The Moran model with selection revisited

Sandra Kluth

joint work with Ellen Baake

Faculty of Technology
Bielefeld University

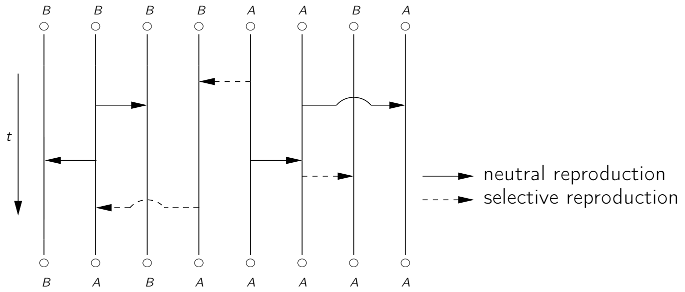
April 2, 2013

Outline

1) Moran model with selection and fixation probabilities
2) Connection to particle representation
   - labelled Moran model
   - defining events
   - simulation algorithm
3) Common ancestor type process

Moran model with selection

- N individuals
- Set of types: $S = \{A, B\}$
- Individuals of type A reproduce at rate $1 + s$, individuals of type B at rate 1
  Decomposition into neutral (types A and B, rate 1) and selective (just type A, rate s) reproductions (Krone/ Neuhauser 1997)

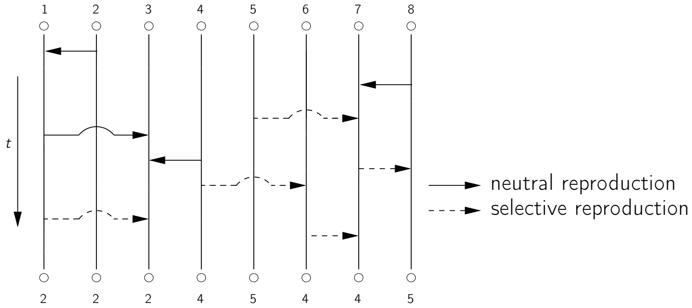


Moran model with selection

- $Z_t :=$ number of individuals of type A at time t, birth-death process with birth rates $\lambda_i = (1+s)i\frac{N-i}{N}$ and death rates $\mu_i = (N-i)\frac{i}{N}$
- $T_k := \min\{t \mid Z_t = k\}$
- Fixation probability:
  $h_i := \mathbb{P}(T_N < T_0 \mid Z_0 = i) = \frac{\sum_{j=N-i}^{N-1}(1+s)^j}{\sum_{j=0}^{N-1}(1+s)^j}$

Labelled Moran model

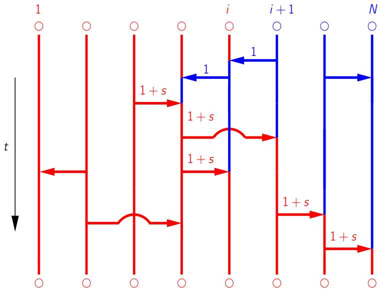
- N individuals, each characterised by label $i \in \{1, \dots, N\}$
- Offspring inherit parent's label
- Neutral arrows at rate $1/N$ (between every pair of labels), selective arrows at rate s/N (emanating from label i, pointing to label $j > i$)
- Spatial structure at time 0: Label i occupies position i



Ancestors and fixation probabilities

$I :=$ label that becomes fixed/ancestor

- $\mathbb{P}(I \leq i) = h_i$



Ancestors and fixation probabilities

$I :=$ label that becomes fixed/ancestor

- $\mathbb{P}(I \leq i) = h_i$
- $\eta_i := \mathbb{P}(I = i) = (1+s)^{N-i}\eta_N$
  with $\eta_N := \mathbb{P}(I = N) = \frac{1}{\sum_{j=0}^{N-1}(1+s)^j}$

Ancestors and fixation probabilities

$I :=$ label that becomes fixed/ancestor

- $\mathbb{P}(I \leq i) = h_i$
- $\eta_i := \mathbb{P}(I = i) = (1+s)^{N-i}\eta_N$
  with $\eta_N := \mathbb{P}(I = N) = \frac{1}{\sum_{j=0}^{N-1}(1+s)^j}$

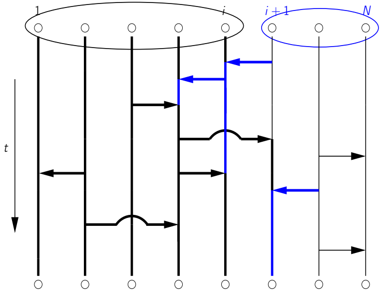
Based on particle *number* representation $\rightarrow$ decode particle representation

Particle representation behind $\eta_i = (1+s)^{N-i}\eta_N$

New descendants of labels in $\mathcal{S}$, $\mathcal{S} \subseteq \{1, \ldots, N\}$:

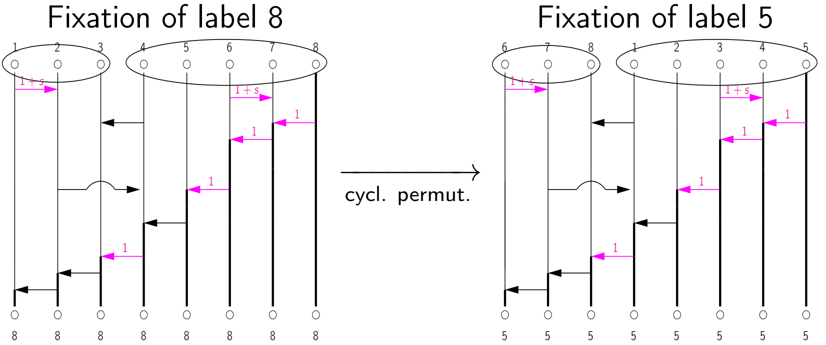
Descendants that increase the number of inidividuals in $\mathcal{S}$.

New descendants of labels in $\{i+1, \ldots, N\}$:



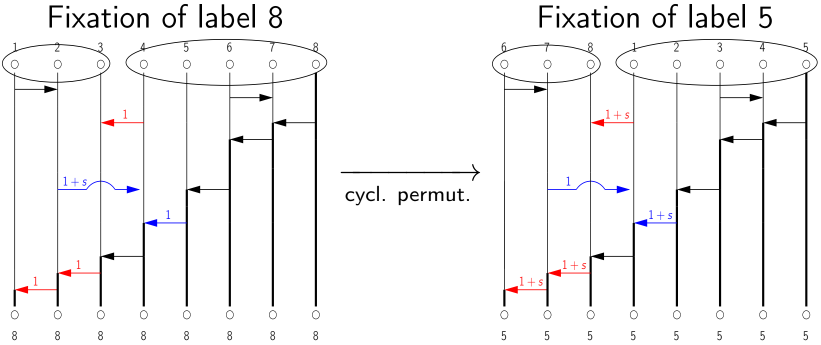
Particle representation behind $\eta_i = (1+s)^{N-i}\eta_N$

$N = 8, i = 5:$



Particle representation behind $\eta_i = (1+s)^{N-i}\eta_N$

$N = 8, i = 5:$

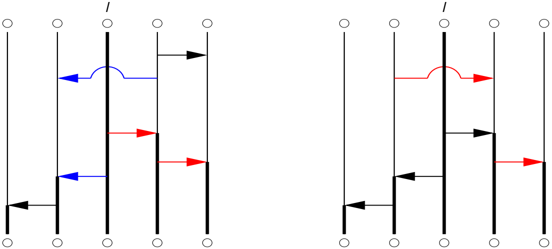


$$\eta_5 = \underbrace{\frac{1}{1+s}(1+s)}_{\text{new desc. of } \{6,7,8\}} \cdot \underbrace{(1+s)^3}_{\{6,7,8\}} \eta_8 = (1+s)^3 \eta_8$$

Definition

Let $I = i$.

Defining events:
Arrows emanating from labels $\{1, \dots, i\}$ and pointing to individuals with labels $\{i+1, \dots, N\}$ that are not new descendants of $\{i+1, \dots, N\}$.



defining event

Selective defining events

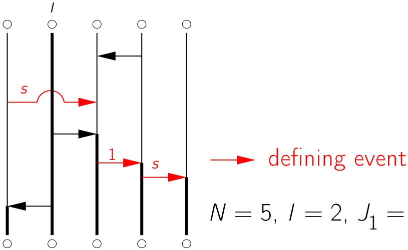
$Y :=$ number of selective defining events

- $\mathbb{P}(Y = n, I = i) = \binom{N-i}{n} s^n \eta_N$
- $\mathbb{P}(Y = n) = \sum_{i=1}^{N-n} \mathbb{P}(I = i, Y = n) = \binom{N}{n+1} s^n \eta_N$
- $h_i = \mathbb{P}(I \leq i) = \sum_{n=0}^{N-1} \mathbb{P}(I \leq i \mid Y = n)\mathbb{P}(Y = n)$

$$= \sum_{n=0}^{N-1} \left[\binom{N}{n+1} - \binom{N-i}{n+1} \right] s^n \eta_N$$

Targets of selective defining events

Let $Y = n$. Define $J_1, \ldots, J_n$ with $J_1 < \cdots < J_n$ as the (random) positions that are hit by selective defining events



$N = 5$, $I = 2$, $J_1 = 3$, $J_2 = 5$

- $\mathbb{P}(I = i, J_1 = j_1, \ldots, J_n = j_n \mid Y = n) = \frac{1}{\binom{N}{n+1}}$

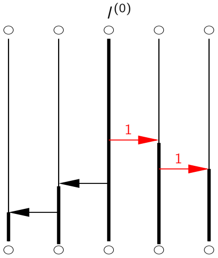
A simulation algorithm

Aim:
Generate $(I, J_1, \ldots, J_n)$

Generate Y. If $Y = n$ stop after step n.

Step 0: Generate $U^{(0)} \sim \mathcal{U}_{\{1,\ldots,N\}}$. Set $I^{(0)} := U^{(0)}$.

Genealogical interpretation:



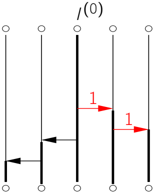
A simulation algorithm

Step 1: Generate $U^{(1)} \sim \mathcal{U}_{\{1,\dots,N\}\setminus U^{(0)}}$ independently of $U^{(0)}$.

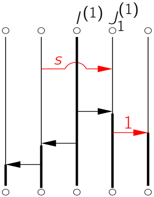
(a) If $U^{(1)} > I^{(0)}$:
Set $I^{(1)} := I^{(0)}$, $J_1^{(1)} := U^{(1)}$

(b) If $U^{(1)} < I^{(0)}$:
Set $I^{(1)} := U^{(1)}$, $J_1^{(1)} := I^{(0)}$

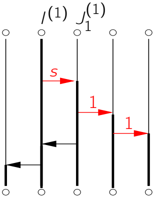
Step 0:



Step 1(a):



Step 1(b):



'shift to the left'

A simulation algorithm

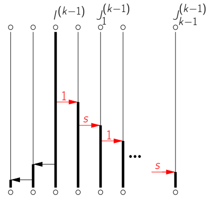
Step k: Generate

$U^{(k)} \sim \mathcal{U}_{\{1,\dots,N\}\setminus\{U^{(0)},\dots,U^{(k-1)}\}}$.

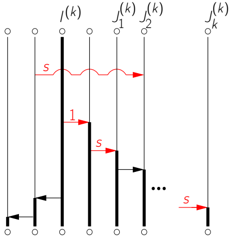
(a) If $U^{(k)} > I^{(k-1)}$: Set $I^{(k)} := I^{(k-1)}$

(b) If $U^{(k)} < I^{(k-1)}$: Set $I^{(k)} := U^{(k)}$

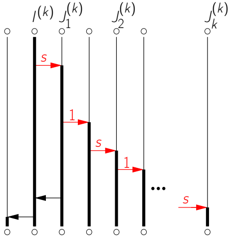
Step $k-1$:



Step k(a):



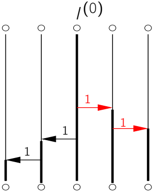
Step k(b):



'shift to the left'

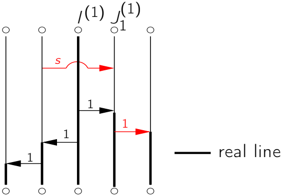
Relation to the ancestral selection graph

Step 0:



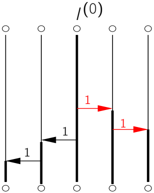
$\xrightarrow{\text{(a)}}$

Step 1:



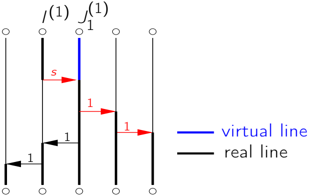
ASG without branching

Step 0:



$\xrightarrow{\text{(b)}}$

Step 1:



ASG with one branching event

Genealogical interpretation

$$
\begin{aligned}
h_i &= \mathbb{P}(I \leq i | Y = 0)\mathbb{P}(Y \geq 0) \\
&\quad + \sum_{n \geq 1} [\mathbb{P}(I \leq i | Y = n) - \mathbb{P}(I \leq i | Y = n-1)]\, \mathbb{P}(Y \geq n) \\
&= \mathbb{P}(I^{(0)} \leq i) + \sum_{n \geq 1} \left[\mathbb{P}(I^{(n)} \leq i) - \mathbb{P}(I^{(n-1)} \leq i)\right] \mathbb{P}(Y \geq n) \\
&= \mathbb{P}(I^{(0)} \leq i) + \sum_{n \geq 1} \mathbb{P}(I^{(n)} \leq i, I^{(n-1)} > i)\mathbb{P}(Y \geq n)
\end{aligned}
$$

Decomposition according to first step in which the ancestor has a label in $\{1, \ldots, i\}$.

Diffusion limit under weak selection ($Ns \xrightarrow{N\to\infty} \sigma$)

- $h_i \xrightarrow{\frac{i}{N}\to x} h(x) = \frac{1-e^{-\sigma x}}{1-e^{-\sigma}}$

Diffusion limit under weak selection ($Ns \xrightarrow{N\to\infty} \sigma$)

- $h_i \xrightarrow{\frac{i}{N}\to x} h(x) = \frac{1-e^{-\sigma x}}{1-e^{-\sigma}}$
- $a_n := \lim_{N\to\infty} \mathbb{P}(Y \geq n)$
  - $1 - a_1 = \lim_{N\to\infty} \mathbb{P}(Y = 0) = \frac{\sigma}{\exp(\sigma)-1}$
  - $a_n - a_{n+1} = \lim_{N\to\infty} \mathbb{P}(Y = n) = \frac{\sigma}{n+1}(a_{n-1} - a_n)$

Diffusion limit under weak selection ($Ns \xrightarrow{N\to\infty} \sigma$)

- $h_i \xrightarrow{\frac{i}{N}\to x} h(x) = \frac{1-e^{-\sigma x}}{1-e^{-\sigma}}$
- $a_n := \lim_{N\to\infty} \mathbb{P}(Y \geq n)$
  - $1 - a_1 = \lim_{N\to\infty} \mathbb{P}(Y = 0) = \frac{\sigma}{\exp(\sigma)-1}$
  - $a_n - a_{n+1} = \lim_{N\to\infty} \mathbb{P}(Y = n) = \frac{\sigma}{n+1}(a_{n-1} - a_n)$
  - $h_i = \mathbb{P}(I^{(0)} \leq i) + \sum_{n\geq 1} \mathbb{P}(I^{(n)} \leq i, I^{(n-1)} > i)\mathbb{P}(Y \geq n)$

$$\xrightarrow{\frac{i}{N}\to x} h(x) = x + \sum_{n\geq 1} x(1-x)^n a_n$$

$$= \frac{1}{\exp(\sigma)-1} \sum_{n\geq 1} \frac{1}{n!}(1-(1-x)^n)\sigma^n$$

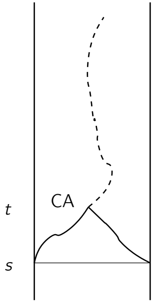
Common ancestor type process

Moran model with mutation and selection:

- N individuals
- Set of types: $S = \{A, B\}$
- Individuals of type A reproduce at rate $1 + s$, individuals of type B at rate 1
- mutations: $i \xrightarrow{u\nu_j} j$, $i, j \in S$
  here:
  - u general mutation rate with $Nu \xrightarrow{N\to\infty} \theta$
  - ν_j probability of mutations to type j ($\nu_A + \nu_B = 1$)
- Stationary density $\pi_X(x) = C(1-x)^{\theta\nu_B - 1} x^{\theta\nu_A - 1} \exp(\sigma x)$ (Wright's formula)

Common ancestor type process

- Population is stationary
- Common ancestor at time t:
  Unique individual (at time t) that is ancestral to the whole population at some time $s > t$
- I_t = type of common ancestor at time t
  $(I_t)_{t \geq 0}$ common ancestor type process



Stationary type distribution $\alpha = (\alpha_i)_{i \in S}$?

Taylor (2007)

- $(I_t, X_t)_{t \geqslant 0}$ with states (i, x), $i \in S$, $x \in [0,1]$
- $h(x) :=$ conditional probability that the common ancestor at time t is of type A, given that the frequency of type-A individuals at time t is x ($h(0) = 0$, $h(1) = 1$)
- Stationary distribution:

$$\pi_T(0, x) = h(x)\,\pi_X(x)$$
$$\pi_T(1, x) = (1 - h(x))\,\pi_X(x)$$

$\Rightarrow$ Stationary type distribution $\alpha_i = \int_0^1 \pi_T(i, x)dx$

Fearnhead (2002)

$$h(x) = x + x \sum_{n \geq 1} a_n (1-x)^n$$

Recursion of Fearnhead's coefficients a_n, $n \geq 0$:

$$(n + \theta\nu_1)\, a_n - (n + \sigma + \theta)\, a_{n-1} + \sigma a_{n-2} = 0, \quad n \geqslant 2,$$

with initial values $a_0 = 1$ and

$$a_1 = \frac{\sigma}{1+\theta\nu_1}(1-\tilde{x}), \text{where } \tilde{x} = \frac{\int_0^1 p^{\theta\nu_0+1}\,(1-p)^{\theta\nu_1}\exp(\sigma p)dp}{\int_0^1 p^{\theta\nu_0}\,(1-p)^{\theta\nu_1}\exp(\sigma p)dp}$$

Thank you for your attention!