Hipster random walks and their ilk

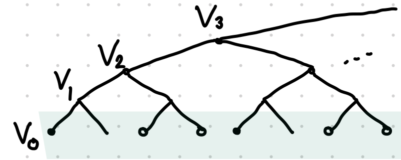
Louigi Addario-Berry, Hannah Cairns, Luc Devroye, Celine Kerriou, Rivka Mitchell

Munich

September 10, 2019

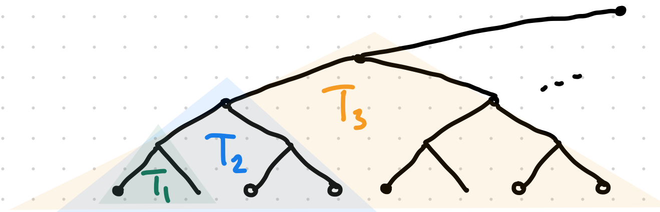
"A recent study found no correlation between the amount of nodding done by the audience during a talk and the amount the audience actually understands." — @FactsOfMath, Twitter, August 8, 2019.

T = infinite binary canopy tree



$\mathcal{L}$ = leaves of T

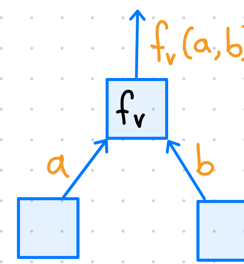
- one-way infinite path $v_0, v_1, v_2, \ldots$
- node v_n is the root of a complete binary tree of depth n



T_n = subtree rooted at v_n

$\mathcal{L}_n$ = leaves of T_n

Functions on T: - input from children
- combination function at nodes
- output to parents



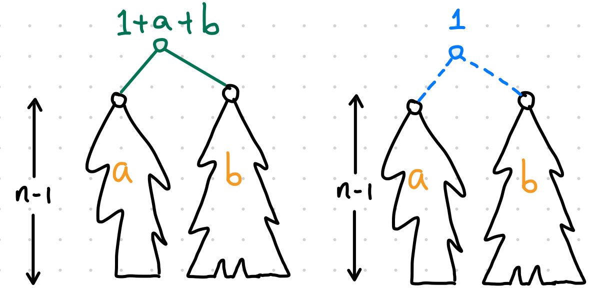
Choose functions $(f_v, v \in T_n \setminus \mathcal{L}_n)$; this turns T_n into a function,

$$x = (x_v, v \in \mathcal{L}_n) \longmapsto T_n(x) \leftarrow \text{output at root } v_n \text{, on input } x.$$

Either or both of x and $(f_v, v \in T_n \setminus \mathcal{L}_n)$ can be random

Examples

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

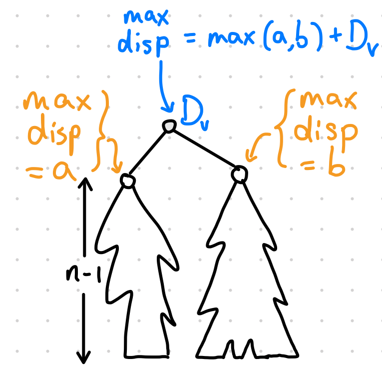


Then $T_n(\vec{1}) \overset{d}{=}$ # nodes at level $\leq n$ in a Galton-Watson tree with offspring dist. $\begin{cases} 2 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$

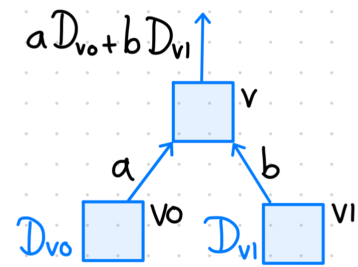
② Let $(D_v, v \in T)$ be IID with law μ, let $f_v(a,b) = \max(a,b) + D_v$

Then $T_n(\vec{0}) \overset{d}{=}$ maximum position in generation $n-1$ of a binary branching random walk with displacement dist μ

(displacements at vertices)



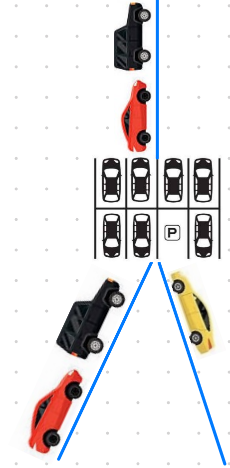
③ Let $(D_v, v \in T)$ be IID with law μ, let $f_v(a,b) = aD_{v0} + bD_{v1}$.



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 $f_v(a,b) = \max(a+b-1, 0)$

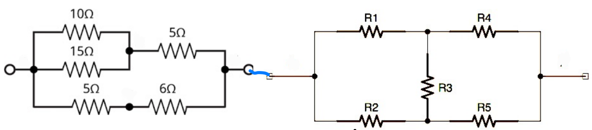
Question: Large-n behaviour of $T_n(X)$ where $X = (X_v, v \in \mathcal{L}_n)$ IID with some law μ

(answer of course depends on μ)

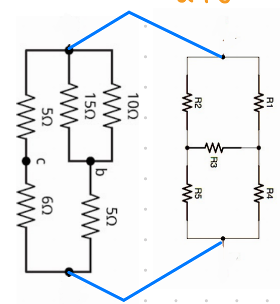
[Refs: Hu, Mallein, Pain, 1811.08749v2 ; Hu, Shi, 1705.03792 ; Goldschmidt, Przykucki, 1610.08786; Chen, Dagard, Derrida, Hu, Lifshits, Shi, 1907.01601]

(5) Random hierarchical lattice.

Series connection Resistance → $a+b$



Parallel connection Resistance → $\frac{ab}{a+b}$



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

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

(6) Pemantle's Min-Plus tree

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

Theorem (A-C) $\dfrac{\log T_n(\vec{1})}{(\pi^2 n/3)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Pemantle conjectured that $\exists c$ s.t. $\dfrac{\log T_n(\vec{1})}{c\, n^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

[Ref: Auffinger-Cable: 1709.07849

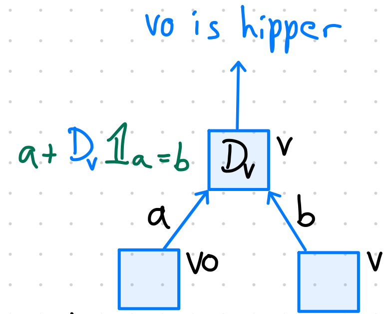
(Open question: universality: what happens for other inputs?)

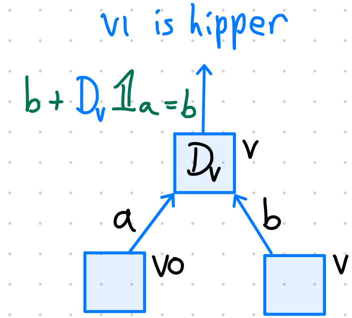
New model   Hipster random walk

Fix $(D_v, v \in \mathcal{L})$ IID. Let f_v be defined by

$(a,b) \xmapsto{f_v} a + D_v \mathbb{1}_{a=b}$ with prob. $\frac{1}{2}$

$(a,b) \xmapsto{f_v} b + D_v \mathbb{1}_{a=b}$ with prob. $\frac{1}{2}$





Idea Think of time as running up the tree

1. One of vo, vi 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 $\longrightarrow$ $D_v = \begin{cases} 1 & \text{w/ prob. } \frac{1}{2} \\ -1 & \text{w/ prob. } \frac{1}{2} \end{cases}$
  SSHRW
- totally asymmetric lazy simple hipster random walk $\longrightarrow$ $D_v = \begin{cases} 1 & \text{w/ prob. } p \\ 0 & \text{w/ prob. } 1-p \end{cases}$ $p \in (0,1)$.
  TALSHRW

Theorem

For SSHRW, for any IID inputs $\vec{X}$, $\dfrac{T_n(\vec{X})}{(36n)^{1/3}} \xrightarrow{d} \text{Beta}(2,2) - \frac{1}{2}$.

For TALSHRW, for any IID inputs $\vec{X}$, $\dfrac{T(\vec{X})}{(4(1-p)n)^{1/2}} \xrightarrow{d} \text{Beta}(2,1)$

Take $\vec{X} = \vec{0}$ below for simplicity.

Note Result for TALSHRW very similar to that of Auffinger-Cable.

Recall Auffinger-Cable:

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

Theorem (A-C) $\dfrac{\log T_n(\vec{1})}{(\pi^2 n/3)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Intuition (connecting min/plus and TALSHRW)

Write L, R for values at children of root of T_n.

If $T_n(\vec{1})$ is growing on a (stretched) exponential scale then it's natural to compare $\log L$ and $\log R$.

Behaviour when $|\log L - \log R|$ small

If $|\log L - \log R| \approx 0$ then $\begin{cases} L+R \approx 2L \\ \min(L,R) \approx L \end{cases}$ $\left.\begin{matrix} \log(L+R) \approx \log(L)+1 \\ \min(\log L, \log R) \approx \log L \end{matrix}\right\}$ This is the common value plus a $\{0,1\}$-valued increment

Behaviour when $|\log L - \log R|$ large

If $|\log L - \log R| \approx \infty$ then $\begin{cases} L+R \approx \max(L,R) \\ \min(L,R) = \min(L,R) \end{cases}$ $\left.\begin{matrix} \log(L+R) \approx \max(\log L, \log R) \\ \log(\min(L,R)) = \min(\log L, \log R) \end{matrix}\right\}$ This is just log(value of a random child)

Similar intuition should work for the hierarchical lattice:

$$f_v = \begin{cases} (a,b) \longmapsto a+b & \text{with prob. } \frac{1}{2} \\ (a,b) \longmapsto \frac{ab}{a+b} & \text{with prob. } \frac{1}{2} \end{cases}$$

<u>Intuition</u>: Suppose $T_n(\vec{1})$ is growing on a (stretched) exponential scale.

Write L, R for values at children of root.

Behaviour when $|\log L - \log R|$ small

If $|\log L - \log R|$ small then $\begin{cases} L+R \approx 2L \\ \frac{LR}{L+R} \approx \frac{1}{2}L \end{cases}$ $\quad \begin{matrix} \log(L+R) \approx \log(L)+1 \\ \log\left(\frac{LR}{L+R}\right) \approx \log(L)-1 \end{matrix}$ $\Big\}$ This is the common value plus a $\{-1,1\}$-valued increment

Behaviour when $|\log L - \log R|$ large

If $|\log L - \log R|$ big then $\begin{cases} L+R \approx \max(L,R) \\ \frac{LR}{L+R} \approx \min(L,R) \end{cases}$ $\quad \begin{matrix} \log(L+R) \approx \max(\log L, \log R) \\ \log\left(\frac{LR}{L+R}\right) = \min(\log L, \log R) \end{matrix}$ $\Big\}$ 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.

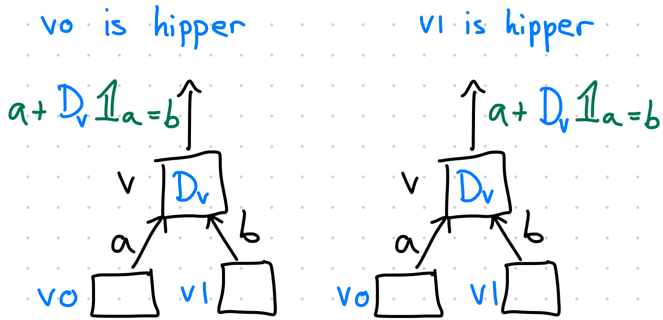
$$\frac{\log T_n(\vec{1})}{(cn)^{1/3}} - \frac{1}{2} \xrightarrow{d} \mathrm{Beta}(2,2)$$

(Disagrees with a conjecture of Hambly-Jordan)

Theorem (Totally asymmetric lazy SHRW) $\dfrac{T(\vec{0})}{(2n)^{1/2}} \xrightarrow{d} \mathrm{Beta}(2,1)$

Proof Idea

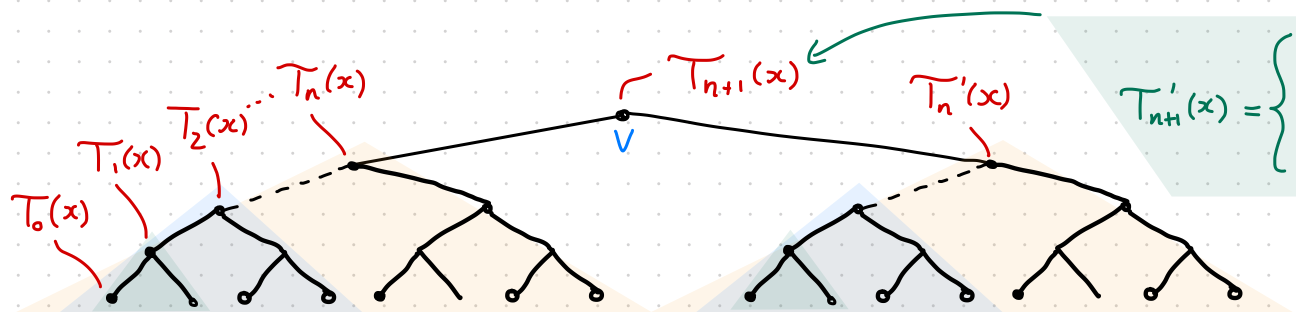
Original dynamics:



$D_v \sim \mathrm{Bernoulli}(\tfrac{1}{2})$

By symmetry, can assume left child is always chosen.

For inputs $x = (x_v, v \in \mathcal{L})$, useful notation: $T_n(x) := T_n((x_v, v \in \mathcal{L}_n))$



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

Proof Idea (Totally asymmetric case)

Let $p_n(k) = \mathbb{P}(T_n(\vec{0}) = k)$

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

$p_n(k)(1-p_n(k))$ ← left child $= k$, right child $\neq k$

$\frac{1}{2}p_n(k-1)^2$ ← both $= k-1$, make a step

$\frac{1}{2}p_n(k)^2$ ← both $= k$, be lazy

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

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

$u_t = -\frac{1}{2}(u^2)_x = -u\,u_x$

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

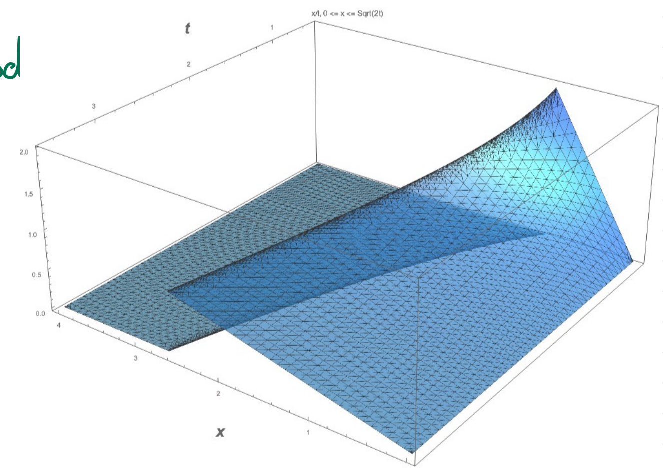
$$\begin{cases} u_t = -u u_x, & t \geq 0,\ x \in \mathbb{R} \\ u_0(x) = \delta_0(x) = \mathbb{1}_{[x=0]} \end{cases}$$

(Dirac mass at 0, understood as a prob. measure)

Ignoring space-time points of discontinuity,

this is solved* by $u : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ given by

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



Note $u(t,x)$ is the density of a scaled Beta(2,1).

Proof Idea (Symmetric simple HRW case)

Let $q_n(k) = \mathbb{P}(T_n(\vec{0}) = k)$

$q_n(k)(1-q_n(k)) \leftarrow$ left child $= k$, right child $\neq k$

$\frac{1}{2} q_n(k-1)^2 \leftarrow$ both $= k-1$, make a $+1$ step

$\frac{1}{2} q_n(k+1)^2 \leftarrow$ both $= k$, make a -1 step

Then $q_{n+1}(k) = q_n(k)(1-q_n(k)) + \frac{1}{2} q_n(k-1)^2 + \frac{1}{2} q_n(k+1)^2$

Rearranging gives $q_{n+1}(k) - q_n(k) = \frac{1}{2}\left(q_n(k+1)^2 - 2q_n(k)^2 + q_n(k-1)^2\right)$

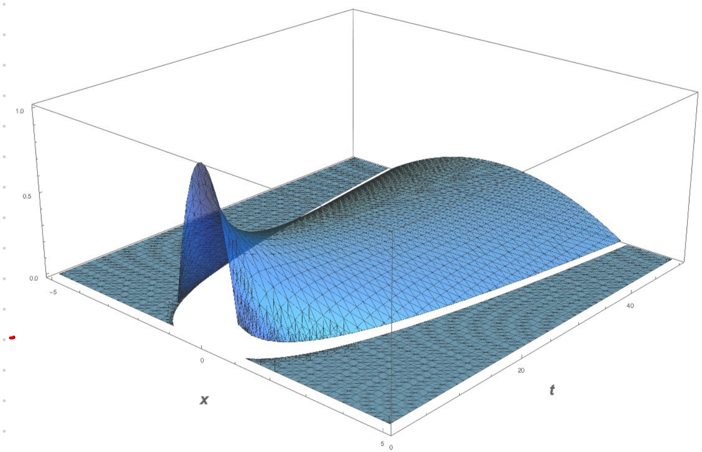
This is a discretization of the porous medium equation $\frac{\partial}{\partial t} u(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(u(x,t)^2\right)$

So we are trying to solve the initial-value problem

$$\begin{cases} u_t = \frac{1}{2}(u^2)_{xx}, & t \geq 0,\ x \in \mathbb{R} \\ u_0(x) = \delta_0(x) = \mathbb{1}_{[x=0]} \end{cases}$$

Target solution: $u(t,x) = \max\left(\left[\frac{3}{4}\left(\frac{2}{9t}\right)^{\frac{1}{3}} - \frac{2x^2}{9t}\right], 0\right)$;

truncated parabola. $\equiv$ Density of a scaled Beta(2,2).



Rest of talk: focus principally on TALSHRW

Inviscid Burgers' equation

Initial value problem $\begin{cases} \frac{\partial}{\partial t} u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\left(u(x,t)^2\right) \\ u(x,0) = \delta_0(x) \end{cases}$ ← Dirac mass at 0.

Note: Any function of the form $u(x,t) = \dfrac{\alpha x + \beta}{\alpha t + \gamma}$

Satisfies $\frac{\partial}{\partial t} u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\left(u(x,t)^2\right)$

Special cases:

$\alpha = 1,\ \beta = \gamma = 0$ $\quad u(x,t) = \dfrac{x}{t}$ $\quad \alpha > 0$: solution flattens out

$\alpha = \beta = 0,\ \gamma = 1$ $\quad u(x,t) = 0$ $\quad \alpha = 0$: flat line

} Target: a mixture of these two solutions.

$\alpha = -1,\ \beta = \alpha = 1$ $\quad u(x,t) = \dfrac{1-x}{1-t}$ $\quad \alpha > 0$: solution steepens with time (problem at $t = 1$ ...)

Why should $((p_n(k), k \in \mathbb{Z}), n \geq 0)$ converge to the claimed solution?
This is by no means obvious. (A bit like Itô vs. Stratonovich integrals.)

Warning example: solve $\boxed{p_{n+1}(k) - p_n(k) = -\frac{1}{2}\left(p_n(k)^2 - p_n(k-1)^2\right)}$ with $p_0(k) = \begin{cases} 2, & k = 0 \\ 0 & k \neq 0 \end{cases}$

Get $p_n(k) = \begin{cases} 2, & k = n \\ 0 & k \neq n \end{cases}$

Heuristic "naturally arising" difference equations pick out "physical" solutions

What does "solution" mean?

$$\frac{\partial}{\partial t} u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\left(u(x,t)^2\right)$$

$$u(x,0) = \delta_0(x)$$

Potential solutions: functions of bounded variation.

"Locally integrable functions u whose generalized derivatives are locally measures." (Volpert 1967)

This means:

$\exists$ a Radon measure $\underbrace{\nabla u}_{\text{Notation!}} = ((\nabla u)_x, (\nabla u)_t)$ on $\mathbb{R}\times[0,\infty)$, taking values in $\mathbb{R}^2$, s.t.

- $|\nabla u|$ is locally finite
- For any C^∞ test f^n $\phi : \mathbb{R}\times[0,\infty) \longrightarrow \mathbb{R}$ with compact support,

$$\left(\int u \frac{\partial}{\partial x}\phi(x,t), \int u \frac{\partial}{\partial t}\phi(x,t)\right) =: \int u(x,t)\,\nabla\phi(x,t)\,dx\,dt = -\int \phi(x,t)\,\nabla u(x,t) =: -\left(\int \phi(x,t)(\nabla u)_x(x,t), \int \phi(x,t)(\nabla u)_t(x,t)\right)$$

What does "physical" mean?

Viscosity solution add Gaussian noise and take a small-noise limit.

Solve $\begin{cases} \frac{\partial}{\partial t} u(x,t) = -\frac{1}{2}\frac{\partial}{\partial x}\left(u(x,t)^2\right) + \varepsilon \frac{\partial^2}{\partial x^2} u(x,t) \\ u(x,0) = \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-x^2/2\varepsilon\right) \end{cases}$; let $\varepsilon \to 0$ and hope for the best.

(Not a helpful perspective in our setting.)

"Entropy"/generalized solution Mathematically formalizes that "in a fluid, shocks increase disorder / have a scattering effect".

Consider a Cauchy problem of the form $\circledast$ $\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2} A(u(x,t)) - \frac{\partial}{\partial x} b\,u((x,t)) \\ u(x,0) = u_0(x) \end{cases}$ $\leftarrow$ bounded measurable $f^{\underline{n}}$.

A generalized solution of $\circledast$ is a weak solution u s.t. for all $c \in \mathbb{R}$, the following holds.

Let Γ_u = set of discontinuities of u. Let $\nu = (\nu_x, \nu_t)$ be the normal to Γ_u.

Then $\boxed{(\mathrm{sign}(u^+ - c) - \mathrm{sign}(u^- - c))((\bar{u} - c)\nu_t + \overline{b(u(x,t))} - b(c))\nu_x] \le 0}$, "Scattering condition near discontinuities"

mean value in direction $\pm\nu$ symmetric mean value

in that the 1-dimensional Hausdorff capacity of the set of points where this fails is zero.

Volpert (2000): Proves uniqueness of the generalized solution under weak conditions on A, b.

In Volpert's result, initial condition must be a f^n; can't start from the measure δ_0.

First step Start Burgers' from a smoother initial condition of the form $u_0(x) = \frac{x}{t_0} \mathbb{1}_{0 \le x \le \sqrt{2t_0}}$

(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 $j \ge 0$ s.t. $\frac{j}{M} \le \sqrt{2t_0}$

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

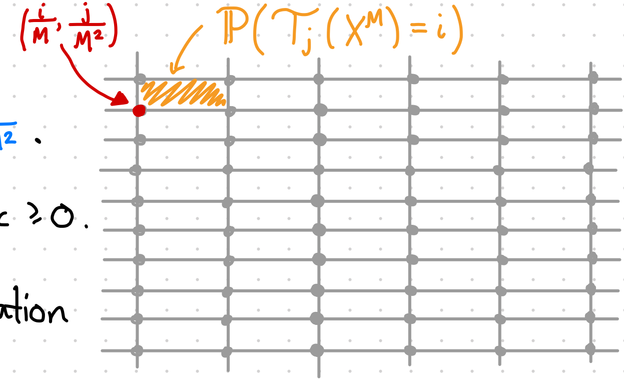
Let $X^M = (X_v^M, v \in \mathbb{Z})$ be vector of IIDs with $\mathbb{P}(X_v^M = j) = u_j^0(M)$ (discretization of u_0 at mesh size $\frac{1}{M}$)

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

Lemma We have $\mathbb{P}(T_n(X^M) = j) = \frac{1}{M} \cdot u_j^n(M)$, where $(u_j^n(M))_{n \ge 0, j \ge 0}$ is defined by the recurrence $M \cdot u_j^{n+1} = M \cdot u_j^n - \frac{1}{2}\left((u_j^n)^2 - (u_{j-1}^n)^2\right)$.

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 $\frac{1}{M^2}$.

$$\mathcal{U}_M(t,x) = u^{\lfloor tM^2 \rfloor}_{\lfloor xM \rfloor}(M) = \mathbb{P}(T_{\lfloor tM^2 \rfloor}(X^M) = \lfloor xM \rfloor) \quad \text{for } t,x \geq 0.$$

Call $\mathcal{U}_M$ 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 generalized solution u of Burgers' equation almost everywhere on $\mathbb{R} \times [0,\infty)$, and for any compact $C \subset \mathbb{R} \times [0,\infty)$, $\int_C |\mathcal{U}^m(x,t) - u(x,t)| \, dx \, dt \longrightarrow 0$.

Generalized solution $\rightarrow$ The correct solution of our problem (this requires verification but is basically technical)

Conclusion $\mathcal{U}_M \rightarrow u$ defined by $u(t,x) = \frac{x}{t+t_0} \mathbb{1}_{0 \leq x \leq \sqrt{2(t+t_0)}}$

Evje & Karlsen in fact prove convergence for general monotone finite difference approximations of Cauchy problems of the form $\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2} Au((x,t)) - \frac{\partial}{\partial x} b(u(x,t))$, with smooth initial condition.

So we can also use their result when we study the SHRW.

Implication for TALSHRW

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

$$\frac{T_{\lfloor UM^2 \rfloor}(X^M)}{\sqrt{2(t_0+U)}\,M} \xrightarrow{d} \mathrm{Beta}(2,1).$$

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

$$\iint_C \left| \mathbb{P}(T_{\lfloor tM^2\rfloor}(X^M) = \lfloor xM\rfloor) - \frac{x}{t+t_0}\mathbb{1}_{0\le x\le\sqrt{2(t+t_0)}} \right| dx\,dt \longrightarrow 0$$

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

$$\int_{[1-\varepsilon,1+\varepsilon]} \left|\mathbb{P}(T_{\lfloor tM^2\rfloor}(X^M)\le a\sqrt{2(t+t_0)}\,M) - \int_0^a \frac{x}{t+t_0}dx\right| \cdot\frac{1}{2\varepsilon}dt \longrightarrow 0 \quad \text{as } M\to\infty$$

(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 $\frac{1}{2\varepsilon}\mathbb{1}_{|t-1|\le\varepsilon}$, the result follows. $\square$

Last step: stochastic domination.

Proposition

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

Proof: A straightforward induction. $\blacksquare$

Corollary 1 For all $n, M\in\mathbb{N}$, $T_n(X^M)-\lfloor\sqrt{2t_0M}\rfloor \preccurlyeq_{st} T_n(\vec{0}) \preccurlyeq_{st} T_n(\vec{0})$ $\blacksquare$

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 $M\in\mathbb{N}$, $T_{(1-\varepsilon)M^2}(X^M)\preccurlyeq_{st} T_{UM^2}(X^M)\preccurlyeq_{st} T_{(1+\varepsilon)M^2}(X^M)$

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 $\frac{T_n(X^M)}{\sqrt{2}\,M}\xrightarrow{d}\text{Beta}(2,1)$.

Stochastic domination argument more delicate for SSHRW as dynamics non-monotone, but core idea of the argument is the same.


THANKS
That's all Folks!