Some mathematical models from population genetics
II. Adding space

Alison Etheridge
University of Oxford

RIMS, September 2023

What we have so far

In time units of N_e generations,

- (Forwards time) The Wright-Fisher diffusion

$$dp_t = \sqrt{p_t(1 - p_t)}dW_t;$$

- (Backwards time) The Kingman coalescent

$$n_t \mapsto n_t - 1 \text{ at rate } \binom{n_t}{2};$$

- Sampling probabilities

$$\mathbb{E}[p(t)^{n(0)}] = \mathbb{E}[p(0)^{n(t)}].$$

Stronger result holds. Kingman coalescent really describes genealogy of random sample from (neutral) population.

Adding spatial structure: subdivided populations

Population subdivided into demes = islands = colonies

- Vertices of graph, $i \in I$;
- $i \sim j$ if i, j neighbours
- $N_i =$ population size in deme i

Structured Wright-Fisher model
Reproduction in discrete generations

- neutral Wright-Fisher within each deme
- proportion m_{ij} of individuals in deme i migrate to deme j

$$N_i \sum_{j \sim i} m_{ij} = \sum_{j \sim i} N_j m_{ji}$$

Genealogy of structured Wright-Fisher model

1. Two lineages sampled from deme i

$$\mathbb{P}\big[\text{coalesce in } j \neq i \text{ in previous generation}\big] = \frac{\binom{m_{ji}N_j}{2}}{\binom{N_i}{2}} \frac{1}{N_j}$$

$$\mathbb{P}\big[\text{coalesce in } i \text{ in previous generation}\big] = \frac{\binom{N_i - \sum_{j \sim i} m_{ji}N_j}{2}}{\binom{N_i}{2}} \frac{1}{N_i}$$

Genealogy of structured Wright-Fisher model

1. Two lineages sampled from deme i

$$\mathbb{P}\big[\text{coalesce in } j \neq i \text{ in previous generation}\big] = \frac{\binom{m_{ji}N_j}{2}}{\binom{N_i}{2}} \frac{1}{N_j}$$

$$\mathbb{P}\big[\text{coalesce in } i \text{ in previous generation}\big] = \frac{\binom{N_i - \sum_{j \sim i} m_{ji}N_j}{2}}{\binom{N_i}{2}} \frac{1}{N_i}$$

2. Two lineages sampled from demes $i \neq j$

$$\mathbb{P}\big[\text{coalesce in } k \notin \{i,j\} \text{ in previous generation}\big] = \frac{m_{ki}N_k}{N_i} \frac{m_{kj}N_k}{N_j} \frac{1}{N_k}$$

$$\mathbb{P}\big[\text{coalesce in } j \text{ in previous generation}\big] = \frac{m_{ji}N_j}{N_i} \frac{\left(N_j - \sum_{l \sim j} m_{lj}N_l\right)}{N_j} \frac{1}{N_j}$$

Scaling limit: the structured coalescent

- $N_i = O(N)$
- $m_{ij} = O(\frac{1}{N})$
- time unit $= N$ generations

Scaling limit: the structured coalescent

▶ $N_i = O(N)$ ▶ $m_{ij} = O(\frac{1}{N})$ ▶ time unit $= N$ generations

$$\mathbb{P}\big[\text{simultaneous migration and coalescence}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{simultaneous or multiple mergers}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{single lineage at } i \text{ migrates}\big] = \sum_{j \sim i} \frac{m_{ji} N_j}{N_i} = O(1/N)$$

Scaling limit: the structured coalescent

▶ $N_i = O(N)$ ▶ $m_{ij} = O(\frac{1}{N})$ ▶ time unit $= N$ generations

$$\mathbb{P}\big[\text{simultaneous migration and coalescence}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{simultaneous or multiple mergers}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{single lineage at } i \text{ migrates}\big] = \sum_{j \sim i} \frac{m_{ji} N_j}{N_i} = O(1/N)$$

The structured coalescent $\underline{n} = (n_i)_{i \in I}$:

▶ $\begin{cases} n_i \mapsto n_i - 1 \\ n_j \mapsto n_j + 1 \end{cases}$ at rate $n_i \frac{N_e(j)}{N_e(i)} m_{ji}$

▶ $n_i \mapsto n_i - 1$ at rate $\frac{1}{2N_e(i)} n_i (n_i - 1)$

Scaling limit: the structured coalescent

- $N_i = O(N)$ - $m_{ij} = O(\frac{1}{N})$ - time unit $= N$ generations

$$\mathbb{P}\big[\text{simultaneous migration and coalescence}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{simultaneous or multiple mergers}\big] = O(1/N^2)$$

$$\mathbb{P}\big[\text{single lineage at } i \text{ migrates}\big] = \sum_{j \sim i} \frac{m_{ji} N_j}{N_i} = O(1/N)$$

The structured coalescent $\underline{n} = (n_i)_{i \in I}$:

- $\begin{cases} n_i \mapsto n_i - 1 \\ n_j \mapsto n_j + 1 \end{cases}$ at rate $n_i \frac{N_e(j)}{N_e(i)} m_{ji}$

- $n_i \mapsto n_i - 1$ at rate $\frac{1}{2N_e(i)} n_i (n_i - 1)$

Ancestral lineages drawn into more populous demes

Forwards in time?

Alleles a, A.　　$p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

Alleles a, A. $p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

$$
\begin{aligned}
\mathbb{E}\big[\Delta p_i\big] &= \frac{1}{N_i}\Big(\big(1 - \sum_{j \sim i} m_{ij}\big) N_i p_i + \sum_{j \sim i} m_{ji} N_j p_j\Big) - p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} p_j - \frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i
\end{aligned}
$$

Alleles a, A. $\qquad p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

$$
\begin{aligned}
\mathbb{E}\big[\Delta p_i\big] &= \frac{1}{N_i}\Big(\big(1 - \sum_{j \sim i} m_{ij}\big) N_i p_i + \sum_{j \sim i} m_{ji} N_j p_j\Big) - p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} p_j - \frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i
\end{aligned}
$$

$$
\frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i = \frac{1}{N_i} \sum_{j \sim i} N_j m_{ji} p_i
$$

Alleles a, A. $p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

$$
\begin{aligned}
\mathbb{E}\big[\Delta p_i\big] &= \frac{1}{N_i}\Big(\big(1 - \sum_{j \sim i} m_{ij}\big) N_i p_i + \sum_{j \sim i} m_{ji} N_j p_j\Big) - p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} p_j - \frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} \big(p_j - p_i\big)
\end{aligned}
$$

$$
\frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i = \frac{1}{N_i} \sum_{j \sim i} N_j m_{ji} p_i
$$

Forwards in time?

Alleles a, A. $p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

$$
\begin{aligned}
\mathbb{E}\big[\Delta p_i\big] &= \frac{1}{N_i}\Big((1 - \sum_{j \sim i} m_{ij}) N_i p_i + \sum_{j \sim i} m_{ji} N_j p_j \Big) - p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} p_j - \frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i \\
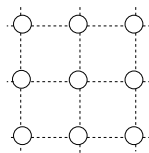&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} (p_j - p_i)
\end{aligned}
$$

$$
\mathbb{E}\big[(\Delta p_i)^2\big] = \frac{1}{N_i}\big(p_i(1-p_i) + O(1/N) \big) \qquad \mathrm{Cov}\big(\Delta p_i, \Delta p_j\big) = O(1/N^2)
$$

Alleles a, A. $\qquad p_i(t) =$ proportion of type a in deme i at time t

Δp_i change across single generation

$$
\begin{aligned}
\mathbb{E}\big[\Delta p_i\big] &= \frac{1}{N_i}\Big(\big(1 - \sum_{j \sim i} m_{ij}\big) N_i p_i + \sum_{j \sim i} m_{ji} N_j p_j\Big) - p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} p_j - \frac{1}{N_i} \sum_{j \sim i} m_{ij} N_i p_i \\
&= \sum_{j \sim i} m_{ji} \frac{N_j}{N_i} \big(p_j - p_i\big)
\end{aligned}
$$

$\mathbb{E}\big[(\Delta p_i)^2\big] = \frac{1}{N_i}\big(p_i(1 - p_i) + O(1/N)\big) \qquad \text{Cov}\big(\Delta p_i, \Delta p_j\big) = O(1/N^2)$

As $N \to \infty$ recover a system of diffusions coupled through migration

Kimura's stepping stone model

$$dp_i = \sum_j \frac{N_e(j)}{N_e(i)} m_{ji}(p_j - p_i)dt + \sqrt{\frac{1}{N_e(i)} p_i(1 - p_i)} dW_i$$

$\{W_i\}_{i \in I}$ independent Brownian motions
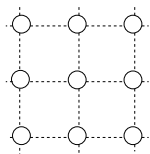
System of W-F diffusions coupled through migration

Kimura's stepping stone model

$$\sum_j N_e(i) m_{ij} = \sum_j N_e(j) m_{ji}$$

$$dp_i = \sum_j \frac{N_e(j)}{N_e(i)} m_{ji}(p_j - p_i) dt + \sqrt{\frac{1}{N_e(i)} p_i(1 - p_i)} dW_i$$
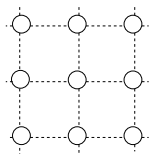
$\{W_i\}_{i \in I}$ independent Brownian motions

System of W-F diffusions coupled through migration

The structured coalescent $\underline{n}$:

- $\begin{cases} n_i \mapsto n_i - 1 \\ n_j \mapsto n_j + 1 \end{cases}$ at rate $n_i \frac{N_e(j)}{N_e(i)} m_{ji}$

- $n_i \mapsto n_i - 1$ at rate $\frac{1}{2N_e(i)} n_i (n_i - 1)$

Kimura's stepping stone model

$$dp_i = \sum_j \frac{N_e(j)}{N_e(i)} m_{ji}(p_j - p_i)dt + \sqrt{\frac{1}{N_e(i)} p_i(1 - p_i)} dW_i$$

$\{W_i\}_{i \in I}$ independent Brownian motions

System of W-F diffusions coupled through migration

The structured coalescent $\underline{n}$:

▶ $\begin{cases} n_i \mapsto n_i - 1 \\ n_j \mapsto n_j + 1 \end{cases}$ at rate $n_i \frac{N_e(j)}{N_e(i)} m_{ji}$

▶ $n_i \mapsto n_i - 1$ at rate $\frac{1}{2N_e(i)} n_i (n_i - 1)$

$\underline{p}^{\underline{n}} := \prod_{i \in I} p_i^{n_i}.$

$\mathbb{E}\left[\underline{p}_t^{\underline{n}_0}\right] = \mathbb{E}\left[\underline{p}_0^{\underline{n}_t}\right].$

Interpretation

$$\mathbb{E}\left[\underline{p}_t^{n_0}\right] = \mathbb{E}\left[\underline{p}_0^{n_t}\right].$$

- Sample $n_i(0)$ individuals from deme i at time t, $\sum_i n_i(0) < \infty$,
- Probability all type a is $\mathbb{E}\left[\underline{p}_0^{n_t}\right]$

Interpretation

$$\mathbb{E}\left[\underline{p}_t^{n_0}\right] = \mathbb{E}\left[\underline{p}_0^{n_t}\right].$$

- Sample $n_i(0)$ individuals from deme i at time t, $\sum_i n_i(0) < \infty$,
- Probability all type a is $\mathbb{E}\left[\underline{p}_0^{n_t}\right]$

Example Suppose $I = \mathbb{Z}^2$ *for simplicity $N_i \equiv N_e$*
For any finite sample, eventually $\underline{n}_t$ is a singleton, so all individuals in the sample are of the same type.

Interpretation

$$\mathbb{E}\left[\underline{p}_t^{n_0}\right] = \mathbb{E}\left[\underline{p}_0^{n_t}\right].$$

▶ Sample $n_i(0)$ individuals from deme i at time t,
  $\sum_i n_i(0) < \infty$,
▶ Probability all type a is $\mathbb{E}\left[\underline{p}_0^{n_t}\right]$

Example Suppose $I = \mathbb{Z}^2$ *for simplicity $N_i \equiv N_e$*
For any finite sample, eventually $\underline{n}_t$ is a singleton, so all individuals
in the sample are of the same type.

Need to account for mutation in our model

Adding mutation

Simplest example:

- Infinitely many alleles model of mutation: each individual in each generation, independently, with small probability μ mutates to a type never before seen in the population
- *Probability of identity by descent* of two individuals, F, = probability no mutation since time T of most recent common ancestor (MRCA)
- Equivalently $F = (1 - 2\mu)^T \approx \exp(-2\mu T)$ is the Laplace transform of the distribution of the time to the MRCA.

The neutral mutation rate dictates the timescales over which we can reconstruct information about genealogies.

Isolation by distance

In a population in which individuals typically migrate to geographically close subpopulations, and new mutations continuously accumulate, $\mathbb{P}[\text{two individuals in same allelic state}]$ declines with increasing separation.

Isolation by distance (Wright 1943)

Isolation by distance

In a population in which individuals typically migrate to geographically close subpopulations, and new mutations continuously accumulate, $\mathbb{P}$[two individuals in same allelic state] declines with increasing separation.

<div align="right">Isolation by distance (Wright 1943)</div>

In $\mathbb{Z}$ with nearest neighbour migration there is an explicit expression for the probability of identity under the stepping stone model. It declines exponentially with distance. But the exact formula is very special.

Isolation by distance

In a population in which individuals typically migrate to geographically close subpopulations, and new mutations continuously accumulate, $\mathbb{P}[\text{two individuals in same allelic state}]$ declines with increasing separation.

Isolation by distance (Wright 1943)

In $\mathbb{Z}^2$, approximate dispersal by Gaussian with variance σ^2,

$$(*) \qquad F(x) = \mathbb{E}_x[e^{-2\mu T}] \approx \frac{K_0(|x|/l_\mu)}{\mathcal{N} + \log(l_\mu/\kappa)} \qquad |x| > \kappa$$

K_0 modified Bessel function of second kind of degree zero, $l_\mu = \sigma/2\mu$, $\mathcal{N} = 2N_e\pi\sigma^2$ is *Wright's neighbourhood size*, κ is a local scale.

$(*)$ is known as the Wright-Malécot formula.

Wright-Malécot approximation for the stepping stone model

$F = \mathbb{P}[\text{identity}]$
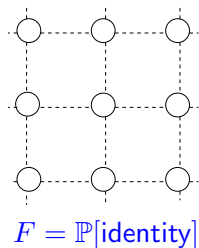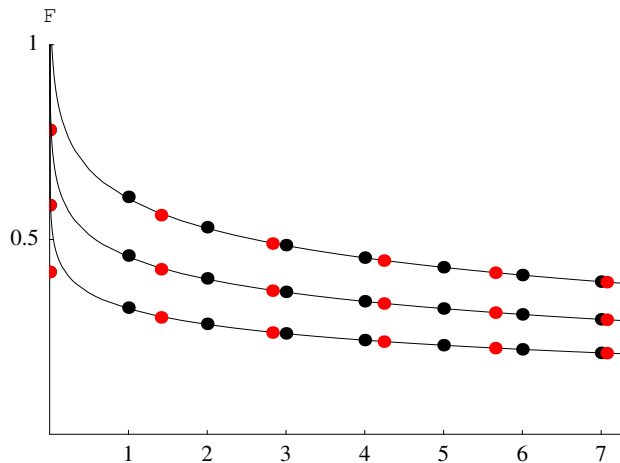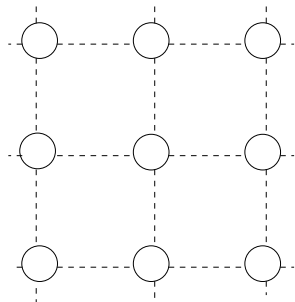
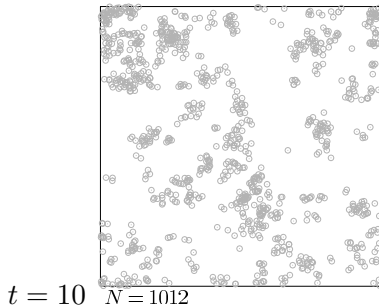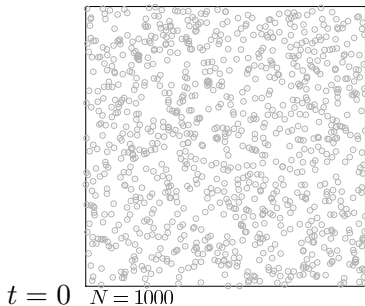# Modelling a spatial continuum: the Wright-Malécot model
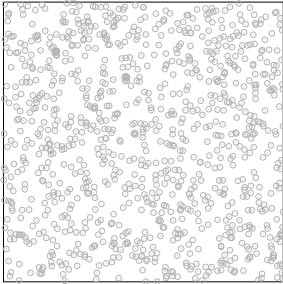
$N = 1000$

- ▶ Individuals are scattered across a two-dimensional space.
- ▶ In each generation, each individual produces a Poisson number of offspring (average one).
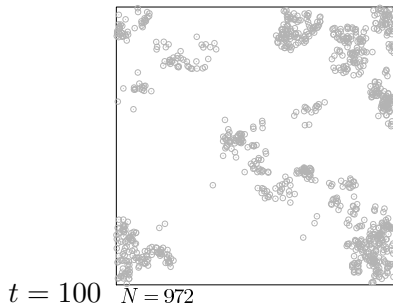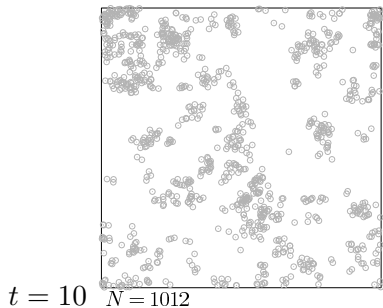- ▶ Offspring are scattered in a Gaussian distribution around their parent.

Poisson distribution with parameter 1
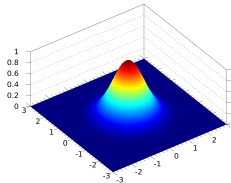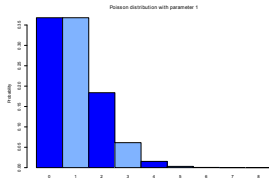
Mitch Gooding
Jerome Kelleher

The pain in the torus (Felsenstein, 1975)

With thanks to Jerome Kelleher
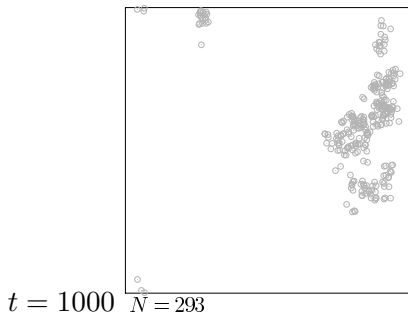

$t = 0$ $N = 1000$

$t = 10$ $N = 1012$

The pain in the torus (Felsenstein, 1975)

With thanks to Jerome Kelleher

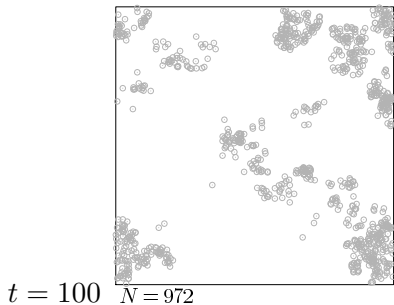$t = 10$ $N = 1012$ $t = 100$ $N = 972$

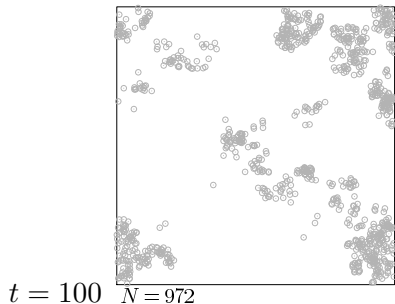# The pain in the torus (Felsenstein, 1975)

With thanks to Jerome Kelleher

$t = 100$ $N = 972$ $t = 1000$ $N = 293$

The pain in the torus (Felsenstein, 1975)

With thanks to Jerome Kelleher

$t = 100$ $N = 972$

$t = 1000$ $N = 293$
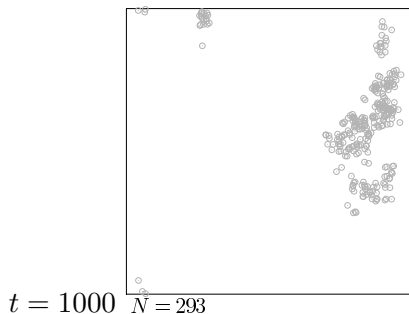
In $d = 1, 2$ population exhibits clumping/extinction

Felsenstein (1975). The pain in the torus: In $d = 1$, 2, independent reproduction $\implies$ clumping;

Mathematical problems

Felsenstein (1975). The pain in the torus: In $d = 1, 2$, independent reproduction $\implies$ clumping;

Local regulation $\implies$ correlated reproduction.

Mathematical problems

Felsenstein (1975). The pain in the torus: In $d = 1, 2$, independent reproduction $\implies$ clumping;

Local regulation $\implies$ correlated reproduction.

What about modifying the stepping stone model?

$$dp_t(x) = \frac{1}{2}\Delta p_t(x) + \sqrt{\frac{1}{N_e}p_t(x)(1 - p_t(x))}dW(t, x)$$

Mathematical problems

Felsenstein (1975). The pain in the torus: In $d = 1, 2$, independent reproduction $\implies$ clumping;

Local regulation $\implies$ correlated reproduction.

What about modifying the stepping stone model?

$$dp_t(x) = \frac{1}{2}\Delta p_t(x) + \sqrt{\frac{1}{N_e}p_t(x)(1 - p_t(x))}dW(t, x)$$

In 2D the diffusion limit fails over small scales

Mathematical problems

Felsenstein (1975). The pain in the torus: In $d = 1, 2$, independent reproduction $\implies$ clumping;

Local regulation $\implies$ correlated reproduction.

What about modifying the stepping stone model?

$$dp_t(x) = \frac{1}{2}\Delta p_t(x) + \sqrt{\frac{1}{N_e}p_t(x)(1 - p_t(x))}dW(t, x)$$

In 2D the diffusion limit fails over small scales ... and so does the obvious backwards model.

Biological problems

Genetic diversity much lower than expected from census numbers

Biological problems

Genetic diversity much lower than expected from census numbers
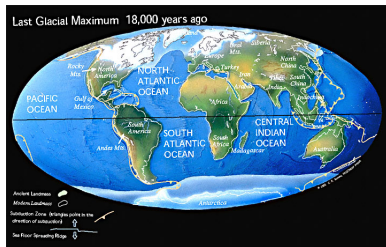
Allele frequencies correlated over long distances

Biological problems

Genetic diversity much lower than expected from census numbers

Allele frequencies correlated over long distances

Demographic history of many species dominated by large scale extinction-recolonisation events
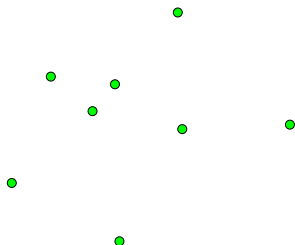
Small neighbourhood size

In a spatial continuum, a single individual can be parent to a significant proportion of the local population.
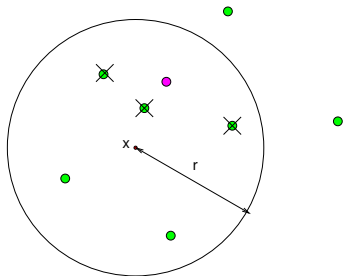
An individual based model

▶ Start with Poisson intensity λdx.
Events rate $dt \otimes dx \otimes \xi(dr, du)$.
Throw down ball $B(x, r)$.

An individual based model

- Start with Poisson intensity λdx.
  Events rate $dt \otimes dx \otimes \xi(dr, du)$.
  Throw down ball $B(x, r)$.

- If region empty, do nothing,
  otherwise:

- Choose parent from $B(x, r)$,
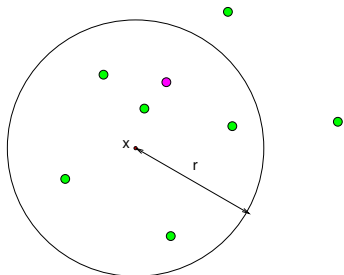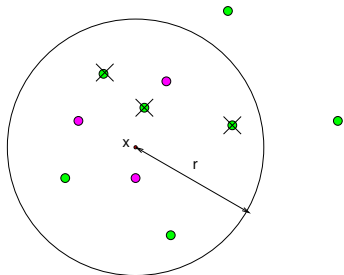
An individual based model

- ▶ Start with Poisson intensity λdx.
  Events rate $dt \otimes dx \otimes \xi(dr, du)$.
  Throw down ball $B(x, r)$.

- ▶ If region empty, do nothing,
  otherwise:

- ▶ Choose parent from $B(x, r)$,

- ▶ Each individual in region dies with
 probability u,

An individual based model

▶ Start with Poisson intensity λdx.
  Events rate $dt \otimes dx \otimes \xi(dr, du)$.
  Throw down ball $B(x, r)$.

▶ If region empty, do nothing,
  otherwise:

▶ Choose parent from $B(x, r)$,

▶ Each individual in region dies with
  probability u,

▶ New individuals born according to
  Poisson intensity $\lambda u \mathbf{1}_{B_r(x)}$.

An individual based model

- ▶ Start with Poisson intensity λdx.
  Events rate $dt \otimes dx \otimes \xi(dr, du)$.
  Throw down ball $B(x, r)$.

- ▶ If region empty, do nothing,
  otherwise:

- ▶ Choose parent from $B(x, r)$,

- ▶ Each individual in region dies with
  probability u,

- ▶ New individuals born according to
  Poisson intensity $\lambda u \mathbf{1}_{B_r(x)}$.

Offspring inherit type of parent
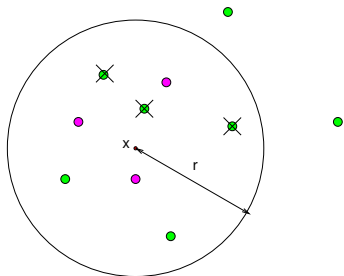
$\lambda \to \infty$ limit (no space)

Start from $\texttt{Poiss}(\lambda)$

If first reproduction event has 'impact' u

- $\texttt{Poiss}((1-u)\lambda)$ 'survivors';
- $\texttt{Poiss}(u\lambda)$ offspring.

As $\lambda \to \infty$ proportion u of individuals die and are replaced by offspring of the type of the parent.

The Λ-Fleming-Viot process

State $\{\rho(t, \cdot) \in \mathcal{M}_1(K), t \geq 0\}$. K space of genetic types.

- Poisson Point Process Π intensity $dt \otimes F(du)$
- if $(t, u) \in \Pi$, individual sampled at random from population at time $t-$ (i.e. choose $k \sim \rho(t-)$)
- proportion u of population replaced by offspring of chosen individual
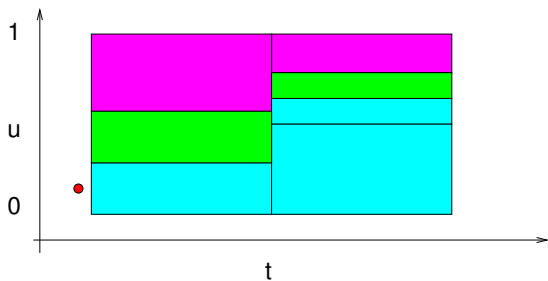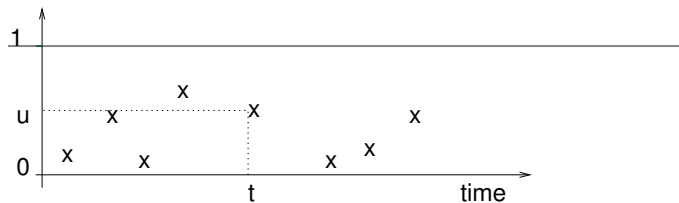
$$\rho(t, \cdot) = (1 - u)\rho(t-, \cdot) + u\delta_k.$$

$F(du) = \frac{\Lambda(du)}{u^2}$, Λ finite measure on $[0, 1]$.
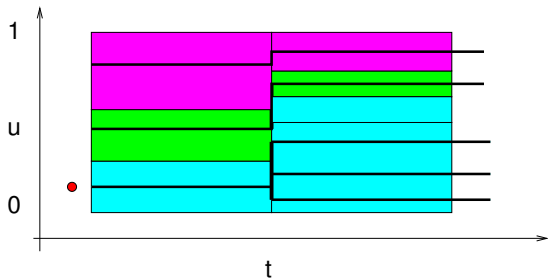
Donnelly & Kurtz (1999)

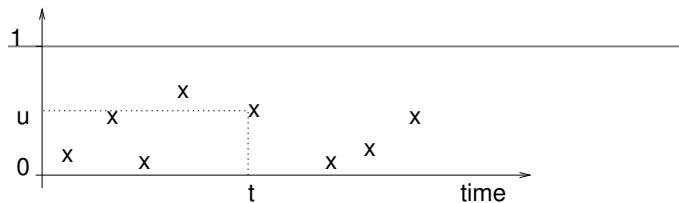('Generalised Fleming-Viot process', Bertoin & Le Gall 2003)

The Λ-Fleming-Viot process

The Λ-Fleming-Viot process

Λ-coalescents

Donnelly & Kurtz (1999), Pitman (1999), Sagitov (1999)

If there are currently n ancestral lineages, each transition involving j of them merging happens at rate

$$\beta_{n,j} = \int_0^1 u^j (1-u)^{n-j} \frac{\Lambda(du)}{u^2}$$

▶ Λ a finite measure on $[0,1]$
▶ Kingman's coalescent, $\Lambda = \delta_0$

State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$.

The spatial Λ-Fleming-Viot process Barton - E - Véber and friends

State $\{\rho(t,x,\cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.

The spatial Λ-Fleming-Viot process Barton - E - Véber and friends

State $\{\rho(t, x, \cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.
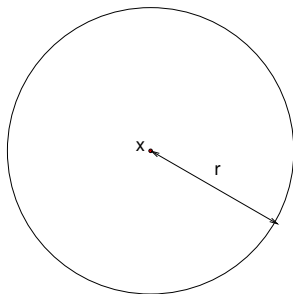
Dynamics: for each $(t, x, r, u) \in \Pi$,

State $\{\rho(t, x, \cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.

Dynamics: for each $(t, x, r, u) \in \Pi$,

The spatial Λ-Fleming-Viot process Barton - E - Véber and friends

State $\{\rho(t, x, \cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.

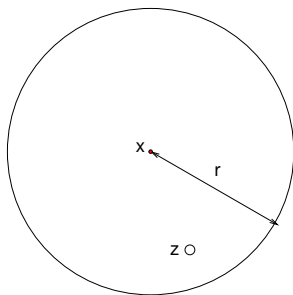Dynamics: for each $(t, x, r, u) \in \Pi$,

▶ $z \sim U(B_r(x))$

The spatial Λ-Fleming-Viot process Barton - E - Véber and friends

State $\{\rho(t, x, \cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.

Dynamics: for each $(t, x, r, u) \in \Pi$,
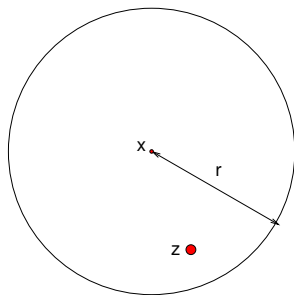
- $z \sim U(B_r(x))$
- $k \sim \rho(t-, z, \cdot)$.

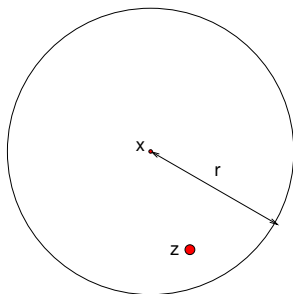# The spatial Λ-Fleming-Viot process Barton - E - Véber and friends

State $\{\rho(t, x, \cdot) \in \mathcal{M}_1(K), x \in \mathbb{R}^2, t \geq 0\}$. Π Poisson point process rate $dt \otimes dx \otimes \xi(dr, du)$ on $[0, \infty) \times \mathbb{R}^2 \times [0, \infty) \times [0, 1]$.

Dynamics: for each $(t, x, r, u) \in \Pi$,

- $z \sim U(B_r(x))$
- $k \sim \rho(t-, z, \cdot)$.

For all $y \in B_r(x)$,

$\rho(t, y, \cdot) = (1 - u)\rho(t-, y, \cdot) + u\delta_k.$

Backwards in time

▶ A *single* ancestral lineage evolves in series of jumps with intensity

$$dt \otimes \int_{(|x|/2, \infty)} \int_{[0,1]} \frac{L_r(x)}{\pi r^2} \, u \, \xi(dr, du) dx$$

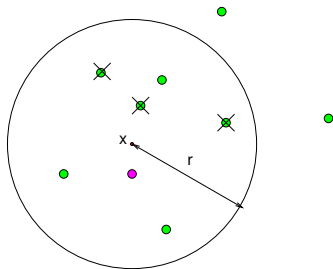on $\mathbb{R}_+ \times \mathbb{R}^2$ where $L_r(x) = |B_r(0) \cap B_r(x)|$.

Backwards in time

▶ A *single* ancestral lineage evolves in series of jumps with intensity

$$dt \otimes \int_{(|x|/2,\infty)} \int_{[0,1]} \frac{L_r(x)}{\pi r^2} \, u \, \xi(dr, du) dx$$

on $\mathbb{R}_+ \times \mathbb{R}^2$ where $L_r(x) = |B_r(0) \cap B_r(x)|$.
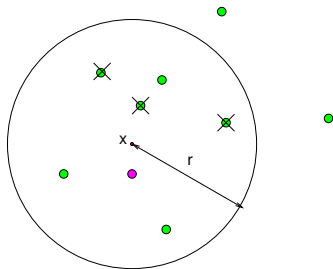
▶ Lineages can coalesce when hit by same 'event'.

Backwards in time
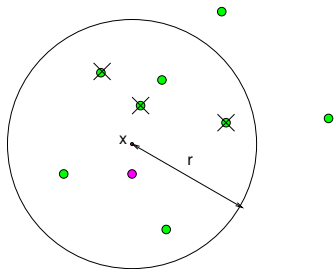
- A *single* ancestral lineage evolves in series of jumps with intensity

$$dt \otimes \int_{(|x|/2,\infty)} \int_{[0,1]} \frac{L_r(x)}{\pi r^2} \, u \, \xi(dr, du) dx$$

  on $\mathbb{R}_+ \times \mathbb{R}^2$ where $L_r(x) = |B_r(0) \cap B_r(x)|$.

- Lineages can coalesce when hit by same 'event'.

Note: If $\xi(dr, du) = \mu(dr) \otimes \delta_u$, rate of jumps $\propto u$.

Duality

Suppose $K = \{a, A\}$. Define $w(t, x) = \rho(t, x, \{a\})$ to be the proportion of the population at site x at time t that are of type a.

Duality

Suppose $K = \{a, A\}$. Define $w(t, x) = \rho(t, x, \{a\})$ to be the proportion of the population at site x at time t that are of type a.

(Only really defined up to a set of Lebesgue measure zero)

Duality

Suppose $K = \{a, A\}$. Define $w(t, x) = \rho(t, x, \{a\})$ to be the proportion of the population at site x at time t that are of type a.

(Only really defined up to a set of Lebesgue measure zero)

- Sample N_0 individuals from locations $\{X_i(0)\}_{i=1}^{N_0}$ from the present day population;
- Let $\{X_i(t)\}_{i=1}^{N_t}$ denote the positions of the random number of individuals ancestral to the sample at time t before the present

Duality

Suppose $K = \{a, A\}$. Define $w(t, x) = \rho(t, x, \{a\})$ to be the proportion of the population at site x at time t that are of type a.

(Only really defined up to a set of Lebesgue measure zero)

▶ Sample N_0 individuals from locations $\{X_i(0)\}_{i=1}^{N_0}$ from the present day population;

▶ Let $\{X_i(t)\}_{i=1}^{N_t}$ denote the positions of the random number of individuals ancestral to the sample at time t before the present

$$\mathbb{E}\left[\prod_{i=1}^{N_0} w\big(t, X_i(0)\big)\right] = \mathbb{E}\left[\prod_{i=1}^{N_t} w\big(0, X_i(t)\big)\right].$$

Direct analogue of our duality in the stepping stone model

Duality

Suppose $K = \{a, A\}$. Define $w(t, x) = \rho(t, x, \{a\})$ to be the proportion of the population at site x at time t that are of type a.

(Only really defined up to a set of Lebesgue measure zero)

► Sample N_0 individuals from locations $\{X_i(0)\}_{i=1}^{N_0}$ from the present day population;

► Let $\{X_i(t)\}_{i=1}^{N_t}$ denote the positions of the random number of individuals ancestral to the sample at time t before the present

$$\mathbb{E}\left[\prod_{i=1}^{N_0} w\big(t, X_i(0)\big)\right] = \mathbb{E}\left[\prod_{i=1}^{N_t} w\big(0, X_i(t)\big)\right].$$

(actually have to sample from random positions and integrate to circumvent issues with sets of Lebesgue measure zero)

Direct analogue of our duality in the stepping stone model

A framework for modelling

- ▶ Different spaces,
- ▶ Different shapes of event,
- ▶ Non-uniform replacement,
- ▶ Non-constant density,
- ▶ Multiple parents,

A framework for modelling

▶ Different spaces,
▶ Different shapes of event,
▶ Non-uniform replacement,
▶ Non-constant density,
▶ Multiple parents,
▶ Selection,
▶ Recombination,

A framework for modelling

▶ Different spaces,

▶ Different shapes of event,

▶ Non-uniform replacement,

▶ Non-constant density,

▶ Multiple parents,

▶ Selection,

▶ Recombination,

Robust results? ⤳ Scaling limits.

Example: Wright and Malécot again

The effect of mixed events on F(x, μ). A mixture of rare large events and frequent small events

OXFORD
UNIVERSITY PRESS