

Modélisation stochastique de populations en expansion

Apolline Louvet (MAP5)

le 22 juin 2021

Journées de probabilité 2021

Sous la direction d'Amandine Véber (MAP5)

Motivations

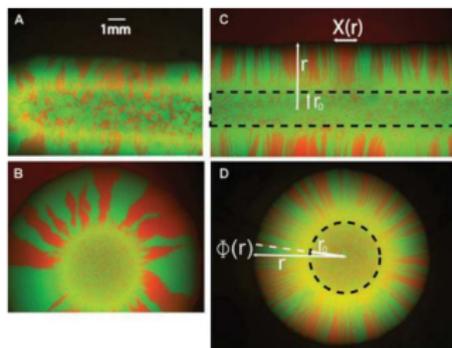


Figure 1: Bactéries fluorescentes vertes et rouges mises en croissance dans une boîte de Pétri. La fluorescence est neutre sélectivement.

Image issue de [Hallatschek2010]¹.

¹Hallatschek, O. and Nelson, D. (2010). [Life at the front of an expanding population.](#)

Evolution: International Journal of Organic Evolution, 64(1):193–206

- 1 Processus Λ -FV spatial à k parents
 - Hypothèses
 - Illustration
 - Problème martingale
- 2 Dual du processus Λ -FV spatial
 - Objectif
 - Zones habitées
 - Etude de la diversité génétique
- 3 Limite $k \rightarrow +\infty$ du Λ -FV spatial à k parents
 - Heuristique
 - Construction - Idée de preuve
 - Problème martingale
 - Illustration de la dynamique

Processus Λ -Fleming Viot spatial à k parents

Hypothèse 1 - Construction "événement-centrée"

Soit $d \geq 1$. Soit μ une mesure σ -finie sur $(0, +\infty)$ telle que

$$\int_0^{\infty} r^d \mu(dr) < +\infty.$$

Soit Π un processus ponctuel de Poisson défini sur $\mathbb{R} \times \mathbb{R}^d \times (0, +\infty)$ d'intensité

$$dt \otimes dx \otimes \mu(dr).$$

$(t, x, \mathcal{R}) \in \Pi$ correspond à un événement de reproduction :

- se produisant à l'instant t ,
- affectant la boule de centre x et de rayon \mathcal{R} .

Hypothèse 2 - Unicité du type en un point

- En chaque point occupé de l'espace, une infinité d'individus sont présents
- Tous ces individus sont du même type

Hypothèse 3 - Individus fantômes



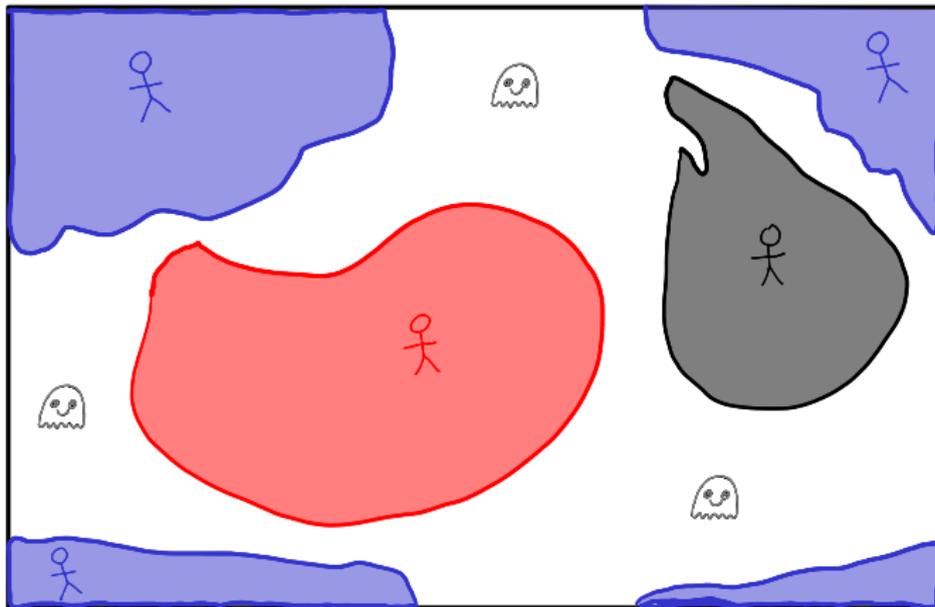
⇒ Remplissage des zones vides par des individus "fantômes", qui peuvent se reproduire aussi, mais avec un fort désavantage sélectif (quantifié par le paramètre k).

Description de la dynamique

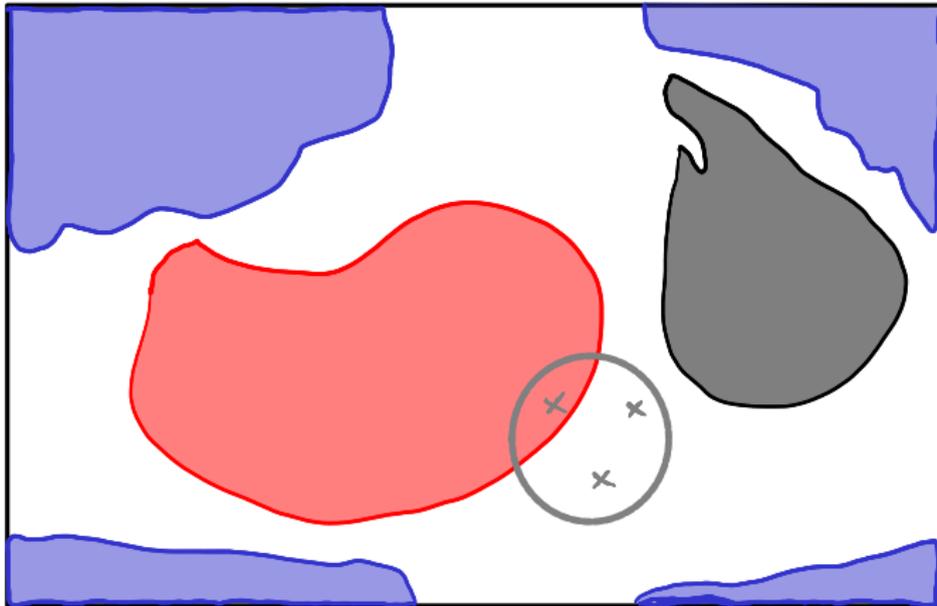
Lors d'un événement de reproduction, k *parents potentiels* sont choisis uniformément au hasard à l'intérieur de la boule:

- Si tous sont des parents fantômes, la boule est remplie d'individus fantômes.
- Si au moins un des parents est réel, le vrai parent est le premier individu réel choisi.

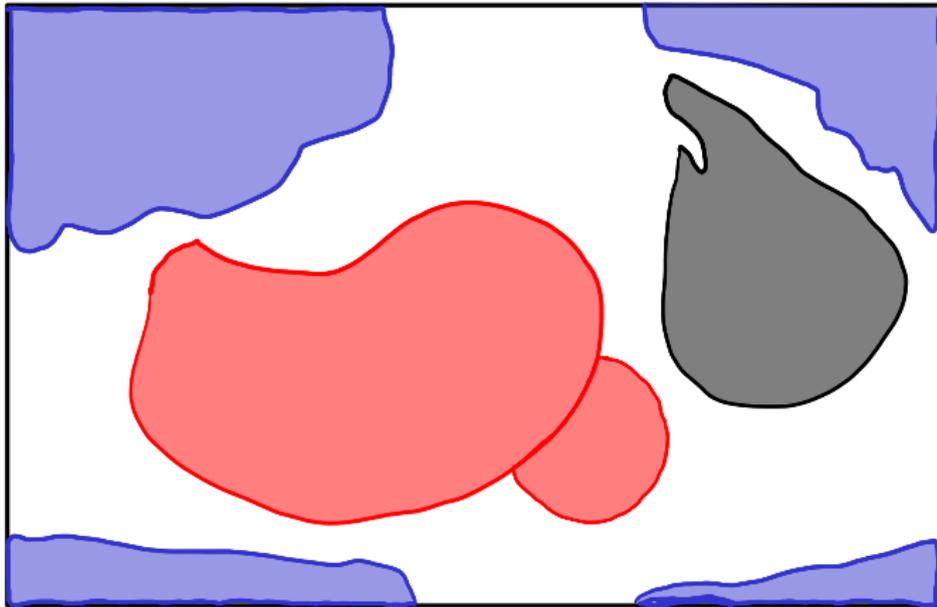
Modélisation ($k = 3$)



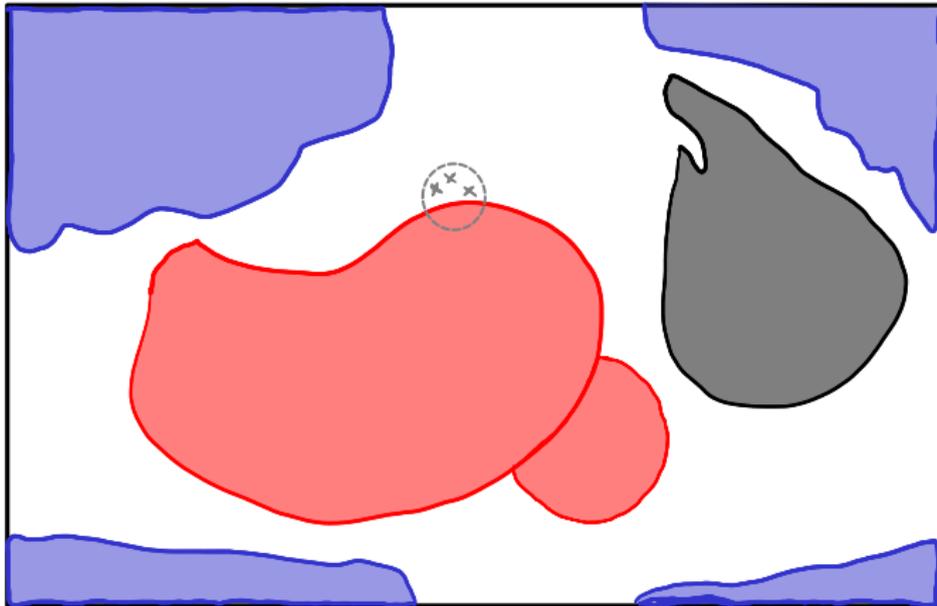
Modélisation ($k = 3$)



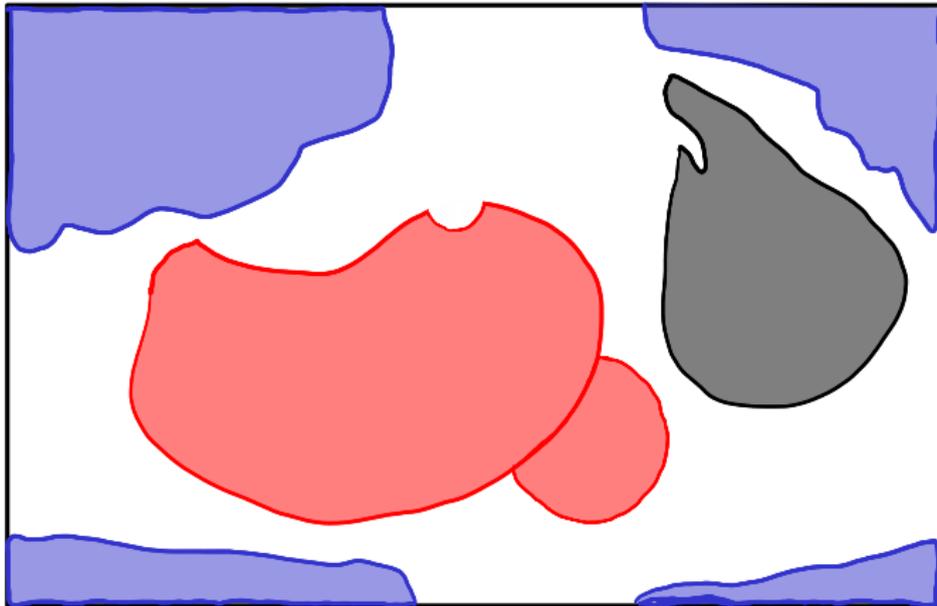
Modélisation ($k = 3$)



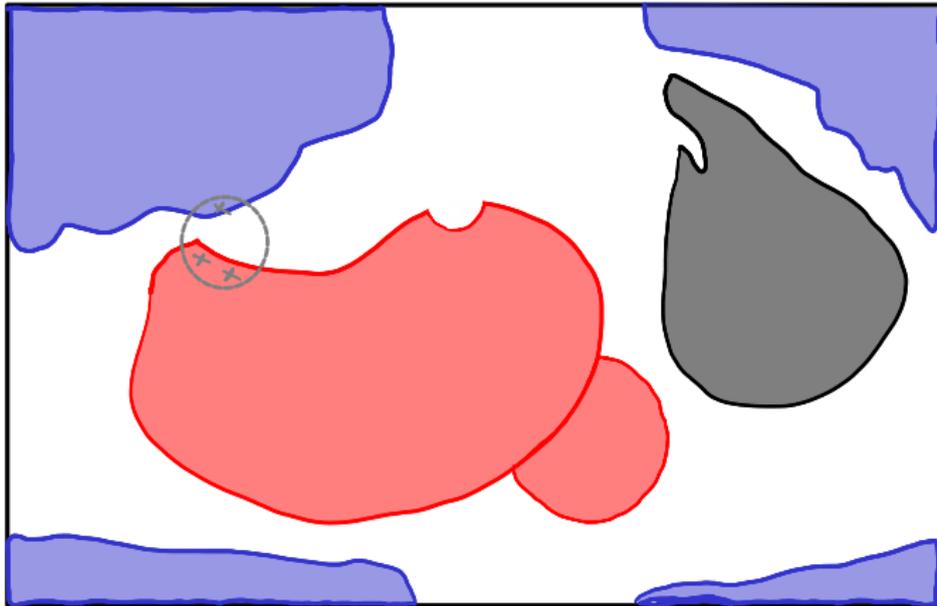
Modélisation ($k = 3$)



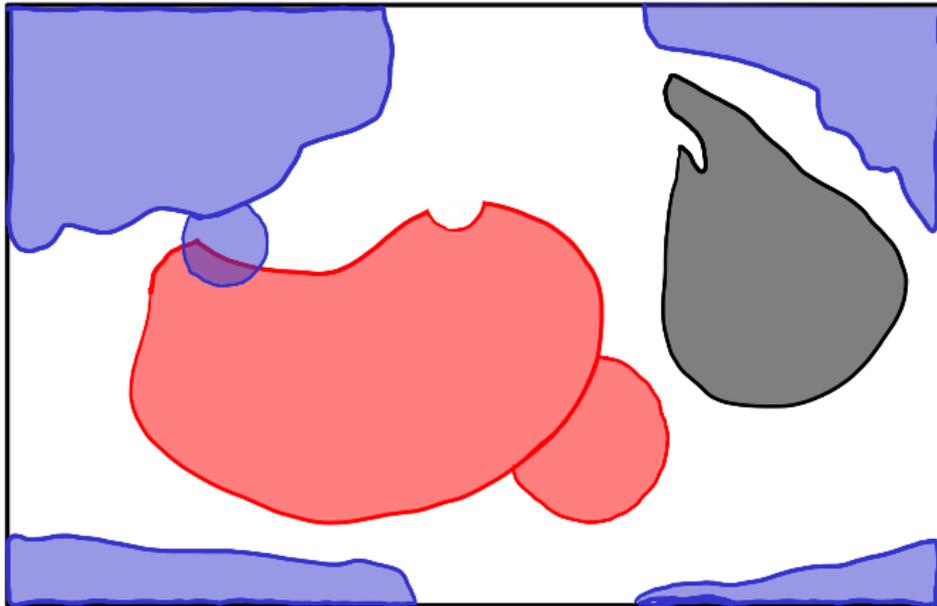
Modélisation ($k = 3$)



Modélisation ($k = 3$)



Modélisation ($k = 3$)



Espace d'états

⚠ Pour alléger les notations, nous considérerons que la population ne comprend que 2 types différents, notés **type 1** et type 0.

\mathcal{M}_λ : ensemble des mesures M sur $\mathbb{R}^d \times \{0, 1\}$ telles qu'il existe $\omega : \mathbb{R}^d \rightarrow \{0, 1\}$ mesurable satisfaisant

$$M(dx, dk) = ((\omega(x)\delta_0(dk) + (1 - \omega(x))\delta_1(dk)) dx,$$

muni de la topologie de la convergence vague.

$\Rightarrow \omega$: *densité en individus de type 0*

⚠ Pour tout $M \in \mathcal{M}_\lambda$, ω n'est pas unique, mais défini à un ensemble de mesure de Lebesgue nulle près.

Espace d'états

$D_{\mathcal{M}_\lambda}[0, +\infty)$: ensemble des chemins càdlàg à valeurs dans \mathcal{M}_λ , muni de la topologie de Skorokhod.

Fonctions test et notations

ω_M : représentant arbitrairement choisi de la mesure $M \in \mathcal{M}_\lambda$

$\Psi_{F,f}$: fonction test

Pour tout $f \in C_c(\mathbb{R}^d)$ et $F \in C^1(\mathbb{R})$, la fonction $\Psi_{F,f} \in C_b(\mathcal{M}_\lambda)$ est définie par

$$\begin{aligned}\forall M \in \mathcal{M}_\lambda, \Psi_{F,f}(M) &:= F\left(\left\langle \int_{\mathbb{R}^d} f(x) \omega_M(x) dx \right\rangle\right) \\ &= F(\langle \omega_M, f \rangle).\end{aligned}$$

$\Theta_{x,\mathcal{R}}^+(\omega) / \Theta_{x,\mathcal{R}}^-(\omega)$: remplissage de la boule $\mathcal{B}(x, \mathcal{R})$ par des individus de type 0/1

Pour tout $\omega : \mathbb{R}^d \rightarrow \{0,1\}$ mesurable, $\mathcal{R} \in \mathbb{R}_+^*$ et $x \in \mathbb{R}^d$,

$$\Theta_{x,\mathcal{R}}^+(\omega) := \mathbb{1}_{\{\mathcal{B}(x,\mathcal{R})^c\}} \times \omega + \mathbb{1}_{\{\mathcal{B}(x,\mathcal{R})\}}$$

$$\Theta_{x,\mathcal{R}}^-(\omega) := \mathbb{1}_{\{\mathcal{B}(x,\mathcal{R})^c\}} \times \omega.$$

Problème martingale (2 types)

\mathcal{L}_μ^k : opérateur agissant sur les fonctions tests $\Psi_{F,f}$.

Pour tout $f \in C_c(\mathbb{R}^d)$, $F \in C_1(\mathbb{R})$ et $M \in \mathcal{M}_\lambda$,

$$\begin{aligned} \mathcal{L}_\mu^k \Psi_{F,f}(M) := & \int_{\mathbb{R}^d} \int_0^\infty \int_{\mathcal{B}(x, \mathcal{R})^k} \frac{1}{V_{\mathcal{R}}^k} \times \left[\left(\prod_{j=1}^k \omega_M(y_j) \right) \times F \left(\langle \Theta_{x, \mathcal{R}}^+(\omega_M), f \rangle \right) \right. \\ & + \left. \left(1 - \left(\prod_{j=1}^k \omega_M(y_j) \right) \right) \times F \left(\langle \Theta_{x, \mathcal{R}}^-(\omega_M), f \rangle \right) \right. \\ & \left. - F(\langle \omega_M, f \rangle) \right] dy_1 \dots dy_k \mu(d\mathcal{R}) dx. \end{aligned}$$

Définition du Λ -FV spatial avec sélection à k parents

Théorème

Pour tout $M^0 \in \mathcal{M}_\lambda$, il existe un unique processus $(M_t)_{t \geq 0}$ à valeurs dans $D_{\mathcal{M}_\lambda}[0, +\infty)$ tel que $M_0 = M^0$ et, pour tout $F \in C^1(\mathbb{R})$, $f \in C_c(\mathbb{R}^d)$,

$$\left(\Psi_{F,f}(M_t) - \Psi_{F,f}(M_0) - \int_0^t \mathcal{L}_\mu^k \Psi_{F,f}(M_s) ds \right)_{t \geq 0}$$

est une martingale. De plus, ce processus est Markovien.

Preuve : voir [Etheridge2020]¹ pour le cas $k = 2$.

¹Etheridge, A., Véber, A., and Yu, F. (2020). [Rescaling limits of the spatial Lambda-Fleming-Viot process with selection.](#)

Définition du Λ -FV spatial avec sélection à k parents

Définition

Le Λ -Fleming Viot spatial avec sélection à k parents de condition initiale M^0 associé à μ est l'unique solution du problème martingale (\mathcal{L}_μ^k, M^0) .

- Cas $k = 2$: cas particulier de [Etheridge2020]
- Cas $k \geq 2$: pour cette définition particulière, voir [L.2021]¹. Pour une définition plus générale, recouvrant plusieurs formes de sélection, voir [Forien2017]².

¹L., A. (2021). [The \$k\$ -parent spatial Lambda-Fleming-Viot process as a stochastic measure-valued model for an expanding population.](#)
arXiv preprint arXiv:2103.02902

²Forien, R. and Penington, S. (2017). [A central limit theorem for the spatial Lambda-Fleming-Viot process with selection.](#)
Electronic Journal of Probability, 22

Références bibliographiques

Utilisation d'individus fantômes pour modéliser des populations en expansion

- [Hallatschek2008]¹ : analyse principalement par simulations
- [Durrett2016]² : faible avantage sélectif des individus réels sur les individus fantômes

⇒ **Ici, nous nous intéressons à un régime de sélection forte, et à une limite dans laquelle la sélection est infiniment grande.**

¹Hallatschek, O. and Nelson, D. (2008). [Gene surfing in expanding populations.](#)

Theoretical Population Biology, 73(1):158–170

²Durrett, R. and Fan, W.-T. (2016). [Genealogies in expanding populations.](#)
Annals of Applied Probability, 26(6):3456–3490

Dual du processus Λ -Fleming Viot spatial à k parents

Notion de processus dual

- Permet de "remonter en arrière" dans le temps
- Utilisé dans les preuves d'unicité de solutions
- En génétique des populations, correspond à reconstruire les généalogies d'un échantillon d'individus

Idee: En l'absence de mutations, si deux individus ont un ancêtre commun, alors ils sont du même type.

\Rightarrow *Permet d'étudier la diversité génétique*

Notion de processus dual

- Permet de "remonter en arrière" dans le temps
- Utilisé dans les preuves d'unicité de solutions
- En génétique des populations, correspond à reconstruire les généalogies d'un échantillon d'individus

Idée: En l'absence de mutations, si deux individus ont un ancêtre commun, alors ils sont du même type.

\Rightarrow Permet d'étudier la diversité génétique

⚠ 2 niveaux de "diversité génétique" :

- individus "réels" de type 1 VS individus "fantômes" de type 0
- diversité génétique parmi les individus de type 1 : types 1.1, 1.2, 1.3,... (ajoutés via techniques de "traceurs" [Hallatschek2008], [Durrett2016])

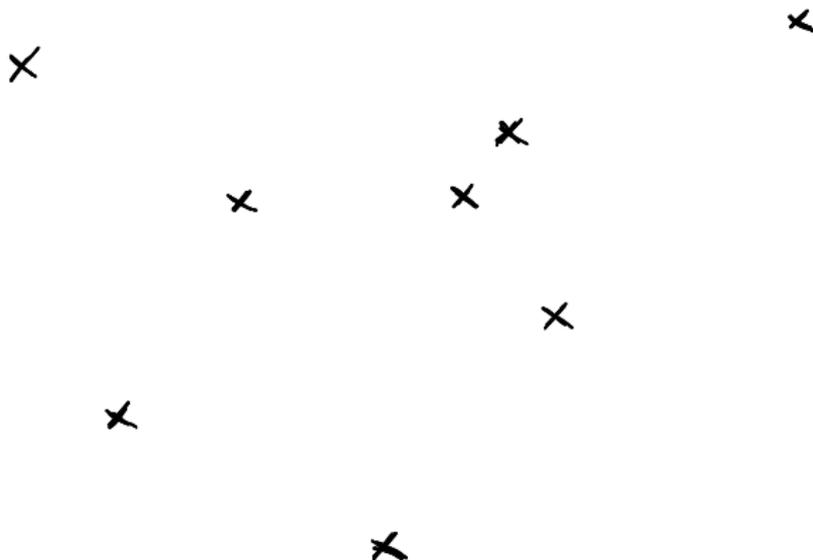
Etude des zones habitées

zones habitées = zones occupées
par des individus de type 1

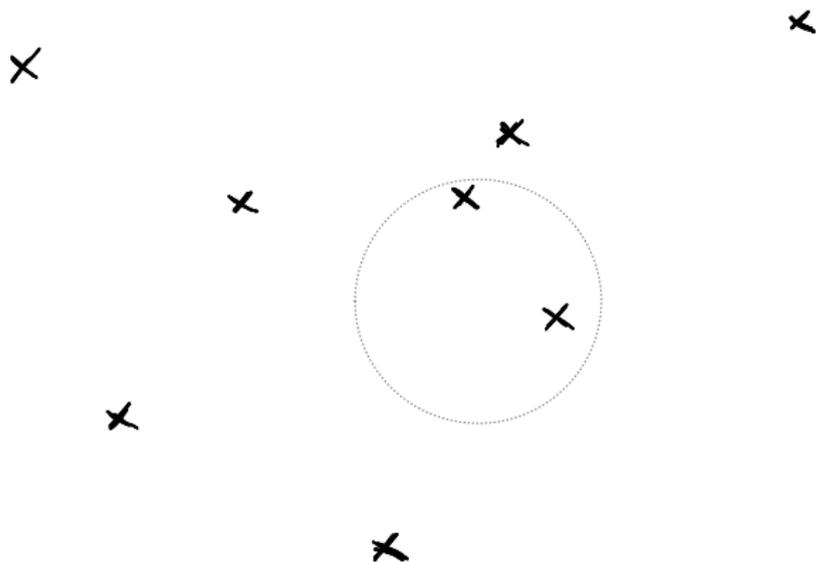
⚠ Pour connaître le parent d'un individu parmi k parents potentiels, il faut connaître le type de tous les k parents potentiels... or c'est ce que l'on cherche à savoir !

⇒ **Les k parents potentiels sont intégrés au dual.**

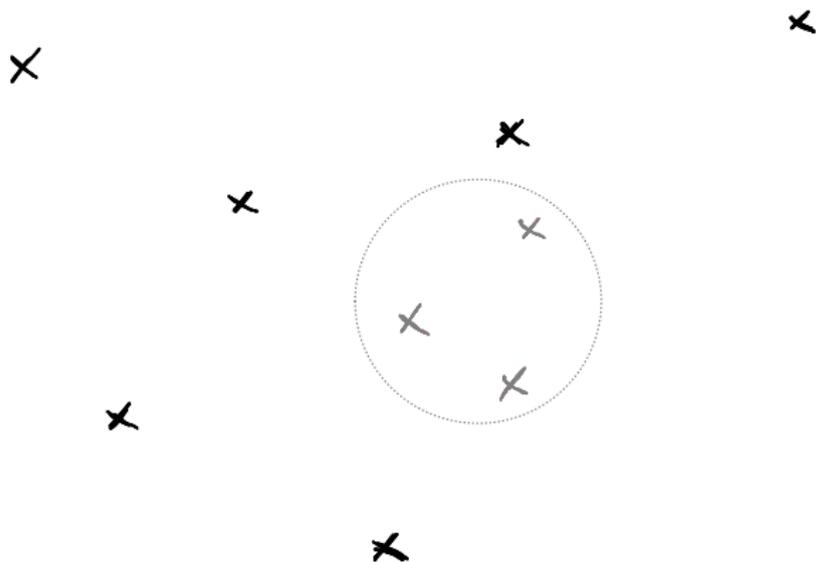
Dual du Λ -FV avec sélection à k parents ($k = 3$)



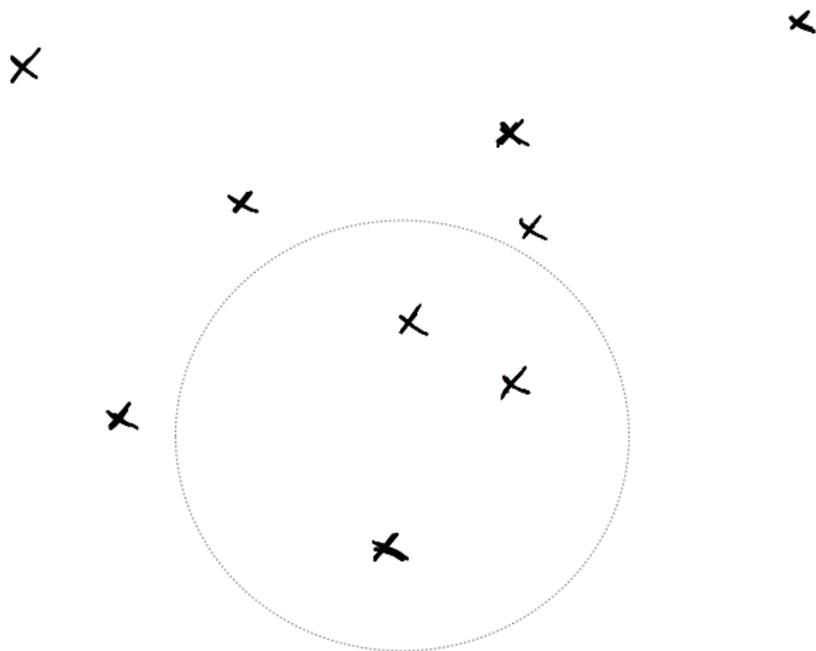
Dual du Λ -FV avec sélection à k parents ($k = 3$)



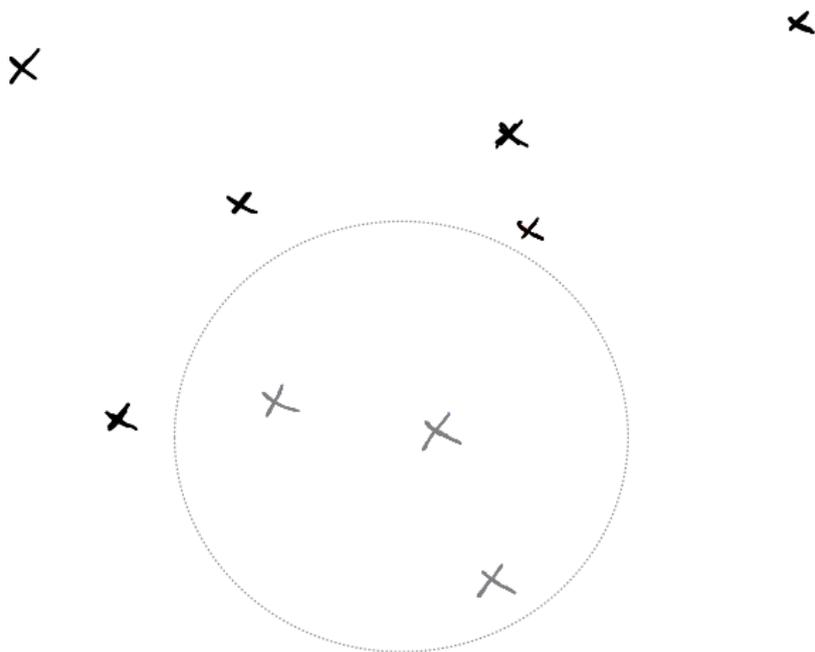
Dual du Λ -FV avec sélection à k parents ($k = 3$)



Dual du Λ -FV avec sélection à k parents ($k = 3$)



Dual du Λ -FV avec sélection à k parents ($k = 3$)



Dual du Λ -FV spatial

$\mathcal{M}_p(\mathbb{R}^d)$: ensemble des mesures ponctuelles finies sur \mathbb{R}^d , muni de la topologie de la convergence faible.

Si $\Xi = \sum_{i=1}^l \delta_{\xi_i} \in \mathcal{M}_p(\mathbb{R}^d)$,

$$I_{x,\mathcal{R}}(\Xi) := \{i \in \llbracket 1, l \rrbracket : \|x - \xi_i\| \leq \mathcal{R}\}.$$

Définition

Soit $\Xi^0 \in \mathcal{M}_p(\mathbb{R}^d)$. Le processus ancestral $(\Xi_t)_{t \geq 0}$ associé à μ de condition initiale Ξ^0 est le processus de Markov à valeurs dans $\mathcal{M}_p(\mathbb{R}^d)$ défini de la façon suivante.

- $\Xi_0 = \Xi^0$.
- Pour tout $(t, x, \mathcal{R}) \in \Pi$, si $I_{x,\mathcal{R}}(\Xi_{t-}) \neq \emptyset$, Ξ_t est construit en retirant tous les atomes de Ξ_{t-} dans $\mathcal{B}(x, \mathcal{R})$, et en ajoutant k atomes pris uniformément au hasard dans $\mathcal{B}(x, \mathcal{R})$.

Relation de dualité

Soient $M^0 \in \mathcal{M}_\lambda$, $l \in \mathbb{N}^*$ et ψ une densité sur $(\mathbb{R}^d)^l$.

$(M_t)_{t \geq 0}$: Λ -FV spatial à k parents de condition initiale M^0 associé à μ .

Proposition

Pour tout $t \geq 0$,

$$\begin{aligned} & \mathbb{E}_{M^0} \left[\int_{(\mathbb{R}^d)^l} \psi(x_1, \dots, x_l) \left\{ \prod_{j=1}^l \omega_{M_t}(x_j) \right\} dx_1 \dots dx_l \right] \\ &= \int_{(\mathbb{R}^d)^l} \psi(x_1, \dots, x_l) E_{\Xi[x_1, \dots, x_l]} \left[\prod_{j=1}^{N_t} \omega_{M_0}(\xi_t^j) \right] dx_1 \dots dx_l, \end{aligned}$$

où $\Xi_t = \sum_{i=1}^{N_t} \delta_{\xi_t^i}$ est le processus ancestral de condition initiale $\sum_{i=1}^l \delta_{x_i}$ est associé à μ .

Relation de dualité - Interprétation

Relation de dualité

$$\begin{aligned} & \mathbb{E}_{M^0} \left[\int_{(\mathbb{R}^d)^I} \psi(x_1, \dots, x_I) \left\{ \prod_{j=1}^I \omega_{M_t}(x_j) \right\} dx_1 \dots dx_I \right] \\ &= \int_{(\mathbb{R}^d)^I} \psi(x_1, \dots, x_I) E_{\Xi[x_1, \dots, x_I]} \left[\prod_{j=1}^{N_t} \omega_{M_0}(\xi_t^j) \right] dx_1 \dots dx_I, \end{aligned}$$

\Rightarrow un échantillon contient **au moins un individu de type 1** si, et seulement si **au moins l'un des ancêtres de l'échantillon au temps 0 est de type 1**.

Etude de la diversité génétique

→ A partir du processus dual, mais en gardant en mémoire l'ordre dans lequel les parents potentiels sont choisis.

Basé sur l'ordered ASG introduit dans [Lenz2015]¹

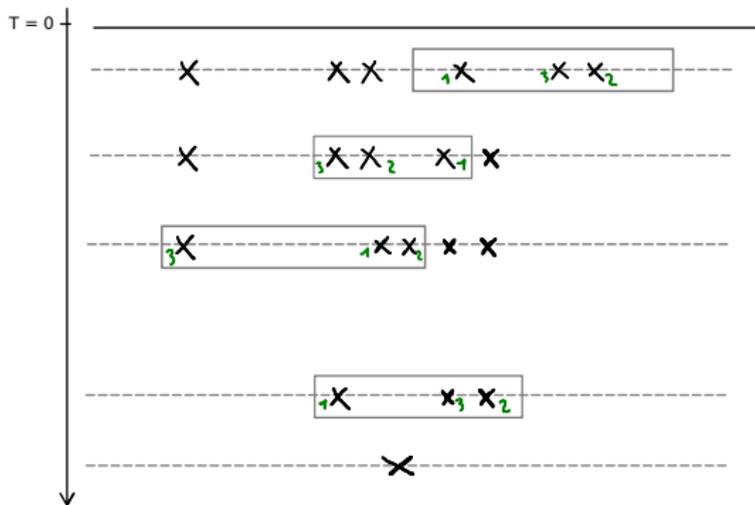
- *Etape 1*: Recherche des ancêtres potentiels d'un échantillon, jusqu'à remonter à la condition initiale
- *Etape 2*: Atteinte de la condition initiale
- *Etape 3*: Résolution des types

¹Lenz, U., Kluth, S., Baake, E., and Wakolbinger, A. (2015). [Looking down in the ancestral selection graph: A probabilistic approach to the common ancestor type distribution.](#)

Theoretical Population Biology, 103:27–37

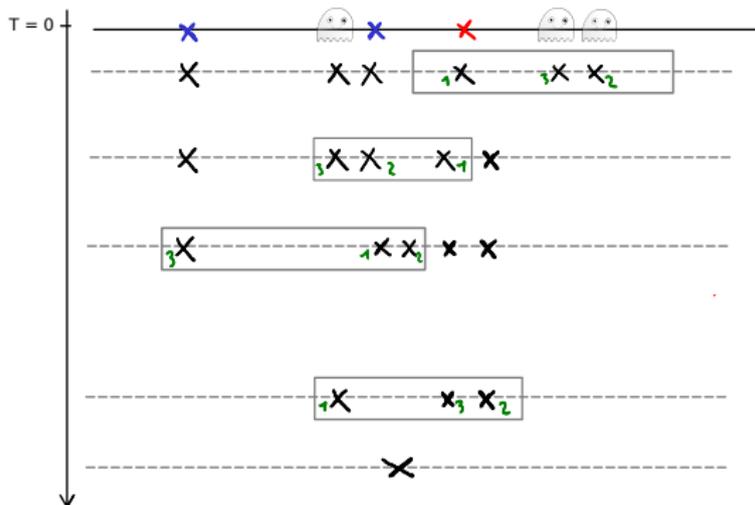
Dual du Λ -FV avec sélection à k parents

Etape 1: Recherche des ancêtres potentiels d'un échantillon, jusqu'à remonter à la condition initiale.



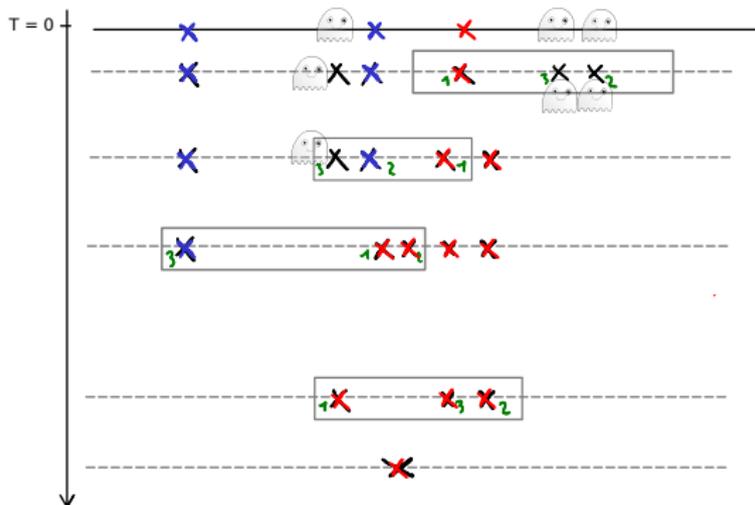
Dual du Λ -FV avec sélection à k parents

Etape 2: Atteinte de la condition initiale.



Dual du Λ -FV avec sélection à k parents

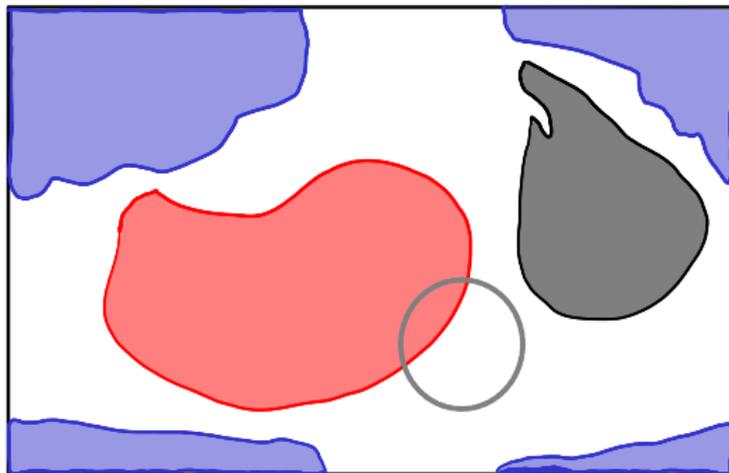
Etape 3: Résolution des types.



Limite $k \rightarrow +\infty$ du Λ -Fleming Viot spatial avec sélection à k parents

Limite $k \rightarrow +\infty$ - Heuristique

Lorsque $k \rightarrow +\infty$, si les individus de type 1 occupent une fraction de volume non nul de la boule de reproduction, alors ils vont finir par être choisis comme parents potentiels.



Construction de la limite $k \rightarrow +\infty$

Idée générale de la construction basée sur [Véber2015]¹

Idée : Extension du processus ponctuel de Poisson, pour y intégrer le choix des parents potentiels

Soit $U = \mathcal{B}(0,1)^{\mathbb{N}}$, et soit \tilde{u} la loi d'une séquence de variables aléatoires i.i.d $(P_n)_{n \geq 1}$ uniformément distribuées sur $\mathcal{B}(0,1)$.

$(p_n)_{n \geq 1} \in U$: *séquence ordonnée de parents potentiels*

Soit Π^c un processus ponctuel de Poisson défini sur $\mathbb{R} \times \mathbb{R}^d \times (0, +\infty) \times U$ d'intensité

$$dt \otimes dx \otimes \mu(dr) \otimes \tilde{u}(d(p_n)_{n \geq 1}).$$

¹Véber, A. and Wakolbinger, A. (2015). [The spatial Lambda-Fleming-Viot process: An event-based construction and a lockdown representation.](#)

Construction de la limite $k \rightarrow +\infty$

\Rightarrow Le processus ponctuel de Poisson Π^c encode toute la stochasticité dans le processus Λ -Fleming Viot spatial à k parents.

Construction alternative du Λ -Fleming Viot spatial à k parents à partir de Π^c

- Notation : $(\omega_t^{\Pi^c, k})_{t \geq 0}$, de condition initiale ω
- A chaque événement de reproduction, les k premiers parents potentiels donnés par Π^c sont utilisés
- Pour tout $t \geq 0$ et $x \in \mathbb{R}^d$, $\omega_t^{\Pi^c, k}(x)$ est défini en utilisant Π^c pour remonter jusqu'aux ancêtres potentiels des individus en x au temps t

Remarque : Conditionnellement à la connaissance de Π^c , $(\omega_t^{\Pi^c, k}(x))$ est totalement déterministe.

Construction de la limite $k \rightarrow +\infty$

Proposition

Pour tout $t \geq 0$ et $x \in \mathbb{R}^d$, la suite

$$(\omega_t^{\Pi^c, k}(x))_{k \geq 2}$$

est décroissante et à valeurs dans $\{0, 1\}$.

La limite $(\omega_t^\infty)_{t \geq 0}$ est alors définie en posant, pour tout $t \geq 0$ et $x \in \mathbb{R}^d$,

$$\omega_t^\infty(x) = \lim_{k \rightarrow +\infty} \omega_t^{\Pi^c, k}(x).$$

Problème martingale (2 types)

Cas particulier : Rayon des événements de reproduction fixé et égal à \mathcal{R}_0 .

\mathcal{L}_μ^∞ : opérateur agissant sur les fonctions de la forme $\Psi_{F,f}$.

Pour tout $f \in C_c(\mathbb{R}^d)$, $F \in C_1(\mathbb{R})$ et $M \in \mathcal{M}_\lambda$,

$$\begin{aligned} \mathcal{L}_\mu^\infty \Psi_{F,f}(M) := & \int_{\mathbb{R}^d} \left[\delta_0 \left(\int_{\mathcal{B}(x, \mathcal{R}_0)} (1 - \omega_M(z)) dz \right) \times F \left(\left\langle \Theta_{x, \mathcal{R}_0}^+(\omega_M), f \right\rangle \right) \right. \\ & + \left(1 - \delta_0 \left(\int_{\mathcal{B}(x, \mathcal{R}_0)} (1 - \omega_M(z)) dz \right) \right) \times F \left(\left\langle \Theta_{x, \mathcal{R}_0}^-(\omega_M), f \right\rangle \right) \\ & \left. - F(\langle \omega_M, f \rangle) \right] dx. \end{aligned}$$

Illustration de la dynamique

Merci pour votre attention !

