistics to which our result is relevant.

**Principal component analysis.** Suppose that  $X = [x_1, \dots, x_n]'$  is a Gaussian data matrix. Write the sample covariance matrix  $S = n^{-1}X'HX$ , where  $H = I - n^{-1}\mathbf{1}\mathbf{1}'$  is the centering matrix, principal component analysis looks for a sequence of standardized vectors  $a_1, \dots, a_p$  in  $\mathbb{R}^p$ , such that for  $i = 1, \dots, p$ , where  $a_i$  successively solves the following optimization problem:

$$\max\{a'Sa : a'a_j = 0, j \leq i\},$$

where  $a_0$  can be taken as the zero vector. The successive sample principal component eigenvalues  $\ell_1 \geq \dots \geq \ell_p$  then satisfy  $\ell_i = a_i'Sa_i$ . From a different perspective, these  $\ell_i$ 's may also be found

as the roots of the determinantal equation

$$\det(S - \lambda I) = 0.$$

One basic question in the application of PCA is testing the hypothesis of isotropic variation, i.e., the hypothesis that all the population principal component eigenvalues are equal. Under this null hypothesis, the population covariance matrix of the row vectors in  $X$  is  $\sigma^2 I$ . For simplicity, let us suppose that  $\sigma^2 = 1$  (if  $\sigma^2$  is an unknown value, we can estimate it by some  $\hat{\sigma}^2$  first and divide  $S$  by  $\hat{\sigma}^2$ ). Then the sample covariance matrix  $S$  satisfies

$$nS \sim W_p(I, n - 1).$$

The largest principal component eigenvalue  $\ell_1$  of  $S$  is a natural test statistic for a union intersection test. Our result applies for  $n\ell_1$ .

**Multidimensional scaling.** Let  $X$  be an  $n \times p$  data matrix. Consider the centered inner product matrix  $B = HXX'H$ , i.e.  $B_{ij} = (x_i - \bar{x})(x_j - \bar{x})$ . In a typical setting of multidimensional scaling, we are usually only given the matrix  $B$  instead of the original observations  $X$ . Let  $\lambda_1 \geq \dots \geq \lambda_p$  be the ordered eigenvalues of  $B$  and  $v_i$  be the corresponding eigenvector. As defined in Mardia et al. (1979, Section 14.3): for fixed  $k$  ( $1 \leq k \leq p$ ), the rows of  $V_k = (v_1, \dots, v_k)$  are called the principal coordinates of  $X$  in  $k$  dimensions, which constitute the classical  $k$ -dimensional solution to the multidimensional scaling problem.

We observe that the matrix  $B$  shares its non-zero eigenvalues with  $nS = X'HX$ . For the principal coordinate method to make sense, it is important that non-zero eigenvalues of  $B$  and hence all the eigenvalues of  $nS$  do not equal a common value. Translated to the population level, the population covariance matrix  $\Sigma \neq \sigma^2 I$ . Assuming  $\sigma^2 = 1$  (or dividing  $B$  by  $\sigma^2$  or its estimate  $\hat{\sigma}^2$ ), the null hypothesis can be written as  $H_0 : \Sigma = I$ . As in the situation of PCA, our result is useful for the test statistic  $\ell_1$ , where  $\ell_1$  is the largest eigenvalue of  $B$ .

**Testing that a covariance matrix equals a specified matrix.** Suppose that we have the Gaussian data matrix  $X$  with the rows  $x_1, \dots, x_n$  be independent  $N_p(\mu, \Sigma)$  random vectors and consider the null hypothesis  $H : \Sigma = \Sigma_0$ , where  $\Sigma_0$  is a specified positive definite matrix.

If the mean vector  $\mu$  is unknown, let  $S = n^{-1}X'HX$  be the sample covariance matrix. The union intersection test uses the largest eigenvalue of the matrix  $\Sigma_0^{-1}S$ , denoted by  $\lambda_1(\Sigma_0^{-1}S)$ , as the test statistic (see Mardia et al., 1979, p.130).

We observe that  $\lambda_1(\Sigma_0^{-1}S) = \lambda_1(\Sigma_0^{-1/2}S\Sigma_0^{-1/2})$ , where under the null hypothesis,

$$n\Sigma_0^{-1/2}S\Sigma_0^{-1/2} \sim W_p(I, n - 1).$$

Hence, our result is available for  $n\lambda_1(\Sigma_0^{-1}S)$ .

**Singular value decomposition.** For  $X$  a real  $n \times p$  matrix, there exists orthogonal matrices  $U(n \times n)$  and  $V(p \times p)$ , such that

$$X = UDV^T,$$

where  $D = \text{diag}(\sigma_1, \dots, \sigma_{\min(n,p)}) \in \mathbb{R}^{n \times p}$ , and  $\sigma_1 \geq \dots \geq \sigma_{\min(n,p)} \geq 0$ . This representation is called the singular value decomposition of  $X$  [See Golub and van Loan (1996, Theorem 2.5.2)]. For  $1 \leq i \leq \min(n, p)$ ,  $\sigma_i$  is called the  $i$ -th singular value of  $X$ . Theorem 1 then provides an accurate distributional approximation for  $\sigma_1^2$  when the entries of  $X$  are independent standard normal.

### 2.3 Other issues

For here, we provide brief remarks on several interesting issues that we come across during the development of this work. More details about them could be found in Appendix B.

**Transformation.** In the analysis of the greatest root statistic, Johnstone (2007) suggested that a nonlinear transformation [the logit transformation:  $\tau(x) = \log[x/(1-x)]$  in his case] helps improve the distributional approximation by the Tracy-Widom law, see Theorem 1, Table 1 and Fig. 1 there. In addition to its numerical effect, the transformation has an geometric explanation and yields a very natural integral representation for the correlation kernel which later appears in the central formula Eq.(50) there; see Johnstone (2007, Section 2.2, also Eq.'s (16) and (46)), Forrester (2004, Proposition 4.11) and Adler et al. (2000, Proposition 4.2).

Following Forrester (2004, Proposition 4.11), if we wanted to employ a comparable transformation for our white Wishart case, it would be the logarithmic transformation:  $\tau(x) = \log x$ . In fact, in our study, we first looked into some depth along this direction and could conclude the following second order accuracy result: under the condition of Theorem 1, let  $\nu_{np} = \log \tilde{\mu}_{np}$  and  $\tau_{np} = \tilde{\sigma}_{np}/\tilde{\mu}_{np}$ , there exists a continuous and nonincreasing function  $C(\cdot)$ , such that for all real  $s_0$ , there exists an integer  $N_0(s_0, \gamma)$  for which we have that for any  $s \geq s_0$  and  $n \wedge p \geq N_0(s_0, \gamma)$ ,

$$|P\{\log \lambda_1 \leq \nu_{np} + \tau_{np}s\} - F_1(s)| \leq C(s_0)(n \wedge p)^{-2/3} \exp(-s/2). \quad (5)$$

Some comments on how this result could be derived are included in B.1.

Although the rates of convergence are the same, numerical experiments suggest that using the nonlinear transformation does not yield as good numerical results in distributional approximation for small to moderate  $n$  and  $p$  as simply rescaling  $\lambda_1$  using (4), especially on the right-hand tail which is of the most statistical interest. When  $n$  and  $p$  grow large, using the transformation or not does not have as much influence, as they approach the same limit.

In consideration of the actual quality of approximation, especially for small to moderate  $n$  and  $p$ , we suggest not using the logarithmic transformation for the largest eigenvalues. However, it is of theoretical interest to know why such natural transformation works for the greatest root statistic in Johnstone (2007) but not for the largest eigenvalue in white Wishart matrices here.

**The smallest eigenvalue.** Following the principle of union intersection tests, the smallest eigenvalue could also serve as the test statistic in some cases, see, for instance, Mardia et al. (1979, Section 5.2.2c). Hence, what we have established for the largest eigenvalue is also worth investigation for the smallest one. Moreover, understanding the deviation of the smallest eigenvalue from its almost sure limit is also of independent interest. For example, it plays an important role in the theory of sparse signal recovery from large underdetermined linear system. See, for example, Donoho (2004) and Candes and Tao (2006). In fact, as we studied the accuracy result for the largest eigenvalue using the logarithmic transformation, we obtained a parallel result for smallest eigenvalues as a pleasant byproduct. We state without proof the result here.

Suppose that  $n-1 \geq p$  and  $n/p \rightarrow \gamma \in (1, \infty)$  and introduce the reflect Tracy-Widom law (Paul, 2006) as

$$G_1(s) = 1 - F_1(-s).$$

Let

$$\mu_{np}^- = \left( \sqrt{n - \frac{1}{2}} - \sqrt{p - \frac{1}{2}} \right)^2, \sigma_{np}^- = \left( \sqrt{n - \frac{1}{2}} - \sqrt{p - \frac{1}{2}} \right) \left( \frac{1}{\sqrt{p - \frac{1}{2}}} - \frac{1}{\sqrt{n - \frac{1}{2}}} \right)^{1/3},$$



and then define

$$\tau_{np}^- = \sigma_{np}^- / \mu_{np}^-, \quad \text{and} \quad \nu_{np}^- = \log \mu_{np}^- + \frac{1}{8} (\tau_{np}^-)^2. \quad (6)$$

We then have that for the smallest eigenvalue  $\lambda_p$  of a  $p \times p$  white Wishart matrix with  $n$  degrees of freedom, there exists a continuous and nondecreasing function  $C(\cdot)$ , such that for all real  $s_0$ , there exists an integer  $N_0(s_0, \gamma)$  for which we have that for any  $s \leq s_0$  and  $p \geq N_0(s_0, \gamma)$ ,

$$|P\{\log \lambda_p \leq \nu_{np}^- + \tau_{np}^- s\} - G_1(s)| \leq C(s_0) p^{-2/3} \exp(s/2). \quad (7)$$

See B.2 for remarks on how to prove this result.

Unlike the case for  $\lambda_1$ , the logarithmic transformation improves the numerical accuracy of the distributional approximation for  $\lambda_p$  significantly, especially when  $p$  is small and  $n/p$  is close to 1. We feel that an intuitive explanation to this phenomenon could be the following: for  $\lambda_p$ , the lower bound at 0 strongly affects the approximation on the original scale, especially when both  $p$  and  $n/p$  are small. However, by transforming  $\lambda_p$  to  $\log \lambda_p$ , one maps the lower bound to  $-\infty$  and hence avoids this ‘hard edge’ effect. The largest eigenvalue does not enjoy such a benefit for it does not have an algebraic upper bound.

As a numerical illustration, in Table 2, we present some simulation results on the Tracy-Widom approximation to smallest eigenvalues transformed as above for two  $n/p$  ratios: 2 : 1 and 4 : 1, both with  $p = 5, 10$  and 100. Again, for each combination of  $n$  and  $p$ , we run  $R = 40,000$  replications. The approximation seems good on the left-hand tail (where traditional significance levels locate) even for  $p$  as small as 5, regardless of the  $n/p$  ratio. Moreover, for both  $n/p$  ratios, when  $p$  grows to 100, the approximation becomes reasonably accurate over the entire range under investigation and is almost perfect on the left-hand tail. Therefore, the numerical results agree well with the theory for the smallest eigenvalues, too.

Percentiles	3.8954	3.1804	2.7824	1.9104	1.2686	0.5923	-0.4501	-0.9793	-2.0234
RTW	.99	.95	.90	.70	.50	.30	<b>.10</b>	<b>.05</b>	<b>.01</b>
10 × 5	1.000	.995	.976	.796	.553	.306	<b>.093</b>	<b>.045</b>	<b>.012</b>
20 × 10	.999	.984	.952	.760	.536	.305	<b>.098</b>	<b>.049</b>	<b>.011</b>
200 × 100	.993	.958	.910	.708	.504	.301	<b>.099</b>	<b>.050</b>	<b>.010</b>
20 × 5	.998	.977	.939	.745	.527	.306	<b>.097</b>	<b>.049</b>	<b>.010</b>
40 × 10	.996	.969	.926	.726	.511	.300	<b>.098</b>	<b>.048</b>	<b>.010</b>
400 × 100	.993	.955	.905	.703	.501	.301	<b>.100</b>	<b>.050</b>	<b>.010</b>
2 × SE	.001	.002	.003	.005	.005	.005	.003	.002	.001

Table 2: Simulations for finite  $n \times p$  vs. Tracy-Widom approximation: the smallest eigenvalue. For each combination of  $n$  and  $p$ , the estimated cumulative probabilities are computed for  $(\log \lambda_p - \nu_{np}^-) / \tau_{np}^-$  with  $R = 40,000$  draws from  $W_p(I, n)$ . The methods of sampling, computing  $F_1$  and obtaining percentiles are the same as in Table 1. The conventional significance levels are highlighted in bold font and the last line gives approximate standard errors based on binomial sampling.

### 3 Random Matrix Theory

The establishment of Theorem 1 relies heavily on results and methods from Random Matrix Theory (RMT) literature. In particular, those about unitary and orthogonal Laguerre matrix

ensembles play an important role. In this section, we first restate our main result using RMT terminology. With a Lipschitz-type bound, we transform the problem into the study of convergence rate of operators with matrix kernels and derive the closed form representation (15) of the top-left entry in the kernel for Laguerre orthogonal ensemble. Finally, we study the effect of scaling on our kernel representation and carefully formulate the analysis problem to be solved in later sections.

### 3.1 Restatement of Theorem 1 in Random Matrix Theory

Suppose  $A$  is an  $N \times N$  matrix following a  $W_N(I, n)$  distribution with  $n > N$ . [Here and after, following the RMT notational convention, we use  $N$  rather than  $p$  to denote the number of features.] The celebrated joint probability density function of the eigenvalues  $x_1 \geq \dots \geq x_N \geq 0$  is given by (Muirhead, 1982):

$$p_N(x_1, \dots, x_N) = d_{n,N}^{-1} \prod_{1 \leq j < k \leq N} (x_j - x_k) \prod_{j=1}^N x_j^{n-N-1} e^{-x_j/2},$$

where  $d_{n,N}$  is a normalizing constant depending only on  $n$  and  $N$ .

On the other hand, RMT people have investigated Laguerre Orthogonal Ensembles (LOE), where ‘ensemble’ stands for distribution of matrices and ‘orthogonal’ refers to the invariance of the distribution under orthogonal transformations. The LOE( $N, \tilde{\alpha}$ ) model ( $\tilde{\alpha} > -1$ ) has the matrix eigenvalue density as

$$\tilde{p}_N(x_1, \dots, x_N) = d_{\tilde{\alpha},N}^{-1} \prod_{1 \leq j < k \leq N} (x_j - x_k) \prod_{j=1}^N x_j^{\tilde{\alpha}} e^{-x_j/2}, \quad (8)$$

where  $x_1 \geq \dots \geq x_N \geq 0$  and  $d_{\tilde{\alpha},N}$  is a normalizing constant depending only on  $N$  and  $\tilde{\alpha}$ .

If we define  $\alpha_N = n - N$ , the joint eigenvalue density of white Wishart matrix  $A$  is exactly the eigenvalue density of the LOE( $N, \alpha_N - 1$ ) model. By this observation, we can formulate Theorem 1 in terms of RMT as the following.

**Theorem 2.** *Let  $x_1$  be the largest eigenvalue in the LOE( $N, \alpha_N - 1$ ) model and  $F_1$  be the orthogonal Tracy-Widom law. Define  $\tilde{\mu}_{n,N}$  and  $\tilde{\sigma}_{n,N}$  as*

$$\tilde{\mu}_{n,N} = \left( \sqrt{n - \frac{1}{2}} + \sqrt{N - \frac{1}{2}} \right)^2, \quad \tilde{\sigma}_{n,N} = \left( \sqrt{n - \frac{1}{2}} + \sqrt{N - \frac{1}{2}} \right) \left( \frac{1}{\sqrt{n - \frac{1}{2}}} + \frac{1}{\sqrt{N - \frac{1}{2}}} \right)^{1/3}. \quad (9)$$

*If  $n > N$ ,  $N \rightarrow \infty, n = n(N) \rightarrow \infty$  and  $n/N \rightarrow \gamma \in [1, \infty)$ , there exists a continuous and nonincreasing function  $C(\cdot)$ , such that for all real  $s_0$ , there exists an integer  $N_0(s_0, \gamma)$  for which we have that for any  $s \geq s_0$  and  $N \geq N_0(s_0, \gamma)$ ,*

$$|P\{x_1 \leq \tilde{\mu}_{n,N} + \tilde{\sigma}_{n,N}s\} - F_1(s)| \leq C(s_0)N^{-2/3} \exp(-s/2).$$

*Remark 2.* The theorem is stated only for situations where  $n > N$ . It works equally well when  $n < N$  by switching  $n$  and  $N$ . This results from the following observations: (a) constants in (9) are symmetric in  $n$  and  $N$  and (b) switching  $n$  and  $N$  does not change the distribution of  $x_1$ .

### 3.2 Operator determinant and kernel representation

We focus on the LOE( $N, \tilde{\alpha}$ ) model in (8) for the moment. For general orthogonal ensembles, Tracy and Widom (1998, Section 9) showed that when  $N$  is even, for  $\chi = I_{x>x'}$ :

$$F_{N,1}(x') \equiv P\{x_1 \leq x'\} = \sqrt{\det(I - K_N \chi)}, \quad (10)$$

with  $K_N$  an operator with a  $2 \times 2$  matrix kernel:

$$K_N(x, y) = \begin{pmatrix} I & -\partial_2 \\ \varepsilon_1 & T \end{pmatrix} S_{N,1}(x, y) - \begin{pmatrix} 0 & 0 \\ \varepsilon(x-y) & 0 \end{pmatrix}, \quad (11)$$

where  $\partial_2$  is the differential operator with respect to the second argument,  $\varepsilon_1$  is the convolution operator acting on the first argument with the kernel  $\varepsilon(x-y) = \frac{1}{2}\text{sgn}(x-y)$  and  $TS(x, y) = S(y, x)$  for any kernel  $S$ . However, no explicit representation of  $S_{N,1}$  was given there.

In a follow-up paper, Widom (1999) derived explicit expression of the kernel  $S_{N,1}$  for Gaussian and Laguerre orthogonal ensembles, which is summarized in Adler et al. (2000, Eq.(4.3)) in a more friendly form. In particular, for the LOE( $N, \tilde{\alpha}$ ) model of our interest, we have [Warning: we need to switch  $x$  and  $y$  in Adler et al. (2000, Eq.(4.3))]:

$$S_{N,1}(x, y) = S_{N,2}(x, y) + \frac{N!}{4\Gamma(N + \tilde{\alpha})} x^{\tilde{\alpha}/2} e^{-x/2} \left( \frac{d}{dx} L_N^{\tilde{\alpha}}(x) \right) \times \int_0^\infty \text{sgn}(y-z) z^{\tilde{\alpha}/2-1} e^{-z/2} [L_N^{\tilde{\alpha}}(z) - L_{N-1}^{\tilde{\alpha}}(z)] dz, \quad (12)$$

where  $L_k^{\tilde{\alpha}}$  ( $k = N-1, N$ ) are Laguerre polynomials defined in Szegö (1975, Chapter V) and  $S_{N,2}(x, y)$  is the kernel related to the Laguerre unitary ensemble (LUE) with parameter  $(N, \tilde{\alpha})$ , which has the following eigenvalue density:

$$p_N(x_1, \dots, x_N) = c_{n,N}^{-1} \prod_{1 \leq j < k \leq N} (x_j - x_k)^2 \prod_{j=1}^N x_j^{\tilde{\alpha}} e^{-x_j}, \quad x_1 \geq \dots \geq x_N \geq 0.$$

With (12), we start to derive an closed form representation for  $S_{N,1}$  after some necessary definitions. As in Johnstone (2001), we define a basis  $\{\phi_k\}_{k=0}^\infty$  on  $L^2([0, \infty))$  with transformed Laguerre polynomials

$$\phi_k(x; \tilde{\alpha}) = \sqrt{\frac{k!}{(k + \tilde{\alpha})!}} x^{\tilde{\alpha}/2} e^{-x/2} L_k^{\tilde{\alpha}}(x). \quad (13)$$

Then calling  $a_N = \sqrt{N(N + \tilde{\alpha})}$ , we follow El Karoui (2006b, Section 2) to introduce for  $x \geq 0$ ,

$$\phi(x; \tilde{\alpha}) = (-1)^N \sqrt{\frac{a_N}{2}} \phi_N(x; \tilde{\alpha} - 1) x^{-1/2}; \quad \psi(x; \tilde{\alpha}) = (-1)^{N-1} \sqrt{\frac{a_N}{2}} \phi_{N-1}(x; \tilde{\alpha} + 1) x^{-1/2}. \quad (14)$$

With the definition in (14), for the first term in (12), Johnstone (2001, Eq.(3.6)) and El Karoui (2006b, Appendix A.5) gave the following integral representation

$$S_{N,2}(x, y) = \int_0^\infty \phi(x+z) \psi(y+z) + \psi(x+z) \phi(y+z) dz.$$

For the second term, we could apply Szegö (1975, Eq.(5.1.13), (5.1.14)) to obtain that it equals

$$-\frac{N!}{4\Gamma(N + \tilde{\alpha})} x^{\tilde{\alpha}/2} e^{-x/2} L_{N-1}^{\tilde{\alpha}+1}(x) \int_0^\infty \text{sgn}(y-z) z^{\tilde{\alpha}/2-1} e^{-z/2} L_N^{\tilde{\alpha}-1}(z) dz = \psi(x) \int_0^\infty \varepsilon(y-z) \phi(z) dz.$$

Hence, we obtain

$$S_{N,1}(x, y) = S_{N,2}(x, y) + \psi(x) \int_0^\infty \varepsilon(y - z)\phi(z)dz.$$

Recall that for the white Wishart matrix  $A \sim W_N(I, n)$ , setting  $\alpha_N = n - N$ , it is connected to the LOE( $N, \tilde{\alpha}$ ) model by the identity  $\tilde{\alpha} = \alpha_N - 1$ . Thus, if we use the parameters  $N$  and  $\alpha_N$ , then the above calculation gives the following representation for  $S_{N,1}$ :

$$S_{N,1}(x, y; \alpha_N - 1) = S_{N,2}(x, y; \alpha_N - 1) + \psi(x; \alpha_N - 1) \int_0^\infty \varepsilon(y - z)\phi(z; \alpha_N - 1)dz. \quad (15)$$

### 3.2.1 Framework for deriving the determinant formula

The determinant formula (10) introduced at the beginning of this subsection provides the foundation for the convergence arguments. However, it is worth clarification under which framework it is derived.

Tracy and Widom (2005) described with care the operator convergence of  $K_N\chi$  to the limit  $K_{GOE}$  for the Hermite finite  $N$  ensemble. We adapt and extend their approach to the Laguerre finite  $N$  ensemble. Therefore, we paraphrase their remarks on the weighted Hilbert spaces and regularized 2-determinants under the current setting.

In the kernel  $K_N$  given in (11), the first term on the right hand side has each of its entries finite rank operators and hence a trace class operator. However, this is not true for  $\varepsilon(x - y)$ . According to Reed and Simon (1980, Theorem VI.23), it is even not Hilbert-Schmidt on  $L^2([x', \infty))$ . One way to take care of this problem is to introduce the weighted  $L^2$  space and to generalize the operator determinant as in Tracy and Widom (2005).

To this end, let  $\rho$  be any weight function which satisfies the following two conditions:

- (1) its reciprocal  $\rho^{-1} \in L^1([0, \infty))$ ; and
- (2) each operator that constitutes elements in the first term on the right hand side of (11) is in  $L^2([x', \infty); \rho) \cap L^2([x', \infty); \rho^{-1})$ .

Then, as remarked in Tracy and Widom (2005),  $\varepsilon: L^2([x', \infty); \rho) \rightarrow L^2([x', \infty); \rho^{-1})$  is Hilbert-Schmidt. Moreover,  $K_N$  could now be regarded as a  $2 \times 2$  matrix kernel on the space  $L^2([x', \infty); \rho) \oplus L^2([x', \infty); \rho^{-1})$ .

We have thus made clear on which space the kernel  $K_N$  acts. In order for the determinant formula (10) to hold, we need a generalization of the usual Fredholm determinant for trace class operators to determinant for Hilbert-Schmidt operators.

By our condition on  $\rho$ , for  $K_N = [K_{ij}]_{1 \leq i, j \leq 2}$ , we regard  $K_{11}$  and  $K_{22}$  as trace class operators on  $L^2([x', \infty); \rho)$  and  $L^2([x', \infty); \rho^{-1})$  respectively and off-diagonal elements as Hilbert-Schmidt operators:

$$K_{12} : L^2([x', \infty); \rho^{-1}) \rightarrow L^2([x', \infty); \rho) \quad \text{and} \quad K_{21} : L^2([x', \infty); \rho) \rightarrow L^2([x', \infty); \rho^{-1}).$$

Hence,  $\text{tr}(K_N) = \text{tr}(K_{11}) + \text{tr}(K_{22})$  is well defined. The regularized 2-determinant of Hilbert-Schmidt operator  $T$  with eigenvalues  $\mu_k$  is defined by

$$\det_2(I - T) = \prod_k (1 - \mu_k) e^{\mu_k}.$$

Then one naturally extends the operator definition of determinants to Hilbert-Schmidt operator matrix  $T$  with trace class diagonal entries by setting

$$\det(I - T) = \det_2(I - T) \exp(-\text{tr}T). \quad (16)$$

Finally, as remarked in Tracy and Widom (2005), the resulting notion of  $\det(I - K_N)$  is independent of the choice of  $\rho$  and allows the derivation in Tracy and Widom (1998) that yields (10), (11) and eventually (15).

Later in Section 4.1.1, we will make a specific choice of  $\rho$ , which not only makes our arguments more explicit but also eases the derivation of the right tail exponential decay in our desired bound.

### 3.3 Scaling the kernel

Fixing any real number  $s_0$  and introducing the linear transformation  $\tau(s) = \tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}$ , we are interested in the convergence rate of  $F_{N,1}(\tau(s'))$  to  $F_1(s')$  for all  $s' \geq s_0$ .

Define the rescaled kernel  $K_\tau$  as the following:

$$K_\tau(s, t) = \sqrt{\tau'(s)\tau'(t)} K_N(\tau(s), \tau(t)) = \tilde{\sigma}_{n,N} K_N(\tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}, \tilde{\mu}_{n,N} + t\tilde{\sigma}_{n,N}).$$

We have  $\det(I - K_N) = \det(I - K_\tau)$  by noticing that  $K_N$  and  $K_\tau$  share the spectrum. We give below an explicit representation of  $K_\tau$  for later use.

Before we proceed, we apply the  $\tau$ -scaling to  $\phi$ ,  $\psi$  and  $S_{N,2}$  and thus define

$$\phi_\tau(s) = \tilde{\sigma}_{n,N} \phi(\tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}), \quad \psi_\tau(s) = \tilde{\sigma}_{n,N} \psi(\tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}) \quad (17)$$

and

$$\begin{aligned} S_\tau(s, t) &= \tilde{\sigma}_{n,N} S_{N,2}(\tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}, \tilde{\mu}_{n,N} + t\tilde{\sigma}_{n,N}) \\ &= \int_0^\infty \phi_\tau(s+z)\psi_\tau(t+z) + \psi_\tau(s+z)\phi_\tau(t+z) dz. \end{aligned} \quad (18)$$

For later convenience,  $\phi_\tau(s)$  and  $\psi_\tau(s)$  are assumed to be 0 when  $\tau(s) = \tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N} < 0$ , and hence they are well-defined on the entire real line.

Finally, we introduce the short notation

$$S_\tau^R(s, t) = S_\tau(s, t) + \psi_\tau(s) \int_{-\infty}^\infty \varepsilon(t-z)\phi_\tau(z) dz = S_\tau(s, t) + \psi_\tau(s) (\varepsilon\phi_\tau)(t). \quad (19)$$

[We remind the reader that in the above discussion, we have dropped the explicit dependence on  $\tilde{\alpha}$  or  $\alpha_N - 1$  to avoid notation nightmare. Henceforth, we mention the explicit dependence only for eliminating ambiguity.]

We further observe that the determinant formula does not change if we modify  $K_\tau$  as

$$K_\tau(s, t) = \begin{pmatrix} K_{\tau,11}(s, t) & \tilde{\sigma}_{n,N} K_{\tau,12}(s, t) \\ \tilde{\sigma}_{n,N}^{-1} K_{\tau,21}(s, t) & K_{\tau,22}(s, t) \end{pmatrix},$$

for the spectrum does not change. Based on this observation and our detailed calculation in A.2, we could represent the entries of  $K_\tau$  as

$$\begin{aligned} K_{\tau,11}(s, t) &= S_\tau^R(s, t), & K_{\tau,12}(s, t) &= -\partial_t S_\tau^R(s, t), \\ K_{\tau,21}(s, t) &= (\varepsilon_1 S_\tau^R)(s, t) - \varepsilon(s-t), & K_{\tau,22}(s, t) &= K_{\tau,11}(t, s). \end{aligned} \quad (20)$$

For the desired limit  $F_1(s')$  of the sequence  $F_{N,1}(s')$ , Tracy and Widom (2005) showed that  $F_1(s') = \sqrt{\det(I - K_{GOE})}$ , where the operator  $K_{GOE}$  has the matrix kernel

$$K_{GOE}(s, t) = \begin{pmatrix} S(s, t) & SD(s, t) \\ IS(s, t) - \varepsilon(s - t) & S(t, s) \end{pmatrix}$$

with the entries given by

$$S(s, t) = S_A(s, t) + \frac{1}{2} \text{Ai}(s) \left( 1 - \int_t^\infty \text{Ai}(u) du \right), \quad SD(s, t) = -\partial_t S_A(s, t) - \frac{1}{2} \text{Ai}(s) \text{Ai}(t)$$

and  $IS(s, t) = -\int_s^\infty S_A(u, t) du + \frac{1}{2} \left( \int_t^s \text{Ai}(u) du + \int_s^\infty \text{Ai}(u) du \int_t^\infty \text{Ai}(u) du \right).$

Here  $S_A(s, t) = \int_0^\infty \text{Ai}(s+u) \text{Ai}(t+u) du$  represents the Airy kernel with  $\text{Ai}(\cdot)$  the Airy function defined in Olver (1974, p.53, Eq.(8.01)).

By our discussion in Section 3.2.1, it is necessary that both  $K_\tau$  and  $K_{GOE}$  belong to the following class  $\mathcal{A}$  of operators

$$\mathcal{A} \equiv \{2 \times 2 \text{ Hilbert-Schmidt operator matrices } A \text{ on } L^2([s', \infty); \rho \circ \tau) \oplus L^2([s', \infty); \rho^{-1} \circ \tau) \text{ with trace class diagonal entries}\}. \quad (21)$$

This fact will be verified after we choose a specific  $\rho$  in Section 4.1.1. For the convenience of argument, let us assume it for the moment.

### 3.4 Lipschitz bound and kernel difference

Let  $p_N = F_{N,1}(s')$  and  $p_\infty = F_1(s')$ , we note that  $|p_N - p_\infty| \leq |p_N^2 - p_\infty^2|/p_\infty \leq C(s_0)|p_N^2 - p_\infty^2|$ , where  $C(s_0) = 1/F_1(s_0)$  which is continuous and non-increasing in  $s_0$ . Thus, we are led to the difference of the determinants

$$|F_{N,1}(s') - F_1(s')| \leq C(s_0) |\det(I - K_\tau) - \det(I - K_{GOE})|. \quad (22)$$

*Remark 3.* Here and after, we use  $C(s_0)$  to denote in general any continuous and non-increasing function of  $s_0$  and  $C$  any universal constant, where the actual function and constant might be different from display to display.

To study the quantity on the right hand side of (22), our basic tool is the following Lipschitz-type bound on the matrix operator determinant for operators in  $\mathcal{A}$ .

**Proposition 1.** *For operators  $A$  and  $B$  in class  $\mathcal{A}$  and determinants of  $I - A$  and  $I - B$  defined as in (16), if  $\sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \leq 1/2$ , then*

$$|\det(I - A) - \det(I - B)| \leq M(B) \left( \sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \right), \quad (23)$$

where  $M(B) = 2 |\det(I - B)| + 2 \exp \left[ 2(1 + \|B\|_2)^2 + \sum_i \|B_{ii}\|_1 \right]$ .

Note that the leading term on the right hand side of (23) depends only on  $B$ . In this sense, Proposition 1 is a refinement of Proposition 3 in Johnstone (2007). Its proof could be found in A.5.

By Proposition 1, if we could control the entry-wise convergence rate of  $K_\tau$  to  $K_{GOE}$ , we will be able to bound the right hand side of (22) and hence prove our theorem. To this end, a convenient expression of the kernel difference  $K_\tau - K_{GOE}$  is helpful. we derive such an expression below by essentially adapting the arguments in Johnstone (2007, Section 8.3) to the current context.

According to Nagao and Forrester (1995, Eq.(4.2)), we could calculate [see A.1 for detail] that when  $N$  is even,

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_\tau(s; \alpha_N - 1) ds &= 0, \quad \text{and} \\ \int_{-\infty}^{\infty} \phi_\tau(s; \alpha_N - 1) ds &= \frac{N^{1/4}(n-1)^{1/4}}{2^{(\alpha_N-3)/2}(N+1)} \frac{\Gamma(\frac{N+3}{2})}{\Gamma(\frac{n+1}{2})} \left[ \frac{\Gamma(n)}{\Gamma(N+1)} \right]^{1/2}. \end{aligned}$$

For later use, we define  $\beta_N = \frac{1}{2} \int_{-\infty}^{\infty} \phi_\tau(s) ds$ .

Bring in the right-tail integration operator  $\tilde{\varepsilon}$  introduced in Johnstone (2007, Section 8.3) as

$$(\tilde{\varepsilon}g)(s) \equiv \int_s^{\infty} g(u) du, \quad (24)$$

we have the identity  $(\varepsilon g)(s) = \frac{1}{2} \int_{-\infty}^{\infty} g(u) du - (\tilde{\varepsilon}g)(s)$  and hence obtain

$$\varepsilon \phi_\tau = \beta_N - \tilde{\varepsilon} \phi_\tau, \quad \text{and} \quad \varepsilon \psi_\tau = -\tilde{\varepsilon} \psi_\tau.$$

For  $S_\tau$  defined in (18), by Fubini's theorem [justified by Lemma 1]

$$\int_{-\infty}^{\infty} S_\tau(u, t) du = 2\beta_N \int_0^{\infty} \psi_\tau(t+z) dz = 2\beta_N (\tilde{\varepsilon} \psi_\tau)(t).$$

Observing that for any kernel  $A(s, t)$ ,  $(\varepsilon_1 A)(s, t) = \frac{1}{2} \int_{-\infty}^{\infty} A(u, t) du - \int_s^{\infty} A(u, t) du$ , and introducing the abbreviation  $a \otimes b$  for rank one operator with kernel  $a(s)b(t)$ , we have  $\varepsilon_1 S_\tau = \beta_N \otimes \tilde{\varepsilon} \psi_\tau - \tilde{\varepsilon}_1 S_\tau$ , and for  $S_\tau^R$  in (19), we have  $S_\tau^R = S_\tau + \psi_\tau \otimes \beta_N - \psi_\tau \otimes \tilde{\varepsilon} \phi_\tau$ , which finally gives

$$\varepsilon_1 S_\tau^R = -\tilde{\varepsilon}_1 (S_\tau - \psi_\tau \otimes \tilde{\varepsilon} \phi_\tau) + \beta_N (1 \otimes \tilde{\varepsilon} \psi_\tau - \tilde{\varepsilon} \psi_\tau \otimes 1).$$

By the explicit expressions for  $K_\tau$  entries in (20),

$$K_\tau(s, t) = \begin{pmatrix} S_\tau^R(s, t) & -\partial_t S_\tau^R(s, t) \\ \varepsilon_1 S_\tau^R(s, t) & S_\tau^R(t, s) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -\varepsilon(s-t) & 0 \end{pmatrix} = L S_\tau^R + K^\varepsilon,$$

where

$$L = \begin{pmatrix} I & -\partial_2 \\ \varepsilon_1 & T \end{pmatrix} \quad \text{and} \quad K^\varepsilon = \begin{pmatrix} 0 & 0 \\ -\varepsilon & 0 \end{pmatrix}.$$

We then decompose  $K_\tau$  and  $K_{GOE}$  as follows:

$$K_\tau = K_\tau^R + K_{\tau,1}^F + K_{\tau,2}^F + K^\varepsilon \quad \text{and} \quad K_{GOE} = K^R + K_1^F + K_2^F + K^\varepsilon, \quad (25)$$

where by defining  $G = \text{Ai}/\sqrt{2}$  and the matrix kernels  $\tilde{L} = \begin{pmatrix} I & -\partial_2 \\ -\tilde{\varepsilon}_1 & T \end{pmatrix}$ ,  $L_1 = \begin{pmatrix} I & 0 \\ -\tilde{\varepsilon}_1 & 0 \end{pmatrix}$ , and

$L_2 = \begin{pmatrix} 0 & 0 \\ \tilde{\varepsilon}_2 & I \end{pmatrix}$ , we could write down the unspecified components in (25) explicitly as

$$\begin{aligned} K_\tau^R &= \tilde{L}[S_\tau - \psi_\tau \otimes \tilde{\varepsilon} \phi_\tau], & K_{\tau,1}^F &= \beta_N L_1[\psi_\tau(s)], & K_{\tau,2}^F &= \beta_N L_2[\psi_\tau(t)], \\ \text{and } K^R &= \tilde{L}[S_A - G \otimes \tilde{\varepsilon} G], & K_1^F &= \frac{1}{\sqrt{2}} L_1[G(s)], & K_2^F &= \frac{1}{\sqrt{2}} L_2[G(t)]. \end{aligned} \quad (26)$$

For  $\Delta_N$  to be defined in (61), we will establish in Lemma 1 that  $\phi_\tau = G + \Delta_N G' + O(N^{-2/3})$ , so set  $G_N = G + \Delta_N G'$ , we write the difference

$$\begin{aligned} K_\tau^R - K^R &= \tilde{L}[S_\tau - \psi_\tau \otimes \tilde{\varepsilon}\phi_\tau - S_A + G \otimes \tilde{\varepsilon}(G_N - \Delta_N G')] \\ &= \tilde{L}[S_\tau - S_A + \Delta_N G \otimes G] - \tilde{L}[\psi_\tau \otimes \tilde{\varepsilon}\phi_\tau - G \otimes \tilde{\varepsilon}G_N] = \delta^R + \delta_0^F. \end{aligned} \quad (27)$$

Set

$$S_{A_N}(s, t) = \int_0^\infty G(s+z)G_N(t+z) + G_N(s+z)G(t+z)dz,$$

since  $\int_0^\infty \text{Ai}(s+z)\text{Ai}'(t+z) + \text{Ai}'(s+z)\text{Ai}(t+z)dz = \int_0^\infty \frac{d}{dz} [\text{Ai}(s+z)\text{Ai}(t+z)] dz = -\text{Ai}(s)\text{Ai}(t)$ , we obtain

$$\delta^R = \tilde{L}[S_\tau - S_{A_N}]. \quad (28)$$

Finally, we organize the components of  $K_\tau - K_{GOE}$  as

$$K_\tau - K_{GOE} = \delta^R + \delta_0^F + \delta_1^F + \delta_2^F \quad (29)$$

where except for  $\delta_0^F$  and  $\delta^R$  given in (27) and (28), we further define  $\delta_i^F = K_{\tau,i}^F - K_i^F$  for  $i = 1, 2$ .

By the bounds (22) and (23), we need entrywise bounds on  $K_\tau - K_{GOE}$  to get our final convergence rate. By the decomposition in (29), the problem reduces to entrywise bounds for each of the  $\delta$ -terms. Since all these entries have explicit representations, this becomes an analysis problem which is to be solved in the next two sections.

## 4 Proof

In this section, we prove Theorem 2 [and hence Theorem 1] by focusing on the entries of the  $\delta$ -terms in (29). Besides the RMT analysis performed in Section 3, the proof needs two additional toolkits: a) asymptotics of transformed Laguerre polynomials, and b) several operator theoretic bounds of Hilbert-Schmidt and trace class norms.

### 4.1 Preliminaries

Here, we introduce some basic results for later repeated use in the proof. Moreover, we make a specific choice of the weight function  $\rho$ .

We start with Laguerre polynomial asymptotics. Recall that with constants  $\tilde{\mu}_{n,N}, \tilde{\sigma}_{n,N}$  in (9) and functions  $\phi, \psi$  defined in (14), we have defined transformed Laguerre polynomials  $\phi_\tau$  and  $\psi_\tau$  in (17). Moreover, for the Airy function, we define

$$G(s) = \frac{1}{\sqrt{2}} \text{Ai}(s). \quad (30)$$

By (25), (26) and (29), the kernels  $K_\tau$  and  $K_{GOE}$  and hence their difference are essentially expressed in terms of  $\phi_\tau, \psi_\tau, G$  and their variants. Therefore, we will find the following set of asymptotic bounds helpful to the analysis of their behavior.

**Lemma 1.** *Let  $\phi_\tau, \psi_\tau$  and  $G$  be defined as in (17) and (30) and  $\Delta_N$  to be defined in (61). If  $n > N, N \rightarrow \infty, n = n(N) \rightarrow \infty$  and  $n/N \rightarrow \gamma \in [1, \infty)$ , there exists a continuous and*



nonincreasing function  $C(\cdot)$ , such that for any real number  $s_0$ , there exists an integer  $N_0(s_0, \gamma)$  for which we have that for all  $s \geq s_0$  and  $N \geq N_0(s_0, \gamma)$ ,

$$|\psi_\tau(s)|, |\psi'_\tau(s)| \leq C(s_0) \exp(-s); \quad (31)$$

$$|\phi_\tau(s)|, |\phi'_\tau(s)| \leq C(s_0) \exp(-s); \quad (32)$$

$$|\psi_\tau(s) - G(s)|, |\psi'_\tau(s) - G'(s)| \leq C(s_0) N^{-2/3} \exp(-s); \quad (33)$$

$$|\psi_\tau(s) - G(s) - \Delta_N G'(s)|, |\psi'_\tau(s) - G'(s) - \Delta_N G''(s)| \leq C(s_0) N^{-2/3} \exp(-s). \quad (34)$$

In order not to distract us from the cause of proving Theorem 2, we defer the proof of Lemma 1 to Section 5. For the rest of Section 4, let us assume temporarily that Lemma 1 is already established.

In addition to the Laguerre polynomial asymptotics, we need some operator theoretic bounds of Hilbert-Schmidt and trace class norms. This set of tools has been previously established in Johnstone (2007, Section 8.4.1). For the sake of completeness, we state them here with some corrections and modifications that are helpful to our context.

From now on, we fix a real number  $s_0$  and consider any  $s' \in [s_0, \infty)$ . In general, let an operator  $T : L^2([s', \infty), \rho_1) \rightarrow L^2([s', \infty), \rho_2)$  defined by

$$f \mapsto Tf : (Tf)(u) = \int_{s'}^{\infty} T(u, v) f(v) dv \quad (35)$$

for some kernel  $T(u, v)$ . We obtain that the Hilbert-Schmidt norm  $\|T\|_2$  of  $T$  satisfies

$$\|T\|_2^2 = \iint_{[s', \infty)^2} |T(u, v)|^2 \rho_1^{-1}(v) \rho_2(u) dudv.$$

Following the notation in Johnstone (2007), we introduce the symbol  $\diamond$  for the following convolution type operator:

$$(a \diamond b)(u, v) \equiv \int_0^{\infty} a(u+z) b(v+z) dz.$$

Among all the operators defined by (35), we are interested in those with kernels  $D$  of the form  $D(u, v) = \alpha(u)\beta(v)$ , or  $D(u, v) = \alpha(u)\beta(v)(a \diamond b)(u, v)$ . We use the following notation for a Laplace-type transform:

$$\mathcal{L}(\rho)[t] \equiv \int_{s'}^{\infty} e^{-tz} \rho(z) dz.$$

For an operator with kernel of the form  $D(u, v) = \alpha(u)\beta(v)(a \diamond b)(u, v)$ , we have the following bound on its operator norm:

**Lemma 2.** *Let  $D$  be an operator taking  $L^2([s', \infty), \rho_1)$  to  $L^2([s', \infty), \rho_2)$  and having kernel  $D(u, v) = \alpha(u)\beta(v)(a \diamond b)(u, v)$ , where we assume, for  $u \geq s'$ , that*

$$|\alpha(u)| \leq \alpha_0 e^{\alpha_1 u}, \quad |\beta(u)| \leq \beta_0 e^{\beta_1 u}, \quad |a(u)| \leq a_0 e^{-a_1 u}, \quad |b(u)| \leq b_0 e^{-b_1 u}. \quad (36)$$

*If both  $\mathcal{L}(\rho_1^{-1})$  and  $\mathcal{L}(\rho_2)$  converge for  $t > c$ , and  $2(a_1 - \alpha_1), 2(b_1 - \beta_1) > c$ , the Hilbert-Schmidt norm satisfies*

$$\|D\|_2 \leq \frac{\alpha_0 \beta_0 a_0 b_0}{a_1 + b_1} \left\{ \mathcal{L}(\rho_2)[2(a_1 - \alpha_1)] \mathcal{L}(\rho_1^{-1})[2(b_1 - \beta_1)] \right\}^{1/2}.$$

*If  $\rho_1 = \rho_2$ , then the trace norm  $\|D\|_1$  satisfies the same bound.*

Next, we investigate rank one operators with kernels of the form  $D(u, v) = \alpha(u)\beta(v)$ . First, a remark taken verbatim from Tracy and Widom (2005): the norm of an operator  $D = \alpha \otimes \beta$  taking  $L^2(\rho_1)$  to  $L^2(\rho_2)$  with kernel  $D(u, v) = \alpha(u)\beta(v)$  is given by  $\|D\| = \|\alpha\|_{2, \rho_2} \|\beta\|_{2, \rho_1^{-1}}$ . Here, the norm can be trace class (if  $\rho_1 = \rho_2$ ) or Hilbert-Schmidt since they agree for rank one operators. Moreover, if  $\alpha$  and  $\beta$  satisfies the bound (36), similar derivation to that for proving Lemma 2 will give the following lemma specific for rank one operators.

**Lemma 3.** *Let  $D = \alpha \otimes \beta$  be a rank one operator taking  $L^2([s', \infty), \rho_1)$  to  $L^2([s', \infty), \rho_2)$  and having kernel  $D(u, v) = \alpha(u)\beta(v)$ , where we assume, for  $u \geq s'$ , that  $|\alpha(u)| \leq \alpha_0 e^{\alpha_1 u}$  and  $|\beta(u)| \leq \beta_0 e^{\beta_1 u}$ . If both  $\mathcal{L}(\rho_1^{-1})$  and  $\mathcal{L}(\rho_2)$  converge for  $t > c$  that  $-2\alpha_1, -2\beta_1 > c$ , the Hilbert-Schmidt norm satisfies*

$$\|D\|_2 \leq \alpha_0 \beta_0 \left\{ \mathcal{L}(\rho_2)[-2\alpha_1] \mathcal{L}(\rho_1^{-1})[-2\beta_1] \right\}^{1/2}.$$

If  $\rho_1 = \rho_2$ , then the trace norm  $\|D\|_1$  satisfies the same bound.

#### 4.1.1 Choice of the weight function $\rho$

In order to make our arguments explicit and to obtain the exponential decay of the right tail in our bound, we feel it convenient to make a specific choice of the weight function  $\rho$ .

In particular, for  $\nu \in (0, 1]$  and to be specified later in (45), on the  $s$ -scale, let

$$\rho \circ \tau(s) = 1 + \exp(\nu|s|). \quad (37)$$

The above definition implies that on the  $x$ -scale, we specify the weight function as

$$\rho(x) = 1 + \exp\left(\frac{\nu|x - \tilde{\mu}_{n,N}|}{\tilde{\sigma}_{n,N}}\right).$$

We remark that on the  $x$ -scale, our choice of  $\rho$  depends on  $N$ .

First of all, we check that our choice of  $\rho$  [on the  $x$ -scale] satisfies the two required conditions spelled out in Section 3.2.1. Condition (1) holds for  $\rho^{-1}(x) \asymp \exp(-\nu x/\tilde{\sigma}_{n,N})$  as  $x \rightarrow \infty$ . Condition (2) holds if  $\phi, \psi, \phi', \psi', \tilde{\varepsilon}\phi$  and  $\tilde{\varepsilon}\psi$  belong to  $L^2([x', \infty); \rho) \cap L^2([x', \infty); \rho^{-1})$ . We take  $\phi$  and  $\psi$  as examples, while the argument for the rest is essentially the same. By the definition of  $\phi_k$  in (13), the right tails of both  $\phi$  and  $\psi$  are bounded by  $\exp(-x/3)$ . On the other hand, as  $x \rightarrow \infty$ ,  $\rho^{\pm 1}(x) \asymp \exp(\pm \nu x/\tilde{\sigma}_{n,N})$  with  $\nu/\tilde{\sigma}_{n,N} \leq 1/\tilde{\sigma}_{n,N} = O(N^{-2/3})$ . These two facts suffice to show that both  $|\phi|^2 \rho^{\pm 1}$  and  $|\psi|^2 \rho^{\pm 1}$  are integrable over the region  $[x', \infty)$ , at least when  $N$  is large. Condition (2) is hence satisfied.

By (37), the operator class  $\mathcal{A}$  in (21) is now concrete. We now make valid all the formal derivation in Section 3 by verifying that  $K_\tau, K_{GOE} \in \mathcal{A}$ . Observing that  $\tau$  is linear, by Reed and Simon (1980, Theorem VI.22(h) and Theorem VI.23), condition (2) on  $\rho$  implies that  $K_\tau - K^\varepsilon \in \mathcal{A}$ . The super exponential decay (52) of the Airy functions, together with the same theorems as above, guarantees that  $K_{GOE} - K^\varepsilon \in \mathcal{A}$ . Hence, we need only to verify that  $\varepsilon : L^2([s', \infty); \rho \circ \tau) \rightarrow L^2([s', \infty); \rho^{-1} \circ \tau)$  is Hilbert-Schmidt, which is an immediate consequence of condition (1) on  $\rho$ .

From now on, we use  $\rho$  to denote  $\rho \circ \tau$  in (37) with no ambiguity, for all the remaining discussion in this paper focuses on the  $s$ -scale.

For the operator-theoretic bounds, by our choice of  $\rho$  in (37), we could adapt Lemma 2 and Lemma 3 into a more convenient form as follows.

**Corollary 1.** *With  $\rho$  as specified in (37), for  $\nu \leq \eta/2$ , we have*

$$\mathcal{L}(\rho^{\pm 1})[\eta] \leq \frac{4}{\eta - \nu} \exp(-\eta s' \pm \nu |s'|) \leq \frac{8}{\eta} \exp(-\eta s' \pm \nu |s'|). \quad (38)$$

*In particular, under the assumption of Lemma 2, if  $\{\rho_1, \rho_2\} \subset \{\rho, \rho^{-1}\}$  and  $a_1 - \alpha_1, b_1 - \beta_1 \geq \nu$ , then*

$$\|D\|_2, \|D\|_1 \leq C \frac{\alpha_0 \beta_0 a_0 b_0}{a_1 + b_1} \exp[-(a_1 + b_1 - \alpha_1 - \beta_1)s' + \nu |s'|], \quad (39)$$

where  $C = C(a_1, \alpha_1, b_1, \beta_1)$ .

*Under the assumptions of Lemma 3, if  $\{\rho_1, \rho_2\} \subset \{\rho, \rho^{-1}\}$  and  $-\alpha_1, -\beta_1 \geq \nu$ , then*

$$\|D\|_2, \|D\|_1 \leq C \alpha_0 \beta_0 \exp[(\alpha_1 + \beta_1)s' + \nu |s'|], \quad (40)$$

where  $C = C(\alpha_1, \beta_1)$ .

The proof of (38) follows directly from the derivation in Johnstone (2007, p.50); see, in particular, Eq.(205), (206) there. Then the operator bounds (39) and (40) are obtained by plugging (38) into the bounds in Lemma 2 and Lemma 3.

## 4.2 Operator convergence

With the tools from the previous subsection, we work out here entrywise bounds for each  $\delta$  term given in the decomposition (29).

**$\delta^R$  term.** Using the  $\diamond$  operator, we have  $\delta^R = \tilde{L}[S_\tau - S_{A_N}]$  with  $S_\tau = \phi_\tau \diamond \psi_\tau + \psi_\tau \diamond \phi_\tau$  and  $S_{A_N} = G_N \diamond G + G \diamond G_N$ . We shall use the abbreviation  $D^{(k)}f$ ,  $k = -1, 0$  and  $1$  to denote  $\tilde{\varepsilon}f$ ,  $f$  and  $f'$  respectively. Regardless of the signs, we have the following unified expression for the entries of  $\delta^R$ :

$$\begin{aligned} \delta_{ij}^R &= D^{(k)}(\phi_\tau - G_N) \diamond D^{(l)}\psi_\tau + D^{(k)}G_N \diamond D^{(l)}(\psi_\tau - G) \\ &\quad + D^{(k)}(\psi_\tau - G) \diamond D^{(l)}\phi_\tau + D^{(k)}G \diamond D^{(l)}(\phi_\tau - G_N), \end{aligned} \quad (41)$$

for  $i, j \in \{1, 2\}$ ,  $k \in \{-1, 0\}$  and  $l \in \{0, 1\}$ . By Lemma 1 and asymptotics of the Airy function [see (52)], we find that for any of the four terms in (41), the condition (36) is satisfied with  $\alpha_0 = \beta_0 = 1$ ,  $\alpha_1 = \beta_1 = 0$ ,  $a_1 = b_1 = 1$  and  $a_0, b_0$  as shown in the following table.

	$a_0$	$b_0$
$D^{(k)}(\phi_\tau - G_N) \diamond D^{(l)}\psi_\tau$	$C(s_0)N^{-2/3}$	$C(s_0)$
$D^{(k)}G_N \diamond D^{(l)}(\psi_\tau - G)$	$C(s_0)$	$C(s_0)N^{-2/3}$
$D^{(k)}(\psi_\tau - G) \diamond D^{(l)}\phi_\tau$	$C(s_0)N^{-2/3}$	$C(s_0)$
$D^{(k)}G \diamond D^{(l)}(\phi_\tau - G_N)$	$C(s_0)$	$C(s_0)N^{-2/3}$

We apply Corollary 1 and obtain that for  $\nu \leq 1$ ,

$$\|\delta_{ij}^R\| \leq C(s_0)N^{-2/3} \exp(-2s' + \nu |s'|). \quad (42)$$

Here and after, the unspecified norm  $\|\cdot\|$  denotes Hilbert-Schmidt norm  $\|\cdot\|_2$  if  $i \neq j$  and trace class norm  $\|\cdot\|_1$  otherwise. We remark that by a simple triangular inequality, we could choose the  $C(s_0)$  function in the last display as the sum of products of continuous and non-increasing functions, which could be seen from the term  $(\alpha_0 \beta_0 a_0 b_0)/(a_1 + b_1)$  in (39). Moreover, the term  $C$  in (39) is a universal constant for fixed  $a_1, \alpha_1, b_1$  and  $\beta_1$  here. Hence, the final  $C(s_0)$  function remains continuous and non-increasing. For the other  $\delta$  terms, we will have the same result by the same arguments and hence will be omitted.

$\delta_0^F$  **term.** We reorganize  $\delta_0^F$  as

$$\delta_0^F = -\tilde{L}[\psi_\tau \otimes \tilde{\varepsilon}\phi_\tau - G \otimes \tilde{\varepsilon}G_N] = -\tilde{L}[\psi_\tau \otimes \tilde{\varepsilon}(\phi_\tau - G_N) + (\psi_\tau - G) \otimes \tilde{\varepsilon}G_N] = \delta_0^{F,1} + \delta_0^{F,2}.$$

The entries of  $\delta_0^{F,i}$ ,  $i = 1, 2$  are all of the form  $\alpha(s)\beta(t)$  with the multipliers chosen from  $D^{(k)}\psi_\tau$ ,  $D^{(k)}(\phi_\tau - G_N)$ ,  $D^{(k)}(\psi_\tau - G)$  and  $D^{(k)}G_N$  for  $k \in \{-1, 0, 1\}$ . For these multipliers, the condition for Lemma 3 holds with the constants  $\alpha_1 = \beta_1 = -1$  and  $\alpha_0$  (or  $\beta_0$ ) specified below.

	$\alpha_0$ (or $\beta_0$ )
$D^{(k)}\psi_\tau$	$C(s_0)$
$D^{(k)}(\phi_\tau - G_N)$	$C(s_0)N^{-2/3}$
$D^{(k)}(\psi_\tau - G)$	$C(s_0)N^{-2/3}$
$D^{(k)}G_N$	$C(s_0)$

We apply Corollary 1 for these rank one terms and obtain that for  $\nu \leq 1$ ,

$$\|\delta_{0,ij}^F\| \leq \|\delta_{0,ij}^{F,1}\| + \|\delta_{0,ij}^{F,2}\| \leq C(s_0)N^{-2/3} \exp(-2s' + \nu|s'|). \quad (43)$$

$\delta_1^F$  and  $\delta_2^F$  **terms.** For these two terms, we have

$$\delta_1^F = L_1 \left[ \psi_\tau \otimes \beta_N - G \otimes \frac{1}{\sqrt{2}} \right] \quad \text{and} \quad \delta_2^F = L_2 \left[ \beta_N \otimes \psi_\tau - \frac{1}{\sqrt{2}} \otimes G \right].$$

By their similarity, we take  $\delta_1^F$  as example and the same analysis applies to  $\delta_2^F$  with obvious modification. For  $\delta_1^F$ , we reorganize it as

$$\delta_1^F = L_1 \left[ (\psi_\tau - G) \otimes \beta_N + G \otimes \left( \beta_N - \frac{1}{\sqrt{2}} \right) \right] = \delta_1^{F,1} + \delta_1^{F,2}.$$

For analysis of the terms here, Corollary 1 no longer works and we give an alternative bound which was derived in full detail in Johnstone (2007). In particular, consider matrices of rank one operators on  $L^2([s', \infty); \rho) \otimes L^2([s', \infty); \rho^{-1})$ , we denote, here and after, the  $L^2$ -norm on  $L^2([s', \infty); \rho)$  and  $L^2([s', \infty); \rho^{-1})$  by  $\|\cdot\|_+$  and  $\|\cdot\|_-$  respectively. Johnstone (2007, Eq.(214)) gives the following bound

$$\begin{pmatrix} \|a_{11} \otimes b_{11}\|_1 & \|a_{12} \otimes b_{12}\|_2 \\ \|a_{21} \otimes b_{21}\|_2 & \|a_{22} \otimes b_{22}\|_1 \end{pmatrix} \leq \begin{pmatrix} \|a_{11}\|_+ + \|b_{11}\|_- & \|a_{12}\|_+ + \|b_{12}\|_+ \\ \|a_{21}\|_- + \|b_{21}\|_- & \|a_{22}\|_- + \|b_{22}\|_+ \end{pmatrix}.$$

By the inequality above and our reorganization of  $\delta_1^F$ , we will see that the essential elements we need to bound are  $\|D^{(k)}(\psi_\tau - G)\|_\pm$ ,  $\|D^{(k)}G\|_\pm$  and  $\|1\|_-$  for  $k = -1$  and 0.

For  $\|D^{(k)}(\psi_\tau - G)\|_\pm$ , we obtain from Lemma 1 and (38) that for  $\nu \leq 1$ :

$$\|D^{(k)}(\psi_\tau - G)\|_\pm^2 \leq C^2(s_0)N^{-4/3}\mathcal{L}(\rho^{\pm 1})[2] \leq C^2(s_0)N^{-4/3} \exp(-2s' + \nu|s'|).$$

For  $\|D^{(k)}G\|_\pm$ , asymptotics of the Airy function and (38) give that for  $\nu \leq 1$ :

$$\|D^{(k)}G\|_\pm^2 \leq C^2(s_0)\mathcal{L}(\rho^{\pm 1})[2] \leq C^2(s_0) \exp(-2s' + \nu|s'|).$$

Finally, for  $\|1\|_-$ , we derive directly that

$$\begin{aligned} \|1\|_-^2 &= \int_{s'}^\infty [1 + \exp(\nu|s|)]^{-1} ds \leq \int_{s'}^\infty \exp(-\nu|s|) ds \\ &\leq \int_0^\infty \exp(-\nu s) ds + \int_{-|s'|}^0 \exp(\nu s) ds = \frac{1}{\nu} + \frac{1}{\nu} - \frac{1}{\nu} \exp(-\nu|s'|) \leq \frac{2}{\nu}. \end{aligned}$$

By definition of the operator  $L_1$  and our reorganization, we have the first column of  $\delta_1^F$  as following while the second column of it are zeros:

$$\begin{pmatrix} \delta_{1,11}^F \\ \delta_{1,21}^F \end{pmatrix} = \begin{pmatrix} (\psi_\tau - G) \otimes \beta_N + G \otimes (\beta_N - 1/\sqrt{2}) \\ -\tilde{\varepsilon}(\psi_\tau - G) \otimes \beta_N - \tilde{\varepsilon}G \otimes (\beta_N - 1/\sqrt{2}) \end{pmatrix}.$$

Assuming  $\beta_N - 1/\sqrt{2} = \mathcal{O}(N^{-1})$  [for a proof, see A.1], we have

$$\begin{aligned} \|\delta_{1,11}^F\|_1 &\leq \|(\psi_\tau - G) \otimes \beta_N\|_1 + \|G \otimes (\beta_N - 1/\sqrt{2})\|_1 \\ &\leq \|(\psi_\tau - G)\|_+ \|\beta_N\|_- + \|G\|_+ \|\beta_N - 1/\sqrt{N}\|_- \\ &\leq C(s_0)N^{-2/3}\nu^{-1/2} \exp(-s' + \nu|s'|/2) + C(s_0)N^{-1}\nu^{-1/2} \exp(-s' + \nu|s'|/2) \\ &\leq C(s_0)N^{-2/3}\nu^{-1/2} \exp(-s' + \nu|s'|/2) \leq C(s_0)N^{-2/3} \exp(-s'/2). \end{aligned}$$

The last inequality holds by fixing  $\nu$ , for example, at 1. By the same calculation, this bound also holds for  $\|\delta_{1,12}^F\|_2$  and those entries of  $\delta_2^F$ . Finally, we conclude our analysis with the following bound on entries of  $\delta_1^F$  and  $\delta_2^F$ : for  $\nu = 1$ , we have

$$\|\delta_{1,ij}^F\|, \|\delta_{2,ij}^F\| \leq C(s_0)N^{-2/3} \exp(-s'/2). \quad (44)$$

### 4.3 Proof of Theorem 2

Throughout the proof, we fix

$$\nu = 1 \quad (45)$$

in the weight function  $\rho$  specified in (37).

By (29) and the bounds (42), (43) and (44), we bound the entries of  $K_\tau - K_{GOE}$  using a simple triangular inequality

$$\|K_{\tau,ij} - K_{GOE,ij}\| \leq C(s_0)N^{-2/3} \exp(-s'/2).$$

Apply Proposition 1 with  $A = K_\tau$  and  $B = K_{GOE}$ ,

$$|\det(I - K_\tau) - \det(I - K_{GOE})| \leq M(K_{GOE})C(s_0)N^{-2/3} \exp(-s'/2), \quad (46)$$

where

$$M(K_{GOE}) = 2 \det(I - K_{GOE}) + 2 \exp \left[ 2(1 + \|K_{GOE}\|_2)^2 + \sum_i \|K_{GOE,ii}\|_1 \right].$$

For the first term in  $M(K_{GOE})$ , we have  $\det(I - K_{GOE}) = F_1^2(s') \leq 1$ . On the other hand, we have

$$\|K_{GOE}\|_2 \leq \sum_{i,j} \|K_{GOE,ij}\|_2 \leq \sum_i \|K_{GOE,ii}\|_1 + \sum_{i \neq j} \|K_{GOE,ij}\|_2.$$

In principle, one could show for each  $i$  and  $j$

$$\|K_{GOE,ij}\| \leq C(s_0),$$

with  $C(s_0)$  continuous and non-increasing. Here, we only take  $\|K_{GOE,11}\|_1$  as an example for the proof of the others is essentially the same. Let  $H_\tau$  and  $G_\tau$  be Hilbert-Schmidt operators with kernels  $\phi_\tau(x+y)$  and  $\psi_\tau(x+y)$  respectively, then as operator

$$K_{GOE,11} = H_\tau G_\tau + G_\tau H_\tau + G \otimes \frac{1}{\sqrt{2}} - G \otimes \tilde{\varepsilon}G.$$

By the relation  $\|AB\|_1 \leq \|A\|_2 \|B\|_2$ ,

$$\|K_{GOE,11}\|_1 \leq 2 \|H_\tau\|_2 \|G_\tau\|_2 + \frac{1}{\sqrt{2}} \|G\|_{2,\rho} \|1\|_{2,\rho^{-1}} + \|G\|_{2,\rho} \|\tilde{\varepsilon}G\|_{2,\rho^{-1}}.$$

Each norm on the right hand side of the above inequality is the square root of an integral of a positive function on  $[s', \infty)$  or  $[s', \infty)^2$  that is bounded by the corresponding integral over  $[s_0, \infty)$  or  $[s_0, \infty)^2$ , which in turn is continuous and non-increasing in  $s_0$ . Hence,  $\|K_{GOE,11}\|_1 \leq C(s_0)$ .

By the above discussion, we could control the second term of  $M(K_{GOE})$  and hence itself by a continuous and non-increasing  $C(s_0)$ . Finally, we complete the proof by combining this fact with the initial bounds (46) and (22).

## 5 Laguerre Polynomial Asymptotics

In this section, our goal is to establish Lemma 1. To this end, we exploit the Liouville-Green approach to study the related asymptotics for Laguerre polynomials of both large order and large degree. This approach has been successfully used in Johnstone (2001), El Karoui (2006b) and more recently, Johnstone (2007) in deriving similar type of results. The novelty here is the establishment of the bounds (33) and (34) for the derivatives of these polynomials.

To start with, let us consider the ‘‘intermediate’’ function  $F_{n,N}$  introduced in El Karoui (2006b, Section 2.2.2) as

$$F_{n,N}(x) \equiv (-1)^N \sigma_{n,N}^{-1/2} \sqrt{N!/n!} x^{(\alpha_N+1)/2} e^{-x/2} L_N^{\alpha_N}(x) \quad (47)$$

with  $\alpha_N = n - N$ . We could then relate  $F_{n,N}$  to  $\phi_N$ ,  $\phi$  and  $\phi_\tau$  as

$$\begin{aligned} \phi_N(x; \alpha_N) &= (-1)^N \sigma_{n,N}^{1/2} x^{-1/2} F_{n,N}(x), \\ \phi(x; \alpha_N - 1) &= \frac{N^{1/4} (n-1)^{1/4}}{\sqrt{2}} \sigma_{n-2,N}^{1/2} F_{n-2,N}(x)/x, \\ \text{and } \phi_\tau(s) &= \frac{1}{\sqrt{2}} \left( \frac{N^{1/4} (n-1)^{1/4} \sigma_{n-2,N}^{1/2} \tilde{\sigma}_{n,N}}{\mu_{n-2,N}} \right) F_{n-2,N}(\tilde{\mu}_{n,N} + s \tilde{\sigma}_{n,N}) \left( \frac{\mu_{n-2,N}}{\tilde{\mu}_{n,N} + s \tilde{\sigma}_{n,N}} \right), \end{aligned}$$

with  $\mu_{n,N}$  and  $\sigma_{n,N}$  defined as

$$\mu_{n,N} = \left( \sqrt{n_+} + \sqrt{N_+} \right)^2 \quad \text{and} \quad \sigma_{n,N} = \left( \sqrt{n_+} + \sqrt{N_+} \right) \left( \frac{1}{\sqrt{n_+}} + \frac{1}{\sqrt{N_+}} \right)^{1/3},$$

using the abbreviations  $n_+ = n + \frac{1}{2}$  and  $N_+ = N + \frac{1}{2}$ . If we replace the subscripts  $(n-2, N)$  in  $\mu_{n-2,N}$ ,  $\sigma_{n-2,N}$  and  $F_{n-2,N}$  by  $(n-1, N-1)$  on the right hand sides of the expressions for  $\phi(x; \alpha_N - 1)$  and  $\phi_\tau(s)$ , we obtain the identities for  $\psi(x; \alpha_N - 1)$  and  $\psi_\tau(s)$ . Due to this close connection of  $\phi_\tau$  and  $\psi_\tau$  to  $F_{n,N}$ , the essential element for proving the desired asymptotic bounds reduces to the understanding of the behavior of  $F_{n,N}$  and its derivative, for which the Liouville-Green approach is instrumental.

In the rest of this section, we first study in detail the Liouville-Green approximation to the  $F_{n,N}$  function and its derivative. Then the result is used to facilitate the derivation of the global bounds and the local as well as global Airy approximation to  $\phi_\tau$ ,  $\psi_\tau$  and their derivatives.

## 5.1 Liouville-Green approach

Many of the arguments in this part have been spelled out in some detail in Johnstone (2001) and El Karoui (2006b). A more complete account of the theory could be found in Olver (1974, Chapter 11). However, for completeness, we state them here briefly with notation similar to that in El Karoui (2006b).

Consider  $w_N(x) = x^{(\alpha_N+1)/2} e^{-x/2} L_N^{\alpha_N}(x)$  as a multiple of  $F_{n,N}$ , we have

$$\frac{d^2 w_N}{dx^2} = \left\{ \frac{1}{4} - \frac{\kappa_N}{x} + \frac{\lambda_N^2 - 1/4}{x^2} \right\} w_N \quad (48)$$

with  $\kappa_N = N + \frac{\alpha_N+1}{2} = \frac{n+N+1}{2}$  and  $\lambda_N = \frac{\alpha_N}{2} = \frac{n-N}{2}$ .

By a change of variable  $\xi = x/\kappa_N$ , we obtain

$$\frac{d^2 w_N}{d\xi^2} = \{ \kappa^2 f(\xi) + g(\xi) \} w_N,$$

where

$$f(\xi) = \frac{(\xi - \xi_-)(\xi - \xi_+)}{4\xi^2} \quad \text{and} \quad g(\xi) = \frac{1}{4\xi^2},$$

with  $\xi_{\pm} = 2 \pm \sqrt{4 - \omega_N^2}$  and  $\omega_N = 2\lambda_N/\kappa_N = \frac{2(n-N)}{n+N+1}$ . The Liouville-Green method introduces the change of independent variable as

$$\frac{2}{3}\zeta^{3/2} = \int_{\xi_+}^{\xi} \sqrt{f(t)} dt \quad (\xi \geq \xi_+) \quad \text{and} \quad \frac{2}{3}(-\zeta)^{3/2} = \int_{\xi}^{\xi_+} \sqrt{-f(t)} dt \quad (\xi \leq \xi_+),$$

and defines a new dependent variable  $W = (d\zeta/d\xi)^{1/2} w_N$ . For the new pair  $(W, \zeta)$ , we have the new differential equation as

$$\frac{d^2 W}{d\zeta^2} = \{ \kappa_N^2 \zeta + v(\omega_N, \zeta) \} W.$$

Let  $\hat{f} = f/\zeta$ , the recessive solution of (48) satisfies (Olver, 1974, p.399, Theorem 3.1)

$$w_N(\kappa_N \xi) \propto \hat{f}^{-1/4}(\xi) \{ \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \},$$

with the following estimates for the error term  $\varepsilon_2$  and its derivative with  $\xi \in [2, \infty)$ :

$$\begin{aligned} |\varepsilon_2(\kappa_N, \xi)| &\leq \mathbf{M}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \left[ \exp\left(\frac{\lambda_0}{\kappa_N} F(\omega_N)\right) - 1 \right], \\ |\partial_{\xi} \varepsilon_2(\kappa_N, \xi)| &\leq \kappa_N^{2/3} \hat{f}^{1/2}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \left[ \exp\left(\frac{\lambda_0}{\kappa_N} F(\omega_N)\right) - 1 \right]. \end{aligned}$$

In the above bounds,  $\mathbf{M}, \mathbf{E}$  are the modulus and weight functions for the Airy function, and  $\mathbf{N}$  the phase function for its derivative (Olver, 1974, pp.394-396). Moreover,  $\lambda_0 \doteq 1.04$  and  $F(\omega_N)$  has been well studied in El Karoui (2006b, A.3).

For the function  $F_{n,N}$  of our interest, we have from El Karoui (2006b, Eq.(5) and A.1) that

$$F_{n,N}(x) = r_N \left( \frac{\kappa_N}{\sigma_{n,N}^3} \right)^{1/6} \hat{f}^{-1/4}(\xi) \{ \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \},$$

with

$$r_N^2 = \frac{2\pi \exp[-(n_+ + N_+)] n_+^{n_+} N_+^{N_+}}{N! n!} = 1 + O(n^{-1}, N^{-1}). \quad (49)$$

For the convenience of argument, we define an auxiliary function  $R_N(\xi) = (\dot{\zeta}(\xi)/\dot{\zeta}_N)^{-1/2}$  with  $\dot{\zeta}_N = \dot{\zeta}(\xi_+)$ . We remark that by our definition, we have  $\sigma_{n,N} = (\kappa_N^{-1/3} \dot{\zeta}_N)^{-1}$  and  $\hat{f} = \dot{\zeta}(\xi)^2$ . Hence,  $F_{n,N}$  could be rewritten as

$$F_{n,N} = r_N R_N(\xi) \{ \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \}. \quad (50)$$

Finally, we conclude this part with some useful bounds and asymptotics of  $\mathbf{M}, \mathbf{E}, \mathbf{N}$  and the Airy function (Olver, 1974, pp.392-397). As  $x \rightarrow \infty$ , we have

$$\mathbf{E}(x) \sim \sqrt{2} e^{\frac{2}{3}x^{3/2}}, \quad \mathbf{M}(x) \sim \pi^{-1/2} x^{-1/4}, \quad \mathbf{N}(x) \sim \pi^{-1/2} x^{1/4}. \quad (51)$$

For all  $x > 0$ , the Airy function and its derivative are bounded as

$$0 \leq \text{Ai}(x) \leq \frac{e^{-\frac{2}{3}x^{3/2}}}{2\pi^{1/2}x^{1/4}}, \quad |\text{Ai}'(x)| \leq \left(1 + \frac{7}{48x^{3/2}}\right) \frac{x^{1/4}e^{-\frac{2}{3}x^{3/2}}}{2\pi^{1/2}}. \quad (52)$$

Finally, for all  $x$ , we have the following bounds

$$|\text{Ai}(x)| \leq \mathbf{M}(x)\mathbf{E}^{-1}(x), \quad |\text{Ai}'(x)| \leq \mathbf{N}(x)\mathbf{E}^{-1}(x), \quad \mathbf{M}(x) \leq 1, \quad \mathbf{E}(x) \geq 1, \quad (53)$$

and finally,  $\mathbf{E}(x)$  is monotone increasing in  $x$  (Olver, 1974, p.395).

## 5.2 Large $N$ asymptotics

We now derive the large  $N$  asymptotics of  $\phi_\tau, \psi_\tau$  and related functions. First, we use the analysis done in Johnstone (2001) and El Karoui (2006b) to obtain bounds for  $|\psi_\tau|$  and  $|\psi_\tau - G|$  without much extra effort. Then we derive the bounds for  $|\psi'_\tau|$  and  $|\psi'_\tau - G|$ , which need some careful analysis to be detailed below and the bound on  $|\psi_\tau - G|$  is then further refined to match the claim in Lemma 1. Finally, corresponding results for quantities related to  $\phi_\tau$  could be obtained by understanding the difference of the centering and scaling constants involved in  $\phi_\tau$  and  $\psi_\tau$ .

### 5.2.1 Bounds for $|\psi_\tau(s)|$ and $|\psi_\tau(s) - G(s)|$

We define  $x_{n,N}(s) = \mu_{n,N} + s\sigma_{n,N}$  and let

$$\theta_{n,N}(x_{n,N}(s)) \equiv F_{n,N}(x_{n,N}(s)) \left( \frac{\mu_{n,N}}{x_{n,N}(s)} \right).$$

Johnstone (2001, A.8) showed that under the condition of Lemma 1,

$$\left| F_{n,N}(x_{n,N}(s)) \sigma_{n,N}^{1/2} N^{-1/6} \right| \leq C \exp(-s), \quad \text{for all } s \geq 0.$$

Simple manipulation gives  $\sigma_{n,N}^{-1/2} N^{1/6} \leq \sigma_{n,N}^{-1/2} N_+^{1/6} \leq (N_+/n_+)^{1/2} < 1$ , and hence for all  $s \geq 0$ ,

$$|F_{n,N}(x_{n,N}(s))| \leq C \exp(-s).$$

If  $s_0 < 0$ , by (49), (51), (73) and El Karoui (2006b, A.3), we obtain that when  $N \geq N_0(s_0, \gamma)$ ,

$$|F_{n,N}(x_{n,N}(s))| \leq r_N |R_N(\xi)| \mathbf{M}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \leq 2 \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \leq 2.$$



If we let  $M(s_0) = \max_{s \in [s_0, 0]} \{2e^s\}$ , and define

$$C(s_0) = \max\{C, M(s_0)I_{s_0 < 0}, M(0)\},$$

it is then continuous and non-increasing in  $s_0$  as desired and we have that when  $N \geq N_0(s_0, \gamma)$ ,  $|F_{n,N}(x_{n,N}(s))| \leq C(s_0) \exp(-s)$  for all  $s \geq s_0$ . Moreover, by noting  $\sigma_{n,N}/\mu_{n,N} = O(N^{-2/3})$ , when  $N$  is larger than some constant that depends only on  $s_0$ ,

$$\frac{\mu_{n,N}}{x_{n,N}(s)} \leq \left(1 + s_0 \frac{\sigma_{n,N}}{\mu_{n,N}}\right)^{-1} \leq 2, \quad \text{for all } s \geq s_0.$$

Hence, under the condition of Lemma 1, we have that when  $N \geq N_0(s_0, \gamma)$ ,

$$|\theta_{n,N}(x_{n,N}(s))| \leq C(s_0) \exp(-s), \quad \text{for all } s \geq s_0.$$

Later on, El Karoui (2006b, Section 3.2) showed that for any constant  $\varrho_N = 1 + O(N^{-1})$ , if we define  $\Delta_{n,N}(x_{n,N}(s)) = |\varrho_N \theta_{n,N}(x_{n,N}(s)) - \text{Ai}(s)|$ , then under the condition of Lemma 1, we have

$$N^{2/3} \Delta_{n,N}(x_{n,N}(s)) \leq C(s_0) \exp(-s/2), \quad \text{for all } s \geq s_0.$$

For  $\psi_\tau(s)$ , we have  $\tilde{\mu}_{n,N} = \mu_{n-1,N-1}$  and  $\tilde{\sigma}_{n,N} = \sigma_{n-1,N-1}$  and hence it is of the form  $\frac{1}{\sqrt{2}} \rho_N \theta_{n-1,N-1}(x_{n-1,N-1}(s))$ . Noting that  $\rho_N = 1 + O(N^{-1})$  [see A.1 for a proof], we apply the bounds for  $\theta_{n,N}$  and  $\Delta_{n,N}$  directly and obtain that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ ,

$$|\psi_\tau(s)| \leq C(s_0) \exp(-s), \quad \text{and} \quad |\psi_\tau(s) - G(s)| \leq C(s_0) N^{-2/3} \exp(-s/2), \quad \text{for all } s \geq s_0.$$

Actually the bound on  $|\psi_\tau(s) - G(s)|$  could be further improved to be that claimed in Lemma 1: see (59) for the refinement. We also remark that we could not apply the results directly to  $\phi_\tau$  since the centering and scaling constants  $(\mu_{n-2,N}, \sigma_{n-2,N})$  specific to  $F_{n-2,N}$  does not agree with the global constants  $(\tilde{\mu}_{n,N}, \tilde{\sigma}_{n,N})$  which we use.

## 5.2.2 Bounds for $|\psi'_\tau(s)|$ and $|\psi'_\tau(s) - G'(s)|$

As we have seen, the analysis of  $\psi_\tau$  depends on our understanding of the function  $\theta_{n,N}(x_{n,N}(s))$ . To investigate the bounds for  $\psi'_\tau$  and its approximation by  $G'$ , we start with a detailed analysis of the quantity  $\partial_s \theta_{n,N}(x_{n,N}(s))$ .

We split  $\partial_s \theta_{n,N}(x_{n,N}(s))$  into two parts:

$$\begin{aligned} |\partial_s \theta_{n,N}(x_{n,N}(s))| &\leq \left| \sigma_{n,N} F'_{n,N}(x_{n,N}(s)) \frac{\mu_{n,N}}{x_{n,N}(s)} \right| + \left| \sigma_{n,N} F_{n,N}(x_{n,N}(s)) \frac{\mu_{n,N}}{x_{n,N}^2(s)} \right| \\ &= T_{N,1}(s) + T_{N,2}(s). \end{aligned}$$

**$T_{N,2}$  term.** This term is relatively easy to bound. Note that  $T_{N,2}(s) = |\theta_{n,N}(x_{n,N}(s)) \sigma_{n,N} / x_{n,N}(s)|$  and that  $\sigma_{n,N} / \mu_{n,N} = O(N^{-2/3})$ . When  $N \geq N_0(s_0)$ , the ratio

$$|\sigma_{n,N} / x_{n,N}(s)| = |s + \mu_{n,N} / \sigma_{n,N}|^{-1} \leq C(s_0) N^{-2/3}, \quad \text{for all } s \geq s_0.$$

Hence, by our previous bound on  $|\theta_{n,N}|$ , we obtain that under the condition of Lemma 1,

$$T_{N,2}(s) \leq C(s_0) N^{-2/3} \exp(-s), \quad \text{for all } s \geq s_0.$$

$T_{N,1}$  **term.** Recalling that  $\mu_{n,N}/x_{n,N}(s)$  could be bounded by 2, we focus on  $\sigma_{n,N}F'_{n,N}$ . Thinking of  $x = x_{n,N}(s)$ , we have from (50) that

$$\begin{aligned}\sigma_{n,N}F'_{n,N}(x) &= r_N \left( \frac{\sigma_{n,N}}{\kappa_N} \right) R'_N(\xi) \left[ \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \right] \\ &\quad + r_N R_N(\xi) \left[ \text{Ai}'(\kappa_N^{2/3} \zeta) R_N^{-2}(\xi) + \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \frac{\partial}{\partial \xi} \varepsilon_2(\kappa_N, \xi) \right].\end{aligned}$$

To facilitate our analysis, on the  $s$ -scale, we divide the whole region  $[s_0, \infty)$  as  $I_{1,N} \cup I_{2,N}$  with  $I_{1,N} = [s_0, s_1 N^{1/6})$  and  $I_{2,N} = [s_1 N^{1/6}, \infty)$ . The choice of  $s_1$  is made explicit in A.4.

*Case  $s \in I_{1,N}$ .* In this case, we first reorganize  $\sigma_{n,N}F'_{n,N}(x)$  as  $\sigma_{n,N}F'_{n,N}(x) = \sum_{i=1}^4 D_{n,N}^i$ , with

$$\begin{aligned}D_{n,N}^1 &= r_N \left( \frac{\sigma_{n,N}}{\kappa_N} \right) R'_N(\xi) \{ \text{Ai}(\kappa_N^{2/3} \xi) + \varepsilon_2(\kappa_N, \xi) \}, & D_{n,N}^2 &= r_N [R_N^{-1}(\xi) - 1] \text{Ai}'(\kappa_N^{2/3} \zeta), \\ D_{n,N}^3 &= r_N \text{Ai}'(\kappa_N^{2/3} \zeta), & D_{n,N}^4 &= r_N \left( \frac{\sigma_{n,N}}{\kappa_N} \right) R_N(\xi) \frac{\partial}{\partial \xi} \varepsilon_2(\kappa_N, \xi).\end{aligned}$$

(i) Consider  $D_{n,N}^1$  first. Direct computation shows  $N^{2/3}(\sigma_{n,N}/\kappa_N) \rightarrow 2(1+1/\sqrt{\gamma})^{1/3}(1+\sqrt{\gamma})(1+\gamma)^{-1}$ . Hence when  $N \geq N_0(\gamma)$ , we have the bound

$$N^{2/3} \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \leq C \left( 1 + \frac{1}{\sqrt{\gamma}} \right)^{1/3} \frac{1 + \sqrt{\gamma}}{1 + \gamma}.$$

Moreover, by the bound (72) for  $R'_N$  and recalling that  $\gamma \geq 1$ , we know that when  $N \geq N_0(s_0, \gamma)$ ,

$$N^{2/3} \left( \frac{\sigma_{n,N}}{\kappa_N} \right) |R'_N(\xi)| \leq C \left( 1 + \frac{1}{\sqrt{\gamma}} \right)^{1/3} \frac{\sqrt{\gamma}}{1 + \sqrt{\gamma}} \leq C, \quad \text{for all } s \in I_{1,N}. \quad (54)$$

On the other hand, by (51) and (74), we obtain

$$\left| \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \right| \leq C \mathbf{M}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \leq C \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta).$$

When  $s \geq 0$ , we know from (74) and the monotonicity of  $\mathbf{E}$  that when  $N \geq N_0(s_0, \gamma)$ ,  $\kappa_N^{2/3} \zeta \geq s/2$  holds, and hence by (51),

$$\left| \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \right| \leq C \mathbf{E}^{-1}(s/2) \leq C \exp\left(-\frac{1}{3\sqrt{2}} s^{3/2}\right) \leq C \exp(-s).$$

If  $s_0 \leq 0$ , for all  $s \in [s_0, 0]$ , we obtain from (74) that when  $N \geq N_0(s_0, \gamma)$ ,  $\kappa_N^{2/3} \zeta \in [3s_0/2, 1]$ . Hence, we have

$$\mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \exp(s) \leq C(s_0) \equiv \max_{s \in [3s_0/2, 1]} e \mathbf{E}^{-1}(s),$$

the right hand side of which is, by its definition, continuous and non-increasing. Therefore, we conclude that when  $N \geq N_0(s_0, \gamma)$ ,

$$\left| \text{Ai}(\kappa_N^{2/3} \zeta) + \varepsilon_2(\kappa_N, \xi) \right| \leq C(s_0) \exp(-s), \quad \text{for all } s \in I_{1,N}. \quad (55)$$

Finally, putting the bounds (54) and (55) together and recalling that  $|r_N|$  could be bounded by 2, we obtain that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ , on  $I_{1,N}$ ,

$$|D_{n,N}^1| \leq C(s_0) N^{-2/3} \exp(-s).$$

(ii) For  $D_{n,N}^2$ , we first split and control  $|r_N R_N^{-1}(\xi) - 1|$  as

$$|r_N R_N^{-1}(\xi) - 1| \leq r_N |R_N^{-1}(\xi) - 1| + |r_N - 1| = r_N |R_N(\xi)|^{-1} |R_N(\xi) - 1| + |r_N - 1|.$$

By (49) and (66), when  $N \geq N_0(s_0, \gamma)$ , we have  $|r_N| \leq 2$ ,  $|R_N(\xi)|^{-1} \leq 2$  and hence

$$|r_N R_N^{-1}(\xi) - 1| \leq C N^{-2/3} s + C N^{-1} \leq C N^{-2/3} s, \quad \text{for all } s \in I_{1,N}. \quad (56)$$

On the other hand, by (53), we obtain

$$\left| \text{Ai}'(\kappa_N^{2/3} \zeta) \right| \leq \mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta).$$

When  $s \geq 0$ , we have from (74) that  $\kappa_N^{2/3} \zeta \in [s/2, 3s/2]$ , and using (51), we obtain

$$\mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \leq C (\kappa_N^{2/3} \zeta)^{1/4} \exp\left(-\frac{1}{3\sqrt{2}} s^{3/2}\right) \leq C s^{1/4} \exp\left(-\frac{1}{3\sqrt{2}} s^{3/2}\right) \leq C \exp(-3s/2).$$

If  $s_0 \leq 0$ , we know that when  $N \geq N_0(s_0, \gamma)$ ,  $\kappa_N^{2/3} \zeta \in [3s_0/2, 1]$  for all  $s \in [s_0, 0]$ . We then have

$$\mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}(\kappa_N^{2/3} \zeta) \exp(3s/2) \leq C(s_0) \equiv \max_{s \in [3s_0/2, 1]} e^{3/2} \mathbf{N}(s) \mathbf{E}(s),$$

the right hand side of which is again continuous and non-increasing in  $s_0$ . As before, this enables us to conclude that when  $N \geq N_0(s_0, \gamma)$ ,

$$\left| \text{Ai}'(\kappa_N^{2/3} \zeta) \right| C(s_0) \exp(-3s/2), \quad \text{for all } s \in I_{1,N}. \quad (57)$$

Assembling (56) and (57), we obtain that when  $N \geq N_0(s_0, \gamma)$ ,

$$|D_{n,N}^2| \leq C(s_0) N^{-2/3} |s| \exp(-3s/2) \leq C(s_0) N^{-2/3} \exp(-s), \quad \text{for all } s \in I_{1,N}.$$

(iii) For  $D_{n,N}^3$ , recalling that  $r_N = 1 + O(N^{-1})$  and we obtain the following bound under the condition of Lemma 1 by using the previously derived bound on  $\text{Ai}'(\kappa_N^{2/3} \zeta)$ :

$$|D_{n,N}^3| \leq C(s_0) \exp(-s).$$

(iv) For  $D_{n,N}^4$ , by the definition of  $R_N$  and  $\dot{\zeta}_N$  as well as the bound for  $\partial_\xi \varepsilon_2(\kappa_N, \xi)$ , we have

$$\begin{aligned} |D_{n,N}^4| &= \left| \left( \frac{\sigma_{n,N}}{\kappa_N} \right) r_N R_N(\xi) \partial_\xi \varepsilon_2(\kappa_N, \xi) \right| \\ &\leq C N^{-2/3} \sigma_{n,N} \kappa_N^{-1/3} r_N R_N(\xi) \dot{\zeta}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta) \\ &= C N^{-2/3} r_N R_N^{-1}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta). \end{aligned}$$

All the terms involved in the last bound have been well studied during our analysis of  $D_{n,N}^2$ , and applying various results established there, we obtain that when  $N \geq N_0(s_0, \gamma)$ ,

$$|D_{n,N}^4| \leq C(s_0) N^{-2/3} \exp(-s), \quad \text{for all } s \in I_{1,N}.$$

Combining the bounds for the four terms, we obtain from a simple triangular inequality that when  $N \geq N_0(s_0, \gamma)$ ,

$$T_{N,1} \leq C(s_0) \exp(-s), \quad \text{for } s \in I_{1,N}.$$

We remark that, here and after, we derive a more stringent bound with the rate term  $N^{-2/3}$  whenever possible. Although it is not necessary here, those bounds with this rate term will become useful in the later study of  $|\psi'_\tau(s) - \text{Ai}'(s)|$ .

Case  $s \in I_{2,N}$ . In this case, we define  $\tilde{D}_{n,N}^1 = D_{n,N}^1$  and  $\tilde{D}_{n,N}^2 = D_{n,N}^2 + D_{n,N}^3 + D_{n,N}^4$ .

(i) To analyze the  $\tilde{D}_{n,N}^1$  term, we first introduce a useful lemma:

**Lemma 4.** *Let  $r > 0$  be fixed. For  $x = x_N(s) = \mu_{n,N} + s\sigma_{n,N}$  and  $\xi = x/\kappa_N$ , when  $s \geq r^2$ , we have*

$$\sigma_{n,N} \sqrt{f(\xi)} \geq r\xi_+/\xi = r\mu_{n,N}/(\mu_{n,N} + s\sigma_{n,N}).$$

For  $\tilde{D}_{n,N}^1$ , we could bound it for large  $N$  as

$$\left| \tilde{D}_{n,N}^1 \right| \leq C r_N \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \left| \frac{R'_N(\xi)}{R_N(\xi)} \right| R_N(\xi) \mathbf{M}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta).$$

We consider first the  $R_N(\xi) \mathbf{M}(\kappa_N^{2/3} \zeta)$  term. Recall that  $R_N(\xi) = \kappa_N^{1/6} \sigma_{n,N}^{-1/2} \hat{f}^{-1/4}(\xi)$  and that  $|\mathbf{M}(\kappa_N^{2/3} \zeta)| \leq C \kappa_N^{-1/6} \zeta^{-1/4}$  when  $N$  is large. Applying Lemma 4, we obtain that when  $N \geq N_0(\gamma)$ ,

$$\begin{aligned} R_N(\xi) \mathbf{M}(\kappa_N^{2/3} \zeta) &\leq C \zeta^{-1/4} \sigma_{n,N}^{-1/2} \hat{f}^{-1/4}(\xi) = C f^{-1/4}(\xi) \sigma_{n,N}^{-1/2} \\ &\leq C r^{-1/2} \left( \frac{\mu_{n,N}}{\mu_{n,N} + s\sigma_{n,N}} \right)^{-1/2} \leq C s, \quad \text{for all } s \in I_{2,N}. \end{aligned}$$

We remark that our choice of  $s_1$  ensures that  $s_1 \geq r^2$  with  $r = 1$ .

Switching to the term  $|R'_N(\xi)/R_N(\xi)|$ , from the definition, we have

$$\frac{R'_N(\xi)}{R_N(\xi)} = -\frac{\ddot{\zeta}(\xi)}{2\dot{\zeta}(\xi)}, \quad \text{and} \quad \frac{\ddot{\zeta}(\xi)}{\dot{\zeta}(\xi)} = \frac{f'(\xi)}{2f(\xi)} - \frac{\sqrt{f(\xi)}}{3I(\sqrt{f})}$$

where  $I(\sqrt{f}) = \int_{\xi_+}^{\xi} \sqrt{f}$ . Simple triangular inequality gives a direct bound as

$$\left| \frac{R'_N(\xi)}{R_N(\xi)} \right| \leq \frac{1}{4} \left| \frac{f'(\xi)}{f(\xi)} \right| + \frac{1}{6} \frac{\sqrt{f(\xi)}}{I(\sqrt{f})}.$$

For the first term on the right hand side, simple manipulation gives us

$$\left| \frac{f'(\xi)}{f(\xi)} \right| = \left| \frac{1}{\xi - \xi_+} + \frac{1}{\xi - \xi_-} - \frac{2}{\xi} \right| \leq \frac{4}{\xi - \xi_+} = \frac{4\kappa_N}{s\sigma_{n,N}} \leq C \frac{\kappa_N}{\sigma_{n,N}}.$$

Moreover, we could bound  $(\xi - \xi_+)[\sqrt{f}/I(\sqrt{f})]$  as

$$\begin{aligned} \frac{(\xi - \xi_+) \sqrt{f(\xi)}}{I(\sqrt{f})} &\leq \frac{(\xi - \xi_+)^{3/2} (\xi - \xi_-)^{1/2}}{2\xi \int_{\xi_+}^{\xi} \frac{\sqrt{t - \xi_+} (t - \xi_-)}{2t} dt} \leq \frac{(\xi - \xi_+)^{3/2}}{(1 - \xi_-/\xi_+) \int_{\xi_+}^{\xi} \sqrt{t - \xi_+} dt} \\ &= \frac{3}{2(1 - \xi_-/\xi_+)} \leq \frac{6}{\xi_+ - \xi_-} \leq \frac{3}{4} \left( 1 + \frac{2n}{N} \right) \end{aligned}$$

Hence, when  $N \geq N_0(\gamma)$ , we obtain the bound for  $\sqrt{f}/I(\sqrt{f})$  as

$$\frac{\sqrt{f(\xi)}}{I(\sqrt{f})} \leq \frac{3}{4} \left( \frac{1 + 2n/N}{\xi - \xi_+} \right) \leq \frac{3}{4} \left( 1 + \frac{2n}{N} \right) \frac{\kappa_N}{\sigma_{n,N}} s_1^{-1} N^{-1/6} \leq C \frac{\kappa_N}{\sigma_{n,N}}.$$

This implies that  $|R'_N(\xi)/R_N(\xi)|$  is bounded by  $C\kappa_N/\sigma_{n,N}$  which further ensures

$$r_N \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \left| \frac{R'_N(\xi)}{R_N(\xi)} \right| \leq C.$$

Finally, using (75) and the fact that  $s_1$  is a fixed constant, we obtain that when  $N \geq N_0(\gamma)$ ,

$$\left| \tilde{D}_{n,N}^1 \right| \leq Cs \exp(-3s/2) \leq Cs^{-4} \exp(-s) \leq CN^{-2/3} \exp(-s), \quad \text{for all } s \in I_{2,N}.$$

(ii) For  $\tilde{D}_{n,N}^2$ , we first recall its definition as

$$\tilde{D}_{n,N}^2 = r_N R_N(\xi) \left[ \text{Ai}'(\kappa_N^{2/3} \zeta) R_N^{-2}(\xi) + \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \frac{\partial}{\partial \xi} \varepsilon_2(\kappa_N, \xi) \right].$$

By definition of  $R_N$  and the large  $N$  bounds on  $r_N$ ,  $\partial_\xi \varepsilon_2(\kappa_N, \xi)$  and  $\text{Ai}'$ , we have

$$\left| \tilde{D}_{n,N}^2 \right| \leq C R_N^{-1}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \mathbf{E}^{-1}(\kappa_N^{2/3} \zeta).$$

The asymptotics of the phase function  $\mathbf{N}$  suggest that

$$R_N^{-1}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \leq C R_N^{-1}(\xi) \kappa_N^{1/6} \zeta^{1/4} = C f^{1/4}(\xi) \sigma_{n,N}^{1/2}.$$

For  $\sigma_{n,N} \sqrt{f(\xi)}$ , we could simply bound it as

$$\sigma_{n,N} \sqrt{f(\xi)} = \frac{\sigma_{n,N} \sqrt{(\xi - \xi_+)(\xi - \xi_-)}}{2\xi} \leq \frac{\sigma_{n,N}}{2}.$$

Observing that for  $s \in I_{2,N}$ ,  $\sigma_{n,N} \leq C(\gamma)N^{1/3} \leq Cs^4$ , we obtain

$$R_N^{-1}(\xi) \mathbf{N}(\kappa_N^{2/3} \zeta) \leq C \sigma_{n,N}^{1/2} f^{1/4}(\xi) \leq C \sigma_{n,N}^{1/2} \leq Cs^2.$$

Once more, by (75) and our choice of  $s_1$  [see A.4], we obtain

$$\left| \tilde{D}_{n,N}^2 \right| \leq Cs^2 \exp(-3s/2) \leq Cs^{-4} \exp(-s) \leq CN^{-2/3} \exp(-s).$$

This finally gives a bound of the form  $C N^{-2/3} \exp(-s)$  for  $T_{N,1}$  on  $I_{2,N}$ .

By a simple triangular inequality, we combine our bounds on  $T_{N,1}$  and  $T_{N,2}$  on both  $I_{1,N}$  and  $I_{2,N}$  together and obtain that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ ,

$$|\partial_s \theta_{n,N}(x_{n,N}(s))| \leq C(s_0) \exp(-s), \quad \text{for all } s \geq s_0.$$

**Bound for  $|\psi'_\tau(s)|$ .** We have pointed out that  $\psi_\tau$  is of the form  $\frac{1}{\sqrt{2}} \rho_N \theta_{n-1, N-1}(x_{n-1, N-1}(s))$  with  $\rho_N = 1 + O(N^{-1})$ . Hence, we have  $\psi'_\tau(s)$  as

$$\psi'_\tau(s) = \frac{1}{\sqrt{2}} \rho_N \partial_s \theta_{n-1, N-1}(x_{n-1, N-1}(s)),$$

for which our bound on  $\sigma_{n,N} \partial_s \theta_{n,N}(s)$  apply directly and we obtain that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ ,

$$|\psi'_\tau(s)| \leq C(s_0) \exp(-s), \quad \text{for all } s \geq s_0.$$

**Bound for  $|\psi'_\tau(s) - G'(s)|$ .** By the expression of  $\psi'_\tau$ , we could split  $|\psi'_\tau(s) - G'(s)|$  as

$$\begin{aligned} |\psi'_\tau(s) - G'(s)| &\leq \frac{1}{\sqrt{2}} |\rho_N - 1| |\partial_s \theta_{n-1, N-1}(x_{n-1, N-1}(s))| \\ &\quad + \frac{1}{\sqrt{2}} |\partial_s \theta_{n-1, N-1}(x_{n-1, N-1}(s)) - \text{Ai}'(s)|. \end{aligned} \quad (58)$$

By our bound on  $|\partial_s \theta_{n-1, N-1}(x_{n-1, N-1}(s))|$  and recalling that  $\rho_N = 1 + O(N^{-1})$ , the first term is then bounded by  $C(s_0)N^{-1} \exp(-s)$ . We focus on the quantity  $|\partial_s \theta_{n, N}(x_{n, N}(s)) - \text{Ai}'(s)|$  to bound the second term.

We split the quantity of interest into two parts as the following:

$$\begin{aligned} |\partial_s \theta_{n, N}(x_{n, N}(s)) - \text{Ai}'(s)| &\leq \left| \sigma_{n, N} F'_{n, N}(x_{n, N}(s)) \frac{\mu_{n, N}}{x_{n, N}(s)} - \text{Ai}'(s) \right| + \left| \sigma_{n, N} F_{n, N}(x_{n, N}(s)) \frac{\mu_{n, N}}{x_{n, N}^2(s)} \right| \\ &= \mathcal{T}_{N,1}(s) + \mathcal{T}_{N,2}(s). \end{aligned}$$

The  $\mathcal{T}_{N,2}(s)$  term is exactly the same as  $T_{N,2}(s)$  defined in the previous study of  $\partial_s \theta_{n, N}(x_{n, N}(s))$  and hence we quote the bound derived there directly as

$$\mathcal{T}_{N,2}(s) \leq C(s_0)N^{-2/3} \exp(-s), \quad \text{for all } s \geq s_0.$$

Switching to the  $\mathcal{T}_{N,1}(s)$  term, we divide the whole region into the two disjoint intervals  $I_{1, N} = [s_0, s_1 N^{1/6})$  and  $I_{2, N} = [s_1 N^{1/6}, \infty)$  again.

*Case  $s \in I_{1, N}$ .* Exploiting a similar strategy in splitting  $\sigma_{n, N} F'_{n, N}(x)$ , on  $I_{1, N}$ , we decompose  $\mathcal{T}_{N,1}(s)$  as  $\mathcal{T}_{N,1}(s) = \sum_{i=1}^5 \mathcal{D}_{n, N}^i$ , with  $\mathcal{D}_{n, N}^i = D_{n, N}^i \mu_{n, N} / x_{n, N}(s)$  for  $i = 1, 2$  and  $4$ ,

$$\mathcal{D}_{n, N}^3 = r_N \frac{\mu_{n, N}}{x_{n, N}(s)} \left[ \text{Ai}'(\kappa_N^{2/3} \zeta) - \text{Ai}'(s) \right], \quad \text{and} \quad \mathcal{D}_{n, N}^5 = \left[ r_N \frac{\mu_{n, N}}{x_{n, N}(s)} - 1 \right] \text{Ai}'(s).$$

For  $i = 1, 2$  and  $4$ , using our previous bounds on  $D_{n, N}^i$  and noting that  $|\mu_{n, N} / x_{n, N}(s)|$  could be bounded by  $2$  on  $I_{1, N}$ , we obtain directly that, when  $N \geq N_0(s_0, \gamma)$ ,

$$|\mathcal{D}_{n, N}^i| \leq C(s_0)N^{-2/3} \exp(-s), \quad \text{for } i = 1, 2 \text{ and } 4, \text{ and all } s \in I_{1, N}.$$

For  $\mathcal{D}_{n, N}^3$ , by a first order Taylor expansion and the identity  $\text{Ai}''(s) = s \text{Ai}(s)$  for all  $s$ , we have that, for some  $s^*$  in the middle of  $\kappa_N^{2/3} \zeta$  and  $s$ ,

$$|\mathcal{D}_{n, N}^3| = r_N \left| \frac{\mu_{n, N}}{x_{n, N}(s)} \right| |s^* \text{Ai}(s^*)| \left| \kappa_N^{2/3} \zeta - s \right| \leq C N^{-2/3} s^2 |s^* \text{Ai}(s^*)|,$$

where the inequality holds when  $N \geq N_0(s_0, \gamma)$  and comes from (74) and the large  $N$  bounds for  $r_N$  and  $\mu_{n, N} / x_{n, N}(s)$ .

When  $s \geq 0$ , we know from the definition of  $\zeta$  that  $\kappa_N^{2/3} \zeta \geq 0$  and hence  $s^* \geq 0$ . Moreover, (74) implies that when  $N$  is large,  $\kappa_N^{2/3} \zeta$  and hence  $s^*$  will be greater than  $s/2$ . Thus, by (51) and the monotonicity of  $\mathbf{E}$ , we obtain

$$|s^* \text{Ai}(s^*)| \leq C s \mathbf{E}^{-1}(s/2) \leq C s \exp\left(-\frac{1}{3\sqrt{2}} s^{3/2}\right) \leq C \exp(-3s/2).$$

If  $s_0 \leq 0$ , as before, we consider all  $s \in [s_0, 0]$ . Once again, we obtain from (74) that for large  $N$ ,  $\kappa_N^{2/3} \zeta \in [3s_0/2, 1]$  and hence  $s^* \in [3s_0/2, 1]$ . Then for all  $s \in [s_0, 0]$ , when  $N \geq N_0(s_0, \gamma)$ ,

$$|s^* \text{Ai}(s^*)| \exp(3s/2) \leq C(s_0) \equiv \max_{s \in [3s_0/2, 1]} e^{3/2} |s \text{Ai}(s)|.$$

This  $C(s_0)$  is continuous and non-increasing in  $s_0$ .

Thus, we could conclude that when  $N \geq N_0(s_0, \gamma)$ , for all  $s \in I_{1,N}$ ,

$$|\mathcal{D}_{n,N}^3| \leq C N^{-2/3} s^2 |s^* \text{Ai}(s^*)| \leq C(s_0) N^{-2/3} s^2 \exp(-3s/2) \leq C(s_0) N^{-2/3} \exp(-s).$$

In  $\mathcal{D}_{n,N}^5$ , recalling  $\sigma_{n,N}/\mu_{n,N} = \mathcal{O}(N^{-2/3})$  and  $r_N = 1 + \mathcal{O}(N^{-1})$ , we have that, when  $N \geq N_0(s_0, \gamma)$ , for all  $s \in I_{1,N}$ ,  $|s_0 + \mu_{n,N}/\sigma_{n,N}| \geq \frac{1}{2}(\mu_{n,N}/\sigma_{n,N})$  and hence

$$\begin{aligned} \left| r_N \frac{\mu_{n,N}}{x_{n,N}(s)} - 1 \right| &\leq r_N \left| \frac{\mu_{n,N}}{x_{n,N}(s)} - 1 \right| + |r_N - 1| = r_N |s| \left| s + \frac{\mu_{n,N}}{\sigma_{n,N}} \right|^{-1} + |r_N - 1| \\ &\leq r_N |s| \left| s_0 + \frac{\mu_{n,N}}{\sigma_{n,N}} \right|^{-1} + |r_N - 1| \leq C N^{-2/3} |s| + C N^{-1}. \end{aligned}$$

For  $\text{Ai}'(s)$ , by (51) and (53), we obtain directly that

$$|\text{Ai}'(s)| \leq C(s_0) |s|^{1/4} \exp\left(-\frac{2}{3}s^{3/2}\right),$$

where  $C(s_0)$  could be chosen as

$$\max_{s \in [s_0, \infty)} |\text{Ai}'(s)| \left(1 + |s|^{1/4}\right)^{-1} \exp\left(\frac{2}{3}s^{3/2}\right),$$

which is continuous and non-increasing.

Putting two parts together, we obtain that for all  $s \in I_{1,N}$ ,

$$|\mathcal{D}_{n,N}^5| \leq C(s_0) N^{-2/3} \left(|s| + C N^{-1/3}\right) |s|^{1/4} \exp\left(-\frac{2}{3}s^{3/2}\right) \leq C(s_0) N^{-2/3} \exp(-s).$$

We could then assemble all the bounds on  $\mathcal{D}_{n,N}^i$  using the triangular inequality and conclude that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ ,

$$\mathcal{T}_{N,1}(s) \leq C(s_0) N^{-2/3} \exp(-s), \quad \text{for all } s \in I_{1,N}.$$

*Case  $s \in I_{2,N}$ .* In this case, we could act more heavy-handedly. In particular, by the asymptotics of  $\mathcal{T}_{N,1}(s)$  on  $I_{2,N}$  and the asymptotics of  $\text{Ai}'$ , we have

$$\begin{aligned} \mathcal{T}_{N,1}(s) &\leq \left| \sigma_{n,N} F'_{n,N}(x_{n,N}(s)) \frac{\mu_{n,N}}{x_{n,N}(s)} \right| + |\text{Ai}'(s)| \leq C N^{-2/3} \exp(-s) + C s^{1/4} \exp(-3s/2) \\ &\leq C N^{-2/3} \exp(-s) + C N^{-2/3} s^{4+1/4} \exp(-3s/2) \leq C N^{-2/3} \exp(-s). \end{aligned}$$

We then obtain the bound  $C(s_0) N^{-2/3} \exp(-s)$  for  $\mathcal{T}_{N,1}(s)$  and hence also for  $|\partial_s \theta_{n,N}(x_{n,N}(s)) - \text{Ai}'(s)|$  for all  $s \in [s_0, \infty)$ . Applying the bound to the second term in (58), we obtain that under the condition of Lemma 1, when  $N \geq N_0(s_0, \gamma)$ ,

$$|\psi'_\tau(s) - G'(s)| \leq C(s_0) N^{-2/3} \exp(-s), \quad \text{for all } s \in [s_0, \infty).$$

**Improved bound for  $|\psi_\tau - G|$ .** The above bound on  $|\psi'_\tau(s) - G'(s)|$  could be used to derive a more stringent bound for  $|\psi_\tau(s) - G(s)|$  as the following:

$$\begin{aligned}
|\psi_\tau(s) - G(s)| &= \left| \int_s^{2s} [\psi'_\tau(t) - G'(t)] dt - [\psi_\tau(2s) - G(2s)] \right| \\
&\leq \int_s^{2s} |\psi'_\tau(t) - G'(t)| dt + |\psi_\tau(2s) - G(2s)| \\
&\leq \int_s^{2s} C(s_0) N^{-2/3} e^{-t} dt + C(s_0) N^{-2/3} \exp(-s) \leq C(s_0) N^{-2/3} \exp(-s).
\end{aligned} \tag{59}$$

This is exactly the bound that we have claimed in Lemma 1.

### 5.2.3 Bounds for quantities related to $\phi_\tau(s)$

In this part, we employ a trick that was first used in Johnstone (2001, p.320) to derive bounds for quantities related to  $\phi_\tau$  from those for quantities related to  $\psi_\tau$ .

Recall that  $\phi_\tau$  could be expressed as

$$\phi_\tau(s) = \frac{1}{\sqrt{2}} \tilde{\rho}_N F_{n-2,N}(x_{n-1,N-1}(s)) \frac{\mu_{n-2,N}}{x_{n-1,N-1}(s)},$$

where  $\tilde{\rho}_N = 1 + O(N^{-1})$  [see A.1 for its proof]. The problem of  $\phi_\tau$  is that the centering and scaling constants  $(\mu_{n-1,N-1}, \sigma_{n-1,N-1})$  in the transformation  $x_{n-1,N-1}(s)$  does not agree with the “optimal” constants  $(\mu_{n-2,N}, \sigma_{n-2,N})$  for the related function  $F_{n-2,N}$ . To circumvent this problem, we introduce a new independent variable  $s'$  as the following [one should not confuse it with the  $s'$  previously appeared in Section 4]:

$$\mu_{n-1,N-1} + s\sigma_{n-1,N-1} = \mu_{n-2,N} + s'\sigma_{n-2,N}. \tag{60}$$

Then  $s' = (\mu_{n-1,N-1} - \mu_{n-2,N})/\sigma_{n-2,N} + s\sigma_{n-1,N-1}/\sigma_{n-2,N}$ . By defining

$$\Delta_N = \frac{\mu_{n-1,N-1} - \mu_{n-2,N}}{\sigma_{n-2,N}}, \tag{61}$$

we have  $s' - s = \Delta_N + [\sigma_{n-1,N-1}\sigma_{n-2,N}^{-1}]s$  and  $\phi_\tau(s)$  could be rewritten as

$$\phi_\tau(s) = \frac{1}{\sqrt{2}} \tilde{\rho}_N F_{n-2,N}(x_{n-2,N}(s')) \frac{\mu_{n-2,N}}{x_{n-2,N}(s')}.$$

Before we proceed, we list two important properties as the following [with proof given in A.1]:

$$\Delta_N = O(N^{-1/3}) \quad \text{and} \quad 1 \leq \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} = 1 + O(N^{-1}). \tag{62}$$

**Bounds for  $|\phi_\tau(s)|$  and  $|\phi'_\tau(s)|$ .** Applying our previous bounds for  $|\theta_{n,N}(x_{n,N}(s))|$  and  $|\partial_s \theta_{n,N}(x_{n,N}(s))|$ , and using (62), we obtain that under the condition of Lemma 1, for all  $s \in [s_0, \infty)$

$$\begin{aligned}
|\phi_\tau(s)| &= \frac{1}{\sqrt{2}} \tilde{\rho}_N |\theta_{n-2,N}(x_{n-2,N}(s'))| \leq C(s_0) \exp(-s') \leq C(s_0) \exp(-s); \\
|\phi'_\tau(s)| &= \frac{1}{\sqrt{2}} \tilde{\rho}_N |\partial_s \theta_{n-2,N}(x_{n-2,N}(s'))| = \frac{1}{\sqrt{2}} \tilde{\rho}_N |\partial_{s'} \theta_{n-2,N}(x_{n-2,N}(s'))| \frac{ds'}{ds} \\
&\leq C(s_0) \exp(-s') \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} \leq C(s_0) \exp(-s).
\end{aligned}$$



**Bounds for  $|\phi_\tau(s) - G(s) - \Delta_N G'(s)|$  and  $|\phi'_\tau(s) - G'(s) - \Delta_N G''(s)|$ .** We consider  $|\phi_\tau(s) - G(s) - \Delta_N G'(s)|$  in detail and the derivation for  $|\phi'_\tau(s) - G'(s) - \Delta_N G''(s)|$  is essentially the same.

By our definition of  $s'$  and recalling that  $\text{Ai}''(s) = s\text{Ai}(s)$ , we have the Taylor expansion of  $G(s')$  as

$$\begin{aligned} G(s') &= G(s) + (s' - s)G'(s) + \frac{1}{2}(s' - s)^2 G''(s^*) \\ &= G(s) + \Delta_N G'(s) + \frac{1}{\sqrt{2}} \left[ \frac{\sigma_{n-1, N-1}}{\sigma_{n-2, N}} - 1 \right] s \text{Ai}'(s) + \frac{1}{2\sqrt{2}} (s' - s)^2 s^* \text{Ai}(s^*), \end{aligned}$$

with  $s^*$  lies at somewhere between  $s$  and  $s'$ .

Hence, by our bound on  $|\psi_\tau(s) - G(s)|$ , we obtain

$$|\phi_\tau(s) - G(s) - \Delta_N G'(s)| \leq C(s_0) N^{-2/3} \exp(-s') + C N^{-1} |s \text{Ai}'(s)| + C (s' - s)^2 |s^* \text{Ai}(s^*)|. \quad (63)$$

On  $[s_0, \infty)$ , we have  $C(s_0) N^{-2/3} \exp(-s') \leq C(s_0) N^{-2/3} \exp(-s)$  for the first term in the above bound. Moreover, by (51) and (53), the second term satisfies

$$C N^{-1} |s \text{Ai}'(s)| \leq C(s_0) N^{-1} |s|^{1+1/4} \exp\left(-\frac{2}{3}s^{3/2}\right) \leq C(s_0) N^{-1} \exp(-s), \quad \text{for all } s \geq s_0.$$

For the last term, we split  $[s_0, \infty)$  into  $I_{1, N} \cup I_{2, N}$  as usual. For  $s \in I_{1, N}$ , when  $N \geq N_0(s_0, \gamma)$ ,

$$(s - s')^2 = \left[ \Delta_N + \left( \frac{\sigma_{n-1, N-1}}{\sigma_{n-2, N}} - 1 \right) s \right]^2 \leq \left[ C N^{-1/3} + C N^{-1} s \right]^2 \leq \left( C N^{-2/3} \right) \wedge 1.$$

We obtain from the above bound that  $|s^* - s| \leq 1$  and hence by (51) and (53),

$$C (s - s')^2 |s^* \text{Ai}(s^*)| \leq C N^{-2/3} (|s| + 1) \mathbf{E}^{-1}(s - 1) \leq C(s_0) N^{-2/3} \exp(-s),$$

where  $C(s_0)$  could be chosen as  $\max_{s \in [s_0, \infty)} C e^s \mathbf{E}^{-1}(s - 1)$ .

On  $I_{2, N}$ , we have  $s' \geq s/2$  from (77) and hence  $s^* \geq s/2$ . By (51) and (53), we obtain that when  $N \geq N_0(s_0, \gamma)$ ,

$$C (s' - s)^2 |s^* \text{Ai}(s^*)| \leq C s^3 \exp\left(-\frac{1}{3\sqrt{2}} s^{3/2}\right) \leq C s^{-4} \exp(-s) \leq C N^{-2/3} \exp(-s).$$

Therefore, we have shown that, for all  $N \geq N_0(s_0, \gamma)$  and  $s \geq s_0$ , the right hand side of (63) is further controlled by  $C(s_0) N^{-2/3} \exp(-s)$ , which is exactly the desired bound for  $|\phi_\tau(s) - G(s) - \Delta_N G'(s)|$ .

## A Technical Details

### A.1 Properties of $\beta_N, \rho_N, \tilde{\rho}_N, \Delta_N$ and $\sigma_{n-1, N-1}/\sigma_{n-2, N}$

#### A.1.1 Property of $\beta_N$

We are to show that

$$\beta_N = \frac{1}{\sqrt{2}} + \mathcal{O}(N^{-1}).$$

First of all, we recall that  $\phi_\tau(s)$  is defined to be 0 when  $\tau(s) = \tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N} < 0$ , i.e., when  $s \in (-\infty, -\tilde{\mu}_{n,N}/\tilde{\sigma}_{n,N})$ . Hence, we have

$$\begin{aligned}\beta_N &= \frac{1}{2} \int_{-\infty}^{\infty} \phi_\tau(s) ds = \frac{1}{2} \int_0^{\infty} \phi(x; \alpha_N - 1) dx \\ &= \frac{N^{1/4}(n-1)^{1/4} \Gamma^{1/2}(N+1)}{2\sqrt{2} \Gamma^{1/2}(n)} \int_0^{\infty} x^{\alpha_N/2-1} e^{-x/2} L_N^{\alpha_N-1}(x) dx \\ &= \frac{2^{-(\alpha_N-1)/2} N^{1/4} (n-1)^{1/4} \Gamma^{1/2}(n) \Gamma\left(\frac{N+3}{2}\right)}{(N+1) \Gamma^{1/2}(N+1) \Gamma\left(\frac{n+1}{2}\right)}.\end{aligned}$$

Applying Sterling's formula

$$\Gamma(z) = \left(\frac{2\pi}{z}\right)^{1/2} \left(\frac{z}{e}\right)^z \left[1 + \mathcal{O}\left(\frac{1}{z}\right)\right],$$

we obtain that

$$\begin{aligned}\beta_N &= \frac{2^{-(\alpha_N-1)/2} N^{1/4} (n-1)^{1/4}}{N+1} \frac{\left(\frac{2\pi}{n}\right)^{1/4} \left(\frac{n}{e}\right)^{n/2} \left(\frac{4\pi}{N+3}\right)^{1/2} \left(\frac{N+3}{2e}\right)^{(N+3)/2}}{\left(\frac{2\pi}{N+1}\right)^{1/4} \left(\frac{N+1}{e}\right)^{(N+1)/2} \left(\frac{4\pi}{n+1}\right)^{1/2} \left(\frac{n+1}{2e}\right)^{(n+1)/2}} (1 + \mathcal{O}(N^{-1})) \\ &= \frac{1}{\sqrt{2e}} \left(1 - \frac{1}{n+1}\right)^{n/2} \left(1 + \frac{2}{N+1}\right)^{(N+1)/2+3/4} (1 + \mathcal{O}(N^{-1})) = \frac{1}{\sqrt{2}} (1 + \mathcal{O}(N^{-1})).\end{aligned}$$

The last equality is exactly the asymptotics that we need for  $\beta_N$ .

### A.1.2 Asymptotics of $\rho_N$ and $\tilde{\rho}_N$

In this part, we show that the asymptotics of  $\rho_N$  and  $\tilde{\rho}_N$  satisfy

$$\rho_N, \tilde{\rho}_N = 1 + \mathcal{O}(N^{-1}).$$

We consider  $\rho_N$  first. By definition, we have

$$\rho_N = \frac{N^{1/4}(n-1)^{1/4} \sigma_{n-1, N-1}^{1/2} \tilde{\sigma}_{n,N}}{\tilde{\mu}_{n,N}} = \frac{N^{1/4}(n-1)^{1/4} \sigma_{n-1, N-1}^{3/2}}{\mu_{n-1, N-1}}.$$

Plugging in the definition of  $\sigma_{n-1, N-1}$  and  $\mu_{n-1, N-1}$ , we obtain that

$$\begin{aligned}\rho_N &= N^{1/4}(n-1)^{1/4} \left(\sqrt{N-\frac{1}{2}} + \sqrt{n-\frac{1}{2}}\right)^{-1/2} \left(\frac{1}{\sqrt{N-\frac{1}{2}}} + \frac{1}{\sqrt{n-\frac{1}{2}}}\right)^{1/2} \\ &= \left(\frac{N}{N-\frac{1}{2}}\right)^{1/4} \left(\frac{n-1}{n-\frac{1}{2}}\right)^{1/4} = 1 + \mathcal{O}(N^{-1}).\end{aligned}$$

For  $\tilde{\rho}_N$ , we have from its definition that

$$\begin{aligned}\tilde{\rho}_N &= \frac{N^{1/4}(n-1)^{1/4}\sigma_{n-2,N}^{1/2}\tilde{\sigma}_{n,N}}{\mu_{n-2,N}} = \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} \frac{N^{1/4}(n-1)^{1/4}\sigma_{n-2,N}^{3/2}}{\mu_{n-2,N}} \\ &= \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} N^{1/4}(n-1)^{1/4} \left( \sqrt{N+\frac{1}{2}} + \sqrt{n-\frac{3}{2}} \right)^{-1/2} \left( \frac{1}{\sqrt{N+\frac{1}{2}}} + \frac{1}{\sqrt{n-\frac{3}{2}}} \right)^{1/2} \\ &= \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} \left( \frac{N}{N+\frac{1}{2}} \right)^{1/4} \left( \frac{n-1}{n-\frac{3}{2}} \right)^{1/4} = 1 + O(N^{-1}).\end{aligned}$$

The last equality holds since  $\sigma_{n-1,N-1}/\sigma_{n-2,N} = 1 + O(N^{-1})$  as claimed in (61), which is to be shown below in A.1.3.

### A.1.3 Properties of $\Delta_N$ and $\sigma_{n-1,N-1}/\sigma_{n-2,N}$

We focus on  $\Delta_N$  first. As a reminder, we recall its definition as

$$\Delta_N = \frac{\mu_{n-1,N-1} - \mu_{n-2,N}}{\sigma_{n-2,N}}.$$

By El Karoui (2006b, A.1.2), we have for the numerator that  $\mu_{n-1,N-1} - \mu_{n-2,N} = O(1)$ . For the denominator, if we let denote  $(n - \frac{3}{2}) / (N + \frac{1}{2})$  by  $\gamma_{n,N}$ , we then have

$$\begin{aligned}\frac{1}{\sigma_{n-2,N}} &= \left( \sqrt{N+\frac{1}{2}} + \sqrt{n-\frac{3}{2}} \right)^{-1} \left( \frac{1}{\sqrt{N+\frac{1}{2}}} + \frac{1}{\sqrt{n-\frac{3}{2}}} \right)^{-1/3} \\ &= \frac{1}{1 + \sqrt{\gamma_{n,N}}} \left( 1 + \frac{1}{\sqrt{\gamma_{n,N}}} \right) \left( N + \frac{1}{2} \right)^{-1/3} = O(N^{-1/3}).\end{aligned}$$

The last equality holds since  $\gamma_{n,N}$  is bounded below for all  $n > N$ . Combining the two estimates, we establish that

$$\Delta_N = O(N^{-1/3}).$$

We now switch to prove that

$$1 \leq \sigma_{n-1,N-1}/\sigma_{n-2,N} = 1 + O(N^{-1}).$$

The fact that  $\sigma_{n-1,N-1}/\sigma_{n-2,N} = 1 + O(N^{-1})$  has been proved in El Karoui (2006b, A.1.3). On the other hand, we have from the second last display of El Karoui (2006b, A.1.3) that

$$\left( \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} \right)^3 = \left[ 1 + \frac{\sqrt{n/N} - \sqrt{N/n}}{n+N} + O(n^{-2}) \right] \left[ 1 + \frac{1}{2} \left( \frac{1}{n} + \frac{1}{N} \right) + O(n^{-2}) \right].$$

Both terms become greater than 1 when  $N \geq N_0(\gamma)$  and hence  $\sigma_{n-1,N-1}\sigma_{n-2,N}^{-1} \geq 1$  for large  $N$ . Actually, the inequality holds for any  $n > N \geq 2$ . However, what we have proved here is sufficient for our argument in Section 5.2.3.

## A.2 Evaluation of the entries of $K_\tau$

In this part, we work out the explicit expressions for the entries of  $K_\tau$  given in (20). To this end, we proceed term by term.

**$K_{\tau,11}$  term.** For  $K_{\tau,11}$ , we have from its definition that

$$\begin{aligned} K_{\tau,11}(s, t) &= \sigma_{n,N} S_{N,1}(\tilde{\mu}_{n,N} + s\tilde{\sigma}_{n,N}, \tilde{\mu}_{n,N} + t\tilde{\sigma}_{n,N}) \\ &= \sigma_{n,N} [S_{N,2}(\tau(s), \tau(t)) + \psi(\tau(s))(\varepsilon\phi)(\tau(t))] \\ &= S_\tau(s, t) + \sigma_{n,N} \psi(\tau(s))(\varepsilon\phi)(\tau(t)). \end{aligned}$$

For the second term in the last expression, we have  $\sigma_{n,N} \phi(\tau(s)) = \psi_\tau(s)$  and

$$\int_y^\infty \phi(z) dz = \int_t^\infty \phi(\tau(u)) \tau'(u) du = \tilde{\sigma}_{n,N} \int_t^\infty \phi(\tau(u)) du = \int_t^\infty \phi_\tau(u) du.$$

Hence, the second term equals  $\psi_\tau(s)(\varepsilon\phi_\tau)(t)$  and we obtain

$$K_{\tau,11}(s, t) = S_\tau(s, t) + \psi_\tau(s)(\varepsilon\phi_\tau)(t).$$

**$K_{\tau,12}$  term.** We first recall the definition of  $K_{\tau,12}$  as

$$K_{\tau,12}(s, t) = -\tilde{\sigma}_{n,N} \sqrt{\tau'(s)\tau'(t)} \partial_2 S_{N,1}(\tau(s), \tau(t)).$$

For the involved partial derivative, we have

$$\partial_2 S_{N,1}(\tau(s), \tau(t)) = \frac{1}{\tau'(t)} \frac{\partial}{\partial t} \frac{K_{\tau,11}(s, t)}{\sqrt{\tau'(s)\tau'(t)}} = \frac{1}{\tilde{\sigma}_{n,N}} \frac{\partial}{\partial t} \frac{S_\tau(s, t) + \psi_\tau(s)(\varepsilon\phi_\tau)(t)}{\tilde{\sigma}_{n,N}} = \frac{1}{\tilde{\sigma}_{n,N}^2} \partial_t S_\tau^R(s, t),$$

with  $S_\tau^R$  defined as in (19). Observing that  $\tau'(s) = \tau'(t) = \tilde{\sigma}_{n,N}$ , we obtain

$$K_{\tau,12}(s, t) = -\partial_t S_\tau^R(s, t).$$

**$K_{\tau,21}$  term.** By its definition, we have

$$K_{\tau,21}(s, t) = \frac{\sqrt{\tau'(s)\tau'(t)}}{\tilde{\sigma}_{n,N}} [\varepsilon S_{N,1}(\tau(s), \tau(t)) - \varepsilon(\tau(s) - \tau(t))].$$

Observing that  $\tau$  is a monotone transformation, we obtain

$$\varepsilon(\tau(s) - \tau(t)) = \varepsilon(s - t).$$

For the quantity  $\varepsilon S_{N,1}(\tau(s), \tau(t))$ , by using the above identity, we have

$$\varepsilon S_{N,1}(\tau(s), \tau(t)) = \int S_{N,1}(\tau(u), \tau(t)) \varepsilon(\tau(s) - \tau(u)) \tau'(u) du = \int S_\tau^R(u, t) \varepsilon(s - u) du = (\varepsilon S_\tau^R)(s, t).$$

Plugging all these identities back into the definition of  $K_{\tau,21}$ , we obtain the expression

$$K_{\tau,21}(s, t) = (\varepsilon S_\tau^R)(s, t) - \varepsilon(s - t).$$

**$K_{\tau,22}$  term.** The formula for  $K_{\tau,22}$  is obtained directly from that of  $K_{\tau,11}$  by switch  $s$  and  $t$ .

### A.3 Behavior of $R_N(\xi)$ , $R'_N(\xi)$ and $\kappa_N^{2/3}\zeta$

In this part, we investigate the behavior of  $R_N(\xi)$ ,  $R'_N(\xi)$  and  $\kappa_N^{2/3}\zeta$  which is essential in deriving the Laguerre asymptotics. Before we start, we remark that throughout our discussion, we consider only the case where  $s \in I_{1,N} = [s_0, s_1 N^{1/6}]$ .

#### A.3.1 Properties of $R_N(\xi)$ and $R'_N(\xi)$

Recall the definition  $R_N(\xi) = (\dot{\zeta}(\xi)/\dot{\zeta}(\xi_+))^{1/2}$ , we obtain that

$$R_N(\xi_+) = 1, \quad \text{and} \quad R'_N(\xi) = -\frac{1}{2}\dot{\zeta}_N^{1/2}\dot{\zeta}(\xi)^{-3/2}\ddot{\zeta}(\xi). \quad (64)$$

By our derivation in the Liouville-Green approximation, we know that  $\xi = \xi_+ + s\sigma_{n,N}/\kappa_N$  and as has been shown before, when  $N \geq N_0(\gamma)$ ,

$$N^{2/3} \left( \frac{\sigma_{n,N}}{\kappa_N} \right) \leq 4(1 + 1/\sqrt{\gamma})^{1/3} (1 + \sqrt{\gamma}) (1 + \gamma)^{-1} \leq C.$$

As  $N \rightarrow \infty$ , we have

$$\sup_{s \in I_{1,N}} |\xi - \xi_+| = O\left(s_1 N^{-1/2}\right) \rightarrow 0. \quad (65)$$

We then have the following first order Taylor expansion

$$R_N(\xi) = R_N(\xi_+) + R'_N(\xi^*)(\xi - \xi_+), \quad \text{for some } \xi^* \in [\xi \wedge \xi_+, \xi \vee \xi_+].$$

Hence, when  $N \geq N_0(s_0, \gamma)$ , we have

$$|R_N(\xi) - 1| \leq |R'_N(\xi^*)| 4(1 + 1/\sqrt{\gamma})^{1/3} (1 + \sqrt{\gamma}) (1 + \gamma)^{-1} N^{-2/3} |s|. \quad (66)$$

In order to bound  $|R_N(\xi) - 1|$  uniformly on  $I_{1,N}$  and also of its own interest, we are to derive a bound for  $|R'_N(\xi)|$  by some constant that does not depend on  $N$  and is uniform for  $s \in I_{1,N}$ . By the definition of  $R'_N(\xi)$  in (64), this relies on the understanding of the quantities  $\dot{\zeta}_N$ ,  $\dot{\zeta}(\xi)$  and  $\ddot{\zeta}(\xi)$ .

First, we consider the asymptotics of  $\dot{\zeta}_N$ . Using the notation  $m_{\pm} = m \pm 1/2$ , we obtain from simple calculation that as  $N \rightarrow \infty$ ,

$$\dot{\zeta}_N = \frac{n_+^{1/6} N_+^{1/6} (n_+ + N_+)^{1/3}}{2^{1/3} (\sqrt{n_+} + \sqrt{N_+})^{4/3}} \longrightarrow \frac{\gamma^{1/6} (1 + \gamma)^{1/3}}{2^{1/3} (1 + \sqrt{\gamma})^{4/3}}. \quad (67)$$

Second, we check the behavior of  $\dot{\zeta}(\xi)$ . For simplicity, we let  $\xi_{\pm}^{\infty} = \lim_{N \rightarrow \infty} \xi_{\pm}$  and simple manipulation gives us

$$\xi_+^{\infty} = 2(1 + \sqrt{\gamma})^2 / (1 + \gamma), \quad \text{and} \quad \xi_+^{\infty} - \xi_-^{\infty} = 8\sqrt{\gamma} / (1 + \gamma).$$

We assume first that  $s_0 \geq 0$ . By the definition of  $\dot{\zeta}(\xi)$  for  $\xi \geq \xi_+$ , we recognize it as

$$\dot{\zeta}(\xi) = \left[ \frac{3}{2} \int_{\xi_+}^{\xi} \left( \frac{z - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{\xi \sqrt{z - \xi_+} dz \sqrt{\xi_+ - \xi_-}}{z \cdot 2\xi} \right]^{-1/3} \frac{\sqrt{(\xi - \xi_+)(\xi - \xi_-)}}{2\xi}.$$

When  $s \in I_{1,N}$  with  $s_0 \geq 0$ , we always have the bounds

$$1 \leq \left( \frac{z - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{\xi}{z} \leq \left( \frac{\xi - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{\xi}{\xi_+}. \quad (68)$$

Plugging these bounds into our modification of  $\dot{\zeta}(\xi)$ , we obtain the lower and upper bounds for  $\dot{\zeta}(\xi)$  as

$$\frac{\xi_+^{1/3} (\xi - \xi_-)^{1/3}}{2^{2/3} \xi} \leq \dot{\zeta}(\xi) \leq \frac{(\xi - \xi_-)^{1/2}}{2^{2/3} \xi^{2/3} (\xi_+ - \xi_-)^{1/6}},$$

where as  $N \rightarrow \infty$ , both bounds converge to the same limit:

$$\frac{(\xi_+^\infty - \xi_-^\infty)^{1/3}}{2^{2/3} (\xi_+^\infty)^{2/3}} = \lim_{N \rightarrow \infty} \dot{\zeta}_N.$$

We remark that because of (65), the convergence is uniform on  $I_{1,N}$ , which is crucial for deriving finite  $N$  bounds from the limit.

If  $s_0 < 0$ , we only need to consider the case where  $s \in [s_0, 0]$ , for the case where  $s \geq 0$  has essentially been considered in the above derivation. When  $s \in [s_0, 0]$ , the definition of  $\dot{\zeta}(\xi)$  is changed to

$$\begin{aligned} \dot{\zeta}(\xi) &= \left( \frac{3}{2} \int_\xi^{\xi_+} \frac{\sqrt{(\xi_+ - z)(z - \xi_-)}}{2z} dz \right)^{-1/3} \frac{\sqrt{(\xi_+ - \xi)(\xi - \xi_-)}}{2\xi} \\ &= \left[ \frac{3}{2} \int_\xi^{\xi_+} \left( \frac{z - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{\xi \sqrt{\xi_+ - z} dz \sqrt{\xi_+ - \xi_-}}{z \cdot 2\xi} \right]^{-1/3} \frac{\sqrt{(\xi_+ - \xi)(\xi - \xi_-)}}{2\xi} \end{aligned}$$

In this case, we have for all  $s \in [s_0, 0]$ ,

$$\left( \frac{\xi - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{\xi}{\xi_+} \leq \left( \frac{z - \xi_-}{\xi_+ - \xi_-} \right)^{1/2} \frac{z}{\xi_+} \leq 1. \quad (69)$$

We notice that all the bounds tend to 1 when  $N \rightarrow \infty$ . Hence, plugging these bounds to our modification of  $\dot{\zeta}(\xi)$ , we obtain the lower and upper bounds for it that tend to the same limit as when  $s_0 \geq 0$ . Thus, we conclude for  $\dot{\zeta}(\xi)$  that when  $N \geq N_0(s_0, \gamma)$ ,

$$C_1 \lim_{N \rightarrow \infty} \dot{\zeta}_N \leq \dot{\zeta}(\xi) \leq C_2 \lim_{N \rightarrow \infty} \dot{\zeta}_N, \quad \text{for all } s \in I_{1,N}. \quad (70)$$

Such a derivation is valid, since the convergence to the limit is uniform for  $s \in I_{1,N}$ .

Finally, we study the behavior of  $\ddot{\zeta}(\xi)$ . To this end, we first derive a convenient representation for it. By the definition of  $\zeta$ , we have  $(\dot{\zeta})^2 = f\zeta^{-1}$ . We then take derivative with respect to  $\xi$  on both sides and collect to get

$$\ddot{\zeta} = \frac{f'\zeta - f\dot{\zeta}}{2\dot{\zeta}\zeta^2}.$$

Furthermore, we plug in  $\zeta = f/\dot{\zeta}^2$  and obtain the final representation as

$$\ddot{\zeta} = \frac{f'\dot{\zeta} - \dot{\zeta}^4}{2f}.$$

Noticing the definition of  $f$ , we could regard the above representation as the product of three factors:  $\dot{\zeta}(\xi)$ ,  $(f'(\xi) - \dot{\zeta}(\xi)^3)/(\xi - \xi_+)$  and  $2\xi^2/(\xi - \xi_-)$ . The first factor  $\dot{\zeta}$  has already been studied. We first investigate the second factor:  $(f'(\xi) - \dot{\zeta}(\xi)^3)/(\xi - \xi_+)$ .

As before, we start with the assumption that  $s_0 \geq 0$ . By the definition of  $f$ , we have

$$f'(\xi) = \frac{\xi - \xi_-}{4\xi^2} + \frac{\xi - \xi_+}{4\xi^2} - \frac{(\xi - \xi_+)(\xi - \xi_-)}{2\xi^3}.$$

For  $f'(\xi) - \dot{\zeta}(\xi)^3$ , we consider first the quantity  $\mathcal{I}(\xi) = (\xi - \xi_-)/4\xi^2 - \dot{\zeta}(\xi)^3$ . By (68) and straightforward calculation, we obtain

$$\left[1 - \left(\frac{\xi - \xi_-}{\xi_+ - \xi_-}\right)^{1/2}\right] \frac{\xi - \xi_-}{4\xi^2} \leq \mathcal{I}(\xi) \leq \left(1 - \frac{\xi_+}{\xi}\right) \frac{\xi - \xi_-}{4\xi^2}.$$

Hence, we obtain that when  $N \geq N_0(s_0, \gamma)$ , for all  $s \in I_{1,N}$

$$\left|\frac{\mathcal{I}(\xi)}{\xi - \xi_+}\right| \leq \frac{\xi - \xi_-}{\xi^3}, \quad \text{and hence} \quad \left|\frac{f'(\xi) - \dot{\zeta}(\xi)^3}{\xi - \xi_+}\right| \leq \frac{1}{4\xi^2} + \frac{2(\xi - \xi_-)}{\xi^3} \leq \frac{9}{4\xi^2} \leq \frac{C}{(\xi_+^\infty)^2}.$$

Moreover, when  $N \geq N_0(s_0, \gamma)$ , we could also have

$$\left|\dot{\zeta}(\xi)\right| \leq (\xi_+^\infty - \xi_-^\infty)^{1/3} (\xi_+^\infty)^{-2/3}, \quad \text{and} \quad \left|\frac{2\xi^2}{\xi_+ - \xi_-}\right| \leq \frac{4(\xi_+^\infty)^2}{\xi_+^\infty - \xi_-^\infty}.$$

Multiplying the three bounds, we finally obtain that when  $N \geq N_0(s_0, \gamma)$ ,

$$\left|\ddot{\zeta}(\xi)\right| \leq C(\xi_+^\infty)^{-2/3}(\xi_+^\infty - \xi_-^\infty)^{-2/3} = C\gamma^{-1/3}(1 + \sqrt{\gamma})^{-4/3}(1 + \gamma)^{4/3}, \quad \text{for all } s \in I_{1,N}. \quad (71)$$

We remark that when  $s_0 < 0$ , we just focus on  $s \in [s_0, 0]$ . In this case, the quantity  $\mathcal{I}(\xi)$  becomes

$$\mathcal{I}(\xi) = \frac{\xi - \xi_-}{4\xi^2} - \left(\frac{\sqrt{(\xi_+ - \xi)(\xi - \xi_-)}}{2\xi}\right)^3 \left[\frac{3}{2} \int_\xi^{\xi_+} \left(\frac{z - \xi_-}{\xi_+ - \xi_-}\right)^{1/2} \frac{\xi \sqrt{\xi_+ - z} dz}{z} \frac{\sqrt{\xi_+ - \xi_-}}{2\xi}\right]^{-1}$$

with (69) holds. Everything else follows just as in the study of  $\dot{\zeta}(\xi)$ . In particular, (71) still holds.

Finally, by the definition of  $R'_N(\xi)$  in (64) and our analysis of  $\dot{\zeta}_N$ ,  $\dot{\zeta}(\xi)$  and  $\ddot{\zeta}(\xi)$ , we have that when  $N \geq N_0(s_0, \gamma)$ ,

$$\left|R'_N(\xi)\right| \leq C\gamma^{-1/2}(1 + \gamma), \quad \text{for all } s \in I_{1,N}. \quad (72)$$

This bound, together with (66), gives

$$\left|R_N(\xi) - 1\right| \leq C\sqrt{\gamma}(1 + \sqrt{\gamma})(1 + 1/\sqrt{\gamma})^{1/3}N^{-2/3}s \leq CN^{-2/3}|s|, \quad \text{for all } s \in I_{1,N}. \quad (73)$$

### A.3.2 Behavior of $\kappa_N^{2/3}\zeta$

Exploiting a simple Taylor expansion at  $\xi_+$  to the second order, we obtain that

$$\begin{aligned} \kappa_N^{2/3}\zeta(\xi) &= \kappa_N^{2/3}\zeta(\mu_{n,N}/\kappa_N + s\sigma_{n,N}/\kappa_N) \\ &= \kappa_N^{2/3}\zeta(\xi_+) + \kappa_N^{-1/3}\sigma_{n,N}\dot{\zeta}_N s + \frac{1}{2}\kappa_N^{-4/3}\sigma_{n,N}^2\ddot{\zeta}(\xi^*)s^2. \end{aligned}$$

Recalling that  $\zeta(\xi_+) = 0$  and that  $\sigma_{n,N}\kappa_N^{-1/3}\dot{\zeta}_N = 1$ , we obtain that

$$\kappa_N^{2/3}\zeta(\xi) - s = \frac{1}{2} \frac{\sigma_{n,N}}{\kappa_N} \frac{\ddot{\zeta}(\xi^*)}{\dot{\zeta}_N} s^2.$$

According to our previous discussion, we have

$$\frac{\sigma_{n,N}}{\kappa_N} = O\left(N^{-2/3}\right) \quad \text{and} \quad \frac{\ddot{\zeta}(\xi^*)}{\dot{\zeta}_N} = O(1), \quad \text{for all } s \in I_{1,N}.$$

Hence for all  $s \in I_{1,N}$ , when  $N \geq N_0(s_0, \gamma)$ , we have

$$\left| \kappa_N^{2/3}\zeta(\xi) - s \right| \leq CN^{-2/3}s^2.$$

Note that on  $I_{1,N}$ ,  $|s| \leq s_1 N^{1/6}$  and hence we could modify the above bound to be

$$\left| \kappa_N^{2/3}\zeta(\xi) - s \right| \leq \left( CN^{-2/3}s^2 \right) \wedge \frac{|s|}{2} \wedge 1, \quad \text{for all } s \in I_{1,N}. \quad (74)$$

#### A.4 Choice of $s_1$ and its consequences

The key point in our choice of  $s_1$  is to ensure that when  $s \geq s_1$ , we have

$$\frac{2}{3}\kappa_N\zeta^{3/2} \geq \frac{3}{2}s. \quad (75)$$

To this end, recall that in Johnstone (2001, A.8), one could choose  $\tilde{s}_1(\gamma) = C(\gamma)(1 + \delta)$  with some  $\delta > 0$ , such that when  $s \geq \tilde{s}_1(\gamma)$ , we have  $\sqrt{f(\xi)} \geq 2/\sigma_{n,N}$  and hence if  $s \geq 4\tilde{s}_1(\gamma)$ ,

$$\frac{2}{3}\kappa_N\zeta^{3/2} = \kappa_N \int_{\xi_+}^{\xi} \sqrt{f(z)} dz \geq \kappa_N \frac{2}{\sigma_{n,N}} (s - \tilde{s}_1(\gamma)) \frac{\sigma_{n,N}}{\kappa_N} = 2(s - \tilde{s}_1(\gamma)) \geq \frac{3}{2}s.$$

Moreover, by the analysis in El Karoui (2006b, A.6.4),  $\tilde{s}_1(\gamma)$  could be chosen independently of  $\gamma$  and hence we could define our  $s_1$  to be

$$s_1 = 4\tilde{s}_1$$

which is independent of  $\gamma$  and such that (75) holds. Moreover, for our convenience of arguments, we could also impose the constraint that  $s_1 \geq 1$ .

After specifying our choice of  $s_1$ , we spell out two of its consequences. The first of them is that when  $s \geq s_1 \geq 1$ ,

$$\mathbf{E}^{-1}(\kappa_N^{2/3}\zeta) \leq C \exp(-3s/2) \leq C \exp(-s). \quad (76)$$

This is from the observation that  $\mathbf{E}(x) \geq C \exp(\frac{2}{3}x^{3/2})$  and hence

$$\mathbf{E}^{-1}(\kappa_N^{2/3}\zeta) \leq C \exp\left(-\frac{2}{3}\kappa_N\zeta^{3/2}\right) \leq C \exp(-3s/2).$$

The other consequence is about the behavior of  $s'$  defined in (60) when  $s \geq s_1$ . Remembering that  $s_1 \geq 1$ , we then have that when  $s \geq s_1$  and  $N \geq N_0(\gamma)$ ,

$$s' - \frac{s}{2} = \Delta_N + \left( \frac{\sigma_{n-1,N-1}}{\sigma_{n-2,N}} - \frac{1}{2} \right) s \geq \Delta_N + \frac{s_1}{2} \geq \Delta_N + \frac{1}{2} \geq 0. \quad (77)$$

The last inequality holds when  $N \geq N_0(\gamma)$  for  $\Delta_N = O(N^{-1/3})$ .



## A.5 Proofs of Proposition 1 and Lemma 4

### A.5.1 Proof of Proposition 1

By the definition (16) of the determinant for operators in class  $\mathcal{A}$ , we have a first decomposition as

$$\begin{aligned} |\det(I - A) - \det(I - B)| &= |\det_2(I - A) \exp(-\operatorname{tr}A) - \det_2(I - B) \exp(-\operatorname{tr}B)| \\ &\leq |\det_2(I - A) - \det_2(I - B)| [|\exp(-\operatorname{tr}A) - \exp(-\operatorname{tr}B)| + \exp(-\operatorname{tr}B)] \\ &\quad + |\det_2(I - B)| |\exp(-\operatorname{tr}A) - \exp(-\operatorname{tr}B)|. \end{aligned}$$

According to Gohberg et al. (2000, p.69, Theorem 7.4), we have the bound for the 2-determinant as

$$|\det_2(I - A) - \det_2(I - B)| \leq \|A - B\|_2 \exp \left[ \frac{1}{2} (1 + \|A\|_2 + \|B\|_2)^2 \right].$$

Moreover, for any  $A, B \in \mathcal{A}$ , the Hilbert-Schmidt norm satisfies

$$\|A - B\|_2 \leq \sum_i \|A_{ii} - B_{ii}\|_2 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \leq \sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2.$$

Recalling that for any trace class operator  $A$ ,  $\operatorname{tr}(A) \leq \|A\|_1$ , we obtain

$$\exp(-\operatorname{tr}B) \leq \exp|\operatorname{tr}B| \leq \exp(|\operatorname{tr}B_{11}| + |\operatorname{tr}B_{22}|) \leq \exp(\|B_{11}\|_1 + \|B_{22}\|_1).$$

Observing that for  $|x| \leq 1/2$ ,  $|e^x - 1| \leq 2|x|$ , we obtain that, when  $\sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \leq 1/2$ ,

$$\begin{aligned} |\exp(-\operatorname{tr}A) - \exp(-\operatorname{tr}B)| &\leq 2 \exp(-\operatorname{tr}B) |\operatorname{tr}A - \operatorname{tr}B| \leq 2 \sum_i \|A_{ii} - B_{ii}\|_1 e^{\|B_{11}\|_1 + \|B_{22}\|_1} \\ &\leq 2 \left( \sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \right) \exp(\|B_{11}\|_1 + \|B_{22}\|_1). \end{aligned}$$

Plugging all these bounds into our first decomposition, we obtain an intermediate bound as  $M(A, B) \left( \sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \right)$ , where

$$M(A, B) = 2 |\det(I - B)| + 2 \exp \left[ \frac{1}{2} (1 + \|A\|_2 + \|B\|_2)^2 + \sum_i \|B_{ii}\|_1 \right].$$

Under the given condition,

$$\begin{aligned} 1 + \|A\|_2 + \|B\|_2 &\leq 1 + 2\|B\|_2 + \|A - B\|_2 \\ &\leq 1 + 2\|B\|_2 + \sum_i \|A_{ii} - B_{ii}\|_1 + \sum_{i \neq j} \|A_{ij} - B_{ij}\|_2 \leq 2 + 2\|B\|_2, \end{aligned}$$

which reduce  $M(A, B)$  to the constant  $M(B)$  claimed.

### A.5.2 Proof of Lemma 4

By definition, we have

$$\zeta_N^3 = \frac{\xi_+ - \xi_-}{4\xi_+^2} = \frac{\kappa_N}{\sigma_{n,N}^3}.$$

Thus, we obtain from direct calculation that

$$\sqrt{f(\xi)} = \frac{(\xi - \xi_+)(\xi - \xi_-)}{2\xi} \geq r \sqrt{\frac{\sigma_{n,N}}{\kappa_N}} \sqrt{\frac{\xi_+ - \xi_-}{2\xi_+}} \frac{\xi_+}{\xi} = r \sqrt{\frac{\sigma_{n,N}}{\kappa_N}} \sqrt{\frac{\kappa_N}{\sigma_{n,N}^3}} \frac{\xi_+}{\xi} = \frac{r\xi_+}{\sigma_{n,N}\xi}.$$

## B Logarithmic Transformation and the Smallest Eigenvalue

In this part, we give a brief account of how one could derive the similar second order accuracy results claimed in (5) and (7) with a logarithmic transformation. In many aspects, the derivation here for Laguerre orthogonal ensembles [as based on Adler et al. (2000, Proposition 4.2)] is parallel to what Johnstone (2007) did for Jacobi orthogonal ensembles.

### B.1 Logarithmic transformation for the largest eigenvalue

For the largest eigenvalue, we assume the same setting as that in the beginning of Section 3.2. With  $\phi_k$  defined in (13), let

$$\bar{\phi}_k(x; \tilde{\alpha}) = (-1)^j \phi_k(x; \tilde{\alpha}) / \sqrt{x}. \quad (78)$$

Then setting  $a_N = \sqrt{N(N + \alpha_N - 1)}$ , we have the following alternative way of expressing  $S_{N,1}$  in term of  $S_{k,2}$ , the correlation kernel occurring in LUE( $k, \tilde{\alpha}$ ) model:

$$S_{N,1}(x, y; \alpha_N - 1) = \sqrt{\frac{y}{x}} S_{N-1,2}(x, y; \alpha_N) + \sqrt{\frac{N-1}{N}} \frac{a_N}{2} \bar{\phi}_{N-1}(x; \alpha_N) (\varepsilon \bar{\phi}_{N-2})(y; \alpha_N). \quad (79)$$

As a comparison, the central formula (15) could be rewritten as

$$S_{N,1}(x, y; \alpha_N - 1) = S_{N,2}(x, y; \alpha_N - 1) + \frac{a_N}{2} \bar{\phi}_{N-1}(x; \alpha_N) (\varepsilon \bar{\phi}_N)(y; \alpha_N - 2).$$

The equivalence of the above two representations is given in Adler et al. (2000, Appendix) and hence omitted here.

We make use of the representation (79) to give an alternative second order accuracy argument with a logarithmic transformation. Recalling  $\alpha_N = n - N$ , we define

$$\mu_k = \left( \sqrt{k + \frac{1}{2}} + \sqrt{k + \alpha_N + \frac{1}{2}} \right)^2, \quad \sigma_k = \left( \sqrt{k + \frac{1}{2}} + \sqrt{k + \alpha_N + \frac{1}{2}} \right) \left( \frac{1}{\sqrt{k + \frac{1}{2}}} + \frac{1}{\sqrt{k + \alpha_N + \frac{1}{2}}} \right)^{1/3}.$$

Then we let

$$\check{\phi}_k(x) = (-1)^k \frac{(N-1)^{1/4} (N-1 + \alpha_N)^{1/4}}{\sqrt{2}} x^{1/2} \phi_k(x; \alpha_N).$$

For  $\hat{S}_{N-1,2}(u, v; \alpha_N) = S_{N-1,2}(e^u, e^v; \alpha_N) e^{u/2} e^{v/2}$ , we could represent it as

$$\hat{S}_{N-1,2}(u, v; \alpha_N) = \int_0^\infty [\check{\phi}_{N-2}(e^{u+w}) \check{\phi}_{N-1}(e^{v+w}) + \check{\phi}_{N-1}(e^{u+w}) \check{\phi}_{N-2}(e^{v+w})] dw.$$

We then define

$$\nu_{n,N} = \log \mu_{N-1}, \quad \tau_{n,N} = \sigma_{N-1}/\mu_{N-1}, \quad \text{and} \quad \tau(s) = \exp(\nu_{n,N} + s\tau_{n,N}).$$

The  $\tau$ -transformation induces the following transformed Laguerre polynomials:

$$\psi_\tau(s) = \tau_{n,N} \check{\phi}_{N-1}(\tau(s)), \quad \phi_\tau(s) = \tau_{n,N} \check{\phi}_{N-2}(\tau(s)).$$

Define  $S_\tau(s, t) = \sqrt{\tau'(s)\tau'(t)} S_{N-1,2}(\tau(s), \tau(t); \alpha_N)$ , we have the following integral representation from the expression for the  $\hat{S}_{N-1,2}$  kernel:

$$S_\tau(s, t) = \int_0^\infty [\phi_\tau(s+z)\psi_\tau(t+z) + \phi_\tau(t+z)\psi_\tau(s+z)] dz. \quad (80)$$

Moreover, if we define the following quantities [fix  $s_0 \in \mathbb{R}$ , with  $s, t \geq s_0$ ]

$$q_N(s) = \sqrt{\tau'(s_0)/\tau'(s)}, \quad \text{and} \quad S_\tau^R(s, t) = S_\tau(s, t) + \psi_\tau(s)(\varepsilon\phi_\tau)(t),$$

we have

$$F_{N,1}(s') = P(x_1 \leq \tau(s')) = P((\log x_1 - \nu_{n,N})/\tau_{n,N} \leq s') = \sqrt{\det(I - K_\tau)},$$

where the new operator  $K_\tau$  has a  $2 \times 2$  matrix kernel with entries given by

$$\begin{aligned} K_{\tau,11}(s, t) &= q_N(s)q_N^{-1}(t)S_\tau^R(s, t); & K_{\tau,12}(s, t) &= -q_N(s)q_N(t)\partial_t S_\tau^R(s, t); \\ K_{\tau,21}(s, t) &= q_N^{-1}(s)q_N^{-1}(t)[\varepsilon_1 S_\tau^R(s, t) - \varepsilon(s-t)]; & K_{\tau,22}(s, t) &= K_{\tau,11}(t, s). \end{aligned} \quad (81)$$

By Proposition 1, we need to obtain entrywise bound for  $K_\tau - K_{GOE}$  here. To this end, a convenient representation of the kernel difference as in Section 3.4 is most helpful.

For the transformed Laguerre polynomials  $\phi_\tau$  and  $\psi_\tau$ , we have

$$\int_{-\infty}^\infty \psi_\tau = 0, \quad \text{and} \quad \int_{-\infty}^\infty \phi_\tau = \frac{(N-1)^{1/4}(n-1)^{1/4}\Gamma^{1/2}(n-1)\Gamma(\frac{N+1}{2})}{2^{\alpha_{N-2}}(N-1)\Gamma^{1/2}(N-1)\Gamma(\frac{n}{2})}.$$

For notational convenience, let  $\tilde{\beta}_N = \frac{1}{2} \int_{-\infty}^\infty \phi_\tau = \frac{1}{\sqrt{2}} + O(N^{-1})$ .

With the replacement of  $\varepsilon$  by  $\tilde{\varepsilon}$  in (24) and the matrices  $\tilde{L}, L_1$  and  $L_2$  introduced in Section 3.4, we obtain that

$$K_\tau = Q_N(s) [K_\tau^R + K_{\tau,1}^F + K_{\tau,2}^F + K^\varepsilon] Q_N^{-1}(t).$$

with the unspecified components given by

$$K_\tau^R = \tilde{L}[S_\tau - \psi_\tau \otimes \tilde{\varepsilon}\phi_\tau], \quad K_{\tau,1}^F = \tilde{\beta}_N L_1[\psi_\tau(s)], \quad K_{\tau,2}^F = \tilde{\beta}_N L_2[\psi_\tau(t)],$$

where  $Q_N(s) = \text{diag}(q_N(s), q_N^{-1}(s))$  and as before  $G(s) = \text{Ai}(s)/\sqrt{2}$ .

For

$$\Delta_N = \frac{\log \mu_{N-1} - \log \mu_{N-2}}{\sigma_{N-2}/\mu_{N-2}},$$

set  $G_N = G + \Delta_N G'$ , we have

$$\begin{aligned} K_\tau^R - K_R &= \tilde{L} [S_\tau - S_A - \psi_\tau \otimes \tilde{\varepsilon}\phi_\tau + G \otimes \tilde{\varepsilon}(G_N - \Delta_N G')] \\ &= \tilde{L} [S_\tau - S_A + \Delta_N G \otimes G] - \tilde{L} [\psi_\tau \otimes \tilde{\varepsilon}\phi_\tau - G \otimes \tilde{\varepsilon}G_N] \\ &= \delta^{R,I} + \delta_0^F. \end{aligned}$$

If we write  $S_{A_N} = G \diamond G_N + G_N \diamond G$ , we have  $\delta^{R,I} = \tilde{L}[S_\tau - S_{A_N}]$ .

Finally, we organize  $K_\tau - K_{GOE}$  as

$$K_\tau - K_{GOE} = \delta^{R,D} + \delta^{R,I} + \delta_0^F + \delta_1^F + \delta_2^F + \delta^\varepsilon,$$

where the unspecified terms are defined as the following:

$$\begin{aligned} \delta^{R,D} &= Q_N(s)K_\tau^R Q_N^{-1}(t) - K_\tau^R, \\ \delta_i^F &= Q_N(s)K_{\tau,i}^F Q_N^{-1}(t) - K_i^F, \quad \text{for } i = 1, 2, \text{ and} \\ \delta^\varepsilon &= Q_N(s)K^\varepsilon Q_N^{-1}(t) - K^\varepsilon. \end{aligned}$$

With the above representation of the kernel difference, we could apply the machineries in Johnstone (2007) to obtain the desired second order accuracy of the Tracy-Widom approximation to the distribution of  $(\log x_1 - \nu_{n,N})/\tau_{n,N}$ . After establishing the result in RMT notation, we replace  $N$  by  $p$  and hence obtain the bound in (5).

## B.2 The smallest eigenvalue

We first restate the claim in (7) in a more friendly way. Let  $\nu_{n,N}^-$  and  $\tau_{n,N}^-$  be the centering and scaling constants defined in (6), with  $p$  replaced by  $N$ . Then for  $x_N$  the smallest eigenvalue in the model (8) [with  $\tilde{\alpha} = \alpha_N - 1$ ], there exists a continuous and nonincreasing function  $C(\cdot)$ , such that for all real  $s_0$ , there is an integer  $N_0(s_0, \gamma)$  for which we have that for any  $s \geq s_0$  and  $N \geq N_0(s_0, \gamma)$ ,

$$\left| P\{\log x_N > \nu_{n,N}^- - s\tau_{n,N}^-\} - F_1(s) \right| \leq C(s_0)N^{-2/3} \exp(-s/2). \quad (82)$$

Fix  $x_0 \geq 0$  and consider any  $x' \in [0, x_0]$ . To prove (82), we first observe that for  $x_N$  in model (8), when  $N$  is even, choosing  $\chi = I_{0 \leq x \leq x'}$ , we have

$$P\{x_N > x'\} = \sqrt{\det(I - K_N \chi)},$$

where  $K_N$  is the same operator as for  $x_1$ , which has the kernel (11). If we think of  $K_N$  as Hilbert-Schmidt operator on  $L^2([0, x']; \rho) \oplus L^2([0, x']; \rho^{-1})$  with  $\rho$  any weight function chosen from some proper class, then the above formula changes to

$$P\{x_N > x'\} = \sqrt{\det(I - K_N)}.$$

Introduce the transformation

$$\tau(s) = \exp(\nu_{n,N}^- - s\tau_{n,N}^-),$$

and let  $s_0 = \tau^{-1}(x_0)$  and  $s_0 \leq s' = \tau^{-1}(x')$ , we have  $\tau^{-1}([0, x']) = [s', \infty)$ . By defining  $\phi_\tau = -\tau_{n,N}^- \check{\phi}_{N-1}(\tau(s))$  and  $\psi_\tau = -\tau_{n,N}^- \check{\phi}_{N-2}(\tau(s))$  and using the alternative representation (79), the formal derivation for the largest eigenvalue in B.1 could be carried out analogously for the smallest eigenvalue. In particular, we have the integral representation (80) for  $S_\tau(s, t) = \sqrt{\tau'(s)\tau'(t)}S_{N-1,2}(\tau(s), \tau(t); \alpha_N)$  and

$$P(x_N > \tau(s')) = P((\log x_N - \nu_{n,N}^-)/\tau_{n,N}^- > -s') = \sqrt{\det(I - K_\tau)},$$

with  $K_\tau$  thought of as Hilbert-Schmidt operator on  $L^2([s', \infty); \rho \circ \tau) \oplus L^2([s', \infty); \rho^{-1} \circ \tau)$  with entries given by (81). We remark that the actual definition of  $\phi_\tau$  and  $\psi_\tau$  used in these formulas have changed, albeit the formal representations remain the same.

The rest of the proof for the smallest eigenvalue becomes the routine procedure of a) finding a representation for the kernel difference  $K_\tau - K_{GOE}$  and b) studying the asymptotic behavior of the transformed Laguerre polynomials  $\phi_\tau$  and  $\psi_\tau$ . The former is very similar to the largest eigenvalue case while the latter could be obtained by applying the Liouville-Green approach to analyze the behavior of the solution to the differential equation (48) around the lower turning point  $\xi_-$ .

## Acknowledgment

The author is most grateful to Professor Iain Johnstone for his indispensable advice during the development of this project. The author would also like to thank Debashis Paul for sharing a draft of his paper on the smallest eigenvalue. This work is supported in part by grants NSF DMS 0505303 and NIH EB R01 EB001988.

## References

- M. Adler, P. J. Forrester, T. Nagao, and P. van Moerbeke. Classical skew orthogonal polynomials and random matrices. *J. Statist. Phys.*, 99:141–170, 2000.
- T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, 3rd edition, 2003.
- E. Candes and T. Tao. Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inform. Theory*, 52:5406–5425, 2006.
- N. G. de Bruijn. On some multiple integrals involving determinants. *J. Indian Math. Soc.*, 19:133–151, 1955.
- D. L. Donoho. For most large underdetermined systems of linear equations the minimal  $\ell^1$ -norm solution is also the sparsest solution. 2004.
- I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.
- A. Edelman and P.-O. Persson. Numerical methods for eigenvalue distributions of random matrices. Technical report, Massachusetts Institute of Technology, 2002.
- N. El Karoui. On the largest eigenvalue of Wishart matrices with identity covariance when  $n, p$  and  $p/n \rightarrow \infty$ . *arXiv:math/0309335v1*, 2006a.
- N. El Karoui. A rate of convergence result for the largest eigenvalue of complex white Wishart matrices. *Ann. Probab.*, 34:2077–2117, 2006b.
- P. J. Forrester. Log-gases and random matrices. Book manuscript, 2004.
- I. Gohberg, S. Goldberg, and N. Krupnik. *Traces and Determinants of Linear Operators*, volume 116 of *Operator Theory, Advances and Applications*. Birkhäuser Verlag, Basel, 2000.
- G. H. Golub and C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd ed. edition, 1996.

- K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209:437–476, 2000.
- I. Johnstone. On the distribution of the largest eigenvalue in principal component analysis. *Ann. Statist.*, 29:295–327, 2001.
- I. Johnstone. High dimensional statistical inference and random matrices. *arXiv:math/0611589*, 2006.
- I. Johnstone. Canonical correlation analysis and Jacobi ensembles: Tracy Widom limits and rates of convergence. Unpublished manuscript, 2007.
- P. Koev and A. Edelman. The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comp.*, 75:833–846, 2006.
- D. Luca, S. Ringquist, L. Klei, A. B. Lee, C. Gieger, H. E. Wichmann, S. Schreiber, M. Krawczak, Y. Lu, A. Styche, B. Devlin, K. Roeder, and M. Trucco. On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am. J. Hum. Genet.*, 82:453–463, 2008.
- K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, 1979.
- R. J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley and Sons, 1982.
- T. Nagao and P. J. Forrester. Asymptotic correlations at the spectrum edge of random matrices. *Nucl. Phys. B*, 435:401–420, 1995.
- F. W. J. Olver. *Asymptotics and Special Functions*. Academic Press, 1974.
- N. Patterson, A. L. Price, and D. Reich. Population structure and eigenanalysis. *PLoS Genet.*, 2:e190, 2006. doi: 10.1371/journal.pgen.0020190.
- D. Paul. Distribution of the smallest eigenvalue of  $\text{Wishart}(N, n)$  when  $N/n \rightarrow 0$ . Unpublished manuscript, 2006.
- A. L. Price, N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, 38:904–909, 2006.
- M. Reed and B. Simon. *Methods of Modern Mathematical Physics. Vol. I: Functional Analysis*. Academic Press, 1980.
- S. N. Roy. On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Stat.*, 24:220–238, 1953.
- G. Szegő. *Orthogonal Polynomials*. Amer. Math. Soc., Providence, RI., 4th edition, 1975.
- C. A. Tracy and H. Widom. Matrix kernels for the Gaussian orthogonal and symplectic ensembles. *Ann. Institut. Fourier, Grenoble*, 55:2197–2207, 2005.
- C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Commun. Math. Phys.*, 159:151–174, 1994.
- C. A. Tracy and H. Widom. On orthogonal and symplectic matrix ensembles. *Commun. Math. Phys.*, 177:727–754, 1996.

- C. A. Tracy and H. Widom. Correlation functions, cluster functions, and spacing distributions for random matrices. *J. Statist. Phys.*, 92:809–835, 1998.
- H. Widom. On the relation between orthogonal, symplectic and unitary matrix ensembles. *J. Statist. Phys.*, 94:347–364, 1999.