Coagulation-transport PDE and nested coalescents

Emmanuel Schertzer.
Joint work with Amaury Lambert.

June 26, 2018

SCIENCES SORBONNE UNIVERSITÉ

LPSM LABORATOIRE DE PROBABILITÉS STATISTIQUE & MODÉLISATION

COLLÈGE DE FRANCE 1530

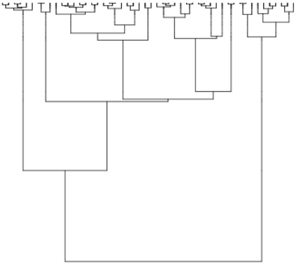
S.M.I.L.E Stochastic Models for the Inference of Life Evolution

Part 1

Coalescents and Nested Coalescents.

Kingman coalescent

- Markov process valued in $\mathcal{P}_{\mathbb{N}}$.
- Pairs of blocks coalesce at rate 1.



Kingman coalescent

- Coming down from ∞:

$$\forall \epsilon > 0, \quad K_\epsilon < \infty \quad \text{a.s.}$$

where K_t is the block counting process at time t ($K_0 = \infty$).

- Speed of coming down from ∞:

$$\left(\frac{1}{n} K_{t/n}; t > 0 \right) \implies (x_t; t > 0)$$

where $(x_t; t > 0)$ is solution of the ODE

$$\dot{x} = -\frac{1}{2} x^2, \quad \text{with} \quad x_0 = \infty$$

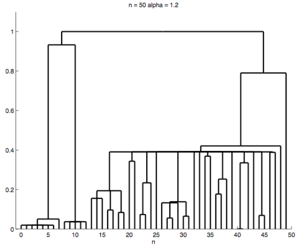
so that $x_t = \frac{2}{t}$.

Lambda coalescent (Pitman, Sagitov)

- Let Λ be a finite measure on $[0,1]$.
- If n blocks present at time t, each k-uplet of blocks ($k \geq 2$) coagulate at rate

$$\int_{[0,1]} x^{k-2}(1-x)^{n-k}\Lambda(dx).$$

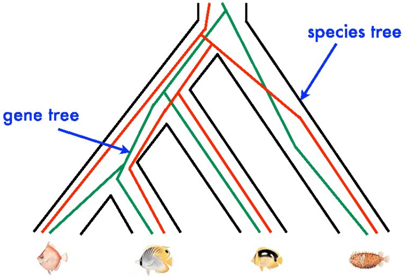
- $\Lambda = \delta_0$ is the Kingman coalescent.



- Define
$$\psi(x) \;=\; \int_{[0,1]} (e^{-rx} - 1 + rx) \frac{\Lambda(dr)}{r^2}$$
which is the Laplace exponent of a critical and spectrally positive Lévy process.
- Under Grey's condition
$$\int^{\infty} 1/\psi(u) su \;<\; \infty$$
the Λ-coalescent comes down from infinity.
-
$$\forall t > 0, \quad K_{t/n}/x_{t/n} \;\Longrightarrow\; 1$$
where $(x_t; t \geq 0)$ is solution of the ODE
$$\dot{x} = -\psi(x), \quad \text{with} \quad x_0 = \infty$$
(Beresticky's, Limic, Schweinsberg)

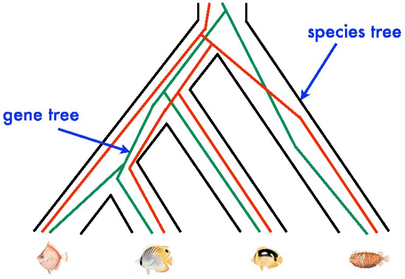
Nested coalescents

- Two levels of coalescence: species and gene trees
- Gene lineages can only coalesce if they belong to the same species
- For simplicity, we will consider the case where
  - (i) The species tree is a Kingman coalescent
  - (ii) The gene lineages coalesce *à la* Λ (taking into account the species constraint).



Nested coalescents

- Special case $\Lambda = \delta_0$: nested Kingman coalescent. Recently considered (with very different methods) by Benitez, Rogers, Schweinsberg, Siri Jégousse.
- General exchangeable nested coalescents considered by Benitez, Lambert, Duschamps, Siri Jégousse.



Coming down from ∞

- Let ρ_t be the number of gene lineages at time t.
- Let s_t be the number of species at time t.
- We say that the nested coalescent c.d.i. iff
  1. $s_0 = \infty$
  2. at least one gene lineage per species.
  3. For every $\epsilon > 0$,

$$\rho_\epsilon < \infty, \quad \text{whereas } \rho_0 = \infty.$$

Lemma

The nested coalescent c.d.i. iff the underlying Λ coalescent c.d.i.

Coming down from ∞

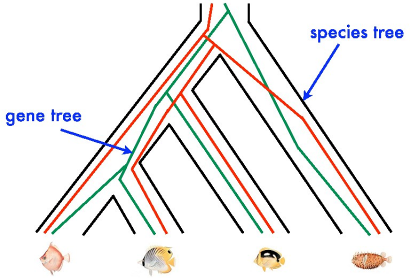
- Different ways of coming down from infinity.
- Question 1: Entrance law at ∞ ?
- Question 2 : Different speeds of coming down from ∞ ?
- This talk will address the second question for nested coalescents c.d.i.
  1. For simple coalescents: speed is determined by a simple ODE with ∞- initial condition.
  2. For nested coalescents: speed is determined by a degenerate PDE, raising non-trivial unicity and existence problems.

Part II

Convergence results (nested Kingman coalescent and more)

Empirical measures

- Sequence of nested Kingman indexed by n.
- For $i \leq s_t^n$, $\Pi_t^n(i) = \#$ gene lineages in species i.
- Π_t^n will be called the genetic composition vector.
- $g_t^n := \frac{1}{s_t^n} \sum_{i=1}^{s_t^n} \delta_{\Pi_t^n(i)}$ the empirical measure associated to the number of gene lineages.



R1: finite but large populations ($O(n)$ species, $O(n)$ genes/species)

- (Renormalization) $\tilde{g}^n_t := \frac{1}{s^n_{t/n}} \sum_{i=1}^{s^n_{t/n}} \delta_{\frac{1}{n}\Pi^n_{t/n}(i)}$

Theorem

Assume that that there exist $r \in (0,\infty)$ and $\nu \in M_p(\mathbb{R}^+)$ deterministic such that

1. $s^n_0/n \to r \in (0,\infty)$ in $L^{2+\epsilon}$
2. $\tilde{g}^n_0 \to \nu$ (weak cv)

Then $(\tilde{g}^n_t; t \geq 0)$ converges in $D([0,T],(M_F(\mathbb{R}^+),w))$ to $(d(t,x)dx; t \geq 0)$ where d is the weak solution of the PDE

$$\partial_t d(t,x) \;=\; \partial_x(\frac{1}{2}x^2 d)(t,x) + \frac{1}{t+\delta}(d \star d(t,x) \;-\; d(t,x)) \quad t,x \geq 0,$$

with initial condition $d(0,x)dx \;= \nu(dx)$ and $\delta = \frac{2}{r}$ (inverse population size).

R2: infinite populations

Theorem

1. Nested Kingman c.d.i., i.e., $s_0^n = \infty$.
2. Only constraint $\forall i \in \mathbb{N}$ $\Pi_0(i) \geq 1$.

Then $(\tilde{g}_t^n; t \geq 0)$ converges (in $D([\tau, T], (M_F(\mathbb{R}^+), w))$ for every $0 < \tau < T$) to $(d(t,x)dx; t \geq 0)$ where d is the unique proper weak solution of the PDE

$$\partial_t d(t,x) \ = \ \partial_x(\frac{1}{2}x^2 d)(t,x) + \frac{1}{t}\left(d \star d(t,x) \ - \ d(t,x)\right) \quad t,x \geq 0,$$

- Degenerate at $t = 0$.
- No prescription of the initial condition ! (only require the solution to be proper, i.e. $\lim_{t \to 0} d(t,x)dx \neq \delta_0$.)

Corollary

1. Nested Kingman c.d.i., i.e., $s_0^n = \infty$.
2. Only constraint $\forall i \in \mathbb{N}$ $\Pi_0(i) \geq 1$.

Let ρ_t be the number of gene lineages at time t. Then

$$\frac{1}{n^2}\rho_{t/n} \Longrightarrow \frac{2}{t}\int_0^\infty x d(t,x) dx =_{\text{scaling}} \frac{2}{t^2}\int_0^\infty x d(1,x) dx < \infty$$

where d is the proper solution of the previous slide.

Benitez, Rogers, Schweinsberg, Siri Jégousse (18) characterizes the same limit in terms of a fixed point problem.

General coalescent. PDE problem

- Recall that for Λ coalescent, the speed of c.d.i. is related to the ODE

$$\dot{x} = -\psi(x), \quad \text{(for Kingman, } \psi(x) = \tfrac{1}{2}x^2\text{)}$$

where

$$\psi(x) \; = \; \int_{[0,1]} (e^{-rx} - 1 + rx) \frac{\Lambda(dr)}{r^2}$$

- For general nested coalescents, the speed of c.d.i should be related to the PDE

$$\partial_t d(t,x) \; = \; \partial_x(\psi(x) d)(t,x) + \frac{1}{t+\delta} (d \star d(t,x) \; - \; d(t,x))$$

with $\delta > 0$ (finite population) and $\delta = 0$ (infinite population).

- For general nested coalescents, the speed of c.d.i should be related to the PDE

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) + \frac{1}{t+\delta}\,(d \star d(t,x) \;-\; d(t,x))$$

- General strategy (Méléard, Fournier, Tran ....)

Step 1 Tightness of $(\tilde{g}^n_t, t > 0)$, and show that any sub-sequential limit solves the PDE.

Step 2 Show that the PDE has a unique solution.

- Question: Existence and uniqueness of the (possibly) degenerate PDE when ψ is the Laplace exponent of a spectrally positive critical Lévy process.

Part III

PDE problem

Finite initial population

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) \;+\; \frac{1}{t+\delta}\left(d \star d(t,x) \;-\; d(t,x)\right)$$

Test functions: $f \in C_b^1(\mathbb{R}^+)$ and $f\psi'$ bounded.

Definition

Let $\delta > 0$ and ν a probability measure on $\mathbb{R}^+$. We say that a probability-valued process $(\mu_t; t \geq 0)$ is a weak solution with initial condition ν if for every test-function f and every $t \geq 0$:

$$\langle \mu_t, f \rangle = \langle \nu, f \rangle - \int_0^t \langle \mu_s, \psi f' \rangle \, ds + \int_0^t \left(\frac{1}{s+\delta} \langle \mu_s \star \mu_s, f \rangle - \langle \mu_s, f \rangle \right) ds$$

Idea: multiply both sides of the original PDE by f, then IPP to transfer the derivatives on the test function, and set $d(t,x)dx = \mu_t(dx)$

Infinite initial population

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) \;+\; \frac{1}{t}\left(d \star d(t,x) \;-\; d(t,x)\right)$$

Test functions: $f \in C_b^1(\mathbb{R}^+)$ and $f\psi'$ bounded.

Definition

Let $\delta = 0$. We say that a probability-valued process $(\mu_t; t > 0)$ is a weak solution if for every test-function f and every $s, t > 0$:

$$\langle \mu_t, f\rangle = \langle \mu_s, f\rangle - \int_s^t \langle \mu_u, \psi f'\rangle \, du + \int_s^t \frac{1}{u}(\langle \mu_u \star \mu_u, f\rangle - \langle \mu_u, f\rangle)\, du,$$

- We say that the solution is a dust solution iff $\mu_t \to_{t\downarrow 0} \delta_0$ (in the weak topology).
- We say that the solution is proper otherwise. (as in the previous convergence result)

Consider the PDE

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) \;+\; \frac{1}{t+\delta}\left(d \star d(t,x) \;-\; d(t,x)\right)$$

where ψ is the Laplace exponent of a critical spectrally ≥ 0 Lévy process.

Theorem

- $\exists!$ weak solution to the PDE problem with inverse pop. $\delta > 0$ and initial condition ν.
- $\exists!$ weak proper solution to the PDE problem when $\delta = 0$.
- $\exists$ infinitely many dust solutions to the PDE problem when $\delta = 0$.

(*Reminiscent of Boltzmann equation*)

Part IV

Construction of a solution. McKean–Vlasov approach

- Let $\delta > 0$ qnd $\nu \in M_P(\mathbb{R}^+)$ and initial condition, and

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) + \frac{1}{t+\delta}(d \star d(t,x) \;-\; d(t,x))$$

  where ψ is the Laplace exponent of a critical spectrally ≥ 0 Lévy process.
-

$$dx_t \;=\; -\psi(x_t)dt \;+\; \Delta J^{\delta} v_t, \quad \mathcal{L}(x_0) = \nu.$$

  where J^{δ} is a Poisson process with (inhomogeneous) rate $1/(t+\delta)$ and $(v_t)_{t\geq 0}$ is a family of independent rv's with $\mathcal{L}(v_t) = \mathcal{L}(x_t)$.

Lemma

If x_t is solution of the MK-V equation then $\mu_t \;:=\; \mathcal{L}(x_t)$ is solution of the corresponding coagulation-transport equation.

Proof: By Itô, for every test function

$$df(x_t) \;=\; -f'(x_t)\psi(x_t)dt \;+\; \Delta J^{\delta}(f(x_t+v_t) - f(x_t))$$

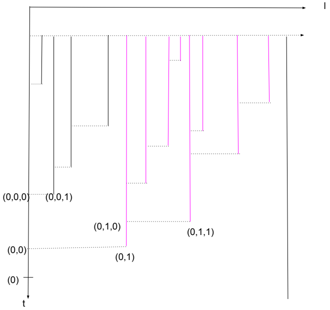
and the result follows by taking the expectation on both sides and integrating with respect to time.

Brownian Coalescent Point Process(Popovic)

- Poisson Point Process on $\mathbb{R}_*^+ \times \mathbb{R}_*^+$ with intensity measure

$$dl \times \frac{dt}{t^2}.$$

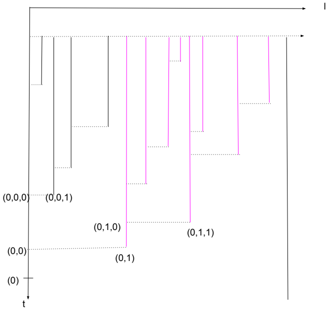
- Let $\mathcal{T}$ be the random ultrametric tree generated by the PPP. (infinite branch at $\{l = 0\}$, leaves at $\{t = 0\}$)



Marking of the tree $\mathcal{T}$ above time horizon δ with initial measure ν.

- Mark each point at level δ with i.i.d. random variables with law ν
- Marks evolve according to the ODE $\dot{x} = -\psi(x)$ along each branch
- When two branches merge, add up the marks.

Define $(m_0(t); t \geq \delta)$ be the marking on the branch $\{l = 0\}$.



Theorem

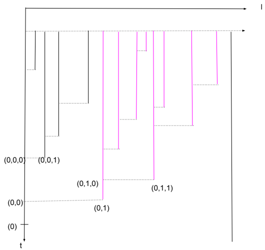
$(\theta_\delta \circ m_0(t); t \geq 0)$ is solution of MK-V

$$dx_t \;=\; -\psi(x_t)dt \;+\; \Delta J^\delta v_t, \quad \mathcal{L}(x_0) = \nu.$$

Proof.

By a simple time translation, enough to show that $(m_0(t); t \geq \delta)$ solves

$$\forall t \geq \delta, \quad dx_t \;=\; -\psi(x_t)dt \;+\; \Delta J^0 v_t, \quad \mathcal{L}(x_\delta) = \nu.$$



Finite population solution

- $\mu_t := \mathcal{L}(x_t)$ is solution of the PDE with inverse pop. δ and initial measure ν.
- $$\mu_t = \mathcal{L}(F(\mathbb{T}^{(t+\delta,\delta)}, (W_i)))$$

  where W_i are i.i.d. rv's with law ν, and $\mathbb{T}^{(t+\delta,\delta)}$ is the tree originated from $(0, t+\delta)$ up to level δ.
- By a martingale arguments, we show that this solution is the unique solution of the PDE. (*Using another probabilistic approach!*)

Infinite population $\delta = 0$

- Assume now that $\delta = 0$.
- For every $s > 0$, on $[s, \infty)$, $(\mu_t; t \geq s)$ is the weak solution of the PDE

$$\partial_t d(t,x) \;=\; \partial_x(\psi(x)d)(t,x) \;+\; \frac{1}{t}\left(d \star d(t,x) \;-\; d(t,x)\right)$$

  with initial condition $d(s,x)dx = \mu_s(dx)$.
- By the previous construction, $(\mu_t; t \geq s)$ can be obtained by marking $\mathcal{T}$ at level s with initial marking μ_s, i.e.,

$$\mu_t = \; \mathcal{L}(F(\mathbb{T}^{(t,s)}, (W_s^i)))$$

  where W_s^i are i.i.d. rv's with law μ_s, and $\mathbb{T}^{(t,s)}$ is the tree originated from $(0,t)$ up to level s.

Infinite population $\delta = 0$

Letting $s \to 0$ and using the fact that the solution is proper, we obtain the following result.

Theorem

Under Grey's condition, there is a unique proper solution. Further

$$\mu_t = \mathcal{L}\left(M^{\mathbb{T}^{(t,0)}}(\mathcal{M}^c)\right)$$

where $M^{\mathbb{T}^{(t,0)}}$ is the 0-entrance measure of a branching CSBP with branching mechanism ψ and genealogy $\mathbb{T}^{(t,0)}$.

Part V

A desintegration formula

Some facts about CSBP

- Let ψ be the Laplace exponent of a critical and spectrally positive Lévy process.
- There exists a Markov process Z_t such that

$$E_x\left(\exp(-\lambda Z_t)\right) = \exp(-x u_t)$$

  where u solves the ODE

$$\dot{u} = -\psi(u), \quad u(0) = \lambda.$$

- Z_t is the CSBP with branching mechanism ψ
- Branching property : $P_{x+y} = P_x \star P_y$.

Entrance law of a CSBP

- Let Z_t be a CSBP with branching mechanism ψ.
- Branching property : $P_{x+y} \ = \ P_x \star P_y$.
- There exists a σ-finite measure N (on the space of càdlàg paths starting from 0) such that under P_x

$$Z_t \ = \ \sum_i z_t^i$$

where the sum is taken over the atoms of a Poisson PP with intensity xN

- N is called the 0-entrance law of the process. Further

$$u_T \ = \ N(1 - \exp(-\lambda Z_T))$$

where u solves the ODE

$$\dot{u} = -\psi(u), \quad u(0) = \lambda.$$

Branching CSBP

- Let $\mathbb{t}$ be an ultrametric tree with depth T.
- Define $\mathcal{Z}^{\mathbb{t}}$ be a branching CSBP with branching mechanism ψ.
  1. Each particle is an indepedent CSBP with branching mechanism ψ.
  2. Particles replicate upon branching.
- Let λ be a vector of dimension the number of leaves. Then

$$E_x(\exp(-\left\langle \lambda, \mathcal{Z}^{\mathbb{t}}_T \right\rangle)) \ = \ F(\mathbb{t}, \lambda).$$

  where F is obtained by marking the leaves with the λ_i's and then by propagating the marks up to the root (as in the CPP).
- Define $M^{\mathbb{t}}$ the 0-entrance law of the branching CSBP.

$$F(\mathbb{t}, \lambda) \ = \ M^{\mathbb{t}} \left(1 - \exp(-\left\langle \mathcal{Z}^{\mathbb{t}}_T, \lambda \right\rangle \right)$$

- When $\delta = 0$,

$$\mu_t = \mathcal{L}(F(\mathbb{T}^{(t,s)}, W_s)) = \mathcal{L}\left(M^{\mathbb{T}^{(t,s)}}(1 - \exp(-\langle \mathcal{Z}_t, W_s \rangle))\right)$$

  with $\mathcal{L}(W_s^i) = \mu_s$.
- $\mathbb{T}^{(t,0)}$ infinitely many leaves.
- If the solution is proper (μ_0 non-degenerate), there exists $a > 0$ such that infinitely many marks are $> a$.
- $\mathcal{M}$ = event of mass extinction.
- Under $\mathcal{M}$: $1 - \exp(-\langle \mathcal{Z}_t, W_s \rangle)$,
- Under $\mathcal{M}^c$: $1 - \exp(-\langle \mathcal{Z}_t, W_s \rangle) \to 1$ as $s \to 0$.
- Thus

$$\mu_t = \mathcal{L}\left(M^{\mathbb{T}^{(t,0)}}(\mathcal{M}^c)\right) < \delta_\infty$$

  under Grey's condition.

Conclusion

- We exposed a relation between the nested coalescent and a degenerate coagulation-transport equation.
- Construction of a finite pop. solution ($\delta > 0$) using the Brownian CPP.
- Using a desintegration formula, under Grey's condition, we proved that there is a unique ∞. pop, solution, and that

$$\mu_t = \mathcal{L}\left(M^{\mathbb{T}^{(t,0)}}(\mathcal{M}^c)\right), \quad \mathcal{M} = \text{mass extinction}$$

Some open problems

- Convergence of nested Λ coalescents (not only nested Kingman). Tightness results needed.
- Uniqueness of the entrance law at ∞.
- What if ψ is not the Laplace exponent of Lévy process ? e.g., $\psi(x) = x^\gamma$ with $\gamma > 2$? Stable case OK using the McKean-Vlasov approach (but with no disintegration formula available which induces non-trivial complications).
- Convergence to dust solutions.

Thank you !