Genealogies of Interacting Particle Systems

IMS, NUS, Singapore, July/August 2017

Learning session on Lookdown Constructions

moderated by Amandine Véber and Anton Wakolbinger

Part 1 (A.W.)

Some features of Donnelly&Kurtz' "Particle Representations for Measure Valued Population Models"

with an outlook towards
Continuum Tree-Valued Processes with Interaction

1. A lookdown construction of tree-valued processes with finite population sizes

(in the footsteps of [DK], Donnelly&Kurtz (1999),
Particle representations for measure-valued poplation models,
Sec. 2)

Although [DK] do not speak of tree-valued processes, these are somehow between their lines.

Thus, tying in with the theme of the program,

and the morning session's topic “Tree valued processes”,

let the state of the process at time t be given by a

matrix of genealogical distances $r = r(\alpha, \beta), \alpha, \beta = 1, 2, \dots, N_t$

and a vector of types $g = g(\alpha), \alpha = 1, 2, \dots, N_t$

$y = (r, g)$ is then called a *marked distance matrix*

(à la Depperschmidt, Greven, Pfaffelhuber (2011))

Usually, it is assumed that the dynamics of y is

symmetric with respect to the “naming” of individuals:

the reproductive success (or death) of an individual may depend on its type and on the entire type configuration but should be invariant under a joint permutation of the individuals’ names and types.

For finite population sizes N_t (possibly varying in time) one may e.g. (as proposed in [DK]) think of a *dynamical naming* of individuals:

- at any time, the set of the individuals' names is $[N_t] := \{1, \dots, N_t\}$

- if the individual named ξ dies, then the name gap is closed: the previous names $\{\xi + 1, \dots, N_{t-}\}$ are shifted down by 1

- if an individual is born, then it gets the name $N_{t-} + 1$

With the aim to obtain more “projectivity” when (later)
considering a sequence of systems
with initial population size $N_0^K \to \infty$,
[DK] suggest an alternative procedure:

order the individuals according to their longevity

by placing them on levels $1, \dots, N_t$, where
the individuals at the lower levels are those that live longer.

In particular, given $N_{t-} = n$ and $N_t = n - 1$, it is always the individual at level N_{t-} (at the currently highest level) that dies.

and

a daughter is always be placed at some higher level than that of its mother.

Specifically, cosider births and deaths of single individuals at times prescribed by given counting paths N^b and N^d.

At a death time t, remove the individual at level N_{t-}

(the highest of all the currently occupied levels)

At a birth time t,

- choose a pair (i, j) with $1 \leq i < j \leq N_{t-} + 1$ at random,
- declare the individual at level i to be the mother,
- place the daughter at level j and
- shift the indiv'ls from levels $j, j + 1, \ldots, N_{t-}$ one level up.

In this way, for a pair $i < j$ of mother-daughter levels,
level j “looks down” to level i
and copies its type.

But alas - this causes a mother-daughter bias!

A first aid:

Place the initial individuals exchangeably on the levels.

Example: Initially two individuals, one red, one blue.

Assume the first event is a birth event.

Given $(X_0(1), X_0(2)) = (r, b)$, after birth the probability of (r, r, b), (r, b, r) and (r, b, b) is 1/3 each.

Given $(X_0(1), X_0(2)) = (b, r)$, after birth the probability of (b, b, r), (b, r, b) and (b, r, r) is 1/3 each.

The overall result is thus a random drawing without replacement from an **urn with the random content** “two red, one blue” or “two blue, one red”, each with probability 1/2.

More generally:

Let, for a given birth (or death) time,

$X_{t-} = ((r(\Pi(i), \Pi(j))_{i,j}, (g(\Pi(i))_i)$

be the state at time $t-$,

with Π a uniformly distributed permutation

(so that X_{t-} is exchangeable).

Then also X_t is exchangeable.

But mind:
Even with an exchangeable placement of the initial individuals, there remains a bias in the lookdown genealogy!

Example:
Consider $N_0 = 1$, with the initial individual placed on level 1. Assume the first two jumps of the population size are due to births.
Then the probability that the second birth comes from the initial individual is $2/3$, whereas the probability that the second birth comes from the individual born in the first event is $1/3$.

An ultimate remedy to cure the bias in the LD genealogy:

Consider an appropriate “dictionary” Θ_t
(a random permutation on $[N_t] = \{1, \dots, N_t\}$)
translating the individuals’ names into the individuals’ levels.

This dictionary is started
with Θ_0 a uniformly distributed permutation on $[N_0]$
and is updated at each birth or death time prescribed by (N^b, N^d):

Let t be a jump time of N^b or of N^d, and put

$\theta := \Theta_{t-}$, $\theta' := \Theta_t$.

We describe how θ' is obtained from θ.

If t is a death time, let

$\xi := \theta^{-1}(N_{t-})$ be the name of the individual dying at time t

and put

$$\theta'(\alpha) := \begin{cases} \theta(\alpha), & \alpha < \xi \\ \theta(\alpha + 1), & \alpha \geq \xi. \end{cases}$$

If t is a birth time,

and the arrow is shot from level i to level $j > i$,

then let σ be a permutation on (i, j) which

with probability $1/2$ is the identity and

with probability $1/2$ is the transposition of i and j.

Let $\xi := \theta^{-1}(i)$ be the mother's name previous to birth,

and put

$$\theta'(\alpha) := \begin{cases} \sigma(i), & \alpha = \xi \\ \sigma(j), & \alpha = n+1, \\ \theta(\alpha), & \alpha \notin \{\xi, n+1\},\ \theta(\alpha) < j, \\ \theta(\alpha)+1, & \alpha \notin \{\xi, n+1\},\ \theta(\alpha) \geq j. \end{cases}$$

In the simple example with one initial individual and two consecutive birth times, we would just flip the mother’s and daughter’s levels (1 and 2) with probability 1/2.

[DK] Lemma (2.1) (Coupling Lemma)

For given N_0 and $(N^b, N^d) = ((N_t^b, N_t^d))_{t \geq 0}$,
let Θ_0 be a uniformly distributed permutation of $[N_0]$
and do the updates of Θ_t at the birth/death times $t_1, \dots, t_m, \dots$,
with $\xi_m :=$ the mother's resp the dyer's name at time t_m-
and σ_m the (probability 1/2) swap (if t_m is a birth time)
of the mother's and daughter's levels.

Then, for any m, the $\xi_1, \dots, \xi_m$ are independent,
and ξ_m is uniformly distributed on $[N_{t_m-}]$.
Moreover, (the updated dictionary) Θ_{t_m} is independent of $\xi_1, \dots, \xi_m$.

Now let X_0 be exchangeable, and independent of Θ_0.

Define $Y_t := X_t(\Theta(t))$.

Since (always given (N^b, N^d))

Y_t is measurable w.r.to $Y_0, \xi_1, \ldots, \xi_m$,

and Θ_t is independent of these, we have that

Θ_t is independent of Y_t.

And since Θ_t was a uniformly distributed permutation,

then $X_t = Y_t(\Theta_t^{-1})$ is exchangeable,

even though Y_t in general is not exchangeable.

Note that the dynamics of Y (even conditional on (N_0, N^b, N^d)) is symmetric.

Indeed, by the Coupling Lemma,
the dying individual's name is chosen uniformly,
the mother's name is chosen uniformly, and
given that α, β are the levels of a mother-daughter pair,
they are swapped with probability 1/2.

We have thus arrived at

[DK] Theorem 1.1 (upgraded to include distance matrices):

Let X_0 be exchangeable and
$(\Theta_t)_{t\geq 0}$ be defined as above.
Then X_t is exchangeable, and
$Y_t := X_t(\Theta_t)$ has a symmetric dynamics.

Thus, together with some additional bits of randomness
(the random transpositions σ_m)
we have obtained a lookdown representation of a process of marked distance matrices whose distribution, after appropriate permutations Θ_t, is the same as that of the process with the symmetric dynamics.

In particular, the process of isomorphy classes of the marked tree-valued process obtained by the lookdown representation is the same as the one obtained from the symmetric dynamics.

Good news: This statement remains true even for a type-dependent reproduction dynamics, e.g. for genic fecundity selection and competitive viability selection.
(More on this might be discussed in the second block.)

One necessary twist: In the proof of the Coupling Lemma, the uniformity of the distributions of ξ_m and of Θ_{t_m} rely on the exchangeability of the type vector at time t_m-. Therefore, the exchangeability of X_{t_m} should (and can!) be taken into the induction statement.

2. From global birth and death counting to the lookdown graph

Given the counting processes N and N^b we want to fill in the lookdown arrows (according to a certain probability kernel):
Given $N_t = n$, $\Delta N_t = 1$, the probability that level 2
looks down to level 1 at time t is

$$\frac{2}{n(n-1)}\,.$$

Thus, if only one individual is born at a time, given N and N^b, the expected number of lookdowns from 2 to 1 up to time t is

$$\sum_{r\le t} 2\Delta N^b_r \,\frac{1}{N_r(N_r-1)}$$

More generally, if there are k births involved at time t, then given $N_t = n$, $\Delta N_t = 1$, the probability that level 2 looks down to level 1 at time t is

$$\frac{\binom{n-2}{k-1}}{\binom{n}{k+1}} = \frac{k(k+1)}{n(n-1)}.$$

Thus, given N and N^b, the expected number of lookdowns from level 2 to level 1 up to time t is

$$\sum_{r \leq t} ((\Delta N_r^b)^2 + \Delta N_r^b) \frac{1}{N_r(N_r - 1)}$$

In the diffusive scaling of branching populations,
one typically has an individual branching rate $K + O(1)$,
and a convergence $N^K/K =: P^K \Rightarrow P$,
provided N_0^K/K converges.
Thus, in the K-th scaling, with

$$U_t^K := \frac{[N^{K,b}]_t + N_t^{K,b}}{K^2},$$

given N^K and $N^{K,b}$ the expected number of lookdowns from

2 to 1 up to time $t < \tau^K$ (the extinction time of N^K) is

$$\int_0^t \frac{1}{P_r^K(P_r^K - 1/K)} dU_r^K =: H_t^K.$$

3. Main limit theorem in [DK]

So far we have discussed the case of finite population sizes (and included the matrix of genealogical distances into the state of the system)

In [DK], X^K, X are processes of type configurations, and do not include distance matrices.

[DK] Theorem 3.2

Assume $(P^K, U^K, H^K, \tau^K) \Rightarrow (P, U, H, \tau)$

and

$X_0^K\big|_{\{1,\dots,N_0\}} = X_0\big|_{\{1,\dots,N_0\}}$

for an infinite exchangeable sequence X_0.

Then $(P^K, U^K, P^K Z^K, X^K) \Rightarrow (P, U, PZ, X)$

where Z^K and Z denote the empirical distributions

of X^K and X, respectively.

The strategy of proof relies

- for fixed t (or finitely many t) on the de Finetti Theorem

- concerning uniformity in t, on an exponential convergence rate in the de Finetti Theorem that is uniform for uniformly bounded exchangeable sequences (see [DK] Lemma A2)

Variants of this theorem have been obtained by

- Birkner et al (2009) for Ξ-Fleming-Viot processes,
- Gufler (2016) for Ξ-Fleming-Viot processes including distance matrices and sampling measures,
- Véber and W. (2013) for spatial Λ-Fleming-Viot processes,

all of these dealing with constant (or locally constant) population sizes.

In cases of - possibly interacting - branching particle systems one has (see DK Sec 3.2)

$$U = [P]$$

E.g. for unit variance branching, even with a population dependent (non-explosive) drift:

$$dU_t = d[P]_t = P_t \, dt,$$

$$dH_t = \frac{1}{P_t^2} \, d[P]_t = \frac{1}{P_t} \, dt, \quad t < \tau.$$

4. Outlook.

Work in progress
(joint with A. Blancas, S. Gufler, S. Kliem and V.C. Tran):

For a (class of) type dependent reproduction dynamics construct a candidate for the limiting process (P, X) where $X = (R, G)$ takes values in the space of marked distance matrices, now indexed by $\mathbb{N}$.

Show $(P^K, X^K) \Rightarrow (P, X)$

Then, for any t, we can obtain from the state R_t, G_t a
marked metric measure space:
Take the completion L_t of $(\mathbb{N}, R_t)$.
Due to exchangeablity, there exists a.s. the weak limit μ_t
of the measures $\frac{1}{n} \sum_{i=1}^{n} \delta_{(i,G_t(i))}$
Denoting the space of types by E, we endow $L_t \times E$ with the measure μ_t and thus obtain a marked metric measure space (L_t, R_t, μ_t)

The isomorphy class of the latter is the state
of the desired process.

Passing to isomorphy classes removes the lookdown bias, rendering a symmetric (albeit type dependent) dynamics also in the infinite population limit.

In the neutral case (including type-independent competition), we are back to time-changed “evolving Kingman coalescents” (cf. S. Evans (2000), Kingman’s coalescent as a random metric space)