Convergence of recursive equations via numerical analysis

Louigi Addario-Berry, Luc Devroye, Celine Kerriou, Rivka Maclaine Mitchell

Analysis of Algorithms, CIRM, Luminy, France

June 24, 2019
\[ T = \text{infinite binary canopy tree} \]

- One-way infinite path \( v_0, v_1, v_2, ... \)
- Node \( v_n \) is the root of a complete binary tree of depth \( n \)

\[ L = \text{leaves of } T \]

\[ T_n = \text{subtree rooted at } v_n \]

\[ L_n = \text{leaves of } T_n \]

**Functions on** \( T \):
- Input from children
- Combination function at nodes
- Output to parents

Choose functions \((f_v, \forall v \in T_n \setminus L_n)\); this turns \( T_n \) into a function,

\[ x = (x_v, \forall v \in L_n) \quad \rightarrow \quad T_n(x) \quad \leftarrow \text{output at root } v_n, \text{ on input } x. \]

Either or both of \( x \) and \((f_v, \forall v \in T_n \setminus L_n)\) can be random.
Examples

1. \( f_v = \begin{cases} (a,b) \mapsto 1 + a + b & \text{with prob. } p \\ (a,b) \mapsto 1 & \text{with prob. } 1 - p \end{cases} \)

Then \( T_n(\mathcal{I}) \overset{d}{=} \# \text{ nodes at level } \leq n \) in a Galton-Watson tree with offspring dist. \( \begin{cases} 0 & \text{with probability } 1 - p \\ 2 & \text{with probability } p \end{cases} \)

2. Let \((D_v, v \in T)\) be IID with law \( \mu \), let \( f_v(a,b) = \max(a,b) + D_v \)

Then \( T_n(\mathcal{I}) \overset{d}{=} \) maximum position in generation \( n \) of a binary branching random walk with displacement dist \( \mu \) (displacements at vertices)

3. Let \((D_v, v \in T)\) be IID with law \( \mu \), let \( f_v(a,b) = aD_{v_0} + bD_{v_1} \).

This is a smoothing transform; fixed points studied by Durrett \& Liggett (1983), many others.

In fact, all these equations have been studied from the perspective of fixed-point equations (sometimes wish to introduce a rescaling or shift).
Examples without a fixed-point theory

4. Derrida-Retaux model / "Parking on trees". Here \( f_v(a,b) = \max(a+b-1,0) \)

Question: Large-\( n \) behaviour of \( T_n(X) \) where \( X=(X_v, v\in \mathbb{L}_n) \) i.i.d with some law \( \mu \)

(answer of course depends on \( \mu \))

(Refs: Hu, Mallein, Pain, 1811.08749v2 ; Hu, Shi, 1705.03792 ; Goldschmidt, Przykucki, 1610.08786)

There is exactly one model where this can be analyzed so far.

5. Random hierarchical lattice.

\[ f_v = \begin{cases} (a,b) \mapsto a+b & \text{with prob. } p \\ (a,b) \mapsto \frac{ab}{a+b} & \text{with prob. } 1-p \end{cases} \]

[Ref: Hambly-Jordan 2004. \( p > \frac{1}{2} \) \( \Rightarrow T_n(I) \) grows exponentially; \( p < \frac{1}{2} \) \( \Rightarrow T_n(I) \) decays exp.]

6. Pemantle's Min-Plus tree

\[ f_v = \begin{cases} (a,b) \mapsto a+b & \text{with prob. } p \\ (a,b) \mapsto \min(a,b) & \text{with prob. } 1-p \end{cases} \]

[Ref: Auffinger-Cable: 1709.07849]

(Open question: universality: what happens for other inputs?)

Theorem (A-C) \[ \frac{\log T_n(\bar{a})}{(\pi^2 n/3)^{1/2}} \xrightarrow{d} \text{Beta}(2,1) \]
New model: Hipster random walk

Fix $(D_v, v \in \mathcal{L})$ IID. Let $f_v$ be defined by
\[
(a, b) \xrightarrow{f_v} a + D_v \mathbb{1}_{a=b} \quad \text{with prob. } \frac{1}{2}
\]
\[
(a, b) \xrightarrow{f_v} b + D_v \mathbb{1}_{a=b} \quad \text{with prob. } \frac{1}{2}
\]

Idea: Think of time as running up the tree

1. One of $v_0, v_1$ is hipper than the other (chosen randomly)
2. If another particle shows up, hipper child takes off.

We will study:
- symmetric simple hipster random walk $SSHRW$
- totally asymmetric lazy simple hipster random walk $TALSHRW$

Theorem:
For $SSHRW$, 
\[
\frac{T_n(\delta)}{(36n)^{\frac{1}{3}}} \xrightarrow{d} \text{Beta}(2, 2) - \frac{1}{2}.
\]
For $TALSHRW$, 
\[
\frac{T(\delta)}{(4(1-p)n)^{\frac{1}{2}}} \xrightarrow{d} \text{Beta}(2, 1).
\]
Note Result for TALSHRW very similar to that of Auffinger-Cable.

Recall Auffinger-Cable:

\[
f_v = \begin{cases} 
(a,b) &\mapsto a+b \quad \text{with prob. } p \\
(a,b) &\mapsto \min(a,b) \quad \text{with prob. } 1-p 
\end{cases}
\]

Theorem (A-C) \[
\frac{\log T_n(\delta)}{(\pi^2 n/3)\frac{1}{2}} \,
\xrightarrow{d}
\,
\text{Beta}(2,1)
\]

Intuition Suppose \(T_n(\delta)\) is growing on a (stretched) exponential scale.

Write \(L,R\) for values at children of root of \(T_n\)

If \(|\log L - \log R|\) small then \[
\begin{cases} 
L+R \approx 2L \\
\min(L,R) \approx L \\
\log(L+R) \approx \log(L) + 1 \\
\min(\log L, \log R) \approx \log L 
\end{cases}
\]

This is the common value plus a \(\pm 1\)-valued increment

If \(|\log L - \log R|\) big then \[
\begin{cases} 
L+R \approx \max(L,R) \\
\min(L,R) = \min(L,R) \\
\log(L+R) \approx \max(\log L, \log R) \\
\log(\min(L,R)) = \min(\log L, \log R) 
\end{cases}
\]

This is just \(\log(\text{value of a random child})\)
Similar intuition should work for the hierarchical lattice:

\[ f_v = \begin{cases} (a, b) \rightarrow a+b & \text{with prob. } p \\ (a, b) \rightarrow \frac{ab}{a+b} & \text{with prob. } 1-p \end{cases} \]

**Intuition:** Suppose \( T_n(\delta) \) is growing on a (stretched) exponential scale.

Write \( L, R \) for values at children of root.

If \( |\log L - \log R| \) small then \( \left\{ \begin{array}{l} L+R \approx 2L \\ \frac{LR}{L+R} \approx L/2 \\ \log (L+R) \approx \log (L) + 1 \\ \log \left( \frac{LR}{L+R} \right) \approx \log (L) - 1 \end{array} \right\} \) This is the common value plus a \( \delta-1, \delta \)-valued increment

If \( |\log L - \log R| \) big then \( \left\{ \begin{array}{l} L+R \approx \max(L, R) \\ \frac{LR}{L+R} \approx \min(L, R) \\ \log (L+R) \approx \max(\log L, \log R) \\ \log \left( \frac{LR}{L+R} \right) \approx \min(\log L, \log R) \end{array} \right\} \) This is just log(value of a random child)

Motivates the following conjecture: in the random hierarchical lattice with \( p=\frac{1}{2}, \exists c > 0 \) s.t.

\[ \frac{\log T_n(\delta)}{(c \cdot n)^{\frac{1}{3}}} \xrightarrow{d} \text{Beta}(2,2) \]
Theorem (Totally asymmetric lazy) \[ \frac{T(0)}{(2n)^{\frac{1}{2}}} \xrightarrow{d} \text{Beta}(2,1) \]

Proof Idea

Original dynamics:

- Vo is hipper
- V1 is hipper

By symmetry, can assume left child is always chosen.

For inputs \( x = (x_v, v \in L) \), useful notation: \( T_n(x) \):=

\[ T_n((x_v, v \in L_n)) \]

\[ T_{n+1}'(x) = \begin{cases} 
T_n(x) & \text{if } T_n(x) \neq T_n'(x) \\
T_n(x) + D_v & \text{if } T_n(x) = T_n'(x)
\end{cases} \]
Proof Idea

(Totally asymmetric case)

Let \( p_n(k) = P_n(T_n(\emptyset) = k) \)

Then \( p_{n+1}(k) = p_n(k)(1 - p_n(k)) + \frac{1}{2} p_n(k-1)^2 + \frac{1}{2} p_n(k)^2 \)

Rearranging gives \( p_{n+1}(k) - p_n(k) = \frac{1}{2} (p_n(k)^2 - p_n(k-1)^2) \)

This is a discretization of the inviscid Burgers' equation \( \frac{\partial}{\partial t} u(x,t) = -\frac{1}{2} \frac{\partial}{\partial x} (u(x,t)^2) \)

So we are trying to solve the (measure-valued) initial-value problem

\[
\begin{cases}
  u_t = -u u_x, & t \geq 0, x \in \mathbb{R} \\
  u_0(x) = \delta_0(x) = 1_{[x=0]} 
\end{cases}
\]

(Dirac mass at 0)

Ignoring space-time points of discontinuity, this is solved \(^*\) by \( u : \mathbb{R} \times [0,\infty) \rightarrow \mathbb{R} \) given by

\[
 u(x,t) = \begin{cases} 
  2t, & 0 \leq x < \sqrt{2t} \\
  0, & \text{otherwise} 
\end{cases}
\]

Note \( u(t,x) \) is always a prob. dist.: the density of a scaled Beta(2,1).

\(^*\)But solution is not unique!
Start Burgers’ from a smooth initial condition of the form \( U_0(x) = \frac{x}{t_0} \mathbf{1}_{0 \leq x \leq \sqrt{2t_0}} \) (think of \( t_0 \) as small).

Probabilistically what does this mean?

\( U_0 \) is density of \( \sqrt{2t_0} \cdot B \) where \( B \sim \text{Beta}(2,1) \).

Fix \( M > 0 \) and define \( U_j^0(M) = M \int_{j/M}^{(j+1)/M} U_0(x) \, dx \). for \( j \geq 0 \) st. \( \frac{j}{M} \leq \sqrt{2t_0} \).

Then \( \sum_j U_j^0(M) = 1 \), so \( (U_j^0(M), j \geq 0) \) defines a probability distribution on \( \{0, 1, \ldots, \lfloor M \cdot \sqrt{2t_0} \rfloor \} \).

Let \( X^M = (X_v^M, v \in \mathbb{Z}) \) be vector of IID\s with \( \text{IP}(X_v^M = j) = U_j^0(M) \) (discretization of \( U_0 \) at mesh size \( \frac{1}{M} \)).

\( T_n(X^M) \) is value of TALSHRW when initial distribution is \( \frac{1}{M} \)-mesh discretization of \( \sqrt{2t_0} \cdot B \).

**Lemma** We have \( \text{IP}(T_n(X^M) = j) = \frac{1}{M} \cdot U_j^0(M) \), where \( (U_j^0(M))_{n \geq 0, j \geq 0} \) is defined by the recurrence \( MU_j^{n+1} = M \cdot U_j^n - \frac{1}{2} (U_j^n)^2 - (U_{j-1}^n)^2 \).

**Proof** Easy induction \( \square \)
Second step  Convergence of the fine-mesh approximation.

The spatial mesh is $\frac{1}{M}$. We take a temporal mesh of $\frac{1}{M^2}$.

$$U_M(t, x) = U_{[x, x+\frac{1}{M}]}(M) = P(T_{L^1(M)}(X^m) = [x, x+\frac{1}{M}]) \text{ for } t, x > 0.$$  

Call $U_m$ a $\frac{1}{M}$ - fine mesh approximation of Burgers' equation.

Theorem (Evje & Karlsen, 2000)

From a bounded variation initial condition, the $\frac{1}{M}$ - fine mesh approximation converges to the BV entropy solution $U$ of Burgers' equation almost everywhere on $\mathbb{R} \times (0, \infty)$, and for any compact $C \subset \mathbb{R} \times [0, \infty)$,

$$\int_C |U_m(t, x) - U(x, t)| \, dx \, dt \to 0.$$  

- **BV** → Bounded variation
- **BV entropy solution** → The correct solution of our problem  
  (verifying this takes some work)

Conclusion  

$U_m \to u$ defined by $u(t, x) = \frac{x}{t+1} \mathbb{1}_{0 < x \leq \frac{1}{2(t+1)}}$
Implication for TALSHRW

**Corollary** For $\varepsilon > 0$ small, if $U = \text{Unif}[1 - \varepsilon, 1 + \varepsilon]$ is independent of $X$, then as $M \to \infty$,

$$
\frac{\text{P}(T_{\text{LIMY}}(X^M) \leq |X^M|)}{\sqrt{2(1+2U)}} M \quad \overset{d}{\to} \quad \text{Beta}(2,1).
$$

**Proof**: For any compact $C \subset \mathbb{R} \times [0,\infty)$,

$$
\int \int_C \left| \text{P}(T_{\text{LIMY}}(X^M) = |X^M|) - \frac{x}{t+t_0} \mathbb{1}_{0 \leq x \leq \sqrt{2(t+t_0)}} \right| \; dx \; dt \to 0.
$$

Taking $C = \{(x,t) : 1t-1 \leq \varepsilon, 0 \leq x \leq \sqrt{2(t+t_0)}\}$, this yields by the triangle inequality that

$$
\int_{[0,1]} \int_{[0,1]} \left| \text{P}(T_{\text{LIMY}}(X^M) \leq \alpha \sqrt{2(t+t_0)} M) - \int_0^\alpha \frac{x}{t+t_0} \; dx \right| \frac{1}{2\varepsilon} \; dt \to 0 \quad \text{as} \quad M \to \infty
$$

(There are "discretization errors" coming from the floors, but it's easy to see these tend to 0 as $M \to \infty$.)

Since $U$ has density $\frac{1}{2\varepsilon} \mathbb{1}_{|t-1| \leq \varepsilon}$, the result follows. □
Last step: stochastic domination.

**Proposition**

If \( x = (x_v, v \in \mathcal{L}) \) and \( y = (y_v, v \in \mathcal{L}) \) are such that \( x_v \in \mathbb{E}, y_v \in \mathbb{E} \) and \( x_v \leq y_v \) for all \( v \in \mathcal{L} \), then \( T_n(x) \preceq_{st} T_n(y) \) for all \( n \geq 1 \).

**Proof:** A straightforward induction. □

**Corollary 1**

For all \( n, M \in \mathbb{N} \), \( T_n(X^M) - [\sqrt{2t_0M}] \preceq_{st} T_n(\sigma) \preceq_{st} T_n(\delta) \).

Allows us to compare all-0 input to random input with \( o(M) \) error (recall \( t_0 > 0 \) is fixed but arbitrarily small).

**Corollary 2**

For all \( M \in \mathbb{N} \), \( T_{(1-\varepsilon)M^2}(X^M) \preceq_{st} T_{UM^2}(X^M) \preceq_{st} T_{(1+\varepsilon)M^2}(X^M) \).

Allows us to compare fixed time near \( M^2 \) to random time \( UM^2 \).

Since \( \frac{1}{\sqrt{2(t_0+U)M}} T_{UM^2}(X^M) \xrightarrow{d} \text{Beta}(2,1) \), corollaries yield that \( \frac{T_n(X^M)}{\sqrt{2M}} \xrightarrow{d} \text{Beta}(2,1) \).