

# Tracy-Widom distribution, $\text{Airy}_2$ process and its sample path properties

SU Zhong-gen      LEI Yu-huan      SHEN Tian

**Abstract.** Tracy-Widom distribution was first discovered in the study of largest eigenvalues of high dimensional Gaussian unitary ensembles (GUE), and since then it has appeared in a number of apparently distinct research fields. It is believed that Tracy-Widom distribution have a universal feature like classic normal distribution.  $\text{Airy}_2$  process is defined through finite dimensional distributions with Tracy-Widom distribution as its marginal distributions. In this introductory survey, we will briefly review some basic notions, intuitive background and fundamental properties concerning Tracy-Widom distribution and  $\text{Airy}_2$  process. For sake of reading, the paper starts with some simple and well-known facts about normal distributions, Gaussian processes and their sample path properties.

## §1 Normal Distributions and Gaussian Processes

This section contains some simple and well-known facts about normal distributions and Gaussian processes. The reader can find them in most of advanced probability and mathematical statistics textbooks for graduate students, see [4, 5] by Billingsley, [12] by Durrett, [23] by Lin, Lu and Su. The purpose of this section is to provide some helpful clues about the next research work around Tracy-Widom distribution and  $\text{Airy}_2$  process from well-studied normal distributions and Gaussian processes. Tracy-Widom distributions are quite novel compared with normal distributions, and have been very hot words in probability and statistics in the past twenty-five years.

Throughout the paper,  $c_1, c_2, \dots$ , stand for positive constants possibly varying from place to place.

### 1.1 Normal Distributions

The normal distribution is arguably the most important distribution in probability theory and mathematical statistics. The normal density function is defined as follows.

$$p(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

---

Received: 2020-9-1.      Revised: 2020-11-17.

MR Subject Classification: 60K35, 60B20.

Keywords:  $\text{Airy}_2$  process, Dyson Brownian motion, Tracy-Widom distribution, sample path property.

Digital Object Identifier(DOI): <https://doi.org/10.1007/s11766-021-4251-2>.

Supported by the National Natural Science Foundation of China(11731012, 11871425) and Fundamental Research Funds for Central Universities grant(2020XZZX002-03).

where  $\mu \in \mathbb{R}$  and  $\sigma > 0$  are two parameters. In particular,  $p(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$  is the so-called standard normal density function. It consist of a symmetric bell-shaped curve asymptotically close to the  $x$ -axis as  $|x| \rightarrow \infty$ . It is an elegant and nontrivial exercise to verify

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$$

The standard normal distribution function  $\Phi(x)$  is defined as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

$\Phi(x)$  is a strictly increasing positive function with bounded derivative. However, there is no closed form for  $\Phi(x)$ , people usually use the approximate values in practice. It is easy to see  $\Phi(0) = \frac{1}{2}$ ,  $\Phi(x) = 1 - \Phi(-x)$ , and

$$1 - \Phi(x) = p(x) \left( \frac{1}{x} - \frac{1}{x^3} + \dots + (-1)^k \frac{(2k-1)!!}{x^{2k+1}} + \dots \right), \quad x > 0. \tag{1}$$

A widely used simple inequality is for each  $x > 0$

$$\frac{x}{x^2 + 1} p(x) < 1 - \Phi(x) < \frac{1}{x} p(x).$$

Let  $X$  be a random variable with probability density function  $p(x)$ . Then it follows

$$EX^{2k-1} = 0, \quad EX^{2k} = (2k-1)!!, \quad k \geq 1$$

$$Ee^{tX} = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

Let  $Y = \mu + \sigma X$ , then  $Y$  has probability density function  $p(x; \mu, \sigma^2)$ .

## 1.2 Central Limit Theorems

The first landmark achievement in probability theory is Bernoulli's law of large numbers for binomial random variables. It gives a completely new and rigorous mathematical interpretation for the belief that frequency has always a limit as the number of times of trials is sufficiently large. In particular, suppose that  $\xi_1, \xi_2, \dots$ , is a sequence of i.i.d.random variables,  $P(\xi_1 = 1) = p, P(\xi_1 = 0) = 1 - p$ , where  $0 < p < 1$ . Let  $S_n = \sum_{i=1}^n \xi_i$ , then as  $n \rightarrow \infty$ ,

$$P\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) = \sum_{k: \left|\frac{k}{n} - p\right| > \varepsilon} \binom{n}{k} p^k (1-p)^{n-k} \rightarrow 0.$$

This is the most important result in Bernoulli's celebrated book *Ars conjectandi* published in 1713. In modern terminology, it reads as follows

$$\frac{S_n}{n} \xrightarrow{P} p, \quad n \rightarrow \infty.$$

Following Bernoulli's work, the second great achievement is the central limit theorem. De Moirve ( $p = \frac{1}{2}$ , 1733) and Laplace ( $p \neq \frac{1}{2}$ , 1812) established

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1), \quad n \rightarrow \infty.$$

In other words, for any real numbers  $a < b$

$$P(a \leq S_n \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right).$$

This can be seen from the asymptotic formula  $n! = \sqrt{2\pi n} n^n e^{-n} (1 + o(1))$  discovered by De Moivre and Stirling.

de Moivre and Laplace discovered the above limit distribution for a very simple and special model. A natural question is whether such a result holds true for other cases. This is absolutely not an easy problem. As a matter of fact, it was not solved until 1920s with the introduction of characteristic functions. The Lévy-Feller central limit theorem states that if  $X_1, X_2, \dots$  is a sequence of i.i.d.r.v.'s and  $S_n = \sum_{i=1}^n X_i$ , then

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} N(0, 1), \quad n \rightarrow \infty$$

if and only if  $EX_1 = \mu$  and  $Var(X_1) = \sigma^2 < \infty$ . There have since then been several different proofs in the literature, among which are characteristic function method (Lévy continuity theorem), Linderberge replacement trick, and Stein's equation (Stein's continuity theorem). The importance of this theorem lies in that there is no assumption about the underlying distribution of random variables  $X_i$ 's. The limit distribution is normal for normalized sums of i.i.d.r.v.'s whenever the second-order moment exists and is finite. Not only this, the above limit theorem is still valid for a more general class of random variables. In fact, it has been proven for sums of non-i.i.d.r.v.'s, martingale differences,  $m$ -dependent r.v.'s, mixing sequences, and even for random matrix eigenvalues, determinantal point processes. In other words, the central limit theorem is really a universal principle in probability theory. As you notice, the linear structure of partial sums plays an *essential* role for the validity of the central limit theorem.

Due to its universality and simplicity, the normal distribution and central limit theorems have been widely used in a lot of research areas, like statistical inferences (hypothesis testing and confidence interval estimation), financial market (stock pricing), insurance risk (premium pricing), particle physics (motion path).

### 1.3 Functional Central Limit Theorems

In 1950-60s, a culminated work in probability theory is the following Donsker invariance principle. Let  $\xi_1, \xi_2, \dots$ , be a sequence of random variables with mean zero and variance 1, let  $S_n = \sum_{i=1}^n \xi_i$ . Define for each  $t \in [0, 1]$

$$X_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} \xi_i + \frac{nt - [nt]}{n} \xi_{[nt]+1}.$$

Then

$$\mathbf{X}_n \Rightarrow \mathbf{B}.$$

where  $\mathbf{X}_n = (X_n(t), 0 \leq t \leq 1)$ ,  $\mathbf{B} = (B(t), 0 \leq t \leq 1)$  is Brownian motion (see Subsection 1.5 below). Namely, for any bounded continuous function  $f : C[0, 1] \mapsto \mathbb{R}$ ,  $Ef(\mathbf{X}_n) \mapsto Ef(\mathbf{B})$ .

This theorem does not only prove the existence of Brownian motion, but also provides a unified frame of solving diverse problems. Let  $h : C[0, 1] \mapsto \mathbb{R}^d$  ( $d \geq 1$ ) be a continuous (or even *slightly* discontinuous) mapping, then

$$h(\mathbf{X}_n) \xrightarrow{d} h(\mathbf{B}), \quad n \rightarrow \infty. \quad (2)$$

This is referred to as functional central limit theorem. Note the limit  $h(\mathbf{B})$  does not depend on the distribution of  $\xi_i$ , either. As an illustration, we consider the following case. Letting  $h(\mathbf{x}) = \sup_{0 \leq t \leq 1} \mathbf{x}(t)$ , then by (2)

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} S_k \leq u\right) = P\left(\sup_{0 \leq t \leq 1} B(t) \leq u\right).$$

In turn, to figure out explicitly the limit  $P(\sup_{0 \leq t \leq 1} B(t) \leq u)$ , we let  $\xi_i$  be a special random

variable  $P(\xi_i = \pm 1) = \frac{1}{2}$ , and apply the formula (2) once again,

$$P\left(\sup_{0 \leq t \leq 1} B(t) \leq u\right) = \lim_{n \rightarrow \infty} P\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} S_k \leq u\right).$$

Now we can apply the reflection principle of simple symmetric random walks to yield

$$P\left(\sup_{0 \leq t \leq 1} B(t) \leq u\right) = P(|N(0, 1)| \leq u) = 2\Phi(u) - 1.$$

Thus we get the limit distribution of  $P\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} S_k \leq u\right)$  for general random variables. This methodology is expounded at length in the classic book by Billingsley [4], and its use is endless. Here we emphasize that such ideas will be repeatedly used in the context of GUE Tracy-Widom distribution and Airy<sub>2</sub> process as well.

### 1.4 Multivariate Normal Distributions

Let  $m \geq 1$ . An  $m$ -dimensional random vector  $\mathbf{X} = (X_1, X_2, \dots, X_m)'$  is normal if the joint probability density function is given by

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{m/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1/2}(\mathbf{x}-\mu)}, \quad \mathbf{x} \in \mathbb{R}^m,$$

where  $\mu = (\mu_1, \mu_2, \dots, \mu_m)'$  and  $\Sigma = (\sigma_{ij})_{m \times m} > 0$  with  $\sigma_{ii} = \sigma_i^2 > 0$ ,  $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ , and  $-1 \leq \rho_{ij} = \rho_{ji} \leq 1$ ,  $1 \leq i \neq j \leq m$ .

It is easy to see  $E\mathbf{X} = \mu$ ,  $Var(X_i) = \sigma_i^2$ , and  $Cov(X_i, X_j) = \rho_{ij}\sigma_i\sigma_j$  for  $i \neq j$ .  $X_i$  and  $X_j$  are uncorrelated if and only if they are independent of each other. For any  $1 \leq l < m$ ,  $(X_{i_1}, X_{i_2}, \dots, X_{i_l})$  is an  $l$ -dimensional normal random vector. It is easy to show that  $\mathbf{X}$  is normal if and only if its any linear combination is normal. More generally, let  $A = (a_{ij})_{l \times m}$  be a real matrix, then  $\mathbf{Y} = A\mathbf{X}$  is an  $l$ -dimensional normal (possibly degenerate) random vector with  $E\mathbf{Y} = A\mu$  and  $Cov(\mathbf{Y}) = A\Sigma A'$ . In particular, letting  $A = \Sigma^{-1/2}$ , then  $\mathbf{Y}$  is a normal vector consisting of i.i.d. standard normal components. Note a normal random vector is uniquely determined by its mean vector and covariance structure.

### 1.5 Gaussian Processes with Sample Paths

A random process  $\mathbf{G} = (G(t), t \geq 0)$  is Gaussian if for any  $k \geq 1$  and any  $t_1, t_2, \dots, t_k \geq 0$ ,  $(G(t_1), G(t_2), \dots, G(t_k))$  is a  $k$ -dimensional normal random vector. Gaussian processes have been well-studied and have a wide range of applications. We only take Brownian motions and Ornstein-Uhlenbeck processes as two representative examples below.

Brownian motion is named after English botanist Brown, who first observed the zigzag motion of pollen power in the liquid. Following Bachalier, Einstein, Frechét, Hilbert and Lebesgue, Wiener (1920s) proved the existence of Brownian motion by a rigorous construction and found many fundamental and nice properties like non-differentiability,  $\frac{1}{2} - \delta$  Hölder continuity, unbounded variation and quadratic variational processes.

A standard Brownian motion  $\mathbf{B} = (B(t), t \geq 0)$  is a Gaussian process with  $EB(t) = 0$  and  $EB(s)B(t) = \min(s, t)$ . It is easy to see

$$E|B(t)| = \sqrt{\frac{2}{\pi}}t, \quad EB(t)^2 = t,$$

and

$$E(B(t) - B(s))^2 = t - s, \quad E(B(t) - B(s))^4 = 3(t - s)^2.$$

According to Kolmogorov's continuity criterion, there exists a continuous version for  $\mathbf{B} = (B(t), t \geq 0)$ . But the sample paths are extremely irregular, in particular, nowhere differentiable, and even have unbounded variation.

Note a Gaussian vector is independent if and only if they are pairwise uncorrelated. Hence  $\mathbf{B} = (B(t), t \geq 0)$  has stationary independent increments, and so is a Lévy process. Further, it is a continuous time strong Markov process starting at 0 with transition density function

$$p_{s,t}(x; y) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{(y-x)^2}{2(t-s)}}, \quad s < t, \quad x, y \in \mathbb{R}.$$

Still,  $\mathbf{B} = (B(t), t \geq 0)$  is also a continuous martingale with linear variation process.

Assume that  $\mathbf{B} = (B(t), t \geq 0)$  is a Brownian motion, then so are the following three random processes.

- (1) For any stopping time  $T$ ,  $X(t) = B(T+t) - B(T)$ ,  $t \geq 0$ ;
- (2) For any constant  $c > 0$ ,  $X(t) = \frac{1}{\sqrt{c}}B(ct)$ ,  $t \geq 0$ ;
- (3)  $X(0) = 0$ ,  $X(t) = tB(\frac{1}{t})$ ,  $t > 0$ .

As noticed above, the sample paths of Brownian motion are extremely irregular. But there have been a number of fine characterizations about the module of continuity in the literature. The remarkable uniform modulus of continuity went back to Lévy as early as in 1930s. It states

$$\limsup_{t \rightarrow \infty} \frac{|B(t)|}{\sqrt{2t \log \log t}} = 1 \quad \text{a.s.}$$

and so by the time reciprocal change,

$$\limsup_{t \rightarrow 0} \frac{|B(t)|}{\sqrt{2t \log |\log t|}} = 1 \quad \text{a.s.}$$

Since for any  $s < t$  and  $x > 0$

$$P(|B(t) - B(s)| > x) \leq 2e^{-x^2/2(t-s)},$$

then a standard chaining argument easily gives

$$\sup_{s,t \in [0,1]} \frac{|B(t) - B(s)|}{\sqrt{|t-s| |\log |t-s||}} \leq \Upsilon \quad \text{a.s.}$$

where  $\Upsilon$  is a random variable with  $P(\Upsilon > x) \leq c_2 e^{-c_1 x^2}$ .

More refinement has been done. In particular, there has been a lot of refined results about how big or how small the increments of Brownian motion (even Gaussian processes) is in the past decades. The reader is referred to a classic book by Csörgő and Révész [9] for more information.

As Brown first observed, Brownian motion walks along a zigzag path. How far does it reach within an interval of time  $[0, 1]$ ? It turns out that

$$P\left(\sup_{0 \leq t \leq 1} B(t) > x\right) = 2(1 - \Phi(x)), \quad x \in \mathbb{R}_+$$

and

$$P\left(\sup_{0 \leq t \leq 1} |B(t)| \leq x\right) = \sum_{k=-\infty}^{\infty} (-1)^k (\Phi((2k+1)x) - \Phi((2k-1)x)), \quad x \in \mathbb{R}.$$

By the scaling property, namely  $(B(Tt), 0 \leq t \leq 1) \stackrel{d}{=} (\sqrt{T}B(t), 0 \leq t \leq 1)$ , one can compute the probability  $P(\sup_{0 \leq t \leq T} B(t) > x)$ . Trivially,  $\sup_{0 \leq t \leq T} B(t)$  increases to infinity as  $T \rightarrow \infty$ .

The maximum of  $B(t) - t^2$  over  $[0, \infty)$  plays an important role in certain recent studies on

the asymptotic distributions of tests for monotone hazards, based on integrals type statistics measuring the distance between the empirical cumulative hazard functions and its greatest convex minorant. Groeneboom and Temme [16] in 2010 analyzed the tail behavior of the maximum and gave an asymptotic expansion for  $P(\sup_{t \geq 0}(B(t) - t^2) > x)$ :

$$P\left(\sup_{t \geq 0}\{B(t) - t^2\} > x\right) \sim \frac{1}{\sqrt{3}} \exp\left(-\frac{8}{3\sqrt{3}}x^{3/2}\right) \sum_{k=0}^{\infty} \tau_k x^{-3/2k}, \quad x \rightarrow \infty \tag{3}$$

where  $\tau_k$  is a sequence of numeric constants. The maximum of a two-sided Brownian motion on  $(-\infty, \infty)$  originated at zero can be obtained by the independence and a simple probabilistic analysis.

Let us turn to the second example, Ornstein-Uhlenbeck process. Let  $\mathbf{B} = (B(t), t \geq 0)$  be a standard Brownian motion,  $\lambda > 0$ . Define

$$X(t) = e^{-\frac{\lambda t}{2}} B(e^{\lambda t}), \quad t \in \mathbb{R}.$$

Obviously,  $EX(t) = 0$ ,  $EX(t)^2 = 1$  for each  $t$ . So each  $X(t)$  is a standard normal random variable. Also, for  $s < t$  it follows

$$Cov(X(s), X(t)) = EX(s)X(t) = e^{-\frac{\lambda(t-s)}{2}}$$

and

$$E|X(t) - X(s)|^2 = 2(1 - e^{-\frac{\lambda(t-s)}{2}}).$$

Hence  $\mathbf{X} = (X(t), t \in \mathbb{R})$  is a stationary Gaussian process, whose correlation decays exponentially as the gap of time increases, and so is ergodic in the sense of mean.

This process can also be induced by a stochastic differential equation

$$dX(t) = -X(t)dt + dB(t), \quad X(0) \sim N(0, 1).$$

The probability density function  $p(x, t)$  satisfies the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} + \frac{1}{2} \frac{\partial}{\partial x}(xp)$$

with initial value  $p(0, x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ .

We also remark that the Ornstein-Uhlenbeck process can be interpreted as a scaling limit of a discrete process, in the same way that Brownian motion is a scaling limit of random walks.

Before concluding the Introduction, we would like to take a quick look at the regularity properties of sample paths of general Gaussian processes. Suppose  $\mathbf{X} = (X(t), t \in T)$  is a centered Gaussian process, where  $T$  is an index set. Define the pseud-metric

$$d_X(s, t) = E(X(t) - X(s))^2, \quad s, t \in T.$$

Let  $N(T, d_X, \delta)$  denote the number of finite  $\delta$ -nets covering the set  $T$ , i.e., metric entropy. Then for any  $\delta > 0$

$$c_3 \log N(T, d_X, \delta) \leq E \sup_{t \in T} X(t) \leq c_4 \int_0^\infty (\log N(T, d_X, \delta))^{1/2} d\delta. \tag{4}$$

The LHS of (4) follows from the Slepian comparison lemma, while the RHS of (4) follows from the chaining argument due to Dudley in the 1960-70s. However, there is a gap between the upper bound and lower bound. This was resolved by Fernique and Talagrand by introducing the concepts of majorizing measures and using the refined generic chaining argument. The final result is as follows. Let

$$\gamma_2(T, d_X) = \inf_{\{T_k\}} \sup_{t \in T} \sum_{k=0}^{\infty} 2^{k/2} d_X(t, T_k),$$

where the infimum is with respect to all admissible sequences. Then

$$c_5\gamma_2(T, d_X) \leq E \sup_{t \in T} X(t) \leq c_6\gamma_2(T, d_X).$$

As for the tail probability estimates, we have the following rough result: Let  $\|X\|_T^2 = \sup_{t \in T} EX(t)^2$ , then

$$P\left(\sup_{t \in T} |X(t)| > x\right) \leq 2e^{-\frac{x^2}{2\|X\|_T^2}}, \quad x > 0.$$

The reader is referred to recent nice graduate textbooks by Vershynin [36] and Wainwright [38] for more information.

## §2 Tracy-Widom Distribution

The Tracy-Widom distribution was first discovered in the study of spectral properties of high dimensional random matrices by two mathematicians C.Tracy and H.Widom from California in 1990s. Their paper *Level-spacing distributions and the Airy kernel*, published in 1994 in *Communications in Mathematical Physics*, was awarded the 2020 Steele prize for seminal contribution to research in analysis/probability theory.

*From a statistician's perspective, the introduction of the Tracy-Widom distributions has been a breakthrough of lasting importance. The Tracy-Widom distributions characterize the limiting distribution of the top eigenvalue in the null hypothesis case of no structure, a challenge for statisticians since the 1950s. In particular, the distribution function  $F(2;s)$  governs the complex-valued data of signal processing.*

We remark that Tracy-Widom distributions usually refer to a family of distributions in literature. However, the present paper shall mainly focus on the Tracy-Widom distribution related to GUE matrix model, and so we simply use the terminology Tracy-Widom distribution without confusion. In Subsection 2.1 we shall start by the definition of the Tracy-Widom distribution, denoted by  $F_2$ , in terms of solution to Painlevé II equation, and then briefly review some basic properties like expectation and variance and tail asymptotics, and finally give the Fredholm determinant expression for  $F_2$ . In Subsection 2.2 we address the issue how the  $F_2$  arises by a quick look at Tracy and Widom's original GUE model. The domain around  $F_2$  is rapidly growing.

### 2.1 Tracy-Widom Distribution

To introduce the Tracy-Widom distribution, we need some notations. Let  $q(x)$  be the unique solution to the following Painlevé II equation

$$\begin{cases} q''(x) = xq(x) + 2q^3(x), \\ q(x) \sim -\frac{1}{2\sqrt{\pi}x^{1/4}}e^{-\frac{2}{3}x^{3/2}}, \quad x \rightarrow \infty \\ q(x) \sim \sqrt{\frac{|x|}{2}}, \quad x \rightarrow -\infty. \end{cases}$$

Now we define the Tracy-Widom distribution

$$F_2(x) = e^{-\int_x^\infty (u-x)q^2(u)du}, \quad x \in \mathbb{R}. \quad (5)$$

It is not clear from the above formula (5) whether  $F_2$  is a distribution function. But it is, namely

$$F_2(-\infty) = 0, \quad F_2(+\infty) = 1, \quad F_2'(x) > 0, \quad x \in \mathbb{R}$$

Figure 1 displays the density curve of  $F_2$ . Obviously, it is no longer symmetric. As  $x \rightarrow \infty$ ,

$$1 - F_2(x) = \frac{1}{32\pi x^{3/2}} e^{-4x^{3/2}/3} (1 + O(x^{-3/2})), \tag{6}$$

$$F_2(-x) = \frac{2^{1/2} e^{\zeta'(-1)}}{x^{1/8}} e^{-x^3/12} \left( 1 + \frac{3}{2^6 x^3} + O(x^{-6}) \right). \tag{7}$$

where  $\zeta$  stands for zeta function. It is easy to see that the left tail probability decays faster than  $e^{-x^2/2}$ , while the right tail probability is slower than  $e^{-x^2/2}$ . We also remark that it takes in contrast to (1) lengthy and laborious effort to obtain the exact constants in the asymptotic expansions (6) and (7).

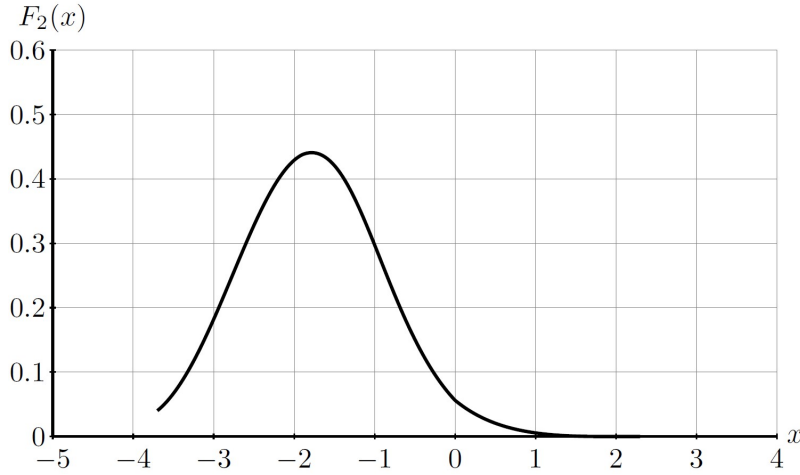


Figure 1: The density curve of  $F_2$ .

Let  $X$  be a random variable distributed as  $F_2$ . Then

$$\mu = EX \approx -1.7710868074, \quad \sigma^2 = Var(X) \approx 0.81319479283,$$

$$Skewness(X) = E\left(\frac{X - \mu}{\sigma}\right)^3 \approx 0.2240842036, \quad Kurtosis(X) \approx 0.09344808.$$

It turns out that  $F_2$  has another Fredholm determinant representation, which is very helpful in manipulation. Let  $Ai(x)$  be a solution to the following equation

$$\begin{cases} Ai''(x) = xAi(x), \\ Ai(x) \rightarrow 0, \quad x \rightarrow \infty. \end{cases}$$

It can be written as

$$Ai(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{1}{3}t^3 + xt\right) dt, \quad x \in \mathbb{R}.$$

The extension to the whole complex plane is given by

$$Ai(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} e^{t^3/3 - zt} dt, \quad z \in \mathbb{C}$$

where  $\mathcal{C}$  denotes the contour in the complex plane consisting of the ray joining  $e^{-i\pi/3}\infty$  to the origin plus the ray joining the origin to  $e^{i\pi/3}\infty$ . The Airy function and associated Airy kernel  $\mathcal{A}(x, y)$ , see (8) below, will repeatedly appear in literature. The reader is referred to the book [37] by Vallée and Soares for commonly used properties. Let  $\mathcal{A}(x, y)$  be the Airy kernel



defined by

$$\begin{aligned} \mathcal{A}(x, y) &= \begin{cases} \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}(y)\text{Ai}'(x)}{x-y}, & x \neq y \\ \text{Ai}'(x)^2 - \text{Ai}(x)\text{Ai}''(x), & x = y \end{cases} \\ &= \int_0^\infty \text{Ai}(x+t)\text{Ai}(y+t)dt. \end{aligned} \tag{8}$$

The Fredholm determinant of  $\mathcal{A}$  in  $L^2(x, \infty)$  is defined by

$$\det(\mathbf{1} - \mathcal{A})|_{L^2(x, \infty)} = \sum_{k=0}^\infty \frac{(-1)^k}{k!} \int_x^\infty \cdots \int_x^\infty \det(\mathcal{A}(x_i, x_j))_{k \times k} dx_1 \cdots dx_k. \tag{9}$$

The sum in the RHS of (9) is convergent by Hadamard's inequality for determinant and the upper bound of  $\text{Ai}(x)$ . The reader is referred to Chapter 3 of the book [1] by Anderson, Guionnet and Zeitouni for a rigorous proof and more information.

It is Tracy and Widom who prove for any  $x \in \mathbb{R}$ ,

$$F_2(x) = \det(\mathbf{1} - \mathcal{A})|_{L^2(x, \infty)}.$$

Indeed, after first taking logarithm and then making second-order derivative at both sides, it reduces to proving

$$\frac{\partial^2}{\partial x^2} \log \det(\mathbf{1} - \mathcal{A})|_{L^2(x, \infty)} = -q^2(x), \quad x \in \mathbb{R}.$$

See [1] for the remaining computation.

## 2.2 How does $F_2$ arise?

Having  $F_2$  and its basic properties, it is curious where  $F_2$  arises from. It turns out that  $F_2$  was first discovered when Tracy and Widom studied the limiting behaviours of spectral gaps of high dimensional Gaussian unitary ensemble (GUE) in their seminal paper [34] in the 1990s. The study of random matrices went back to 1930s in applied multivariate statistics, and reached its golden epoch in physics in 1950-60s, and entered a rapidly developing new era since the discovery of  $F_2$ .

GUE is a prototype of random matrix models. Let  $Z_{ii}, i \geq 1$ , be a sequence of i.i.d. standard normal random variables,  $Z_{ij}, 1 \leq i < j$ , a sequence of i.i.d. complex standard normal random variables. All of these random variables are defined in a common probability space, and are independent of each other. Let  $H_n = (Z_{ij})_{n \times n}$  be an  $n \times n$  complex conjugate random matrix, where  $Z_{ji} = Z_{ij}^*$ . It is easy to see

$$P(H_n) = \frac{2^{n(n-1)/2}}{(2\pi)^{n^2/2}} e^{-\text{tr}H_n^2/2}.$$

Let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be its  $n$  eigenvalues. They are almost surely distinct real numbers since  $H_n$  is a conjugate matrix with continuous entries. The remarkable result due to Weyl is that the distributions of eigenvalues are absolutely continuous with respect to the Lebesgue measure and have joint probability density function

$$p_n(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} \prod_{k=1}^n k!} \prod_{1 \leq i < j \leq n} |x_i - x_j|^2 \prod_{k=1}^n e^{-x_k^2/2}, \quad \mathbf{x} = (x_1, x_2, \dots, x_n). \tag{10}$$

As you notice, the density function  $p_n$  consists of two distinct parts, one is the square of van de Monde determinant,  $\prod_{1 \leq i < j \leq n} |x_i - x_j|^2$  (the power 2 corresponds to the complex conjugate matrix), and the other is the product of independent normal densities,  $\prod_{k=1}^n e^{-x_k^2/2}$ . So the

eigenvalues repel each other and keep a certain gap between.

Let  $h_0(x) = 1$ ,  $h_l(x)$ ,  $l \geq 1$  be a sequence of Hermite orthogonal polynomials, namely

$$h_l(x) = (-1)^l e^{x^2/2} \frac{d^l}{dx^l} e^{-x^2/2},$$

and let  $\varphi_l(x)$  be the normalized Hermite wave function

$$\varphi_l(x) = (2\pi)^{-1/4} (l!)^{-1/2} h_l(x) e^{-x^2/4}.$$

It is easy to check

$$\int_{-\infty}^{\infty} \varphi_l(x) \varphi_m(x) dx = \delta_{l,m}, \quad l, m \geq 0.$$

By a series of transformations of van de Monde determinant, (10) can be written as

$$p_n(\mathbf{x}) = \frac{1}{n!} \det(K_n(x_i, x_j))_{n \times n}, \tag{11}$$

where  $K_n(x, y) = \sum_{l=0}^{n-1} \varphi_l(x) \varphi_l(y)$ .

(11) plays an important role in the study of GUE. In fact, this is a prototype of determinantal point processes coined by Borden and Olshanski around 2000. It is easy to derive from (11) any  $k$ -dimensional marginal density of  $(\lambda_1, \lambda_2, \dots, \lambda_k)$ :

$$p_{n,k}(x_1, x_2, \dots, x_k) = \frac{(n-k)!}{n!} \det(K_n(x_i, x_j))_{k \times k}. \tag{12}$$

In particular, 1-dimensional probability density is

$$p_{n,1}(x) = \frac{1}{n} K_n(x, x), \quad x \in \mathbb{R}. \tag{13}$$

Now one can derive the celebrated Wigner semicircle law for empirical spectral measure from (13) and asymptotic behaviours of Hermite polynomials. For any  $a < b$ ,

$$\frac{1}{n} \#\{1 \leq i \leq n : \lambda_i \in \sqrt{n}(a, b)\} \longrightarrow \int_a^b \rho_{sc}(x) dx, \quad n \rightarrow \infty,$$

where  $\rho_{sc}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{[-2,2]}(x)$ .

The above semicircle law basically tells that the majority of eigenvalues lies in  $[-2\sqrt{n}, 2\sqrt{n}]$  with high probability. In fact, set

$$\lambda_{(1)} = \min_{1 \leq k \leq n} \lambda_k, \quad \lambda_{(n)} = \max_{1 \leq k \leq n} \lambda_k,$$

then by the concentration inequality for a Gaussian random variable and the Borel-Cantelli lemma,

$$\frac{\lambda_{(1)}}{\sqrt{n}} \rightarrow -2, \quad \frac{\lambda_{(n)}}{\sqrt{n}} \rightarrow 2, \quad \text{a.s.}$$

Next, a very interesting and challenging issue: what is the asymptotic distribution of  $\lambda_{(n)}$  around  $2\sqrt{n}$ ? It follows from (12)

$$\begin{aligned} P(\lambda_{(n)} \leq x) &= P\left(\max_{1 \leq k \leq n} \lambda_k \leq x\right) = E \prod_{k=1}^n (1 - \mathbf{1}_{(\lambda_k > x)}) \\ &= \sum_{k=0}^n \frac{(-1)^k}{k!} \int_x^\infty \cdots \int_x^\infty \det(K_n(x_i, x_j))_{k \times k} dx_1 \cdots dx_k \\ &= \det(\mathbf{1} - K_n)_{L^2[x, \infty)}, \end{aligned} \tag{14}$$

where the last term in (14) is a finite series of Fredholm determinant. And again by asymptotic properties of Hermite polynomials,

$$\frac{1}{n^{1/6}} K_n\left(2\sqrt{n} + \frac{1}{n^{1/6}}x, 2\sqrt{n} + \frac{1}{n^{1/6}}y\right) \rightarrow \mathcal{A}(x, y), \quad n \rightarrow \infty,$$

where  $\mathcal{A}(x, y)$  is Airy kernel given by (8).

In summary, we have sketched out the proof that for  $x \in \mathbb{R}$ ,

$$P\left(n^{1/6}(\lambda_{(n)} - 2\sqrt{n}) \leq x\right) \xrightarrow{d} \det(\mathbf{1} - \mathcal{A})_{L^2[x, \infty)}, \quad n \rightarrow \infty.$$

As noticed above, Tracy and Widom [34] only considered a very special case of GUE so that they could apply Weyl's joint probability density representation and nice determinantal structure. How universal is the  $F_2$  distribution?

In fact, soon after the work of Tracy-Widom, Baik, Deift and Johansson [2] proved that  $F_2$  is the asymptotic distribution of the lengths of longest increasing subsequences for a random permutation. Specifically speaking, let  $n \geq 1$ , and let  $\mathcal{S}_n$  be a symmetric permutation group of  $\{1, 2, \dots, n\}$ . Given at random a permutation  $\pi = (\pi(1), \pi(2), \dots, \pi(n)) \in \mathcal{S}_n$ , define the length of longest increasing subsequences of  $\pi$  by

$$L_n(\pi) = \max\{1 \leq k \leq n : \pi(i_1) < \pi(i_2) < \dots < \pi(i_k), i_1 < i_2 < \dots < i_k\}.$$

Logan and Shepp [24], Vershik and Kerov [35] almost at the same time in 1977 proved

$$\frac{L_n}{\sqrt{n}} \xrightarrow{P} 2, \quad n \rightarrow \infty.$$

A longstanding problem is what the fluctuation around  $2\sqrt{n}$  looks like. It turns out

$$\frac{L_n - 2\sqrt{n}}{n^{1/6}} \xrightarrow{d} F_2, \quad n \rightarrow \infty.$$

This is so amazing. As reader might notice, there is no link between GUE and random permutations in terms of finite models, but they have a common limit distribution  $F_2$ . This is an encouraging and enlightening result. There have since then been an intensive research activities around the  $F_2$ , and a dozen of apparently distinct models have actually been found to have  $F_2$  as the asymptotic fluctuation law of a certain suitably scaled statistic. Below is a short (absolutely not exhaustive) list of the models, detailed references are readily available in the literature. The interested reader is referred to a survey [33] for model patterns and main results.

- (1) Complex Wishart matrices;
- (2) Hermitian Wigner matrices;
- (3) Invariant matrix ensembles;
- (4) Directed last passage percolation time in the planar lattice;
- (5) Random Plancherel integers partitions;
- (6) Log-gamma directed polymer models;
- (7) Semi-discrete directed polymer models;
- (8) Totally asymmetric simple exclusion processes;
- (9) Kardar-Parisi-Zhang stochastic dynamic equations.

We remark that a seemingly big difference trivially exists between the above models. In fact, each individual model has its own feature in either algebraic or analytic structure. There is no unified approach working well with all models.

Deift presented a series of unsolved problems in random matrix and integrable systems in the 2006 conference, where *Problem 10* was called a Tracy-Widom central limit theorem, see [3]. We rephrase it as follows.

Let  $\xi_1, \xi_2, \dots$ , be a sequence of i.i.d.r.v.'s, perform an operation  $f$  on them,

$$(\xi_1, \xi_2, \dots) \mapsto (X_1, X_2, \dots).$$

What is the operation  $f$  so that the  $X_n$ 's converges in distribution to  $F_2$ ?

Important progress towards answering this question has been made recently, and independently by Baik and Suidan, Bodineau and Martin, but the full problem remains wide open and very challenging.

### §3 Airy<sub>2</sub> process

The Airy<sub>2</sub> process was first introduced by Prähofer and Spohn [27] in the study of random fields, and has attracted a lot of attentions recently. In Subsection 3.1 we shall define the Airy<sub>2</sub> process through finite dimensional distributions and give two alternative representations. In Subsection 3.2 we give the short range correlation due to Prähofer and Spohn [27], which implies the process is continuous in probability, as well as the long range correlation due to Widom [39], which decays polynomially. In Subsection 3.3 we describe the amazing local Brownian motion behaviors of Airy<sub>2</sub> process. In Subsection 3.4 we explain how the Airy<sub>2</sub> process arise from Dyson’s Brownian motions, while in Subsection 3.5 we introduce an extended Airy point process, whose largest particle locations induce the Airy<sub>2</sub> process.

#### 3.1 Airy<sub>2</sub> process

The precise definition of Airy<sub>2</sub> process goes as follows. The extended Airy kernel is defined by

$$\mathcal{A}^{ext}(t, x; t', x') = \begin{cases} \int_0^\infty e^{-\lambda(t-t')} \text{Ai}(x + \lambda) \text{Ai}(x' + \lambda) d\lambda, & t \geq t' \\ - \int_{-\infty}^0 e^{-\lambda(t-t')} \text{Ai}(x + \lambda) \text{Ai}(x' + \lambda) d\lambda, & t < t' \end{cases} \quad (15)$$

Obviously,  $\mathcal{A}^{ext}(t, x; t, x') = \mathcal{A}(x, x')$ . Note by a result due to Okounkov [25], for  $t > t'$

$$\int_{-\infty}^\infty e^{-\lambda(t-t')} \text{Ai}(x + \lambda) \text{Ai}(y + \lambda) d\lambda = \frac{1}{\sqrt{4\pi(t-t')}} e^{-\frac{(x-y)^2}{4(t-t')} - \frac{t-t'}{2}(x+y) + \frac{(t-t')^3}{12}}. \quad (16)$$

We remark that the RHS of (16) is a form of normal density function for fixed  $t, t'$ , and it follows

$$\begin{aligned} \mathcal{A}^{ext}(t, x; t', x') &= \int_0^\infty e^{-\lambda(t-t')} \text{Ai}(x + \lambda) \text{Ai}(x' + \lambda) d\lambda \\ &\quad - \frac{1}{\sqrt{4\pi(t-t')}} e^{-\frac{(x-x')^2}{4(t-t')} - \frac{t-t'}{2}(x+x') + \frac{(t-t')^3}{12}} \mathbf{1}_{(t < t')}. \end{aligned} \quad (17)$$

The Airy process  $\mathcal{A}_2 = (\mathcal{A}_2(t), t \in \mathbb{R})$  is defined through its finite dimensional distributions, which are given by a Fredholm determinantal formula: given  $x_1, x_2, \dots, x_m$  and  $t_1 < t_2 < \dots < t_m$  in  $\mathbb{R}$ ,

$$P(\mathcal{A}_2(t_1) \leq x_1, \dots, \mathcal{A}_2(t_m) \leq x_m) = \det \left( \mathbf{1} - \chi^{1/2} \mathcal{A}^{ext} \chi^{1/2} \right)_{L^2(\{t_1, \dots, t_m\} \times \mathbb{R})}, \quad (18)$$

where we have counting measure on  $\{t_1, \dots, t_m\}$  and Lebesgue measure on  $\mathbb{R}$ ,  $\chi$  is defined on  $\{t_1, \dots, t_m\} \times \mathbb{R}$  by

$$\chi(t_i, x) = \mathbf{1}_{(x_i, \infty)}(x).$$

We remark that the operator  $\chi^{1/2} \mathcal{A}^{ext} \chi^{1/2}$  in the RHS of (18) is actually a trace class operator on  $L^2(\{t_1, \dots, t_m\} \times \mathbb{R})$ , and the determinant is well-defined and equals

$$\sum_{n=0}^\infty \frac{(-1)^n}{n!} \sum_{i_1, \dots, i_n=1}^m \int_{-\infty}^\infty \dots \int_{-\infty}^\infty \prod_{k=1}^n \chi(t_{i_k}, u_k) \det(\mathcal{A}^{ext}(t_{i_k}, u_k; t_{i_l}, u_l))_{1 \leq k, l \leq n} du_1 \dots du_n.$$

Trivially, for any  $t \in \mathbb{R}$ ,  $\mathcal{A}_2(t) \sim F_2$ .

The above Fredholm determinant allows another expression. Let  $H$  denote the Airy Hamiltonian operator:

$$H = -\frac{\partial^2}{\partial x^2} + x \tag{19}$$

and  $\bar{P}_a$  denote the projection onto the interval  $(-\infty, a]$ :

$$P_a f(x) = \mathbf{1}_{(a, \infty)}(x) f(x), \quad \bar{P}_a = 1 - P_a. \tag{20}$$

Prähofer and Spohn [27] asserted

$$\begin{aligned} & \det(\mathbf{1} - \chi^{1/2} \mathcal{A}^{ext} \chi^{1/2})_{L^2(\{t_1, \dots, t_m\} \times \mathbb{R})} \\ &= \det\left(\mathbf{1} - \mathcal{A} + \bar{P}_{x_1} e^{(t_1-t_2)H} \bar{P}_{x_2} e^{(t_2-t_3)H} \dots \bar{P}_{x_m} e^{(t_m-t_1)H} \mathcal{A}\right)_{L^2(\mathbb{R})}. \end{aligned} \tag{21}$$

Indeed, let

$$\begin{aligned} \mathbf{P} &= \mathbf{P}(x_1, \dots, x_m) = \text{diag}(P_{x_1}, P_{x_2}, \dots, P_{x_m})_{m \times m}, \\ \mathbf{K}^0(x, y) &= \text{diag}(\mathcal{A}(x, y), \mathcal{A}(x, y), \dots, \mathcal{A}(x, y))_{m \times m}, \\ \mathbf{T}^-(x, y) &= \begin{pmatrix} \mathbf{1} & 0 & \dots & 0 & 0 \\ e^{(t_2-t_1)H} & \mathbf{1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ e^{(t_{m-1}-t_1)H} & e^{(t_{m-1}-t_2)H} & \dots & \mathbf{1} & 0 \\ e^{(t_m-t_1)H} & e^{(t_m-t_2)H} & \dots & e^{(t_m-t_{m-1})H} & \mathbf{1} \end{pmatrix}_{m \times m}, \\ \mathbf{T}^+(x, y) &= \begin{pmatrix} 0 & e^{-(t_2-t_1)H} & \dots & e^{-(t_{m-1}-t_1)H} & e^{-(t_m-t_1)H} \\ 0 & 0 & \dots & e^{-(t_{m-1}-t_2)H} & e^{-(t_m-t_2)H} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & e^{-(t_m-t_{m-1})H} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}_{m \times m}. \end{aligned}$$

Define

$$\mathbf{K} = \mathbf{P}(\mathbf{T}^- \mathbf{K}^0 + \mathbf{T}^+ \mathbf{K}^0 - \mathbf{T}^+) \mathbf{P}.$$

Note  $H\mathcal{A} = \mathcal{A}H$ , and

$$e^{(t-s)H} \mathcal{A}(x, y) = \begin{cases} \mathcal{A}^{ext}(s, x; t, y), & s < t \\ \mathcal{A}^{ext}(s, x; t, y) - \mathbf{1}, & s \geq t. \end{cases}$$

Then it follows

$$\det(\mathbf{1} - \chi^{1/2} \mathcal{A}^{ext} \chi^{1/2})_{L^2(\{t_1, \dots, t_m\} \times \mathbb{R})} = \det(\mathbf{1} - \mathbf{K})_{L^2(\mathbb{R})}.$$

On the other hand, note  $\mathbf{T}^+$  is an upper triangular matrix, and

$$(\mathbf{P}\mathbf{T}^+)^m = \mathbf{0}, \quad \text{tr}((\mathbf{P}\mathbf{T}^+)^k) = 0, \quad 1 \leq k \leq m-1,$$

and hence

$$(\mathbf{1} + \mathbf{P}\mathbf{T}^+)^{-1} = \mathbf{1} - \mathbf{P}\mathbf{T}^+ + (\mathbf{P}\mathbf{T}^+)^2 + \dots + (-1)^{m-1} (\mathbf{P}\mathbf{T}^+)^{m-1}.$$

We have

$$\begin{aligned} \det(\mathbf{1} - \mathbf{K})_{L^2(\mathbb{R})} &= \det(\mathbf{1} + \mathbf{P}\mathbf{T}^+ \mathbf{P} - \mathbf{P}(\mathbf{T}^- + \mathbf{T}^+) \mathbf{K}^0 \mathbf{P})_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} + \mathbf{P}\mathbf{T}^+)_{L^2(\mathbb{R})} \cdot \det(\mathbf{1} - (\mathbf{1} + \mathbf{P}\mathbf{T}^+)^{-1} \mathbf{P}(\mathbf{T}^- + \mathbf{T}^+) \mathbf{K}^0)_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - (\mathbf{1} + \mathbf{P}\mathbf{T}^+)^{-1} \mathbf{P}(\mathbf{T}^- + \mathbf{T}^+) \mathbf{K}^0)_{L^2(\mathbb{R})}, \end{aligned}$$

where we have used the fact that  $\det(\mathbf{1} + \mathbf{P}\mathbf{T}^+)_{L^2(\mathbb{R})} = 1$ .

Furthermore,

$$(\mathbf{T}^- + \mathbf{T}^+)(\mathbf{1} + \mathbf{T}^+)^{-1} = \begin{pmatrix} \mathbf{1} & 0 & \cdots & 0 & 0 \\ e^{(t_2-t_1)H} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ e^{(t_{m-1}-t_1)H} & 0 & \cdots & 0 & 0 \\ e^{(t_m-t_1)H} & 0 & \cdots & 0 & 0 \end{pmatrix}_{m \times m}$$

and

$$(\mathbf{1} + \mathbf{T}^+)(\mathbf{1} + \mathbf{PT}^+)^{-1} =$$

Then letting

$$\mathbf{B} = (\mathbf{1} + \mathbf{T}^+)(\mathbf{1} + \mathbf{PT}^+)^{-1} \mathbf{P}(\mathbf{T}^- + \mathbf{T}^+)(\mathbf{1} + \mathbf{T}^+)^{-1} \mathbf{K}^0,$$

it follows

$$(\mathbf{B})_{11} = \mathcal{A} - \bar{P}_{x_1} e^{(t_1-t_2)H} \bar{P}_{x_2} e^{(t_2-t_3)H} \cdots \bar{P}_{x_m} e^{(t_m-t_1)H} \mathcal{A}$$

and

$$(\mathbf{B})_{ij} = 0, \quad 1 \leq i \leq m, \quad 2 \leq j \leq m.$$

In combination, we have

$$\begin{aligned} & \det(\mathbf{1} - \mathbf{K})_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - (\mathbf{1} + \mathbf{PT}^+)^{-1} \mathbf{P}(\mathbf{T}^- + \mathbf{T}^+) \mathbf{K}^0)_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - \mathbf{B})_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - (\mathbf{B})_{11})_{L^2(\mathbb{R})}, \end{aligned}$$

which is as desired.

Both (18) and (21) are very useful in the exploration of Airy<sub>2</sub> process. The operator in the formula (18) acts on the extended space  $L^2(\{t_1, \dots, t_m\} \times \mathbb{R})$ , while the Fredholm determinant in (21) is computed on the Hilbert space  $L^2(\mathbb{R})$ . The latter formula avoids a big difficulty when the number  $m$  of times  $t_i$  goes to infinity. Another advantage of (21) is that it makes apparent that  $\mathcal{A}_2$  is a stationary process since the time increments  $t_i - t_{i+1}$  appear explicitly in the joint distribution.

In summary, the Airy<sub>2</sub> process  $(\mathcal{A}_2(t), t \in \mathbb{R})$  is a stationary stochastic process whose f.d.d.'s are given through a Fredholm determinant with extended Airy kernel  $\mathcal{A}^{ext}$  as in (18), and 1-dimensional distribution is  $F_2$ .

### 3.2 Short and Long Range Correlations

In this subsection we look at the short and long range correlation for  $\mathcal{A}_2(t)$ . Since it is stationary, we focus on  $Cov(\mathcal{A}_2(0), \mathcal{A}_2(t))$  where  $t > 0$ .

The short range correlation was first investigated by Prähofer and Spohn [27], in which the Airy<sub>2</sub> process was introduced. Their result is as follows

$$Cov(\mathcal{A}_2(0), \mathcal{A}_2(t)) = Var(\mathcal{A}_2(0)) - 2t + O(t^2), \quad t \rightarrow 0$$

In particular, we have

$$E|\mathcal{A}_2(t) - \mathcal{A}_2(0)|^2 = 2t + O(t^2), \quad t \rightarrow 0 \tag{22}$$

which immediately implies that  $\mathcal{A}_2(t)$  is continuous in probability. However, (22) does not directly imply the a.s. continuity of sample paths of  $\mathcal{A}_2(t)$ , see Section 4.

The long range correlation was first studied by Widom [39], who got the following result:

$$Cov(\mathcal{A}_2(0), \mathcal{A}_2(t)) = \frac{1}{t^2} + \frac{2E\mathcal{A}_2(0)}{t^4} + O\left(\frac{1}{t^5}\right), \quad t \rightarrow \infty \tag{23}$$

Later on, Shinault and Tracy [32] improved this result to any finite orders.

To see (23), let  $f_2(x) = F_2'(x)$  be the probability density function of GUE Tracy-Widom distribution, and define

$$C_2(x_1, x_2) = f_2(x_1)f_2(x_2)$$

and

$$C_4(x_1, x_2) = (x_1 + x_2)f_2(x_1)f_2(x_2) + \frac{1}{2}f_2'(x_1)f_2'(x_2).$$

By the integration formula by parts and noting  $F_2$  and  $f_2$  exponentially decays at  $\pm\infty$ , it is easy to check

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1x_2dC_2(x_1, x_2) = 1$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1x_2dC_4(x_1, x_2) = 2E\mathcal{A}_2(0).$$

So it is sufficient to show as  $t \rightarrow \infty$

$$\begin{aligned} &P(\mathcal{A}_2(0) \leq x_1, \mathcal{A}_2(t) \leq x_2) - P(\mathcal{A}_2(0) \leq x_1)P(\mathcal{A}_2(t) \leq x_2) \\ &= \frac{C_2(x_1, x_2)}{t^2} + \frac{C_4(x_1, x_2)}{t^4} + O\left(\frac{1}{t^5}\right). \end{aligned} \tag{24}$$

In turn, let

$$\begin{aligned} P &= \text{diag}(P_{x_1}, P_{x_2}), \quad K_1(x, y) = \text{diag}(\mathcal{A}(x, y), \mathcal{A}(x, y)) \\ K_2(x, y) &= \begin{pmatrix} 0 & \mathcal{A}^{ext}(0, x; t, y) \\ \mathcal{A}^{ext}(t, x; 0, y) & 0 \end{pmatrix}_{2 \times 2}, \end{aligned}$$

and denote  $K(x, y) = K_1(x, y) + K_2(x, y)$ . Then

$$\begin{aligned} P(\mathcal{A}_2(0) \leq x_1, \mathcal{A}_2(t) \leq x_2) &= \det(\mathbf{1} - PKP)_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - PK_1P - PK_2P)_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - PK_1P)_{L^2(\mathbb{R})} \det(\mathbf{1} - (\mathbf{1} - PK_1P)^{-1}PK_2P)_{L^2(\mathbb{R})}. \end{aligned}$$

Note

$$\det(\mathbf{1} - PK_1P)_{L^2(\mathbb{R})} = P(\mathcal{A}_2(0) \leq x_1)P(\mathcal{A}_2(t) \leq x_2).$$

So, we turn to analyze the determinant  $\det(\mathbf{1} - (\mathbf{1} - PK_1P)^{-1}PK_2P)_{L^2(\mathbb{R})}$ . Repeatedly applying the integration by part formula, we get

$$\begin{aligned} \mathcal{A}^{ext}(0, x; t, y) &= \int_0^{\infty} e^{t\lambda} \text{Ai}(x + \lambda)\text{Ai}(y + \lambda)d\lambda \\ &= \frac{1}{t} \text{Ai}(x)\text{Ai}(y) - \frac{1}{t^2} (\text{Ai}'(x)\text{Ai}(y) + \text{Ai}(x)\text{Ai}'(y)) \\ &\quad + \frac{1}{t^2} \int_0^{\infty} e^{t\lambda} \frac{d^2}{d\lambda^2} (\text{Ai}(x + \lambda)\text{Ai}(y + \lambda))d\lambda \end{aligned}$$

and

$$\begin{aligned} \mathcal{A}^{ext}(t, x; 0, y) &= - \int_{-\infty}^0 e^{-t\lambda} \text{Ai}(x + \lambda) \text{Ai}(y + \lambda) d\lambda \\ &= \frac{1}{t} \text{Ai}(x) \text{Ai}(y) - \frac{1}{t^2} (\text{Ai}'(x) \text{Ai}(y) + \text{Ai}(x) \text{Ai}'(y)) \\ &\quad + \frac{1}{t^2} \int_{-\infty}^0 e^{-t\lambda} \frac{d^2}{d\lambda^2} (\text{Ai}(x + \lambda) \text{Ai}(y + \lambda)) d\lambda. \end{aligned}$$

For clarity of writing, we introduce the following notations. Denote  $f \otimes g(x, y) = f(x)g(y)$ , and let

$$\begin{aligned} \mathbf{T}_{12}(x, y) &= \frac{1}{t} (\mathbf{1} - P_{x_1} \mathcal{A} P_{x_1})^{-1} P_{x_1} \text{Ai} \otimes \text{Ai} P_{x_2}(x, y) \\ &\quad + \frac{1}{t^2} (\mathbf{1} - P_{x_1} \mathcal{A} P_{x_1})^{-1} P_{x_1} (\text{Ai}' \otimes \text{Ai} + \text{Ai} \otimes \text{Ai}') P_{x_2}(x, y) + O\left(\frac{1}{t^3}\right) \end{aligned}$$

and

$$\begin{aligned} \mathbf{T}_{21}(x, y) &= \frac{1}{t} (\mathbf{1} - P_{x_2} \mathcal{A} P_{x_2})^{-1} P_{x_2} \text{Ai} \otimes \text{Ai} P_{x_1}(x, y) \\ &\quad + \frac{1}{t^2} (\mathbf{1} - P_{x_2} \mathcal{A} P_{x_2})^{-1} P_{x_2} (\text{Ai}' \otimes \text{Ai} + \text{Ai} \otimes \text{Ai}') P_{x_1}(x, y) + O\left(\frac{1}{t^3}\right). \end{aligned}$$

Then

$$\mathbf{1} - (\mathbf{1} - \text{PK}_1 \text{P})^{-1} \text{PK}_2 \text{P} = \begin{pmatrix} \mathbf{1} & -\mathbf{T}_{12} \\ -\mathbf{T}_{21} & \mathbf{1} \end{pmatrix}$$

and so

$$\begin{aligned} &\det(\mathbf{1} - (\mathbf{1} - \text{PK}_1 \text{P})^{-1} \text{PK}_2 \text{P})_{L^2(\mathbb{R})} \\ &= \det(\mathbf{1} - \mathbf{T}_{12} \mathbf{T}_{21})_{L^2(\mathbb{R})} \\ &= e^{tr \ln(\mathbf{1} - \mathbf{T}_{12} \mathbf{T}_{21})} \\ &= 1 - tr(\mathbf{T}_{12} \mathbf{T}_{21}) - \frac{1}{2} (tr(\mathbf{T}_{12} \mathbf{T}_{21})^2 - (tr(\mathbf{T}_{12} \mathbf{T}_{21}))^2) + \dots \end{aligned}$$

It remains to evaluate traces  $tr(\mathbf{T}_{12} \mathbf{T}_{21})$  and  $tr(\mathbf{T}_{12} \mathbf{T}_{21})^2$ . To this end, we use the fact  $(f \otimes g)(h \otimes k) = \langle g, h \rangle_{L_2} f \otimes k$  whose trace is  $\langle g, h \rangle_{L_2} \langle f, k \rangle_{L_2}$ . Shinault and Tracy [32] found that

$$tr(\mathbf{T}_{12} \mathbf{T}_{21}) = \frac{f_2(x_1) f_2(x_2)}{t^2}$$

and

$$tr(\mathbf{T}_{12} \mathbf{T}_{21})^2 = 2 \frac{f_2(x_1) f_2(x_2)}{t^4}.$$

In combination, we have proved (24), and so (23).

Since  $(\mathcal{A}_2(t), t \in \mathbb{R})$  is a stationary asymptotically uncorrelated stochastic process, then it is ergodic in the sense of mean. Namely,

$$\frac{1}{2T} \int_{-T}^T \mathcal{A}_2(t) dt \xrightarrow{L^2} E \mathcal{A}_2(0) \approx -1.7710868074, \quad T \rightarrow \infty.$$

As a by-product of (23),  $(\mathcal{A}_2(t), t \in \mathbb{R})$  is *not* Markovian. In fact, for a stationary Markov process, its correlation must exponentially decay if it decay.



### 3.3 Local Brownian Motion Phenomena

As observed above, a normal distribution usually arises as a limiting distribution of normalized sums for i.i.d.r.v.'s with finite second-order moments, while the  $F_2$  is typically used to described a new class of extremal phenomena like the largest eigenvalue of Hermite matrices with i.i.d. random entries under certain moment conditions. Brownian motion is a Gaussian process with stationary independent increments (*Markovian*), while Airy process is a stationary *non-Markovian* process with 1-dimensional distribution  $F_2$ . At a first glance, they are two totally different objects appearing in different settings. However, as the following result shows, the  $\text{Airy}_2$  process behaves locally like a Brownian motion. It is so amazing. How does a normal distribution appear in the extremal phenomena?

Define for  $\varepsilon > 0$

$$Z_\varepsilon(t) = \frac{\mathcal{A}_2(t\varepsilon) - \mathcal{A}_2(0)}{\sqrt{\varepsilon}}, \quad t \geq 0$$

and consider the process  $\mathbf{Z}_\varepsilon = (Z_\varepsilon(t), t \geq 0)$ . The so-called local Brownian motion property mathematically reads as follows:

$$\mathbf{Z}_\varepsilon \Rightarrow \mathbf{B}, \quad \varepsilon \rightarrow 0 \tag{25}$$

The proof of (25) consists of two parts as usual: finite dimensional distributions weakly converges, and  $\mathbf{Z}_\varepsilon$  is uniformly tight in  $C[0, \infty)$ . The uniform tightness of  $\mathbf{Z}_\varepsilon$  was first established by Pimentel [26] using the comparison theorem on last passage percolation models. In fact, it now readily follows from the tail probability estimate (40) on modulus of continuity for  $\text{Airy}_2$  process.

In the following paragraphs we will briefly outline the basic ideas for finite dimensional convergence due to Hägg [17]. It suffices to show for any  $m \geq 1$  and  $t_1 < t_2 < \dots < t_m$ ,

$$(Z_\varepsilon(t_1), Z_\varepsilon(t_2), \dots, Z_\varepsilon(t_m)) \xrightarrow{d} (B(t_1), B(t_2), \dots, B(t_m)), \quad \varepsilon \rightarrow 0.$$

Namely, for any  $z_1, z_2, \dots, z_m \in \mathbb{R}$

$$\begin{aligned} &P(Z_\varepsilon(t_1) = z_1, Z_\varepsilon(t_2) = z_2, \dots, Z_\varepsilon(t_m) = z_m) \\ &\rightarrow P(B(t_1) = z_1, B(t_2) = z_2, \dots, B(t_m) = z_m) \\ &= \prod_{i=1}^m \frac{1}{\sqrt{2\pi s_i}} \exp\left(-\frac{x_i^2}{2s_i}\right), \end{aligned}$$

where  $s_1 = t_1, x_1 = z_1, s_i = t_i - t_{i-1}, x_i = z_i - z_{i-1}, 2 \leq i \leq m$ .

Note

$$\begin{aligned} &P(Z_\varepsilon(t_1) = z_1, Z_\varepsilon(t_2) = z_2, \dots, Z_\varepsilon(t_m) = z_m | \mathcal{A}_2(0) = x_0) \\ &= \lim_{\delta_0, \dots, \delta_m \rightarrow 0} \frac{P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0], \mathcal{A}_2(\varepsilon t_i) \in (y_i - \sqrt{2\varepsilon}\delta_i, y_i], 1 \leq i \leq m)}{(\prod_{i=1}^m \delta_i) P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0])} \end{aligned}$$

where  $y_i = x_0 + \sqrt{2\varepsilon}z_i$ .

Hence, it is sufficient to prove for any  $x_0, x_1, \dots, x_m \in \mathbb{R}$  and  $s_1, s_2, \dots, s_m > 0$ ,

$$\begin{aligned} &\lim_{\delta_0, \dots, \delta_m \rightarrow 0} \frac{P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0], \mathcal{A}_2(\varepsilon t_i) \in (y_i - \sqrt{2\varepsilon}\delta_i, y_i], 1 \leq i \leq m)}{(\prod_{i=1}^m \delta_i) P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0])} \\ &\rightarrow \prod_{i=1}^m \frac{1}{\sqrt{2\pi s_i}} \exp\left(-\frac{x_i^2}{2s_i}\right), \quad \varepsilon \rightarrow 0 \end{aligned} \tag{26}$$

where  $y_i = x_0 + \sqrt{2\varepsilon}x_i, t_1 = s_1, t_i = t_{i-1} + s_i, 2 \leq i \leq m$ .

Let us start with the probability  $P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0])$  in the denominator of (26). Denote  $N_0 = \sharp(x_0 - \delta_0, x_0]$  and  $\bar{N}_0 = \sharp[x_0, \infty)$  and  $\tilde{N}_0 = N_0 - \mathbf{1}_{(N_0 \geq 1)}$ . Then

$$\begin{aligned} P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0]) &= E\mathbf{1}_{(N_0 \geq 1)}\mathbf{1}_{(\bar{N}_0=0)} \\ &= EN_0\mathbf{1}_{(\bar{N}_0=0)} - E\tilde{N}_0\mathbf{1}_{(\bar{N}_0=0)}. \end{aligned}$$

Note an elementary calculus gives

$$\begin{aligned} \mathbf{1}_{(\bar{N}_0=0)} &= \lim_{\beta \rightarrow \infty} e^{-\beta \bar{N}_0} \\ &= \lim_{\beta \rightarrow \infty} (1 - (1 - e^{-\beta}))^{\bar{N}_0} \\ &= \lim_{\beta \rightarrow \infty} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} N_0(\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1))(1 - e^{-\beta})^k \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} N_0(\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1)) \lim_{\beta \rightarrow \infty} (1 - e^{-\beta})^k \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \bar{N}_0(\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1)). \end{aligned}$$

Thus it follows

$$EN_0\mathbf{1}_{(\bar{N}_0=0)} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} EN_0\bar{N}_0(\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1)). \tag{27}$$

On the other hand,

$$\begin{aligned} &EN_0\bar{N}_0(\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1)) \\ &= \int_{x_0 - \delta_0}^{x_0} \int_{[x_0, \infty)^k} \rho_{k+1}(v_0, u_1, \dots, u_k) dv_0 du_1 \cdots du_k. \end{aligned} \tag{28}$$

Here we used the relation (38) and the correlation formula (37) for extended Airy point process in Subsection 3.5, the time line was  $\mathbb{R}_0$ . In the sequel we will also use such an approach to moment estimates, the time lines will be clear from the context. For brevity we omit the subscript in  $\rho$ .

Hence we have

$$\lim_{\delta_0 \rightarrow 0} \frac{1}{\delta_0} EN_0\mathbf{1}_{(\bar{N}_0=0)} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{[x_0, \infty)^k} \rho_{k+1}(x_0, u_1, \dots, u_k) du_1 \cdots du_k.$$

Similarly, since  $\tilde{N}_0 \leq N_0(N_0 - 1)$ , then it follows

$$\begin{aligned} \lim_{\delta_0 \rightarrow 0} \frac{1}{\delta_0} E\tilde{N}_0\mathbf{1}_{(\bar{N}_0=0)} &\leq \lim_{\delta_0 \rightarrow 0} \frac{1}{\delta_0} EN_0(N_0 - 1)\mathbf{1}_{(\bar{N}_0=0)} \\ &= 0. \end{aligned}$$

Let us turn to the probability in the numerator of (26). Denote  $N_i = \sharp(y_i - \delta_i, y_i]$  and  $\bar{N}_i = \sharp[y_i, \infty)$  and  $\tilde{N}_i = N_i - \mathbf{1}_{(N_i \geq 1)}$ . Then

$$\begin{aligned} &P(\mathcal{A}_2(0) \in (x_0 - \delta_0, x_0], \mathcal{A}_2(\varepsilon t_i) \in (y_i - \delta_i, y_i], 1 \leq i \leq m) \\ &= P(N_0 \geq 1, \bar{N}_0 = 0, N_1 \geq 1, \bar{N}_1 = 0, \dots, N_m \geq 1, \bar{N}_m = 0) \\ &= P(N_0 \geq 1, N_1 \geq 1, \dots, N_m \geq 1, \bar{N}_0 = 0) \\ &\quad - P(N_0 \geq 1, N_1 \geq 1, \dots, N_m \geq 1, \bar{N}_0 = 0, \cup_{i=1}^m \{\bar{N}_i \neq 0\}). \end{aligned}$$

The major term is

$$\begin{aligned}
 & P(N_0 \geq 1, N_1 \geq 1, \dots, N_m \geq 1, \bar{N}_0 = 0) \\
 &= E \prod_{i=0}^m \mathbf{1}_{(N_i \geq 1)} \mathbf{1}_{(\bar{N}_0=0)} \\
 &= E \prod_{i=0}^m (N_i - \tilde{N}_i) \mathbf{1}_{(\bar{N}_0=0)} \\
 &= E \left( \prod_{i=0}^m N_i \right) \mathbf{1}_{(\bar{N}_0=0)} + EU(\tilde{N}_0, \tilde{N}_1, \dots, \tilde{N}_m) \mathbf{1}_{(\bar{N}_0=0)}, \tag{29}
 \end{aligned}$$

where  $U(\tilde{N}_0, \tilde{N}_1, \dots, \tilde{N}_m) = \prod_{i=0}^m (N_i - \tilde{N}_i) - \prod_{i=0}^m N_i$ . And the term  $E(\prod_{i=0}^m N_i) \mathbf{1}_{(\bar{N}_0=0)}$  will play a major role in estimating (29).

Analogously to (27) and (28), it follows

$$E \left( \prod_{i=0}^m N_i \right) \mathbf{1}_{(\bar{N}_0=0)} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} E \left( \prod_{i=0}^m N_i \right) \bar{N}_0 (\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1))$$

and

$$\begin{aligned}
 & E \left( \prod_{i=0}^m N_i \right) \bar{N}_0 (\bar{N}_0 - 1) \cdots (\bar{N}_0 - (k - 1)) \\
 &= \int_{x_0 - \delta_0}^{x_0} \int_{y_1 - \sqrt{2\varepsilon}\delta_1}^{y_1} \cdots \int_{y_m - \sqrt{2\varepsilon}\delta_m}^{y_m} \int_{[x_0, \infty)^k} \rho(v_0, v_1, \dots, v_m; u_1, \dots, u_k) \prod_{i=0}^m dv_i \prod_{i=1}^k du_i.
 \end{aligned}$$

It follows as  $\delta_0, \dots, \delta_m \rightarrow 0$

$$\begin{aligned}
 & \frac{1}{\prod_{i=0}^m \delta_i} \int_{x_0 - \delta_0}^{x_0} \int_{y_1 - \sqrt{2\varepsilon}\delta_1}^{y_1} \cdots \int_{y_m - \sqrt{2\varepsilon}\delta_m}^{y_m} \int_{[x_0, \infty)^k} \rho(v_0, v_1, \dots, v_m; u_1, \dots, u_k) \prod_{i=0}^m dv_i \prod_{i=1}^k du_i \\
 & \rightarrow (\sqrt{2\varepsilon})^m \int_{[x_0, \infty)^k} \rho(x_0, y_1, \dots, y_m; u_1, \dots, u_k) \prod_{i=1}^k du_i.
 \end{aligned}$$

In combination, we have

$$\begin{aligned}
 & \lim_{\delta_0, \dots, \delta_m \rightarrow 0} \frac{1}{\prod_{i=0}^m \delta_i} E \left( \prod_{i=0}^m N_i \right) \mathbf{1}_{(\bar{N}_0=0)} \\
 &= (\sqrt{2\varepsilon})^m \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{[x_0, \infty)^k} \rho(x_0, y_1, \dots, y_m; u_1, \dots, u_k) \prod_{i=1}^k du_i.
 \end{aligned}$$

It remains to prove

$$\lim_{\varepsilon \rightarrow 0} (\sqrt{2\varepsilon})^m \rho(x_0, y_1, \dots, y_m; u_1, \dots, u_k) = \rho(x_0; u_1, \dots, u_k) \prod_{i=1}^m \frac{1}{\sqrt{2\pi s_i}} \exp\left(-\frac{x_i^2}{2s_i}\right). \tag{30}$$

This is quite technical and lengthy. For an illustration, we only consider a simple case, that is  $m = 1$  and  $k = 1$ . Note

$$\rho(x_0, y_1; u_1) = \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & \mathcal{A}^{ext}(0, x_0; \varepsilon t_1, y_1) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & \mathcal{A}(y_1, y_1) & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & \mathcal{A}^{ext}(0, u_1; \varepsilon t_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix},$$

where by (17)

$$\begin{aligned} \mathcal{A}^{ext}(0, x_0; \varepsilon t_1, y_1) &= - \int_{-\infty}^0 e^{\varepsilon t_1 \lambda} \text{Ai}(x_0 + \lambda) \text{Ai}(y_1 + \lambda) d\lambda \\ &= \int_0^{\infty} e^{\varepsilon t_1 \lambda} \text{Ai}(x_0 + \lambda) \text{Ai}(y_1 + \lambda) d\lambda \\ &\quad - \frac{1}{\sqrt{4\pi\varepsilon t_1}} \exp\left(-\frac{(y_1 - x_0)^2}{4\varepsilon t_1} - \frac{\varepsilon t_1}{2}(x_0 + y_1) + \frac{(\varepsilon t_1)^3}{12}\right) \\ &=: \tilde{\mathcal{A}}^{ext}(0, x_0; \varepsilon t_1, y_1) - \phi_{0, \varepsilon t_1}(x_0, y_1). \end{aligned}$$

Hence a simple algebra yields

$$\begin{aligned} \rho(x_0, y_1; u_1) &= \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & \mathcal{A}^{ext}(0, x_0; \varepsilon t_1, y_1) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & \mathcal{A}(y_1, y_1) & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & \mathcal{A}^{ext}(0, u_1; \varepsilon t_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix} \\ &= \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & \tilde{\mathcal{A}}^{ext}(0, x_0; \varepsilon t_1, y_1) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & \mathcal{A}(y_1, y_1) & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & \tilde{\mathcal{A}}^{ext}(0, u_1; \varepsilon t_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix} \\ &\quad - \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & -\phi_{0, \varepsilon t_1}(x_0, y_1) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & 0 & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & -\phi_{0, \varepsilon t_1}(u_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix}. \end{aligned} \tag{31}$$

We need to compute the limit of each entry in the matrix of RHS of (31) as  $\varepsilon \rightarrow 0$ . Note  $y_1 = x_0 + \sqrt{2\varepsilon}x_1$ , and so it follows

$$\begin{aligned} \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) &= \int_0^{\infty} e^{-\varepsilon t_1 \lambda} \text{Ai}(x_0 + \lambda) \text{Ai}(y_1 + \lambda) d\lambda \\ &\rightarrow \mathcal{A}(x_0, x_0), \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) &= \int_0^{\infty} e^{-\varepsilon t_1 \lambda} \text{Ai}(u_1 + \lambda) \text{Ai}(y_1 + \lambda) d\lambda \\ &\rightarrow \mathcal{A}(u_1, x_0). \\ \mathcal{A}^{ext}(0, u_1; \varepsilon t_1, y_1) &= \int_0^{\infty} e^{\varepsilon t_1 \lambda} \text{Ai}(u_1 + \lambda) \text{Ai}(y_1 + \lambda) d\lambda \\ &\quad - \frac{1}{\sqrt{4\pi\varepsilon t_1}} e^{-\frac{(y_1 - u_1)^2}{4\varepsilon t_1} - \frac{\varepsilon t_1}{2}(u_1 + y_1) + \frac{(\varepsilon t_1)^3}{12}} \\ &\rightarrow \mathcal{A}(u_1, x_0), \end{aligned}$$

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \tilde{\mathcal{A}}^{ext}(0, x_0; \varepsilon t_1, y_1) &= \lim_{\varepsilon \rightarrow 0} \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) = \mathcal{A}(x_0, x_0), \\ \lim_{\varepsilon \rightarrow 0} \tilde{\mathcal{A}}^{ext}(0, u_1; \varepsilon t_1, y_1) &= \lim_{\varepsilon \rightarrow 0} \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) = \mathcal{A}(u_1, x_0) \end{aligned}$$

and

$$\lim_{\varepsilon \rightarrow 0} \mathcal{A}(y_1, y_1) = \mathcal{A}(x_0, x_0).$$

Also,  $t_1 = s_1$ , and  $u_1 > x_0$ , so it follows

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \sqrt{2\varepsilon} \phi_{0, \varepsilon t_1}(x_0, y_1) &= \frac{1}{\sqrt{2s_1}} \exp\left(-\frac{x_1^2}{2s_1}\right), \\ \lim_{\varepsilon \rightarrow 0} \phi_{0, \varepsilon t_1}(u_1, y_1) &= 0. \end{aligned}$$

In combination, we have

$$\lim_{\varepsilon \rightarrow 0} \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & \tilde{\mathcal{A}}^{ext}(0, x_0; \varepsilon t_1, y_1) & \mathcal{A}(x_0, u) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & \mathcal{A}(y_1, y_1) & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & \tilde{\mathcal{A}}^{ext}(0, u_1; \varepsilon t_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix} = 0$$

and

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \sqrt{2\varepsilon} \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & -\phi_{0, \varepsilon t_1}(x_0, y_1) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, x_0) & 0 & \mathcal{A}^{ext}(\varepsilon t_1, y_1; 0, u_1) \\ \mathcal{A}(u_1, x_0) & -\phi_{0, \varepsilon t_1}(u_1, y_1) & \mathcal{A}(u_1, u_1) \end{pmatrix} \\ &= \frac{1}{\sqrt{2s_1}} \exp\left(-\frac{x_1^2}{2s_1}\right) \det \begin{pmatrix} \mathcal{A}(x_0, x_0) & \mathcal{A}(x_0, u_1) \\ \mathcal{A}(u_1, x_0) & \mathcal{A}(u_1, u_1) \end{pmatrix}. \end{aligned}$$

Thus we have completed the proof of (30) in the case of  $m = k = 1$ .

To conclude this subsection, we mention a simple result about Brownian bridge. Let  $B^0(t) = B(t) - tB(1)$ ,  $0 \leq t \leq 1$ . Trivially,  $B^0(0) = B^0(1) = 0$ . Then as  $\varepsilon \rightarrow 0$

$$\frac{1}{\sqrt{\varepsilon}}(B^0(\varepsilon t), 0 \leq t \leq 1) \Rightarrow (B(t), 0 \leq t \leq 1).$$

### 3.4 How does Airy<sub>2</sub> process arise?

The Airy<sub>2</sub> process  $\mathcal{A}_2(t)$  was first introduced by Prähofer and Spohn [27] in exploring polynuclear growth processes. It turns out that  $\mathcal{A}_2(t)$  has since been found to appear in a dozen of apparently distinct fields. Only recently, it is proved that the directed landscape has KPZ fixed point as its marginal, while the KPZ fixed point has Airy processes as its marginals. We shall give a quick introduction from the eigenvalues of stationary random matrix-valued processes in this subsection.

Let  $\mathbf{B}_{ij} = (B_{ij}(t), t \geq 0)$ ,  $i, j \geq 1$  be an array of i.i.d. standard real Brownian motions, let  $(\tilde{\mathbf{B}}_{ij}, i, j \geq 1)$  an independent copy of  $(\mathbf{B}_{ij}, i, j \geq 1)$ . Define

$$A_{ij}(t) = B_{ij}(t) + i\tilde{B}_{ij}(t)$$

and then

$$H_{ij}(t) = \frac{1}{2}(A_{ij}(t) + A_{ij}^*(t)).$$

The Hermitian matrix-valued Brownian motion is defined by

$$\mathbf{H}_n(t) = (H_{ij}(t))_{n \times n}, \quad t \geq 0 \tag{32}$$

In particular,  $\mathbf{H}_n(1)$  is a well-known GUE. What we are concerned with in the following paragraphs are stationary Ornstein-Uhlenbeck matrix valued random processes defined by SDE

$$d\mathbf{M}_n(t) = -\mathbf{M}_n(t)dt + d\mathbf{H}_n(t), \quad t > 0$$

The stationary distribution associated with  $\mathbf{M}_n(t)$  is GUE. For any finite times  $t_1 < t_2 < \dots < t_m$ , the joint probability density function for  $(\mathbf{M}_n(t_1), \mathbf{M}_n(t_2), \dots, \mathbf{M}_n(t_m))$  is

$$\begin{aligned} & P(\mathbf{M}_n(t_1) = \mathbf{A}_1, \mathbf{M}_n(t_2) = \mathbf{A}_2, \dots, \mathbf{M}_n(t_m) = \mathbf{A}_m) \\ &= P(\mathbf{M}_n(t_1) = \mathbf{A}_1) \prod_{i=2}^m P(\mathbf{M}_n(t_i) = \mathbf{A}_i | \mathbf{M}_n(t_{i-1}) = \mathbf{A}_{i-1}) \\ &= \frac{1}{Z_{n,m}} e^{-\frac{1}{2}tr(\mathbf{A}_1^2)} \prod_{i=2}^m \exp\left(-\frac{tr((\mathbf{A}_i - \rho_{t_{i-1}, t_i} \mathbf{A}_{i-1})^2)}{2(1 - \rho_{t_{i-1}, t_i}^2)}\right), \end{aligned}$$

where  $\rho_{s,t} = e^{-(t-s)/2}$ ,  $Z_{n,m}$  is a normalized constant.

Let  $\lambda_n(t) = (\lambda_{n,1}(t), \lambda_{n,2}(t), \dots, \lambda_{n,n}(t))$  be the eigenvalues of  $\mathbf{M}_n(t)$  in an increasing order. Then as a result of Itô's formula,  $\lambda_n(t)$  is a diffusion process and satisfies the following SDE

$$d\lambda_{n,i}(t) = dW_i(t) + \left( -\lambda_{n,i}(t) + \sum_{j \neq i} \frac{1}{\lambda_{n,i}(t) - \lambda_{n,j}(t)} \right) dt, \quad 1 \leq i \leq n$$

where  $W_i, 1 \leq i \leq n$  are  $n$  independent standard Brownian motions. The corresponding diffusion operator is

$$L = \frac{1}{2} \Delta + \sum_{i=1}^n \left( -x_i + \sum_{j \neq i} \frac{1}{x_i - x_j} \right) \frac{\partial}{\partial x_i}.$$

Furthermore, the equilibrium distribution associated with  $L$  is given by  $d\nu(\mathbf{x}) = p(\mathbf{x})d\mathbf{x}$ , where

$$p(\mathbf{x}) = \frac{1}{Z_n} \prod_{1 \leq i < j \leq n} (x_j - x_i)^2 \prod_{i=1}^n e^{-x_i^2/2}, \quad \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

This is the joint probability density function of eigenvalues of GUE, which is same up to a constant factor as in (10)

The above diffusion process ( $\lambda_n(t), t \geq 0$ ) was first introduced by Dyson [13] to describe the diffusion of  $n$  mutually repelling particles with positions. Now it is named Dyson's Brownian motion in the literature. According to a classic Karlin-McGregor formula for non-intersecting diffusion process, it follows for  $s < t$

$$P(\lambda_n(t) = \mathbf{y} | \lambda_n(s) = \mathbf{x}) = \frac{\prod_{1 \leq i < j \leq n} (y_j - y_i)}{\prod_{1 \leq i < j \leq n} (x_j - x_i)} \det \left( \exp \left( -\frac{(y_j - \rho_{s,t} x_i)^2}{2(1 - \rho_{s,t}^2)} \right) \right),$$

and so by the Markovian property we have

$$\begin{aligned} & P(\lambda_n(t_1) = \mathbf{x}^1, \lambda_n(t_2) = \mathbf{x}^2, \dots, \lambda_n(t_m) = \mathbf{x}^m) \\ &= P(\lambda_n(t_1) = \mathbf{x}^1) \prod_{i=2}^m P(\lambda_n(t_i) = \mathbf{x}^i | \lambda_n(t_{i-1}) = \mathbf{x}^{i-1}) \\ &= \frac{1}{Z_{n,m}} \Delta(\mathbf{x}^1) \prod_{i=1}^n e^{-(x_i^1)^2} \prod_{i=2}^m \det \left( \exp \left( -\frac{(x_k^i - \rho_{t_i, t_{i-1}} x_l^{i-1})^2}{2(1 - \rho_{t_i, t_{i-1}}^2)} \right) \right)_{1 \leq k, l \leq m} \Delta(\mathbf{x}^m). \end{aligned} \quad (33)$$

This is a basic example of extended determinantal point processes. Indeed, let  $\mathcal{X}_0 = \mathcal{X}_m = \{0, 1, 2, \dots, n-1\}$ ,  $\mathcal{X}_r = \mathbb{R}, 1 \leq r \leq m-1$ , and equip  $\mathcal{X}_r$  with the Lebesgue measure  $d\mu_r$ . Let  $\mathbf{x}_i^0 = \mathbf{x}_i^m = i-1$  for  $i = 1, 2, \dots, n$ . Set  $\phi_{0,1}(i, x) = h_i(x)e^{-x^2}$ ,  $\phi_{m-1,m}(x, i) = h_i(x)$  where  $h_i(x)$  is the normalized Hermitian polynomials (see (11)), and

$$\phi_{r,r+1}(x, y) = \frac{1}{\sqrt{\pi(1 - q_r^2)}} \exp(-(y - q_r x)^2 / (1 - q_r^2)).$$

Then (33) is written as follows

$$\frac{1}{Z_{n,m}} \prod_{r=0}^{m-1} \det \left( \phi_{r,r+1}(x_i^r, x_j^{r+1}) \right)_{1 \leq i, j \leq n} d\mu_1^n(\mathbf{x}^1) \cdots d\mu_1^n(\mathbf{x}^{m-1}). \quad (34)$$

It is proved in [20, 21] that the probability measure (34) has determinantal correlation function, that is, the probability density with respect to the reference measure  $d\mu_{r_1}(y_1) \cdots d\mu_{r_k}(y_k)$  of finding particles at  $z_1 = (r_1, y_1), \dots, z_k = (r_k, y_k)$  is given by

$$\det \left( K_{n,m}(z_i, z_j) \right)_{1 \leq i, j \leq k},$$

where  $K_{n,m}$  is the so called correlation kernel. This kernel is given by

$$K_{n,m}(r, x; s, y) = -\phi_{r,s}(x, y) + \sum_{i,j=1}^n \phi_{r,m}(x, x_i^m)(A^{-1})_{i,j}\phi_{0,s}(x_j^0, y),$$

where  $A = (\phi_{0,m}(x_i^0, x_j^m))_{1 \leq i,j \leq n}$ ,  $\phi_{r,s} : \mathcal{X}_r \times \mathcal{X}_s \mapsto \mathbb{R}$  for  $r < s$  is given by

$$\phi_{r,s}(x, y) = \int \phi_{r,r+1}(x, z_1) \cdots \phi_{s-1,s}(z_{r-s-1}, y) d\mu_{r+1}(z_1) \cdots d\mu_{s-1}(z_{r-s-1}).$$

Return to Dyson's Brownian motion. Note the following expansion

$$\frac{1}{\sqrt{\pi(1-q^2)}} \exp\left(-\frac{(y-qx)^2}{1-q^2}\right) = \sum_{k=0}^{\infty} h_k(x)h_k(y)q^k e^{-y^2},$$

where  $0 < q < 1$ . Repeated use of this identity gives

$$\phi_{0,s}(j, y) = e^{-j(t_s-t_1)} h_j(y) e^{-y^2}, \quad \phi_{r,m}(x, j) = e^{-j(t_m-r-t_r)} h_j(x) \tag{35}$$

and

$$\begin{aligned} \phi_{r,s}(x, y) &= \frac{1}{\sqrt{\pi(1-e^{(t_r-t_s)})}} \exp\left(-\frac{(y-e^{t_r-t_s/2}x)^2}{1-e^{t_r-t_s}}\right) \\ &= \sum_{k=0}^{\infty} h_k(x)h_k(y)e^{k(t_r-t_s)} e^{-y^2}. \end{aligned} \tag{36}$$

Set  $\chi_{t,s} = 1$  if  $t < s$  and  $\chi_{t,s} = 0$  if  $t \geq s$ . We get the extended Hermite kernel

$$\begin{aligned} K_{Herm}^{ext}(t, x; s, y) &= \frac{1}{\sqrt{\pi(1-e^{(t-s)})}} \exp\left(-\frac{(y-e^{t-s/2}x)^2}{1-e^{t-s}}\right) \chi_{t,s} \\ &\quad + \sum_{k=0}^{n-1} h_k(x)h_k(y)e^{k(t-s)} e^{-y^2}. \end{aligned}$$

Using the second equality in (36) we obtain the alternative formula

$$K_{Herm}^{ext}(t, x; s, y) = \begin{cases} \sum_{k=0}^{n-1} h_k(x)h_k(y)e^{k(t-s)} e^{-y^2}, & t \geq s \\ -\sum_{k=n}^{\infty} h_k(x)h_k(y)e^{k(t-s)} e^{-y^2}, & t < s \end{cases}$$

Let  $\gamma$  be a positively oriented circle around the origin with radius  $r > 0$ , and  $\Gamma$  the line  $\mathbb{R} + iL$  with  $L > \gamma$ . Using the integral formulas,

$$\begin{aligned} H_n(x) &= \frac{2^n}{i\sqrt{\pi}} e^{x^2} \int_{\Gamma} e^{w^2-2xw} w^n dw, \\ H_n(x) &= \frac{n!}{2\pi i} \int_{\gamma} e^{-z^2+2xz} \frac{1}{z^{n+1}} dz, \end{aligned}$$

it is not difficult to show that

$$\begin{aligned} K_{Herm}^{ext}(t, x; s, y) &= \frac{1}{\sqrt{\pi(1-e^{(t-s)})}} \exp\left(-\frac{(y-e^{t-s/2}x)^2}{1-e^{t-s}}\right) \chi_{t,s} \\ &\quad + \frac{2}{(2\pi i)^2} \int_{\Gamma} dw \int_{\gamma} dz \frac{w^n}{z^n} \frac{1}{w-z} e^{w^2-2yw-e^{t-s}z^2+2e^{t-s/2}xz}. \end{aligned}$$

The double contour integral representation can be useful for asymptotic computations in the proof of convergence to the extended Airy kernel when we have the edge scaling.

Also, note the kernel is not unique. We can multiply it by  $\frac{\psi(r,x)}{\psi(s,y)}$  for an arbitrary function  $\psi \neq 0$  and get the same correlation functions. In particular, we can get the extended Hermite

kernel multiplying with  $\exp(-x^2/2 + y^2/2)$ ,

$$K_{Herm}^{ext}(t, x; s, y) = \begin{cases} \sum_{k=0}^{n-1} \varphi_k(x)\varphi_k(y)e^{k(t-s)}, & t \geq s \\ -\sum_{k=n}^{\infty} \varphi_k(x)\varphi_k(y)e^{k(t-s)}, & t < s \end{cases}$$

where  $\varphi_k(x) = h_k(x)e^{-x^2/2}$ . Note

$$P(\lambda_{n,n}(t_1) \leq x_1, \dots, \lambda_{n,n}(t_m) \leq x_m) = \det(\mathbf{1} - \chi^{1/2} K_{Herm}^{ext} \chi^{1/2})_{L^2(\{t_1, \dots, t_m\} \times \mathbb{R})}.$$

Now consider the rescaled largest eigenvalue given by

$$\tilde{\lambda}_{n,n}(u) = \sqrt{2}n^{1/6} \left( \lambda_{n,n}\left(\frac{u}{n^{1/3}}\right) - \sqrt{2}n \right), \quad u \in \mathbb{R}.$$

Recall the Plancherel-Rotach formula for asymptotics of Hermitian polynomials

$$\lim_{n \rightarrow \infty} 2^{-1/4} n^{1/12} \varphi_n \left( \sqrt{2}n + \frac{x}{\sqrt{2}n^{1/6}} \right) = \text{Ai}(x) \quad x \in \mathbb{R}.$$

It follows that for  $t > s$

$$\begin{aligned} & K_{Herm}^{ext} \left( \frac{t}{n^{1/3}}, \sqrt{2}n + \frac{x}{\sqrt{2}n^{1/6}}; \frac{s}{n^{1/3}}, \sqrt{2}n + \frac{y}{\sqrt{2}n^{1/6}} \right) \\ &= \frac{1}{\sqrt{2}n^{1/6}} \sum_{k=0}^{n-1} \phi_k \left( \sqrt{2}n + \frac{x}{\sqrt{2}n^{1/6}} \right) \phi_k \left( \sqrt{2}n + \frac{y}{\sqrt{2}n^{1/6}} \right) e^{k(t-s)/n^{1/3}} \\ &= \frac{1}{\sqrt{2}n^{1/6}} \sum_{k=0}^{n-1} \phi_{n-1-k} \left( \sqrt{2}n + \frac{x}{\sqrt{2}n^{1/6}} \right) \phi_{n-1-k} \left( \sqrt{2}n + \frac{y}{\sqrt{2}n^{1/6}} \right) e^{(n-1-k)(t-s)/n^{1/3}} \\ &\sim e^{(n-1)(t-s)/n^{1/3}} \int_0^\infty e^{-(t-s)\lambda} \text{Ai}(x+\lambda) \text{Ai}(y+\lambda) d\lambda \\ &\sim e^{n^{2/3}(t-s)} \mathcal{A}^{ext}(t, x; s, y). \end{aligned}$$

Similarly, for  $t < s$

$$\begin{aligned} & K_{Herm}^{ext} \left( \frac{t}{n^{1/3}}, \sqrt{2}n + \frac{x}{\sqrt{2}n^{1/6}}; \frac{s}{n^{1/3}}, \sqrt{2}n + \frac{y}{\sqrt{2}n^{1/6}} \right) \\ &\sim e^{n^{2/3}(t-s)} \left( \frac{1}{\sqrt{2\pi(s-t)}} e^{-\frac{(y-x)^2}{2(s-t)}} + \int_0^\infty e^{-(t-s)\lambda} \text{Ai}(x+\lambda) \text{Ai}(y+\lambda) d\lambda \right) \\ &\sim e^{n^{2/3}(t-s)} \mathcal{A}^{ext}(t, x; s, y). \end{aligned}$$

As a consequence, we have so far proved the following

$$\tilde{\lambda}_{n,n}(t) \Rightarrow \mathcal{A}_2(t), \quad n \rightarrow \infty$$

in the sense of convergence of finite dimensional distributions.

### 3.5 The extended Airy point process

Motivated by the Dyson's Brownian motions, we can give the third representation of Airy<sub>2</sub> process as a marginal process of an extended Airy point process.

Let  $m \geq 1$  be an arbitrary integer, and let  $t_1 < t_2 < \dots < t_m$  points in  $\mathbb{R}$  which we shall think of as times. Define

$$\mathcal{E} = \mathbb{R}_{t_1} \cup \mathbb{R}_{t_2} \cup \dots \cup \mathbb{R}_{t_m}.$$

$\mathbb{R}_{t_j}$  is referred to as the time line  $t_j$ . Assume that  $\mathcal{X}$  is the space of all locally finite countable configurations of points (or particles) in  $\mathcal{E}$ . Let  $\Sigma$  be the minimal  $\sigma$ -algebra that contains all cylinder sets. The so-called *extended Airy point process* is such a measure on the space  $(\mathcal{E}, \Sigma)$  that its  $k$ -point ( $1 \leq k \leq m$ ) correlation function has the extended Airy kernel  $\mathcal{A}^{ext}$ , see (15).



In particular, let  $z_1 = (t_{r_1}, x_1), \dots, z_k = (t_{r_k}, x_k)$  be  $k$  points from  $\mathcal{E}$ , where  $r_1, \dots, r_k$  are possibly same, then

$$\begin{aligned} \rho_k(z_1, z_2, \dots, z_k) &= P(\{z_1, z_2, \dots, z_k\} \subseteq \mathcal{X}) \\ &= \det(\mathcal{A}^{ext}(t_{r_i}, x_i; t_{r_j}, x_j))_{1 \leq i, j \leq k}. \end{aligned} \tag{37}$$

It is a well-known fact that at each time line  $\mathbb{R}_{t_i}$  there is almost surely a largest particle  $\lambda(t_i)$ . Moreover, it follows

$$(\lambda(t_1), \dots, \lambda(t_m)) \stackrel{d}{=} (\mathcal{A}_2(t_1), \dots, \mathcal{A}_2(t_m)). \tag{38}$$

The determinantal point process has become a very hot object of study in probability theory since Borodin and Olshanski coined around 2000. Interestingly, both Poisson processes and Brownian motions are determinantal point processes.

### §4 Sample Path Properties

#### 4.1 Modulus of Continuity for Airy<sub>2</sub> process

As we observed, the determinantal formula (18) and its equivalent expression (21) are very useful in the definition of Airy<sub>2</sub> process and in the study of long range correlations, stationarity, and local Brownian motion behaviours. However, it is difficult to derive even the most basic path properties, such as continuity, from it directly, see Prähofer and Spohn [27]. There has been a good advance in the exploration of sample path properties recently. In particular, a useful technique, called *the Brownian Gibbs property*, was developed by [6–8, 10, 11].

The main result of this subsection is the modulus of continuity for Airy<sub>2</sub> process. It reads as follows.

$$\sup_{s, t \in [0, 1]} \frac{|\mathcal{A}_2(t) - \mathcal{A}_2(s)|}{\sqrt{|t - s|} |\log |t - s||} < \infty \quad a.s. \tag{39}$$

More precisely, there exist two positive constants  $c_7$  and  $c_8$ , and a random variable  $\Xi$  so that

$$|\mathcal{A}_2(t) - \mathcal{A}_2(s)| \leq \Xi \sqrt{|t - s|} |\log |t - s||$$

for all  $s, t \in [-T, T]$ , and

$$P(\Xi > x) \leq c_7 e^{-c_8 x^2}.$$

In addition, for any  $t \in \mathbb{R}$ ,  $0 < \delta \leq 1$ , it follows

$$P\left(\sup_{0 < h < \delta} |\mathcal{A}_2(t+h) - \mathcal{A}_2(t)| > x \delta^{1/2}\right) \leq c_9 e^{-c_{10} x^2} \tag{40}$$

for any  $x > 0$ .

It is enough to prove that there are constants  $c_{11}, c_{12}$  such that for every  $t \in \mathbb{R}$ ,  $\delta \in (0, 1)$

$$P(|\mathcal{A}_2(t) - \mathcal{A}_2(t + \delta)| > x \sqrt{\delta}) \leq c_{11} e^{-c_{12} x^2}. \tag{41}$$

Indeed, (39) directly follows from (41) according to a classical Lévy’s technique (i.e. a generic chaining argument).

Proceed with the proof of (41). We need to introduce the concept of a line ensemble and the Brownian Gibbs property.

A line ensemble  $\mathcal{L} = (L(i, t), i \geq 1, t \in \mathbb{R})$  is a collection of random continuous curves indexed by  $\mathbb{N}$ , each  $L(i, \cdot)$  of which maps  $\mathbb{R}$  into  $\mathbb{R}$ . If for all  $i < j$ ,  $L(i, t) > L(j, t)$  for all  $t \in \mathbb{R}$ , then  $\mathcal{L}$  is said to be a non-intersecting line ensemble.

Let  $k \geq 1$ ,  $\bar{x} = (x_1, x_2, \dots, x_k)$  and  $\bar{y} = (y_1, y_2, \dots, y_k)$  are two points of  $k$ -decreasing lists.

Let  $a < b$ , and let  $f, g : [a, b] \rightarrow \bar{\mathbb{R}}$  be two continuous functions that satisfy  $f(t) > g(t)$  for all  $t \in [a, b]$  as well as the boundary conditions  $f(a) > x_1, f(b) > y_1$  and  $g(a) < x_k, g(b) < y_k$ . A line ensemble  $\mathcal{Q}$  with index  $\{1, 2, \dots, k\}$  is called a  $(f, g)$ -avoiding Brownian line ensemble on  $[a, b]$  with entrance data  $\bar{x}$  and exit data  $\bar{y}$  if the law of  $\mathcal{Q}$  is equal to the law of  $k$ -independent Brownian bridges  $\{B_i : (a, b) \rightarrow \mathbb{R}\}$  from  $B_i(a) = x_i$  to  $B_i(b) = y_i$  conditioned on the event that for all  $t \in [a, b]$

$$f(t) > B_1(t) > B_2(t) > \dots > B_k(t) > g(t).$$

For any  $k_1 < k_2$  and any  $a < b$ , write  $D_{k_1, k_2}^{a, b} = \{k_1, k_1 + 1, \dots, k_2\} \times [a, b]$  and  $\bar{D}_{k_1, k_2}^{a, b} = \mathbb{N} \times \mathbb{R} \setminus D_{k_1, k_2}^{a, b}$ . A line ensemble  $\mathcal{L}$  is said to have the *Brownian Gibbs property* if

$$\mathcal{L}|_{D_{k_1, k_2}^{a, b}} \text{ conditional on } \mathcal{L}|_{\bar{D}_{k_1, k_2}^{a, b}} \stackrel{d}{=} \mathcal{Q}, \tag{42}$$

where  $\mathcal{Q}_i = \tilde{\mathcal{Q}}_{i-k_1+1}$  and  $\tilde{\mathcal{Q}}_{i-k_1+1}$  is the  $(f, g)$ -avoiding Brownian line ensemble on  $[a, b]$  with  $\bar{x} = (L_{k_1}(a), \dots, L_{k_2}(a))$  and  $\bar{y} = (L_{k_1}(b), \dots, L_{k_2}(b))$ ,  $f = L_{k_1-1}$  and  $g = L_{k_2+1}$  by convention  $L_0 = -\infty$ .

A line ensemble  $\mathcal{R} = (R(i, t), i \geq 1, t \in \mathbb{R})$  is called the Airy line ensemble if its finite dimensional distributions are given by the extended Airy point process (see Subsection 3.5). Corwin and Hammond [6] proved the existence of the Airy line ensemble: There exists a unique continuous non-intersecting  $\mathbb{N}$ -indexed line ensemble with finite dimensional distributions given by the extended Airy point process. Moreover, they also proved that the  $\mathbb{N}$ -indexed line ensemble  $\mathcal{L}$  given by  $L(i, t) = 2^{-1/2}(R(i, t) - t^2)$  for each  $i \in \mathbb{N}$  has the *Brownian Gibbs property*. Some authors refer to  $\mathcal{L}$  as the parabolic Airy line ensemble. Trivially,  $R(1, t) = \mathcal{A}_2(t)$ .

To prove (41), we may and do assume  $x > 7\delta^{3/2}$ , otherwise, just change the values of constants. Set  $b = \delta^2 + x\sqrt{\delta}$ . Since  $\mathcal{A}_2(t), t \in \mathbb{R}$  is stationary, then

$$\begin{aligned} P(|\mathcal{A}_2(t) - \mathcal{A}_2(t + \delta)| > x\sqrt{\delta}) &= 2P(\mathcal{A}_2(0) - \mathcal{A}_2(\delta) > x\sqrt{\delta}) \\ &= 2P(L(1, 0) - L(1, \delta) > b). \end{aligned}$$

Set  $r = b/8\delta$ , then trivially  $r > \delta$ . By the *Brownian Gibbs property*, conditionally on the set  $\{L(1, 0), L(1, \delta)\} \cup \{L(2, u), u \in [0, r]\}$ ,

$$(L(1, u), u \in [0, r]) \stackrel{d}{=} (B^0(u), u \in [0, r]),$$

where  $(B^0(u), u \in [0, r])$  is a Brownian bridge of variance parameter 2 with  $B^0(0) = L(1, 0)$ ,  $B^0(r) = L(1, r)$  conditioned to avoid  $(L(2, u), u \in [0, r])$ .

Construct a linear function  $y(u)$  with  $y(0) = L(1, 0)$  and  $y(r) = L(1, r)$ :

$$y(u) = L(1, 0) + \frac{L(1, r) - L(1, 0)}{r}u, \quad u \in [0, r] \tag{43}$$

Then it follows

$$\begin{aligned} &P(L(1, 0) - L(1, \delta) > b) \\ &= E(P(L(1, 0) - L(1, \delta) > b | \{L(1, 0), L(1, \delta)\} \cup \{L(2, u), u \in [0, r]\})) \\ &= E(P(B^0(0) - B^0(\delta) > b | \{L(1, 0), L(1, \delta)\} \cup \{\text{No touch } (L(2, u), u \in [0, r])\})) \\ &\leq P((\tilde{B}^0(0) + y(0)) - (\tilde{B}^0(\delta) + y(\delta)) > b), \end{aligned}$$

where  $(\tilde{B}^0(u), u \in [0, r])$  is a standard Brownian bridge and  $B^0(u) = \tilde{B}^0(u) + y(u)$ .

On the other hand, we have

$$\begin{aligned} &P((\tilde{B}^0(0) + y(0)) - (\tilde{B}^0(\delta) + y(\delta)) > b) \\ &\leq P(y(0) - y(\delta) > \frac{b}{2}) + P(\tilde{B}^0(\delta) < -\frac{b}{2}). \end{aligned}$$

Note it follows

$$\begin{aligned}
 P\left(y(0) - y(\delta) > \frac{b}{2}\right) &= P\left(L(1, 0) - L(1, \delta) > \frac{br}{2\delta}\right) \\
 &= P\left(\mathcal{A}_2(0) - \mathcal{A}_2(\delta) + \delta^2 > \frac{br}{2\delta}\right) \\
 &= P\left(\mathcal{A}_2(0) \geq E\mathcal{A}_2(0) + \frac{br}{4\delta}\right) + P\left(\mathcal{A}_2(\delta) \leq E\mathcal{A}_2(\delta) + \delta^2 - \frac{br}{4\delta}\right) \\
 &\leq P\left(\mathcal{A}_2(0) \geq E\mathcal{A}_2(0) + \frac{b^2}{32\delta^2}\right) + P\left(\mathcal{A}_2(\delta) \leq E\mathcal{A}_2(\delta) - \frac{b^2}{64\delta^2}\right),
 \end{aligned}$$

where in the last inequality we used the fact  $b > 8\delta^2$ .

It is known that there exist  $c_{13}, c_{14}$  such that for any  $x > c_{14}$

$$P(|\mathcal{A}_2(0) - E\mathcal{A}_2(0)| > x) \leq c_{13}e^{-c_{14}x} \tag{44}$$

and

$$P(|\tilde{B}^0(\delta)| > x\sqrt{\delta}) \leq c_{15}e^{-c_{16}x^2}. \tag{45}$$

We refer to Dauvergne and Virág [10] for the proofs of (44) and (45). Putting these together concludes the proof of (41).

As a corollary, we have for any  $x > c_{17} > 0$

$$P\left(\sup_{t \in [0, 1]} |\mathcal{A}_2(t) - E\mathcal{A}_2(t)| > x\right) \leq c_{18}e^{-c_{19}x}.$$

We remark that the proof of (39) do depend on the intricacies of the jump ensemble method, a method that exploits the *Brownian Gibbs property* of the parabolic Airy line ensemble in order to make inferences about its curves, notably the ones at its edge, such as  $\mathcal{A}_2$  after parabolic shift. As Hammond pointed out (private communication): *the method may seem complex and indirect, though it is in its fundamentals attractively probabilistic and geometric. I am not aware that the result can be proved by viewing the Airy process in isolation. It is by embedding it in the richer Airy line ensemble that this valuable Brownian Gibbs method becomes available.*

### 4.2 Maximum of Airy<sub>2</sub> process

In this subsection, we turn to the distribution of the maximum of  $\mathcal{A}_2(t)$  over the real line  $\mathbb{R}$ . Define

$$\mathcal{M} = \max_{t \in \mathbb{R}} \{\mathcal{A}_2(t) - t^2\}, \quad \mathcal{T} = \arg \max_{t \in \mathbb{R}} \{\mathcal{A}_2(t) - t^2\}.$$

As an analogue of (3), we have the following remarkable result:

$$P(\mathcal{M} \leq x) = F_1(4^{1/3}x), \quad x \in \mathbb{R} \tag{46}$$

Here  $F_1$  is the limiting distribution of the largest eigenvalue of Gaussian Orthogonal Ensembles (GOE), and has a close link with  $F_2$  through

$$F_1(x) = F_2^{1/2}(x)e^{-\frac{1}{2} \int_x^\infty q(u)du}$$

The formula (46) was first conjectured by Johansson [20], which explored the discrete growth models like polynuclear growth models (PNG) and last passage percolation (LPP) times in the planar lattice. Indeed,  $F_2$  and  $F_1$  are the limiting distributions of point-to-point LPP times and point-to-line LPP times after properly scaled, respectively, see Johansson [20], Prähofer and Spohn [27] for detailed information. on the other hand, the point-to-line LPP times can be simply computed as the maximum of the point-to-point LPP times. Thus (46) can be proved, though in a very indirect way.

A direct proof was later provided by Corwin, Quastel and Remenik [8]. It largely relies on the following two fundamental facts, which are of interest in their own right. The first one is the determinantal representation of  $F_1$  obtained by Ferrari and Spohn [15]:

$$F_1(x) = \det(\mathbf{1} - P_0 B_x P_0)|_{L^2(\mathbb{R})},$$

where  $B_x$  is the kernel

$$B_x(u, v) = \text{Ai}(u + v + x).$$

The second fact is, if letting  $H$  be as in (19),  $R$  a reflection operator with kernel given by

$$R_T(u, v) = \frac{1}{\sqrt{8\pi T}} e^{-(u+v-2x-2T^2)^2/8T-(u+v)T+2T^3/3}$$

and defining

$$\Theta_T = \bar{P}_{x+T^2}(e^{-2TH} - R_T)\bar{P}_{x+T^2},$$

then it follows

$$P\left(\sup_{t \in [-T, T]} \{\mathcal{A}_2(t) - t^2\} \leq x\right) = \det(\mathbf{1} - \mathcal{A} + e^{TH} \mathcal{A} \Theta_T e^{TH} \mathcal{A})|_{L^2(\mathbb{R})}. \tag{47}$$

Note  $e^{TH} R_T e^{TH} = \varrho_x$ , where  $\varrho_x$  is the reflection operator, i.e.,  $\varrho_x f(u) = f(2x - u)$ . Hence

$$\begin{aligned} P(\mathcal{M} \leq x) &= \lim_{T \rightarrow \infty} P\left(\sup_{t \in [-T, T]} \{\mathcal{A}_2(t) - t^2\} \leq x\right) \\ &= \det(\mathbf{1} - \mathcal{A} \varrho_x \mathcal{A})|_{L^2(\mathbb{R})}. \end{aligned}$$

Finally, to see the identity (46), we need only to note  $\mathcal{A} = B_0 P_0 B_0$  and  $B_0 \varrho_x B_0(u, v) = 2^{-1/3} \text{Ai}(2^{-1/3}(u + v + 2x))$ . The reader is referred to the survey on Airy<sub>2</sub> processes and variational problems by Quastel and Remenik [29] for explicit derivation.

Proceed with  $\mathcal{T}$ . The study of  $\mathcal{T}$  has received quite a bit of recent interest in the physics literature since it is believed that the distribution of  $\mathcal{T}$  is universal so that it governs the rescaled endpoint of directed polymers in 1+1 dimensions for large time or temperature. Flores, Quastel and Remenik [14] obtained an explicit expression for the distribution of  $\mathcal{T}$ . Later on, Quastel and Remenik [30] start from a probabilistic argument and use the continuum statistics formula (47) to establish the following tail decay: There exists a  $c_{20}$  such that for every  $c_{21} > \frac{32}{3}$  and large enough  $t$

$$e^{-c_{20}t^3} \leq P(|\mathcal{T}| > t) \leq c_{21}e^{(-4/3)t^3+2t^2+O(t^{3/2})}. \tag{48}$$

The order  $e^{-ct^3}$  confirms a prediction made in the physics literature. It is believed that the  $-4/3$  in the upper bound is the correct exponent. In particular, it is open to show

$$\frac{\log P(|\mathcal{T}| > t)}{t^3} \rightarrow -\frac{4}{3}, \quad t \rightarrow \infty.$$

The key step for the proof of (48) is to prove that there exists a constant  $c_T > 0$  for each  $T \geq 1$  such that for every  $x > 0$

$$P\left(\sup_{-T \leq t \leq T} \mathcal{A}_2(t) > x + 1\right) \leq c_T e^{-4/3x^{3/2}}.$$

### 4.3 Concluding Remarks

We have restricted ourselves to the Airy<sub>2</sub> process  $\mathcal{A}_2$  with its sample paths. As seen above,  $\mathcal{A}_2$  is a comparatively new stationary non-Markovian continuous random process, and is very worthy of further attention. In addition to the Airy<sub>2</sub> process  $\mathcal{A}_2$ , there have appeared other new processes very alike, say  $\mathcal{A}_1$ ,  $\mathcal{A}_{stat}$ ,  $\mathcal{A}_{2 \rightarrow 1}$ ,  $\mathcal{A}_{2 \rightarrow BM}$ ,  $\mathcal{A}_{1 \rightarrow BM}$  in the recent study around

random growth models. A detailed survey is beyond the scope of this paper. Indeed, it is widely open to estimate precisely the moduli of continuity for their sample paths. Hopefully, a satisfactory progress would be made in the near future.

### Acknowledgement

Z.G.Su wishes to express his best gratitude to Yimin Xiao for his instructive comments on sample paths properties of random processes. The authors are grateful to the anonymous referee for careful reading.

### References

- [1] M Anderson, A Guionnet, O Zeitouni. *An Introduction to Random Matrices*, Cambridge University Press, 2008.
- [2] J Baik, P Deift, K Johansson. *On the distribution of the length of the longest increasing subsequence of random permutations*, J Amer Math Soc, 1999, 12 (4): 1119-1178.
- [3] J Baik, T Kriecherbauer, L C Li. *Integrable Systems and Random Matrices: In Honor of Percy Deift*, Contemporary Mathematics, Amer Math Soc, 2008.
- [4] P Billingsley. *Convergence of Probability Measures*, Wiley Series in Probability and Statistics, Wiley, 2009.
- [5] P Billingsley. *Probability and Measure*, Wiley Series in Probability and Statistics, Wiley, 2012.
- [6] I Corwin, A Hammond. *Brownian Gibbs property for Airy line ensembles*, Invent Math, 2014, 195 (2): 441-508.
- [7] I Corwin, A Hammond. *KPZ line ensemble*, Probab Theory Rel Fieldsh, 2016, 166 (1-2): 67-185.
- [8] I Corwin, J Quastel, D Remenik. *Continuum statistics of the Airy<sub>2</sub> process*, Comm Math Phys, 2013, 317(2): 347-362.
- [9] M Csörgő, P Révész. *Strong Approximations in Probability and Statistics*, 1981, Academic Press.
- [10] D Dauvergne, B Virag. *Basic properties of the line ensemble*, arXiv:1812.00311, 2018.
- [11] D Dauvergne, J Ortmann, B Virag. *The directed landscape*, arXiv:1812.00309, 2018.
- [12] R Durrett. *Probability: Theory and Examples*, 5th edition, Cambridge University Press, 2018.
- [13] F J Dyson. *A Brownian-motion model for the eigenvalues of a random matrix*, J Math Phys, 1962, 3: 1191-1198.
- [14] G M Flores, J Quastel, D Remenik. *Endpoint distribution of directed polymers in 1+1 dimensions*, Comm Math Phys, 2013, 317(2): 363-380.
- [15] P L Ferrari, H Spohn. *A determinantal formula for the GOE GUE Tracy-Widom distribution*, J Phys A, 2005, 38(33): L557-L561.

- [16] P Groeneboom, N M Temme. *The tail of the maximum of Brownian motion minus a parabola*, Electron Commun Probab, 2011,16: 458-466.
- [17] J Hägg. *Local Gaussian fluctuations in the Airy and discrete PNG processes*, Ann Probab, 2008, 36: 1059-1092.
- [18] A Hammond. *Brownian regularity for the Airy line ensemble, and multi-polymer watermelons in Brownian last passage percolation*, Mem Amer Math Soc, arXiv:1609.02971, 2016.
- [19] A Hammond, *Modulus of continuity of polymer weight profiles in Brownian last passage percolation*, Ann Probab, 2019, 47: 3911-3962.
- [20] K Johansson. *Discrete polynucleat growth and determinantal process*, Comm Math Phys, 2003, 242(1-2): 277-329.
- [21] K Johansson. *Random Matrices and Determinantal Processes*, Ecole de Physique, Les Houches, 2005, 141-162.
- [22] I Karatzas, S E Shreve. *Brownian Motion and Stochastic Calculus*, 2nd edition, Graduate Texts in Mathematics, 113, Springer, New York, 1998.
- [23] Z Y Lin, C R Lu, Z G Su. *Foundations of Probability Limit Theory*, 2nd edition, Higher Education Press, 2015.
- [24] B F Logan, L A Shepp. *A variational problem for random Young tableaux*, Adv Math, 1977, 26(2): 206-222.
- [25] A Okounkov. *Generating functions for intersection numbers on moduli spaces of curves*, Int Math Res Not, 2002, 18: 933-957.
- [26] L Pimentel. *Local behavior of Airy<sub>2</sub> processes*, arXiv:1704.01903, 2017.
- [27] M Prähofer, H Spohn. *Scale invariance of the PNG droplet and the Airy<sub>2</sub> process*, J Stat Phys, 2002, 108(5-6): 1071-1106.
- [28] S Prohac, H Spohn. *The one-dimensional KPZ equation and the Airy<sub>2</sub> process*, J Stat Mech Theory Exp, 2011, 2011(03): 1-15.
- [29] J Quatesl, D Remenik. *Airy<sub>2</sub> process and variational problems*, In Topics in Percolative and Disordered Systems (A Rámiez, G Ben Arous, P A Ferrari, C Newman, V Sidoravicius, M E Vares, eds.), Springer Proc Math Stat, Springer New York, 2014, 69: 121-171.
- [30] J Quastel, D Remenik. *Tails of the endpoint distribution of directed polymers*, Ann Inst Henri Poincaré Probab Stat, 2015, 51 (1): 1-17.
- [31] D Revuz, M Yor. *Continuous Martingales and Brownian Motion*, Grundlehren der mathematischen Wissenschaften, Springer, 1999.
- [32] G Shinault, C Tracy. *Asymptotics for the covariance of the Airy<sub>2</sub> process*, J Stat Phys, 2011, 143, 60-71.
- [33] Z G Su. *GUE Tracy-Widom distribution with Applications*, Journal of Applied Probability and Statistics , 2016, 32:551-580.(in chinese)

- [34] C A Tracy, H Widom. *Level-spacing distribution and the Airy kernel*, Comm Math Phys, 1994, 159 (1): 151-174.
- [35] A M Vershik, S V Kerov. *Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux*, Dokl Akad Nauk SSSR, 1977, 233(6): 1024-1027.
- [36] R Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2018.
- [37] O Vallée, M Soares. *Airy Functions and Applications to Physics*, Imperial College Press, 2004.
- [38] M J Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, 2019.
- [39] H Widom. *On asymptotics for the Airy process*, J Stat Phys, 2004, 115 (3-4): 1129-1234.

School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China.

Email: suzhonggen@zju.edu.cn