



**HAL**  
open science

## Random Young diagrams in a Rectangular Box

Dan Beltoft, Cédric Boutillier, Nathanaël Enriquez

► **To cite this version:**

Dan Beltoft, Cédric Boutillier, Nathanaël Enriquez. Random Young diagrams in a Rectangular Box. 2010. hal-00508535v1

**HAL Id: hal-00508535**

**<https://hal.science/hal-00508535v1>**

Preprint submitted on 4 Aug 2010 (v1), last revised 14 Dec 2010 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Random Young diagrams in a Rectangular Box

Dan Beltoft\*, Cédric Boutillier†, Nathanaël Enriquez‡,‡

August 4, 2010

## Abstract

We exhibit the limit shape of random Young diagrams having a distribution proportional to the exponential of their area, and confined in a rectangular box. The Ornstein-Uhlenbeck bridge arises from the fluctuations around the limit shape.

**Keywords:** Young diagrams, Gauss polynomials, Ornstein-Uhlenbeck process

**Math. Subject Classification:** 60C05, 60F17, 05A10, 05A15, 05A30

## Introduction

A *partition* of an integer  $n$  is a finite non-increasing sequence of integers

$$\pi = (\pi_1, \pi_2, \dots, \pi_k),$$

with  $\pi_1 + \dots + \pi_k = n$ . A convenient graphical representation in  $(\mathbb{Z}^+)^2$  of the partition  $\pi$  is its *Young diagram*, also denoted by  $\pi$ . The Young diagram consists of a stack of  $\pi_i$  unit squares on the  $i$ -th column. See Figure 1. The *area*  $|\pi|$  of the Young diagram  $\pi$  is the total number of squares, which is equal to  $n$ .

If the number of summands  $k$  is less than some integer  $a$  and all the summands  $\pi_j$  are less than  $b$ , then the Young diagram stays in the box  $[0, a] \times [0, b]$ .

In this paper we study the limit shape and fluctuations of large random Young diagrams assigned to stay inside a (large) rectangular box, when the probability measure is proportional to  $q^{|\pi|}$ , and  $q$  is a suitable parameter. It is easy to see that if one doesn't want the system to degenerate in the limit, one has to make  $q$  depend on the size of the box. Namely  $1 - q$  has to be taken of the order of the inverse of the side of the box.

The study of the combinatorics of partitions of integers goes back to Hardy and Ramanujan and then Erdős in the 40's. The statistical physics point of view was introduced by Vershik [9] in his study of the typical shape of the partition of a large integer and obtained what we will call Vershik's curve. Recently, Funaki and Sasada obtained

---

\*Aarhus University, Department of Math. Sciences, Ny Munkegade 118, 8000 Aarhus C, Denmark

†Université Pierre et Marie Curie, Laboratoire LPMA, 4 Place Jussieu 75005 Paris, France

‡Université Paris-Ouest, Laboratoire MODAL'X, 200 Av. de la République 92001 Nanterre, France

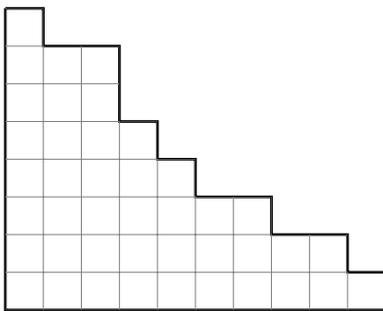


Figure 1: The Young diagram for the partition  $\pi = (8, 7, 7, 5, 4, 3, 3, 2, 2, 1)$  of  $n = 42$ .

Vershik's result, in [6], as a by-product of an hydrodynamic result for the corresponding particle system model.

Our study lies at the intersection of classical topics of probability theory, combinatorics, and statistical physics.

Along the paper, we make an extensive use of the classical Gauss polynomials, well-known in combinatorics as the generating functions for the number of Young diagrams with given area (see e.g. [1]). In order to get to the asymptotic regime, we are led to state a  $q$ -analogue for Stirling's formula.

For all values of the parameters of the problem (limiting aspect ratio of the box, and  $c = \lim n(1 - q)$ ), the limit shape obtained turns out to be a restriction of Vershik's curve as it was recently noticed by Petrov in [7].

The main part of the paper deals with the fluctuations around this limit shape and requires a fine understanding of the boundary of the random diagram. From a probabilistic point of view, the boundary is nothing but the  $q$ -analogue of the bridge of the simple random walk. In the classical case ( $c = 0$ ), the properly rescaled interface converges to the Brownian bridge. In the general case, the limiting process is an Ornstein-Uhlenbeck bridge. This sheds a new light on this process. Let us mention that, in the framework of unlimited partitions, and under some specifications on the summands, the limiting finite-dimensional marginals of the fluctuations has been computed by Vershik, and Yakubovich [8, 10]. However, in contrast with the limit shape problem, the problem of the fluctuations can't be deduced from the unlimited case.

The paper is organized as follows. After a presentation of the combinatorics of the problem (Section 2), we perform its asymptotic analysis through a  $q$ -version of Stirling's formula (Section 3). Section 4 is devoted to the derivation of the limit shape phenomenon. In Section 5, we study the fluctuations around the limit shape: we first compute the limiting 2-correlation function and deduce the convergence of finite-dimensional marginals. This section ends with the delicate proof of the tightness of the fluctuations.

# 1 Presentation of the model

We study the asymptotic distribution of random Young diagrams fitting in a large rectangular box with dimensions  $a \times b$ . Given a real number  $q > 0$ , we assign to each diagram  $\pi$  the probability

$$\mathbb{P}_{a,b}^q(\pi) = \frac{q^{|\pi|}}{Z_{a,b}(q)},$$

where  $|\pi|$  is the number of boxes of  $\pi$  and  $Z_{a,b}(q)$  is the *partition function*, the sum of all  $q^{|\pi|}$ . For the sake of clarity, the probability measure  $\mathbb{P}_{a,b}^q$  will be simply denoted, by  $\mathbb{P}$ .

We fix a parameter  $\rho \in (0, 1)$ , and choose the dimensions of the box for each  $n$  to be  $a_n \times b_n$ , where  $(a_n)$  and  $(b_n)$  are sequences of positive integers satisfying:

$$a_n + b_n = 2n, \quad \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\rho}{1 - \rho}.$$

We are interested in the limiting behavior when  $n$  goes to  $\infty$ . To obtain a non-degenerate limit,  $q$  must go to 1 as  $n$  goes to infinity. We fix a real parameter  $c$  and pose  $q = e^{-\frac{c}{n}}$ .

Note that  $c = 0$  corresponds to a uniform probability measure. We will assume that  $c \neq 0$ . The results for  $c = 0$  can be obtained by taking limits. The physical meaning of the parameter  $c$  is that of a pressure, since it is the variable conjugated to the (two-dimensional) volume.

# 2 Combinatorics of the partitions

Let us start with a fixed box with dimensions  $a \times b$ . The partition function  $Z_{a,b}(q)$  is expressed in terms of Gaussian polynomials or  $q$ -binomial coefficients, where integers  $j$  are replaced by their  $q$ -analogues  $(j)_q = \frac{1-q^j}{1-q}$ .

**Definition 1.** Let  $q > 0$ ,

$$n!_q = \prod_{j=1}^n (j)_q = \prod_{j=1}^n \frac{1-q^j}{1-q}$$

$$\binom{n}{m}_q = \frac{n!_q}{(n-m)!_q m!_q} = \prod_{j=0}^{m-1} \frac{1-q^{n-j}}{1-q^{m-j}}.$$

**Lemma 1.** For all  $(a, b) \in (\mathbb{N}^*)^2$ , and all  $q > 0$ , the partition function  $Z_{a,b}(q)$  is equal to

$$Z_{a,b}(q) = \binom{a+b}{a}_q = \binom{a+b}{b}_q.$$

*Proof.*  $\binom{a+b}{a}_q$  and  $Z_{a,b}(q)$  both follow the recursion relation

$$Z_{a,b}(q) = Z_{a,b-1}(q) + q^b Z_{a-1,b}(q),$$

where the first term corresponds to those diagrams with all parts strictly smaller than  $b$ , and the second term to the diagrams with at least one part of size  $b$ .  $\square$

We use the following coordinates. The bounding rectangle is the rectangle in the plane with corners at the points  $(0, 0)$  and  $(a + b, b - a)$  and sides with slopes  $\pm 1$ . The boundary of a diagram is encoded as a lattice path  $(X_k)_{0 \leq k \leq n}$  from the origin to the point  $(a + b, b - a)$  and such that  $X_{k+1} = X_k \pm 1$  for all  $k$ .

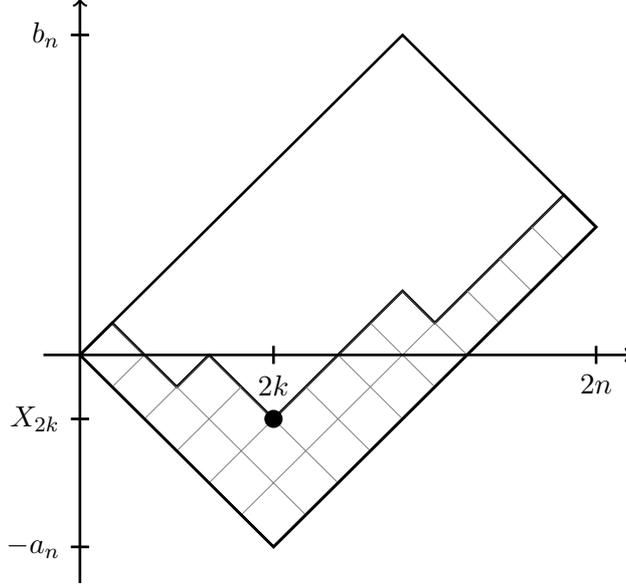


Figure 2: A partition as a lattice path.

Computing the probability  $\mathbb{P}(X_k = \ell)$  is straightforward. A lattice path passing through  $(k, \ell)$  is composed of a path from  $(0, 0)$  to  $(k, \ell)$  and a path from  $(k, \ell)$  to  $(a + b, b - a)$ . Since  $X_k$  has the same parity as  $k$ , the exact formulas also depend on this parity.

**Proposition 1.** *The 1-dimensional marginal of  $X$  under  $\mathbb{P}$  is given by*

$$\mathbb{P}_{a,b}^q(X_{2k} = 2i) = \frac{q^{(k+i)(a-k+i)}}{Z_{a,b}(q)} Z_{k-i, k+i}(q) Z_{a-k+i, b-k-i}(q) \quad (1)$$

and

$$\mathbb{P}_{a,b}^q(X_{2k+1} = 2i + 1) = \frac{q^{(k+i+1)(a-k+i)}}{Z_{a,b}(q)} Z_{k-i, k+i+1}(q) Z_{a-k+i, b-k-i-1}(q) \quad (2)$$

We focus our attention to the behaviour of the process at even times, which is sufficient to study its scaling limit, since the process has bounded steps  $(+1/-1)$ .

We now mention a unimodality result for the distribution above which turns out to be useful later.

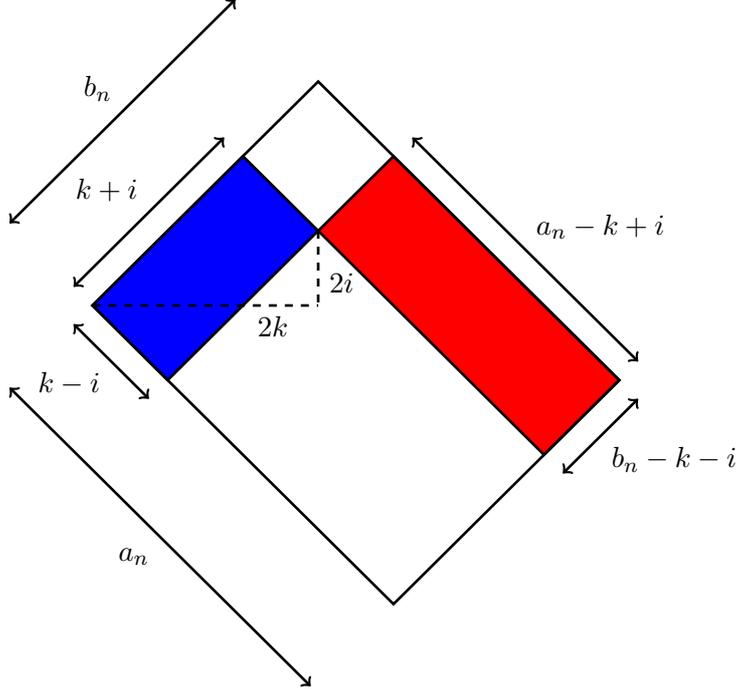


Figure 3: Illustration of Equation (1).

**Lemma 2.** *The function  $\ell \mapsto \mathbb{P}_{a,b}^q(X_{2k} = 2\ell)$  is unimodal: there exists an integer  $L^\sharp(k) = L_{a,b,q}^\sharp(k)$  such that*

$$\frac{\mathbb{P}_{a,b}^q(X_{2k} = 2(\ell + 1))}{\mathbb{P}_{a,b}^q(X_{2k} = 2\ell)} \leq 1 \iff \ell \geq L^\sharp(k). \quad (3)$$

*Proof.* Writing the ratio

$$\begin{aligned} \frac{\mathbb{P}_{a,b}^q(X_{2k} = 2(\ell + 1))}{\mathbb{P}_{a,b}^q(X_{2k} = 2\ell)} &= \frac{Z_{k-\ell-1, k+\ell+1}(q) Z_{a-k+\ell+1, b-k-\ell-1}(q)}{Z_{k-\ell, k+\ell}(q) Z_{a-k+\ell, b-k-\ell}(q)} q^{(a-k+\ell+1)(k+\ell+1) - (a-k+\ell)(k+\ell)} \\ &= \frac{(1 - q^{k-\ell})(1 - q^{b-k-\ell})}{(1 - q^{k+\ell+1})(1 - q^{a-k+\ell+1})} q^{a+2\ell+1} \end{aligned}$$

This ratio is smaller than 1 if and only if

$$\begin{aligned} (1 - q^{k-\ell})(1 - q^{b-k-\ell})q^{a+2\ell+1} &\leq (1 - q^{k+\ell+1})(1 - q^{a-k+\ell+1}) \\ \iff (q^{a+b+1} - 1) &\leq q^{a+1}(q-1)q^{2\ell} + \left( q^{k+1}(q^a - 1) + q^{a-k+1}(q^b - 1) \right) q^\ell \end{aligned}$$

Both terms on the right hand side are increasing continuous functions of  $\ell$ , proving the existence of an integer  $L^\sharp(k)$  with the asserted property.  $\square$

**Proposition 2.** *The 2-dimensional marginal of  $(X_{2k})_{0 \leq k \leq n}$  is given by*

$$\mathbb{P}_{a,b}^q(X_{2k} = 2i, X_{2l} = 2j) = \frac{Z_{k+i, k-i}(q) Z_{j+l-i, l+i-k-j}(q) Z_{a-l+j, b-l-j}(q)}{Z_{a,b}(q)} q^{(a-(k-i))(k+i)+(l+j-k-i)(a-l+j)}.$$

Now we have exact expressions for probabilistic quantities, we let the dimensions of the box depend on  $n$  and investigate the limiting behaviour of the random Young diagram under  $\mathbb{P}_{a_n, b_n}^q$  in the regime

$$a_n + b_n = 2n \rightarrow +\infty, \quad \frac{a_n}{2n} \rightarrow \rho \in (0, 1), \quad -n \log q \rightarrow c \in \mathbb{R}.$$

### 3 Asymptotics of $q$ -factorials

In order to study the limit of the process as the size of the box goes to infinity, we first need the asymptotic behavior of the  $q$ -factorial.

**Proposition 3** ( $q$ -Stirling's Formula). *Let  $c \in \mathbb{R}_+^*$  and fix  $\varepsilon > 0$ . In the limit when  $n$  goes to infinity, with  $q = e^{-\frac{c}{n}}$ , the following asymptotics hold for all  $\ell > \varepsilon n$*

$$\ell!_q = \sqrt{2\pi n} \sqrt{\frac{e^{\frac{c\ell}{n}} - 1}{c}} n^\ell \exp(nS_c(\frac{\ell}{n})) (1 + O(\frac{1}{\ell})),$$

where

$$S_c(\alpha) = \int_0^\alpha \log\left(\frac{1 - e^{-cx}}{c}\right) dx = \alpha S_{ac}(1).$$

In particular, there exist positive constants  $m$  and  $M$  such that for all  $n$ , and all  $\ell$  between 1 and  $n$ ,

$$m\sqrt{2\pi n} \sqrt{\frac{e^{\frac{c\ell}{n}} - 1}{c}} n^\ell \exp(nS_c(\frac{\ell}{n})) < \ell!_q < M\sqrt{2\pi n} \sqrt{\frac{e^{\frac{c\ell}{n}} - 1}{c}} n^\ell \exp(nS_c(\frac{\ell}{n})).$$

*Proof.* Consider  $\log \ell!_q = \sum_{k=1}^\ell \log\left(\frac{1 - e^{-\frac{ck}{n}}}{c}\right) - \ell \log\left(\frac{1 - e^{-\frac{c}{n}}}{c}\right)$ . By the Euler-Maclaurin formula, we have

$$\begin{aligned} \log \ell!_q &= \int_1^\ell \log\left(\frac{1 - e^{-\frac{c}{n}t}}{c}\right) dt + \frac{1}{2} \left( \log\left(\frac{1 - e^{-\frac{c\ell}{n}}}{c}\right) + \log\left(\frac{1 - e^{-\frac{c}{n}}}{c}\right) \right) \\ &\quad + \int_1^\ell B_1(t) \frac{\frac{c}{n} e^{-\frac{c}{n}t}}{1 - e^{-\frac{c}{n}t}} dt - \ell \log\left(\frac{1 - e^{-\frac{c}{n}}}{c}\right) \end{aligned} \quad (4)$$

where  $B_1(t) = t - [t] - \frac{1}{2}$  is the first periodic Bernoulli polynomial. The first term is

$$\begin{aligned} \int_1^\ell \log \left( \frac{1 - e^{-\frac{c}{n}t}}{c} \right) dt &= nS_c\left(\frac{\ell}{n}\right) - \int_0^1 \log \left( \frac{1 - e^{-\frac{c}{n}t}}{c} \right) dt \\ &= nS_c\left(\frac{\ell}{n}\right) + 1 + \log n + O\left(\frac{1}{n}\right). \end{aligned} \quad (5)$$

Next, the terms involving  $\log \left( \frac{1 - e^{-\frac{c}{n}}}{c} \right)$ :

$$-(\ell - \frac{1}{2}) \log \left( \frac{1 - e^{-\frac{c}{n}}}{c} \right) = (\ell - \frac{1}{2}) \log n + \frac{c\ell}{2n} + O\left(\frac{1}{n}\right). \quad (6)$$

For the second integral in (4), we add and subtract  $\frac{1}{t}$  to get

$$\int_1^\ell B_1(t) \frac{\frac{c}{n}e^{-\frac{c}{n}t}}{1 - e^{-\frac{c}{n}t}} dt = \int_1^\ell B_1(t) \frac{1}{t} dt + c \int_{\frac{1}{n}}^{\frac{\ell}{n}} B_1(nt) \left( \frac{e^{-ct}}{1 - e^{-ct}} - \frac{1}{ct} \right) dt \quad (7)$$

The function  $f(t) = \frac{e^{-t}}{1 - e^{-t}} - \frac{1}{t}$  is continuous on  $[0, 1]$  and its derivative is bounded, say  $|f'(t)| \leq M$  for all  $t$ . On the interval  $[\frac{k-1}{n}, \frac{k}{n}]$ , the function  $B_1(nt)$  is given by  $B_1(nt) = nt - k + \frac{1}{2}$  and has the antiderivative  $\tilde{B}_{n,k}(t) = \frac{n}{2}t^2 - (k - \frac{1}{2})t + \frac{k(k-1)}{2n}$  with  $\tilde{B}_{n,k}(\frac{k-1}{n}) = \tilde{B}_{n,k}(\frac{k}{n}) = 0$ . By partial integration we get

$$c \left| \int_{\frac{1}{n}}^{\frac{\ell}{n}} B_1(nt) f(ct) dt \right| \leq c^2 M \sum_{k=1}^{\ell-1} \int_{\frac{k}{n}}^{\frac{k+1}{n}} |\tilde{B}_{n,k}(t)| dt = \frac{Mc^2\ell}{12n^2}$$

proving that this term is  $O(\frac{1}{n})$  as  $n \rightarrow \infty$ . The first term of (7) is found to be

$$\begin{aligned} \int_1^\ell B_1(t) \frac{1}{t} dt &= \ell - 1 - (\ell + \frac{1}{2}) \log \ell + \log \ell! \\ &= -1 + \frac{1}{2} \log 2\pi + O\left(\frac{1}{\ell}\right) \end{aligned} \quad (8)$$

by the classical Stirling approximation  $\ell! = \sqrt{2\pi\ell} \left(\frac{\ell}{e}\right)^\ell (1 + O(\frac{1}{\ell}))$ . Combining (5), (6) and (8) and the term  $\frac{1}{2} \log \left( \frac{1 - e^{-\frac{c\ell}{n}}}{c} \right)$  from (4), we have

$$\log \ell!_q = nS_c\left(\frac{\ell}{n}\right) + (\ell + \frac{1}{2}) \log n - (\ell + \frac{1}{2}) \log c + \frac{1}{2} \log \left( \frac{1 - e^{-\frac{c\ell}{n}}}{c} \right) + \frac{c\ell}{2n} + \frac{1}{2} \log 2\pi + O\left(\frac{1}{\ell}\right)$$

proving the proposition.  $\square$

We turn now to the asymptotics of the 1- and 2-dimensional marginals of  $X$  under  $\mathbb{P}$  given by Propositions 1 and 2. For this purpose, define

$$f_c(x, y) = S_c(x + y) - S_c(x) - S_c(y) \quad \text{and} \quad h_c(x, y) = \sqrt{\frac{c(e^{c(x+y)} - 1)}{(e^{cx} - 1)(e^{cy} - 1)}}.$$

If  $x = \frac{\ell}{n}$  and  $y = \frac{k}{n}$ , the asymptotics of the  $q$ -binomial coefficient is given by

$$Z_{l,k}^{(q)} = \binom{\ell+k}{\ell}_q = \sqrt{\frac{1}{2\pi n}} h_c(x, y) \exp(n f_c(x, y)) (1 + O(\frac{1}{n})).$$

As a consequence,

**Corollary 1.** *For all  $\varepsilon > 0$ , for all  $k$  between  $\varepsilon n$  and  $(1-\varepsilon)n$ , with  $s = k/n$  and  $x = i/n$ , we have*

$$\mathbb{P}(X_{2k} = 2i) = \sqrt{\frac{1}{2\pi n}} H_{\rho,c}^{(1)}(s, x) \exp(n F_{\frac{\rho}{2n},c}^{(1)}(s, x)) (1 + O(\frac{1}{n})),$$

and there exists a constant  $M > 0$  such that for all  $n, i$  and  $k$ ,

$$\mathbb{P}(X_{2k} = 2i) \leq M \sqrt{\frac{1}{2\pi n}} H_{\rho,c}^{(1)}(s, x) \exp(n F_{\rho,c}^{(1)}(s, x)),$$

where

$$\begin{aligned} F_{\rho,c}^{(1)}(s, x) &= -c(2\rho - s + x)(s + x) + f_c(s - x, s + x) + f_c(2\rho - s + x, 2 - 2\rho - s - x) - f_c(2\rho, 2 - 2\rho), \\ H_{\rho,c}^{(1)}(s, x) &= \frac{h_c(s + x, s - x) h_c(2\rho - s + x, 2 - 2\rho - s - x)}{h_c(2\rho, 2 - 2\rho)}. \end{aligned}$$

In the first equality, when replacing  $\frac{\rho}{2n}$  by  $\rho$  in the indices of  $H^{(1)}$ , we make an error of order  $O(\frac{1}{n})$ , which is absorbed in the factor  $(1 + O(\frac{1}{n}))$ . Making the same substitution in the indices of  $F^{(1)}$  would change the multiplicative constant in the asymptotics.

**Corollary 2.** *For all  $\varepsilon > 0$ , for all  $k < l$  between  $\varepsilon n$  and  $(1-\varepsilon)n$  then, with  $s = k/n$ ,  $t = l/n$  and  $x = i/n$ ,  $y = j/n$*

$$\mathbb{P}(X_{2k} = 2i, X_{2l} = 2j) = \frac{1}{2\pi n} H_{\rho,c}^{(2)}(s, t, x, y) \exp(n F_{\frac{\rho}{2n},c}^{(2)}(s, t, x, y)) (1 + O(\frac{1}{n})),$$

where

$$\begin{aligned} F_{\rho,c}^{(2)}(s, t, x, y) &= -c((2\rho - s + x)(s + x) + (t - s + y - x)(2\rho - t + y)) \\ &\quad + f_c(s - x, s + x) + f_c(t - s + y - x, t - s - y + x) \\ &\quad + f_c(2\rho - t + y, 2 - 2\rho - t - y) - f_c(2\rho, 2 - 2\rho), \\ H_{\rho,c}^{(2)}(s, t, x, y) &= \frac{h_c(s + x, s - x) h_c(t - s + y - x, t - s - y + x) h_c(2\rho - t + y, 2 - 2\rho - t - y)}{h_c(2\rho, 2 - 2\rho)}. \end{aligned}$$

We give now a slight refinement of Corollary 2 that will be useful in the fluctuations.

**Corollary 3.** *For all  $\varepsilon > 0$ , for all  $k < l$  between  $\varepsilon n$  and  $(1-\varepsilon)n$  then, with  $s = k/n$ ,  $t = l/n$  and  $x = i/n$ ,  $y = j/n$ , for all  $s' = s + o(1)$ ,  $t' = t + o(1)$ ,  $x' = x + o(1)$ ,  $y' = y + o(1)$ :*

$$\mathbb{P}_n(X_{2k} = 2i, X_{2l} = 2j) = \frac{1}{2\pi n} H_{\rho,c}^{(2)}(s', t', x', y') \exp(n F_{\frac{\rho}{2n},c}^{(2)}(s, t, x, y)) (1 + o(1)).$$

## 4 Limit Shape

We associate to the lattice path  $(X_k)$  the continuous piecewise linear function,  $X : s \mapsto X_s$  defined on  $[0, 2n]$ , which coincides with  $X_k$  when  $s = k$ . The graph of the function  $X$  is the boundary of the random Young diagram we consider.

Let  $L_{\rho,c}$  be the function on  $[0, 1]$  defined by

$$\begin{aligned} \forall t \in [0, 1], \quad L_{\rho,c}(t) &= 1 - 2\rho + \frac{1}{c} \log \frac{\sinh(ct) + e^{c(2\rho-1)} \sinh(c(1-t))}{\sinh c} \\ &= \frac{1}{c} \log \frac{e^{-ct} - e^{ct} + e^{-c(2-2\rho-t)} - e^{-c(t-2\rho)}}{e^{-c(2-2\rho)} - e^{2c\rho}}. \end{aligned} \quad (9)$$

This is the limit shape of the rescaled random process, in the sense made precise in Theorem 1 below. First we need to estimate the proximity between the value of the function  $L_{\rho,c}$  and the most probable value for  $X_{2k}$  denoted by  $L_n^\sharp(k) = L_{a_n, b_n, e^{-c/n}}(k)$  in Lemma 2.

**Lemma 3.** *For all  $n$  and all  $0 \leq k \leq n$ ,  $\left| \frac{1}{n} L_n^\sharp(k) - L_{\rho,c}\left(\frac{k}{n}\right) \right| \leq \frac{1}{n} + \left| \frac{a_n}{2n} - \rho \right|$ , where  $q = e^{-\frac{c}{n}}$ .*

Consequently, in the limit when  $n$  goes to infinity, and  $a_n/2n$  goes to  $\rho$ ,

$$\left| \frac{1}{n} L_n^\sharp(k) - L_{\rho,c}\left(\frac{k}{n}\right) \right| \rightarrow 0. \quad (10)$$

*Proof.* The ratio of probabilities (3) used to define  $L^\sharp(k) = L_n^\sharp(k)$  can be rewritten in terms of the function

$$(\rho, c, t, x) \mapsto R_{\rho,c}(t, x) = e^{-c} \frac{\sinh\left(\frac{c}{2}(t-x)\right) \sinh\left(\frac{c}{2}(2-2\rho-t-x)\right)}{\sinh\left(\frac{c}{2}(t+x)\right) \sinh\left(\frac{c}{2}(2\rho-t+x)\right)},$$

namely as  $R_{\frac{a_n}{2n}, c}\left(\frac{k}{n}, \frac{\ell}{n}\right)$ . The function  $x \mapsto R_{\rho,c}(t, x)$  is decreasing, and by definition  $L_n^\sharp(k)$  is the smallest integer  $\ell$  such that  $R_{\frac{a_n}{2n}, c}\left(\frac{k}{n}, \frac{\ell}{n}\right) \leq 1$ .

On the other hand, a computation shows that for each  $t \in [0, 1]$ , the equation

$$R_{\rho,c}(t, x) = 1$$

has the (unique) solution  $x = L_{\rho,c}(t)$ . Substituting  $\frac{a_n}{2n}$  for  $\rho$ , we conclude that

$$\left| \frac{1}{n} L_n^\sharp(k) - L_{\frac{a_n}{2n}, c}\left(\frac{k}{n}\right) \right| \leq \frac{1}{n} \quad (11)$$

for all  $n$  and all  $k \leq n$ .

Differentiating  $L_{\rho,c}$  with respect to  $\rho$ , we find that

$$\left| \frac{\partial L_{\rho,c}}{\partial \rho}(t) \right| \leq 1$$

for all  $t$ , and so by the mean value theorem,

$$\left| L_{\frac{a_n}{2n}, c}\left(\frac{k}{n}\right) - L_{\rho,c}\left(\frac{k}{n}\right) \right| \leq \left| \frac{a_n}{2n} - \rho \right| \quad (12)$$

for all  $k \leq n$ . From (11), (12) and the triangle inequality, we get the expected result.  $\square$

*Remark 1.* It turns out that  $\log R_{\rho,c}(t, x)$  coincides with the derivative of  $F_{\rho,c}^{(1)}(s, x)$  with respect to  $x$ . As a consequence,  $L_{\rho,c}(s)$  can also be viewed as the argmax of  $F_{\rho,c}^{(1)}(s, \cdot)$ , which is nonpositive and vanishes at  $L_{\rho,c}(s)$ . See Corollary 5.

**Theorem 1.** *The boundary of the rescaled random Young diagram converges in probability, for the uniform topology, to the curve of  $t \mapsto L_{\rho,c}(t)$ .*

$$\forall \varepsilon > 0, \lim_{n \rightarrow 0} \mathbb{P}_{a_n, b_n}^{e^{-c/n}} \left( \sup_{t \in [0,1]} \left| \frac{1}{2n} X_{2tn} - L_{\rho,c}(t) \right| > \varepsilon \right) = 0,$$

where

$$L_{\rho,c}(t) = \frac{1}{c} \log \frac{e^{-ct} - e^{ct} + e^{-c(2-2\rho-t)} - e^{-c(t-2\rho)}}{e^{-c(2-2\rho)} - e^{2c\rho}}.$$

Remark that in the case of a square box ( $\rho = 1/2$ ), the expression for  $L_{\rho,c}$  boils down to:

$$L_{\frac{1}{2},c}(t) = \frac{1}{c} \log \frac{\cosh(c(t - \frac{1}{2}))}{\cosh \frac{c}{2}}$$

*Proof.* Fix  $\varepsilon > 0$ . For  $t < \frac{\varepsilon}{2}$  or  $t > 1 - \frac{\varepsilon}{2}$ , the difference  $|\frac{1}{2n} X_{2tn} - L_{\rho,c}(t)|$  is always smaller than  $\varepsilon$ . We have to control what happens for  $t \in (\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2})$ . Using the fact that  $L_{\rho,c}$  is differentiable, and its derivative with respect to  $t$  is bounded by 1, we have that

$$\left| L_{\rho,c} \left( \frac{1}{n} \lfloor tn \rfloor \right) - L_{\rho,c}(t) \right| \leq \frac{1}{n},$$

which is smaller than  $\frac{\varepsilon}{3}$  for  $n$  sufficiently large. The same is true of  $|\frac{1}{2n} X_{2tn} - \frac{1}{2n} X_{2 \lfloor tn \rfloor}|$ . Thus, by an  $\frac{\varepsilon}{3}$ -argument, to control the sup over  $(\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2})$ , it is sufficient to control what happens at points of the form  $t = \frac{k}{n}$ :

$$\begin{aligned} & \mathbb{P} \left( \sup_{t \in (\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2})} \left| \frac{1}{2n} X_{2tn} - L_{\rho,c}(t) \right| > \varepsilon \right) \\ & \leq \sum_{k \in \mathbb{N} \cap n(\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2})} \mathbb{P}(X_{2k} > 2n(L_{\rho,c}(\frac{k}{n}) + \varepsilon) \text{ or } X_{2k} < 2n(L_{\rho,c}(\frac{k}{n}) - \varepsilon)) \end{aligned} \quad (13)$$

But it follows from (10), that for  $n$  sufficiently large,  $2n(L_{\rho,c}(\frac{k}{n}) + \varepsilon) \geq 2L_n^\#(k)$  for all  $k \in \mathbb{N} \cap n(\frac{\varepsilon}{2}, 1 - \frac{\varepsilon}{2})$ , and thus using the unimodality of the law of  $X_{2k}$

$$\begin{aligned} \mathbb{P}(X_{2k} > 2n(L_{\rho,c}(\frac{k}{n}) + \varepsilon)) &= \sum_{l > n(L_{\rho,c}(\frac{k}{n}) + \varepsilon)} \mathbb{P}(X_{2k} = 2l) \\ &\leq n \times \mathbb{P}(X_{2k} = 2 \lfloor n(L_{\rho,c}(\frac{k}{n}) + \varepsilon) \rfloor). \end{aligned}$$

which by Corollary 1 is exponentially small, uniformly in  $k$ , as  $n$  goes to infinity. Thus for  $n$  sufficiently large, the sum on the RHS of (13) can be made smaller than any positive number.  $\square$

*Remark 2.* From the proof and Lemma 3, we see that the convergence statement in the theorem holds uniformly for a family of sequences  $(a_n)$  and  $(b_n)$ , so long as the convergence  $\frac{a_n}{2n} \rightarrow \rho$  is uniform for the family.

## 5 Fluctuations

We now study the fluctuations of the interface around the limit shape. We define for  $n \in \mathbb{N}^*$  a new rescaled process:

$$\forall t \in [0, 1], \quad \tilde{X}_t = \tilde{X}_t^{(n)} = \sqrt{n} \left( \frac{1}{2n} X_{2nt} - L_{\rho,c}(t) \right). \quad (14)$$

We place ourselves in the space  $D$  of càd-làg paths on  $[0, 1]$  endowed with its usual topology. We state now our second main result for the convergence of the fluctuations of the interface to the Ornstein Uhlenbeck bridge. See the appendix for the definition and some properties of this process.

Define

$$f(s) = \frac{1}{\sqrt{2}} \frac{\sqrt{(e^{2c\rho} - 1)(1 - e^{-2c(1-\rho)})}}{\sinh cs + e^{c(2\rho-1)} \sinh(c(1-s))} = \frac{\sqrt{2 \sinh(c\rho) \sinh(c(1-\rho))}}{e^{c(\frac{1}{2}-\rho)} \sinh(cs) + e^{c(\rho-\frac{1}{2})} \sinh(c(1-s))}.$$

**Theorem 2.** *The sequence  $(\tilde{X}_s^{(n)} / f(s))_n$ , converges weakly in  $D$  to the Ornstein-Uhlenbeck bridge  $(Y_s)_{s \in [0,1]}$ , which is the Gaussian process on  $[0, 1]$  with covariance*

$$\mathbb{E}[Y_s Y_t] = \frac{\sinh(cs) \sinh(c(1-t))}{c \sinh(c)},$$

for  $0 \leq s < t \leq 1$ .

### 5.1 Two-point correlations

To prove convergence of the two-dimensional marginal to a Gaussian process, we apply a saddle-point method, ie. we need to show that the function  $F^{(2)}$ , which governs the exponential decay of 2-dimensional marginals (see Corollary 2), has a critical point ‘on the limit shape’, and that it takes its maximal value of 0 at this point. We will prove that this is indeed the case, but first we need a lemma describing the limit shape in a subrectangle.

**Lemma 4.** *The limit shape  $L$  satisfies the relations*

$$\frac{1}{1-s} (L_{\rho,c}(t) - L_{\rho,c}(s)) = L_{\rho',c'} \left( \frac{t-s}{1-s} \right) \quad (15)$$

and

$$\frac{1}{t} L_{\rho,c}(s) = L_{\rho'',c''} \left( \frac{s}{t} \right) \quad (16)$$

for  $0 < s < t < 1$ , with

$$\rho' = \frac{2\rho - s + L_{\rho,c}(s)}{2(1-s)}, \quad c' = c(1-s)$$

and

$$\rho'' = \frac{t - L_{\rho,c}(t)}{2t}, \quad c'' = tc.$$

*Proof.* These relations can be checked analytically. We sketch now a less computational argument for Equation (16). The same strategy applies for Equation (15). Fix  $t \in (0, 1)$ , and take a sequence of boxes with sidelengths  $[0, a_n] \times [0, b_n]$  (with  $a_n + b_n = 2n$ ,  $a_n/2n \rightarrow \rho$ ). Theorem 1 states that, when  $n$  goes to infinity, the probability under  $\mathbb{P}_{a_n, b_n}^{e^{-c/n}}$  that  $X_{nt}/n$  converges to  $L_{\rho,c}(t)$  goes to 1. As a consequence, the restriction of the limit shape  $L_{\rho,c}$  to the, rescaled by  $n$ , lower left subbox  $[0, \frac{t-L_{\rho,c}(t)}{2}] \times [0, \frac{t+L_{\rho,c}(t)}{2}]$  is the rescaled limit shape for boxes with ratio  $\frac{t-L_{\rho,c}(t)}{t-L_{\rho,c}(t)+t+L_{\rho,c}(t)} = \frac{t-L_{\rho,c}(t)}{2t}$ . The rescaling factor is the perimeter of the subbox  $t$ . The value of the parameter  $q$  remains the same. Since  $q = e^{-c/n} = e^{-\frac{c''}{nt}}$ , we get  $c'' = tc$ .  $\square$

This lemma basically describes the limit shape of the process restricted to a subrectangle defined by a point  $(s, L_{\rho,c}(s))$  and either the right or left corner of the original rectangle.

**Lemma 5.** For all  $0 < s < t < 1$ , the functions  $F^{(1)}$  and  $F^{(2)}$  satisfy the relations

$$F_{\rho,c}^{(2)}(s, t, x, y) = F_{\rho,c}^{(1)}(s, x) + (1-s)F_{\rho',c'}^{(1)}(t', y') \quad (17)$$

and

$$F_{\rho,c}^{(2)}(s, t, x, y) = F_{\rho,c}^{(1)}(t, y) + tF_{\rho'',c''}^{(1)}(s'', x'') \quad (18)$$

with

$$\rho' = \frac{2\rho - s + x}{2(1-s)}, \quad t' = \frac{t-s}{1-s}, \quad y' = \frac{y-x}{1-s}, \quad c' = c(1-s)$$

and

$$\rho'' = \frac{t-y}{2t}, \quad s'' = \frac{s}{t}, \quad x'' = \frac{x}{t}, \quad c'' = tc.$$

*Proof.* The relation (17) follows from the identity

$$\mathbb{P}(X_{2k} = 2i, X_{2\ell} = 2j) = \mathbb{P}(X_{2k} = 2i)\mathbb{P}(X_{2\ell} = 2j | X_{2k} = 2i).$$

allied to scaling arguments similar to those used in the proof of the previous lemma. Equation (18) follows from the conditioning on  $X_{2\ell}$  instead of  $X_{2k}$ .  $\square$

**Lemma 6.** The partial derivatives of  $F^{(1)}$  with respect to  $\rho$ ,  $s$  and  $x$  all vanish at  $x = L_{\rho,c}(s)$ .

**Corollary 4.** *The partial derivatives of  $F^{(2)}$  with respect to  $\rho$ ,  $s$ ,  $t$ ,  $x$  and  $y$  all vanish at  $x = L_{\rho,c}(s)$  and  $y = L_{\rho,c}(t)$ .*

*Proof.* For the partial derivatives with respect to  $\rho$ , use (18). For  $\frac{\partial F^{(2)}}{\partial s}$  and  $\frac{\partial F^{(2)}}{\partial x}$ , use (18) together with (16). Similarly, for  $\frac{\partial F^{(2)}}{\partial t}$  and  $\frac{\partial F^{(2)}}{\partial y}$ , use (17) together with (15).  $\square$

**Corollary 5.**  $F_{\rho,c}^{(1)}(s, L_{\rho,c}(s)) = 0$  and  $F_{\rho,c}^{(2)}(s, t, L_{\rho,c}(s), L_{\rho,c}(t)) = 0$ .

*Proof.* By the chain rule and Lemma 6 we get

$$\frac{d}{ds}(F_{\rho,c}^{(1)}(s, L_{\rho,c}(s))) = 0.$$

Taking the limit  $s \rightarrow 0$  yields zero, proving the claim. For the second identity, use the first one together with (18) and (16).  $\square$

**Proposition 4.** *Assume  $\frac{an}{2n} = \rho + o(\frac{1}{\sqrt{n}})$ . Let  $s < t$  in  $[0, 1]$ . The joint law of  $(\tilde{X}_s^{(n)}, \tilde{X}_t^{(n)})$  converges to the law of the centered 2-dimensional Gaussian vector*

$$(\tilde{X}_s, \tilde{X}_t) = (f(s)Y_s, f(t)Y_t),$$

where

$$f(s) = \frac{e^{cs} \sqrt{2(e^{2c\rho} - 1)(1 - e^{-c(2-2\rho)})}}{-1 + e^{2cs} - e^{-c(2-2\rho-2s)} + e^{2c\rho}} = \frac{1}{\sqrt{2}} \frac{\sqrt{(e^{2c\rho} - 1)(1 - e^{-2c(1-\rho)})}}{\sinh cs + e^{c(2\rho-1)} \sinh(c(1-s))},$$

and  $(Y_t)_{t \in [0,1]}$  is an Ornstein-Uhlenbeck bridge on the interval  $[0, 1]$  with parameter  $c$  (see Appendix).

*Remark 3.* In the case of a square ( $\rho = \frac{1}{2}$ ), the expression of  $f$  simplifies drastically to

$$f(x) = \frac{1}{\sqrt{2} \cosh(c(s - \frac{1}{2}))}.$$

*Proof.* Let  $a_1, b_1, a_2, b_2$  be four real numbers such that  $a_1 < b_1$  and  $a_2 < b_2$ . In the following, we write  $L$  for  $L_{\rho,c}$ . Clearly,  $\tilde{X}_s^{(n)} \in [a_1, b_1]$  if and only if  $X_{2ns}$  is in the interval

$$2nL(s) + [2a_1\sqrt{n}, 2b_1\sqrt{n}].$$

Since  $|X_{2ns} - X_{2\lfloor ns \rfloor}| \leq 2$  and  $|X_{2nt} - X_{2\lfloor nt \rfloor}| \leq 2$ , we just need to compute the limit of

$$\mathbb{P}((X_{2\lfloor ns \rfloor}, X_{2\lfloor nt \rfloor}) \in A_1 \times A_2).$$

as  $n$  goes to infinity.

Set  $k_n = \lfloor ns \rfloor$  and  $\ell_n = \lfloor nt \rfloor$ . One has:

$$\begin{aligned} \mathbb{P}((X_{2\lfloor ns \rfloor}, X_{2\lfloor nt \rfloor}) \in A_1 \times A_2) &= \sum_{(i,j) \in \frac{1}{2}A_1 \times \frac{1}{2}A_2} \mathbb{P}(X_{2k_n} = 2i, X_{2\ell_n} = 2j) \\ &= \sum_{(i,j) \in \frac{1}{2}A_1 \times \frac{1}{2}A_2} \frac{1}{2\pi n} H_{\rho,c}^{(2)}(s, t, L(s), L(t)) \exp\left(n F_{\frac{an}{2n},c}^{(2)}\left(\frac{k_n}{n}, \frac{\ell_n}{n}, \frac{i}{n}, \frac{j}{n}\right)\right) (1 + o(1)). \end{aligned}$$

Here,  $F^{(2)} \leq 0$  everywhere and the number of terms is  $O(n)$ , so the error terms  $o(1)$  can be replaced with an  $o(1)$  term outside the sum. Thus, we are left with a Riemann sum for the double integral

$$\frac{1}{2\pi n} H_{\rho,c}^{(2)}(s, t, L(s), L(t)) \int_{nL(s)+a\sqrt{n}}^{nL(s)+b\sqrt{n}} \int_{nL(t)+c\sqrt{n}}^{nL(t)+d\sqrt{n}} \exp \left[ nF_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \frac{u}{n}, \frac{v}{n} \right) \right] dv du \quad (19)$$

in which we make the substitution

$$u = x\sqrt{n} + nL(s), \quad v = y\sqrt{n} + nL(t)$$

to get

$$\frac{1}{2\pi} H_{\rho,c}^{(2)}(s, t, L(s), L(t)) \int_a^b \int_c^d \exp \left[ nF_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \frac{x}{\sqrt{n}} + L(s), \frac{y}{\sqrt{n}} + L(t) \right) \right] dy dx$$

We make a second degree Taylor expansion of the function  $F_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \cdot, \cdot \right)$  at the point  $(L(s), L(t))$ . Let  $z_n$  be the point  $(\frac{k_n}{n}, \frac{\ell_n}{n}, L(s), L(t))$  in  $\mathbb{R}^4$ .

$$\begin{aligned} & nF_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \frac{x}{\sqrt{n}} + L(s), \frac{y}{\sqrt{n}} + L(t) \right) \\ &= nF_{\frac{a_n}{2n},c}^{(2)}(z_n) + \frac{\partial F_{\frac{a_n}{2n},c}^{(2)}}{\partial x}(z_n)x\sqrt{n} + \frac{\partial F_{\frac{a_n}{2n},c}^{(2)}}{\partial y}(z_n)y\sqrt{n} \\ &+ \frac{1}{2} \frac{\partial^2 F_{\frac{a_n}{2n},c}^{(2)}}{\partial x^2}(z_n)x^2 + \frac{1}{2} \frac{\partial^2 F_{\frac{a_n}{2n},c}^{(2)}}{\partial y^2}(z_n)y^2 + \frac{\partial^2 F_{\frac{a_n}{2n},c}^{(2)}}{\partial x \partial y}(z_n)xy + O(n^{-\frac{1}{2}}) \end{aligned} \quad (20)$$

We must find the limit of this expression as  $n \rightarrow \infty$ . The first term on the right has limit zero, which can be seen by Taylor expansion around  $(s, t, L(s), L(t))$ , using Corollaries 4 and 5.

Since  $F^{(2)}$  is an analytic function, we have in general

$$F_{\rho+o(n^{-\frac{1}{2}}),c}^{(2)} \left( s + o(n^{-\frac{1}{2}}), t + o(n^{-\frac{1}{2}}), x, y \right) = F_{\rho,c}^{(2)}(s, t, x, y) + o(n^{-\frac{1}{2}}) \quad (21)$$

and similarly for all partial derivatives of  $F^{(2)}$ . Since  $\frac{a_n}{2n} = \rho + o(n^{-\frac{1}{2}})$ ,  $\frac{k_n}{n} = s + o(n^{-\frac{1}{2}})$  and  $\frac{\ell_n}{n} = t + o(n^{-\frac{1}{2}})$ , this shows that the second and third terms on the right in (20) tend to zero as  $n \rightarrow \infty$ . Thus, if  $z = (s, t, L(s), L(t))$ ,

$$nF_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \frac{x}{\sqrt{n}} + L(s), \frac{y}{\sqrt{n}} + L(t) \right) \rightarrow \frac{1}{2} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x^2}(z)x^2 + \frac{1}{2} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial y^2}(z)y^2 + \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x \partial y}(z)xy \quad (22)$$

as  $n \rightarrow \infty$ . From (21) it also follows that the sequence  $F_{\frac{a_n}{2n},c}^{(2)} \left( \frac{k_n}{n}, \frac{\ell_n}{n}, \cdot, \cdot \right)$  is equicontinuous, hence by Ascoli's theorem that the sequence converges uniformly, hence that the double integral (19) converges to

$$\frac{1}{2\pi} H_{\rho,c}^{(2)}(s, t, L(s), L(t)) \int_a^b \int_c^d \exp \left[ \frac{1}{2} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x^2}(z)x^2 + \frac{1}{2} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial y^2}(z)y^2 + \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x \partial y}(z)xy \right] dy dx$$

To find these double derivatives, we use (17) and (18). Hence we need

$$\frac{\partial^2 F_{\rho,c}^{(1)}}{\partial x^2}(s, L(s)) = c(1 - e^{-2c})e^{2cs} \frac{(1 - e^{-c(2-2\rho)} - e^{-2cs} + e^{-c(2s-2\rho)})^2}{(1 - e^{2c\rho})(1 - e^{-2cs})(1 - e^{-c(2-2\rho)})(1 - e^{-c(2-2s)})}$$

and exploit the fact that

$$\begin{aligned} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x^2}(s, t, L(s), L(t)) &= \frac{1}{t} \frac{\partial^2 F_{\rho'',c''}^{(1)}}{\partial x^2}(s'', L_{\rho'',c''}(s'')) \\ \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial y^2}(s, t, L(s), L(t)) &= \frac{1}{1-s} \frac{\partial^2 F_{\rho',c'}^{(1)}}{\partial x^2}(t', L_{\rho',c'}(t')) \end{aligned}$$

This way, we find the double derivatives of  $F_{\rho,c}^{(2)}(s, t, \cdot, \cdot)$ , evaluated at the critical point  $(x, y) = (L(s), L(t))$ , to be

$$\begin{aligned} \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x^2}(s, t, L(s), L(t)) &= \frac{c(1 - e^{2ct})(1 - e^{2cs} + e^{-c(2-2\rho-2s)} - e^{2c\rho})^2}{(1 - e^{2cs})(e^{2ct} - e^{2cs})(1 - e^{2c\rho})(1 - e^{-c(2-2\rho)})} \\ &= \frac{-c \sinh ct}{\sinh cs \sinh c(t-s) f(s)^2}, \\ \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial y^2}(s, t, L(s), L(t)) &= \frac{c(e^{2cs} - e^c)(1 - e^{2ct} + e^{-c(2-2\rho-2t)} - e^{2c\rho})^2}{(e^{2cs} - e^{2ct})(e^c - e^{2ct})(1 - e^{2c\rho})(1 - e^{-c(2-2\rho)})} \\ &= \frac{-c \sinh c(1-s)}{\sinh c(1-t) \sinh c(t-s) f(t)^2} \\ \frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x \partial y}(s, t, L(s), L(t)) &= \frac{c(1 - e^{2cs} + e^{-c(2-2\rho-2s)} - e^{2c\rho})(1 - e^{2ct} + e^{-c(2-2\rho-2t)} - e^{2c\rho})}{(e^{2cs} - e^{2ct})(1 - e^{2c\rho})(1 - e^{-c(2-2\rho)})} \\ &= \frac{c}{\sinh c(t-s) f(s) f(t)} \end{aligned}$$

The last one, the mixed derivative, was calculated from scratch, i.e. by calculating

$$\frac{\partial^2 F_{\rho,c}^{(2)}}{\partial x \partial y}(s, t, x, y) = \frac{c(e^{2ct} - e^{2cs})}{(e^{c(t+y)} - e^{c(s+x)})(e^{c(t-y)} - e^{c(s-x)})}$$

and evaluating at the critical point  $(x, y) = (L(s), L(t))$ . Thus, the Hessian of  $F_{\rho,c}^{(2)}(s, t, \cdot, \cdot)$  at the critical point is

$$\begin{aligned} H(F_{\rho,c}^{(2)}) &= \frac{-c}{\sinh cs \sinh c(1-t) \sinh c(t-s)} \times \\ &\begin{pmatrix} \frac{1}{f(s)} & 0 \\ 0 & \frac{1}{f(t)} \end{pmatrix} \begin{pmatrix} \sinh ct \sinh c(1-t) & -\sinh cs \sinh c(1-t) \\ -\sinh cs \sinh c(1-t) & \sinh cs \sinh c(1-s) \end{pmatrix} \begin{pmatrix} \frac{1}{f(s)} & 0 \\ 0 & \frac{1}{f(t)} \end{pmatrix}. \end{aligned}$$

The covariance matrix of the limiting Gaussian distribution is the negative of the inverse of  $H(F^{(2)})$ , which we compute to be

$$\Sigma = \frac{1}{c \sinh c} \begin{pmatrix} f(s) & 0 \\ 0 & f(t) \end{pmatrix} \cdot \begin{pmatrix} \sinh cs \sinh c(1-s) & \sinh cs \sinh c(1-t) \\ \sinh cs \sinh c(1-t) & \sinh ct \sinh c(1-t) \end{pmatrix} \cdot \begin{pmatrix} f(s) & 0 \\ 0 & f(t) \end{pmatrix} \quad (23)$$

The matrix in the middle together with the factor  $\frac{1}{c \sinh c}$  is the covariance matrix of the Ornstein-Uhlenbeck bridge on the interval  $[0, 1]$  with parameter  $c$  (see Appendix). Further computations reveal that

$$\frac{1}{\sqrt{\det(\Sigma)}} = H_{\rho,c}^{(2)}(s, t, L(s), L(t))$$

since both sides equal

$$\frac{c}{f(s)f(t)} \sqrt{\frac{\sinh c}{\sinh cs \sinh c(t-s) \sinh c(1-t)}}.$$

This completes the proof.  $\square$

We explain now that for  $m$  points  $0 < t_1 < t_2 < \dots < t_m < 1$ , the limit of the corresponding  $m$ -dimensional marginal  $(\tilde{X}_{t_1}^{(n)}, \tilde{X}_{t_2}^{(n)}, \dots, \tilde{X}_{t_m}^{(n)})$  is Gaussian, with covariance matrix  $\Sigma = (\sigma_{i,j})$  defined by

$$\Sigma_{i,j} = \frac{1}{c \sinh c} f(t_i) f(t_j) \sinh ct_i \sinh c(1-t_j)$$

for all  $i \leq j$ .

A first approach would be to repeat similar computations as for the 2-dimensional case, and observe that the matrix  $\sigma$  is the inverse of the  $m$ -dimensional Hessian of  $F_{\rho,c}^{(m)}(t_1, \dots, t_m)$ .

Another way takes advantage of the Markov property satisfied by the process  $(\tilde{X}_n)$ . Denote by  $P_{s,t}^{(n)}(x, y)$  the transition kernel of this Markov chain. The law of the  $m$ -tuple  $(\tilde{X}_{t_1}^{(n)}, \tilde{X}_{t_2}^{(n)}, \dots, \tilde{X}_{t_m}^{(n)})$  is given by

$$P_{0,t_1}^{(n)}(0, x_1) P_{t_1,t_2}^{(n)}(x_1, x_2) \dots P_{t_{m-1},t_m}^{(n)}(x_{m-1}, x_m).$$

Proposition 4 implies the convergence of each  $\sqrt{n} P_{t_i,t_{i+1}}^{(n)}(x_n, y_n)$  to the density of the kernel of the Ornstein-Uhlenbeck bridge as  $n \rightarrow \infty$ , and  $(x_n), (y_n)$  converge. The convergence in law follows from Lebesgue dominated convergence theorem. Domination is ensured by Lemma 8 below and unimodality stated in Lemma 2.

## 5.2 Tightness and proof of Theorem 2

We already proved the convergence of the finite-dimensional distributions. We need to show that the sequence of the distribution is tight. For this we use Theorem 13.5 p.142 of [3]. The criterion is checked in Lemma 10.

Before entering the proof, we need a geometric definition.

**Definition 2.** Let  $ABCD$  be a rectangle with sides having slopes  $\pm 1$  in the  $(s, x)$  coordinates. For  $\varepsilon \in (0, 1)$ , the  $\varepsilon$ -parallelogram of the rectangle  $ABCD$  is the unique parallelogram with diagonal  $AC$  and sides with slopes  $1 - \varepsilon$  and  $-1 + \varepsilon$  in the  $(s, x)$  coordinates. The  $\varepsilon$ -interior of the rectangle is the interior of the  $\varepsilon$ -parallelogram. The complement of the  $\varepsilon$ -interior in the rectangle is called the  $\varepsilon$ -boundary of the rectangle. See Figure 5.2. For  $s_0 \in (0, 1)$ , the sides of the parallelogram intersect the straight line  $s = s_0$  in two points. We denote the ordinates of these two points by  $g_\varepsilon^+(s_0)$  and  $g_\varepsilon^-(s_0)$ .

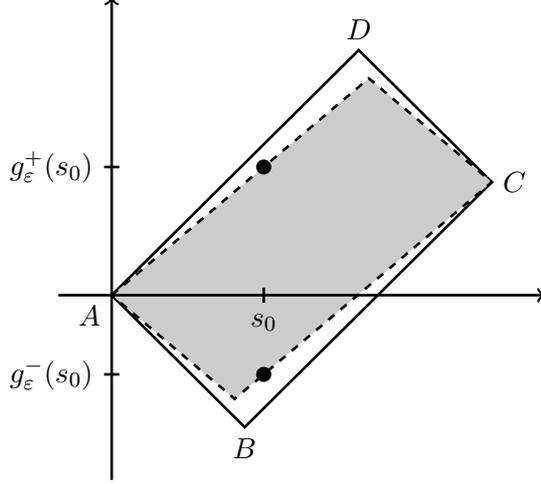


Figure 4: The  $\varepsilon$ -parallelogram of a rectangle.

We first state the following useful fact.

**Fact 1.** For all  $\delta \in (0, 1)$ , for all  $A > 0$ , there exists  $\varepsilon$ , such that for all rectangular boxes of side lengths  $a, b$  satisfying  $\rho = \frac{a}{a+b} \in (\delta, 1 - \delta)$ , and all  $c \in [-A, A]$ , the limit shape  $L_{\rho,c}$  is entirely included in the  $\varepsilon$ -interior of the box.

The following lemma controls the function  $F_{\rho,c}^{(1)}$  in an  $\varepsilon$ -boundary of a rectangular box with side lengths  $\rho$  and  $1 - \rho$ .

**Lemma 7.** Let  $\delta \in (0, 1)$  and  $A > 0$ . Take  $\varepsilon$  as in Fact 1. Then there exists a positive constant  $C$  such that for all  $\rho \in [\delta, 1 - \delta]$ , and all  $c \in [-A, A]$ ,

$$\forall (s, x) \in B_\rho^\varepsilon, \quad F_{\rho,c}^{(1)}(s, x) \leq -Cs(1 - s),$$

where  $B_\rho$  is the (macroscopic) box with perimeter 1 and aspect ratio  $\rho$ .

*Proof.* For a fixed  $s$ , the function  $x \mapsto F_{\rho,c}^{(1)}(s, x)$  is concave, and its maximum is reached at  $x = L_{\rho,c}(s)$ , and the point  $(s, L_{\rho,c}(s))$  is inside the  $\varepsilon$ -interior of the box. Therefore  $F_{\rho,c}^{(1)}(s, x) \leq \max\{F_{\rho,c}^{(1)}(s, g_\varepsilon^+(s)), F_{\rho,c}^{(1)}(s, g_\varepsilon^-(s))\}$ .

In a neighborhood of 0,  $g_\varepsilon^+(s) = (1 - \varepsilon)s$ , and  $s \mapsto F_{\rho,c}^{(1)}(s, (1 - \varepsilon)s) \asymp s$  because it is differentiable and vanishes at  $s = 0$  with a non zero derivative. The same argument

applied to a neighborhood of  $s = 1$ , and for  $g_\varepsilon^-$  gives the result. The uniform bound on  $C$  results from a compactness argument.  $\square$

**Lemma 8** (L.U.F.). *Let  $\delta \in (0, 1)$  and  $A > 0$ . Take  $\varepsilon$  as in Fact 1. For this  $\varepsilon$ , there exist two constants  $M$  and  $\kappa$  such that if  $a + b = 2n$  and  $\frac{a}{a+b} \in (\delta, 1 - \delta)$  and for all  $c \in [-A, A]$ , then*

$$\forall 0 \leq k \leq n, \quad \mathbb{P}_{a,b}^{e^{-c/n}}(X_{2k} = 2 \lfloor nL_{\rho,c}(\frac{k}{n}) + y\sqrt{n} \rfloor) \leq \frac{M}{\sqrt{n}} \frac{\exp\left(-\frac{y^2}{2\kappa s(1-s)}\right)}{\sqrt{2\pi\kappa s(1-s)}}.$$

as soon as  $(2k, \lfloor nL_{\rho,c}(\frac{k}{n}) + y\sqrt{n} \rfloor)$  is in the  $\varepsilon$ -interior of the box.

*Proof.* Use Corollary 1. Notice that  $h_c(x, y) \asymp \sqrt{\frac{x+y}{xy}}$  uniformly in  $c$ . Therefore

$$H_{\rho,c}(s, x) \asymp \sqrt{\rho(1-\rho)} \sqrt{\frac{s}{(s-x)(s+x)}} \sqrt{\frac{1-s}{(2\rho-s+x)(2-2\rho-s-x)}}.$$

The first term is bounded. The second term is of order  $\frac{1}{\sqrt{s}}$  in a neighborhood of  $(0, 0)$  in the  $\varepsilon$ -interior of the box. A simple change of variable  $(s, x) \mapsto (1-s, 2\rho-1-x)$  exchanges the second and the third term, which shows that the third term is of order  $\frac{1}{\sqrt{1-s}}$  in a neighborhood of  $(1, 1-\rho)$  in the  $\varepsilon$ -interior of the box. Therefore,

$$H_{\rho,c}(s, x) \asymp \frac{1}{\sqrt{s(1-s)}},$$

as long as  $(s, x)$  is in the  $\varepsilon$ -interior of the box.

In order to bound the exponential term in Corollary 1, we bound from below the absolute value of  $\frac{\partial^2 F_{\rho,c}^{(1)}(s, x)}{\partial x^2}$ . Using that  $S_c''(u) = \frac{c}{e^{cu}-1}$ , we get

$$\frac{\partial^2 F_{\rho,c}^{(1)}(s, x)}{\partial x^2} = -2c - \frac{c}{e^{c(s+x)}-1} - \frac{c}{e^{c(s-x)}-1} - \frac{c}{e^{c(2\rho-s+x)}-1} - \frac{c}{e^{c(2-2\rho-s-x)}-1}.$$

For  $s$  close to 0, the main contribution comes from the second and the third term, that are both negative, and of order  $s^{-1}$ . For  $s$  close to 1, the main contribution comes from the fourth and the fifth term, that are both negative, and of order  $(1-s)^{-1}$ .  $\square$

**Lemma 9.** *Let  $\delta \in (0, 1)$  and  $A > 0$ , and  $\varepsilon \in (0, 1)$  given by Fact 1. Then there exists a constant  $C > 0$  such that: for all  $n \geq 1$ , for all  $c \in [-A, A]$ , for all sequences of boxes  $(B_n)$  with sides  $a_n, b_n$  such that  $a_n + b_n = 2n$ ,  $\frac{a_n}{2n} \in (\delta, 1 - \delta)$ ,*

$$\forall \lambda > 0, \forall s \in [0, 1], \quad \mathbb{P}(\tilde{X}_s \notin [-\lambda, \lambda]; (X_{2ns}, 2ns) \in B_n^\varepsilon) \leq C \frac{(s(1-s))^2}{\lambda^4}$$

*Proof.* Lemma 8 gives the following bound:

$$\mathbb{P}(\tilde{X}_s \notin [-\lambda, \lambda]; (X_{2ns}, 2ns) \in B_n^\varepsilon) \leq \sum_{|y| \geq \lambda} \frac{M}{\sqrt{n}} \frac{\exp\left(-\frac{y^2}{2\kappa s(1-s)}\right)}{\sqrt{2\pi\kappa s(1-s)}},$$

where the index of the sum  $y$  runs on the set  $\frac{1}{\sqrt{n}}\mathbb{Z} - \sqrt{n}L_{\rho,c}(s)$ .

Comparing the sum on the right-hand side and the corresponding integral, we get

$$\mathbb{P}(\tilde{X}_s \notin [-\lambda, \lambda]; (X_{2ns}, 2ns) \in B_n^\varepsilon) \leq M' \int_{\mathbb{R} \setminus (-\lambda, \lambda)} \frac{\exp\left(-\frac{y^2}{2\kappa s(1-s)}\right)}{\sqrt{2\pi\kappa s(1-s)}} ds.$$

This last integral is equal to  $\mathbb{P}(|\mathcal{N}| \geq \frac{\lambda}{\sqrt{\kappa s(1-s)}})$ , where  $\mathcal{N}$  is a standard Gaussian variable. Conclude by using Markov inequality for the fourth moment.  $\square$

We can now verify Billingsley's condition for tightness [3].

**Lemma 10.** *Let  $((a_n, b_n))_n$  be a sequence such that  $a_n + b_n = 2n$ , and  $\frac{a_n}{2n} = \rho + o(\frac{1}{\sqrt{n}})$ . Then there exists a constant  $C$  such that for all  $n > 0$ , for all  $0 \leq r \leq s \leq t \leq 1$ , for all  $\lambda > 0$ ,*

$$\mathbb{P}_{a_n, b_n}^{e^{-c/n}}(|\tilde{X}_s - \tilde{X}_r| \geq \lambda; |\tilde{X}_t - \tilde{X}_s| \geq \lambda) \leq \frac{C(t-r)^2}{\lambda^4}. \quad (24)$$

*Proof.* Inequality (24) is automatically satisfied, as soon as  $|r-s|$  or  $|t-s|$  is less than  $1/n$ , or  $\lambda$  is greater than  $\sqrt{n}(|r-s| \wedge |t-s|)$ . We suppose now that none of these conditions are satisfied.

Introduce now the three following  $\varepsilon$ -interiors:  $B^\varepsilon$ , the  $\varepsilon$ -interior of the box  $B := [0, a_n] \times [0, b_n]$ ,  $B_L^\varepsilon$  the  $\varepsilon$ -interior of the box  $B_L := [0, ns - \frac{1}{2}X_{2sn}] \times [0, sn + \frac{1}{2}X_{2sn}]$  and  $B_R^\varepsilon$  the  $\varepsilon$ -interior of the box  $B_R := [ns - \frac{1}{2}X_{2sn}, a_n] \times [sn + \frac{1}{2}X_{2sn}, b_n]$ . See Fig. 5.

Let us first control the probability that the random interface exits the  $\varepsilon$ -interior of the box. We use unimodality of the distribution of  $X_{2ns}$ , Corollary 1, and Lemma 7 to get:

$$\begin{aligned} \mathbb{P}((2ns, X_{2ns}) \notin B^\varepsilon) &\leq 4\varepsilon ns(1-s) \mathbb{P}(X_{2ns} = ng_\pm^\varepsilon(s)) \\ &\leq M \frac{ns(1-s)}{\sqrt{n}} H_{\rho,c}(s, g_\pm^\varepsilon(s)) \exp(-nC s(1-s)). \end{aligned}$$

Using the uniform bound of  $H_{\rho,c} \asymp \frac{1}{\sqrt{s(1-s)}}$  obtained in the beginning of the proof of Lemma 8, we get

$$\mathbb{P}((2ns, X_{2ns}) \notin B^\varepsilon) \leq M \sqrt{ns(1-s)} \exp(-nC(s(1-s))) \leq \frac{M}{(nC s(1-s))^2}.$$

Using our assumption that  $n \geq \frac{\lambda^2}{|s-r|^2 \wedge |t-s|^2}$ , we obtain:

$$\mathbb{P}((2ns, X_{2ns}) \notin B^\varepsilon) \leq M \frac{|r-s|^4 \wedge |t-s|^4}{\lambda^4 s^2 (1-s)^2} \leq M \frac{|r-s|^2 \wedge |t-s|^2}{\lambda^4} \leq M \frac{|t-r|^2}{\lambda^4}. \quad (25)$$

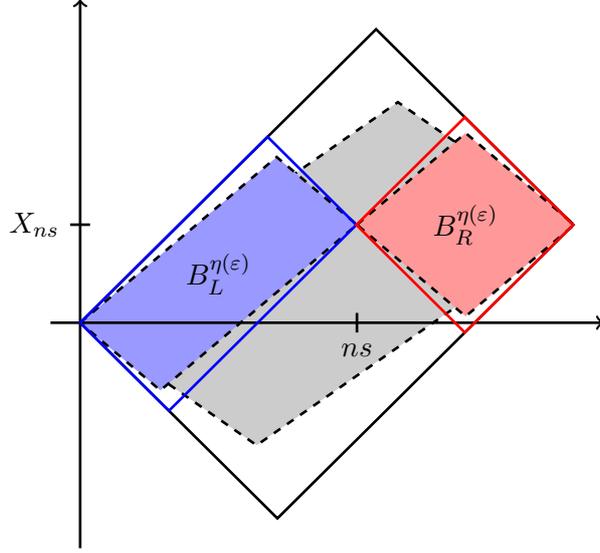


Figure 5: The  $\epsilon$ -interiors  $B^\epsilon$  (grey),  $B_L^{\eta(\epsilon)}$  (blue) and  $B_R^{\eta(\epsilon)}$  (red).

*Remark 4.* The weaker condition  $n \geq \frac{\lambda^2}{s^2 \wedge (1-s)^2}$  yields

$$\mathbb{P}((2ns, X_{2ns}) \notin B^\epsilon) \leq M \frac{s^2 \wedge (1-s)^2}{\lambda^4}. \quad (26)$$

Now we consider three cases depending on the values of  $r$  and  $t$ .

**Case 1:**  $r < \frac{\lambda}{4\sqrt{n}}$ . Then  $|\tilde{X}_r| \leq 2r\sqrt{n} < \frac{\lambda}{2}$ . Thus,

$$\begin{aligned} \mathbb{P}(|\tilde{X}_s - \tilde{X}_r| \geq \lambda; |\tilde{X}_t - \tilde{X}_s| \geq \lambda) &\leq \mathbb{P}_{a_n, b_n}^{e^{-c/n}}(|\tilde{X}_s - \tilde{X}_r| \geq \lambda) \leq \mathbb{P}(|\tilde{X}_s| \geq \frac{\lambda}{2}) \\ &\leq \mathbb{P}(|\tilde{X}_s| \geq \frac{\lambda}{2}, (ns, X_{ns}) \in B^\epsilon) + \mathbb{P}((ns, X_{ns}) \notin B^\epsilon). \end{aligned}$$

But we just saw in (25) that

$$\mathbb{P}((2ns, X_{2ns}) \notin B^\epsilon) \leq M \frac{|t-r|^2}{\lambda^4}.$$

Moreover, according to Lemma 9,

$$\mathbb{P}(|\tilde{X}_s| \geq \frac{\lambda}{2}, (2ns, X_{2ns}) \in B^\epsilon) \leq C \frac{s^2(1-s)^2}{\lambda^4} \leq \frac{Cs^2}{\lambda^4}.$$

Since  $r < \frac{\lambda}{4\sqrt{n}} < \frac{s}{4}$ , we get that  $s < \frac{4}{3}(s-r)$ , and thus

$$\mathbb{P}(|\tilde{X}_s| \geq \frac{\lambda}{2}, (2ns, X_{2ns}) \in B^\epsilon) \leq \frac{C(s-r)^2}{\lambda^4} \leq \frac{C(t-r)^2}{\lambda^4}.$$

And the tightness condition (24) is satisfied in that case.

**Case 2:**  $1 - t < \frac{\lambda}{4\sqrt{n}}$ . This case is treated similarly as the previous one, and (24) is again satisfied.

**Case 3:**  $r \geq \frac{\lambda}{4\sqrt{n}}$  and  $1 - t \geq \frac{\lambda}{4\sqrt{n}}$ . This is the generic situation.

Conditional on  $\{(2ns, X_{2ns}) \in B^\varepsilon\}$ , the boxes  $B_L$  and  $B_R$  have an aspect ratio in the interval  $(\varepsilon, 1 - \varepsilon)$ . Applying Lemma 9 with  $\delta = \varepsilon$  in the subboxes  $B_L$  and  $B_R$ , we get the existence of an  $\eta(\varepsilon)$  such that with sufficient probability  $(2nr, X_{2nr})$  (resp.  $(2nt, X_{2nt})$ ) is in  $B_L^{\eta(\varepsilon)}$  (resp.  $B_R^{\eta(\varepsilon)}$ ). More precisely, the assumptions on  $r$  and  $t$  allow to repeat the argument to derive (26) with the proper scaling in the subboxes  $B_L$  and  $B_R$  and get for some constant  $M$ :

$$\mathbb{P}((2nr, X_{2nr}) \notin B_L^{\eta(\varepsilon)} | (2ns, X_{2ns}) \in B^\varepsilon) \leq M \frac{\left(\frac{r}{s}\right)^2 \wedge \left(\frac{s-r}{s}\right)^2}{(\lambda/\sqrt{s})^4} \leq M \frac{(s-r)^2}{\lambda^4}, \quad (27)$$

$$\mathbb{P}((2nt, X_{2nt}) \notin B_R^{\eta(\varepsilon)} | (2ns, X_{2ns}) \in B^\varepsilon) \leq M \frac{\left(\frac{t-s}{1-s}\right)^2 \wedge \left(\frac{1-t}{1-s}\right)^2}{(\lambda/\sqrt{1-s})^4} \leq M \frac{(t-s)^2}{\lambda^4}. \quad (28)$$

By the Markov property of  $X$ , the variables  $X_r$  and  $X_t$  are independent conditional on the value of  $X_s$ . Hence we can write

$$\begin{aligned} \mathbb{P}(|\tilde{X}_s - \tilde{X}_r| \geq \lambda; |\tilde{X}_t - \tilde{X}_s| \geq \lambda) &= \\ &= \sum_{\substack{j,k,l \\ |j-k| \geq \lambda \\ |l-k| \geq \lambda}} \mathbb{P}(\tilde{X}_s = k) \mathbb{P}(\tilde{X}_r = j | \tilde{X}_s = k) \mathbb{P}(\tilde{X}_t = l | \tilde{X}_s = k), \end{aligned} \quad (29)$$

where  $j, k, l$  runs through all possible values for  $\tilde{X}_r, \tilde{X}_s, \tilde{X}_t$  (two successive values of  $j, k, l$  differ by  $1/\sqrt{n}$ ).

Let  $E_{r,s,t}^\varepsilon = \{(2ns, X_{2ns}) \in B^\varepsilon\} \cap \{(2nr, X_{2nr}) \in B_L^{\eta(\varepsilon)}\} \cap \{(2nt, X_{2nt}) \in B_R^{\eta(\varepsilon)}\}$ . From Inequalities (25), (27), (28), there exists a constant  $M$  such that:

$$1 - \mathbb{P}(E_{r,s,t}^\varepsilon) \leq M \frac{(t-r)^2}{\lambda^4}.$$

As a consequence, it suffices to bound the sum in (29) for values of  $(j, k, l)$  such that

$$\{\tilde{X}_r = j, \tilde{X}_s = k, \tilde{X}_t = l\} \subset \{|\tilde{X}_s - \tilde{X}_r| \geq \lambda\} \cap \{|\tilde{X}_t - \tilde{X}_s| \geq \lambda\} \cap E_{r,s,t}^\varepsilon.$$

By a scaling argument, when looking at what happens on the left of  $s$ ,

$$\mathbb{P}_{a_n, b_n}^{e^{-c/n}}(\tilde{X}_r = j | X_{ns}) = \mathbb{P}_{\frac{2ns-X_{2ns}}{2}, \frac{2ns+X_{2ns}}{2}}^{e^{-c/n}} \left( \tilde{X}_{\frac{r}{s}} = \frac{j}{\sqrt{s}} + \sqrt{\frac{n}{s}} L_{\rho,c}(r) - \sqrt{ns} L_{\frac{2ns-X_{2ns}}{4ns}, sc} \left( \frac{r}{s} \right) \right).$$

Note that on the right hand side,  $\tilde{X}_{r/s}$  is defined with respect to the limit shape inside the box  $[0, \frac{2ns-X_{2ns}}{4ns}] \times [0, \frac{2ns+X_{2ns}}{4ns}]$ . See Fig. 6.

When summing over the values of  $j$  such that  $|j - \tilde{X}_s| \geq \lambda$ , one can get an upper bound for  $\mathbb{P}_{a_n, b_n}(|\tilde{X}_r - \tilde{X}_s| \geq \lambda | \tilde{X}_s)$  using Lemma 9 as long as the limit shape in the

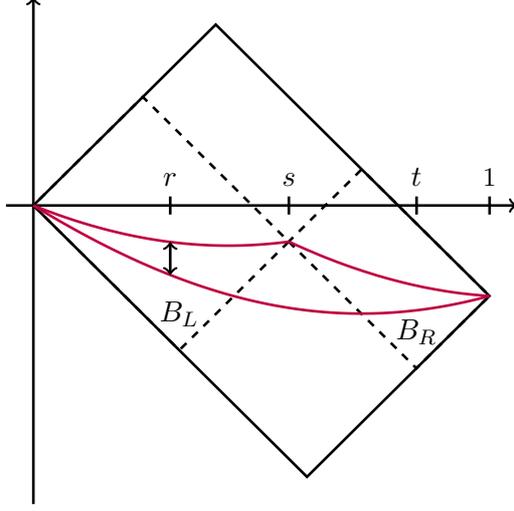


Figure 6: Limit shapes in the box  $B$  and the subboxes  $B_L$  and  $B_R$ . The double arrow represents the difference of the limit shapes in  $B$  and  $B_L$  at time  $r$ , whose expression,  $2nL_{\rho,c}(r) - 2nsL_{\frac{2ns-X_{2ns}}{4ns},sc}(\frac{r}{s})$ , appears in the sticking condition (30).

original box  $B$  and the one in  $B_L$  are sufficiently close to each other. We express this proximity by the following sticking condition for  $B_L$ :

$$\frac{1}{\sqrt{s}}[\tilde{X}_s - \lambda, \tilde{X}_s + \lambda] + \sqrt{\frac{n}{s}}L_{\rho,c}(r) - \sqrt{ns}L_{\frac{2ns-X_{2ns}}{2ns},sc}(\frac{r}{s}) \supset \left[-\frac{\lambda}{\sqrt{2s}}, \frac{\lambda}{\sqrt{2s}}\right]. \quad (30)$$

Indeed, when the sticking condition (30) is satisfied,  $\mathbb{P}_{a_n, b_n}^{e^{-c/n}}(|\tilde{X}_r - \tilde{X}_s| \geq \lambda|\tilde{X}_s)$  takes the form:

$$\begin{aligned} & \mathbb{P}_{\frac{2ns-X_{2ns}}{2}, \frac{2ns+X_{2ns}}{2}}^{e^{-c/n}} \left( \tilde{X}_r \notin \frac{1}{\sqrt{2s}}[\tilde{X}_s - \lambda, \tilde{X}_s + \lambda] + \sqrt{\frac{n}{s}}L_{\rho,c}(r) - \sqrt{ns}L_{\frac{2ns-X_{2ns}}{4ns},sc}(\frac{r}{s}) \right) \\ & \leq \mathbb{P}_{\frac{2ns-X_{2ns}}{2}, \frac{2ns+X_{2ns}}{2}}^{e^{-c/n}} \left( \tilde{X}_r \notin \left[-\frac{\lambda}{\sqrt{2s}}, \frac{\lambda}{\sqrt{2s}}\right] \right) \leq M \frac{(\frac{r}{s} - \frac{s-r}{s})^2}{(\lambda/\sqrt{2s})^4} \leq 4M \frac{(t-r)^2}{\lambda^4} \end{aligned}$$

by Lemma 9 in the small box.

We follow the same idea for the right box  $B_R$ . Whenever the sticking condition for  $B_R$  is satisfied,

$$\mathbb{P}_{a_n, b_n}^{e^{-c/n}}(|\tilde{X}_t - \tilde{X}_s| \geq \lambda|\tilde{X}_s) \leq 4M \frac{(t-r)^2}{\lambda^4}.$$

Therefore, conditional on one of the sticking conditions to be satisfied, the probability of  $\{|\tilde{X}_s - \tilde{X}_r| \geq \lambda\} \cap \{|\tilde{X}_t - \tilde{X}_s| \geq \lambda\}$  is bounded by  $4M \frac{(t-r)^2}{\lambda^4}$ .

Let us now investigate the probability that both of the sticking conditions fail. It is bounded by the probability that the one for  $B_L$  fails.

Using Equation (16) of Lemma 4, we replace  $\frac{1}{s}L_{\rho,c}(r)$  in (30) by  $L_{\frac{s-L_{\rho,c}(s)}{2s},sc}(\frac{r}{s})$ .

Condition (30) is equivalent to

$$\left| \tilde{X}_s + s\sqrt{n} \left( L_{\frac{s-L_{\rho,c}(s)}{2s}, sc} \left( \frac{r}{s} \right) - L_{\frac{s-L_{\rho,c}(s)-\tilde{X}_s n^{-1/2}}{2s}, sc} \left( \frac{r}{s} \right) \right) \right| \leq \frac{\lambda}{2}.$$

By the mean value theorem applied to  $\rho \mapsto L_{\rho, sc}(r/s)$ , there exists a  $\bar{\rho}$  in the interval  $(\frac{s-L_{\rho,c}(s)}{2s}, \frac{s-L_{\rho,c}(s)-\tilde{X}_s n^{-1/2}}{2s})$  such that

$$s\sqrt{n} \left( L_{\frac{s-L_{\rho,c}(s)}{2s}, sc} \left( \frac{r}{s} \right) - L_{\frac{s-L_{\rho,c}(s)-\tilde{X}_s n^{-1/2}}{2s}, sc} \left( \frac{r}{s} \right) \right) = \frac{1}{2} \tilde{X}_s \left. \frac{\partial L_{\rho, cs}(r/s)}{\partial \rho} \right|_{\rho=\bar{\rho}}.$$

But differentiating (9) with respect to  $\rho$ , for generic values of  $\rho, c, t$  yields

$$\begin{aligned} 0 \leq 1 + \frac{1}{2} \frac{\partial L_{\rho, c}(t)}{\partial \rho} &= \frac{e^{c(2\rho-1)} \sinh(c(1-t))}{\sinh(ct) + e^{c(2\rho-1)} \sinh(c(1-t))} \\ &\leq 1 \wedge e^{c(2\rho-1)} \frac{\sinh(c(1-t))}{\sinh(ct)} \leq 1 \wedge K \frac{1-t}{t} \end{aligned}$$

with a constant  $K$  which works for all  $c \in [-A, A]$ , all  $\rho \in (0, 1)$ , and all  $t \in [0, 1]$ .

As a consequence, the sticking condition (30) is satisfied as soon as

$$|\tilde{X}_s| (1 \wedge K \frac{s-r}{r}) \leq \frac{\lambda}{2}.$$

Therefore the probability that (30) is not verified is less than

$$P(|\tilde{X}_s| (1 \wedge K \frac{s-r}{r}) > \frac{\lambda}{2}, (2sn, X_{2sn}) \in B^\varepsilon),$$

which, by Lemma 9 is bounded by  $\lambda^{-4} s^2 (1-s)^2 (1 \wedge K^4 (\frac{s-r}{r})^4)$ . This bound is less than some constant times  $\lambda^{-4} (s-r)^2 \leq \lambda^{-4} (t-r)^2$ , as one can check in both regimes  $(\frac{s-r}{r}) \leq K$ . This finishes the proof of Lemma 10.  $\square$

## 6 Link with the infinite shape

When writing this paper, we discovered that Petrov recently discussed in [7] the link between the limit shapes obtained above and the infinite shape of the unbounded problem. We explain it in this short section.

The family of limit shapes obtained above has the following property: if we fix  $s_0 \in (0, 1)$  and take the point  $(s_0, L_{\rho, c}(s_0))$  as the right corner of a new family of bounding boxes, then the limit shape of this new problem is simply the restriction of  $L_{\rho, c}$  to the interval  $[0, s_0]$ , rescaled by  $s_0$  (See Lemma 4).

We are therefore led to the idea that there exists some ‘inverse limit’ of these curves, from which they can all be recovered by restriction. Indeed, this is the case. The universal limit shape is Vershik’s curve [9]

$$L_\infty(s) = \frac{2\pi}{\sqrt{3}} \log 2 \cosh \left( \frac{\sqrt{3}s}{2\pi} \right),$$

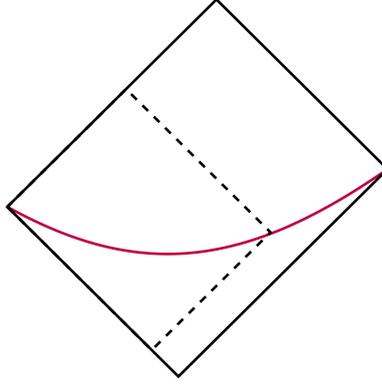


Figure 7: The limit shape in a subbox.

the solution of  $e^{-\frac{\sqrt{3}}{2\pi}(x-s)} + e^{-\frac{\sqrt{3}}{2\pi}(x+s)} = 1$ . For every  $\rho \in (0, 1)$  and  $c \in \mathbb{R}_+^*$ , one can find two real numbers  $s_0 < s_1$  such that the restriction of  $L_\infty$  to  $[s_0, s_1]$  is (up to an affine transformation)  $L_{\rho,c}$ , more precisely.

$$\forall t \in [0, 1], \quad L_{\rho,c}(t) = \frac{L_\infty(s_0 + t(s_1 - s_0)) - L_\infty(s_0)}{s_1 - s_0}. \quad (31)$$

For given values of  $s_0 < s_1$ , the parameter  $\rho$  is given by

$$\rho = \frac{1}{2} - \frac{L_\infty(s_1) - L_\infty(s_0)}{2(s_1 - s_0)},$$

and  $c$  is found for example by comparing the slopes of the limit shape in the corner and identifying  $\frac{\partial L_\infty}{\partial s} \Big|_{s=s_0} (s_1 - s_0)$  and  $\frac{\partial L_{\rho,c}}{\partial s} \Big|_{s=0} = \frac{e^{c(2\rho-1)} - \cosh c}{\sinh c}$ , which is an increasing function of  $c$ .

The limit case  $c \rightarrow 0$  correspond to  $s_1 \rightarrow s_0$ . For negative values of  $c$  one can take advantage of the symmetry of the model.

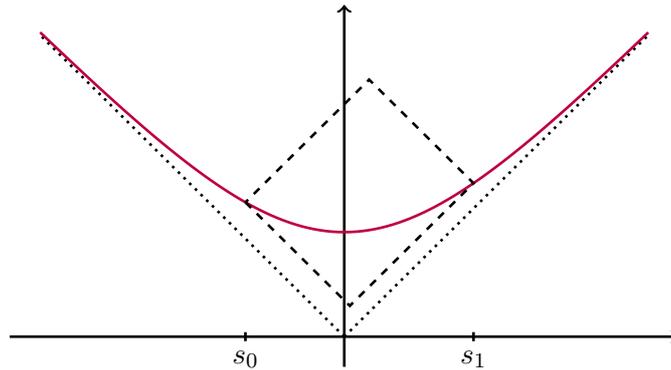


Figure 8: Limit shapes are restrictions of the universal curve.

## 7 Concluding remarks and open questions

Another twin model of random partitions in a box is obtained by taking for the probability measure  $\mathbb{P}$  the uniform measure on diagrams with a given area. It can be shown that in that context, the limit shape is one of the curves  $t \mapsto L_{\rho,c}(t)$  for a suitable value of  $c$ .

When the diagrams fill half of the box, the corresponding value of  $c$  is 0, and the limit shape is a straight line. However we expect the fluctuations to be of a different nature. Indeed, the integral of the fluctuations should be 0 almost-surely, whereas the integral of the Brownian bridge is different from 0 almost surely. We conjecture that in that case, the fluctuations would be described by the Brownian bridge on  $[0, 1]$  conditioned to have an integral equal to 0. In terms of a standard Brownian motion  $(B_t)$ , this process  $(Y_t)$  can be expressed as

$$\forall t \in [0, 1], \quad Y_t = B_t - (3t^2 - 2t)B_1 - 6(t - t^2) \int_0^1 B_s ds. \quad (32)$$

Formula (32) can be easily obtained in the same way as the formula for the Ornstein-Uhlenbeck bridge (see Appendix). It is a special case of the general tool developed by Biane and Yor [2] for extensions of Lévy's formula about the characteristic function of the Lévy area.

We are not able to formulate an analogous conjecture when  $c \neq 0$ .

## 8 Appendix: The Ornstein-Uhlenbeck Bridge

In this short appendix, we give a short description of the Ornstein-Uhlenbeck bridge. Let  $(B_t)$  be a Brownian motion on  $[0, \infty)$ . The Ornstein-Uhlenbeck process  $(Z_t)$  is the Gaussian random process defined by

$$Z_t = B_t - c \int_0^t Z_s dB_s = e^{-ct} \int_0^t e^{cs} dB_s. \quad (33)$$

Using the right-hand side of (33), we can also represent  $Z_t$  as:

$$Z_t = e^{-ct} B_{\int_0^t e^{cs} ds} = e^{-ct} B_{\frac{e^{2ct}-1}{2c}}. \quad (34)$$

From (34) we derive the covariance. For  $0 \leq s \leq t$ ,

$$\mathbb{E}[Z_s Z_t] = e^{-c(s+t)} \frac{e^{2cs} - 1}{2c} = e^{-ct} \frac{\sinh(cs)}{c}. \quad (35)$$

'Tying down' this process at  $t = 1$  gives the Ornstein-Uhlenbeck  $(Y_t)_{t \in [0,1]}$  bridge of length 1. Set

$$h(t) = \frac{\sinh(ct)}{\sinh c} \quad (36)$$

and define

$$Y_t = Z_t - h(t)Z_1. \quad (37)$$

Observe that, for  $0 \leq t \leq 1$ ,

$$\begin{aligned} \mathbb{E}[Y_t Z_1] &= \mathbb{E}[Z_t Z_1 - h(t)Z_1^2] \\ &= e^{-c} \frac{\sinh(ct)}{c} - \frac{\sinh(ct)}{\sinh c} e^{-c} \frac{\sinh c}{c} \\ &= 0, \end{aligned}$$

showing that  $Y_t$  is independent of  $Z_1$ . Now let  $f$  be any test function and consider

$$\begin{aligned} \mathbb{E}[f(X_t) | X_1 = 0] &= \mathbb{E}[f(Y_t + h(t)X_1) | X_1 = 0] \\ &= \mathbb{E}[f(Y_t) | X_1 = 0] \\ &= \mathbb{E}[f(Y_t)], \end{aligned}$$

where the last equality is a consequence of the independence just proved. Thus, the process  $(Y_t)$  has the probability law of the Ornstein-Uhlenbeck process conditional on  $X_1 = 0$ . Its covariance is

$$\begin{aligned} \mathbb{E}[Y_s Y_t] &= \mathbb{E}[Y_s(Z_t - h(t)Z_1)] = \mathbb{E}[Y_s Z_t] \\ &= \mathbb{E}[Z_s Z_t] - h(s)\mathbb{E}[Z_1 Z_t] = \frac{\sinh(cs)}{c \sinh c} (e^{-ct} \sinh c - e^{-c} \sinh(ct)) \\ &= \frac{\sinh(cs) \sinh(c(1-t))}{c \sinh c}, \end{aligned}$$

cf. the middle matrix in (23).

Notice that this covariance is also the Green function of a Brownian motion on  $[0, 1]$  killed an exponential rate  $c^2$  [4]. Let us mention also, that this process was used by C. Donati in her solution of Buffon-Synge problem concerning the typical distance between the extremities of a string of given length, thrown at random [5].

**Acknowledgements.** The authors are pleased to thank Nikolai Reshetikhin for having proposed the question solved in this paper, Marc Yor for having recognized the covariance we obtained as that of the Ornstein-Uhlenbeck bridge, and Richard Kenyon for showing us the way to get the limit shapes we found from the infinite limit shape. The work of the third author was partially supported by the ANR MEMEMO grant.

## References

- [1] George E. Andrews and Kimmo Eriksson. *Integer partitions*. Cambridge University Press, Cambridge, 2004.

- [2] Ph. Biane and M. Yor. Variations sur une formule de Paul Lévy. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2, suppl.):359–377, 1987.
- [3] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [4] T. Chan, David S. Dean, Kalvis M. Jansons, and L. C. G. Rogers. On polymer conformations in elongational flows. *Comm. Math. Phys.*, 160(2):239–257, 1994.
- [5] C. Donati-Martin. Le problème de Buffon-Synge pour une corde. *Adv. in Appl. Probab.*, 22(2):375–395, 1990.
- [6] Tadahisa Funaki and Makiko Sasado. Hydrodynamic limit for an evolutionary model on two-dimensional young diagrams. *Preprint*, 2009.
- [7] F Petrov. Limit shapes of young diagrams. two elementary approaches. *Journal of Mathematical Sciences*, 166(1), 2010.
- [8] A. Vershik and Yu. Yakubovich. The limit shape and fluctuations of random partitions of naturals with fixed number of summands. *Mosc. Math. J.*, 1(3):457–468, 472, 2001.
- [9] A. M. Vershik. Statistical mechanics of combinatorial partitions, and their limit configurations. *Funktsional. Anal. i Prilozhen.*, 30(2):19–39, 96, 1996.
- [10] Yu. V. Yakubovich. The central limit theorem for normalized Young diagrams of partitions into different summands. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 256(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 3):212–223, 267–268, 1999.