A Growth Model

in a Random Environment

Janko Gravner, UC Davis

Craig Tracy, UC Davis

Harold Widom, UC Santa Cruz

Oriented Digital Boiling (ODB)

Interface:

$$\{(x,y) : x \in \mathbb{Z}, y \le h_t(x)\}, t = 0, 1, 2, \ldots$$

Height function rules:

1. $h_t \le h_{t+1}$.
2. If $h_t(x-1) > h_t(x)$,
   then $h_{t+1}(x) = h_t(x-1)$.
3. Else, $h_{t+1}(x) = h_t(x) + 1$ with prob p_x.

Alternatively, toss p_x–coins in advance to get indep. Bernoulli rv's $\varepsilon_{x,t}$. Think of the points (x,t) for which $\varepsilon_{x,t} = 1$ as *marked*. Then

$$\mathbf{h_t(x) = \max\{h_{t-1}(x-1), h_{t-1}(x) + \varepsilon_{x,t-1}\}}.$$

We will assume that the initial state is $h_0(x) = 0$ if $x = 0$, otherwise $-\infty$.

Eventually: p_x i.i.d., with d.f. F.

Path description

A space–time point $(x, t), x \leq t$, has backwards lightcone:

$$\mathcal{L}(x,t) = \{(x', t') : 0 \leq x' \leq x, x' \leq t' < x' + t - x\}.$$

Let H be the longest sequence $(x_1, t_1), \ldots, (x_k, t_k)$ of marked points such that

1. $x_{i-1} \leq x_i$,
2. $x_i - x_{i-1} + 1 \leq t_i - t_{i-1}$.

Alternatively, let $m = t - x$ and $n = x + 1$, and A a random $m \times n$ matrix with Bernoulli entries $\varepsilon_{i,j}$, where $P(\varepsilon_{i,j} = 1) = p_j$. Label columns as usual, but rows started at the bottom. Then $H = H(m,n)$ is the *longest* sequence of 1's in A, with

column index non–decreasing and row index strictly increasing

Then

$$h_t(x) = H(m,n)$$

This is often called a *last passage property*. From now on, we formulate all the results for H, with $n = \alpha m$.

$$\begin{array}{cccccccc}
-\infty & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
-\infty & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
-\infty & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\
-\infty & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
-\infty & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
m \uparrow & n \rightarrow & & & & & &
\end{array}$$

“Remembrance of Things Past”

Ulam’s problem of estimating the longest increasing subsequence in a random permutation of length n.

Strong Law Type Results: Hammersley (1972), Logan–Shepp, Vershik–Kerov (1977), & Aldous–Diaconis (1995)

Fluctuations: Baik–Deift–Johansson (1999)

Methods: subadditivity, exclusion process representation, random Young tableaux, RMT techniques (including Riemann-Hilbert).

The largest increasing sequence in a random (0,1)–matrix: Seppäläinen (1998), limiting shape:

$$\lim_{t\to\infty} \frac{h_t(x)}{t}, \quad x/t \quad \text{constant}$$

Johansson (1999–2000) computed the fluctuations in (universal regime of) this limit law, by a RMT approach

The disordered case, when p_x are initially chosen at random, is related to the Seppäläinen–Krug model (1999).

The main theorem for the homogeneous case

Assume $p_x \equiv p$.

If $0 < \alpha < (1-p)/p$, then define

$$\begin{aligned} c &= 2\sqrt{\alpha}\sqrt{p(1-p)} + (1-\alpha)p, \\ g &= \alpha^{-1/2}\,(\,p(1-p)\,)^{1/6} \times \\ &\quad \Big(\,(1-\alpha)\sqrt{p(1-p)} + (1-2p)\sqrt{\alpha}\,\Big)^{2/3}. \end{aligned}$$

Then, as $m \to \infty$,

$$P\left(\frac{H - c\,m}{g \cdot m^{1/3}} \le s\right) \to F_2(s),$$

where

$$F_2(s) = \exp\left(-\int_s^\infty (x-s)q(x)^2\,dx\right)$$

and q solves

$$q'' = sq + 2q^3, \quad q(s) \sim \mathrm{Ai}(s) \text{ as } s \to \infty.$$

Main steps in proving the theorem

1. dual RSK algorithm gives a bijection between (0,1)–matrices with k 1's and pairs (P, Q) such that P^t and Q are semistandard Young tableaux (of the same shape) of size k. Most importantly, the length of the first row in P is $H = h_t(x)$. This gives, with $r = p/(1 - p)$,

$$P(h_t(x) \le h) = (1-p)^{mn} \sum_{\substack{\lambda \in \mathcal{P} \\ \ell(\lambda) \le h}} r^{|\lambda|} d_\lambda(m) d_{\lambda'}(n).$$

$$d_\lambda(m) = \#\text{SSYT's of shape } \lambda$$

using integers $\{1, \ldots, m\}$

2. Gessel's theorem (1990) & Borodin–Okounkov identity (1999) then establish the connections between the sum above and determinants of matrices and operators, the final result being

$$P(h_t(x) \le h) = \det(I - K_h),$$

where $K_h : \ell^2 \to \ell^2$ is given by its (j,k)–entry

$$\sum_{\ell=0}^{\infty} (\varphi_-/\varphi_+)_{h+j+\ell+1} (\varphi_+/\varphi_-)_{-h-k-\ell-1}.$$

K_h product of two matrices: (j,k)–entries

$$a_{jk}^{+}(h) = \frac{1}{2\pi i} \int (1+rz)^n \, (z-1)^m \times z^{-m+h+j+k} \, dz,$$

$$a_{jk}^{-}(h) = \frac{1}{2\pi i} \int (1+rz)^{-n} \, (z-1)^{-m} \times z^{m-h-j-k-2} \, dz.$$

The contours for both integrals go around the origin once counterclockwise; in the second integral 1 is inside and $-r^{-1}$ is outside.

3. Scaling:

$$h = cm + sm^{1/3}, j = m^{1/3}x, k = m^{1/3}y$$

and compute integrals asymptotically. Use the **steepest descent method**:

- Two saddles points coincide. Location determines c

- *Double zero*, hence the $m^{1/3}$ scaling.

- Variance normalization determined by coefficient of third derivative

- Limit is a Fredholm determinant with **Airy kernel**.

- The main technical effort is in establishing trace–class convergence of the approximations.

Another Connection with Random Matrices.

In **GUE** (Tracy-Widom, 1993) The largest eigenvalue $\lambda_{\max}$ obeys the limit law

$$P\left((\lambda_{\max} - \sqrt{2n}\,) \cdot \sqrt{2}\, n^{1/6} \leq s\right) \to F_2(s),$$

as $s \to \infty$.

The limit laws for the largest eigenvalue in **GOE** and **GSE** (Tracy-Widom, 1996) also arise as limit laws for **increasing path problems** (Baik-Rains, 2000) and associated **growth processes** (Baik-Rains, Prähofer-Spohn)

No known intuitive connection between largest eigenvalues and increasing paths without using the **RSK correspondence**. With RSK Johansson (2000) has given a **discrete orthogonal polynomial ensemble** approach to increasing subsequence problems. In this formulation one has discrete analogues of the distribution of the largest eigenvalue.

Inhomogeneous ODB

Now assume that A is an $m \times n$ random matrix with $P(\varepsilon_{ij} = 1) = p_j$. Here p_j are i.i.d., with $P(p_j \leq x) = F(x)$, where $F : [0,1] \to [0,1]$ is a distribution function. (H is the longest increasing path of 1's in A.)

This corresponds to a **random environment** version of ODB: every $x \in \mathbb{Z}$ decides before the dynamics starts, at random according to F, on the probabilities of its coin flips.

- Time constant can be explicitly determined in terms of F.
- Quenched and annealed fluctuations differ.
- If the right tails of F are sufficiently thin, there is a *composite* (or *glassy*) regime for small $\alpha = n/m$. This regime can be identified with a different fluctuations scaling.

Lemma: Once $p_1, \ldots, p_n$ are determined, the distribution of H does not depend on their order.

Time Constant

p has distr fn F and $\langle\cdot\rangle$ is integration w.r.t. dF.

$$b := \max \operatorname{supp} dF,$$
$$c := c(\alpha, F) = \lim_{m\to\infty} \frac{H}{m}.$$

Define the following **critical values**:

$$\alpha_c := \left\langle \frac{p}{1-p} \right\rangle^{-1}$$
$$\alpha_c' := \left\langle \frac{p(1-p)}{(b-p)^2} \right\rangle^{-1}.$$

Theorem: If $b = 1$, then $c(\alpha, F) = 1$ for all α, while if $b < 1$, then

$$c(\alpha, F) = \begin{cases} b + \alpha(1-b)\left\langle \frac{p}{b-p} \right\rangle \text{ if } \alpha \le \alpha_c', \\ a + \alpha(1-a)\left\langle \frac{p}{a-p} \right\rangle \text{ if } \alpha_c' \le \alpha \le \alpha_c, \\ 1 \text{ if } \alpha_c \le \alpha. \end{cases}$$

Here $a = a(\alpha, F) \in [b, 1]$ is the unique solution to

$$\alpha \left\langle \frac{p(1-p)}{(a-p)^2} \right\rangle = 1.$$

Fluctuations, quenched case, pure regime:
Theorem: Assume that $b < 1$ and $\alpha'_c < \alpha < \alpha_c$. Then there exists a sequence of random variables $c_n \in \sigma\{p_1, \dots, p_n\}$ and a constant $g \neq 0$ (both depending on α) such that, as $m \to \infty$,

$$P\left(\frac{H - c_n m}{g \cdot m^{1/3}} \leq s \mid p_1, \dots, p_n\right) \to F_2(s),$$

almost surely, for any fixed s. The proof is a uniform version of the proof for fixed p.

Fluctuations, annealed case, pure regime:
Theorem Assume that $b < 1$ and $\alpha'_c < \alpha < \alpha_c$. Let a be as before and

$$\tau^2 = \mathrm{Var}\left(\frac{(1-a)p}{a-p}\right).$$

Then, as $m \to \infty$,

$$\frac{H - cm}{\tau\sqrt{\alpha} \cdot m^{1/2}} \xrightarrow{d} N(0,1).$$

Fluctuations, composite regime: Assume (a technical condition and) that

$$1 - F(b - x) \sim Kx^{\eta}, \text{ as } x \to 0,$$

for some K and $\eta > 2$. Then $\alpha'_c > 0$. Assume also that $b < 1$ and $\alpha < \alpha'_c$, and let

$$\tau^2 = b(1-b)\left(\frac{1}{\alpha} - \frac{1}{\alpha'_c}\right).$$

Theorem: As $m \to \infty$,

$$P\left(\frac{H - c_n m + 2\tau\sqrt{n}}{\tau \cdot \sqrt{n}} \le s \mid p_1, \ldots, p_n\right) \to \Phi(s),$$

almost surely, for any fixed s.

Theorem: For $s > 0$, as $m \to \infty$,

$$P\left(\frac{H - cm}{\gamma \cdot n^{1-1/\eta}} \le s\right) \to e^{-s^{\eta}},$$

where $\gamma = (K\alpha)^{-1/\eta}\,(1 - \alpha/\alpha'_c)$.

Why are the fluctuations increased?

The maximal increasing path has a nearly vertical segment of length asymptotic to $(1 - \alpha/\alpha'_c)m$ in (or near) the column of A which uses the largest probability p_1. Therefore, this vertical part of the path dominates the fluctuations, as the rest presumably has $o(\sqrt{m})$ fluctuations. (These are most likely *not* of the order exactly $m^{1/3}$ as they correspond to the critical case $\alpha = \alpha'_c$.) The variables in the p_1–column are Bernoulli with variances about $b(1-b)$, thus the contribution of the vertical part to the variance is about

$$(b(1-b)(1-\alpha/\alpha'_c)m)^{1/2} = \tau\sqrt{n}.$$

Annealed fluctuations are governed by p_1 since

$$c_n = c - \left(1 - \alpha/\alpha'_c\right)(b - p_1) + o(b - p_1).$$

Future directions, open problems

- What happens in either critical case?

- Is this approach suitable for determining large deviation rates?

- What happens for different growth models or different initial states? For example, nothing is known about the (two–sided) DB given by

  $h_{t+1}(x) = \max\{h_t(x-1), h_t(x+1), h_t(x)+\varepsilon_{x,t}\}.$