The algorithmic hardness threshold for the continuous random energy model

Louigi Addario-Berry
McGill

Pascal Maillard
Paris Sud

Branching Structures
Zurich, Sep 3 2018
Branching random walk

Generation $n$: $T_n$ (assume at least binary so $|T_n| \geq 2^n$)

Position of node $v$: $X_v$

Minimum position in generation $n$: $M_n$

Fairly generic fact: $\exists c \in \mathbb{R}$ s.t. $\lim_{n \to \infty} M_n \overset{a.s.}{\to} c$

and moreover $\mathbb{E} M_n = (1 + o(1)) c n$

Fairly generic proof of a.s. conv. a) lower bound:

$\forall c < c^*,$ $\mathbb{E} \# \{ v \in T_n : X_v < c n \} = O(e^{-\delta n})$, some $\delta = \delta(c^*) > 0.$

Then use Borel-Cantelli.

b) upper bound: Fix $c^* > c,$ then $\exists k \in \mathbb{N}$ st

$\mathbb{E} \# \{ v \in T_k : \forall j < k, X_{a(v,j)} \leq c^* k \} > 1$

gen. $j$ ancestor of $v$

Define a renormalized BRW

$T^{(i)} = \{ v \in T_k : \forall j < k, X_{a(v,j)} \leq c^* k \}$

$\mathbb{E} \# T^{(i)} > 1$

$T^{(i+1)} = \{ v \in T^{(i+1)} : \forall j \in [ik, (i+1)k), X_{a(v,j)} - X_{a(v,ik)} \leq c^* k \}$

Then $\hat{T} := (T^{(i)}, i \geq 1)$ is a supercritical branching process

and on the event $\{ \hat{T} \text{ survives} \}$, have $\limsup_{n \to \infty} n^{-1} M_n \leq c^*.$

So $\Pr \{ \limsup_{n \to \infty} n^{-1} M_n \leq c^* \} \geq \Pr \{ \hat{T} \text{ survives} \} = \Pr(c^*) > 0.$
Branching random walk

Generation $n$: $T_n$ (assume at least binary so $|T_n| \geq 2^n$)

Position of node $v$: $X_v$

Minimum position in generation $n$: $M_n$

Fairly generic fact: $\exists c \in \mathbb{R}$ st. $n^t M_n \xrightarrow{a.s.} c$ and moreover $\mathbb{E} M_n = (1+o(1)) c n$

Remark: In fair generality, $c$ is also the expectation threshold, in that

$$\lim_{n \to \infty} n^{-1} \log \mathbb{E} \#\{v \in T_n : X_v < x n\} = \begin{cases} >0, & x > c \\ =0, & x = c \\ <0, & x < c \end{cases}$$
Finding near-minimal states

Hereafter assume binary branching, sub-Gaussian displacements.

\[ \exists c,C>0 \text{ s.t. } \Pr(|X_v - X_{\text{parent}(v)}| > y) \leq C e^{-cy^2}, \]

for all nodes \( v \).

Write \( \sigma^2 \) for offspring variance.

How can one find nodes \( v \in T_n \) with \( X_v = cn \)?

Bootstrap the law of large numbers:

- Given \( \varepsilon > 0 \), fix \( K = K(\varepsilon) \) large enough that \( \varepsilon M_K < (c+\varepsilon)K \)
- Let \( v(1) \in T_K \) have minimal position among depth-\( K \) nodes:
  \[ X_{v(1)} = M_K \]
- For \( j = 2, \ldots, n/K \), let \( v(j) \in T_{jK} \) have minimal position among depth-\( jK \) descendants of \( v(j-1) \):
  \[ X_{v(j)} - X_{v(j-1)} = M_{(v(j-1)} \]
- Then \( \sum E X_{v(n/K)} = \sum E X_{v(i)} < (c+\varepsilon)n \)

Claim \( \text{Var} X_{v(n/K)} \leq n \cdot 2^K \sigma^2 \).

Proof: By the branching property,

\[ \text{Var} X_{v(n/K)} = \sum_{v \in T_K} \text{Var}(X_{v(1)}), \]

and

\[ \text{Var} X_{v(i)} = \sum_{v \in T_K} \text{Var}(X_{v(i)}), \]

with probability \( (1 - o(1)) \) as \( n \to \infty. \)

Follows that \( X_{v(n/K)} \leq (c + 2\varepsilon)n \)

# node value queries = \( 2^K \cdot n/K = O_\varepsilon(n) \): linear-time algorithm.
CREM and its minima

- Setting: continuous random energy model (CREM) \( \text{CREM}(A, n) \)

- \( A \): Cumulative dist. \( f^n \) of a finite measure on \([0,1] \); so \( A(0) = 0 \), \( A(1) \in (0, \infty) \).

- \( n \): number of levels

- Gaussian process \( (X_v, v \in T_n) \) indexed by \( T_n = T_1, T_2, \ldots, T_n \)

Displacement laws:
- If \( v \in T_k \) then \( X_v - X_{\text{parent}(v)} \) is \( \mathcal{N}(0, n(A(k) - A(k-1)/n)) \).

Displacements mutually independent.

- Idea: along any root-to-leaf path, observe an inhomogeneous Brownian motion whose infinitesimal variance at time \( z \) is \( A'(z) \).
CREM and its minima: Examples

**Standard (binary) Gaussian BRW.**

\[ A(z) = \begin{cases} 0, & z \leq 0 \\ z, & z \in (0, 1) \\ 1, & z \geq 1 \end{cases} \]

\[ \mathbb{E}\{v \in T_n : X_v \leq -x n\} \approx 2^n \exp\left(-\frac{x^2}{2n}\right) \]

\[ n^{-1}M_n \xrightarrow{a.s.} (2 \log 2)^{\frac{1}{2}} \]

**Two-speed concave Gaussian BRW.**

\[ A(z) = \begin{cases} 0, & z \leq 0 \\ 2z, & z \in (0, \frac{1}{2}) \\ 1+4z, & z \in (\frac{1}{2}, 1) \\ 3, & z \geq 1 \end{cases} \]

\[ \mathbb{E}\{v \in T_n : X_v \leq -x n\} \approx 2^n \exp\left(-\frac{x^2}{4n}\right) \]

\[ n^{-1}M_n \xrightarrow{a.s.} -(a \cdot 2 \log 2)^{\frac{1}{2}} \]

**Speed-a Gaussian BRW.**

\[ A(z) = \begin{cases} 0, & z \leq 0 \\ a, & z \in (0, 1) \\ a, & z \geq 1 \end{cases} \]

\[ \mathbb{E}\{v \in T_n : X_v \leq -x n\} \approx 2^n \exp\left(-\frac{x^2}{a^2n}\right) \]

\[ n^{-1}M_n \xrightarrow{a.s.} -(a \cdot 2 \log 2)^{\frac{1}{2}} \]

**Two-speed convex Gaussian BRW.**

\[ A(z) = \begin{cases} 0, & z \leq 0 \\ 4z, & z \in (0, \frac{1}{2}) \\ 2+2z, & z \in (\frac{1}{2}, 1) \\ 3, & z \geq 1 \end{cases} \]

\[ \mathbb{E}\{v \in T_n : X_v \leq -x n\} \approx 2^n \exp\left(-\frac{x^2}{4n}\right) \]

\[ n^{-1}M_n \xrightarrow{a.s.} -(a \cdot 2 \log 2)^{\frac{1}{2}} \]

The needed trajectories do not exist.
Proposition (Bovier-Kurkova; Mallein; LAB-Mallard)

Suppose $A$ is absolutely continuous wrt Lebesgue measure, and has a Riemann-integrable derivative $a$.

Let $\hat{A}$ be the concave hull of $A$, let $\hat{a}$ be the left-derivative of $\hat{A}$.

Then $n^t M_n \xrightarrow{a.s.} \int_0^t \sqrt{\hat{a}(t)} \, dt = -c$

Moreover, $c = \sup \left\{ \int_0^t v(s) \, ds : v : [0,1] \to \mathbb{R} \text{ measurable,} \right. \
\forall t \in (0,1], \int_0^t \frac{v(s)^2}{2a(s)} \, ds \leq t \log 2 \left. \right\}$.

The supremum is attained via the $f^n \, v_{\max} : [0,1] \to \mathbb{R}$ with

$v_{\max}(s) = a(s) \cdot \left( \frac{2\log 2}{\hat{a}(s)} \right)^{\frac{1}{2}}$

N.B.: The value $c$ only depends on $A$ through $\hat{A}$; but the trajectory $v_{\max}$ followed to reach $-c \, n$ within $T$ depends sensitively on $A$. 

CREM: The minimum position
CREM and its minima: The algorithmic barrier

**Def (Mallein):** The natural speed path for $A$ is the function

$$Z(t) = Z_A(t) = \int_0^t (z \log 2)^{\frac{1}{2}} a(t)^{\frac{1}{2}} dt$$

**Theorem:** (LAB, Maillard 2018+)

If $A$ abs. continuous, $a = A'$ Riemann-integrable, then with $Z_* = Z_A(1)$, we have:

1. For all $x < Z_*$, there is a linear-time algorithm that finds $v \in T_n$ with $X_v \leq -x n$ with high probability.
2. For all $x > Z_*$, there is $\gamma = \gamma(A, x) > 0$ s.t. for $n$ large, for any algorithm, the expected # of queries before finding a node $v \in T_n$ with $X_v \leq -x n$ stochastically dominates a $\text{Geometric}(\exp(-\gamma n))$ random variable.

**Proof Idea:**

1. In the inhomogeneous setting, the renormalization search follows the natural speed path.
2. For every node $v \in T_n$ with $X_v \leq -x n$, there is a linear-size subsection of the ancestral trajectory of $v$ along which the slope is unnatural (different from the slope of the natural speed path).

The branching property + Gaussian tail estimates $\Rightarrow$ exponentially unlikely to find such a segment on any single query, even conditionally given past queries.