

Horizontal Gene Transfer:

numerical comparison between stochastic and deterministic approaches

Vincent Calvez¹ Susely Figueroa Iglesias² H      Hivert³
Sylvie M     ⁴ Anna Melnykova⁵ Samuel Nordmann⁶

Seminar GdT MathBio,
12 November, 2018

¹CNRS & Institut Camille Jordan, Lyon. Vincent.Calvez@math.cnrs.fr

²Univ. Toulouse, IMT, UPS. Susely.Figueroa@math.univ-toulouse.fr

³Univ. Lyon,       centrale de Lyon. helene.hivert@ec-lyon.fr

⁴Ecole Polytechnique, France. sylvie.meleard@polytechnique.edu

⁵Univ. Cergy-Pontoise & Univ. Grenoble Alpes. anna.melnykova@u-cergy.fr

⁶CAMS, EHESS, PSL Universit  , Paris. samuel.nordmann@ehess.fr

Introduction

Definition

Horizontal Gene Transfer (HGT) — exchange of genetic material between unicellular/multicellular organisms not through a vertical transmission.

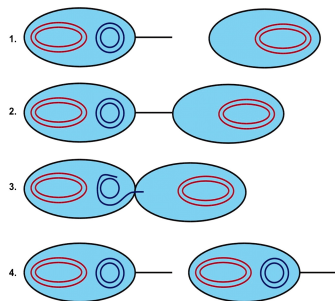
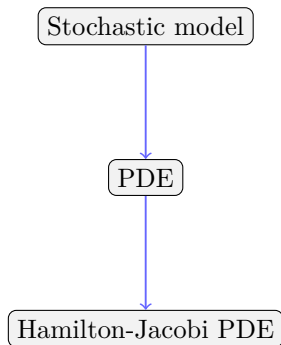


Figure 1: Unilateral HGT

HGT plays a key role in:

- ▶ the evolution of organisms
- ▶ transmission of pathogens
- ▶ resistance to antibiotics in bacteria

Picture from Raz, Y. & Tannenbaum, E. (2010). The Influence of Horizontal Gene Transfer on the Mean Fitness of Unicellular Populations in Static Environments. *Genetics*, 185(1), 327–337.



Our goal:

Study in silico the population dynamics using:

- ▶ Individual-based model (stochastic jump process)
- ▶ Large population limit (PDE)
- ▶ Large population on evolutionary time scale (Hamilton-Jacobi PDE)

Model

The demography of the population is regulated by:

- ▶ $x_j^t \in \mathbb{R}^d$ is the trait of j -th individual living at t
- ▶ K : scale parameter (carrying capacity), N^t — size of the population at time t
- ▶ Birth: $b(x)$ with a mutation kernel $m(x,y)$.
- ▶ Death: $d(x) + C \frac{N^t}{K}$, where C is a competition rate
- ▶ Horizontal Gene Transfer (HGT):

$$h_K(x, y, N^t) = \tau_0 \alpha(x - y) \frac{K}{N^t}$$

where α a smooth bounded function (for example: heaviside, arctan, etc..).

Stochastic model

For a population described by point measure $\nu^t(dx) = \frac{1}{K} \sum_{i=0}^{N^t} \delta_{X_i(t)}(dx)$ the generator of the process is:

$$\begin{aligned} L^K F(\nu) = & \sum_{i=1}^N b(x_i) \int_{\mathbb{R}^d} \left(F\left(\nu + \frac{1}{K} \delta_y\right) - F(\nu) \right) m(x_i, dy) \\ & + \sum_{i=1}^N \left(d(x_i) + C \frac{N}{K} \right) \left(F\left(\nu - \frac{1}{K} \delta_{x_i}\right) - F(\nu) \right) \\ & + \sum_{i,j=1}^N h_K(x_i, x_j, \nu) \left(F\left(\nu + \frac{1}{K} \delta_{x_i} - \frac{1}{K} \delta_{x_j}\right) - F(\nu) \right). \end{aligned}$$

Further reading:

Billiard et al. (2016a, 2015)

Stochastic simulations: algorithm

Algorithm of simulation on $[0, \Delta, \dots, i\Delta, \dots, N\Delta = T]$:

0 Randomly initialize the population $X^0 := \mathcal{N}(x_{mean}^0, \sigma^0) \times N^0$

1 Set: $b \equiv const$, $d = d_r x^2 + C \frac{N^t}{K}$, $\alpha(x) = heav(x)$

2 For $i \in [1, N - 1]$:

(i) Copy X^{i-1} -th population to X^i

(ii) Compute exponential arrival times for each individual:

$$T_b := \lambda(b), T_d := \lambda(d), T_{HGT} := \lambda \left(\sum_{y \in X^i} h_K(x, y, N^{i-1}) \right)$$

(iii) $\forall x \in X^i$:

* If $\mathbf{T}_d \leq \Delta$: remove x -th individual

* If $\mathbf{T}_b \leq \Delta$: add a new individual with a trait from $\mathcal{N}(x, \sigma^2)$

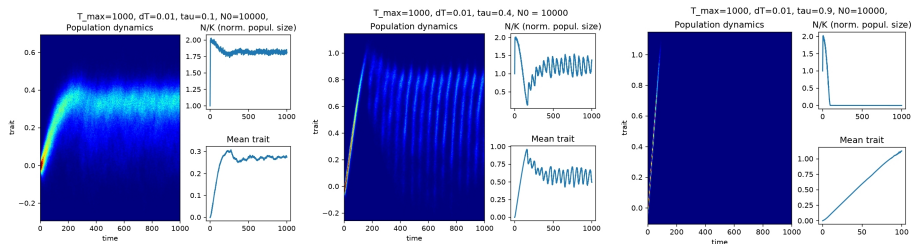
* If $\mathbf{T}_{HGT} \leq \Delta$: pick a trait $y \in X^i$ according to the law

$$\frac{h_K(x, y, N^{i-1})}{\sum_{y \in X^i} h_K(x, y, N^{i-1})}$$

then remove individual with trait x and add a new individual with trait y .

(iv) Return X^i

Stochastic simulations



(a) Stabilization: $\tau = 0.1$

(b) Cycles: $\tau = 0.4$

(c) Extinction: $\tau = 0.9$

Figure 2: Behavior of the population dynamics as the mutation rate τ is changing, ($b_r = d_r = 1$, $\sigma = 0.01$, $T_{max} = 1000$, $\Delta = 0.01$).

All codes and more pictures:

https://github.com/melnyashka/horizontal_gene_transfer

Stochastic model

Advantages

- + Very precise
- + Allows to track the heritage

Limitations

- Computational cost
- Hard to analyze from theoretical point of view

Stochastic model

Advantages

- + Very precise
- + Allows to track the heritage

Limitations

- Computational cost
- Hard to analyze from theoretical point of view

Solution

Pass to the limit with PDE!

PDE: Model and notations

The sequence of stochastic processes gives a macroscopic approximation when $K \rightarrow +\infty$.

$$\left\{ \begin{array}{l} \partial_t f(t, x) = -(d(x) + C\rho(t))f(t, x) + \int_{\mathbb{R}^d} m(x - y)b(y)f(t, y)dy \\ \quad + f(t, x) \int_{\mathbb{R}^d} \tau(x - y) \frac{f(t, y)}{\rho(t)} dy, \\ \rho(t) = \int_{\mathbb{R}^d} f(t, x) dx, \quad f(0, x) = f^0(x) > 0. \end{array} \right.$$

- ▶ $f(t, x) := \lim_{K \rightarrow +\infty} \nu^K$: density of the population at time t with trait x ,
- ▶ $b(x)$, $d(x)$, C : birth, death and competition rate,
- ▶ $m(x - y)$: mutation kernel,
- ▶ $\tau(y - x) := \tau_0 [\alpha(x - y) - \alpha(y - x)]$: HT kernel

Further reading:

Billiard et al. (2016b,a), Ferrière and Tran (2009)

PDE simulations

Explicit Euler scheme: the integrals are computed with a left-point quadrature rule.

► Competition term: $\Delta_x \sum_{j=0}^{N_x-1} f_j^n$

$$\frac{f_i^{n+1} - f_i^n}{\Delta_t} = \left(- \underbrace{C \rho^n}_{\text{orange circle}} - \underbrace{D_i}_{\text{red square}} + \underbrace{T_i^n}_{\text{green square}} \right) f_i^n + \underbrace{B_i^n}_{\text{blue circle}}$$

► Death rate term

► Transfer rate term: $\Delta_x \sum_{j=0}^{N_x-1} \tau(x_i - x_j) \frac{f_j^n}{\rho^n}$

► Birth rate term: $[m * (bf)]_i^n = \Delta_z \sum_{k=0}^{N_z-1} m(z_k) b(x_i + z_k) f^n(x_i + z_k).$

PDE simulations

Explicit Euler scheme: the integrals are computed with a left-point quadrature rule.

- Competition term: $\Delta_x \sum_{j=0}^{N_x-1} f_j^n$
- $$\frac{f_i^{n+1} - f_i^n}{\Delta_t} = \left(- \underbrace{C \rho^n}_{\text{orange circle}} - \underbrace{D_i}_{\text{pink square}} + \underbrace{T_i^n}_{\text{green square}} \right) f_i^n + \underbrace{B_i^n}_{\text{purple circle}}$$
- Death rate term
- Transfer rate term: $\Delta_x \sum_{j=0}^{N_x-1} \tau(x_i - x_j) \frac{f_j^n}{\rho^n}$
- Birth rate term: $[m * (bf)]_i^n = \Delta_z \sum_{k=0}^{N_z-1} m(z_k) b(x_i + z_k) f^n(x_i + z_k).$

Explicit Euler scheme: the integrals are computed with a left-point quadrature rule.

► Competition term: $\Delta_x \sum_{j=0}^{N_x-1} f_j^n$

$$\frac{f_i^{n+1} - f_i^n}{\Delta_t} = \left(-C\rho^n - D_i + T_i^n \right) f_i^n + B_i^n$$

► Death rate term

► Transfer rate term: $\Delta_x \sum_{j=0}^{N_x-1} \tau(x_i - x_j) \frac{f_j^n}{\rho^n}$

► Birth rate term: $[m * (bf)]_i^n = \Delta_z \sum_{k=0}^{N_z-1} m(z_k) b(x_i + z_k) f^n(x_i + z_k).$

Explicit Euler scheme: the integrals are computed with a left-point quadrature rule.

► Competition term: $\Delta_x \sum_{j=0}^{N_x-1} f_j^n$

$$\frac{f_i^{n+1} - f_i^n}{\Delta_t} = \left(-C\rho^n - D_i + T_i^n \right) f_i^n + B_i^n$$

► Death rate term

► Transfer rate term: $\Delta_x \sum_{j=0}^{N_x-1} \tau(x_i - x_j) \frac{f_j^n}{\rho^n}$

► Birth rate term: $[m * (bf)]_i^n = \Delta_z \sum_{k=0}^{N_z-1} m(z_k) b(x_i + z_k) f^n(x_i + z_k).$

PDE simulations

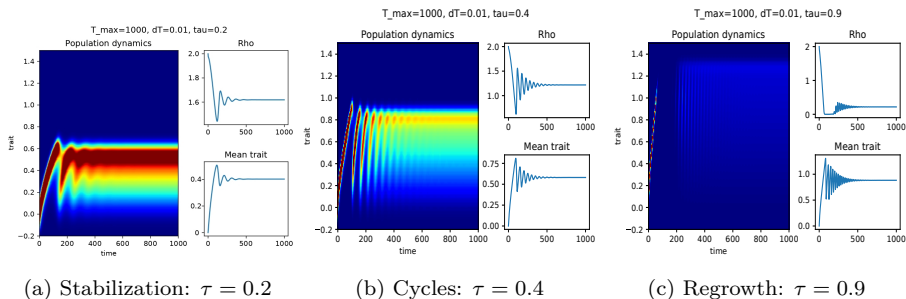


Figure 3: Behavior of the population dynamics as the mutation rate τ is changing.

PDE simulations

PDE model

- + Fast simulation
- Not very well adapted to small values

	$\Delta = 0.1, T = 10$	$\Delta = 0.01, T = 10$
SM ($N = 1000$)	3.883s	38.145s
SM ($N = 10000$)	15.805s	153.255s
PDE	0.186s	1.673s

Table 1: Elapsed time for simulation of population dynamics for stochastic and PDE model.

Limit of PDE model

Evolutionary scale:

- ▶ Time scale $T \rightarrow \infty$
- ▶ Mutation probability is small: $m(y - y') \rightarrow \mathcal{N}(0, \sigma^2)$, where $\sigma \rightarrow 0$

Limit of PDE model

Evolutionary scale:

- ▶ Time scale $T \rightarrow \infty$
- ▶ Mutation probability is small: $m(y - y') \rightarrow \mathcal{N}(0, \sigma^2)$, where $\sigma \rightarrow 0$

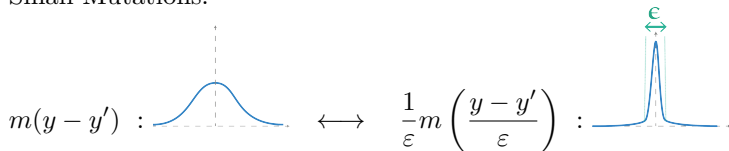
Solution:

Introduce a rescaling parameter $\varepsilon > 0$:

- ▶ $t \rightarrow \frac{t}{\varepsilon}$
- ▶ $m(y - y') \rightarrow \frac{1}{\varepsilon} m\left(\frac{y - y'}{\varepsilon}\right)$

Shape of solutions

► Small Mutations:



► Large time:

t
Ecological scale



$\frac{t}{\epsilon}$
Evolutionary scale

$\partial_t \longleftrightarrow \epsilon \partial_t$

► Competition \longrightarrow Saturation: total population $\rho(t)$ is bounded.

► Selection of the fittest trait $\bar{x} \implies f(t, x)$ converges to a Gaussian centered at \bar{x} , with same variance as m (mutation kernel)

PDE: The rescaled model

Rescaled PDE model is given by:

$$\left\{ \begin{array}{l} \varepsilon \partial_t f_\varepsilon(t, x) = -(d(x) + C\rho_\varepsilon(t))f_\varepsilon(t, x) + \int_{\mathbb{R}} \overbrace{m(z)b(x + \varepsilon z)}^{z = \frac{x-y}{\varepsilon}} f_\varepsilon(t, x + \varepsilon z) dz \\ \quad + f_\varepsilon(t, x) \int_{\mathbb{R}} \tau(x - y) \frac{f_\varepsilon(t, y)}{\rho_\varepsilon(t)} dy, \\ \rho_\varepsilon(t) = \int_{\mathbb{R}} f_\varepsilon(t, x) dx, \quad + \text{init. conditions.} \end{array} \right.$$

Hopf-Cole transformation

Main assumption:

f_ε is Gaussian with a variance ε .

$$\boxed{f_\varepsilon(t, x) = e^{\frac{u_\varepsilon(t, x)}{\varepsilon}}.}$$



$$\begin{aligned} \partial_t u_\varepsilon = & -(d(x) + \rho_\varepsilon(t)) + \int_{\mathbb{R}^d} m(z) b(x + \varepsilon z) e^{\frac{u_\varepsilon(t, x + \varepsilon z) - u_\varepsilon(t, x)}{\varepsilon}} dz \\ & + \int_{\mathbb{R}} \tau(x - y) \frac{f_\varepsilon(t, y)}{\rho_\varepsilon(t)} dy. \end{aligned}$$

The limiting Hamilton-Jacobi equation

►
$$f_\varepsilon(t, x) = e^{\frac{u_\varepsilon(t, x)}{\varepsilon}}.$$

► $u_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} u$, where u is the viscosity solution of:

$$\partial_t u = -(d(x) + \rho(t)) + b(x) \int_{\mathbb{R}} M(z) e^{z \cdot \nabla_x u} dz + \tau(x - \bar{x}(t))$$

► $\sup_{x \in \mathbb{R}} u(t, \cdot) \leq 0$, with " $<$ " \Leftrightarrow extinction.

► Population concentrates on $\bar{x}(t) = \operatorname{argmax} u(t, \cdot)$

► Non-standard drift term

Hamilton-Jacobi simulations

Target equation:

$$\begin{aligned}\partial_t u_\varepsilon = & -(d(x) + \rho_\varepsilon(t)) + \int_{\mathbb{R}^d} m(z)b(x + \varepsilon z)e^{\frac{u_\varepsilon(t, x + \varepsilon z) - u_\varepsilon(t, x)}{\varepsilon}} dz \\ & + \int_{\mathbb{R}} \tau(x - y) \frac{f_\varepsilon(t, y)}{\rho_\varepsilon(t)} dy.\end{aligned}$$

$$\boxed{\varepsilon \rightarrow 0}$$



$\boxed{\text{Refine grid in } x : \Delta_x \rightarrow 0}$



Explicit scheme not consistent

Hamilton-Jacobi simulations

Asymptotic Preserving (AP) scheme (Jin, 1999, Klar, 1998, 1999)

- avoid the increase of computational cost,
- ensure the scheme approaches the limit H-J equation for small ε

$$\begin{array}{ccc} P_\varepsilon & \xrightarrow{\varepsilon \rightarrow 0} & P_0 \\ \uparrow \scriptstyle \begin{smallmatrix} 0 \\ \hline \uparrow \\ h \end{smallmatrix} & & \uparrow \scriptstyle \begin{smallmatrix} 0 \\ \hline \uparrow \\ h \end{smallmatrix} \\ S_\varepsilon^h & \xrightarrow{\varepsilon \rightarrow 0} & S_0^h \end{array}$$

Hamilton-Jacobi simulations: numerical AP scheme

For $u(t, x) = \varepsilon \ln f(t, x)$ we compute:

- Competition term: **ρ is implicit.** Recall: $\rho^{n+1} = \sum_i e^{\frac{u_i^{n+1}}{\varepsilon}}$.

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = -\rho^{n+1} - D_i + B_i^n + T_i^n$$

- Death rate term: **explicit**
► Birth rate term: **linear interpolation**
► Transfer rate term

Remarking that for the exponential computations we make

$$\frac{f^\varepsilon(t^n, y)}{\rho^\varepsilon(t^n)} = \frac{e^{u^\varepsilon(t^n, y)/\varepsilon}}{\int_{\mathbb{R}} e^{u^\varepsilon(t^n, z)/\varepsilon} dz} = \frac{e^{(u^\varepsilon(t^n, y) - \max_x u^\varepsilon(t^n, x))/\varepsilon}}{\int_{\mathbb{R}} e^{(u^\varepsilon(t^n, z) - \max_x u^\varepsilon(t^n, x))/\varepsilon} dz},$$

Hamilton-Jacobi simulations: numerical AP scheme

For $u(t, x) = \varepsilon \ln f(t, x)$ we compute:

- Competition term: **ρ is implicit.** Recall: $\rho^{n+1} = \sum_i e^{\frac{u_i^{n+1}}{\varepsilon}}$.

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = -\rho^{n+1} - D_i + B_i^n + T_i^n$$

- Death rate term: **explicit**
- Birth rate term: **linear interpolation**
- Transfer rate term

Remarking that for the exponential computations we make

$$\frac{f^\varepsilon(t^n, y)}{\rho^\varepsilon(t^n)} = \frac{e^{u^\varepsilon(t^n, y)/\varepsilon}}{\int_{\mathbb{R}} e^{u^\varepsilon(t^n, z)/\varepsilon} dz} = \frac{e^{(u^\varepsilon(t^n, y) - \max_x u^\varepsilon(t^n, x))/\varepsilon}}{\int_{\mathbb{R}} e^{(u^\varepsilon(t^n, z) - \max_x u^\varepsilon(t^n, x))/\varepsilon} dz},$$

Hamilton-Jacobi simulations: numerical AP scheme

For $u(t, x) = \varepsilon \ln f(t, x)$ we compute:

- Competition term: **ρ is implicit.** Recall: $\rho^{n+1} = \sum_i e^{\frac{u_i^{n+1}}{\varepsilon}}$.

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = -\rho^{n+1} - D_i + B_i^n + T_i^n$$

- Death rate term: **explicit**
- Birth rate term: **linear interpolation**
- Transfer rate term

Remarking that for the exponential computations we make

$$\frac{f^\varepsilon(t^n, y)}{\rho^\varepsilon(t^n)} = \frac{e^{u^\varepsilon(t^n, y)/\varepsilon}}{\int_{\mathbb{R}} e^{u^\varepsilon(t^n, z)/\varepsilon} dz} = \frac{e^{(u^\varepsilon(t^n, y) - \max_x u^\varepsilon(t^n, x))/\varepsilon}}{\int_{\mathbb{R}} e^{(u^\varepsilon(t^n, z) - \max_x u^\varepsilon(t^n, x))/\varepsilon} dz},$$

Hamilton-Jacobi simulations: numerical AP scheme

For $u(t, x) = \varepsilon \ln f(t, x)$ we compute:

- Competition term: **ρ is implicit.** Recall: $\rho^{n+1} = \sum_i e^{\frac{u_i^{n+1}}{\varepsilon}}$.

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = -\rho^{n+1} - D_i + B_i^n + T_i^n$$

- Death rate term: **explicit**
- Birth rate term: **linear interpolation**
- Transfer rate term

Remarking that for the exponential computations we make

$$\frac{f^\varepsilon(t^n, y)}{\rho^\varepsilon(t^n)} = \frac{e^{u^\varepsilon(t^n, y)/\varepsilon}}{\int_{\mathbb{R}} e^{u^\varepsilon(t^n, z)/\varepsilon} dz} = \frac{e^{(u^\varepsilon(t^n, y) - \max_x u^\varepsilon(t^n, x))/\varepsilon}}{\int_{\mathbb{R}} e^{(u^\varepsilon(t^n, z) - \max_x u^\varepsilon(t^n, x))/\varepsilon} dz},$$

Hamilton-Jacobi equation: computation⁷ of ρ

1. Recall: $\rho^{n+1} = \sum_i e^{\frac{u_i^{n+1}}{\varepsilon}}$
2. Define: $A_i^n := u_i^n + \Delta_t(-d(x_i) + B_i^n + T_i^n)$ and $A_{i_0}^n = \max_i A_i^n$
3. By definition of a scheme: $\rho^{n+1} = \Delta_x e^{-\Delta_t \rho^{n+1/\varepsilon}} \sum_{i=0}^{N_x-1} e^{A_i^n/\varepsilon}$
4. Note that ρ^{n+1} is a solution of the following equivalent equations:

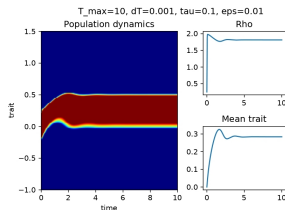
$$g_1(y) = ye^{\Delta_t y/\varepsilon} - \Delta_x e^{A_i^n/\varepsilon} \sum_{i=0}^{N_x-1} e^{(A_i^n - A_{i_0}^n)/\varepsilon}$$

$$g_2(y) = -\varepsilon \ln(y) - \Delta_t y + \varepsilon \ln(\Delta_x) + A_{i_0}^n + \varepsilon \ln \left(\sum_{i=0}^{N_x-1} e^{(A_i^n - A_{i_0}^n)/\varepsilon} \right)$$

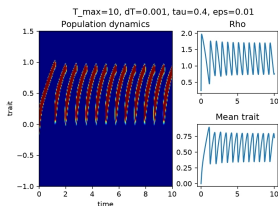
5. Then we find ρ^{n+1} by solving either g_1 or g_2 by a Newton's method (**with derivatives computed analytically**).

⁷Courtesy of H  l  ne Hivert

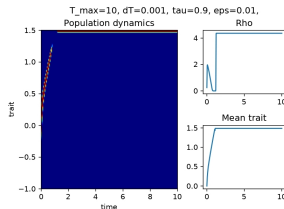
Hamilton-Jacobi simulations (PDE for $\varepsilon \ll 1$):



(a) Stabilization: $\tau = 0.1$



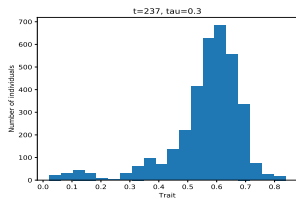
(b) Cycles: $\tau = 0.4$



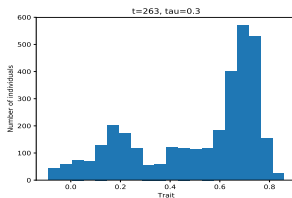
(c) Extinction: $\tau = 0.9$

Figure 4: Behavior of the population dynamics as the mutation rate τ is changing.

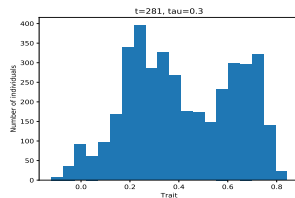
Evolutionary rescue



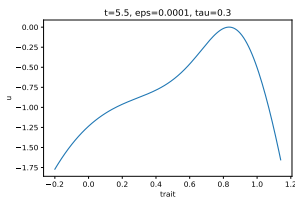
(a) $t = 237$



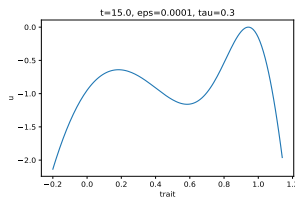
(b) $t = 263$



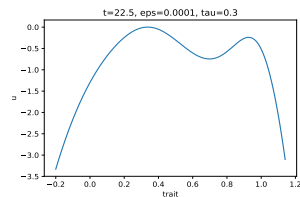
(c) $t = 281$



(d) $t = 5.5$



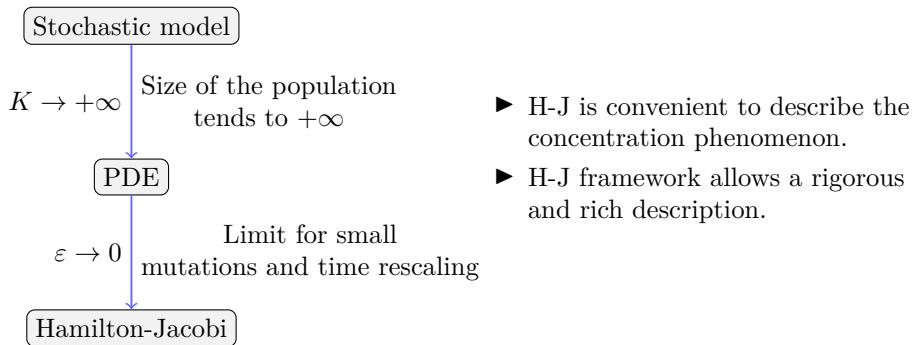
(e) $t = 15.0$



(f) $t = 22.5$

Figure 5: Stochastic/HJ. $\epsilon = 10^{-4}$ and $\tau = 0.3$

Conclusion

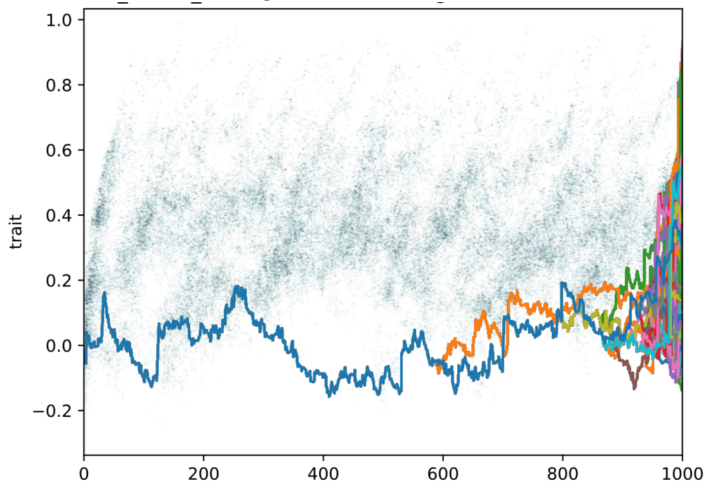


Paper in preparation:

V. Calvez, S.F. Iglesias, H. Hivert, S. Méléard, A. Melnykova, S. Nordmann (2018) "Horizontal gene transfer: numerical comparison between stochastic and deterministic approaches"

Open questions

1. Theoretical results: threshold for oscillation, extinction, stabilization, computations of period of oscillations etc.
2. Reconstruction of lineages.



References

- Billiard, S., Collet, P., Ferrière, R., Méléard, S., and Tran, V. C. (2016a). The effect of competition and horizontal trait inheritance on invasion, fixation, and polymorphism. *J. Theoret. Biol.*, 411:48–58.
- Billiard, S., Collet, P., Ferrière, R., Méléard, S., and Tran, V. C. (2016b). Stochastic dynamics for adaptation and evolution of microorganisms. *arXiv preprint arXiv:1610.00983*.
- Billiard, S., Ferrière, R., Méléard, S., and Tran, V. C. (2015). Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *J. Math. Biol.*, 71(5):1211–1242.
- Ferrière, R. and Tran, V. C. (2009). Stochastic and deterministic models for age-structured populations with genetically variable traits. In *CANUM 2008*, volume 27 of *ESAIM Proc.*, pages 289–310. *EDP Sci., Les Ulis*.
- Jin, S. (1999). Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454.
- Klar, A. (1998). An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit. *SIAM J. Numer. Anal.*, 35(3):1073–1094.
- Klar, A. (1999). An asymptotic preserving numerical scheme for kinetic equations in the low mach number limit. *SIAM J. Numer. Anal.*, 36(5):1507–1527.



This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 639638).

Thank you for your attention!