Modelling the evolution of asexuals: PDE, integro-differential and stochastic approaches

L Roques, BioSP, INRAE

With :
M-E Gil, F Hamel, F Lavigne, G Martin

1st Joint Meeting Brazil-France in Mathematics /
Organization Principles for Living Systems /
IMPA, Rio de Janeiro, July 15 – 19, 2019

Modelling evolutionary dynamics in asexuals

- **Main objectives:**

  - To **predict** the evolution of asexual organisms (viruses, bacterias, cancer lineages)

  - To **understand** complex interplay of drift, selection and mutation in asexuals

- **Challenge:** Management strategies of resistance emergence,

  World Health Organization describes antibiotic resistance as one of the biggest threats to global health, food security, and development today (WHO 2016).

- **ANR Project RESISTE:** *Evolutionary rescue, stochastic effects and interactions with environmental stress.* **Partnership** with Montpellier Institute of Evolutionary Sciences (experimental evolution of bacterias, theoretical models)

Agence Nationale de la Recherche ANR

isem Institut des Sciences de l'Evolution-Montpellier

Comments and basic definitions

- Our **method**. We follow the ***expected fitness distribution*** in asexual populations.

- **Fitness**. Genotype (=individual) reproductive success.

- **Drift.** *Random* sampling process of the individuals to form next generation

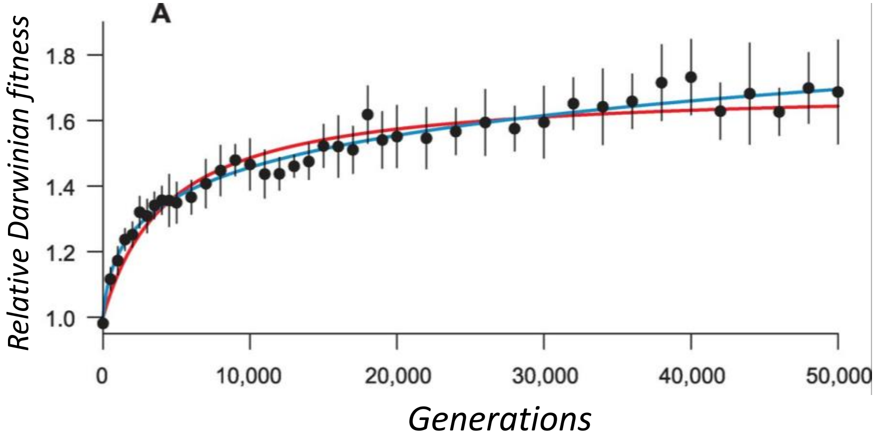
This is a blackboard

Long term evolution experiment (Lenski's experiment)

Fitness is measured in 12 initially identical populations of asexual *Escherichia coli* bacteria since 24 February 1988 (>70000 generations)

Each day, 1% of each population is transferred to a flask of fresh growth medium

Relative fitness is measured by competition with (frozen) ancestral bacteria



Dynamics of fitness distribution:
2 main biological frameworks / 3 approaches

A. Context-intependent effect of mutations (non-epistatic mutation effects on fitness)

- **I. Wright-Fisher stochastic individual-based model (used as a benchmark)**
- **II. Nonlocal transport PDEs and cumulant generating functions**
- **III. Integro-differential equations**

B. Context-dependent effect of mutations (epistatic effects induced by an optimum)

- **I. Fisher's Geometrical model and Wright-Fisher stochastic IBM**
- **II. Nonlocal nonlinear transport PDEs and cumulant generating functions**
- **III. Integro-differential equations**

C. Discussion/extensions: anisotropic effects of mutations

Part A: context-independent effect of mutations

A.I. Wright-Fisher microscopic model

- Standard model. Many references: see e.g. [*Rice 2004, Lambert 2008*]
- Interplay of selection, drift and mutation, in a population of constant size N

Generation t

N individuals

Fitness of the indiv.: $\mathbf{m}_t = (m_t^1, \ldots, m_t^N)$

Reproduction/selection

N individuals are sampled with replacement

Nb offsprings: $\sim \mathrm{Multi}(N, \exp(\mathbf{m}_t))$

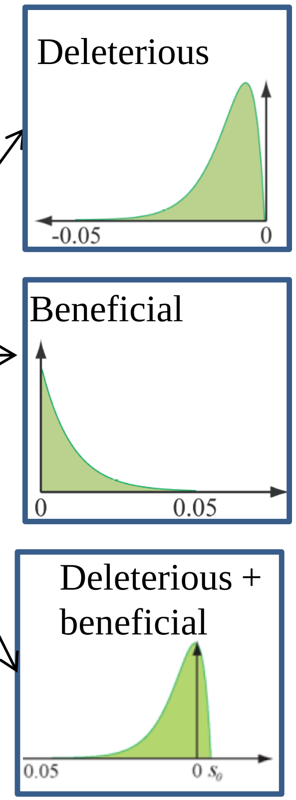
Mutation

Nb mutations/indiv. : $k \sim \mathcal{P}(U)$

Mutation effect on fitness: $m \to m + s$

Generation t+1

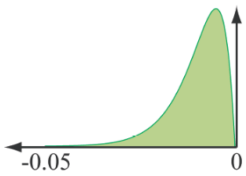
Mutation Kernel *J*



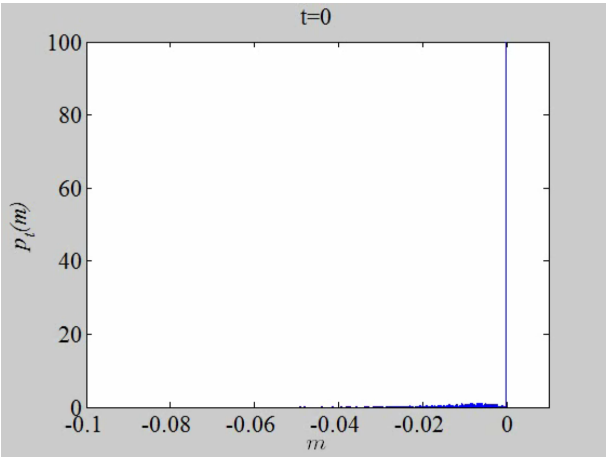
A.I. Wright-Fisher microscopic model

Numerical simulations with $N = 10^5$

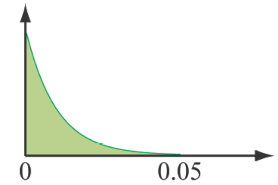
Deleterious mutation kernel



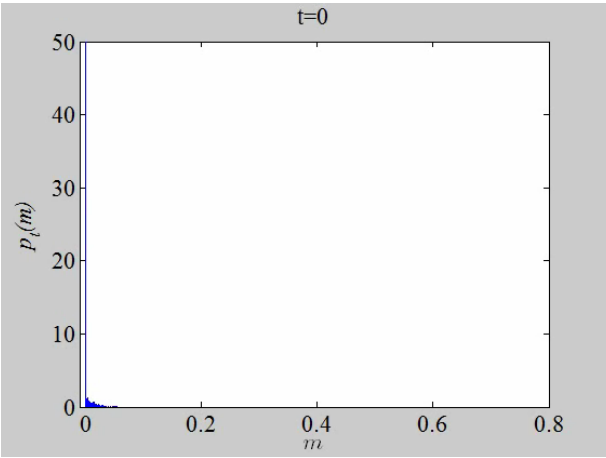
Distribution of fitnesses:



Beneficial mutation kernel



Distribution of fitnesses:



A.II. Cumulant generating functions approach
Basics

Main objective: to obtain an analytic description of the microscopic WF model

Cumulant generating function (CGF).

$$C(t,z) = \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{m_i z}\right), \quad z \geq 0.$$

Important properties.

Mean $\overline{m}_t = \partial_z C(t,0)$ and variance $V_t = \partial_{zz} C(t,0)$.

A.II. Cumulant generating functions approach

Effect of mutations on $C(t,z)$: formal derivation

- Given the distribution of fitness at time t, we define

$$M(t,z) = \frac{1}{N}\sum_{i=1}^{N} e^{m_i z}$$

- $\Delta_{mut} M(t,z)$ =expected variation in M due to mutation during δt

- Poisson number of mutations: $P(\text{nb mut=k}) = e^{-N\,U\,\delta t}\frac{(N\,U\,\delta t)^k}{k!}$

- As $\delta t \ll 1$, $P(\text{nb mut=1}) \approx N\,U\,\delta t$, $P(\text{nb mut} > 1) \approx 0$.

- Mutation of effect s, parent of fitness m:
  $\Delta_{mut} M(t,z|s,m) \approx N\,U\,\delta t\left(e^{(s+m)z} - e^{m\,z}\right)/N = U\,\delta t\, e^{m\,z}\,(e^{s\,z} - 1)$

A.II. Cumulant generating functions approach

Effect of mutations on $C(t,z)$: formal derivation

- Taking the expectation over the distribution of mutation effects on fitness (mutation kernel J):

$$\Delta_{mut} M(t, z|m) = \int_{-\infty}^{\infty} \Delta_{mut} M(t, z|s, m)\, J(s)\, ds \approx U\, \delta t\, e^{m\,z} \left(\int_{-\infty}^{\infty} J(s) e^{s\,z}\, ds - 1 \right)$$

- Summing over the parents, we get :

$$\Delta_{mut} M(t, z) \approx U\, \delta t\, M(t, z) \left(\int_{-\infty}^{\infty} J(s) e^{s\,z}\, ds - 1 \right)$$

- Mutational contribution to the CGF $C(t,z)$ (formally):

$$\boxed{\frac{\Delta_{mut} C(t, z)}{\delta t} \approx U \left(\int_{-\infty}^{\infty} J(s) e^{s\,z}\, ds - 1 \right)}$$

A.II. Cumulant generating functions approach

Formal derivation of a PDE on $E(C(t,z))$

Effects of selection and drift described by a system of stochastic differential equations with branching (large N approximation, [*Ewens 2004*]):

$$dp_i = (m_i - \overline{m}(t))p_i dt + \sum_{j \neq i} \sqrt{\frac{1}{N} p_i \, p_j} dB_t^{ij}$$

Infinitesimal generator:

$$\mathcal{D}_\varphi(\mathbf{p}) = \sum_{i=1}^N p_i(m_i - \overline{m}_t)\frac{\partial \varphi(\mathbf{p})}{\partial p_i} + \frac{1}{2N}\left(\sum_{i=1}^N p_i(1-p_i)\frac{\partial^2 \varphi(\mathbf{p})}{\partial p_i^2} - 2\sum_{i=1}^N \sum_{j=i+1}^N p_i p_j \frac{\partial^2 \varphi(\mathbf{p})}{\partial p_i \partial p_j}\right).$$

PDE obtained by Feynman-Kac theorem with $\mathcal{C}(t,\boldsymbol{p},z) = E^{\boldsymbol{p}_0}[C(t,\boldsymbol{p},z)]$ + mutation:

$\mathcal{C}(t,z)$: *expected* CGF among replicate populations

$$\partial_t \mathcal{C}(t,z) = \underbrace{\partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0)}_{\textit{selection}} + U \underbrace{\left(\int_{\mathbb{R}} J(s) e^{sz} ds - 1\right)}_{\textit{mutation}} + \frac{1}{2N}\underbrace{\left(1 - E^{\mathbf{p}_0}\left[e^{C(t,2z)-2C(t,z)}\right]\right)}_{\textit{genetic drift}}.$$

A.II. Cumulant generating functions approach
Fisher's fundamental theorem

- **Fisher's fundamental theorem (without mutations):**

« The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time » (Fisher 1930)

Derivate w.r.t. z the expression

$$\partial_t \mathcal{C}(t,z) = \partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0) + U\left(\int_{\mathbb{R}} J(s)e^{sz}ds - 1\right) + \frac{1}{2N}\left(1 - E^{\mathbf{P}_0}\left[e^{C(t,2z)-2C(t,z)}\right]\right).$$

leads to:

$$\partial_t \partial_z \mathcal{C}(t,z) = \partial_{zz}\mathcal{C}(t,z) + U\int_{\mathbb{R}} sJ(s)e^{sz}ds + \frac{1}{2N}E^{\mathbf{P}_0}\left[(2\partial_z C(t,2z) - 2\partial_z C(t,z))e^{C(t,2z)-2C(t,z)}\right]$$

Computing the result at $z = 0$:

$$\boxed{\underbrace{\partial_t \overline{m}(t)}_{\text{Rate of increase in expected mean fitness}} = \underbrace{\mathrm{Var}(p(t,\cdot))}_{\text{Expected variance In fitness}} + \underbrace{U\int_{\mathbb{R}} sJ(s)ds}_{\text{Mean effect of mutations on fitness}}}$$

A.II. Cumulant generating functions approach
Solution of the PDE

$\mathcal{C}(t,z)$: *expected* CGF among replicate populations

Large pop size assumption

$$\partial_t \mathcal{C}(t,z) = \underbrace{\partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0)}_{\textit{selection}} + U \underbrace{\left(\int_{\mathbb{R}} J(s) e^{sz} ds - 1 \right)}_{\textit{mutation}} + \frac{1}{2N} \underbrace{\cancel{\left(1 - E^{\mathbf{P}_0} \left[e^{C(t,2z) - 2C(t,z)} \right] \right)}}_{\textit{genetic drift}}.$$

Theorem (Martin, R, 2016)

The unique solution is: $\mathcal{C}(t,z) = C_0(z+t) - C_0(t) + U \int_0^t \beta(z+v) - \beta(v) dv,$

with: $\beta(z) = \int_{\mathbb{R}} J(s) e^{sz} ds.$

Consequences:

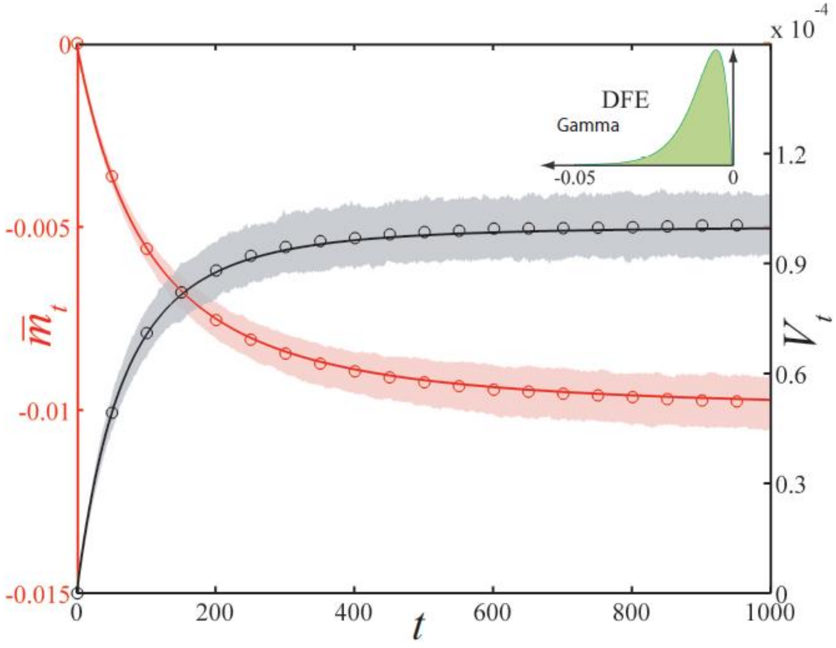
$$E^{\mathbf{P}_0}(\overline{m}_t) = C_0'(t) + U(\beta(t) - 1) \text{ and } E^{\mathbf{P}_0}(V_t) = C_0''(t) + U\,\beta'(t) - U\,\mu_J,$$

with: $\mu_J = \int_{\mathbb{R}} s J(s) ds.$

A.II. Cumulant generating functions approach
Numerical computations

Purely deleterious mutation kernels: very accurate results

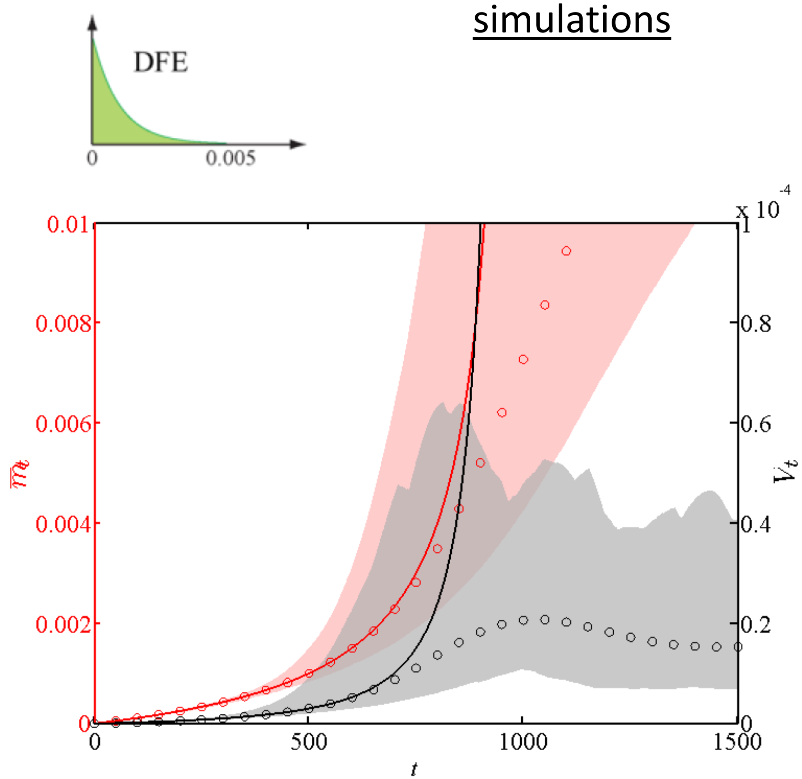
Mean, variance: PDE vs stochastic simulations

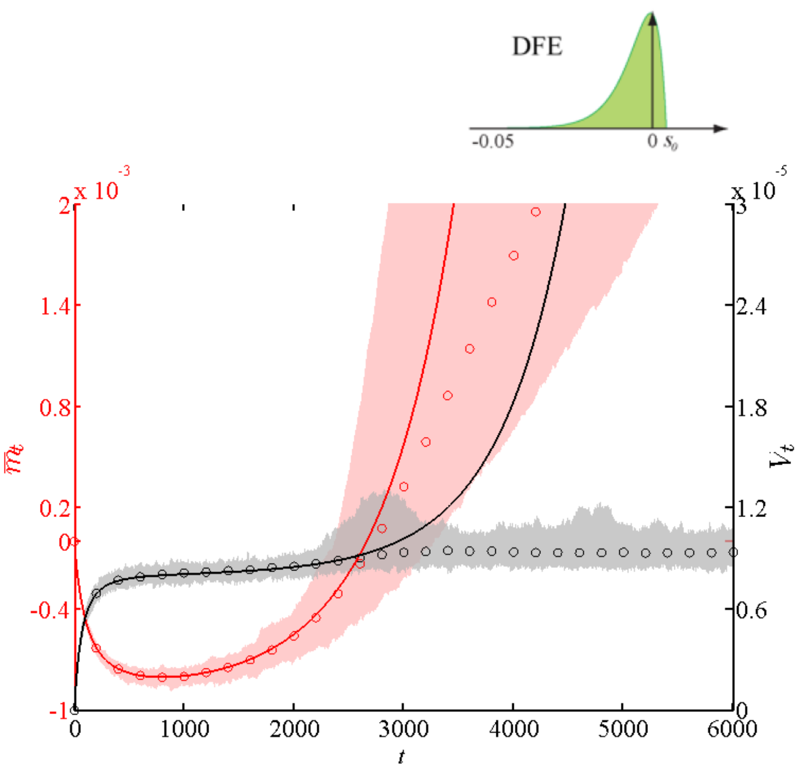


A.II. Cumulant generating functions approach
Numerical computations

Mutation kernels including benefical mutations: accurate at the beginning,

<u>Mean, variance: PDE vs stochastic simulations</u>





Incorrect behaviour after some time due to the neglected drift term

A.III. Integro-differential approach

Main objective: to connect our formal results on CGFs with the theory of IDEs

Standard model for the distribution of the fitness *m* in an infinite population (Gerrish et al., 2007; Sniegowski and Gerrish, 2010; Desai and Fisher, 2011):

$$\boxed{\partial_t p(t,m) = U\left(J \star p - p\right)(t,m) + p(t,m)\left(m - \overline{m}(t)\right),}$$

with $(J \star p - p)(t,m) = \displaystyle\int_{\mathbb{R}} J(m-y)(p(t,y) - p(t,m))\, dy,$

and $\overline{m}(t) = \int_{\mathbb{R}} y p(t,y)\, dy.$

Close to the ``replicator-mutator" eq. studied by [*Alfaro and Carles (2014)*]:

$$\partial_t p(t,m) = \partial_{mm} p(t,m) + p(t,m)\left(m - \overline{m}(t)\right).$$

Assumptions (necessary for the existence of a time-global solution):

$$\lim_{m \to \pm\infty} p_0(m)\, \mathrm{e}^{\alpha|m|} = 0, \int_{\mathbb{R}} J(y) e^{\alpha|y|} dy < +\infty \text{ for all } \alpha > 0.$$

INRA SCIENCE & IMPACT

A.III. Integro-differential approach

Results - existence

$$\partial_t p(t,m) = U\left(J \star p - p\right)(t,m) + p(t,m)\left(m - \overline{m}(t)\right),$$

Theorem (Gil, Hamel, Martin, R, 2017)

- Existence of a unique time-global solution
- the CGF is well-defined $\mathcal{C}(t,z) := \ln\left(\int_{\mathbb{R}} p(t,m)\mathrm{e}^{zm}\, dm\right)$,
- it satisfies the same equation as in part A.II of the talk:

$$\partial_t \mathcal{C}(t,z) = \partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0) + U\left(\int_{\mathbb{R}} J(s)e^{sz}ds - 1\right).$$

- the unique solution satisfies:

$$\mathcal{C}(t,z) = C_0(z+t) - C_0(t) + U\int_0^t \beta(z+v) - \beta(v)dv,$$

with: $\beta(z) = \int_{\mathbb{R}} J(s)e^{sz}ds.$

A.III. Integro-differential approach: results – large time

Theorem (Gil, Hamel, Martin, R, 2017)

Case 1: deleterious mutations only

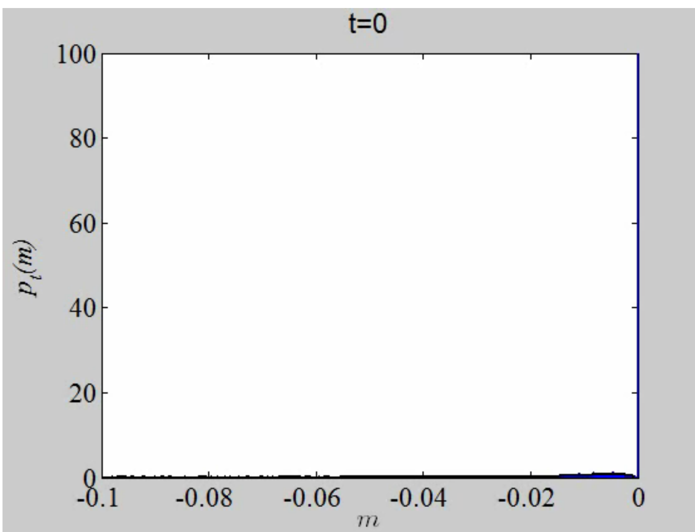
The distribution $p(t,.)$ converges weakly to a distribution p_∞ (known CGF)

Case 2: presence of beneficial mutations

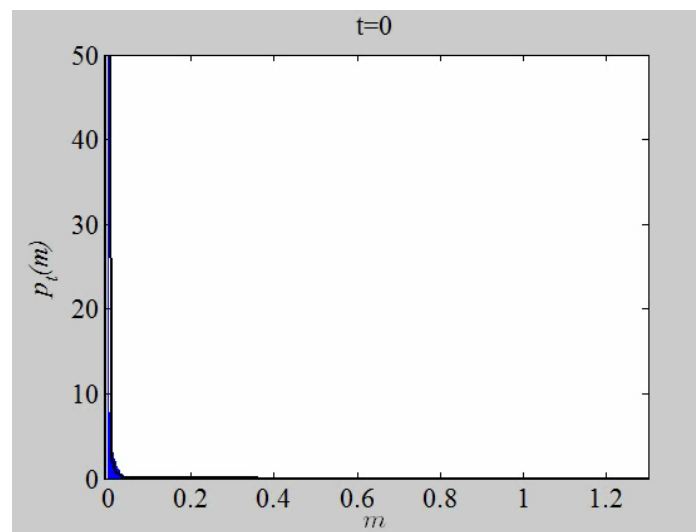
The mean fitness $\overline{m}(t)$ increases exponentially fast. $V(t) \to +\infty$.

$\to$ unrealistic

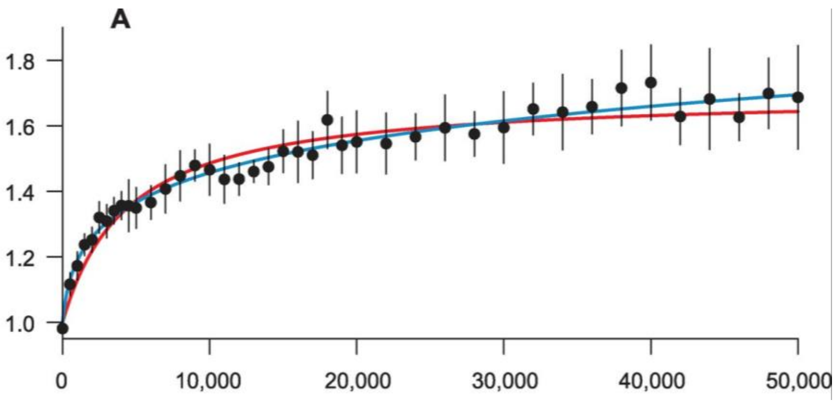
Case 1:



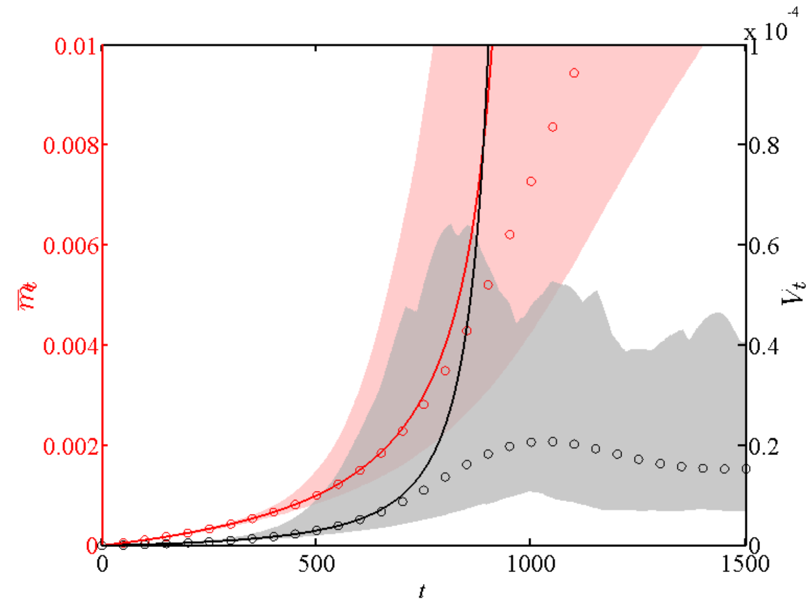
Case 2:



Consistence with Lenski's experiment



≠



Part B: context-dependent effect of mutations

B.I. Fisher's Geometrical Model (FGM)

The FGM is a phenotype-fitness landscape model:

- **Describes the relationships between (n-dimensional) phenotype and fitness**

- **Takes into account the existence of a unique phenotype optimum $\boldsymbol{g}^* \in \mathbb{R}^{\boldsymbol{n}}$ (here, $\boldsymbol{g}^* = \boldsymbol{0}$ for the sake of simplicity)**

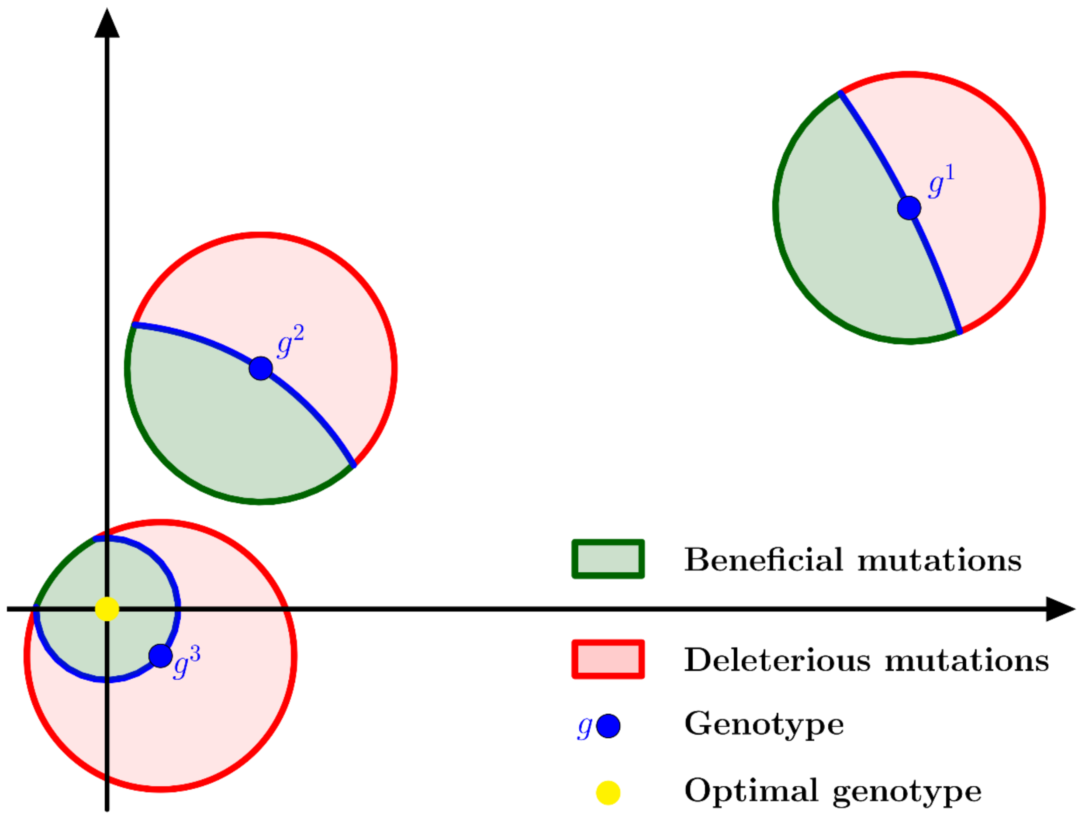
Assumptions:

- **Each individual is characterized by a « phenotype » $\boldsymbol{g} \in \mathbb{R}^{\boldsymbol{n}}$**

- **The relative fitnesses are $\boldsymbol{m}(\boldsymbol{g}) = -\frac{\|\boldsymbol{g}-\boldsymbol{g}^*\|^2}{2}$**

B.I. Fisher's Geometrical Model (FGM)

Mutation effects on fitness:



B.I. Fisher's Geometrical Model (FGM)

Example: Gaussian FGM. One of the most standard models of evolutionary quantitative genetics [Kimura 1965, Lande 1980]

$$\mathbf{g}_i \to \mathbf{g}_i + d\mathbf{g}_i \text{ with } d\mathbf{g}_i \sim \mathcal{N}(0, \lambda\, I_n)$$

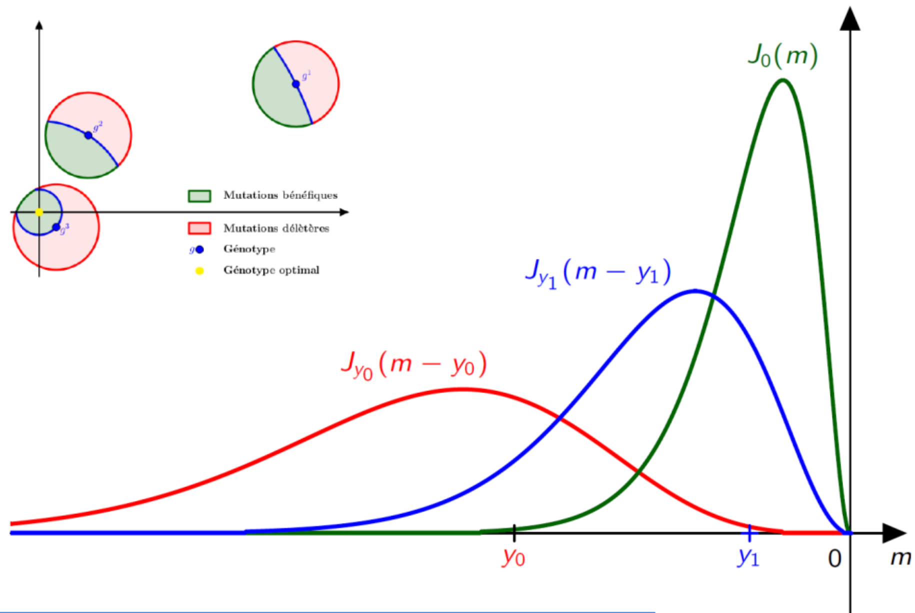
Stochastic representation of the mutation effects on fitness:

$$m \to m + s$$

with

$$s|m \sim -m - \frac{\lambda}{2}\,\chi^2\left(n, -\frac{2m}{\lambda}\right).$$

[Martin, Lenormand 2015]



Mutation kernel J_m, represented by its moment generating function:

$\int_{-\infty}^{\infty} J_m(s) e^{s\,z} ds = M_\star(z) e^{\omega(z) m}$,

with $M_\star(z) = \frac{1}{(1+\lambda\, z)^{n/2}}$ and $\omega(z) = \frac{-\lambda z^2}{1+\lambda z}$. [Martin, Lenormand 2015]

B.I. FGM + Wright-Fisher microscopic model

Generation t

N individuals

Phenotypes: $\boldsymbol{g_i} \in \mathbb{R}^n$

Fitnesses:

$m_t^i = -\|\boldsymbol{g_i} - \boldsymbol{g^*}\|^2/2$

Fitness of the indiv. :

$\mathbf{m}_t = (m_t^1, \ldots, m_t^N)$

Reproduction/selection

N individuals are sampled with replacement in the previous generation

Nb offsprings:

$\sim \mathrm{Multi}(N, \exp(\mathbf{m}_t))$

Mutation

Nb mutations/indiv. :

$k \sim \mathcal{P}(U)$

Mutation effect on phenotype:

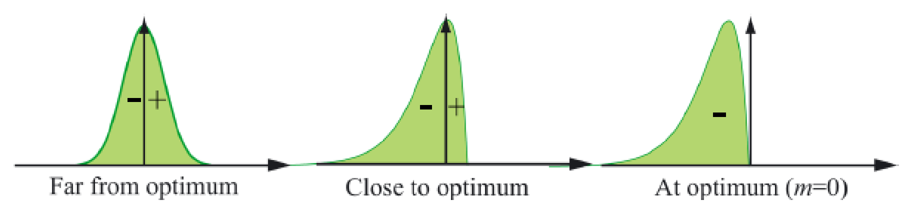
$\boldsymbol{g_i} \rightarrow \boldsymbol{g_i} + \boldsymbol{dg_i}$

Ex: $\boldsymbol{dg_i} \sim \mathcal{N}(0, \lambda\, \boldsymbol{I_n})$

Mutation effect on fitness: $m_i \rightarrow m_i + s$

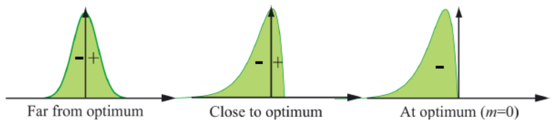
Generation t+1

Context-dependent mutation kernel

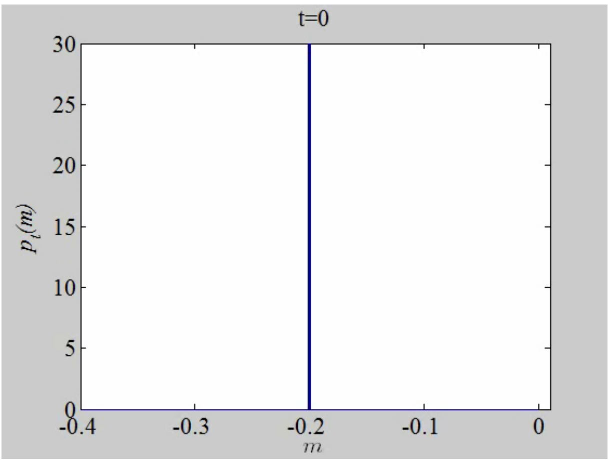


B.I. Numerical simulations

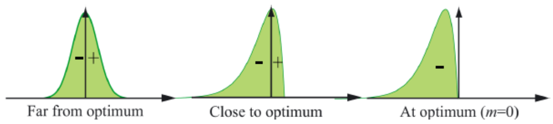
Context-dependent kernel



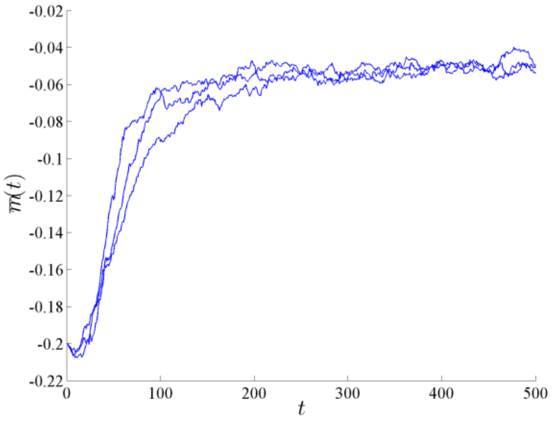
Distribution of fitness:



Context-dependent kernel



Mean fitness: $\overline{m}(t)$



B.II. Cumulant generating functions approach

Cumulant generating function (CGF).

$$C(t,z) = \ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{m_i z}\right), \; z \geq 0.$$

-Mean $\overline{m}_t = \partial_z C(t,0)$ and variance $V_t = \partial_{zz} C(t,0)$.

-Weight of the optimal phenotype: $p(t,0) = \exp(C(t,+\infty))$.

-Upper bound of the support: $\sup \operatorname{supp} p(t,\cdot) = \partial_z C(t,+\infty)$.

B.II. Cumulant generating functions approach

Effect of mutations

Same type of arguments as in the context-independent case, with

$$M(t,z) = \tfrac{1}{N}\sum_{i=1}^{N} e^{m_i z}$$

and the assumption: $\displaystyle\int_{-\infty}^{\infty} J_m(s) e^{s\,z}\, ds = M_\star(z)\, e^{\omega(z)\,m}.$

we get: $$\boxed{\frac{\Delta_{mut} C(t,z)}{\delta t} \approx U\left(M_\star(z)\, e^{C(t,z+\omega(z))-C(t,z)} - 1\right)}$$

B.II. Cumulant generating functions approach

Nonlocal nonlinear PDE on the expected CGF

$\mathcal{C}(t,z)$: *expected* CGF among replicate populations

Deterministic approximation: $N >> 1$.

$$\partial_t \mathcal{C}(t,z) = \underbrace{\partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0)}_{\textit{selection}} + \underbrace{U\left(M_\star(z) e^{\mathcal{C}(t,z+\omega(z)) - \mathcal{C}(t,z)} - 1\right)}_{\textit{mutation}}.$$

Approximate equation in the large U/λ regime:

$$\partial_t \mathcal{C}(t,z) \approx \alpha(z) \partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0) + \beta(z),$$

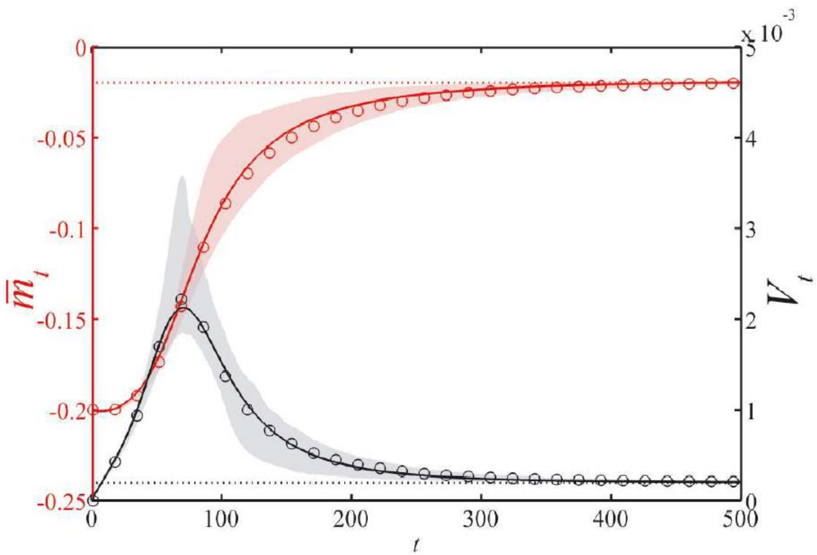
with $\alpha(z) = 1 + U M_\star(z)\omega(z)$ and $\beta(z) = U\left(M_\star(z) - 1\right)$

→ Analytic expression for $\mathcal{C}(t,z)$

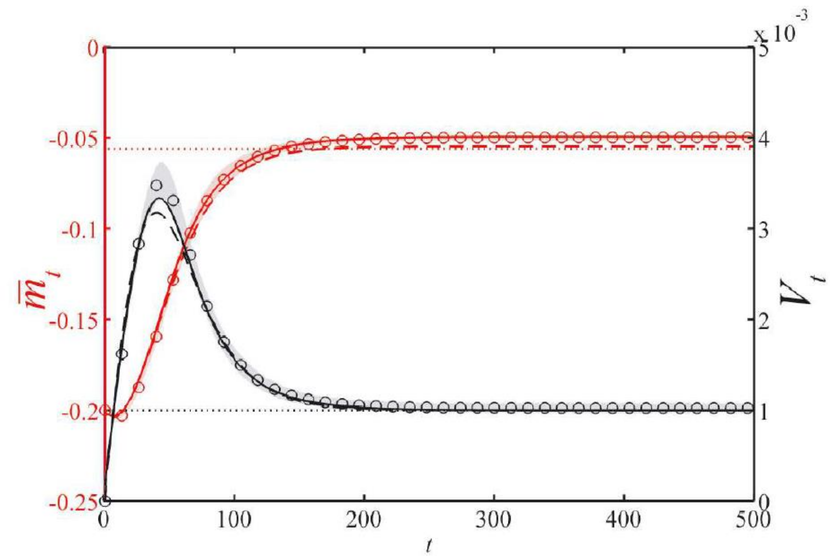
Gaussian FGM, Clonal initial pop $\quad \partial_z \mathcal{C}(t,0) = \overline{m}(t) \approx \frac{n}{2}\sqrt{U\lambda}\,\tanh(t\sqrt{U\lambda}) + m_0/\cosh^2(t\sqrt{U\lambda})$

B.II. Cumulant generating functions approach
Numerical computations

Numerical solution of the PDE
vs
approximate analytic solution
vs
Individual-based stochastic simulations



Small mutation rate (U=0.02)

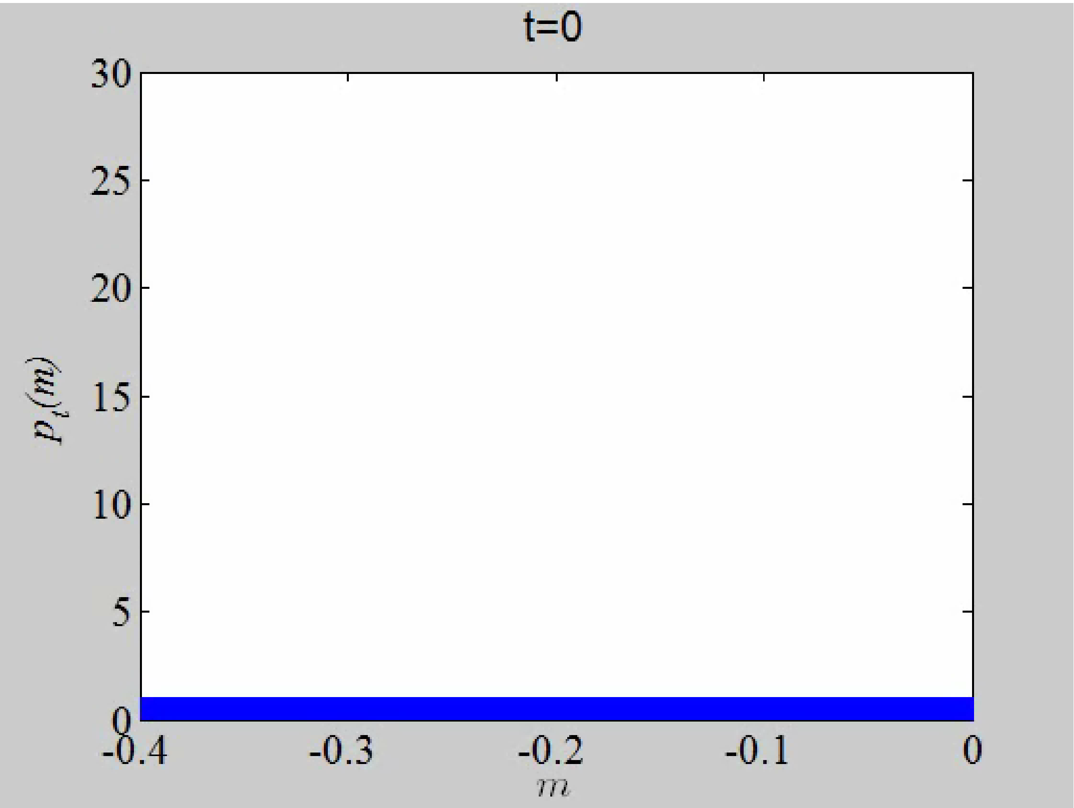


High mutation rate (U=0.1)

B.II. Cumulant generating functions approach

Numerical computations

Distribution: approximate analytic solution (large U/λ)
vs
Individual-based stochastic simulations



B.III. Integro-differential approach

The equation

Main objective: to connect our formal results on CGFs with the theory of IDEs

We consider the equation:

$$\partial_t p = \underbrace{U\left(J_y \circledast p - p\right)}_{\text{mutation}} + \underbrace{(m - \overline{m}(t))p}_{\text{selection}}$$

with $J_y \circledast p(t,m) = \displaystyle\int_{\mathbb{R}} J_y(m-y)p(t,y)\,dy$ and $\overline{m}(t) = \displaystyle\int_{\mathbb{R}} y p(t,y)\,dy.$

Assumptions: $\operatorname{supp} p_0 \subset (-\infty, 0]$, $\operatorname{supp} J_y \subset (-\infty, -y]$, for all $y \leq 0$.

B.III. Integro-differential approach

Results - existence

Theorem (Gil, Hamel, Martin, R, 2018)

- Existence of a unique time-global solution
- the CGF is well-defined $\mathcal{C}(t,z) := \ln\left(\int_{\mathbb{R}} p(t,m) \mathrm{e}^{zm}\, dm\right),$
- if $$\int_{-\infty}^{\infty} J_m(s) e^{s\,z}\, ds = M_\star(z)\, e^{\omega(z)\, m},$$

as in the Gaussian FGM, it satisfies the same equation as in part B.II of the talk:

$$\partial_t \mathcal{C}(t,z) = \partial_z \mathcal{C}(t,z) - \partial_z \mathcal{C}(t,0) + U\left(M_\star(z) e^{\mathcal{C}(t,z+\omega(z)-\mathcal{C}(t,z))} - 1\right).$$

B.III. Integro-differential approach

Results – stationary states

Convergence (not proved, this is an assumption here):

We assume that $(p(t,\cdot))_{t\geq 0}$ converges weakly towards a measure p_∞:

$$\lim_{t\to+\infty}\int_{\mathbb{R}}\phi(m)\,p(t,m)\,dm=\int_{\mathbb{R}}\phi(m)\,dp_\infty(m)$$

for all continuous function ϕ, s.t. $\phi(m)=O(m)$ as $m\to-\infty$.

Stationary equation

We define $\mathcal{C}_\infty(z)=\ln\left(\int_{-\infty}^{0}e^{zm}dp_\infty(m)\right)$, for all $z\geq 0$,

Then $\mathcal{C}'_\infty(z)-\mathcal{C}'_\infty(0)+U\left(e^{\mathcal{C}_\infty(z+\omega(z))-\mathcal{C}_\infty(z)}M_\star(z)-1\right)=0$

B.III. Integro-differential approach

Results – stationary states

Proposition (Gil, Hamel, Martin, R, 2018)

$$\overline{m}_\infty = C'_\infty(0) \geq -U.$$

Furthermore, writing p_∞ as a sum of two measures:

$$p_\infty(m) = (1-\rho)p^\star(m) + \rho\delta_{\{m=0\}}(m), \text{ for all } m \in \mathbb{R},$$

we get,

Proposition (Gil, Hamel, Martin, R, 2018)

$$\rho = 0 \text{ or } \overline{m}_\infty = -U$$

B.III. Integro-differential approach

Results – stationary states

Let $s_H^\star = \left(\int_{-\infty}^{0} \frac{J_0(s)}{|s|} ds \right)^{-1}$ the harmonic mean of J_0

(mutation kernel at the optimum)

Theorem (Gil, Hamel, Martin, R, 2018)

- if $s_H^\star = 0$
  - then $\rho = 0$,
  - additionnally, if $\lim_{z \to +\infty} (z + \omega(z)) < +\infty$ then $\overline{m}_\infty > -U$,
- if $s_H^\star \neq 0$ then
  - if $U \leq s_H^\star$ then $\rho > 0$ and $\overline{m}_\infty = -U$;
  - let $\tilde{U} := \inf \{ U \geq s_H^\star \text{ s.t. } \exists z_1 > 0 \text{ with } 1 + U\omega(z_1) M_\star(z_1) = 0 \}$,
    if $U > U_c$ then $\rho = 0$ and $\overline{m}(+\infty) > -U$.

B.III. Integro-differential approach

Results – stationary states, Gaussian FGM

We recall that:

$$\int_{-\infty}^{\infty} J_m(s) e^{s\,z}\, ds = M_\star(z)\, e^{\omega(z)\, m}.$$

In all cases, we have $s_H^\star = \left(\int_0^{+\infty} M_\star(z) dz \right)^{-1}$

For the Gaussian FGM: $M_\star(z) = \frac{1}{(1+\lambda\, z)^{n/2}}$ and $\omega(z) = \frac{-\lambda z^2}{1+\lambda z}$.

$s_H^\star = 0 \iff \int_0^{+\infty} M_\star(z) dz = +\infty \iff n \leq 2.$

$$\lim_{z \to +\infty} (z + \omega(z)) = 1/\lambda < \infty$$

For $n > 2$, $s_H^\star = \lambda(n-2)/2$

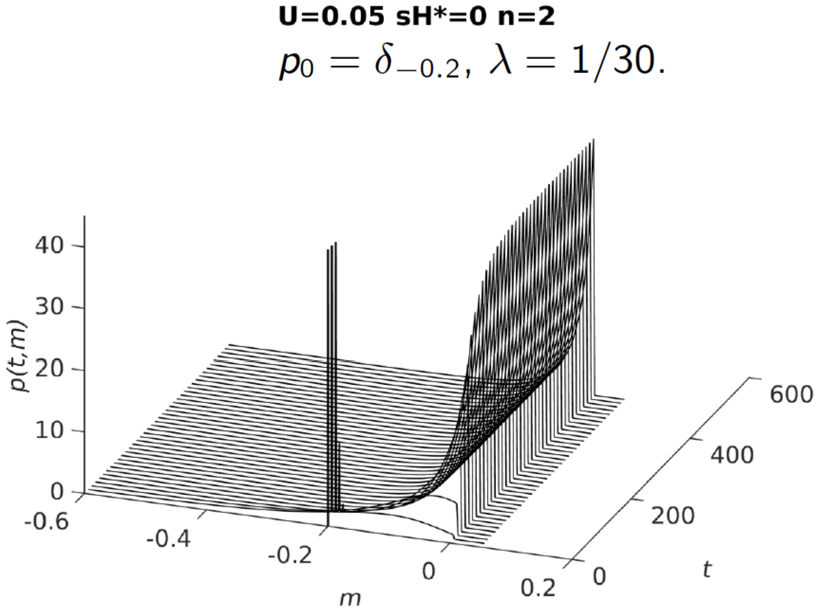
For $n > 2$, $\tilde{U} = \frac{\lambda}{4} \frac{(n/2+1)^{n/2+1}}{(n/2-1)^{n/2-1}}$

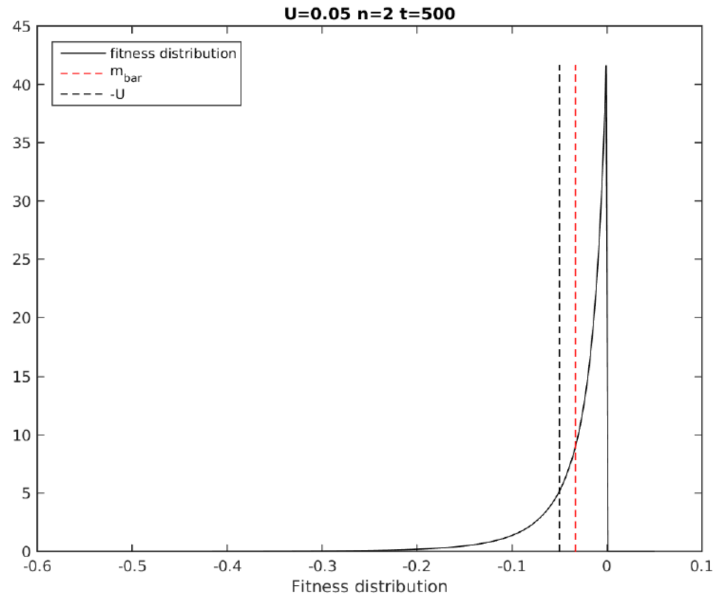
B.III. Integro-differential approach

Results – stationary states, Gaussian FGM

Corollary (Gil, Hamel, Martin, R, 2018)

Gaussian FGM, $n \leq 2$: $\rho = 0$ and $\overline{m}_\infty > -U$.



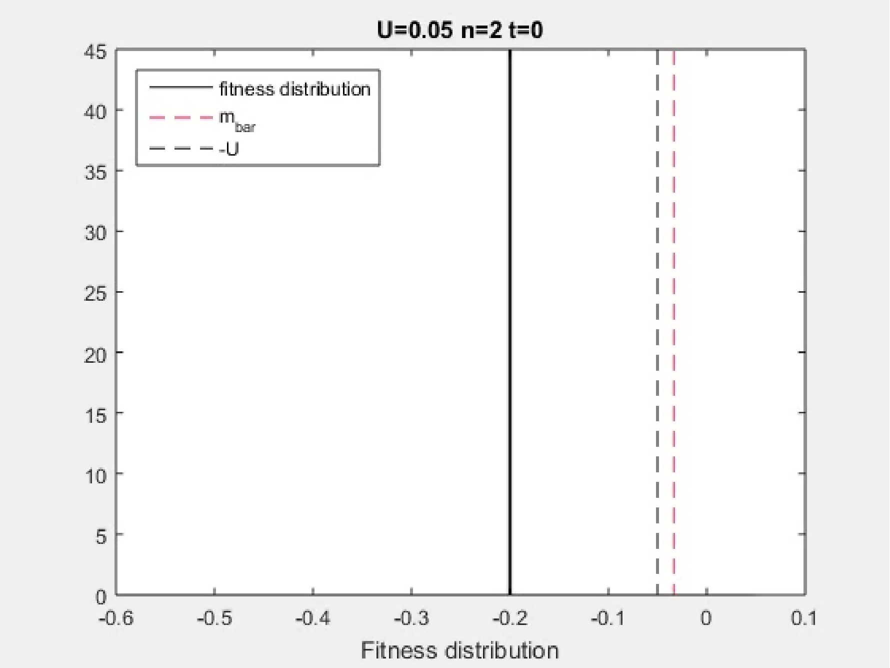


B.III. Integro-differential approach

Results – stationary states, Gaussian FGM

Corollary (Gil, Hamel, Martin, R, 2018)

Gaussian FGM, $n \leq 2$: $\rho = 0$ and $\overline{m}_\infty > -U$.

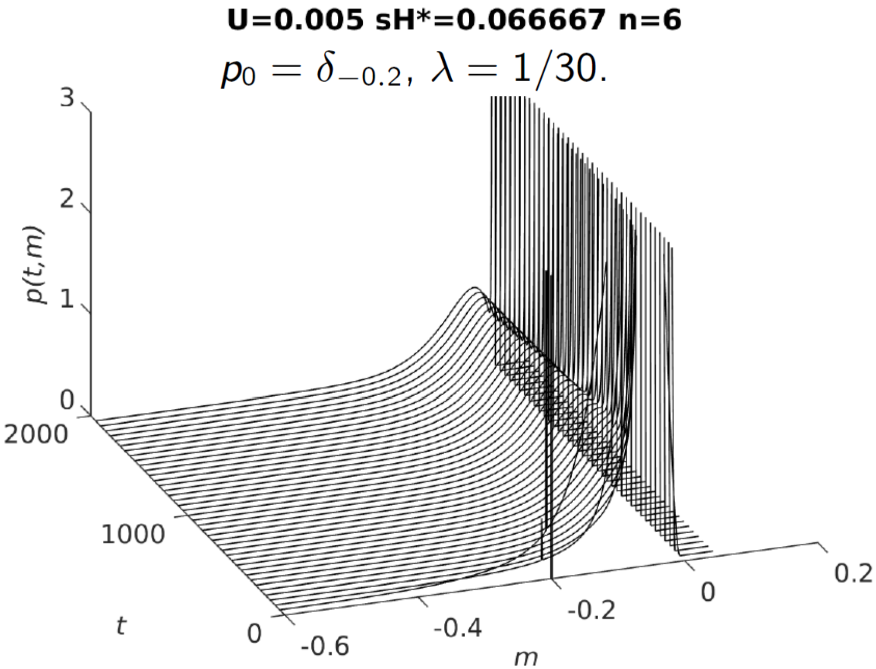


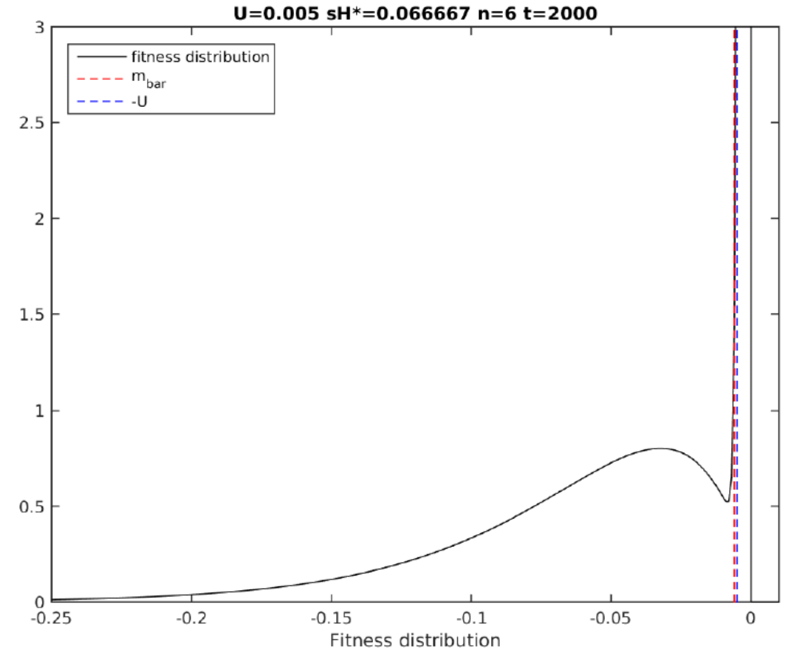
B.III. Integro-differential approach

Results – stationary states, Gaussian FGM

Corollary (Gil, Hamel, Martin, R, 2018)

Gaussian FGM, $n > 2$ and $U < s_H^\star$: $\quad \rho > 0$ and $\overline{m}_\infty = -U.$





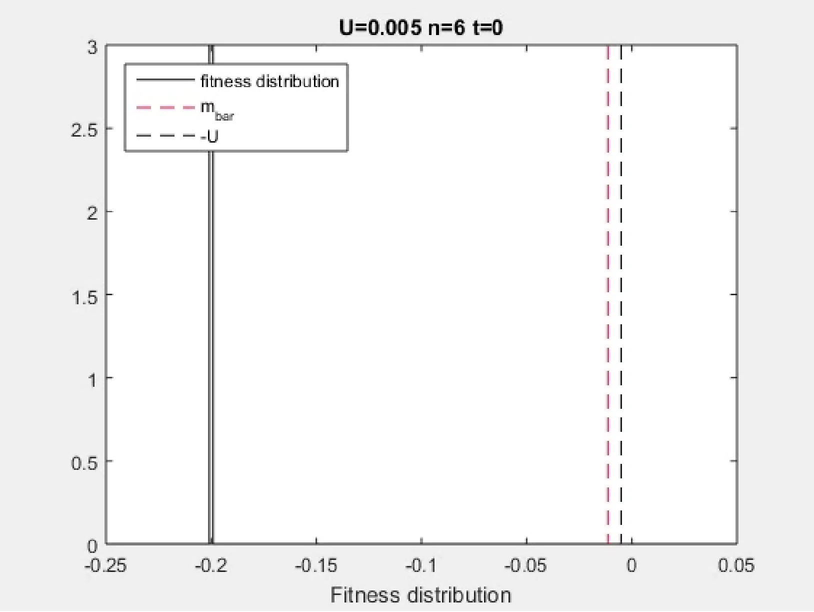
Reminiscent of *Pleiotropy and the preservation of perfection*, Waxman and Peck, Science, 1998

B.III. Integro-differential approach

Results – stationary states, Gaussian FGM

Corollary (Gil, Hamel, Martin, R, 2018)

Gaussian FGM, $n > 2$ and $U < s_H^\star$: $\rho > 0$ and $\overline{m}_\infty = -U$.



Reminiscent of *Pleiotropy and the preservation of perfection*, Waxman and Peck, Science, 1998

INRA SCIENCE & IMPACT

B.III. Integro-differential approach

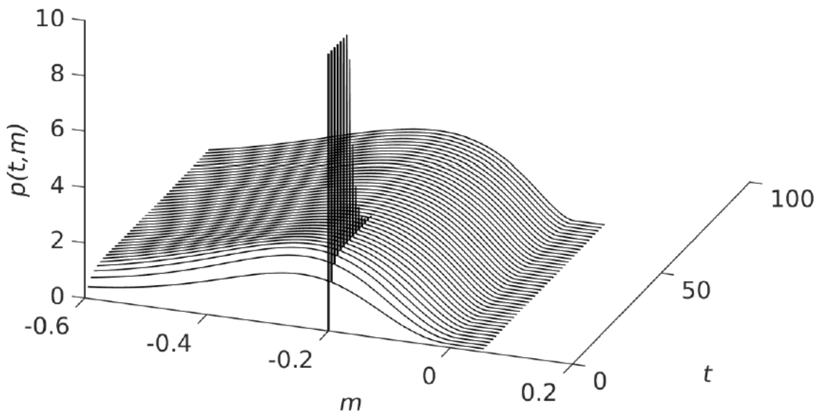
Results – stationary states, Gaussian FGM

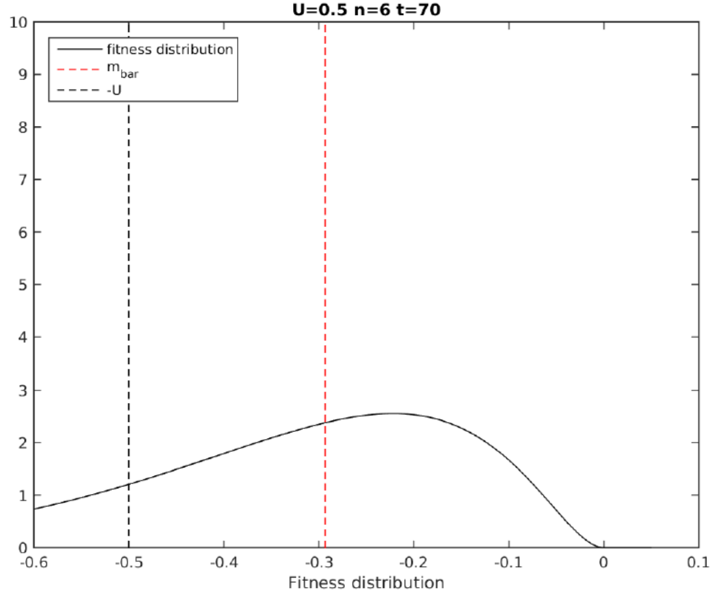
Corollary (Gil, Hamel, Martin, R, 2018)

Gaussian FGM, $n > 2$ and $U > \widetilde{U}$: $\rho = 0$ and $\overline{m}_\infty > -U.$

U=0.5 sH*=0.066667 n=6

$p_0 = \delta_{-0.2}, \ \lambda = 1/30.$



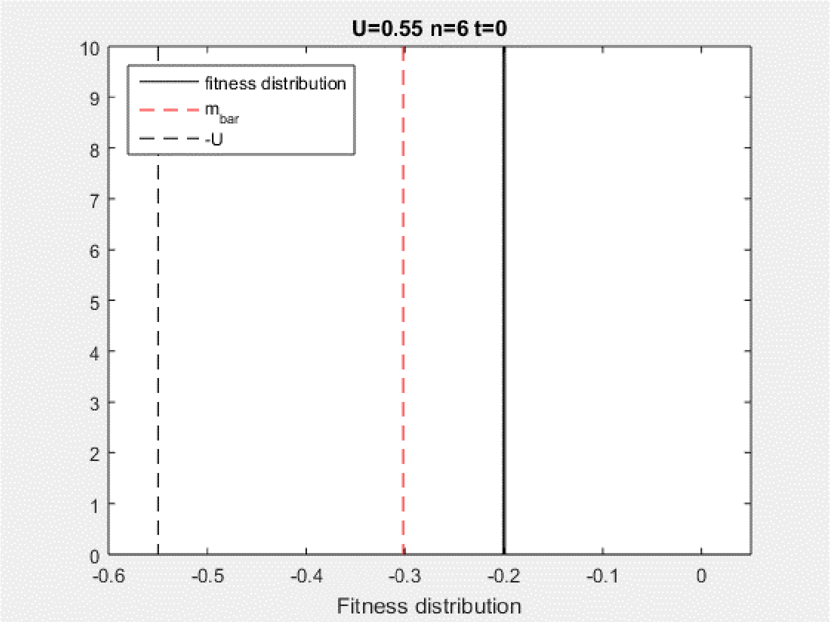


B.III. Integro-differential approach

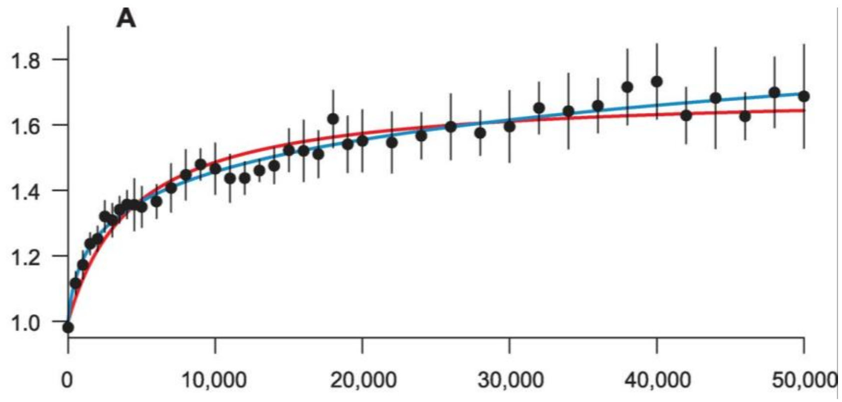
Results – stationary states, Gaussian FGM

Corollary (Gil, Hamel, Martin, R, 2018)

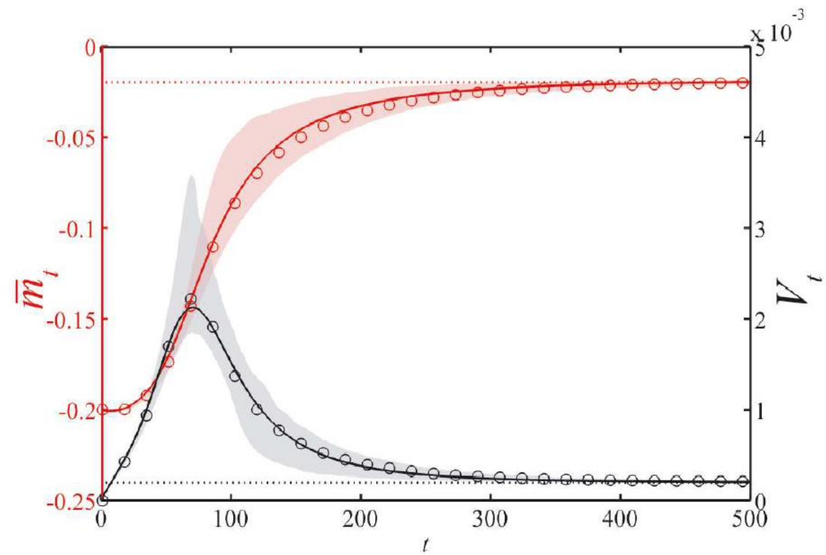
Gaussian FGM, $n > 2$ and $U > \widetilde{U}$: $\rho = 0$ and $\overline{m}_\infty > -U$.



Consistence with Lenski's experiment

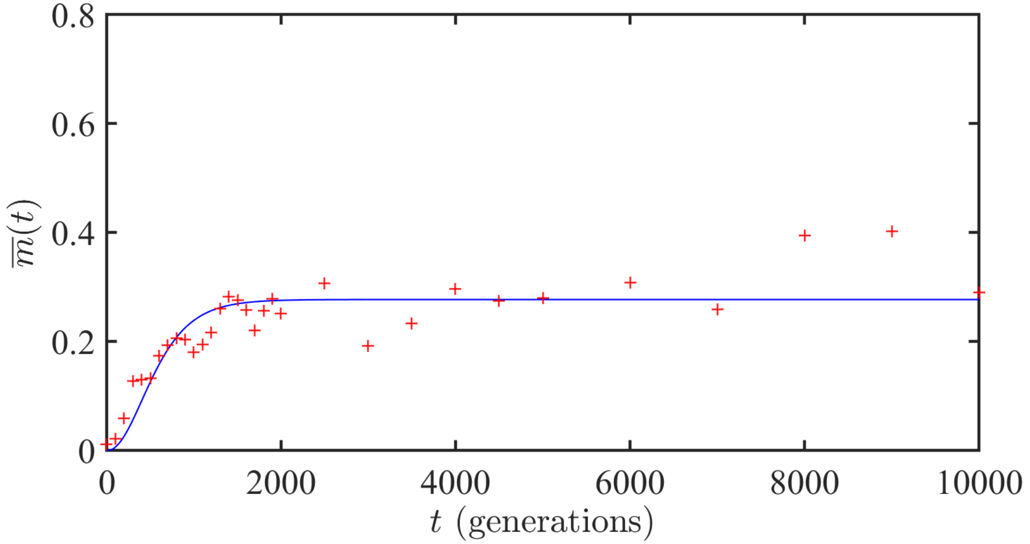


≈

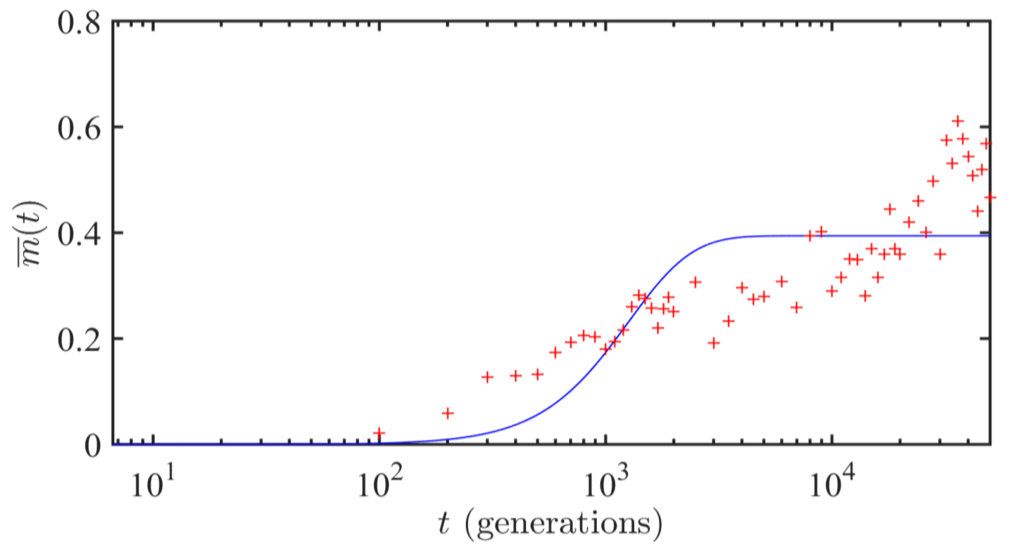


Model fit

Good fit for the first 10^4 generations…



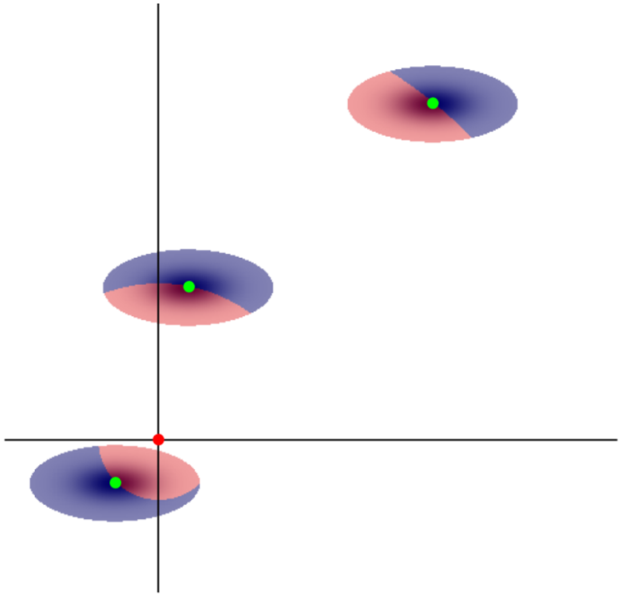
… but not later



Part C: Extension - anisotropic mutation effects

Equation driving the phenotype distribution

Mutation effects on fitness:



Dynamics of phenotype distribution, with anisotropic mutation effects:

$$\partial_t q(t,\mathbf{x}) = \sum_{i=1}^{n} \frac{\mu_i^2}{2}\, \partial_{ii} q(t,\mathbf{x}) + \left(m(\mathbf{x}) - \overline{m}(t)\right) q(t,\mathbf{x}),\ \ t>0,\ \ \mathbf{x}\in\mathbb{R}^n$$

with:

$$m\left(\mathbf{x}\right) = -\frac{\|\mathbf{x}\|^2}{2}.$$

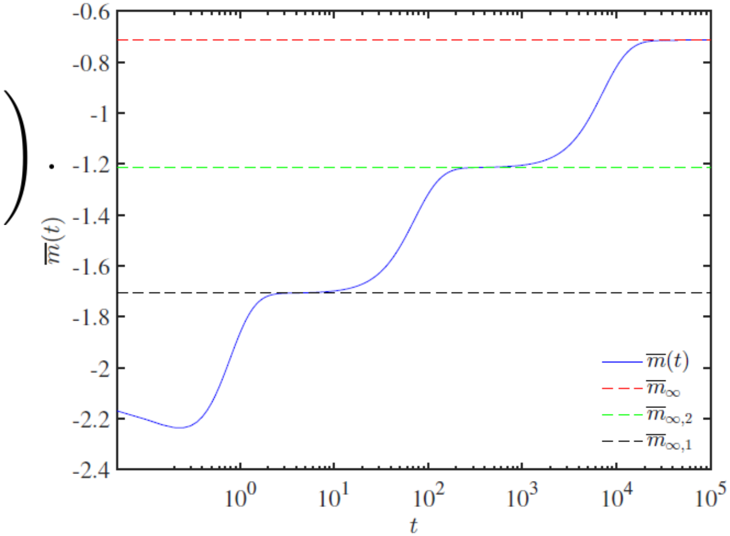
Mathematical results

(Hamel, Lavigne, Martin, R, 2019)

- Existence, uniqueness of the solution
- Equation for the fitness distribution (degenerate parabolic)
- Expression for the mean fitness
- Existence of 'plateaus'

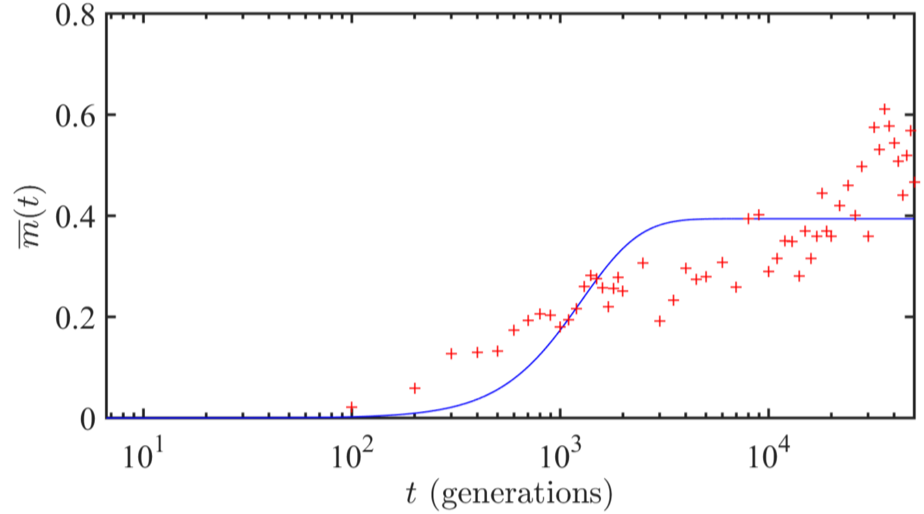
Initially clonal population leads to

$$\overline{m}(t) = \sum_{i=1}^{n} \left(\frac{x_{0,i}^2}{2} \left(\tanh^2(\mu_i t) - 1 \right) - \frac{\mu_i}{2} \tanh(\mu_i t) \right).$$

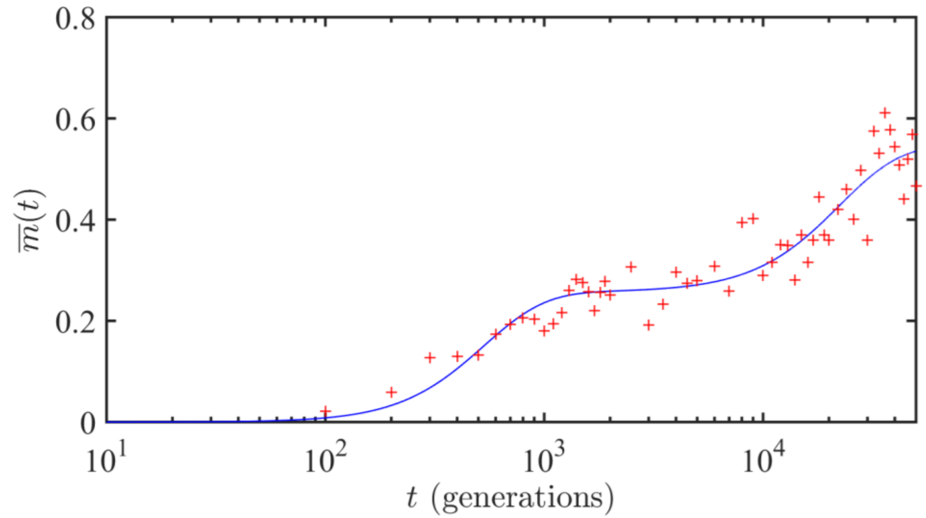


Model fit

Isotropic:



Anisotropic:



Thanks a lot for your attention!

References:

- Martin and Roques (2016) *The Non-stationary Dynamics of Fitness Distributions: Asexual Model with Epistasis and Standing Variation.* Genetics, 204,1541-1558
- Gil, Hamel, Martin and Roques (2017) *Mathematical properties of a class of integro-differential models from population genetics.* SIAM J. Appl. Math., 77, 1536-1561
- Roques, Garnier, Martin (2017) *Beneficial mutation-selection dynamics in finite asexual populations: a free boundary approach*, Scientific reports, 7, Article number: 17838
- Gil, Hamel, Martin and Roques (2019) *Dynamics of fitness distributions in the presence of a phenotypic optimum: an integro-differential approach*, Nonlinearity, in press..
- Hamel, Lavigne, Martin, Roques (2019) *Dynamics of adaptation in an anisotropic phenotype-fitness landscape*, bioRxiv 623330

A.IV. Free boundary approach - presentation

Main objective: to build a model at a mesoscopic scale:
- captures some features of the microscopic model (support of the solution remains bounded, finite speed of adaptation)
- analytically tractable PDE framework

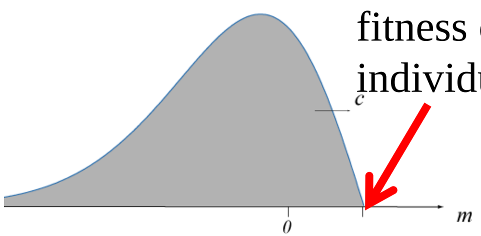
Inspired from the physics of ice melting.

Free boundary

Equation

$$\begin{cases} \partial_t T = D\Delta T, \ t > 0, \ x \in \Omega_t, \\ T = 0, \ t > 0, \ x \in \partial\Omega_t, \\ \partial_t T = \rho|\nabla T|^2, \ t > 0, \ x \in \partial\Omega_t, \\ T(0,x) = T_0(x), \ x \in \Omega_0, \end{cases}$$

Free boundary: fitness of the best individual



Introduce a free boundary in the PDE:

$$\partial_t p = D\partial_m^2 p + (m - \overline{m}(t))p, \ t > 0, m \in \mathbb{R}$$

with

$$\overline{m}(t) = \int_{\mathbb{R}} s\, p(t,s)ds \text{ and } D = M_2 U/2, \ M_2 = \int_{\mathbb{R}} s^2 J(s)ds$$

A.IV. Free boundary approach - TW solutions

We introduce the FBP, for each parameter $\mu > 0$:

$$\begin{cases} \partial_t p = D\partial^2_m p + (m - X(t))p, \ t > 0, m \in (-\infty, s(t)) \\ p(t, s(t)) = 0, \ t > 0 \\ s'(t) = -\mu\, D\, \partial_m p(t, s(t)), \ t > 0. \end{cases} \qquad \text{with } X(t) = \overline{m}(t) - \frac{1}{\mu}\, s'(t).$$

See [Du and Guo, 2012] for related eqs with KPP nonlinearity.

Theorem (Garnier, Martin, R, 2017)

For each $\mu > 0$, existence of a unique travelling wave solution:
$p(t,m) = \phi(m - v\,t)$ and $s(t) = v\,t$ with $\phi > 0$ on $(-\infty, 0)$ and $\int_{-\infty}^{0} \phi(y)\, dy = 1.$

Explicit solution: $\phi(z) = \frac{v}{\mu\, D^{2/3}\, \mathrm{Ai}'(-z_0)}\, e^{-\frac{v\,z}{2D}}\, \mathrm{Ai}\Big(-z_0 - z/D^{1/3}\Big)$

Ai : Airy function solving $\mathrm{Ai}''(z) - z\mathrm{Ai}(z) = 0$, and $\mathrm{Ai}(-z_0) = 0$.

Speed: $\mu = \frac{v}{D^{1/3}\, \mathrm{Ai}'(-z_0)} \int_0^{\infty} e^{\frac{v\,z}{2D^{2/3}}}\, \mathrm{Ai}(z - z_0)\, dz.$

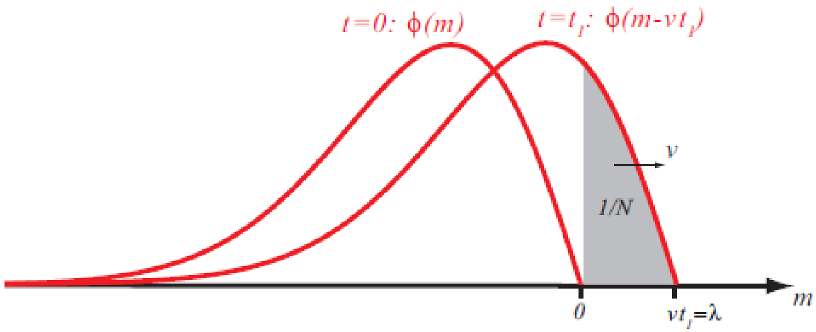
A.IV. Free boundary approach – population size N

Objective: find a relationship connecting μ, v and N.

Define:

t_1: expected time to establish a new beneficial mutation beyond the best fitness class

λ: expected increase in the best fitness class due to this beneficial mutation



$$t_1 \approx \frac{1}{v}\sqrt{\frac{2\,\mu\,D}{N\,v}}, \quad \lambda \approx \int_0^\infty s\,J(s)ds \Big/ \left(\int_0^\infty J(s)ds\right)$$

Using $v = \frac{\lambda}{t_1}$ and the previous relationship $\mu = f(v)$, we get:

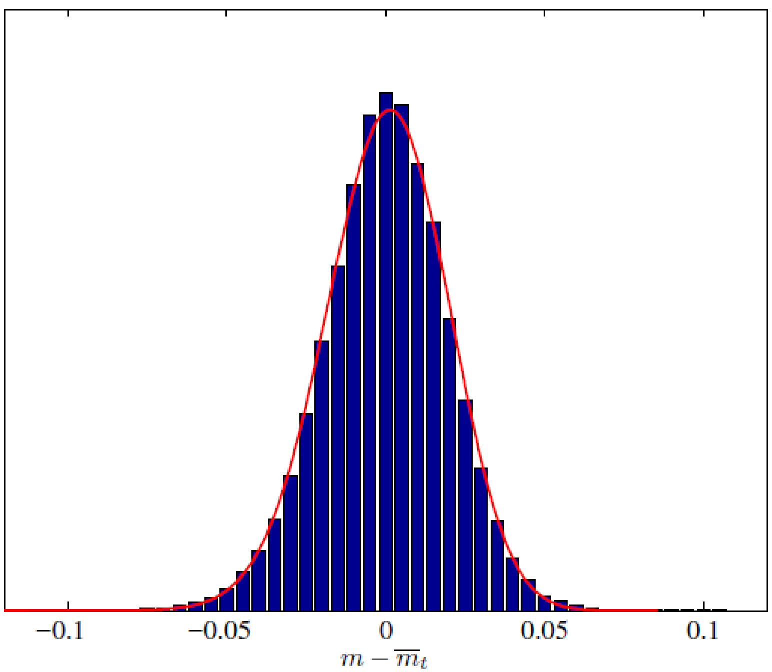
$$\boxed{v \approx K(D,\lambda)[\log(N)]^{1/3}} \text{ (explicit formulas for } v \text{ and } \mu)$$

→ consistent with a formula of *Neher and Hallatschek (2013)* for stochastic integro-differential eqs

A.IV. Free boundary approach – numerical computations

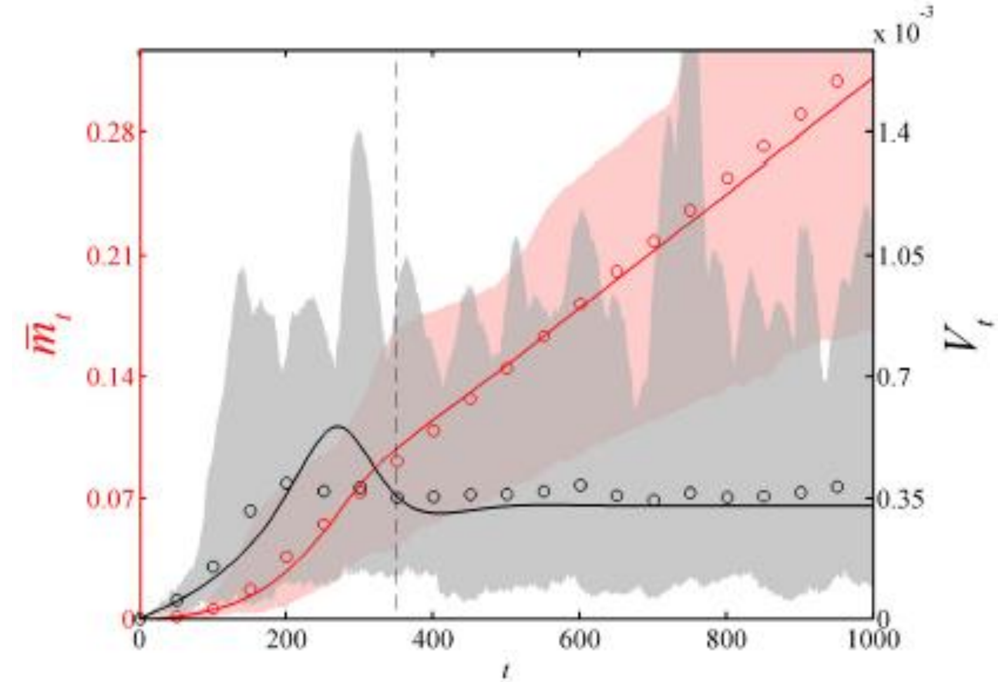
Free boundary approach vs stochastic simulations

Asymptotic distribution



(b) $N = 10^4$, $U = 0.01$

Trajectory of adaptation

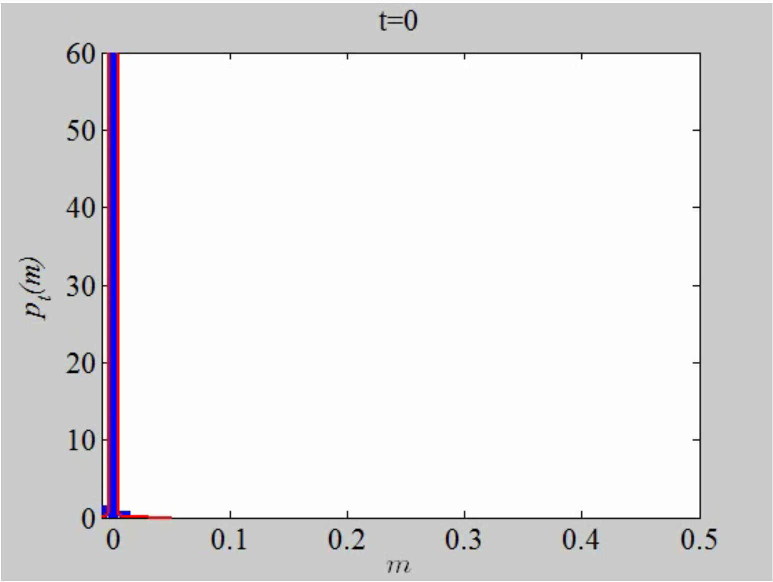


(b) $N = 10^4$, $U = 0.01$

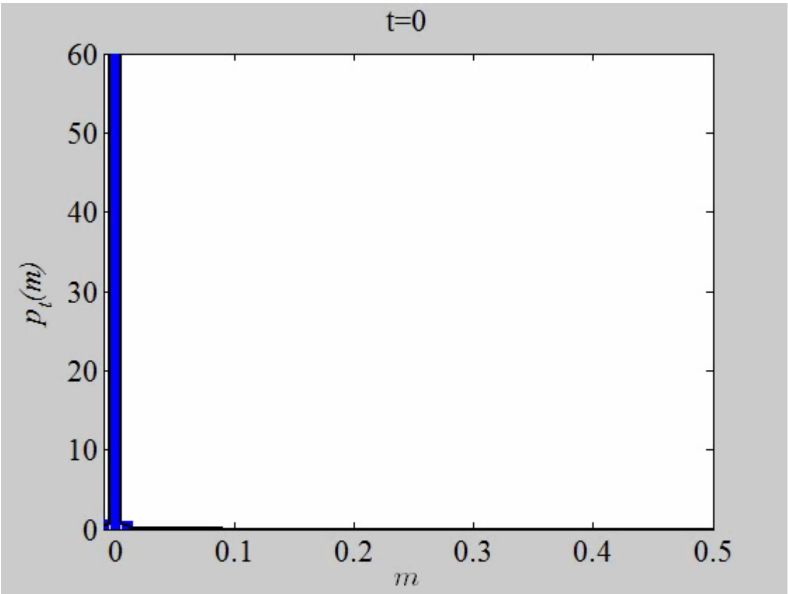
A.IV. Free boundary approach – numerical computations

Free boundary approach vs Integro-differential approach vs stochastic simulations

Distribution, free boundary



Distribution, integro-differential



Mutation kernel: