

# Processus de branchement pour des populations structurées et estimateurs pour la division cellulaire

Aline Marguet

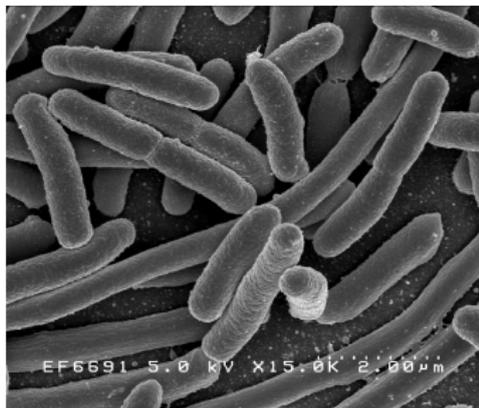
Sous la direction de V. Bansaye et M. Hoffmann

27 novembre 2017



## Décrire la dynamique d'une population de cellules

Modèle individu-centré où chaque cellule est caractérisée par un trait (âge, taille, nombre de parasites, ...).



- ▶ Comprendre le rôle de certaines caractéristiques dans la division cellulaire.
- ▶ Étudier le mécanisme de vieillissement cellulaire.
- ▶ Inférence statistique.

# Étude du mécanisme de vieillissement cellulaire

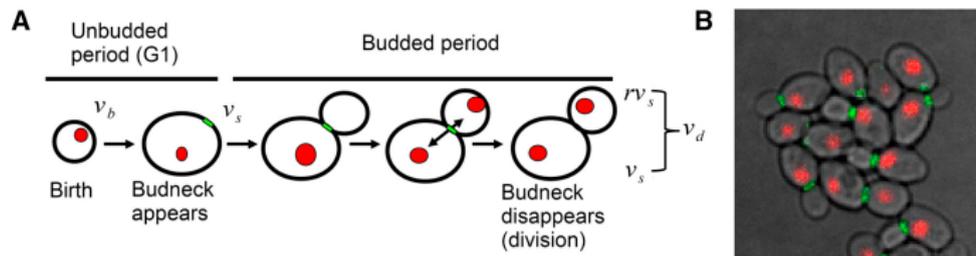


Figure: Image : Soifer, Robert & Amir, 2016.

- ▶ Division asymétrique : répartition asymétrique du trait
- ▶ Phénomène de rajeunissement, plateau de mortalité

# Plan

Échantillonnage uniforme

Loi des grands nombres

Estimation statistique dans des populations structurées

## Description du modèle

- ▶ Le trait  $(X_t^u)_{t \geq 0}$  de chaque individu  $u$  suit un processus de Markov de générateur infinitésimal  $\mathcal{G}$ .

## Description du modèle

- ▶ Le trait  $(X_t^u)_{t \geq 0}$  de chaque individu  $u$  suit un processus de Markov de générateur infinitésimal  $\mathcal{G}$ .
- ▶ Un individu  $u$  meurt au temps  $t$  au taux  $B(X_t^u)$ , i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (X_s^u, \alpha(u) \leq s \leq t)) = \exp\left(-\int_{\alpha(u)}^t B(X_s^u) ds\right),$$

où  $\beta(u)$  désigne la durée de vie de  $u$  et  $\alpha(u)$  son temps de naissance.

## Description du modèle

- ▶ Le trait  $(X_t^u)_{t \geq 0}$  de chaque individu  $u$  suit un processus de Markov de générateur infinitésimal  $\mathcal{G}$ .
- ▶ Un individu  $u$  meurt au temps  $t$  au taux  $B(X_t^u)$ , i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (X_s^u, \alpha(u) \leq s \leq t)) = \exp\left(-\int_{\alpha(u)}^t B(X_s^u) ds\right),$$

où  $\beta(u)$  désigne la durée de vie de  $u$  et  $\alpha(u)$  son temps de naissance.

- ▶ À sa mort, un individu de trait  $x$  est remplacé par  $A_u$  descendants. Les traits à la naissance sont choisis suivant une loi de probabilité dépendant de  $x$  et  $A_u$ .

## Description du modèle

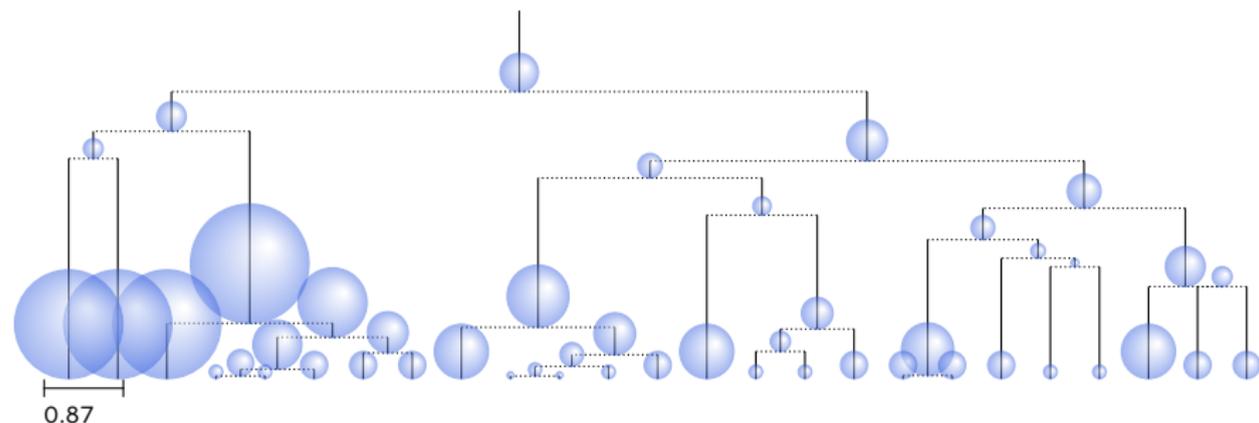
- ▶ Le trait  $(X_t^u)_{t \geq 0}$  de chaque individu  $u$  suit un processus de Markov de générateur infinitésimal  $\mathcal{G}$ .
- ▶ Un individu  $u$  meurt au temps  $t$  au taux  $B(X_t^u)$ , i.e.

$$\mathbb{P}(\beta(u) > t | \alpha(u), (X_s^u, \alpha(u) \leq s \leq t)) = \exp\left(-\int_{\alpha(u)}^t B(X_s^u) ds\right),$$

où  $\beta(u)$  désigne la durée de vie de  $u$  et  $\alpha(u)$  son temps de naissance.

- ▶ À sa mort, un individu de trait  $x$  est remplacé par  $A_u$  descendants. Les traits à la naissance sont choisis suivant une loi de probabilité dépendant de  $x$  et  $A_u$ .
- ▶ Conditionnellement au trait de leur ancêtre, les descendants évoluent ensuite indépendamment en suivant cette même dynamique.

# Réalisation du processus structuré en taille



- ▶ taille initiale :  $X_0^\emptyset = 1$ ,
- ▶ croissance exponentielle à taux 1,
- ▶ temps final :  $T = 3$ ,
- ▶ division à taux  $B(x) = x$ .

## Description avec des processus à valeurs mesures

 Fournier et Méléard 2004, Tran 2006, Bansaye et Tran 2011, etc...

$$Z_t = \sum_{u \in V_t} \delta X_t^u,$$

- ▶  $V_t$  : individus en vie au temps  $t$ ,

## Description avec des processus à valeurs mesures



Fournier et Méléard 2004, Tran 2006, Bansaye et Tran 2011, etc...

$$Z_t = \sum_{u \in V_t} \delta X_t^u,$$

- ▶  $V_t$  : individus en vie au temps  $t$ ,
- ▶  $N_t = \#V_t$ .

## Description avec des processus à valeurs mesures

 Fournier et Méléard 2004, Tran 2006, Bansaye et Tran 2011, etc...

$$Z_t = \sum_{u \in V_t} \delta X_t^u,$$

- ▶  $V_t$  : individus en vie au temps  $t$ ,
- ▶  $N_t = \#V_t$ .

Pour définir rigoureusement le processus comme unique solution forte d'une EDS (et pour éviter l'explosion de la population en temps fini), on considère deux types d'hypothèses:

- ▶ sur les évènements de division,
- ▶ sur la dynamique du trait.

## Hypothèses sur les évènement de division

1. Il existe  $b_1, b_2 \geq 0$  et  $\gamma \geq 1$  tels que pour tout  $x \in \mathcal{X}$ ,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

## Hypothèses sur les évènement de division

1. Il existe  $b_1, b_2 \geq 0$  et  $\gamma \geq 1$  tels que pour tout  $x \in \mathcal{X}$ ,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

2. Il existe  $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  telle que pour tout  $x = (y, t) \in \mathcal{X}$  et  $k \in \mathbb{N}$  :

$$\sum_{i=1}^k F_i^{(k)}(x) \leq x \vee \ell(t), \text{ coordonnées par coordonnées,}$$

où  $F_i^{(k)}(x)$  est le trait du  $i$ ème descendant parmi  $k$  d'un individu de trait  $x$  à la division.

## Hypothèses sur les évènements de division

1. Il existe  $b_1, b_2 \geq 0$  et  $\gamma \geq 1$  tels que pour tout  $x \in \mathcal{X}$ ,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

2. Il existe  $\ell : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  telle que pour tout  $x = (y, t) \in \mathcal{X}$  et  $k \in \mathbb{N}$  :

$$\sum_{i=1}^k F_i^{(k)}(x) \leq x \vee \ell(t), \text{ coordonnées par coordonnées,}$$

où  $F_i^{(k)}(x)$  est le trait du  $i$ ème descendant parmi  $k$  d'un individu de trait  $x$  à la division.

3. Il existe  $\bar{m} \geq 0$  tel que pour tout  $x \in \mathcal{X}$ ,

$$m(x) = \sum_k k p_k(x) \leq \bar{m}.$$

## Hypothèse sur la dynamique du trait

Il existe  $c_1, c_2 \geq 0$  tels que pour tout  $x \in \mathcal{X}$ :

$$\mathcal{G}h_\gamma(x) \leq c_1 h_\gamma(x) + c_2,$$

avec  $\gamma$  défini précédemment et  $h_\gamma(x) = |x|^\gamma = \left(\sum_{i=1}^d x_i\right)^\gamma$ , pour tout  $x \in \mathbb{R}_+^d$ .

## Hypothèse sur la dynamique du trait

Il existe  $c_1, c_2 \geq 0$  tels que pour tout  $x \in \mathcal{X}$ :

$$\mathcal{G}h_\gamma(x) \leq c_1 h_\gamma(x) + c_2,$$

avec  $\gamma$  défini précédemment et  $h_\gamma(x) = |x|^\gamma = \left(\sum_{i=1}^d x_i\right)^\gamma$ , pour tout  $x \in \mathbb{R}_+^d$ .

En particulier, si on considère :

$$\mathcal{G}f(x) = r(x, t)f'(x) + \sigma^2(x, t)f''(x),$$

l'hypothèse précédente est vérifiée si :

$$r(x, t) \leq a_r x + b_r \text{ et } \sigma(x, t) \leq a_\sigma x + b_\sigma.$$

# Plan

Échantillonnage uniforme

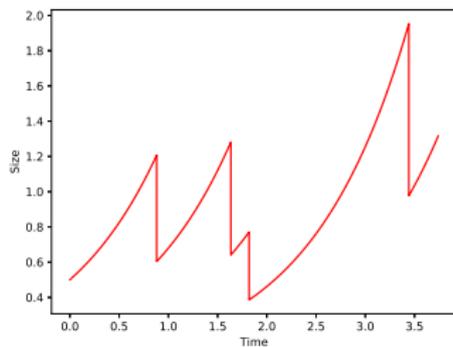
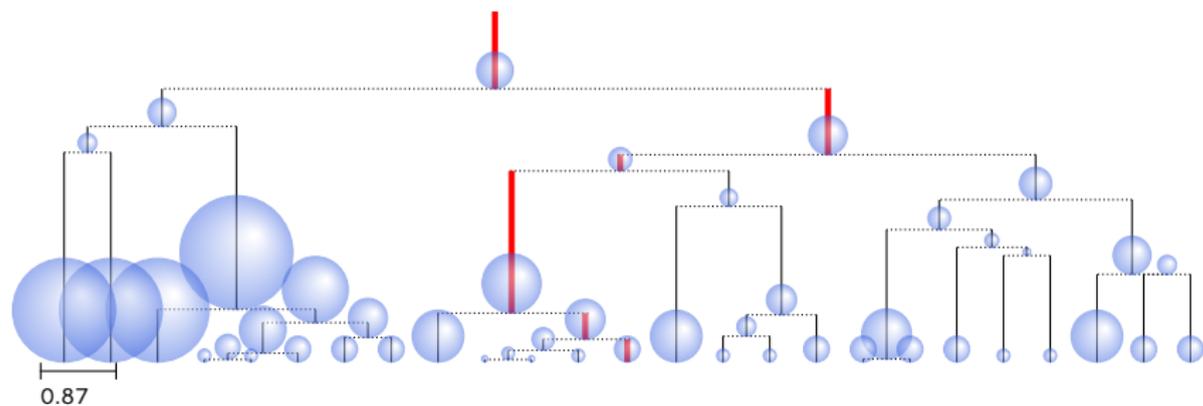
Loi des grands nombres

Estimation statistique dans des populations structurées



"Uniform sampling in a structured branching population",  
Marguet 2017, accepté pour publication dans Bernoulli.

# Trait d'un individu échantillonné uniformément au hasard



## Le trait d'un individu échantillonné

Soit  $U(t)$  la variable aléatoire correspondant au trait d'un individu échantillonné uniformément dans la population au temps  $t$ . On considère

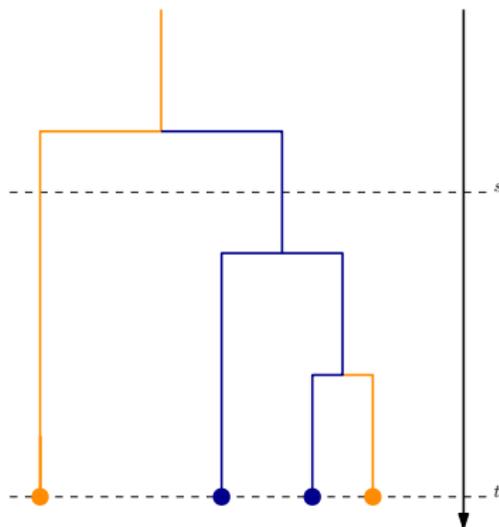
$$\left( X_s^{U(t)} \right)_{s \leq t} .$$

## Le trait d'un individu échantillonné

Soit  $U(t)$  la variable aléatoire correspondant au trait d'un individu échantillonné uniformément dans la population au temps  $t$ . On considère

$$\left( X_s^{U(t)} \right)_{s \leq t}.$$

Ce processus est différent du processus de la cellule étiquetée.



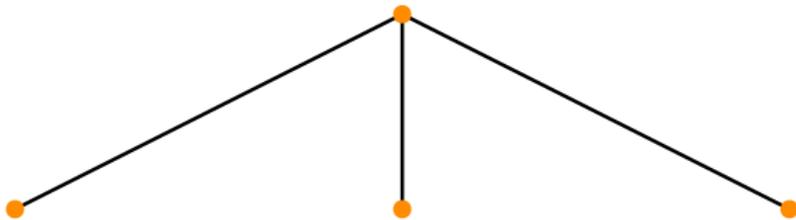
# Arbre de Galton-Watson

Soit  $(p_k, k \geq 0)$  une loi de reproduction.



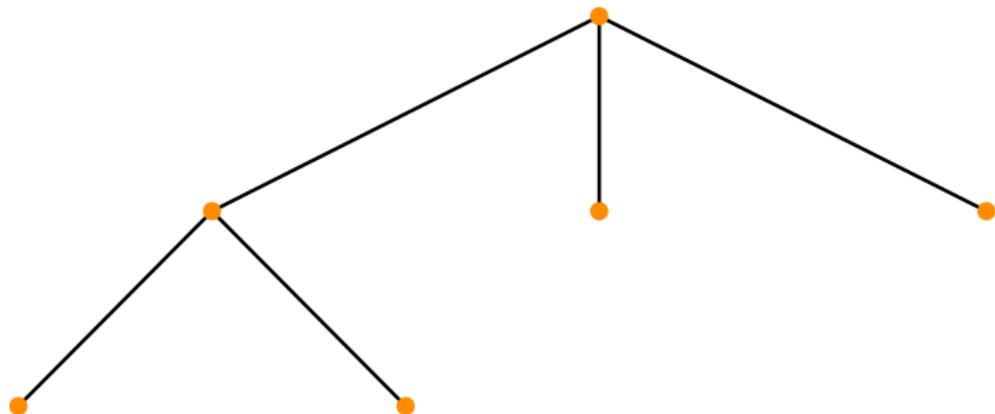
## Arbre de Galton-Watson

Soit  $(p_k, k \geq 0)$  une loi de reproduction.



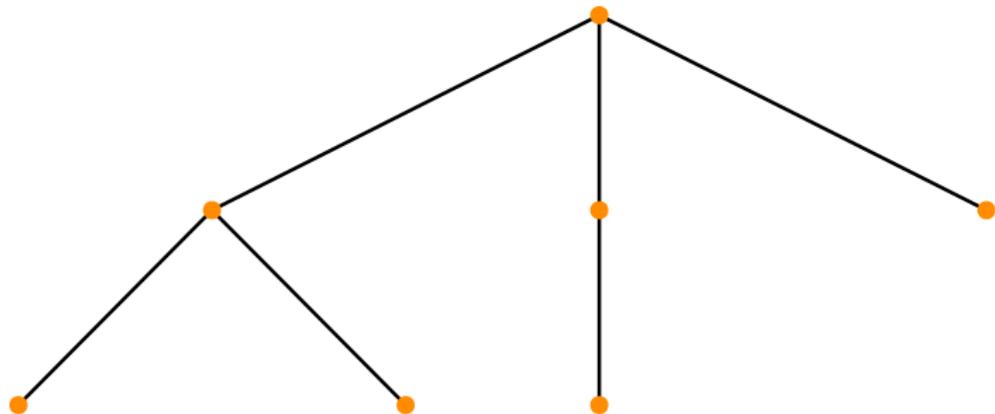
## Arbre de Galton-Watson

Soit  $(p_k, k \geq 0)$  une loi de reproduction.



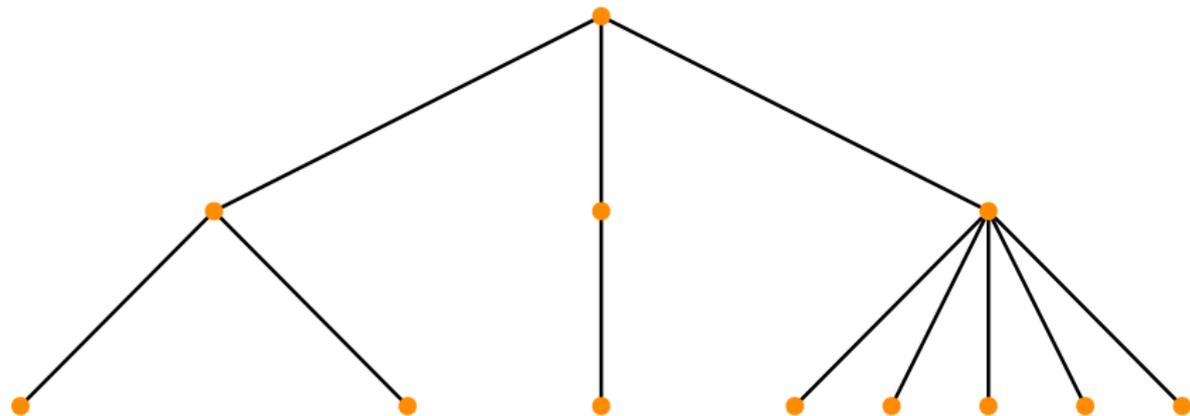
## Arbre de Galton-Watson

Soit  $(p_k, k \geq 0)$  une loi de reproduction.



## Arbre de Galton-Watson

Soit  $(p_k, k \geq 0)$  une loi de reproduction.



# Arbre de Galton-Watson biaisé

Soit:

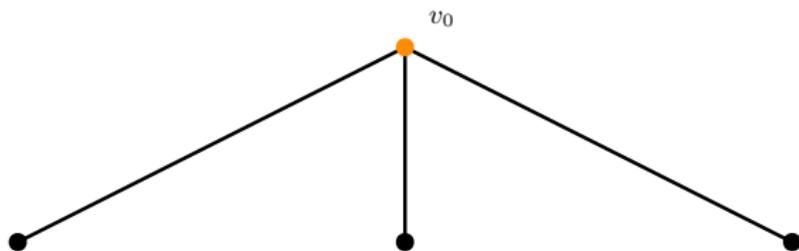
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \text{ avec } m = \sum_{k \geq 0} kp_k.$$

$v_0$   


# Arbre de Galton-Watson biaisé

Soit:

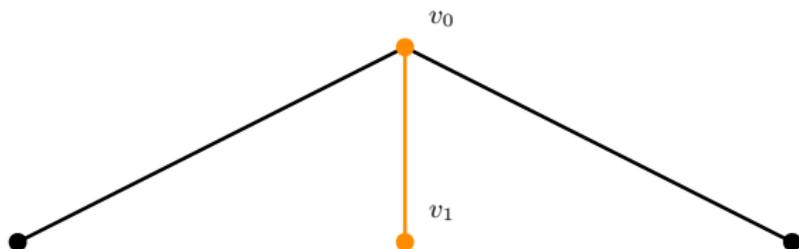
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$



# Arbre de Galton-Watson biaisé

Soit:

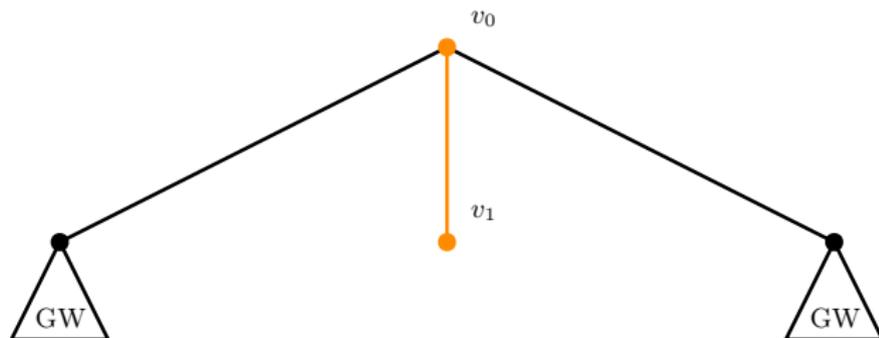
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$



# Arbre de Galton-Watson biaisé

Soit:

$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$

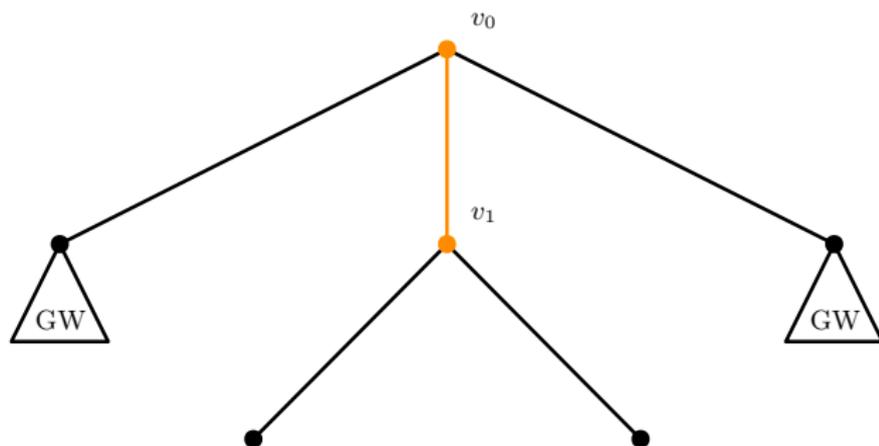


Lyons, Pemantle, Peres 1995.

# Arbre de Galton-Watson biaisé

Soit:

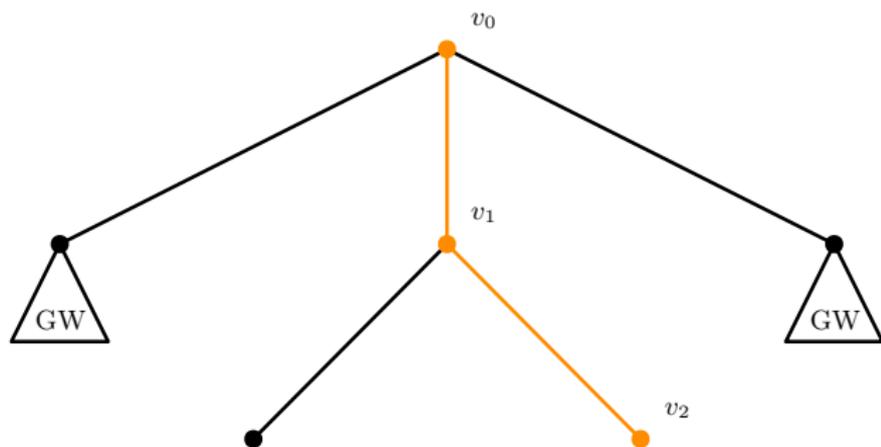
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \text{ avec } m = \sum_{k \geq 0} kp_k.$$



# Arbre de Galton-Watson biaisé

Soit:

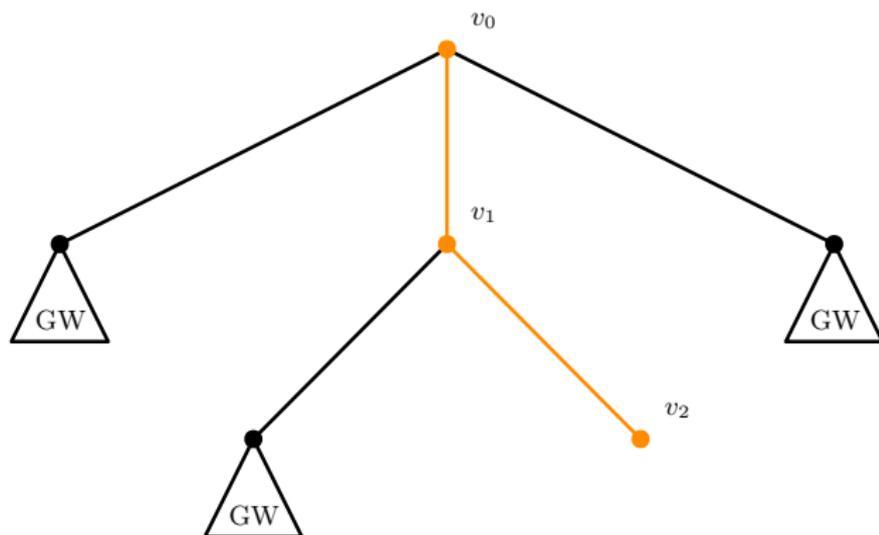
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$



# Arbre de Galton-Watson biaisé

Soit:

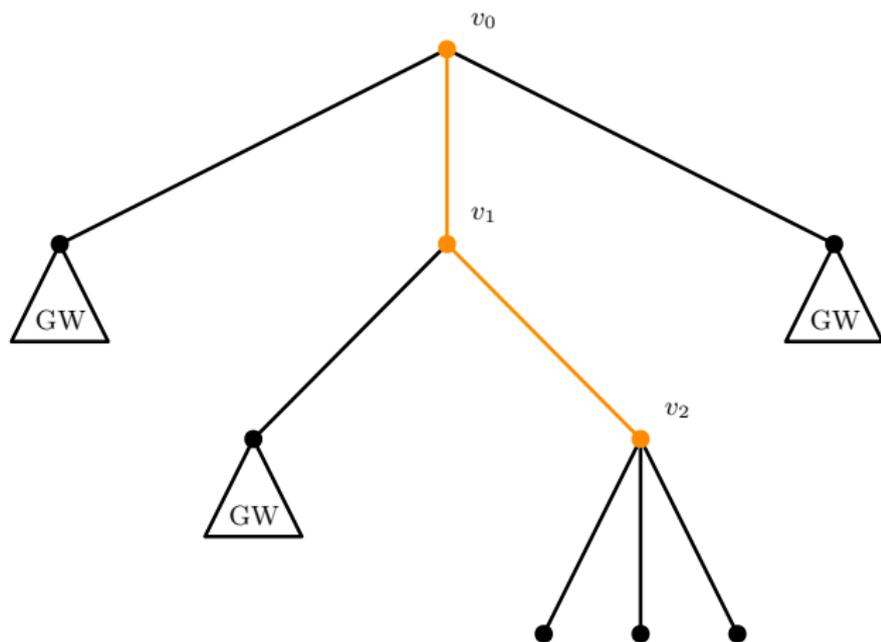
$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$



# Arbre de Galton-Watson biaisé

Soit:

$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \text{ avec } m = \sum_{k \geq 0} kp_k.$$

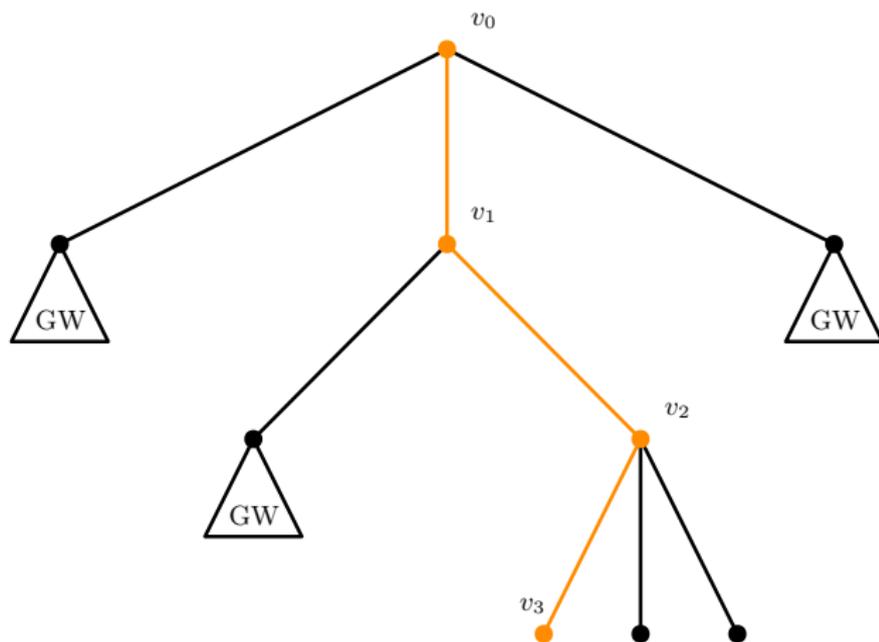


Lyons, Pemantle, Peres 1995.

# Arbre de Galton-Watson biaisé

Soit:

$$\hat{p}_k = \frac{kp_k}{m}, \quad k \geq 0, \quad \text{avec } m = \sum_{k \geq 0} kp_k.$$



Lyons, Pemantle, Peres 1995.

# Épine et arbre biaisé par la taille

La suite de labels  $(v_k, k \geq 0)$  représente l'épine.

Individu échantillonné

$$\mathbb{E}(H(\mathcal{A}_n, \mathcal{U}_n)N_n) = m^n \mathbb{E}\left(H(\widehat{\mathcal{A}}_n, \xi_n)\right).$$

avec :

- $\mathcal{A}_n$  : arbre de Galton-Watson à la génération  $n$ ,
- $\mathcal{U}_n$  : choix uniforme d'un individu à la génération  $n$ ,
- $N_n$  : nombre d'individus en vie à la génération  $n$ ,
- $\widehat{\mathcal{A}}_n$  : arbre de Galton-Watson biaisé à la génération  $n$ ,
- $\xi_n$  : épine à la génération  $n$ .



Lyons, Pemantle, Peres 1995.

## Arbre biaisé par la taille

- Chauvin, Rouault (1988), Chauvin, Rouault, Wakolbinger (1991), Gorostiza, Roelly, Wakolbinger (1991),
- Lyons, Pemantle, Peres, (1995) : théorèmes limites
- Kurtz, Lyons, Pemantle, Peres (1997), Athreya (2000) : extensions multitypes
- Georgii et Baake (2003) : temps continu, trait discret,
- Hardy et Harris (2006) : généralisation de la formalisation de l'épine,
- Bansaye, Delmas, Marsalle, Tran (2011): modèle continu dans le cas neutre, loi des grands nombres,
- Cloez (2011): extension au cas non-neutre à l'aide de la théorie spectrale.

## Semi-groupe associé au processus de l'épine

On pose  $m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x)$ .



Del Moral 2004, Bansaye 2013, Cloez 2017.

## Semi-groupe associé au processus de l'épine

On pose  $m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x)$ .

1. Pour  $f : \mathcal{X} \rightarrow \mathbb{R}$  et  $x \in \mathcal{X}$ , on considère :

$$P_{s,t}f(x) := \frac{R_{s,t}f(x)}{m(x, s, t)} = \frac{\mathbb{E} [\sum_{u \in V_t} f(X_t^u) | Z_s = \delta_x]}{m(x, s, t)}, \quad \forall 0 \leq s \leq t.$$

- ▶ opérateurs conservatifs
- ▶ propriété de semi-groupe



Del Moral 2004, Bansaye 2013, Cloez 2017.

## Semi-groupe associé au processus de l'épine

On pose  $m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x)$ .

1. Pour  $f : \mathcal{X} \rightarrow \mathbb{R}$  et  $x \in \mathcal{X}$ , on considère :

$$P_{s,t}f(x) := \frac{R_{s,t}f(x)}{m(x, s, t)} = \frac{\mathbb{E} [\sum_{u \in V_t} f(X_t^u) | Z_s = \delta_x]}{m(x, s, t)}, \quad \forall 0 \leq s \leq t.$$

- ▶ opérateurs conservatifs
- ▶ propriété de semi-groupe

2. On considère :

$$P_{r,s}^{(t)}f(x) = \frac{R_{r,s}(fm(\cdot, s, t))}{m(x, r, t)}, \quad \text{avec } 0 \leq r \leq s \leq t.$$

- ▶ opérateurs conservatifs
- ▶ propriété de semi-groupe



Del Moral 2004, Bansaye 2013, Cloez 2017.

# Formule Many-to-One

## Théorème

Pour tout  $t > 0$ ,  $x \in \mathcal{X}$ , pour toute fonction mesurable positive  $F : \mathbb{D}([0, t], \mathcal{X}) \rightarrow \mathbb{R}$  on a :

$$\mathbb{E}_{\delta_{x_0}} \left[ \sum_{u \in V_t} F(X_s^u, s \leq t) \right] = \mathbb{E}_{\delta_{x_0}}(N_t) \mathbb{E}_{x_0} \left[ F(Y_s^{(t)}, s \leq t) \right],$$

avec  $(Y_s^{(t)})_{s \leq t}$  processus de Markov inhomogène en temps de générateurs infinitésimaux  $(\mathcal{A}_s^{(t)})_{s \leq t}$  donnés par :

$$\mathcal{A}_s^{(t)} f(x) = \widehat{\mathcal{G}}_{s,t} f(x) + \widehat{B}_{s,t}(x) \left( \int_{\mathcal{X}} (f(y) - f(x)) \widehat{P}_{s,t}(x, dy) \right).$$

## Le processus auxiliaire

$$m(x, dy) = \sum_{k \geq 0} p_k(x) \sum_{j=1}^k P_i^{(k)}(x, dy), \quad m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x).$$

## Le processus auxiliaire

$$m(x, dy) = \sum_{k \geq 0} p_k(x) \sum_{j=1}^k P_i^{(k)}(x, dy), \quad m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x).$$

- ▶ le processus saute à un taux accéléré :

$$\hat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

## Le processus auxiliaire

$$m(x, dy) = \sum_{k \geq 0} p_k(x) \sum_{j=1}^k P_i^{(k)}(x, dy), \quad m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x).$$

- ▶ le processus saute à un taux accéléré :

$$\widehat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

- ▶ le trait de l'unique descendant est choisi selon le noyau de probabilité :

$$\widehat{P}_{s,t}(x, dy) = \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \left( \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \right)^{-1}.$$

## Le processus auxiliaire

$$m(x, dy) = \sum_{k \geq 0} p_k(x) \sum_{j=1}^k P_i^{(k)}(x, dy), \quad m(x, s, t) = \mathbb{E}(N_t | Z_s = \delta_x).$$

- ▶ le processus saute à un taux accéléré :

$$\hat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

- ▶ le trait de l'unique descendant est choisi selon le noyau de probabilité :

$$\hat{P}_{s,t}(x, dy) = \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \left( \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \right)^{-1}.$$

- ▶ la dynamique du trait est donnée par :

$$\hat{G}_{s,t}f(x) = \frac{\mathcal{G}(m(\cdot, s, t)f)(x) - f(x)\mathcal{G}(m(\cdot, s, t))(x)}{m(x, s, t)}$$

## Exemples

- ▶ Dans le cas neutre,  $B \equiv b$  :

$$\mathcal{A}f(x) = \mathcal{G}f(x) + 2b \int_{\mathcal{X}} (f(y) - f(x)) \frac{m(x, dy)}{2}$$

## Exemples

- ▶ Dans le cas neutre,  $B \equiv b$  :

$$\mathcal{A}f(x) = \mathcal{G}f(x) + 2b \int_x (f(y) - f(x)) \frac{m(x, dy)}{2}$$

- ▶ Dans le cas d'une croissance des cellules exponentielle à taux  $a$  et d'une division binaire à taux  $B(x) = x$ , on a :

$$m(x, s, t) = 1 + \frac{x}{a} \left( e^{a(t-s)} - 1 \right),$$

et :

$$\mathcal{A}_s^{(t)} f(x) = axf'(x) + x \left( 1 + \frac{1}{1 + \frac{x}{a} (e^{a(t-s)} - 1)} \right) \left( f\left(\frac{x}{2}\right) - f(x) \right).$$

## Approximation en grande population

Soit  $\nu \in \mathcal{M}_P(\mathcal{X})$  vérifiant :

$$\mathbb{P}_\nu(N_t > 0) > 0.$$

On note  $U(t)$  une v.a. de loi uniforme sur  $V_t$  conditionnellement à la non-extinction. On pose

$$\nu_n := \sum_{i=1}^n \delta_{X_i}, \quad X_i \stackrel{\text{i.i.d.}}{\sim} \nu \quad \forall i = 1, \dots, n.$$

### Théorème

Sous les hypothèses précédentes + hypothèses techniques, on a en loi dans  $\mathbb{D}([0, t], \mathcal{X})$ :

$$X_{[0,t]}^{U(t), \nu_n} \xrightarrow{n \rightarrow +\infty} Y_{[0,t]}^{(t), \pi_t}, \quad \text{avec } \pi_t(dx) = \frac{m(x, 0, t)\nu(dx)}{m(\nu, 0, t)}.$$

# Plan

Échantillonnage uniforme

Loi des grands nombres

Estimation statistique dans des populations structurées



" A law of large numbers for branching Markov processes by the ergodicity of ancestral lineages", Marguet 2017, pré-publication.

## Ergodicité du processus auxiliaire

1. Il existe une fonction  $V : \mathcal{X} \rightarrow \mathbb{R}_+$  et  $c, d > 0$  tels que

$$\mathcal{A}_s^{(t)} V(x) \leq -cV(x) + d, \quad \forall x \in \mathcal{X}, 0 \leq s \leq t.$$

2. Pour tout  $0 < r < s$ , il existe  $\alpha_{s-r} \in (0, 1)$ ,  $\nu_{r,s} \in \mathcal{M}_P(\mathcal{X})$  t.q.

$$\inf_{x \in \{x \in \mathcal{X} : V(x) \leq R\}} P_{r,s}^{(t)}(x, \cdot) \geq \alpha_{s-r} \nu_{r,s}(\cdot), \quad \forall t \geq s.$$

avec pour un certain  $R > \frac{2d}{c}$ .

## Ergodicité du processus auxiliaire

1. Il existe une fonction  $V : \mathcal{X} \rightarrow \mathbb{R}_+$  et  $c, d > 0$  tels que

$$\mathcal{A}_s^{(t)} V(x) \leq -cV(x) + d, \quad \forall x \in \mathcal{X}, 0 \leq s \leq t.$$

2. Pour tout  $0 < r < s$ , il existe  $\alpha_{s-r} \in (0, 1)$ ,  $\nu_{r,s} \in \mathcal{M}_P(\mathcal{X})$  t.q.

$$\inf_{x \in \{x \in \mathcal{X} : V(x) \leq R\}} P_{r,s}^{(t)}(x, \cdot) \geq \alpha_{s-r} \nu_{r,s}(\cdot), \quad \forall t \geq s.$$

avec pour un certain  $R > \frac{2d}{c}$ .

### Proposition

Sous les hypothèses précédentes, il existe  $\bar{c} > 0$  tel que pour tout  $x, y \in \mathcal{X}$ , pour toute fonction mesurable bornée  $f : \mathcal{X} \rightarrow \mathbb{R}$  et tout  $0 \leq r \leq s \leq t$ ,

$$\left| P_{r,s}^{(t)} f(x) - P_{r,s}^{(t)} f(y) \right| \leq C e^{-\bar{c}(s-r)} d(x, y) \|f\|_\infty,$$

où  $d$  est une distance sur  $\mathcal{X}$ ,  $C$  est une constante positive.

## Comportement en temps long

### Théorème

Sous les hypothèses précédentes + hypothèses techniques, pour toute fonction mesurable bornée  $f : \mathcal{X} \rightarrow \mathbb{R}$ , et tout  $x_0, x_1 \in \mathcal{X}$ , on a :

$$\mathbb{E}_{\delta_{x_0}} \left[ \left( \frac{\sum_{u \in V_t} \left( f(X_t^u) - \mathbb{E}_{x_1} [f(Y_t^{(t)})] \right)}{m(x_0, 0, t)} \right)^2 \right] \xrightarrow[t \rightarrow +\infty]{} 0.$$

De plus, la vitesse de convergence est plus grande que :

$$v(t) = \exp \left( \min \left( \bar{c}, \frac{c(x_0) - \alpha_1}{2} \right) t \right).$$

## Hypothèse

Pour tout  $x \in \mathcal{X}$ ,

$$\sup_{t \geq 0} \mathbb{E}_{\delta_x} \left( \left( \frac{N_t}{m(x, 0, t)} \right)^2 \right) < \infty.$$

## Corollaire

Sous les hypothèses précédentes + hypothèses techniques, pour toute fonction mesurable bornées  $f : \mathcal{X} \rightarrow \mathbb{R}$ , pour tout  $x_0, x_1 \in \mathcal{X}$ , on a :

$$\mathbb{E}_{x_0} \left[ \left( \frac{\sum_{u \in V_t} f(X_t^u)}{N_t} - \mathbb{E}_{x_1} [f(Y_t^{(t)})] \right)^2 \right] \xrightarrow[t \rightarrow +\infty]{} 0.$$

## Hypothèse

Pour tout  $x \in \mathcal{X}$ ,

$$\sup_{t \geq 0} \mathbb{E}_{\delta_x} \left( \left( \frac{N_t}{m(x, 0, t)} \right)^2 \right) < \infty.$$

## Corollaire

Sous les hypothèses précédentes + hypothèses techniques, pour toute fonction mesurable bornée  $f : \mathcal{X} \rightarrow \mathbb{R}$ , pour tout  $x_0, x_1 \in \mathcal{X}$ , on a :

$$\mathbb{E}_{x_0} \left[ \left( \frac{\sum_{u \in V_t} f(X_t^u)}{N_t} - \mathbb{E}_{x_1} [f(Y_t^{(t)})] \right)^2 \right] \xrightarrow[t \rightarrow +\infty]{} 0.$$

- Convergence pour les lignées ancestrales.

## Hypothèse

Pour tout  $x \in \mathcal{X}$ ,

$$\sup_{t \geq 0} \mathbb{E}_{\delta_x} \left( \left( \frac{N_t}{m(x, 0, t)} \right)^2 \right) < \infty.$$

## Corollaire

Sous les hypothèses précédentes + hypothèses techniques, pour toute fonction mesurable bornée  $f : \mathcal{X} \rightarrow \mathbb{R}$ , pour tout  $x_0, x_1 \in \mathcal{X}$ , on a :

$$\mathbb{E}_{x_0} \left[ \left( \frac{\sum_{u \in V_t} f(X_t^u)}{N_t} - \mathbb{E}_{x_1} [f(Y_t^{(t)})] \right)^2 \right] \xrightarrow{t \rightarrow +\infty} 0.$$

- ▶ Convergence pour les lignées ancestrales.
- ▶ Hypothèses vérifiées pour de la croissance-fragmentation inhomogène.

Échantillonnage uniforme

Loi des grands nombres

Estimation statistique dans des populations structurées

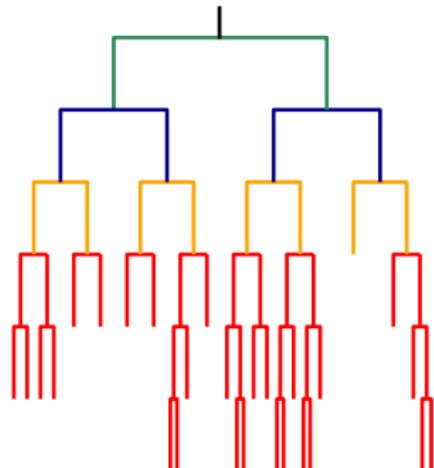
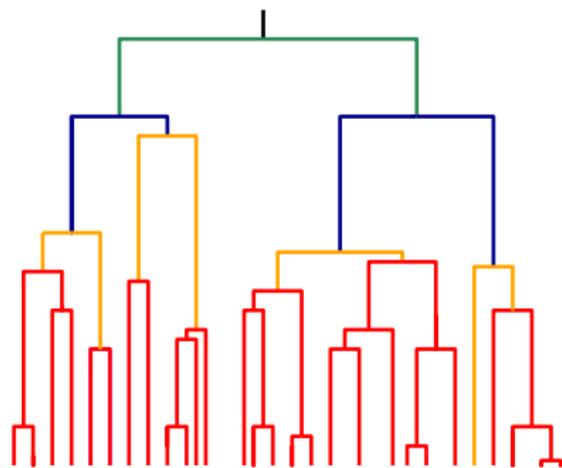
Estimation non-paramétrique

Estimation paramétrique du taux de division



Travail en collaboration avec Marc Hoffmann.

# Temps généalogique



 Thèse A. Olivier 2015.

## Description du modèle

- ▶ Le trait  $(\phi_x(t))_{t \geq 0}$  de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où  $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$  sont des fonctions mesurables et  $(W_t)_{t \geq 0}$  est un mouvement brownien standard.

## Description du modèle

- ▶ Le trait  $(\phi_x(t))_{t \geq 0}$  de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où  $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$  sont des fonctions mesurables et  $(W_t)_{t \geq 0}$  est un mouvement brownien standard.

- ▶ Chaque individu se divise au temps  $t$  à un taux  $B(\phi_x(t))$ ,

## Description du modèle

- ▶ Le trait  $(\phi_x(t))_{t \geq 0}$  de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où  $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$  sont des fonctions mesurables et  $(W_t)_{t \geq 0}$  est un mouvement brownien standard.

- ▶ Chaque individu se divise au temps  $t$  à un taux  $B(\phi_x(t))$ ,
- ▶ À sa mort, un individu de trait  $x$  est remplacé par 2 descendants de traits à la naissance  $\theta x$  et  $(1 - \theta)x$ , où  $\theta \in [0, 1]$  est une v.a. de densité associée  $\kappa(\theta)$  symétrique.

## Description du modèle

- ▶ Le trait  $(\phi_x(t))_{t \geq 0}$  de chaque individu évolue suivant une diffusion d'équation :

$$d\phi_x(t) = r(\phi_x(t))dt + \sigma(\phi_x(t))dW_t, \quad \phi_x(0) = x,$$

où  $r, \sigma : \mathbb{R} \rightarrow \mathbb{R}$  sont des fonctions mesurables et  $(W_t)_{t \geq 0}$  est un mouvement brownien standard.

- ▶ Chaque individu se divise au temps  $t$  à un taux  $B(\phi_x(t))$ ,
- ▶ À sa mort, un individu de trait  $x$  est remplacé par 2 descendants de traits à la naissance  $\theta x$  et  $(1 - \theta)x$ , où  $\theta \in [0, 1]$  est une v.a. de densité associée  $\kappa(\theta)$  symétrique.
- ▶ Conditionnellement au trait de leur ancêtre, les descendants évoluent indépendamment les uns des autres.

## Chaîne de Markov bifurcante

On considère les notations de Ulam-Harris-Neveu, pour  $n, m \geq 0$ , on note :

$$\mathbb{T} = \bigcup_{m \in \mathbb{N}} \{0, 1\}^m, \quad \mathbb{T}_n = \bigcup_{m=0}^n \{0, 1\}^m, \quad \mathbb{T}_n^* = \mathbb{T}_n \setminus \{\emptyset\}.$$

On note  $X_u$  le trait à la naissance d'un individu  $u \in \mathbb{T}$ . On suppose que l'on dispose des observations

$$\mathbb{X}^n = (X_u)_{u \in \mathbb{T}_n}.$$

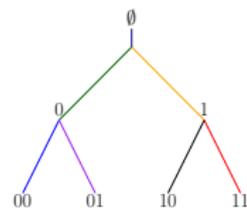
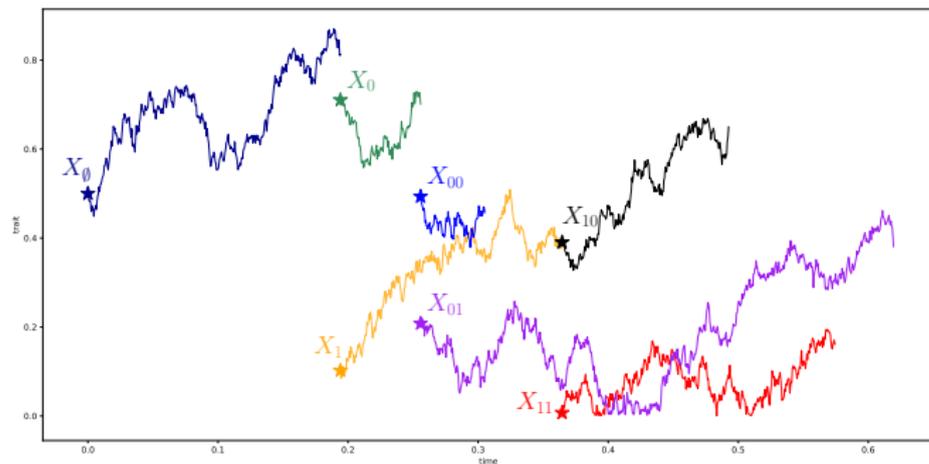
Le processus  $(X_u, u \in \mathbb{T})$  est alors une chaîne de Markov bifurcante de noyau de transition  $\mathcal{P}$  de  $\mathbb{R}$  dans  $\mathbb{R} \times \mathbb{R}$  tel que :

$$\mathbb{E} \left[ \prod_{u \in \mathbb{G}_m} \psi_u(X_u, X_{u0}, X_{u1}) \mid \mathcal{F}_m \right] = \prod_{u \in \mathbb{G}_m} \mathcal{P} \psi_u(X_u),$$

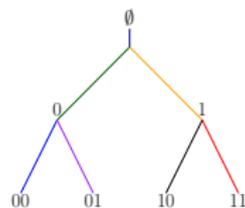
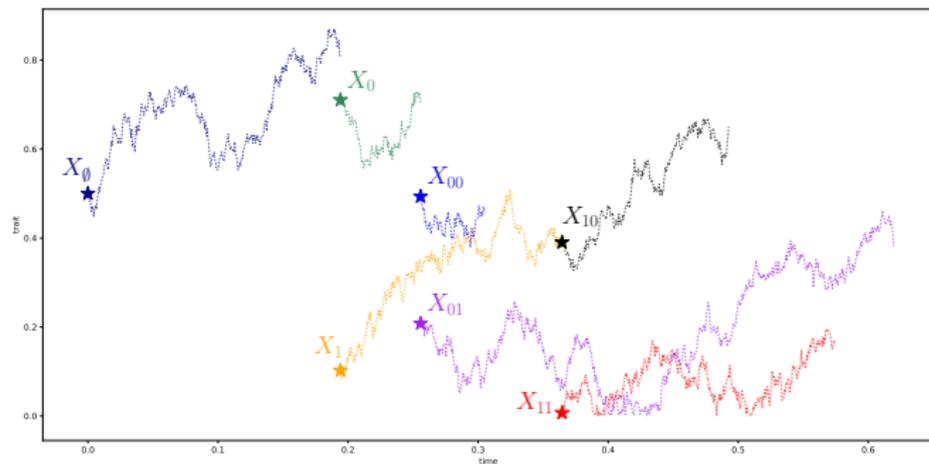
pour tout  $m \geq 0$ .

 Guyon 2007.

# Exemple de trajectoire

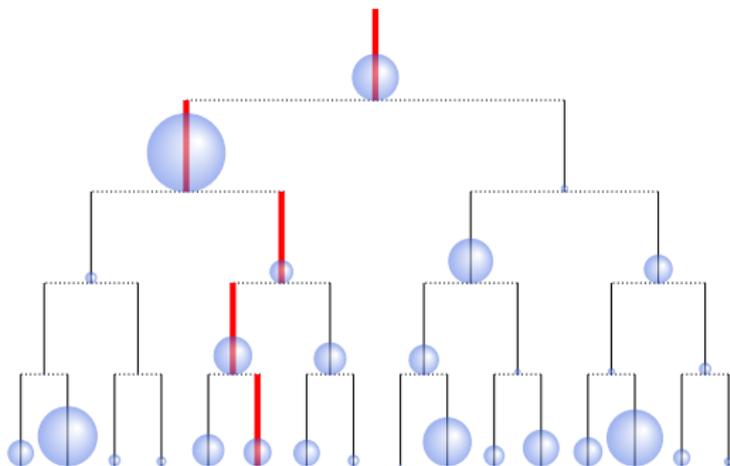


# Exemple de trajectoire



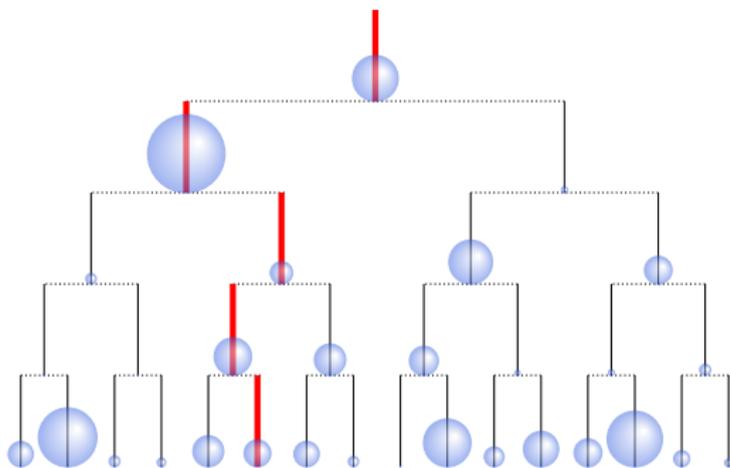
## Processus de la cellule étiqueté

Soit  $Y$  la chaîne de Markov correspondant au trait de la cellule étiquetée.



## Processus de la cellule étiqueté

Soit  $Y$  la chaîne de Markov correspondant au trait de la cellule étiquetée.



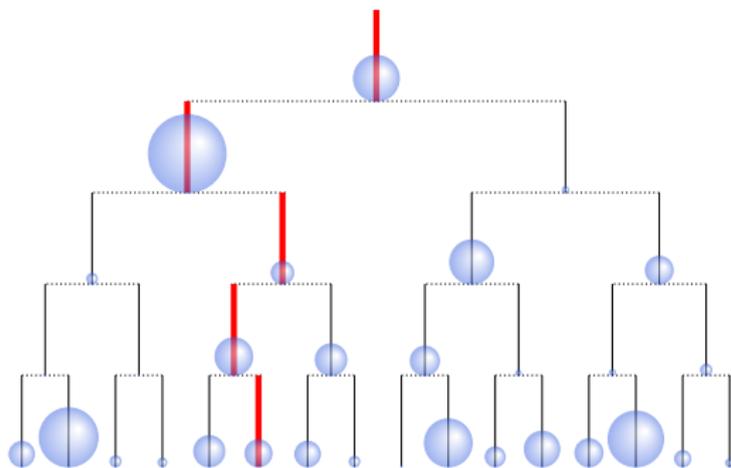
La transition de  $Y$  est donnée par  $\mathcal{Q} = (\mathcal{P}_1 + \mathcal{P}_2)/2$ , avec

$$\mathcal{P}_1(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy_2) = \mathcal{P}_2(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy),$$

car  $\kappa$  est symétrique.

## Processus de la cellule étiqueté

Soit  $Y$  la chaîne de Markov correspondant au trait de la cellule étiquetée.



La transition de  $Y$  est donnée par  $\mathcal{Q} = \mathcal{P}_1$ , avec

$$\mathcal{P}_1(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy_2) = \mathcal{P}_2(x, dy) = \int_{\mathbb{R}} \mathcal{P}(x, dy_1 dy),$$

car  $\kappa$  est symétrique.

## Objectifs

À partir de  $(X_u, u \in \mathbb{T}_n)$ ,

- ▶ estimer  $\nu(dx)$  (mesure invariante),
- ▶ estimer  $Q(dx)$ ,
- ▶ estimer  $B$ .

## Comportement asymptotique de la mesure empirique

$$\mathcal{M}_n(\psi) = \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \psi(X_{u^-}, X_u).$$

## Comportement asymptotique de la mesure empirique

$$\mathcal{M}_n(\psi) = \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \psi(X_{u^-}, X_u).$$

### Théorème

Soit  $\mu \in \mathcal{M}_P(\mathbb{R})$  telle que  $\mu(V^2) < \infty$  et  $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  bornée telle que  $\psi_*$  soit à support compact. Sous des hypothèses portant sur l'ergodicité de  $Y$ ,

$$\mathbb{E}_\mu [(\mathcal{M}_n(\psi) - \nu(Q\psi))^2] \lesssim |\mathbb{T}_n|^{-1} C(\psi), \quad \forall n \in \mathbb{N},$$

où  $\lesssim$  signifie à une constante près dépendant de  $Q$  et  $\text{supp}(\psi_*)$  et

$$C(\psi) = |\psi^2|_{\mu+\nu} + |\psi^* \psi|_\mu + \left(1 + \mu(V^2)^{1/2}\right) |\psi_*|_1 |\psi|_\nu.$$

# Comportement asymptotique de la mesure empirique

$$\mathcal{M}_n(\psi) = \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \psi(X_{u-}, X_u).$$

## Théorème

Soit  $\mu \in \mathcal{M}_P(\mathbb{R})$  telle que  $\mu(V^2) < \infty$  et  $\psi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  bornée telle que  $\psi_*$  soit à support compact. Sous des hypothèses portant sur l'ergodicité de  $Y$ ,

$$\mathbb{E}_\mu [(\mathcal{M}_n(\psi) - \nu(Q\psi))^2] \lesssim |\mathbb{T}_n|^{-1} C(\psi), \quad \forall n \in \mathbb{N},$$

où  $\lesssim$  signifie à une constante près dépendant de  $Q$  et  $\text{supp}(\psi_*)$  et

$$C(\psi) = |\psi^2|_{\mu+\nu} + |\psi^* \psi|_\mu + \left(1 + \mu(V^2)^{1/2}\right) |\psi_*|_1 |\psi|_\nu.$$

$$|\psi|_\mu = \int_{\mathbb{R} \times \mathbb{R}} |\psi(x, y)| \mu(dx) dy + |\psi_*|_1 \wedge |\psi|_{\wedge 1}, \quad \psi^*(x) = \sup_{y \in \mathbb{R}} |\psi(x, y)|$$

$$|\psi|_{\wedge 1} = \int_{\mathcal{X} \times \mathcal{X}} |\psi(x, y)| dx dy \wedge |\psi_*(y)|_1, \quad \psi_*(y) = \sup_{x \in \mathbb{R}} |\psi(x, y)|.$$

# Plan

Échantillonnage uniforme

Loi des grands nombres

Estimation statistique dans des populations structurées

Estimation non-paramétrique

Estimation paramétrique du taux de division

# Noyau

## Définition

Une fonction  $G : \mathbb{R} \rightarrow \mathbb{R}$  est un noyau d'ordre  $k$  si elle satisfait  $\int_{\mathbb{R}} x^l G(x) dx = \mathbf{1}_{\{l=0\}}$  pour tout  $l = 0, \dots, k$ .

Exemples :

- $x \rightarrow \mathbf{1}_{\{x \leq 1\}}$  est un noyau d'ordre 0,
- $x \rightarrow 1/\sqrt{2\pi} e^{-x^2/2}$  est un noyau d'ordre 1,
- on peut construire des noyaux de tout ordre à l'aide des polynômes de Legendre.

## Estimateur à noyau de la mesure invariante

Soit  $G$  un noyau d'ordre  $k$  et  $h > 0$  un paramètre de lissage. On pose

$$G_h(y) = h^{-1} G(h^{-1}y), \quad \forall y \in \mathbb{R}.$$

Comme  $\mathcal{M}_n(\varphi)$  converge vers  $\nu(\varphi)$  (pour  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ ), on a

$$\mathcal{M}_n(G_h(\cdot - x_0)) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} G_h(x - x_0) \nu(x) dx.$$

## Estimateur à noyau de la mesure invariante

Soit  $G$  un noyau d'ordre  $k$  et  $h > 0$  un paramètre de lissage. On pose

$$G_h(y) = h^{-1} G(h^{-1}y), \quad \forall y \in \mathbb{R}.$$

Comme  $\mathcal{M}_n(\varphi)$  converge vers  $\nu(\varphi)$  (pour  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ ), on a

$$\mathcal{M}_n(G_h(\cdot - x_0)) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}} G_h(x - x_0) \nu(x) dx.$$

Un estimateur de  $\nu(x_0)$ , pour  $x_0 \in \mathbb{R}$  est donc donné par

$$\hat{\nu}_n(x_0) = \mathcal{M}_n(G_h(\cdot - x_0)).$$

## Estimateur à noyau de la densité

De la même façon, on considère  $G$  un noyau d'ordre  $k$  et  $h_1, h_2 > 0$  deux paramètres de lissage. On pose

$$G_{h_1, h_2}^{\otimes 2}(x, y) = h_1^{-1} h_2^{-1} G(h_1^{-1} x) G(h_2^{-1} y), \quad \forall x, y \in \mathbb{R}.$$

On a alors la convergence suivante :

$$\begin{aligned} & \mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0)) \\ & \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R} \times \mathbb{R}} G_{h_1, h_2}^{\otimes 2}(x - x_0, y - y_0) q(x, y) \nu(x) dy dx. \end{aligned}$$

## Estimateur à noyau de la densité

De la même façon, on considère  $G$  un noyau d'ordre  $k$  et  $h_1, h_2 > 0$  deux paramètres de lissage. On pose

$$G_{h_1, h_2}^{\otimes 2}(x, y) = h_1^{-1} h_2^{-1} G(h_1^{-1} x) G(h_2^{-1} y), \quad \forall x, y \in \mathbb{R}.$$

On a alors la convergence suivante :

$$\begin{aligned} & \mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0)) \\ & \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R} \times \mathbb{R}} G_{h_1, h_2}^{\otimes 2}(x - x_0, y - y_0) q(x, y) \nu(x) dy dx. \end{aligned}$$

Un estimateur de la densité  $q(x_0, y_0)$ , pour  $x_0, y_0 \in \mathbb{R}$  est alors donné par

$$\hat{q}_n(x_0, y_0) = \frac{\mathcal{M}_n(G_{h_1, h_2}^{\otimes 2}(\cdot - x_0, \cdot - y_0))}{\mathcal{M}_n(G_h(\cdot - x_0)) \vee \varpi}.$$

## Convergence des estimateurs

Soit  $\alpha, \beta > 0$  et  $G$  un noyau d'ordre  $k > \max(\alpha, \beta)$ . On pose

$$h = |\mathbb{T}_n|^{\frac{-1}{2\beta+1}}, \quad h_1 = |\mathbb{T}_n|^{\frac{-s(\alpha,\beta)}{(\alpha \wedge \beta)(2s(\alpha,\beta)+1)}}, \quad h_2 = |\mathbb{T}_n|^{\frac{-s(\alpha,\beta)}{\beta(2s(\alpha,\beta)+1)}}, \quad \varpi_n \rightarrow 0.$$

### Théorème

Sous des hypothèses assurant l'ergodicité de la chaîne  $Y$ , on a

$$\begin{aligned} \left( \mathbb{E} [(\hat{\nu}_n(x_0) - \nu(x_0))^2] \right)^{1/2} &\lesssim |\mathbb{T}_n|^{-\beta/(2\beta+1)}, \\ \left( \mathbb{E} [(\hat{q}_n(x_0, y_0) - q(x_0, y_0))^2] \right)^{1/2} &\lesssim \varpi_n^{-1} |\mathbb{T}_n|^{-s(\alpha,\beta)/(2s(\alpha,\beta)+1)}, \end{aligned}$$

avec  $s(\alpha, \beta)^{-1} = (\alpha \wedge \beta)^{-1} + \beta^{-1}$ , uniformément en  $\mathcal{Q}$  pour  $\mathcal{Q}$  dans une certaine classe de régularité de Hölder dépendant de  $\alpha$  et  $\beta$ .

# Plan

Échantillonnage uniforme

Loi des grands nombres

Estimation statistique dans des populations structurées

Estimation non-paramétrique

Estimation paramétrique du taux de division

## Estimation du taux de division

La dépendance en  $B$  de la transition est complexe mais explicite :

$$q(x, y) = \int_0^1 \frac{\kappa(z)}{z} B(y/z) \sigma(y/z)^{-2} \mathbb{E} \left[ \int_0^\infty e^{-\int_0^t B(\phi_x(s)) ds} dL_t^{y/z}(\phi_x) \right] dz,$$

avec  $L_t^y(\phi_x)$  le temps local au temps  $t$  au point  $y$  de  $(\phi_x(t), t \geq 0)$ .

## Estimation du taux de division

La dépendance en  $B$  de la transition est complexe mais explicite :

$$q(x, y) = \int_0^1 \frac{\kappa(z)}{z} B(y/z) \sigma(y/z)^{-2} \mathbb{E} \left[ \int_0^\infty e^{-\int_0^t B(\phi_x(s)) ds} dL_t^{y/z}(\phi_x) \right] dz,$$

avec  $L_t^y(\phi_x)$  le temps local au temps  $t$  au point  $y$  de  $(\phi_x(t), t \geq 0)$ .

→ Contraste de vraisemblance dans un cadre paramétrique avec  $(r, \sigma, \kappa)$  connus.

## Modèle paramétrique

On suppose que le taux de division  $B$  appartient à une classe

$$\mathcal{B} = \{B : [0, L] \rightarrow \mathbb{R}, B(x) = B_0(\vartheta, x), x \in [0, L], \vartheta \in \Theta\},$$

où  $x \mapsto B_0(x, \vartheta)$  est connu à un paramètre  $\vartheta \in \Theta$  près, et  $\Theta \subset \mathbb{R}^d$  est compact.

But

Estimer  $\vartheta$  à partir de  $(X_u, u \in \mathbb{T}_n)$ .

## Contraste de vraisemblance

Un contraste de vraisemblance est donné par :

$$\mathcal{L}_n(\vartheta, (X_u, u \in \mathbb{T}_n)) = \prod_{u \in \mathbb{T}_n^*} q_{\vartheta}(X_{u^-}, X_u),$$

où  $X_{u^-}$  correspond au trait de l'ancêtre de  $u$ . On considère alors l'estimateur de  $\vartheta$  donné par :

$$\hat{\vartheta}_n \in \operatorname{argmax}_{\vartheta \in \Theta} \left\{ \frac{1}{|\mathbb{T}_n^*|} \sum_{u \in \mathbb{T}_n^*} \log(q_{\vartheta}(X_{u^-}, X_u)) \right\}.$$

# Consistance

## Théorème

Sous des hypothèses de régularité du taux de division,  $\hat{\vartheta}_n$  converge en probabilité vers  $\vartheta$  lorsque  $n$  tend vers l'infini.

## Normalité asymptotique

On définit  $\Psi(\vartheta)$  la matrice d'information de Fisher dont les coefficients sont donnés pour tout  $1 \leq i, j \leq d$  par :

$$\Psi(\vartheta)_{i,j} = \nu_{\vartheta} \mathcal{Q}(\vartheta) \left( \frac{\partial_{\vartheta_i} \mathbf{q}_{\vartheta} \partial_{\vartheta_j} \mathbf{q}_{\vartheta}}{q_{\vartheta}^2} \right).$$

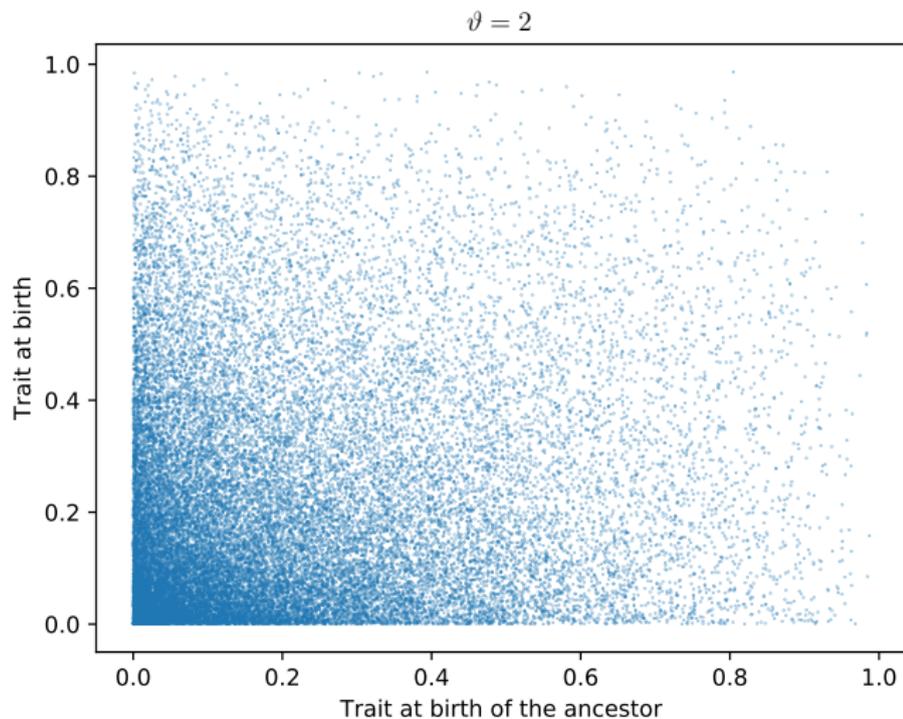
### Théorème

Sous des hypothèses de régularité et si  $\Psi(\vartheta)$  est inversible, pour tout  $\vartheta$  dans l'intérieur de  $\Theta$ , on a :

$$\mathbb{T}_n^{1/2} \left( \widehat{\vartheta}_n - \vartheta \right) \rightarrow \mathcal{N} \left( 0, \Psi(\vartheta)^{-1} \right),$$

en loi lorsque  $n$  tend vers l'infini, où  $\mathcal{N} \left( 0, \Psi(\vartheta)^{-1} \right)$  désigne la loi normale  $d$ -dimensionnelle de moyenne 0 et de matrice de covariance l'inverse de la matrice de Fisher  $\Psi(\vartheta)$ .

# Simulations



**Figure:** Données simulées avec  $B(x, \vartheta) \equiv \vartheta$ ,  $\phi_x(t) = W_t$ .

# Protocole

Pour chaque  $v$  dans une grille fixée,

# Protocole

Pour chaque  $\vartheta$  dans une grille fixée,

- ▶ simulation de données avec le paramètre  $\vartheta$ ,

# Protocole

Pour chaque  $\vartheta$  dans une grille fixée,

- ▶ simulation de données avec le paramètre  $\vartheta$ ,
- ▶ estimation non-paramétrique de  $q_{\vartheta}$  à partir de ces données sur une grille de taille  $200 \times 200$ ,

# Protocole

Pour chaque  $\vartheta$  dans une grille fixée,

- ▶ simulation de données avec le paramètre  $\vartheta$ ,
- ▶ estimation non-paramétrique de  $q_{\vartheta}$  à partir de ces données sur une grille de taille  $200 \times 200$ ,

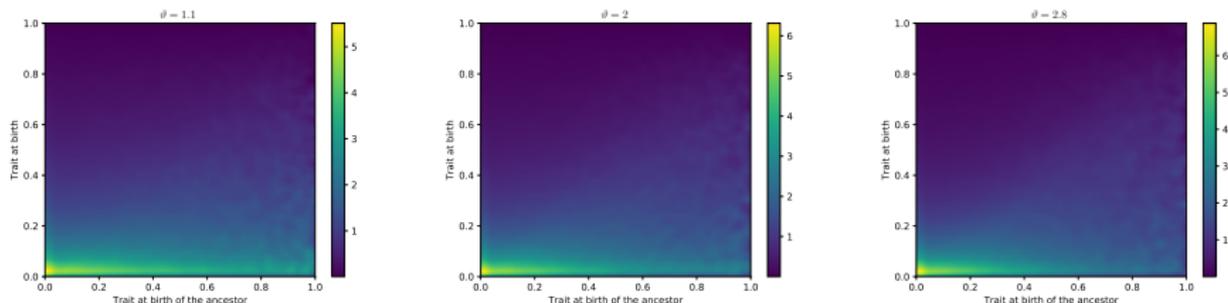


Figure: Densité de la transition pour  $B(x, \vartheta) \equiv \vartheta$ .

## Résultats

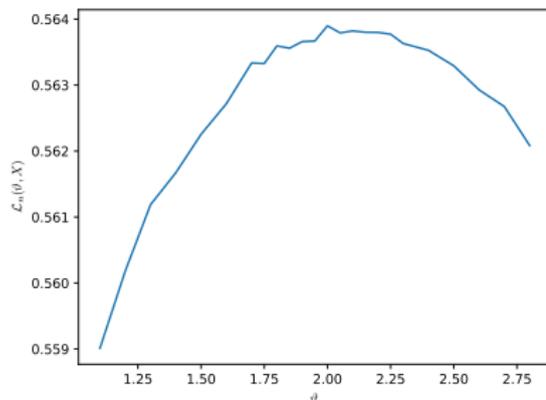
- ▶ simulation 300 jeux de données de taille  $2^{15} - 1 = 32765$  avec le paramètre  $\vartheta_0$ ,

## Résultats

- ▶ simulation 300 jeux de données de taille  $2^{15} - 1 = 32765$  avec le paramètre  $\vartheta_0$ ,
- ▶ pour chaque jeu de données, calcul du contraste de vraisemblance pour les différentes valeurs de  $\vartheta$  à partir des densités estimées.

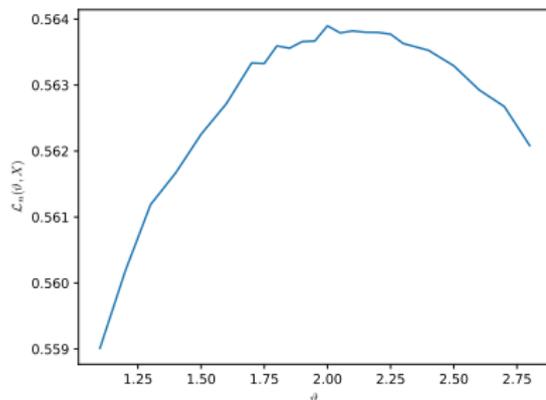
# Résultats

- ▶ simulation 300 jeux de données de taille  $2^{15} - 1 = 32765$  avec le paramètre  $\vartheta_0$ ,
- ▶ pour chaque jeu de données, calcul du contraste de vraisemblance pour les différentes valeurs de  $\vartheta$  à partir des densités estimées.

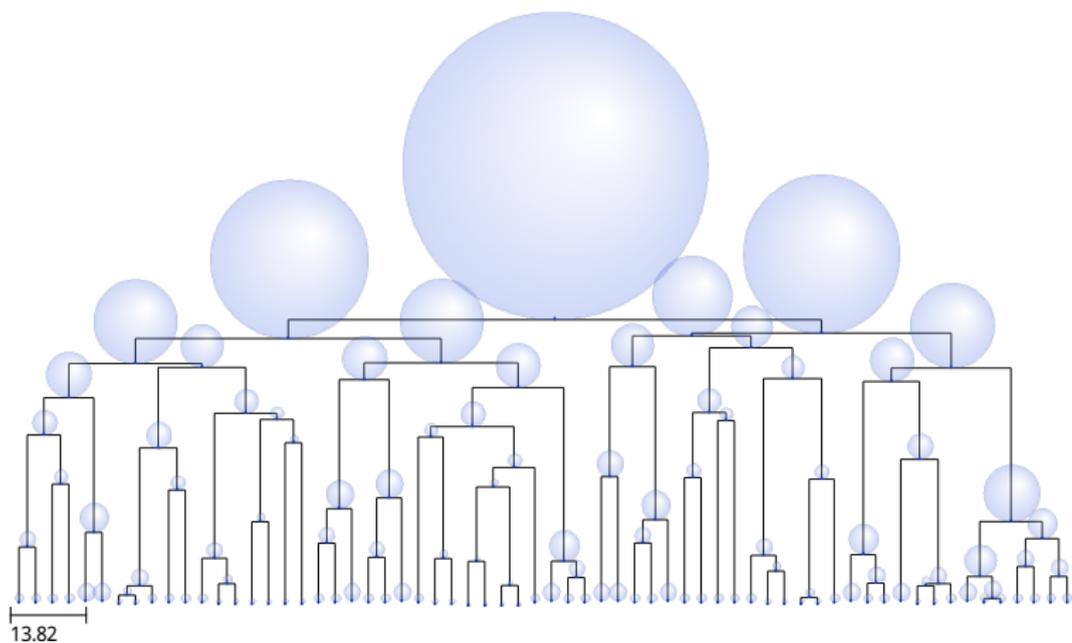


# Résultats

- ▶ simulation 300 jeux de données de taille  $2^{15} - 1 = 32765$  avec le paramètre  $\vartheta_0$ ,
- ▶ pour chaque jeu de données, calcul du contraste de vraisemblance pour les différentes valeurs de  $\vartheta$  à partir des densités estimées.



$\vartheta_0$	Moyenne	Écart-type	$\vartheta_{\min}$	$\vartheta_{\max}$	$\Delta\vartheta$
2	1.991	0.085	1.8	2.2	0.05
15	15.29	0.476	14	15.75	0.25



Merci pour votre attention !