

# Reproductive systems

## Reproductive selection in a Moran model

Raphaël Tran-Thanh

Summer Seminar - Université du Québec à Montréal

June 11, 2025

This work was advised by Diala Abu Awad (Université Paris-Saclay) and Camille Coron (INRAE)

## I - Biological and mathematical framework

- **Biological context**
  - \* Global background
  - \* Biparentality and reproductive selection
- **Modeling**
  - \* Mating preferences in a Moran model
  - \* Subpopulation dynamics

**General problem:** estimate the influence of the reproductive system of a species on its genetic diversity

- **Biological parameters:**

- n-Parentality: biparentality, self-fertilization
- Ploidity: haploid, diploid
- Selection factors: genetic advantage/disadvantage

- **Our tools:**

- Biparental Moran model<sup>\*</sup>
- Convergence to a deterministic system
- Infinite selection hypothesis

<sup>\*</sup> **Adapted from:** Coron C, Le Jan Y. Pedigree in the biparental Moran model. J Math Biology. 2022 May 9;84(6):51

# Our idea - Hypotheses and relevancy

**Our problem:** How can we quantify the impact of genetically induced mating preferences on genetic diversity?

## - Hypotheses:

- **Structured population:** two subpopulations *A* and *B* of a **haploid** species distinguished by their genome.
  - Assume that a single gene (or locus) codes for the type.
  - Each offspring receives half of each parent's genome.
- **Mating preferences:** the individual's type induces potentially advantageous mating preferences.
- **Biparentality:** not only a matter of scaling!

# Two scenarios

We considered two sorts of reproductive advantage:

- **Attractive subpopulation:** all individuals preferably mate with advantaged individuals.
- **Homogamous subpopulation:** advantaged individuals preferably mate within their subpopulation.

In each scenario, we aimed to study how these advantaged individuals invade the population, and to compute the proportion of the present genome that originates from advantaged ancestors.

## Reproductive process in a Moran model

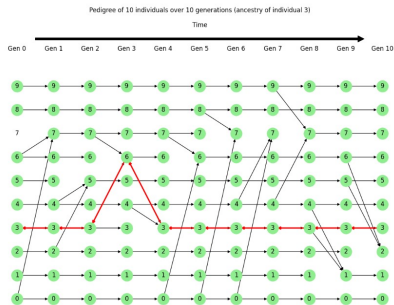
- Consider two genetic types  $A$  and  $B$  within a population of fixed size  $N$ .
- At each time step (generation)  $n \in \mathbb{N}$ :
  - Choose a mother  $\mu_n$  uniformly at random.
  - Choose a father  $\pi_n$  according to the mating preferences of  $\mu_n$ .
  - Each parent passes on half of their genome to the offspring  $\beta_n$ . In particular, if we denote by  $T$  the type:

$$\mathbb{P}[T(\beta_n) = T(\mu_n)] = \mathbb{P}[T(\beta_n) = T(\pi_n)] = \frac{1}{2}$$

- The offspring  $\beta_n$  replaces an individual  $\kappa_n$  chosen uniformly at random.

**⚠** Genes are inherited independently from one another!

# Model simulation - The neutral case



- The Moran model defines the pedigree of a population as a random oriented graph  $\mathcal{G}$ :
- We define the probability, given the graph  $\mathcal{G}$ , that a gene currently observed in the genome of individual  $i$  originates from individual  $j$ :

$$W_n(i, j) = \mathbb{P}(X_n^{(n)} = j | X_0^{(n)} = i, \mathcal{G})$$

# The case of a selective advantage

Introduce a selective advantage in a minority of the population in the form of mating preferences, encoded by:

$$\begin{bmatrix} 1 + s_{11} & 1 \\ 1 + s_{21} & 1 \end{bmatrix} \quad s_{11}, s_{21} \in [-1, \infty)$$

We aim to study the weight of this initially advantaged subpopulation in the asymptotic genome of the population. Biological interpretation:

- **Homogamy:**  $s_{11} > 0, s_{21} = 0$ . Subpopulation 1 tends to be homogamous.
- **Attractive subpopulation:**  $s_{11}, s_{21} > 0$ . All individuals preferably mate with individuals from subpopulation 1. For the sake of simplicity and symmetry, assume that  $s_{11} = s_{21}$  and denote it by  $s$ .
- Example ! With the following matrix:  $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$



## II - Our results

- **For each model (scenario):**
  - \* Stochastic dynamics
  - \* Convergence to a deterministic system
  - \* Infinite selection and long time behavior
- **Illustration of path convergence**

# Quantities of interest

- Given the structure of our population, it is more informative to study the following quantities:

$$U_n = \sum_{l \in \mathcal{Y}_n} \sum_{l' \in \mathcal{Y}_0} W_n(l, l') \quad \text{and} \quad V_n = \sum_{l \notin \mathcal{Y}_n} \sum_{l' \in \mathcal{Y}_0} W_n(l, l'),$$

for which we will seek a deterministic limit at the right time scale.

- The study of these quantities and their convergence was done by Coron and Le Jan<sup>1</sup> in a different setting (selection on death).
- We introduce the following process which we will study throughout this work:

$$Z_n = \left( \frac{Y_n}{N}, \frac{U_n}{N}, \frac{V_n}{N} \right)_{n \in \mathbb{N}}$$

**⚠**  $(Z_n)_n$  is not a Markov chain!

---

<sup>1</sup>Coron, Le Jan. Genetic contribution of an advantaged mutant in the biparental Moran model - finite selection (2025)

# Dynamics of $Z_n$ (attractive subpopulation)

We can split the dynamics of  $Z_n$  into four parts depending on whether  $\kappa_n$  and  $\beta_n$  belong to  $\mathcal{Y}_n$  or not. Taking the expectation yields the average behavior of the process from one generation to the next:

## Proposition

The stochastic process  $(Z_n)_{n \in \mathbb{N}}$  satisfies the following dynamics:

$$\begin{aligned}\mathbb{E}[Z_{n+1} - Z_n | \mathcal{F}_n^Z] &= \frac{1}{N} \times \left[ \left( \frac{(s/2)(Y_n/N)(1 - Y_n/N)}{1 + s(Y_n/N)} \right. \right. \\ &\quad - \frac{U_n}{N} + \frac{(1+s)(U_n/N)(Y_n/N)}{1 + s(Y_n/N)} + \frac{(1+s/2)}{2[1 + s(Y_n/N)]} \left( \frac{Y_n}{N} \frac{V_n}{N} + \left(1 - \frac{Y_n}{N}\right) \frac{U_n}{N} \right), \\ &\quad \left. \left. - \frac{V_n}{N} + \frac{(V_n/N)(1 - Y_n/N)}{1 + s(Y_n/N)} + \frac{(1+s/2)}{2[1 + s(Y_n/N)]} \left( \frac{Y_n}{N} \frac{V_n}{N} + \left(1 - \frac{Y_n}{N}\right) \frac{U_n}{N} \right) \right) \right], \\ &=: \frac{1}{N} G(Z_n).\end{aligned}\tag{1}$$

# Convergence (attractive subpopulation)

## Theorem

Let  $a \in (0, 1)$ . If the proportion of initially advantaged individuals  $\frac{Y_0}{N}$  converges in probability to  $a$  when  $N \rightarrow \infty$ , then for all  $c \geq 0$ ,

$$\sup_{0 \leq t \leq c} \|Z_{[Nt]} - z_t\| \xrightarrow{N \rightarrow \infty} 0$$

in probability, where  $(z_t)_{t \geq 0} = (y_t, u_t, v_t)$  is the solution of the following differential system:

$$\begin{cases} y' = \frac{(s/2)y(1-y)}{(1+s)y+(1-y)} \\ u' = -u + \frac{uy(1+s)}{(1+s)y+(1-y)} + \frac{(1+s/2)(vy+(1-y)u)}{2[(1+s)y+(1-y)]} \\ v' = -v + \frac{v(1-y)}{(1+s)y+(1-y)} + \frac{(1+s/2)(vy+(1-y)u)}{2[(1+s)y+(1-y)]} \end{cases} \quad (2)$$

with initial condition  $(a, a, 0)$ .

## Indications for proving the theorem

- **Martingale decomposition:**

- Introduce the increments:

$$A_{n+1} = Z_{n+1} - \mathbb{E}[Z_{n+1} | \mathcal{F}_n^Z]$$

- Consequently,  $Z_n = Z_0 + \sum_{k=1}^n A_k + \frac{1}{N} \sum_{k=1}^{n-1} G(Z_k)$ , where  $(\sum_{k=1}^n A_k)_{n \geq 0}$  is a martingale.

- **Control via Doob's maximal inequality:**

- Doob's maximal inequality yields a control over the sum of the  $A_k$ .
- The rest of the proof uses Gronwall's lemma to bound the following quantity:

$$\|Z_{\lfloor Nt \rfloor} - z_t\|$$

# Illustrating the convergence (attractive subpopulation)

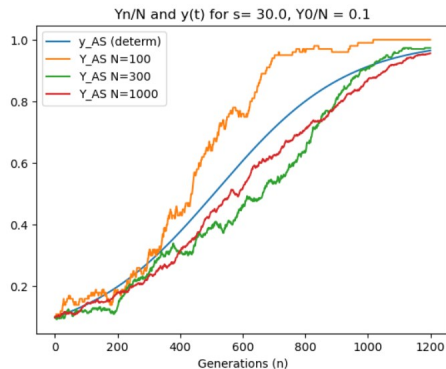


Figure:  $Y_n$  and  $y(t)$ ,  $s = 30$

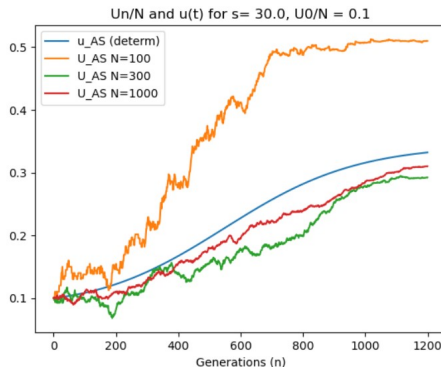


Figure:  $U_n$  and  $u(t)$ ,  $s = 30$

- We need to rescale each plot in order to observe convergence!

# Solution to the limiting ODE (attractive subpopulation)

- The following differential system:

$$\begin{cases} y' = \frac{(s/2)y(1-y)}{(1+s)y + (1-y)} \\ u' = -u + \frac{uy(1+s)}{(1+s)y + (1-y)} + \frac{(1+s/2)(vy + (1-y)u)}{2[(1+s)y + (1-y)]} \\ v' = -v + \frac{v(1-y)}{(1+s)y + (1-y)} + \frac{(1+s/2)(vy + (1-y)u)}{2[(1+s)y + (1-y)]} \end{cases}, \quad (3)$$

has a unique solution  $z_t = (y_t, u_t, v_t)_{t \geq 0}$  with initial condition  $(a, a, 0)$ ,  $a \in (0, 1)$ , which reads as follows:

$$\begin{cases} y_t = \sigma^{-1} \left( \frac{a}{(1-a)^{1+s}} \exp\left(\frac{s}{2}t\right) \right) \\ u_t = \frac{y_t}{2} \left[ 1 + \left( \frac{(1-y_t)a}{y_t(1-a)} \right)^{\frac{2+s}{2s}} - \int_a^{y_t} \frac{(1+s/2)(1-2w)}{sw(1-w)} \left( \frac{(1-w)a}{w(1-a)} \right)^{\frac{2+s}{2s}} dw \right] \\ v_t = \frac{1-y_t}{2} \left[ 1 - \left( \frac{(1-y_t)a}{y_t(1-a)} \right)^{\frac{2+s}{2s}} - \int_a^{y_t} \frac{(1+s/2)(1-2w)}{2(1+sw)} \left( \frac{(1-w)a}{w(1-a)} \right)^{\frac{2+s}{2s}} dw \right], \end{cases}$$

where  $\sigma : (0, 1) \rightarrow (0, \infty)$ ,  $y \mapsto \frac{y}{(1-y)^{1+s}}$ , is a  $\mathcal{C}^1$ -diffeomorphism.

# Infinite selection (attractive subpopulation)

- **Infinite selection:** letting  $s \rightarrow \infty$ , yields a simpler differential system.
  - This asymptotic case allows us to compare the different scenarios with the same selection intensity.

$$\begin{cases} y' = \frac{1}{2}(1 - y) \\ u' = \frac{1}{4y}(u(1 - y) + vy) \\ v' = -v + \frac{1}{4y}(u(1 - y) + vy). \end{cases} \quad (4)$$

- After explicitly solving the system, we can compute the asymptotic weight of advantaged ancestors in the population at time  $t \rightarrow \infty$ :

$$u_t \xrightarrow[t \rightarrow \infty]{} \sqrt{\frac{a}{1-a}} \left[ \arctan \left( \sqrt{\frac{a}{1-a}} \right) \right] := l_a^{AS}$$



# The case of a homogamous subpopulation

- We proved analogous results in the homogamous setting.

$$\Pi_H = \begin{bmatrix} 1+s & 1 \\ 1 & 1 \end{bmatrix} \quad s > 0.$$

- The limiting differential system then becomes:

$$\begin{cases} y' = \frac{(s/2)y^2(1-y)}{(1+s)y + (1-y)} = \frac{(s/2)y^2(1-y)}{1+sy} \\ u' = -u + uy \frac{1+s}{1+sy} + \frac{1}{4} \left( \frac{2+sy}{1+sy} \right) (u(1-y) + vy) \\ v' = -vy + \frac{1}{4} \left( \frac{2+sy}{1+sy} \right) (u(1-y) + vy), \end{cases} \quad (5)$$

- And for infinite selection  $s = \infty$ , we can again compute the asymptotic weight:

$$u_t \xrightarrow[t \rightarrow \infty]{} \sqrt{\frac{a}{1-a}} \left[ \frac{\pi}{2} - \arctan\left(\sqrt{\frac{a}{1-a}} e^{1/4}\right) \right] := l_a^H \quad (6)$$

## III- Discussion

- **Comparing mating preference scenarios**
  - \* Invasion front
  - \* Genetic weight of advantaged ancestors
- **Perspectives**
  - \* Fixation of an advantageous mutation
  - \* Limiting distribution of  $U_n/N$

# Comparing mating preference scenarios

Let us study the invasion fronts:

$$y'_{attirant} = \frac{(s/2)y(1-y)}{(1+s)y + (1-y)}, \quad y'_{homo} = \frac{(s/2)y^2(1-y)}{(1+s)y + (1-y)}$$

$$y'_{mort} = \frac{sy(1-y)}{(1+s)(1-y) + y}$$

- In the death-based selection model, the decay rate of unattractive individuals is twice the growth rate of advantaged individuals.
- Additional factor  $y$  in  $y'_h$  compared to  $y'_a$  : sensitivity to the initial population.

# Comparing differential systems

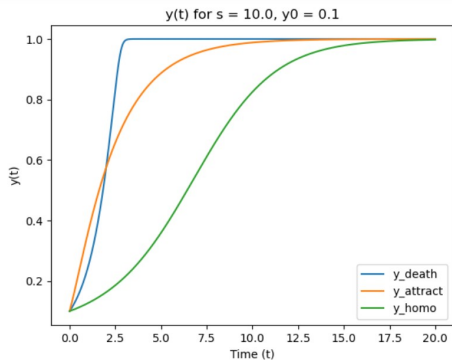


Figure:  $y(t)$  with  $s = 10$ ,  $a = 0.1$

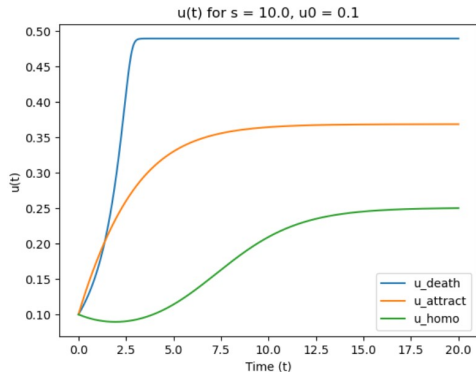


Figure:  $u(t)$  with  $s = 10$ ,  $a = 0.1$

- Death-based selection seems to be the strongest of all three (in terms of invasion front and asymptotic weight).
- Different front profiles (not simply a scaling issue!)

# Comparing advantages when $a \approx 1$

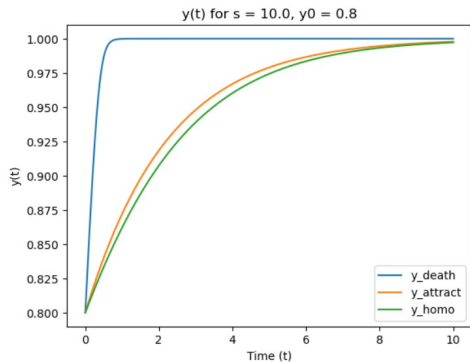


Figure:  $y(t)$  with  $s = 10$ ,  $a = 0.8$

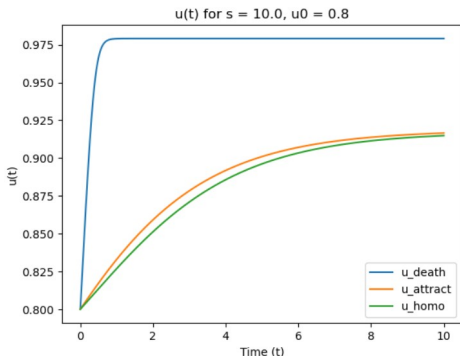


Figure:  $u(t)$  with  $s = 10$ ,  $a = 0.8$

- When  $a$  is close to 1, the curves in the homogenous and the attractive subpopulation cases converge.

# Comparing scenarios under weak selection pressure

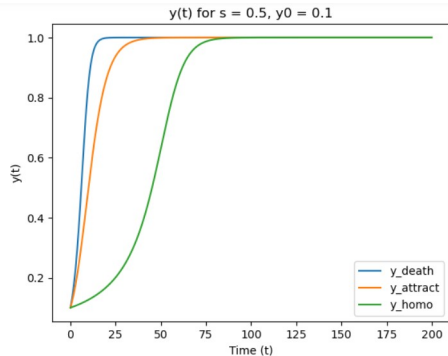


Figure:  $y(t)$  with  $s = 0.5$ ,  $a = 0.1$

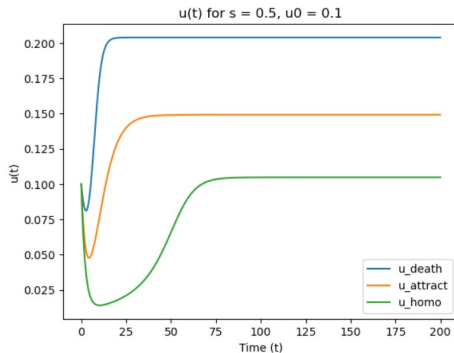


Figure:  $u(t)$  with  $s = 0.5$ ,  $a = 0.1$

- For a weak selection coefficient, we observe an initial drop in genetic weights in all three cases.

# Comparing asymptotic weights

Under infinite selection  $s = \infty$ , with an initial proportion  $a = 1\%$  of advantaged ancestors, what proportion of their genes will they pass onto the future population over a long time?

- $u_t \xrightarrow[t \rightarrow \infty]{} l_{1\%}^{AS} \approx 14.7\%$  (attractive subpopulation)
- $u_t \xrightarrow[t \rightarrow \infty]{} l_{1\%}^H \approx 14.5\%$  (homogamous subpopulation)
- $u_t \xrightarrow[t \rightarrow \infty]{} l_{1\%}^D \approx 19\%$  (death-based selection)

# Perspectives and future developments

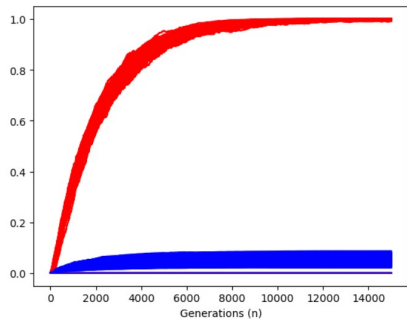


Figure: Attractive subpopulation

- Fixation/extinction time,  $\mathbb{P}[T_N < T_0] \rightarrow 1$  ? (work in progress)
- Studying the asymptotic distribution of the weights  $U_n/N$ ,  $V_n/N$
- Selection in a more general setting (nasty)

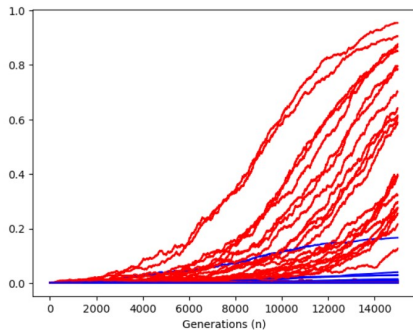


Figure: Homogamous subpopulation