Internal Diffusion Limited Aggregation

David Jerison (MIT)

May 2011

Elias M. Stein Conference, Princeton
Joint work with Lionel Levine and Scott Sheffield

Talk Outline

- Internal Diffusion Limited Aggregation (internal DLA) and its deterministic scaling limit
- **Fluctuations:**
  Mean field theory = variant of Gaussian free field (GFF)
  Maximum fluctuation

Talk Outline

- Internal Diffusion Limited Aggregation (internal DLA) and its deterministic scaling limit
- **Fluctuations:**
  Mean field theory = variant of Gaussian free field (GFF)
  Maximum fluctuation
- Continuum limit Hele-Shaw flow
- analogue of conserved quantities for Hele-Shaw flow: Martingales

Internal DLA with Multiple Sources

- Finite set of points $x_1, \ldots, x_k \in \mathbb{Z}^d$.
- Start with m_i particles at site x_i.

Internal DLA with Multiple Sources

- Finite set of points $x_1, \ldots, x_k \in \mathbb{Z}^d$.
- Start with m_i particles at site x_i.
- Each particle performs **simple random walk** in $\mathbb{Z}^d$ until reaching an unoccupied site.

Internal DLA with Multiple Sources

- Finite set of points $x_1, \ldots, x_k \in \mathbb{Z}^d$.
- Start with m_i particles at site x_i.
- Each particle performs **simple random walk** in $\mathbb{Z}^d$ until reaching an unoccupied site.
- Get a **random set** of $\sum_i m_i$ occupied sites in $\mathbb{Z}^d$.
- The distribution of this random set does not depend on the order of the walks (**Diaconis-Fulton '91**).

Limiting shape

- Fix a single source at the origin in $\mathbb{Z}^d$.
- Run internal DLA on $\frac{1}{n}\mathbb{Z}^d$ with Cn^d particles.

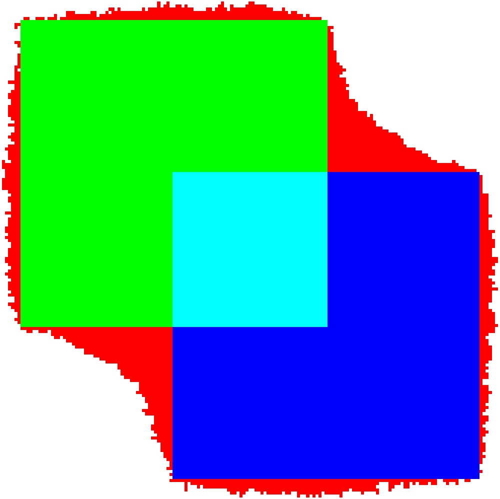
Limiting shape

- Fix a single source at the origin in $\mathbb{Z}^d$.
- Run internal DLA on $\frac{1}{n}\mathbb{Z}^d$ with Cn^d particles.
- **Lawler-Bramson-Griffeath '92**
  Theorem. The limit shape is a ball.

Limiting shape

- Fix a single source at the origin in $\mathbb{Z}^d$.
- Run internal DLA on $\frac{1}{n}\mathbb{Z}^d$ with Cn^d particles.
- **Lawler-Bramson-Griffeath '92**
  Theorem. The limit shape is a ball.

- **Diaconis-Fulton smash sum**: overlapping clusters



Diaconis-Fulton sum of two squares in $\mathbb{Z}^2$ overlapping in a smaller square.

Levine-Peres Theorem 2008: The rescaled limit shape of any distribution is given by the solution to an obstacle problem.

Levine-Peres Theorem 2008: The rescaled limit shape of any distribution is given by the solution to an obstacle problem.
Divisible Sandpile Model: "Toppling" process
initial distribution $\longmapsto$ final distribution

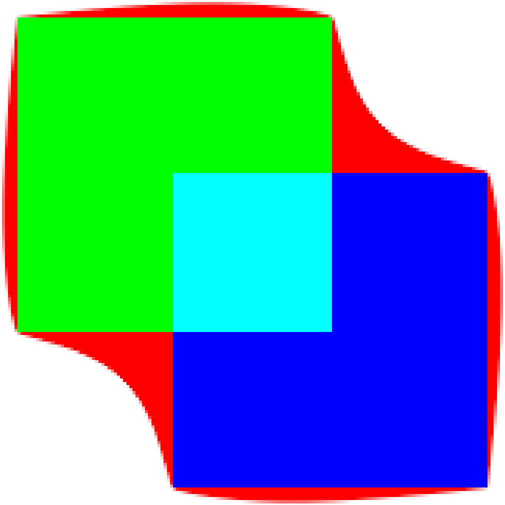
$$\mu(x) \longmapsto \nu(x) \qquad (\text{e.g., } 1_{\Omega_1} + 1_{\Omega_2} \longmapsto 1_D)$$

- Replace $w(x) > 1$ by $w(x) = 1$
- Donate $(w(x) - 1)/2d$ of the excess to each nearest neighbor.
- Repeat until $\nu(x) \leq 1$.

Levine-Peres Theorem 2008: The rescaled limit shape of any distribution is given by the solution to an obstacle problem.
Divisible Sandpile Model: "Toppling" process
initial distribution $\longmapsto$ final distribution

$$\mu(x) \longmapsto \nu(x) \quad (\text{e.g., } 1_{\Omega_1} + 1_{\Omega_2} \longmapsto 1_D)$$

- Replace $w(x) > 1$ by $w(x) = 1$
- Donate $(w(x) - 1)/2d$ of the excess to each nearest neighbor.
- Repeat until $\nu(x) \leq 1$.
- Abelian property: $\nu(x)$ is the same regardless of order of toppling.



Divisible Sandpile starting from the sum of two squares in $\mathbb{Z}^2$.

Pavel Etingof, Max Rabinovich

Hexagonal lattice
Balanced random walk

Pavel Etingof, Max Rabinovich

Hexagonal lattice
Balanced random walk

Jim Propp Question: **Random walk with drift**

Answer: Parabolic obstacle problem: “Heat ball”

Corollary of Levine-Peres The scaled limit of the internal DLA evolution is **Hele-Shaw flow:**

$$\Omega(t) \subset \mathbb{R}^d \quad \text{grows according to} \quad V = \text{normal velocity of } \partial\Omega(t)$$

$$\text{given by } V = |\nabla G_t|$$

($G_t =$ Green's function for $\Omega(t)$ with pole at 0 and $V\, d\sigma_t$ is the hitting probability of Brownian motion.)

Meakin-Deutch 1986 Proposed internal DLA as model for electropolishing, etching, and corrosion. What grows is the fluid region.

Meakin-Deutch 1986 Proposed internal DLA as model for electropolishing, etching, and corrosion. What grows is the fluid region.

How smooth? Numerical simulation of t-particle blob has average fluctuations

$$O(\sqrt{\log t}) \qquad d = 2$$

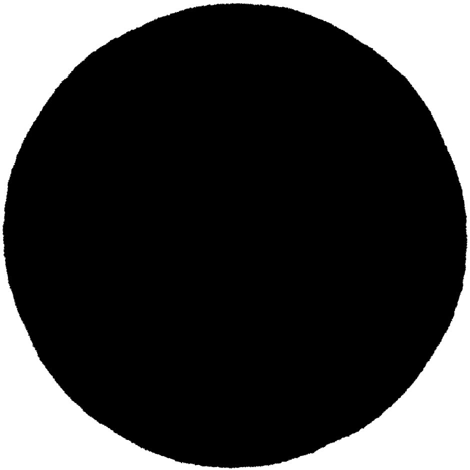
$$O(1) \qquad d = 3$$

Meakin-Deutch 1986 Proposed internal DLA as model for electropolishing, etching, and corrosion. What grows is the fluid region.
How smooth? Numerical simulation of t-particle blob has average fluctuations

$$O(\sqrt{\log t}) \qquad d = 2$$

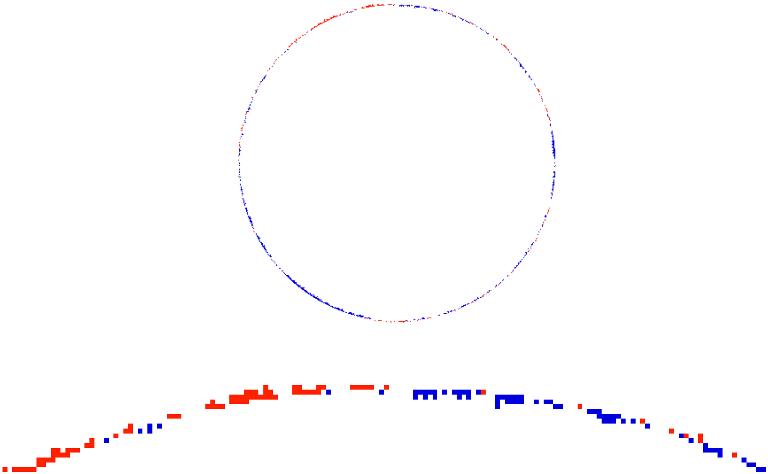
$$O(1) \qquad d = 3$$

Lawler 1995: Max fluctuation $O(r^{1/3})$, $\quad r = t^{1/d}$, almost surely.



Internal DLA blob A_t with $t = 10^6$ particles in $\mathbb{Z}^2$

Detail of boundary of the 1 million particle blob



$t = 10^5$ Symmetric difference: Early = red; Late = blue

Theorem 1 (J–, Levine, Sheffield) Let $d = 2$.
As $t = \pi r^2 \to \infty$, the rescaled discrepancy function

$$X_t = r^{-1} \sum_{z \in \mathbb{Z}^2} (1_{A_t} - 1_{B(r)}) \delta_{z/r} \longrightarrow X \, d\theta$$

X is a Gaussian random distribution on unit circle S^1, associated with the Hilbert space $H^{1/2}(S^1)$.

Theorem 1 (J–, Levine, Sheffield) Let $d = 2$.
As $t = \pi r^2 \to \infty$, the rescaled discrepancy function

$$X_t = r^{-1} \sum_{z \in \mathbb{Z}^2} (1_{A_t} - 1_{B(r)}) \delta_{z/r} \longrightarrow X\, d\theta$$

X is a Gaussian random distribution on unit circle S^1, associated with the Hilbert space $H^{1/2}(S^1)$.

$$X(\theta) = c \sum_{k=1}^{\infty} \alpha_k \frac{\cos k\theta}{\sqrt{k+1}} + \beta_k \frac{\sin k\theta}{\sqrt{k+1}}$$

where α_k and β_k are independent $N(0,1)$.

Gaussian random variables in general

$$X(x) = \sum_k \alpha_k \phi_k(x)$$

where ϕ_k is an orthonormal basis of a Hilbert space and α_k are independent N(0,1) real-valued coefficients. In our case, the Hilbert spaces are Sobolev spaces and the random variables are identified with distributions.

Dual Formulation

Central limit theorem for finite-dimensional projections:

$$\int_0^{2\pi} \int_0^{\infty} X_t \, \phi(\rho,\theta) \, \rho \, d\rho d\theta \longrightarrow N(0,V) \qquad t \to \infty$$

$$\phi(1,\theta) = a_0 + \sum_{k=1}^{N} a_k \cos k\theta + b_k \sin k\theta;$$

$$V = \pi c \sum_{k=1}^{N} (a_k^2 + b_k^2)/(k+1) \qquad = \text{variance}$$

Theorem 2, J–, Levine, Sheffield Let $d = 2$.
Let $T(t) = \#$ points in unit intensity Poisson process in $[0, t]$.

$$F(x) = \inf\{t : x \in A_{T(t)}\} = \text{ arrival time}$$

$$L(x) = \sqrt{F(x)/\pi} - |x| = \text{ lateness}$$

Then as $R \to \infty$

$$L(\lfloor Rx_1 \rfloor, \lfloor Rx_2 \rfloor)$$

tends to a variant of the Gaussian free field (GFF), a random distribution for the Hilbert space H^1 of the plane.

$$\|f\|_D^2 = \int_{\mathbb{R}^2} |\nabla_x f|^2 dx = \int\int (|\partial_\rho f|^2 + \rho^{-2}|\partial_\theta f|^2)\rho \, d\rho d\theta$$

$$\|g(\rho)e^{ik\theta}\|_D^2 = 2\pi \int_0^\infty (|\rho g'(\rho)|^2 + k^2|g(\rho)|^2) d\rho/\rho$$

Variant norm in our theorem:

$$\|g(\rho)e^{ik\theta}\|^2 = 2\pi \int_0^\infty (|\rho g'(\rho)|^2 + (|k|+1)^2|g(\rho)|^2) d\rho/\rho$$

Dual Formulation

As $r \to \infty$,

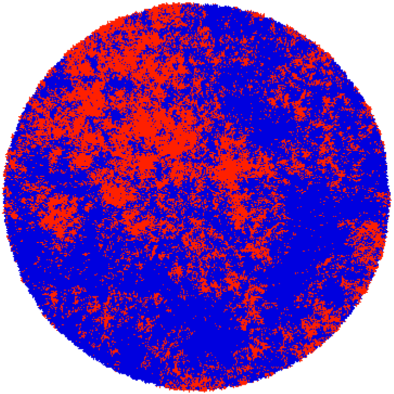
$$\frac{1}{r^2} \sum_{x \in \mathbb{Z}^2/r} L(rx) \frac{\phi(x)}{|x|^2} \longrightarrow N(0, V)$$

where $z = x_1 + ix_2 = \rho e^{i\theta}$,

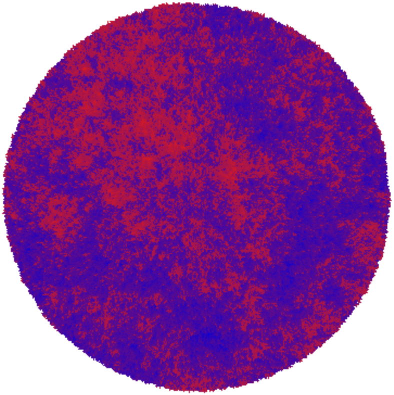
$$\phi(z) = \psi_0(\rho) + \mathrm{Re} \sum_{k=1}^{N} \psi_k(\rho) e^{ik\theta}$$

where ψ_k are smooth and compactly supported on $0 < \rho < \infty$ and the variance V is the square of the dual norm to H^1 above

$$V = 2\pi \sum_{k=0}^{N} \int_0^\infty \left| \int_\eta^\infty \psi_k(\rho)(s/\rho)^{|k|+1} d\rho/\rho \right|^2 ds/s$$



(a)



(b)

(a) $A_{T(t)}$ for $t = 10^5$; Early $= L < 0 =$ red; Late $= L > 0 =$ blue.
(b) Same cluster representing L by red-blue shading.

Heuristics: Same predictions as Meakin-Deutch. At scale r,

$$\text{Variance}(X(\theta)) \approx \sum_{k=1}^{r} \frac{\sin^2 k\theta + \cos^2 k\theta}{(\sqrt{k+1})^2} \approx \log r$$

$$\implies \quad \text{Standard Deviation}(X(\theta)) \approx \sqrt{\log r} \qquad (d=2)$$

Heuristics: Same predictions as Meakin-Deutch. At scale r,

$$\text{Variance}\,(X(\theta)) \approx \sum_{k=1}^{r} \frac{\sin^2 k\theta + \cos^2 k\theta}{(\sqrt{k+1})^2} \approx \log r$$

$$\implies \quad \text{Standard Deviation}\,(X(\theta)) \approx \sqrt{\log r} \qquad (d=2)$$

Maximum fluctuation: r "distinct" values of θ

$$\text{worst of } r \text{ trials} \quad O(\log r) \qquad (d=2)$$

Heuristics: Same predictions as Meakin-Deutch. At scale r,

$$\text{Variance}(X(\theta)) \approx \sum_{k=1}^{r} \frac{\sin^2 k\theta + \cos^2 k\theta}{(\sqrt{k+1})^2} \approx \log r$$

$$\Longrightarrow \quad \text{Standard Deviation}(X(\theta)) \approx \sqrt{\log r} \qquad (d=2)$$

Maximum fluctuation: r "distinct" values of θ

$$\text{worst of } r \text{ trials} \quad O(\log r) \qquad (d=2)$$

In dimension $d = 3$, the prediction is

$$\text{typical fluctuation} = O(1); \qquad \text{worst} = O(\sqrt{\log r})$$

Theorem 3, J–, Levine, Sheffield Almost surely with $t = \omega_d r^d$,

$$B(r - C\log r) \cap \mathbb{Z}^2 \subset A_t \subset B(r + C\log r) \qquad (d = 2)$$

$$B(r - C\sqrt{\log r}) \cap \mathbb{Z}^d \subset A_t \subset B(r + C\sqrt{\log r}) \qquad (d > 2)$$

Asselah, Gaudillière, 2009 Significant improvement of Lawler's power law.

Asselah, Gaudillière, 2010 Theorem 3 also follows from their methods.

How to guess the form of Theorems 1 and 2
Rescale to unit size; pretend the domain is $r < 1 + \varepsilon f(\theta)$.

$$f(\theta) = \alpha_0 + \sum_k \alpha_k \cos k\theta + \beta_k \sin k\theta$$

Under Hele-Shaw flow, f changes in proportion to

$$|\nabla G| \approx \frac{\partial}{\partial r}\left[\alpha_0 + \sum_k \alpha_k r^k \cos k\theta + \beta_k r^k \sin k\theta\right]$$

Restoring force:

$$d\alpha_k = -(k+1)\alpha_k d\rho + dB \quad (\rho = \log r)$$

Conserved harmonic moments (quadrature domains)

If $\Delta\phi = 0$, then for Hele-Shaw flow $\Omega(t)$

$$\frac{d}{dt}\int_{\Omega(t)} \phi\, dx = \int_{\partial\Omega(t)} \phi V d\sigma_t = \phi(0)$$

Discrete analogue:

$$M(t) = \sum_{x \in A(t)} (\phi(x) - \phi(0)) \quad \text{(discrete harmonic } \phi\text{)}$$

is a martingale.

Theorems 1 and 2 are proved using the **martingale central limit theorem**

$$S_n = \sum_1^n X_i, \quad \mathbf{E}(X_n | X_1, \ldots, X_{n-1}) = 0$$

$$Q_n = \sum_1^n X_i^2; \quad s_n^2 = \mathbf{E} S_n^2 = \mathbf{E} Q_n$$

If $Q_n / s_n^2 \to 1$ and $\max_{i=1}^n X_i^2 / s_n^2 \to 0$ almost surely, then

$$S_n / s_n \longrightarrow N(0,1) \quad \text{in law}$$

(Martingale representation theorem: $S_n = B_{t(n)}$, coupling with a Brownian motion.)

Are there enough harmonic polynomial test functions?

Yes. Given a polynomial P defined on $\mathbb{R}^d$ and a radius r, there is a unique harmonic polynomial Q that agrees with P on the sphere of radius r. (Indeed, spherical harmonics can be used as a basis for $L^2(S^{d-1})$.)

Sketch of proof of of Theorem 1.

$$\phi(z) = a_0 + \mathrm{Re}\sum_{k=1}^{N} \alpha_k z^k; \quad \alpha_k = a_k + ib_k$$

$$\Phi_R(z) = a_0 + \mathrm{Re}\sum_{k=1}^{N} \alpha_k P_k(z)/R^k$$

$P_k(z) = z^k + O(|z|^{k-2})$ and P_k is discrete harmonic,

$M(t) := \sum_{z \in A_t} (\Phi_R(z) - a_0), \quad 0 \le t \le T = \pi R^2$ IS A MARTINGALE

$$Q(T) = \sum_{z \in A_T} |\Phi_R(z) - a_0|^2$$
$$= T \left[\frac{1}{2} \sum_{k=1}^{N} |\alpha_k|^2/(k+1) \right] (1 + O(T^{-1/3}))$$

Hence, the martingale convergence theorem $\implies$

$$M(T)/T^{1/2} \longrightarrow N(0, V) \quad \text{as} \, T \to \infty$$

with

$$V = \frac{1}{2} \sum_{k=1}^{N} |\alpha_k|^2/(k+1)$$

$$X_T = R^{-1} \sum_{z \in \mathbb{Z}^2} (1_{A_T} - 1_{B(R)}) \phi(z/R)$$

$$M(T)/\sqrt{T} = T^{-1/2} \sum_{z \in \mathbb{Z}^2} (1_{A_T} - 1_{B(R)}) \Phi_R(z)$$

and $\Phi_R(z) - \phi(z/R)$ is small near $|z| = R$.

Higher dimensional Central Limit Theorems for Fluctuations

$$X_T = R^{-d/2} \sum_{x \in \mathbb{Z}^d} (1_{A_T} - 1_{B(R)}) \phi(x/R), \quad T = \omega_d R^d$$

Fails in dimensions $d \geq 4$. Best possible mean value properties:

$$\frac{1}{R^d} \sum_{x \in B(R)} (\Phi(x) - \Phi(0)) = \Omega(R^{-2}) \qquad (d \geq 5)$$

Higher dimensional Central Limit Theorems for Fluctuations

$$X_T = R^{-d/2} \sum_{x \in \mathbb{Z}^d} (1_{A_T} - 1_{B(R)}) \phi(x/R), \quad T = \omega_d R^d$$

Fails in dimensions $d \geq 4$. Best possible mean value properties:

$$\frac{1}{R^d} \sum_{x \in B(R)} (\Phi(x) - \Phi(0)) = \Omega(R^{-2}) \qquad (d \geq 5)$$

Sandpile w_T is rounder than a ball!

$$\sum_{x \in \mathbb{Z}^d} (\Phi(x) - \Phi(0)) w_T = 0$$

To rescue theorem in all dimensions, use

$$\tilde{X}_T = R^{-d/2} \sum_{x \in \mathbb{Z}^d} (1_{A_T} - w_T) \Phi(x), \quad T = \omega_d R^d$$

Sketch of proof of Theorem 3.

Early/late point detector martingale using test function = discrete Green's function in a ball.

Fix $z_0 \in B$, a ball in $\mathbb{Z}^d$. Define

$$G_{z_0}(z) = \mathbf{P}(\text{random walk from } z \text{ reaches } z_0 \text{ before exiting } B)$$

Discrete harmonic function in punctured ball
Boundary values $= 0$ on ∂B and $= 1$ at z_0.

Lemma

No Tentacles

Let $|x| > m$, $x \in \mathbb{Z}^d$,

$$\mathbb{P}(x \in A_t \text{ and } \#(A_t \cap B(x,m)) \le m^2/C) \le Ce^{-m/C} \quad d = 2$$

$$\mathbb{P}(x \in A_t \text{ and } \#(A_t \cap B(x,m)) \le m^d/C) \le Ce^{-m^2/C} \quad d > 2$$

Lemma

No late points implies no early points

Proof: Assume by contradiction there is an early point. This forces a large positive value of a point detector martingale.

Lemma

No early points implies no late points

Proof: Assume by contradiction, there is a late point. This forces a large negative value of a point detector martingale.

Early and late point detector

$$H_\zeta(z) \approx \frac{\zeta/\rho}{\zeta - z} \quad \rho = |\zeta|$$

$H(0) = 1/\rho$, $H(\zeta) \approx 1$, and discrete harmonic except at ζ

$$M_\zeta(n) = \sum_{z \in A'(n)} (H_\zeta(z) - 1/\rho)$$

$A'(n)$ stops at $\partial B_\rho(0)$, where $H_\zeta(z) = 1/2\rho$. Because $A'(n)$ is stopped before reaching the singularities, M_ζ is a martingale.

LEMMA 1. No ℓ-late implies no $m = C\ell$-early

Event $Q[z,n]$:

- z is the nth point
- z is m-early ($z \in A(\pi r^2)$, $r \approx |z| - m$)
- No previous point is ℓ-late

We will use M_ζ for $\zeta = (1+4m/r)z$ to show for $0 < n \leq T$,

$$\mathbb{P}(Q[z,n]) \leq T^{-10}$$

There is a Brownian motion such that

$$M_\zeta(n) = B_\zeta(s_\zeta(n)), \quad s_\zeta(n) = \mathrm{Var}(M_\zeta(n))$$

On the event $Q[z, n]$

$$\mathbb{P}(s_\zeta(n) < 100 \log T) > 1 - T^{-10}$$

$$\mathbb{P}(M_\zeta(n) > c_0 m) > 1 - T^{-10}$$

On the other hand, $(s = 100 \log T)$

$$\mathbb{P}(\sup_{s' \in [0,s]} B_\zeta(s') \geq s) \leq e^{-s/2} = T^{-50}$$

LEMMA 2. No m-early implies no $\ell = \sqrt{C(\log T)m}$-late

Event $B[\zeta]$: $0 < n \leq T$

- ζ is the nth point
- ζ is ℓ-late
- no point of $A(n)$ is m-early

We will use M_ζ to show that for $0 < n \leq T$,

$$\mathbb{P}(B[\zeta]) \leq T^{-10}$$

On the event $B[\zeta]$, $\rho = |\zeta|$,

$$M_\zeta(T_1) \le -\ell \quad (T_1 = \pi(\rho + \ell)^2)$$

$$\mathbb{P}(s_\zeta(n) < 100m + 100 \log T) > 1 - T^{-10}$$

On the other hand, (with $k \approx \ell/m$, $s = m$)

$$\mathbb{P}(\inf_{s' \in [0,s]} B_\zeta(s') \le -ks) \le e^{-k^2 s/2} \approx T^{-50}$$

Happy Birthday Eli!