Choices and Intervals : an interval fragmentation with dependence

Pascal Maillard (Univ. Paris-Sud, Université Paris-Saclay)

Paris-Bath branching structure meeting, June 28, 2016

joint work with

Elliot Paquette
(Weizmann Institute of Science)

to appear in *Israel J. Math.*
+ work in progress

Power of choices

Consider a random structure formed by adding objects **one after the other** according to some randomized rule. Examples:

1. balls-and-bins model: n bins, place balls one after the other into bins, for each ball choose bin uniformly at random (maybe with size-biasing)
2. random graph growth: n vertices, add (uniformly chosen) edges one after the other.
3. interval fragmentation: unit interval $[0,1]$, add uniformly chosen points one after the other $\to$ fragmentation of the unit interval.

Power of choices (2)

Aim: Changing behaviour of model by applying a different rule when adding objects

1. balls-and-bins model: n bins, at each step choose two bins uniformly at random and place ball into bin with fewer/more balls.
   Azar, Broder, Karlin, Upfal '99; D'Souza, Krapivsky, Moore '07; Malyshkin, Paquette '13
2. random graph growth: n vertices, at each step uniformly sample two possible edges to add, choose the one that (say) minimizes the product of the sizes of the components of its endvertices.
   Achlioptas, D'Souza, Spencer '09; Riordan, Warnke '11+'12
3. interval fragmentation: unit interval $[0,1]$, at each step, uniformly sample two possible points to add, choose the one that falls into the larger/smaller fragment determined by the previous points.
   $\rightarrow$ **this talk**

Balls-and-bins model

n bins, place n balls one after the other into bins.

- Model A: For each ball, choose bin uniformly at random.
- Model B: For each ball, choose two bins uniformly at random and place ball into bin with more balls.
- Model C: For each ball, choose two bins uniformly at random and place ball into bin with fewer balls.

How many balls in bin with largest number of balls?

- Model A:
- Model B:
- Model C:

Balls-and-bins model

n bins, place n balls one after the other into bins.

- Model A: For each ball, choose bin uniformly at random.
- Model B: For each ball, choose two bins uniformly at random and place ball into bin with more balls.
- Model C: For each ball, choose two bins uniformly at random and place ball into bin with fewer balls.

How many balls in bin with largest number of balls?

- Model A: $\approx \log n / \log\log n$
- Model B:
- Model C:

Balls-and-bins model

n bins, place n balls one after the other into bins.

- Model A: For each ball, choose bin uniformly at random.
- Model B: For each ball, choose two bins uniformly at random and place ball into bin with more balls.
- Model C: For each ball, choose two bins uniformly at random and place ball into bin with fewer balls.

How many balls in bin with largest number of balls?

- Model A: $\approx \log n / \log \log n$
- Model B: $\approx \log n / \log \log n$
- Model C:

Balls-and-bins model

n bins, place n balls one after the other into bins.

- Model A: For each ball, choose bin uniformly at random.
- Model B: For each ball, choose two bins uniformly at random and place ball into bin with more balls.
- Model C: For each ball, choose two bins uniformly at random and place ball into bin with fewer balls.

How many balls in bin with largest number of balls?

- Model A: $\approx \log n / \log \log n$
- Model B: $\approx \log n / \log \log n$
- Model C: $O(\log \log n)$

Balls-and-bins model, Model C

Model C: For each ball, choose two bins uniformly at random and place ball into bin with fewer balls.

Theorem (Azar, Broder, Karlin, Upfal '99)

Maximum number of balls in a bin: $O(\log \log n)$

Heuristic reason: Let N_k be the number of bins containing at least k balls at the end of the game. Then one should expect

$$N_{k+1} \asymp Cn \left(\frac{N_k}{n} \right)^2,$$

This yields

$$N_k \asymp \rho^{2^k} n$$

for some $\rho < 1$. Solving $N_k = 1$ yields $k = O(\log \log n)$.

Ψ-process: definition

X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

Ψ-process: definition

X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

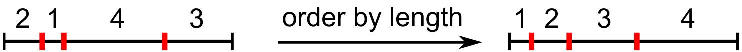
1. Step 1: empty unit interval $[0,1]$

Ψ-process: definition

X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

1. Step 1: empty unit interval $[0,1]$
2. Step n: $n-1$ points in interval, splitting it into n fragments

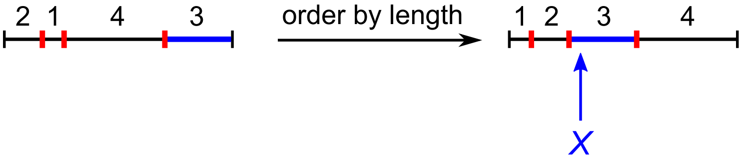
Ψ-process: definition



X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

1. Step 1: empty unit interval $[0,1]$
2. Step n: $n-1$ points in interval, splitting it into n fragments
3. Step $n+1$:
   - Order intervals/fragments according to length

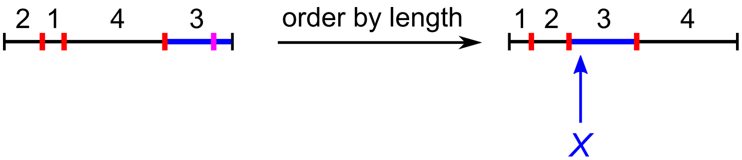
Ψ-process: definition



X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

1. Step 1: empty unit interval $[0,1]$
2. Step n: $n-1$ points in interval, splitting it into n fragments
3. Step $n+1$:
   - Order intervals/fragments according to length
   - Choose an interval according to (copy of) random variable X

Ψ-process: definition



X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

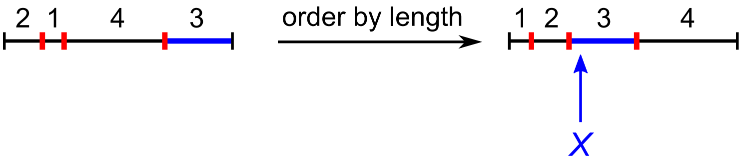
1. Step 1: empty unit interval $[0,1]$
2. Step n: $n-1$ points in interval, splitting it into n fragments
3. Step $n+1$:
   - Order intervals/fragments according to length
   - Choose an interval according to (copy of) random variable X
   - Split this interval at a uniformly chosen point.

Ψ-process: definition

X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

1. Step 1: empty unit interval $[0,1]$
2. Step n: $n-1$ points in interval, splitting it into n fragments
3. Step $n+1$:
   - Order intervals/fragments according to length
   - Choose an interval according to (copy of) random variable X
   - Split this interval at a uniformly chosen point.

Ψ-process: examples



X: random variable on $[0,1]$, $\Psi(x) = P(X \leq x)$.

- $\Psi(x) = x$: uniform process
- $\Psi(x) = \mathbf{1}_{x \geq 1}$: *Kakutani process*
- $\Psi(x) = x^k$, $k \in \mathbb{N}$: max-k-process (maximum of k intervals)
- $\Psi(x) = 1 - (1-x)^k$, $k \in \mathbb{N}$: min-k-process (minimum of k intervals)

Uniform process

$\Psi(x) = x$. Same as n iid points $X_1, \ldots, X_n$ distributed according to Unif$(0,1)$. Trivial fact (law of large numbers / Glivenko–Cantelli): the points $X_1, \ldots, X_n$ are **asymptotically equidistributed**, i.e.:

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i} \to \mathrm{Unif}(0,1), \quad \text{almost surely as } n \to \infty.$$

Furthermore, if $I_1^{(n)}, \ldots, I_{n+1}^{(n)}$ denote the lengths of the intervals at time n, then (Weiss '55, Blum '55)

$$\frac{1}{n+1}\sum_{i=1}^{n+1} \delta_{(n+1)I_i^{(n)}} \to \mathrm{Exp}(1), \quad \text{almost surely as } n \to \infty.$$

Easy proof using the fact that if $0 < T_1 < T_2 < \ldots$ are the arrival times of a Poisson process with intensity 1, then $(T_1/T_{n+1}, \ldots, T_n/T_{n+1})$ have the same law as the order statistics of $X_1, \ldots, X_n$.

Kakutani process

$\Psi(x) = \mathbf{1}_{x \geq 1}$. Always the largest interval splits. Known:

- The splitting points are asymptotically equidistributed
  Lootgieter '77, van Zwet '78, Slud '78,'82
- The length of a typical interval (rescaled by $n+1$) converges to $\mathrm{Unif}(0,2)$ Pyke '80

Basic property exploited in the proofs: Given the intervals at time n, the intervals evolve independently at future times (modulo time-change). Best exploited by stopping the process at the time

$$N_t = \inf\{n : \text{all intervals at time } n \text{ are of length} \leq t\}$$

Basically everything can be proven by studying first and second moment of N_t van Zwet '78.

Not adaptable to Ψ-process.

Other interval fragmentations

There is a large literature (see e.g. Bertoin '06) on interval fragmentation processes which process a certain branching structure: an interval of length L splits at rate $f(L)$ into pieces of random length, independently of the other intervals.

Example: a self-similar fragmentation (Brennan, Durrett '86)

An interval of length L splits at rate L^α into two pieces of respective lengths LU and $L(1-U)$ where $U \sim \mathrm{Unif}(0,1)$, $\alpha \geq 0$.

- $\alpha = 0$: homogenenous interval fragmentation
- $\alpha = 1$: uniform process
- "$\alpha = \infty$": Kakutani process

Known: asymptotic equidistribution + length of typical interval. Branching structure + self-similarity allows for exact calculations, e.g. of moments of the length of a typical interval. Not possible for Ψ-process.

Main result

$I_1^{(n)}, \ldots, I_n^{(n)}$: lengths of intervals after step n.

$$\mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{n \cdot I_k^{(n)}}$$

Main theorem

Assume Ψ is continuous + polynomial decay of $1 - \Psi(x)$ near $x = 1$.

1. μ_n (weakly) converges almost surely as $n \to \infty$ to a deterministic probability measure μ^Ψ on $(0, \infty)$.
2. Set $F^\Psi(x) = \int_0^x y\, \mu^\Psi(dy)$. Then F^Ψ is C^1 and

$$(F^\Psi)'(x) = x \int_x^\infty \frac{1}{z}\, d\Psi(F^\Psi(z)).$$

Properties of limiting distribution

Write $\mu^{\Psi}(dx) = f^{\Psi}(x)\,dx$.

max-k-process ($\Psi(x) = x^k$)

$$f^{\Psi}(x) \sim C_k \exp(-kx), \qquad \text{as } x \to \infty.$$

min-k-process ($\Psi(x) = 1 - (1-x)^k$)

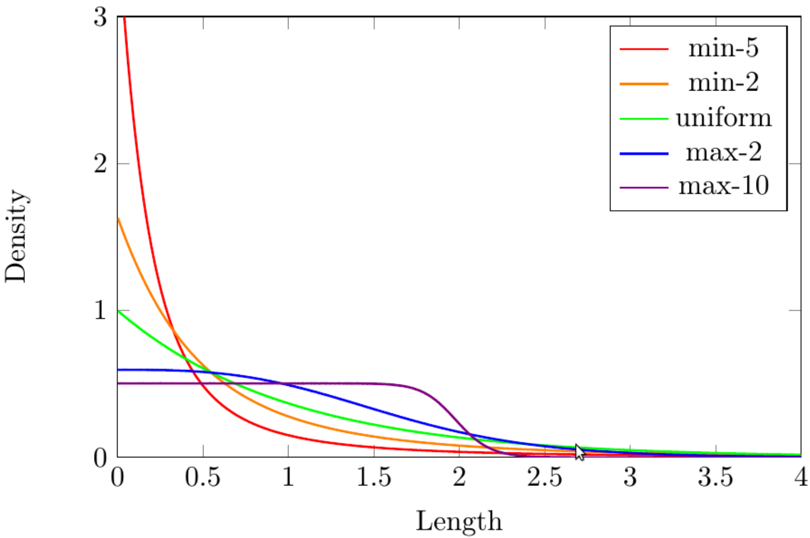
$$f^{\Psi}(x) \sim \frac{c_k}{x^{2+\frac{1}{k-1}}}, \qquad \text{as } x \to \infty.$$

convergence to Kakutani

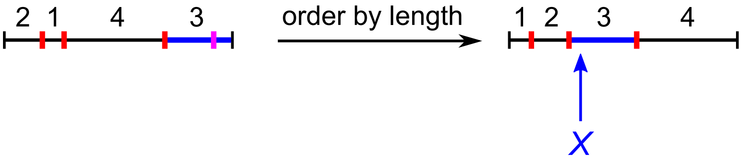
If $(\Psi_n)_{n\geq 0}$ s.t. $\Psi_n(x) \to \mathbf{1}_{x\geq 1}$ pointwise, then

$$f^{\Psi_n}(x) \to \frac{1}{2}\mathbf{1}_{x\in[0,2]}, \qquad \text{as } n \to \infty.$$

Properties of limiting distribution (2)



A good observable: the size-biased distribution function



Define the distribution function of the size-biased interval distribution:

$$\widetilde{A}_n(x) = \sum_{k=1}^{n} I_k^{(n)} \mathbf{1}_{(I_k^{(n)} \le x)}.$$

Suppose we split at time $n+1$ an interval determined by $X \sim d\Psi$. Chance of splitting an interval of length $\in [z, z+dz]$: $d\Psi(\widetilde{A}_n(z))$. Hence

$$\mathbb{E}[\widetilde{A}_{n+1}(x) - \widetilde{A}_n(x)] = x^2 \int_x^\infty \frac{1}{z} d\Psi(\widetilde{A}_n(z)).$$

Embedding in continuous time

Points arrive according to Poisson process with rate e^t.
N_t: number of intervals at time t
$I_1^{(t)}, \dots, I_{N_t}^{(t)}$: lengths of intervals at time t.
Size-biased distribution function

$$\widetilde{A}_t(x) = \sum_{k=1}^{N_t} I_k^{(t)} \mathbf{1}_{(I_k^{(t)} \le x)}$$

Key fact: $\widetilde{\boldsymbol{A}} = (\widetilde{A}_t)_{t\ge 0}$ satisfies the following stochastic integral equation:

$$\widetilde{A}_t(x) = \widetilde{A}_0(x) + \int_0^t e^s x^2 \left[\int_x^\infty \frac{1}{z} d\Psi(\widetilde{A}_s(z)) \right] ds + \widetilde{M}_t(x),$$

for some centered noise $\widetilde{M}_t$.
Note: From now on, **boldface** for *processes* of distribution functions.

A stochastic evolution equation

Set $A_t(x) = \widetilde{A}_t(e^{-t}x)$. Then $\boldsymbol{A} = (A_t)_{t\geq 0}$ satisfies the following stochastic evolution equation:

$$A_t(x) = A_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^{\infty} \frac{1}{z} d\Psi(A_s(z))\right] ds + M_t(x),$$

with $M_t(x) = \widetilde{M}_t(e^{-t}x)$, or

$$\partial_t A_t(x) = -x\partial_x A_t(x) + x^2 \int_x^{\infty} \frac{1}{y} d\Psi(A_t(y)) + \dot{M}_t(x).$$

A stochastic evolution equation

Set $A_t(x) = \widetilde{A}_t(e^{-t}x)$. Then $\boldsymbol{A} = (A_t)_{t\geq 0}$ satisfies the following stochastic evolution equation:

$$A_t(x) = A_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^{\infty} \frac{1}{z} d\Psi(A_s(z)) \right] ds + M_t(x),$$

with $M_t(x) = \widetilde{M}_t(e^{-t}x)$, or

$$\partial_t A_t(x) = -x\partial_x A_t(x) + x^2 \int_x^{\infty} \frac{1}{y} d\Psi(A_t(y)) + \dot{M}_t(x).$$

Claim

A_t converges almost surely to a deterministic limit as $t \to \infty$.

Deterministic evolution

Let $\boldsymbol{F} = (F_t)_{t\geq 0}$ be solution of

$$F_t(x) = F_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^\infty \frac{1}{z} d\Psi(F_s(z))\right] ds$$
$$=: \mathscr{S}^\Psi(\boldsymbol{F})_t.$$

($\Longrightarrow$ solution = fixed point of $\mathscr{S}^\Psi$). Define the following norm:

$$\|f\|_{x^{-2}} = \int_0^\infty x^{-2}|f(x)|\,dx.$$

Lemma

Let $\boldsymbol{F}$ and $\boldsymbol{G}$ be solutions of the above equation. For every $t \geq 0$,

$$\|F_t - G_t\|_{x^{-2}} \leq e^{-t}\,\|F_0 - G_0\|_{x^{-2}}.$$

In particular: $\exists! F^\Psi : F_t \to F^\Psi$ as $t \to \infty$.

Deterministic evolution

Let $\boldsymbol{F} = (F_t)_{t\geq 0}$ be solution of

$$F_t(x) = F_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^\infty \frac{1}{z} d\Psi(F_s(z))\right] ds$$
$$=: \mathscr{S}^\Psi(\boldsymbol{F})_t.$$

($\Longrightarrow$ solution = fixed point of $\mathscr{S}^\Psi$). Define the following norm:

$$\|f\|_{x^{-2}} = \int_0^\infty x^{-2}|f(x)|\,dx.$$

Note: if $f(x) = \int_0^x y\mu(dy)$, then $\|f\|_{x^{-2}} = \mu(\mathbb{R}_+)$

Lemma

Let $\boldsymbol{F}$ and $\boldsymbol{G}$ be solutions of the above equation. For every $t \geq 0$,

$$\|F_t - G_t\|_{x^{-2}} \leq e^{-t}\|F_0 - G_0\|_{x^{-2}}.$$

In particular: $\exists! F^\Psi : F_t \to F^\Psi$ as $t \to \infty$.

Stochastic evolution - stochastic approximation

$$\begin{aligned} A_t(x) &= A_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^{\infty} \frac{1}{z} d\Psi(A_s(z))\right] ds + M_t(x) \\ &= \mathscr{S}^{\Psi}(\boldsymbol{A})_t + M_t(x). \end{aligned}$$

Problem

Cannot control noise M_t using the norm $\|\cdot\|_{x^{-2}}$!

$\Longrightarrow$ cannot use previous Lemma (or the like) for $\boldsymbol{A}$

$\Longrightarrow$ no quantitative estimates to prove convergence.

Stochastic evolution - stochastic approximation

$$
\begin{aligned}
A_t(x) &= A_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^\infty \frac{1}{z} d\Psi(A_s(z))\right] ds + M_t(x) \\
&= \mathscr{S}^\Psi(\boldsymbol{A})_t + M_t(x).
\end{aligned}
$$

Problem

Cannot control noise M_t using the norm $\|\cdot\|_{x^{-2}}$!

$\Longrightarrow$ cannot use previous Lemma (or the like) for $\boldsymbol{A}$

$\Longrightarrow$ no quantitative estimates to prove convergence.

Still possible to prove convergence by adaptation of **Kushner-Clark method** for stochastic approximation algorithms.

Kushner–Clark method (an ∞-dimensional version)

Shifted evolutions $\boldsymbol{A}^{(n)} = (A_{t+n})_{t\geq 0}$. Show:

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact in a suitable functional space

   (the space of maps from $[0,\infty)$ to {the space of subdistribution functions endowed with local L^1 convergence} endowed with local uniform convergence)

2. $\mathscr{S}^{\Psi}$ is continuous in this functional space.
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^{\Psi}(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.

This entails that every subsequential limit $\boldsymbol{A}^{(\infty)}$ of $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is a fixed point of $\mathscr{S}^{\Psi}$.

Kushner–Clark method (an ∞-dimensional version)

Shifted evolutions $\boldsymbol{A}^{(n)} = (A_{t+n})_{t\geq 0}$. Show:

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact in a suitable functional space
   (the space of maps from $[0,\infty)$ to {the space of subdistribution functions endowed with local L^1 convergence} endowed with local uniform convergence)
2. $\mathscr{S}^\Psi$ is continuous in this functional space.
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^\Psi(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.

This entails that every subsequential limit $\boldsymbol{A}^{(\infty)}$ of $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is a fixed point of $\mathscr{S}^\Psi$. If in addition:

4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Then $\left\|A_0^{(\infty)}\right\|_{x^{-2}} = 1$. By previous lemma: $A_t^{(\infty)} \to F^\Psi$, $t \to \infty$ (in fact, $\boldsymbol{A}^{(\infty)} \equiv F^\Psi$). Hence, $A_t \to F^\Psi$, $t \to \infty$.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
2. $\mathscr{S}^\Psi$ is continuous in this functional space.
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^\Psi(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
   equicontinuity + Arzelà-Ascoli
2. $\mathscr{S}^\Psi$ is continuous in this functional space.
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^\Psi(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
   equicontinuity + Arzelà-Ascoli
2. $\mathscr{S}^{\Psi}$ is continuous in this functional space.
   continuity of Ψ + $\mathscr{S}^{\Psi}$ is integral operator
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^{\Psi}(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
   equicontinuity + Arzelà-Ascoli
2. $\mathscr{S}^\Psi$ is continuous in this functional space.
   continuity of Ψ + $\mathscr{S}^\Psi$ is integral operator
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^\Psi(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
   control of noise M_t in norm $\|\cdot\|_{x^{-3}}$
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
   equicontinuity + Arzelà-Ascoli
2. $\mathscr{S}^{\Psi}$ is continuous in this functional space.
   continuity of Ψ + $\mathscr{S}^{\Psi}$ is integral operator
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^{\Psi}(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
   control of noise M_t in norm $\|\cdot\|_{x^{-3}}$
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
   $\|A_t\|_{x^{-2}} = e^{-t}N_t \to 1$, $t \to \infty$
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.

Kushner–Clark method (an ∞-dimensional version)

1. Almost surely, the family $(\boldsymbol{A}^{(n)})_{n\in\mathbb{N}}$ is precompact.
   equicontinuity + Arzelà-Ascoli
2. $\mathscr{S}^{\Psi}$ is continuous in this functional space.
   continuity of Ψ + $\mathscr{S}^{\Psi}$ is integral operator
3. $\boldsymbol{A}^{(n)} - \mathscr{S}^{\Psi}(\boldsymbol{A}^{(n)}) \to 0$ almost surely as $n \to \infty$.
   control of noise M_t in norm $\|\cdot\|_{x^{-3}}$
4. $\|A_t\|_{x^{-2}} \to 1$ a.s., $t \to \infty$.
   $\|A_t\|_{x^{-2}} = e^{-t}N_t \to 1,\ t \to \infty$
5. The family of distribution functions $(A_t)_{t\geq 0}$ is tight.
   next slide

Tightness

Tightness of $(A_t)_{t\geq 0}$ is shown by entropy bounds. Define

$$H_t = \int_0^\infty (\log x)\, dA_t(x) = \sum_{k=1}^n I_k^{(t)} \log(e^t I_k^{(t)}).$$

$\exists$ simple expression for this quantity (valid for *any* interval splitting process), already used by Lootgieter '77; Slud '78.

Lemma

The semimartingale decomposition of $(H_t)_{t\geq 0}$ *is*

$$dH_t = (1 - D_t)dt + dM_t, \quad \text{with } D_t = \frac{1}{2}\int_0^\infty z\, d\Psi(A_t(z)),$$

where $M = (M_t)_{t\geq 0}$ *is a martingale with previsible quadratic variation* $d\langle M\rangle_t \leq L_t D_t\, dt$, L_t = *length of the longest interval at time* t.

Properties of limiting distribution: max-k-process

$\Psi(x) = x^k$. $F^\Psi(x) = \int_0^x y\,\mu^\Psi(dy)$. Recall:

$$(F^\Psi)'(x) = x\int_x^\infty \frac{1}{z}\,d\Psi(F^\Psi(z)).$$

This yields the differential equation

$$(F^\Psi)''(x) = \left(\frac{1}{x} - \Psi'(F^\Psi(x))\right)(F^\Psi)'(x).$$

At $x = \infty$: $\Psi'(F^\Psi(\infty)) = \Psi'(1) = k$. Integrating this equation yields

$$\begin{aligned}(F^\Psi)'(x) &= (F^\Psi)'(1)\exp\left(\int_1^x \left[\frac{1}{y} - \Psi'(F^\Psi(y))\right] dy\right)\\ &\sim C_k x\exp(-kx), \qquad \text{as } x\to\infty,\end{aligned}$$

for some (implicit) constant C_k.

Properties of limiting distribution: min-k-process

$\Psi(x) = 1 - (1-x)^k$. $\Psi'(1) = 0$. The point $x = \infty$ is a critical point for the previous differential equation. Way out: substitute

$$1 - F^{\Psi}(x) = G(\log x^{1/(k-1)})/x^{1/(k-1)}.$$

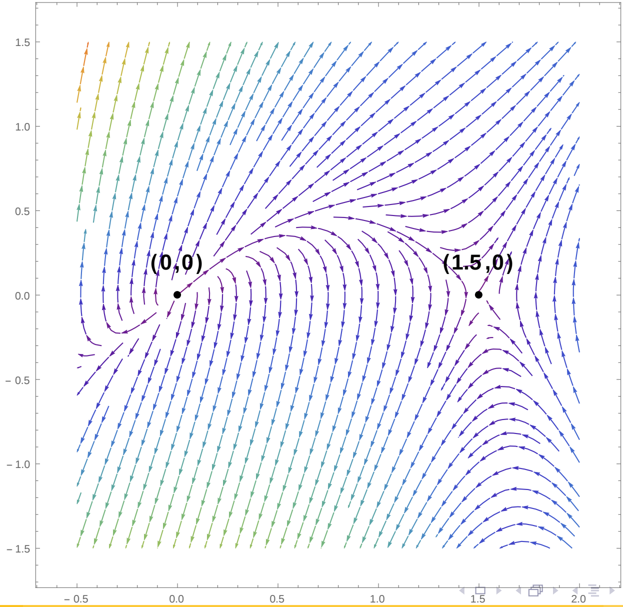
Then G solves the autonomous equation

$$G'' - G' - (2k-1-k(k-1)G^{k-1})(G' - G) = 0.$$

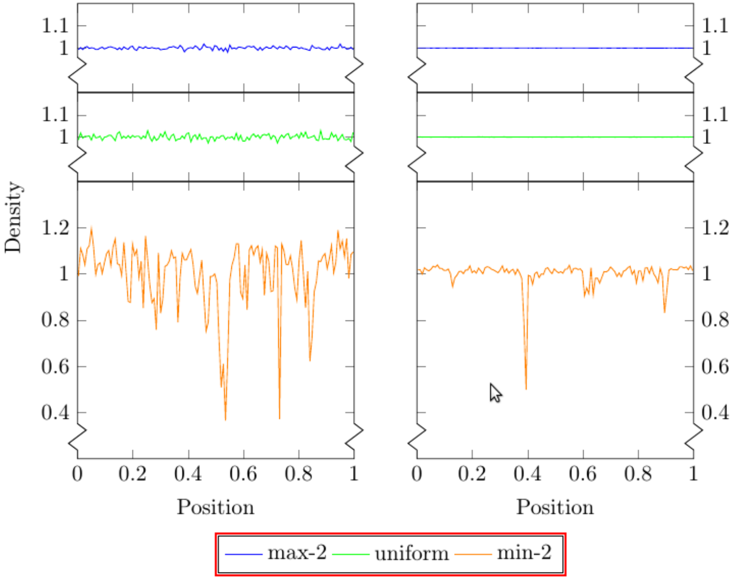
Phase plane analysis yields $G(t) \to c_k = ((2k-1)/k(k-1))^{1/(k-1)}$ as $t \to \infty$, whence

$$(F^{\Psi})'(x) \sim \frac{c_k}{(k-1)x^{1+\frac{1}{k-1}}}, \qquad \text{as } x \to \infty.$$

Properties of limiting distribution: min-k-process



Empirical distribution of points



Empirical distribution of points (2)

Recent progress by Matthew Junge (University of Washington): Considers evolution of $(A_t^\alpha)_{t\ge 0}$, which is defined as $(A_t)_{t\ge 0}$ but counting only the intervals in the interval $[0,\alpha]$. Formally, it satisfies

$$A_t^\alpha(x) = A_0^\alpha(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^\infty \frac{1}{z}\frac{dA_s^\alpha(z)}{dA_s(z)} d\Psi(A_s(z))\right] ds + M_t^\alpha(x),$$

and if Ψ is C^1 with derivative $\psi = \Psi'$, then

$$\frac{dA_s^\alpha(z)}{dA_s(z)} d\Psi(A_s(z)) = \psi(A_s(z)) dA_s^\alpha(z).$$

In the limit leads to the linear evolution equation

$$F_t^\alpha(x) = F_0^\alpha(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \left[\int_{e^{s-t}x}^\infty \frac{\psi(F^\Psi(z))}{z} dF_s^\alpha(z)\right] ds.$$

Junge (2015): Proof of equidistribution for restricted class of processes including max-2 process (but excluding min-2 and even max-k, $k \ge 3$). Key point: extension of contraction lemma to the above evolution equation.

Empirical distribution of points (3)

Theorem (M–Paquette, work in progress)

The Ψ-process is asymptotically equidistributed for all $\Psi \in C^1$ (maybe more...).

Empirical distribution of points (3)

Theorem (M–Paquette, work in progress)

The Ψ-process is asymptotically equidistributed for all $\Psi \in C^1$ (maybe more...).

Thank you for your attention!