GROWTH MODELS AND RANDOM ENVIRONMENTS

Janko Gravner (Univ. of California, Davis)
Craig Tracy (Univ. of California, Davis)
Harold Widom (Univ. of California, Santa Cruz)

Oriented Digital Boiling (ODB).

A two–dimensional interface, which changes in time, is described by $\{(x, y) : x \in \mathbf{Z}, y \leq h_t(x)\}$, $t = 0, 1, 2 \ldots$ and the height function h_t evolves according to the following rule:

(1) $h_t \leq h_{t+1}$.

(2) If $h_t(x-1) > h_t(x)$, then $h_{t+1}(x) = h_t(x-1)$.

(3) Otherwise, $h_{t+1}(x) = h_t(x) + 1$ with probability p_x (independently of other locations and times).

Alternatively, one can toss the p_x–coins in advance to get independent Bernoulli random variables $\epsilon_{x,t}$, $x \geq 0, t \geq 0$. Think of the points (x, t) for which $e_{x,t} = 1$ as *marked*. Then

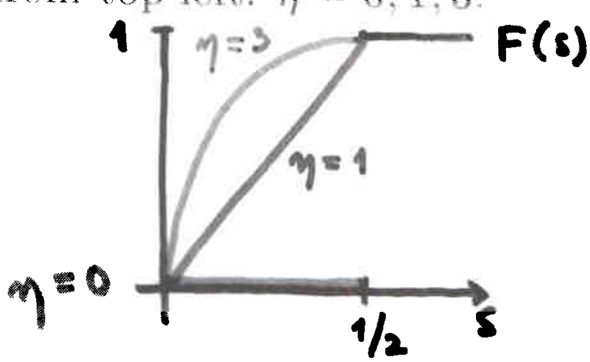
$$h_t(x) = \max\{h_{t-1}(x-1), h_{t-1}(x) + \epsilon_{x,t-1}\}.$$

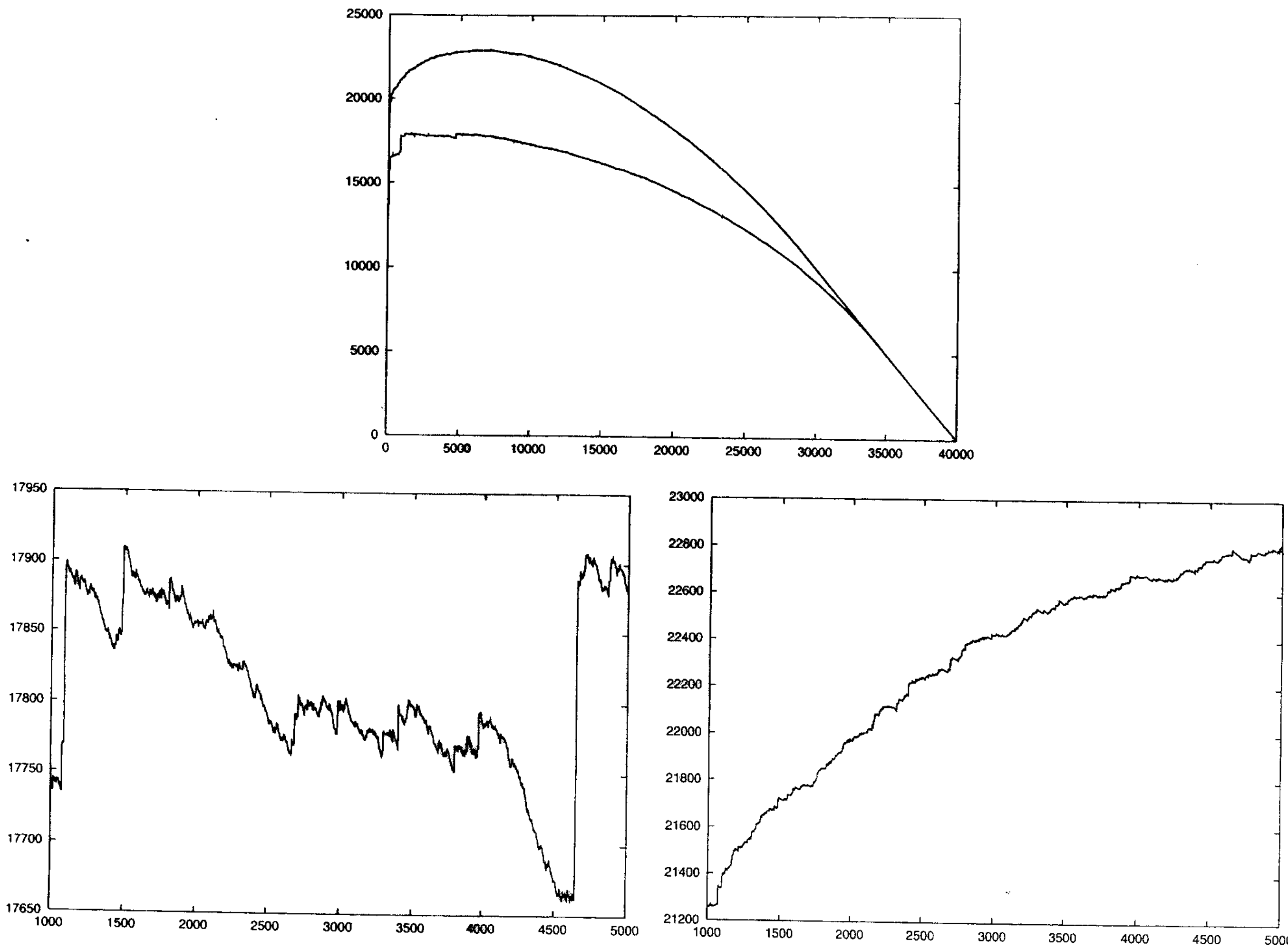
We will assume that the initial state is

$$h_0(x) = \begin{cases} 0, & \text{if } x = 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

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

Three ODB simulations, h_0 as in text, $F(s) = 1 - (1 - 2s)^\eta$.
Clockwise from top left: $\eta = 0, 1, 3$.





Two ODB simulations for $\eta = 1, \eta = 3$, at large times.

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 $\epsilon_{i,j}$, where $P(\epsilon_{i,j} = 1) = p_j$. 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

(1) column index non–decreasing,

(2) row index strictly increasing.

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

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

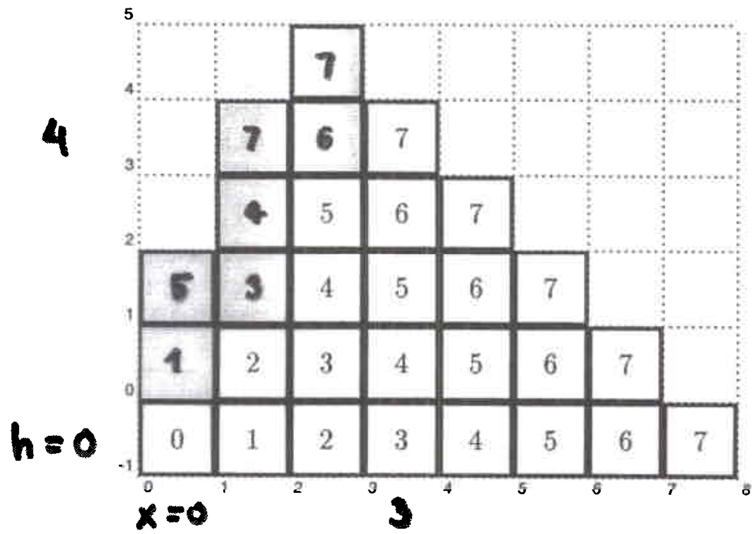


Figure 2: Oriented Digital Boiling Process. The number in a box is the time this box was added and if the box is colored, then the box was added stochastically according to Rule 3.

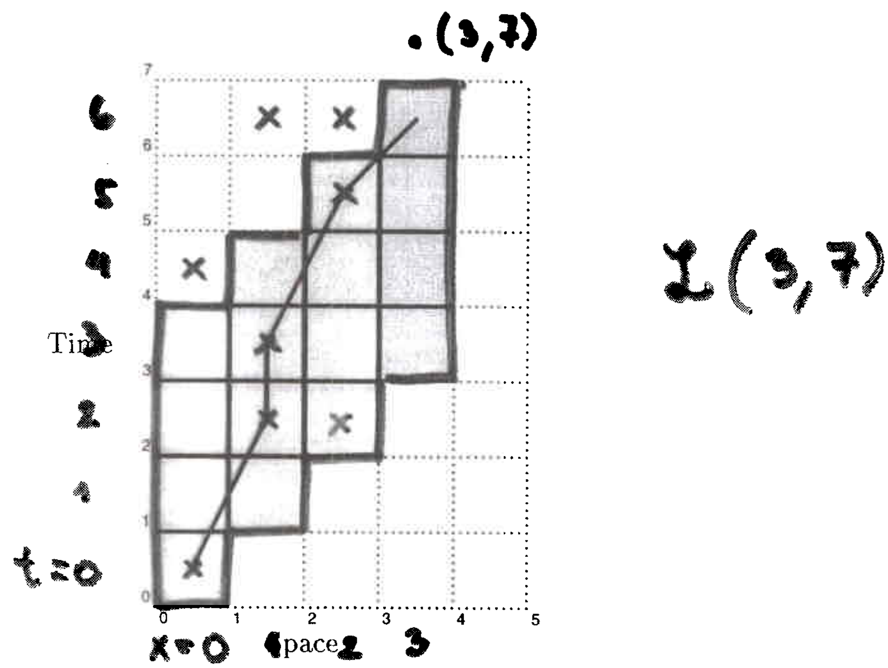


Figure 3: The backwards lightcone of the point $(x,t) = (3,7)$ for the process shown in Fig. 2. The $\times$'s denote the marked points and polyogonal line gives a longest increasing path. The length of this path is equal to the number of $\times$'s in the path. This length equals $h_t(x)$.

A little history.

Longest increasing sequences in random matrices are related to Ulam's problem of estimating the longest increasing subsequence in a random permutation of length n. This has been studied by Hammersley (1972), Logan–Shepp and Vershik–Kerov (1977), Aldous–Diaconis (1995), and Baik–Deift–Johansson (1999). Methods: subadditivity, exclusion process representation, random Young tableaux and random matrices.

The largest increasing sequence in a random 01–matrix has been first studied by Seppäläinen (1998), who used a particle system approach to compute the limiting shape; that is,

$$\lim_{t\to\infty} \frac{h_t(x)}{t},$$

when x/t is constant. Johansson (1999) computed the fluctuations in (universal regime of) this limit law, by a random matrix approach, using orthogonal polynomials.

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 $p_x \equiv p$.

"universal regime"

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

$$c = 2\sqrt{\alpha}\,\sqrt{p(1-p)} + (1-\alpha)p,$$
$$g = \alpha^{-1/2}\left(p(1-p)\right)^{1/6} \times \left((1-\alpha)\sqrt{p(1-p)} + (1-2p)\sqrt{\alpha}\right)^{2/3}.$$

Then, as $m \to \infty$,

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

where

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

and q solves

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

Fluctuations are also known for fixed n (order $\sqrt{m}$), for $n = ((1-p)/p)\cdot m + o(\sqrt{m})$ (tight), and $\alpha > (1-p)/p$ (nonexistent).

Main steps in proving the theorem.

Step 1: Combinatorics and algebra. The dual RSK algorithm, Gessel's theorem (1990) and Borodin–Okounkov identity (1999) establishes a connection between 01–matrices and determinants of operators, the final result being

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

where $K_h : \ell^2 \to \ell^2$ is the product of two matrices, given by (j,k)–entries

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

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

$$r = \frac{p}{1-p}$$

Step 2: Analysis. In the universal regime, we take $h = cm + sm^{1/3}$, $j = m^{1/3}x$, $k = m^{1/3}y$. Then, e.g.,

$$a^{+}_{jk}(h) = \frac{1}{2\pi i}\int \psi(z)(-z)^{m^{1/3}(x+y+s)}\,dz.$$

The asymptotics of the integrals are computed by the steepest descent method. To get a nontrivial limit, we need to choose $c = c_1$ so that the $\frac{d}{dz}\log\psi(z)$ has a double 0, the third derivative of $\log\psi(z)$ then determines c_2. The limit is another Fredholm determinant, of an operator on $L^2[0,1]$ with the Airy kernel.

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

Inhomogeneous ODB.

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

This corresponds to a random environment version of ODB: every $x \in \mathbf{Z}$ decides before the dynamics starts, at random according to F, on the probabilities of its coin flips. In this case, the following can be proved:

(1) Time constant can be explicly determined in terms of F.

(2) Quenched and annealed fluctuations differ.

(3) 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. That is, the distribution of H is a function of the e.d.f.*

Time constant.

Let p have dist. funct. F and $\langle\cdot\rangle$ be the integration w.r.t. dF. Denote

$$b = \max \operatorname{supp} dF,$$

and

$$c = c(\alpha, F) = \lim_{m\to\infty} \frac{H}{m}.$$

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

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

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

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

Variational formula for the time constant.

Let $c(\alpha, x) = c(\alpha, \delta_x)$ and $\zeta(y, x) = y \cdot c(1/y, x)$. Then

$$\zeta(y,x) = \begin{cases} 2\sqrt{y}\sqrt{x(1-x)} + (y-1)x, & x/(1-x) < y, \\ y, & x/(1-x) \geq y. \end{cases}$$

The fact that $c(\alpha, F)$ is given by the most advantageous choice of transition points between the columns of A leads to

$$c(\alpha, F) = \max_{\psi} \int_0^1 \zeta(\psi'(\alpha F(x)), x) \cdot \alpha\, dF(x),$$

where $\psi : [0, \alpha] \to [0, 1]$ are nondecreasing functions with $\psi(0) = 0$ and $\psi(\alpha) = 1$.

This variational problem can be solved. If $\alpha < \alpha_c'$, the minimizer has a jump of size $1 - \alpha/\alpha_c'$ at α, hence one would expect that largest probabilities have a significant influence.

A variational approach was used by Deuschel–Zeitouni (1995), to study a variant of Ulam's problem in which a number of points in the unit square is chosen independently according to some distribution with a density, then a longest sequence, increasing in both coordinates, is extracted from this sample.

Fluctuations, quenched case, pure regime.

Theorem 2. *Assume that* $b < 1$ *and* $\alpha_c' < \alpha < \alpha_c$. *Then there exists a sequence of random variables* $c_n \in \sigma\{p_1, \ldots, p_n\}$ *and a constant* $g \neq 0$ *(both depending on* α*) such that, as* $m \to \infty$,

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

almost surely, for any fixed s.

In fact, c_n are obtained the same way as in Theorem 1, except that the e.d.f. is used to compute $\langle \cdot \rangle$. Furthermore, g is given by an appropriate third derivative.

The proof is a uniform version of the proof for fixed p.

Fluctuations, annealed case, pure regime.

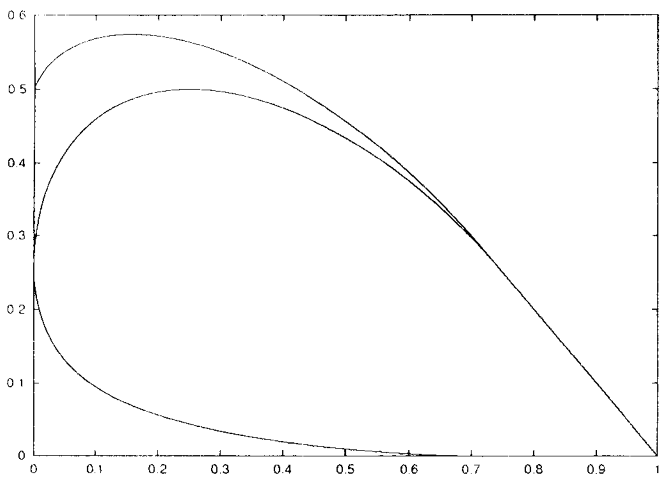
Theorem 3. *Assume that* $b < 1$ *and* $\alpha_c' < \alpha < \alpha_c$*. Let* a *be as in Theorem 1 and*

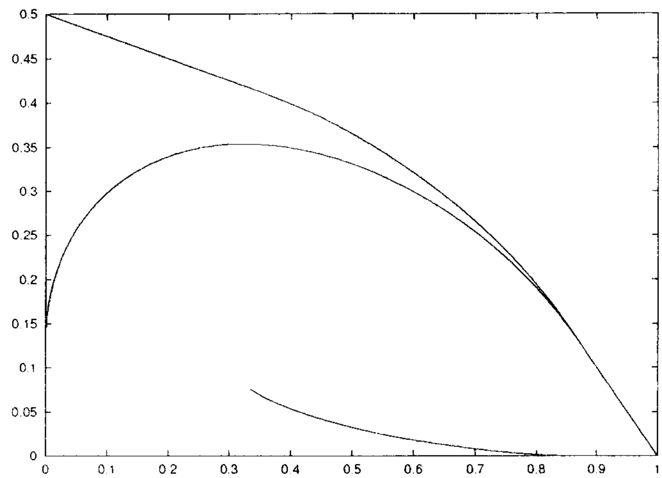
$$\tau^2 = Var\left(\frac{(1-a)p}{a-p}\right).$$

Then, as $m \to \infty$*,*

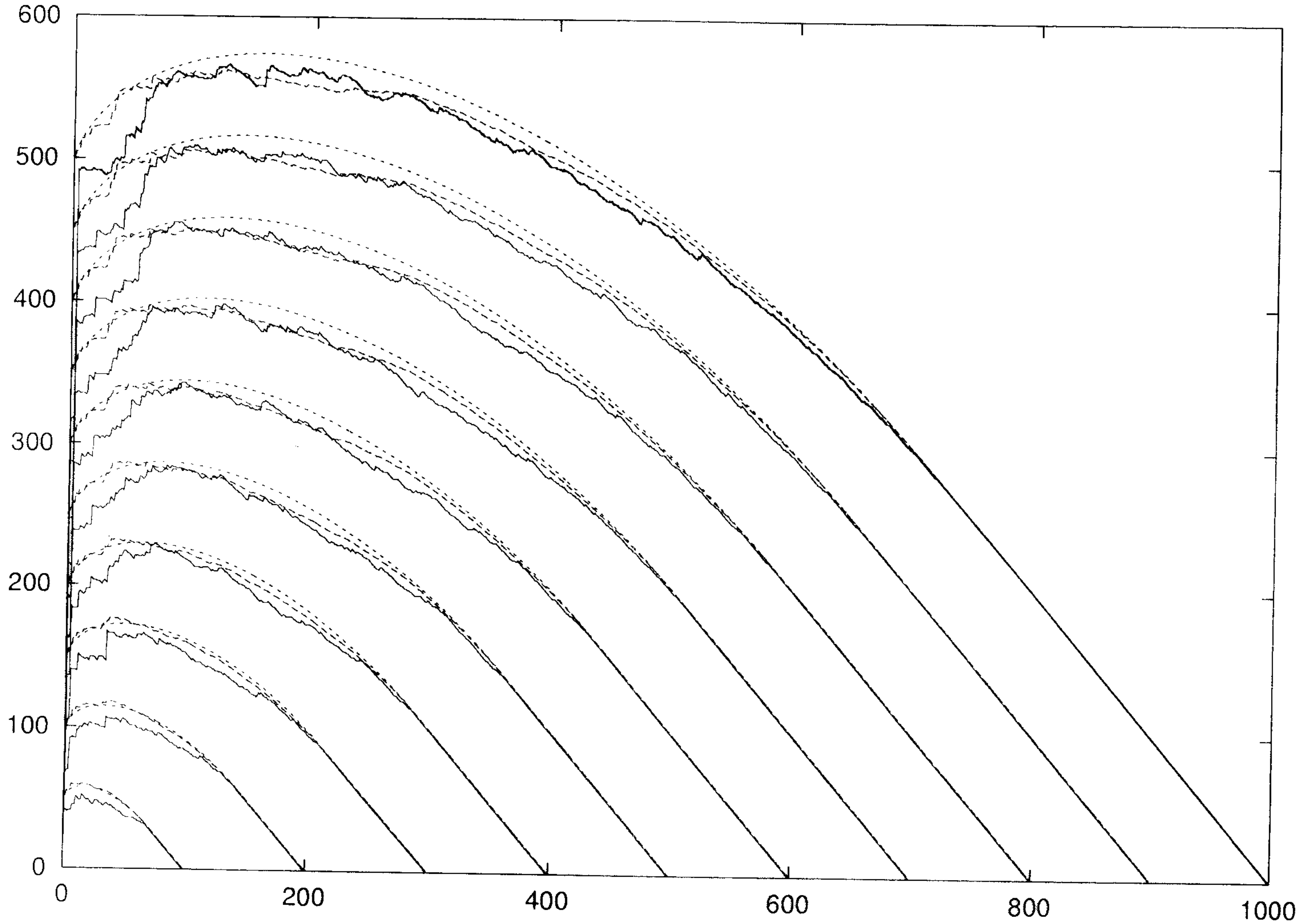
$$\frac{H - cm}{\tau\sqrt{\alpha} \cdot m^{1/2}} \xrightarrow{d} N(0,1).$$

From the Brownian bridge convergence, it follows that c_n satisfy a CLT: $\sqrt{m}(c - c_n)$ converges to a Gaussian variable. Theorem 3 then follows from Theorem 2.





Annealed scaling: shape (top), variance (bottom), and the shape for $p_x \equiv \langle p \rangle$ (middle) vs. x/t, for $\eta = 1, \eta = 3$.



Approximations for $\eta = 1$: h_t is solid curve, dotted curve is based on the annealed normalization, dashed curve on the quenched one.

Fluctuations, composite regime.

Assume (a technical condition and) that

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

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 4. *As $m \to \infty$,*

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

almost surely, for any fixed s.

Theorem 5. *For $s > 0$, as $m \to \infty$,*

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

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

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

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

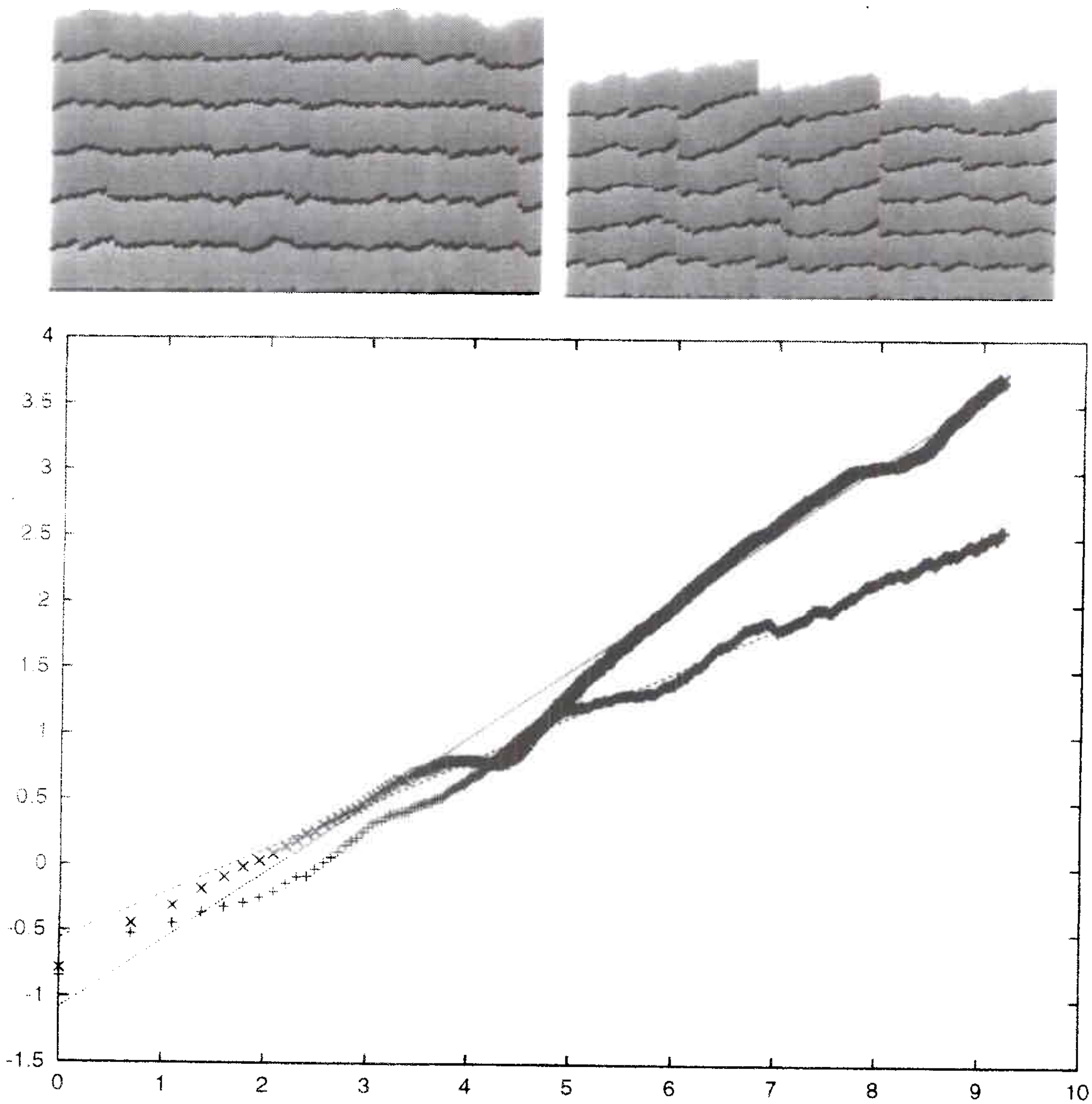
Annealed fluctuations are governed by p_1 since

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

Three questions.

- What happens in either critical case?
- What happens if ODB is started from a flat initial state $h_0 \equiv 0$?
- Do answers change for the (two–sided) DB given by

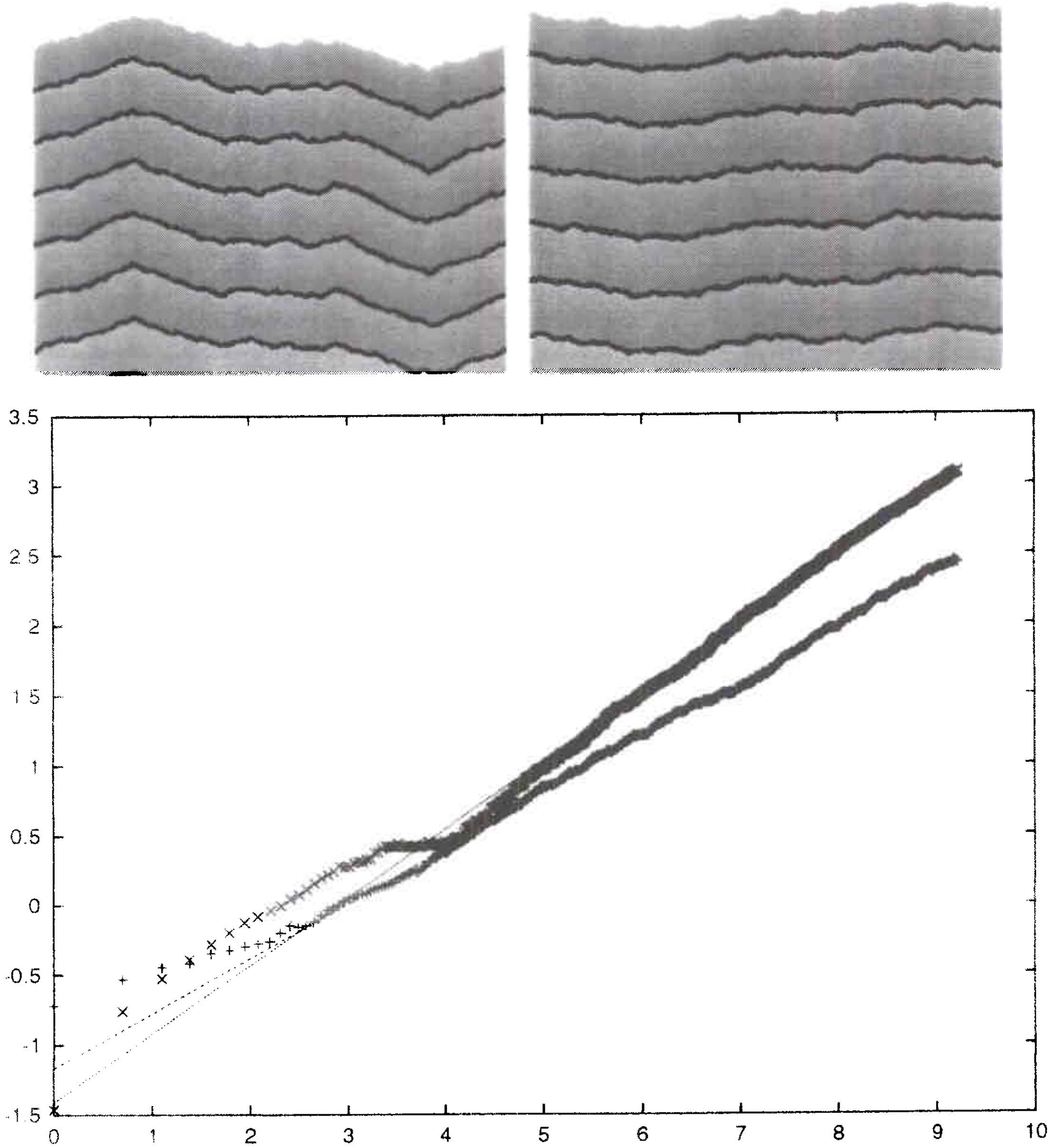
$$h_{t+1}(x) = \max\{h_t(x-1), h_t(x+1), h_t(x) + \epsilon_{x,t}\}?$$



Here, ODB, $h_0 \equiv 0$, $F(s) = 1 - (1 - 2s)^\eta$. The top of Figure 1 illustrates the ODB on 600 sites run until time 600. On the left, $\eta = 1$ (i.e., p is uniform on $[0, 1/2]$), while $\eta = 3$ on the right. In the pure regime, the boundary of the growing set reaches a local equilibrium, while in the composite regime the boundary apparently divides into domains, which are populated by different equilibria and grow sublinearly. The bottom of the figure features a log–log plot of quenched standard deviation (estimated over 1000 independent trials) of $h_t(0)$ vs. t up to $t = 10,000$. The $\eta = 1$ case is drawn with $+$'s (thinner curve) and the $\eta = 3$ case with $\times$'s (thicker curve); the two least squares approximations lines (with slopes 0.339 and 0.517, respectively) are also drawn.

Previous page: ODB, $h_0 \equiv 0$, $F(s) = 1 - (1 - 2s)^\eta$.

The top of Figure 1 illustrates the ODB on 600 sites, run until time 600. On the left, $\eta = 1$ (i.e., p is uniform on $[0, 1/2]$), while $\eta = 3$ on the right. In the pure regime, the boundary of the growing set reaches a local equilibrium, while in the composite regime the boundary apparently divides into domains, which are populated by different equilibria and grow sublinearly. The bottom of the figure features a log–log plot of quenched standard deviation (estimated over 1000 independent trials) of $h_t(0)$ vs. t up to $t = 10,000$. The $\eta = 1$ case is drawn with $+$'s (thinner curve) and the $\eta = 3$ case with $\times$'s (thicker curve); the two least squares approximations lines (with slopes 0.339 and 0.517, respectively) are also drawn.



Here: DB, given by

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

On top: $\eta = 0.2$ (left) and $\eta = 1$ (right). The figures only show evolution near time $t = 5000$. The plot of quenched deviations is analogous to the one in the previous figure, the least squares lines have slopes 0.395 ($\eta = 0.2$) and 0.49 ($\eta = 1$).

Previous page: DB, given by

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

Initial state: $h_0 \equiv 0$. On top: $\eta = 0.2$ (left) and $\eta = 1$ (right). The figures only show evolution near time $t = 5000$. The plot of quenched deviations is analogous to the one in the previous figure, the least squares lines have slopes 0.395 ($\eta = 0.2$) and 0.49 ($\eta = 1$).

Connections with other growth models.

Consider the *Threshold Growth Model (TGM)* with

$$x + \mathcal{N} = \begin{array}{cc} \bullet & x \\ & \bullet \end{array}$$

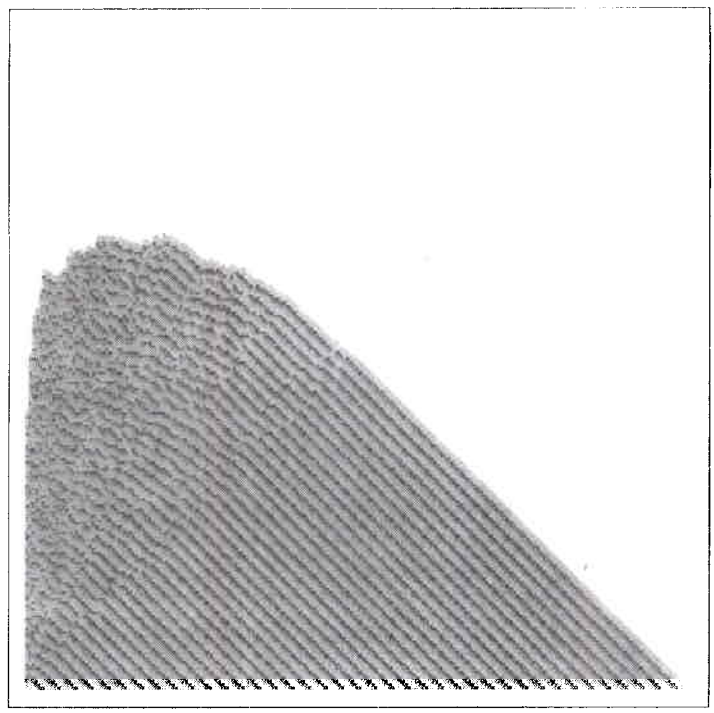
$\theta = 1$, $p_1 = p$, and $p_2 = 1$.

Assume $A_t \subset \mathbf{Z}^2$ is given. Then the updated set A_{t+1} is obtained by adjoining a site $x \in \mathbf{Z}^2$

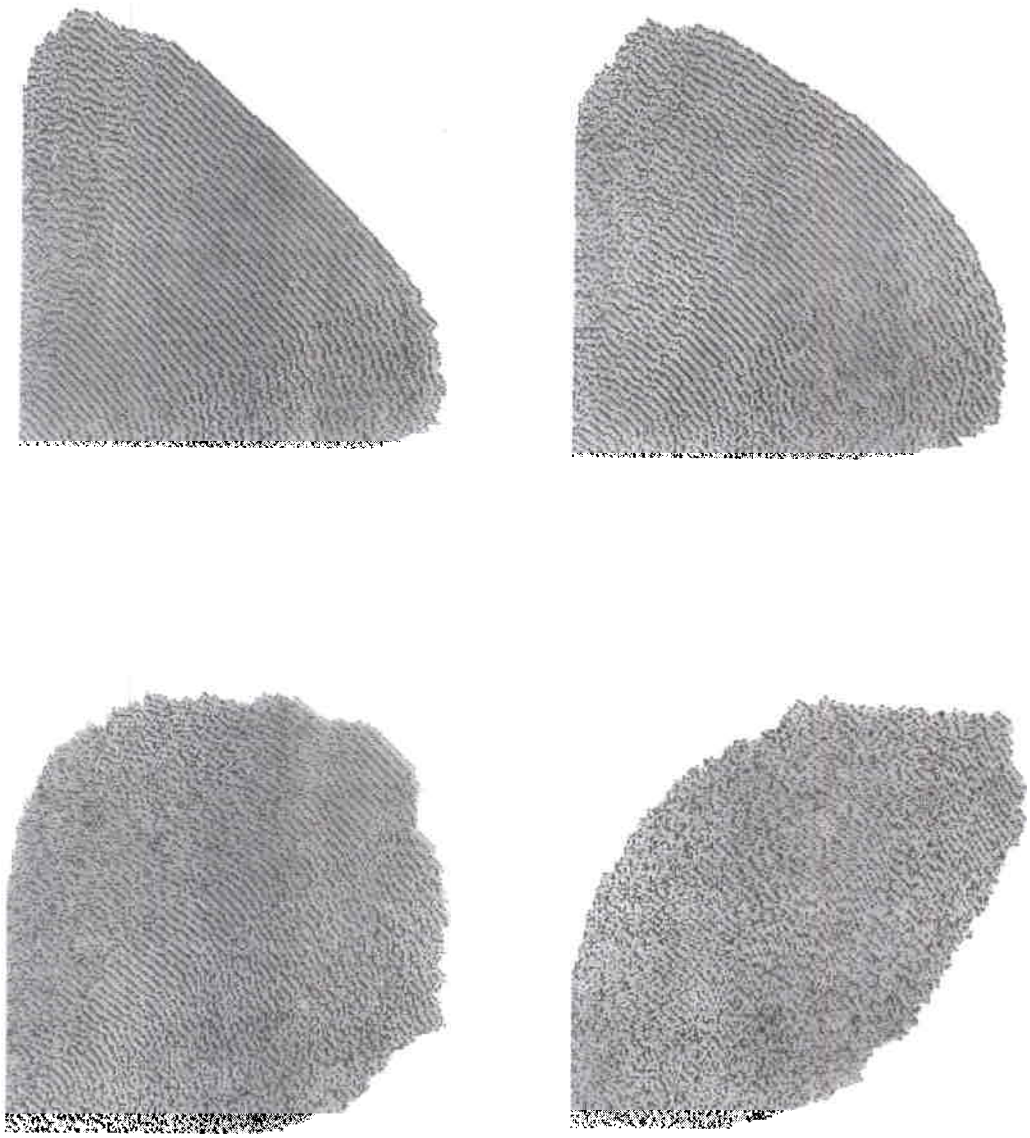
- automatically if $|(x + \mathcal{N}) \cap \mathbf{Z}^2| \geq 2$, and
- independently with prob. p if $|(x + \mathcal{N}) \cap \mathbf{Z}^2| = 1$.

This TGM is equivalent to ODB when $A_0 = \{(x, y) : x \geq 0, y \leq -x\}$. What if $A_0 = \{0\}$ (or any other finite set)?

The asymptotic shape $L_p = \lim_{t\to\infty} A_t/t$ can be completely described. So can the fluctuations of A_t around $t \cdot L_p$ in any direction $\gamma = x/t$ *except in the diagonal direction when* $p \leq 1/2$. Such fluctuations are not known for any local growth model in a direction where the asymptotic shape has a kink. (*Except* in cases when the shape is the same as deterministic: $L_p = L_1$, G–Griffeath, 2002.)



ODB, with $p = 0.5$, started from the wedge:



ODB shapes for $p = 0.6, 0.5, 0.3, 0.1$.

Connections with other growth models, cont.

Construct $G(M,N)$ on $\mathbf{Z}^2_+$ by

$$G(M,N) = \max\{G(M-1,N), G(M,N-1)\} + g(M,N),$$

where $g(M,N)$ are i.i.d. and $P(g(M,N)=k) = p^k(1-p)$, $k \geq 0$.

An observation by Prähofer couples $h(x,t)$ and $G(M,N)$ so that

$$h(M, G(M,N)+M+N) = G(M,N),$$

and so at least the time constant follows from the classic result by Rost (1981). It is not clear whether the Prähofer coupling can connect the fluctuation result to Johansson's (2000) result on G.

The disordered version of ODB corresponds to the case when $p = p_M$ are independently randomly chosen (but constant for a fixed M). In the particle interpretation, this gives rise totally asymmetric simple exclusion with random jump probabilities associated with particles, thus it provides an alternative approach to the Seppäläinen–Krug traffic model (1999).