The dynamics of fitness under selection, and context-dependent mutation

Guillaume Martin & Lionel Roques

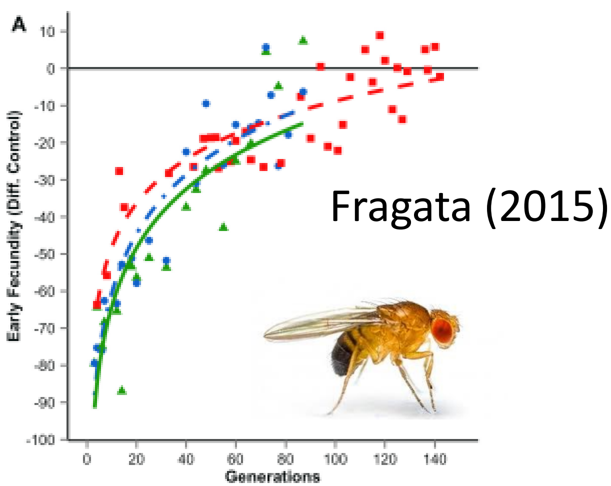
Models describing the rate of adaptation

Typical adaptation models: **stable regime assumptions** **=>** $\partial_t E(\overline{m})$ *= constant*

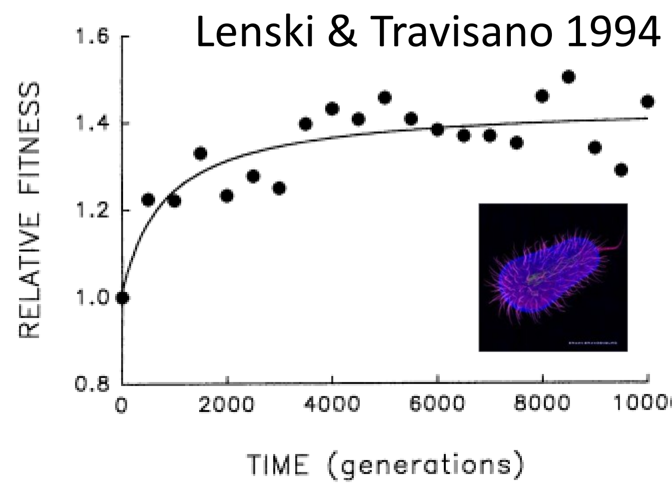
<u>fitness</u>: substitutions *(Gerrish & Lenski 1998)*, Travelling waves *(Desai & Fisher 2007)*

<u>traits</u>: Quantitative genetics *(Lande 1979) etc.*

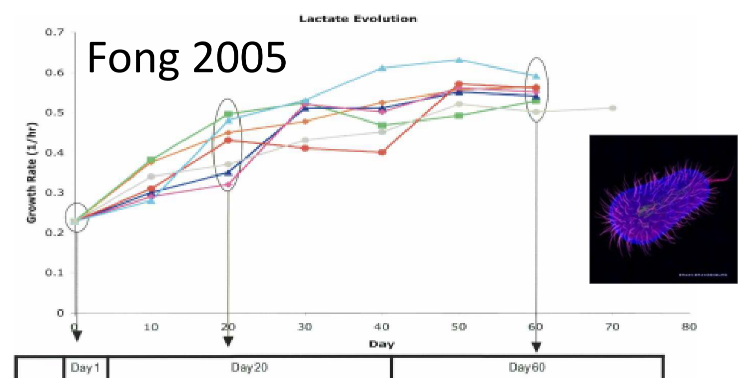
Typical adaptation data: **saturation**



Fragata (2015)



Lenski & Travisano 1994



Fong 2005

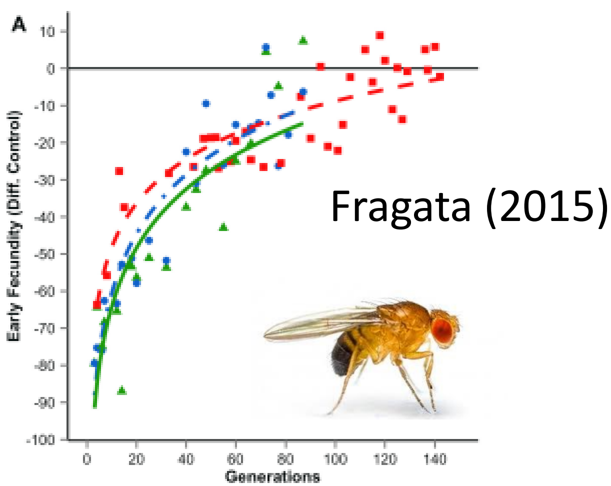
Models describing the rate of adaptation

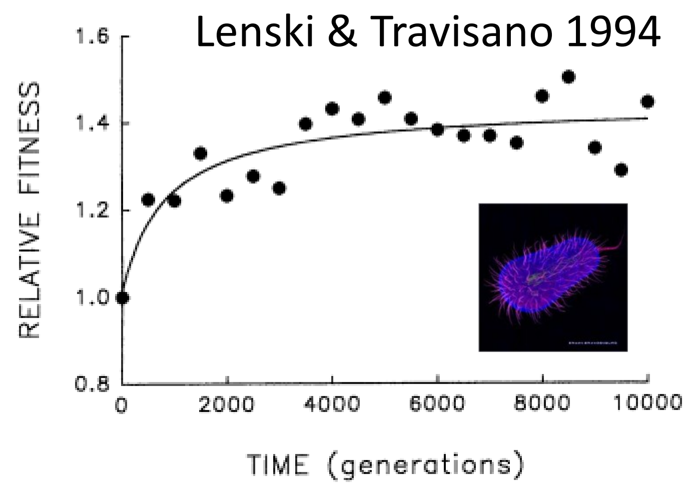
Typical adaptation models: **stable regime assumptions** **=>** $\partial_t E(\overline{m})$ = *constant*

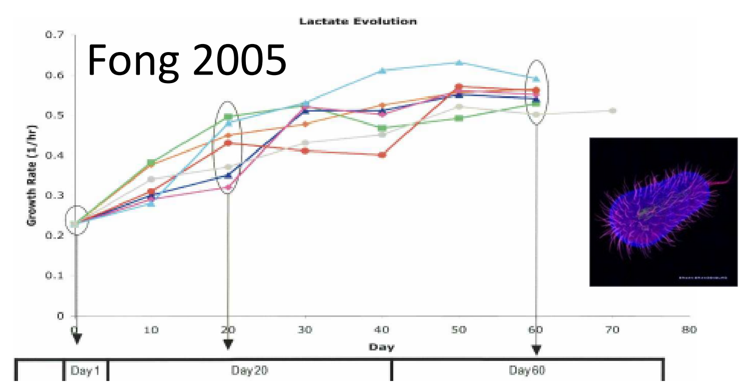
<u>fitness</u>: substitutions *(Gerrish & Lenski 1998)*, Travelling waves *(Desai & Fisher 2007)*

<u>traits</u>: Quantitative genetics *(Lande 1979) etc.*

Typical adaptation data: **saturation**







- **Erosion of initial variance**
- **Epistasis**

Fitness distribution dynamics with epistasis + standing variance generating function of the fitness distribution

Generating functions can prove handy:
R. Bürger (1991) : trait mutation
Johnson (1999),Desai & Fisher (2011) : deleterious mutation
Gerrish & Sniegowski (2012): mutation + short term
Rattray & Shapiro (2001): biallelic locus + mutation + drift
Good & Desai (2013): fitness-based , deleterious mutation + drift

Often: Infinite set of moment equations (but see Johnson 1999)
No epistasis

Fitness trajectories with epistasis:
Kryazhimskiy (2009) Dwyer (2012), Good & Desai (2015)
Analytic progress = away from clonal interference regime

Definitions / Assumptions

Asexuals, multitype wright fisher diffusion (continuous time approx)

m_i: malthusian fitness of genotype i (no frequency/density dependence)

Mutation : poisson process at rate U per unit time per capita

Distribution of Fitness Effects (DFE): $f(s|m_i)$ in background m_i

Cumulant Generating Function (CGF) :

$C_t(z) = \log(\sum p_i(t) e^{z\, m_i})$ at time t ($z \in \mathbb{R}^+$ if $m < 0$)

Derivatives in z => cumulants moments: $C_t'(0) = \bar{m}_t$: mean fitness

Dynamics of C_t ?

Selection + drift

$\mathcal{C}_t(z) = \langle C_t(z) \rangle$: expected CGF over stochastic process:

$\mathbf{p}(t) = \{p_i(t)\}_{i \in [1,K]} \in [0,1]^K \sim K$ −type Wright-Fisher diffusion

e.g. apply Feynman-Kac Theorem:

$$\partial_t \mathcal{C}_t(z) = \boxed{\mathcal{C}_t'(z) - \mathcal{C}_t'(0)} - \boxed{\frac{1 - \left\langle e^{C_t(2z) - 2C_t(z)} \right\rangle}{2N_e}}$$

Selection (clonal interference) **drift**

See also Rattray & Shapiro (2001) Good & Desai (2013)

When can we neglect drift here ?

$$\partial_t \mathcal{C}_t(z) = \boxed{\mathcal{C}_t'(z) - \mathcal{C}_t'(0)} - \frac{1 - \left\langle e^{\mathcal{C}_t(2z) - 2\mathcal{C}_t(z)} \right\rangle}{2N_e}$$

Selection **Drift effect $\boldsymbol{\delta}$**

- Var(within) >> Var(between)

p_{max}: frequency of the fittest class

$$0 \leq |\boldsymbol{\delta}| \leq \frac{1}{2N_e} \left|1 - \langle p_{max}^{-1} \rangle\right|$$

OK in models where the fittest class remains or quickly becomes substantial ($p_{max} \gg 1/N$)

= fitness upper bound (optimum, purely deleterious before ratchet)

Not OK on unbounded fitness sets (travelling waves) or with Muller's ratchet (late effect)

Background-dependent Mutation

Mutation effects CGF: $C_s(z,m) = \log\left(\int e^{s\,z} f(s|m)ds\right)$

Assume: linear context-dependence: $C_s(z,m) \approx \omega(z)m + C_*(z)$

$C_*(z)$: CGF of DFE in background $m = 0$ (e.g. optimum)

$\omega(z) = \partial_m C_s(z,m)|_{m=0}$: « context-dependence function »

Background-dependent Mutation

Mutation effects CGF: $C_s(z,m) = \log\left(\int e^{s\,z} f(s|m)ds\right)$

Assume: linear context-dependence: $C_s(z,m) \approx \omega(z)m + C_*(z)$

$C_*(z)$: CGF of DFE in background $m = 0$ (e.g. optimum)

$\omega(z) = \partial_m C_s(z,m)|_{m=0}$: « context-dependence function »

Small m approx when nearing a maximum fitness set at $m = 0$:

=> mutation – selection balance with diminishing returns epistasis

Exact at all times in some particular models:

- $\omega(z) = 0$: any non-epistatic model (finite moments)
- $\omega(z) = -z$: House Of Cards model (absolute effect independent of m)
- Fisher's geometrical model (quadratic trait- fitness function), see next

Closed approximate dynamics

$$\partial_t \mathcal{C}_t(z) \approx \mathcal{C}_t'(z) - \mathcal{C}_t'(0) + U\Big(\exp\Big(C_*(z) + \mathcal{C}_t\big(z + \omega(z)\big) - \mathcal{C}_t(z)\Big) - 1\Big)$$

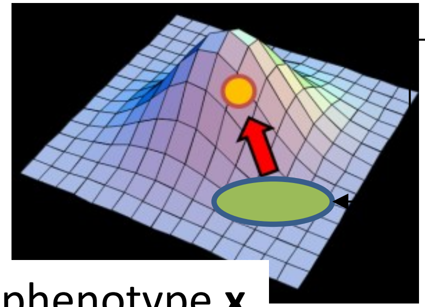
selection **Background-dependent mutation**

$\mathcal{C}_0(z)$: initial condition (standing fitness variance)

Nonlinear nonlocal PDE
can be solved (at least numerically) for given $C_*(z), \omega(z), C_0(z)$

Example : Fisher's model

fitness $e^{m(\mathbf{x})}$



phenotype **x**

Mutant cloud around background $\mathbf{x}$

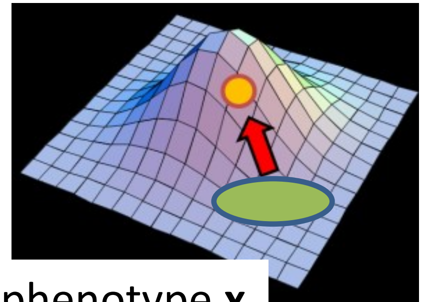
- variance λ per trait (isotropy)
- dimension n
- Normally distributed $\mathbf{dx} \sim N(\mathbf{0}, \lambda\, \mathbf{I}_n)$ (relaxed in some cases)

$$\Rightarrow C_s(z,m) = \omega(z)\, m + C_*(z)$$

Martin (2014)

$C_*(z) = -n/2 \log(1 + \lambda\, z)$: neg. gamma DFE at optimum

$\omega(z) = -\lambda\, z^2/(1 + \lambda\, z)$: epistasis



Example : Fisher's model

Example: at $t = 0$: a clone with fitness m_0
=> **expected mean fitness trajectory**: $\mathcal{C}_t'(0) = \langle \overline{m}_t \rangle$

- <u>Strong Selection Weak Mutation SSWM</u> ($U \ll s$)

$$\begin{cases} \partial_t \langle \overline{m}_t \rangle = U\, e^{\langle \overline{m}_t \rangle\, \omega(t)} \big(\langle \overline{m}_t \rangle\, M_*(t) \omega'(t) + M_*'(t) \big) \\ \langle \overline{m}_0 \rangle = m_0 \end{cases}$$

: ODE easy to solve, fit etc.

- <u>Weak selection strong mutation WSSM</u> ($U \gg s$)

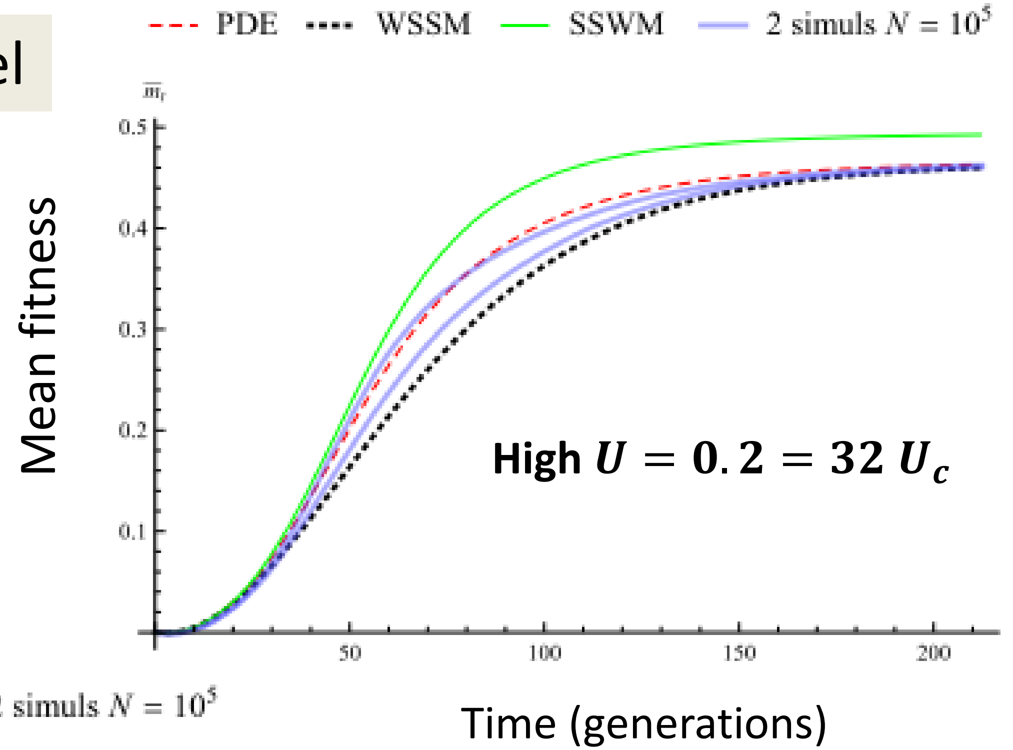
Leading order in λ => analytic solution to linearized PDE => $\langle \overline{m}_t \rangle = \mathcal{C}_t'(0)$

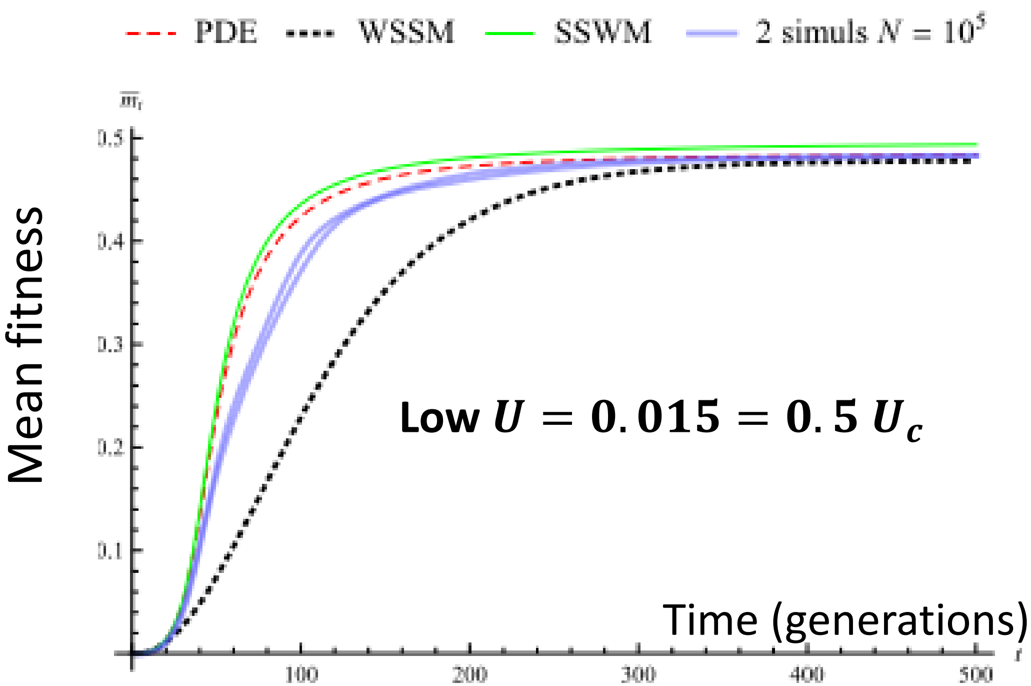
$$\langle \overline{m}_t \rangle - m_0 = -m_0 \tanh(\mu\, t)^2 - n/2\, \mu \tanh(\mu\, t)$$
$$\mu = \sqrt{U\, \lambda}$$

Valid when $U \gg U_c = n^2 \lambda / 4$

Fisher's Geometrical Model

$n = 5, \lambda = 0.001, N = 10^5$





Strategies for empirical testing

Obvious quantitative test: high power to reject/compare models

$$\langle \bar{m}_t \rangle = f(U, n, \lambda, m_0, t)$$

PDE + analytical approximations
Fisher's model and House of cards

1. Fit trajectories of $\bar{m}_t$
2. Compare fitted parameters to direct estimates (U, n, λ)

challenge: precision/stability of estimates across environments/labs etc.

Heursitic test: Use WSSM approx trajectory

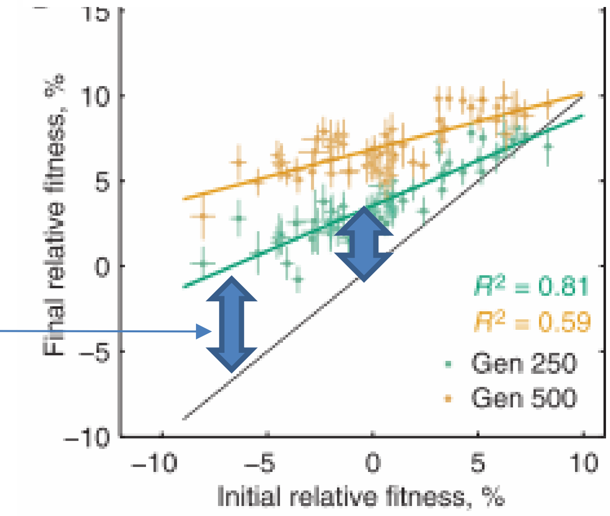
$$\langle \bar{m}_t \rangle - m_0 = -m_0 \tanh(\mu\, t)^2 - n/2\, \mu \tanh(\mu\, t)$$

Expected cumulated improvement $\propto$ initial fitness

1/ check linearity
2/ predict regression coefficients (depend on time + mutational parameters)

NB: Linearity with m_0 seems to hold outside the WSSM approx (formal proof ?)

Kryazhimskiy et al. Science, 2014



Yeast in YPD 30°C:
Founders with various m_0
Mean fitness after $t = 250$ or 500 generation

m_{ref}: fitness of reference strain for competitions

1/ Good linearity ($R^2 \geq 60\%$):

2/ Observed regression coefs at $t = 250$

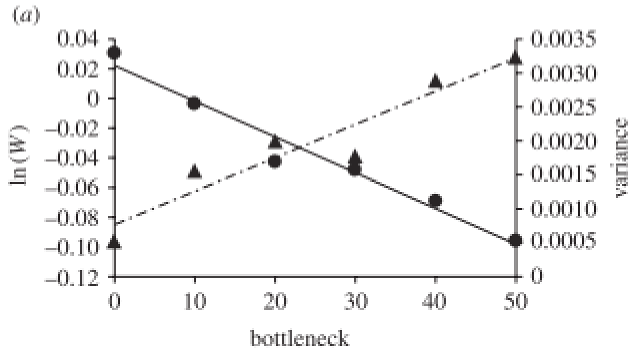
$\hat{\mu} = 0.0034$, $\hat{m}_{ref} = 0.09$

~ predicts regression coefs at $t = 500$

coef.	predicted	observed
β_{500}	-0.87	-0.67
α_{500}	7.86	7.88

NB: rejects context indep. model (slope = 0) and unique optimum reached (slope = -1)

Mutation Accumulation data



E.coli in LB, fitness in competition

Trindade et al. 2010
Perfeito et al. *Evolution* (2013)

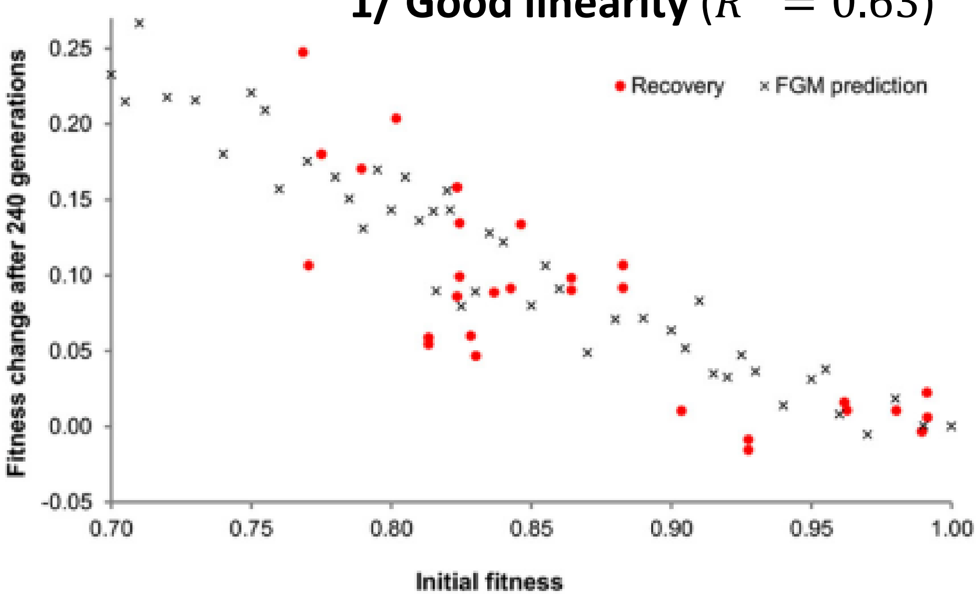
$\mu_{pred} \sim 0.007$

Perfeito et al. *Evolution* (2013): already tested FGM by ABC

Dynamics of fitness recovery in LB after 240 generations, same conditions

Against ancestor ($m_{ref} = 0$)

1/ Good linearity ($R^2 = 0.63$)



2/ ~ Predicts regresison coefficients

coef.	predicted	observed
β_{240}	- 0.87	-0.77
α_{240}	- 0.019	- 0.02

Conclusions

- Null models for experimental evolution
  +
  Classic methods to parameterize them from independent data

  } quantitatively test population genetics

- Drift need not be modelled sometimes, even in asexuals and with clonal interf.
  => Prediction robust to drift parameters (birth death rates, in each geno etc.)

- • Heuristic test uses WSSM which is not optimal here a priori (just orders)
  • Accounting for drift sensitive regimes (bias + enveloppe around expectations)

- some predictable patterns overlooked in experimental evolution (e.g. standing variance)
  Extensions: e.g. coupling with rescue dynamics, Yoann Anciaux, poster 1

- sequences data / coalescent: fitness distribution with predictable
  shape/scale/variance over time. Implement into MMC Coalescents ?

Thanks

The organizers and you

Ongoing tests / extensions :

Thomas Lenormand (CEFE Montpellier)

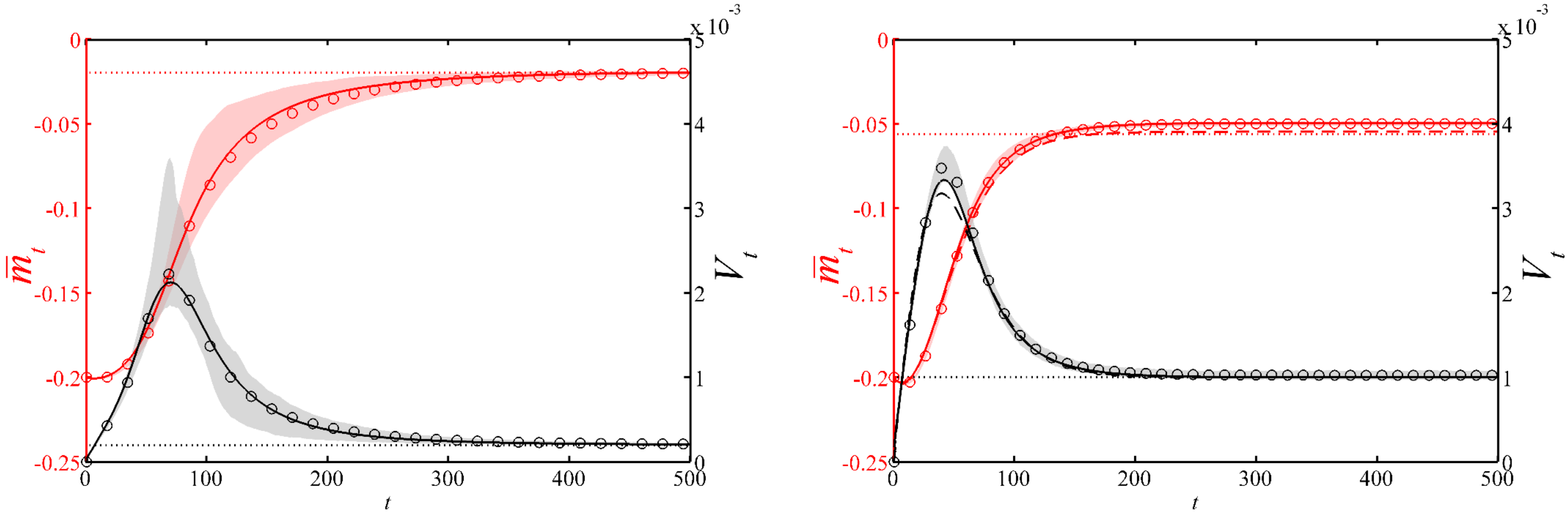
Yoann Anciaux (ISEM Montpellier)

Ophelie Ronce (ISEM Montpellier)

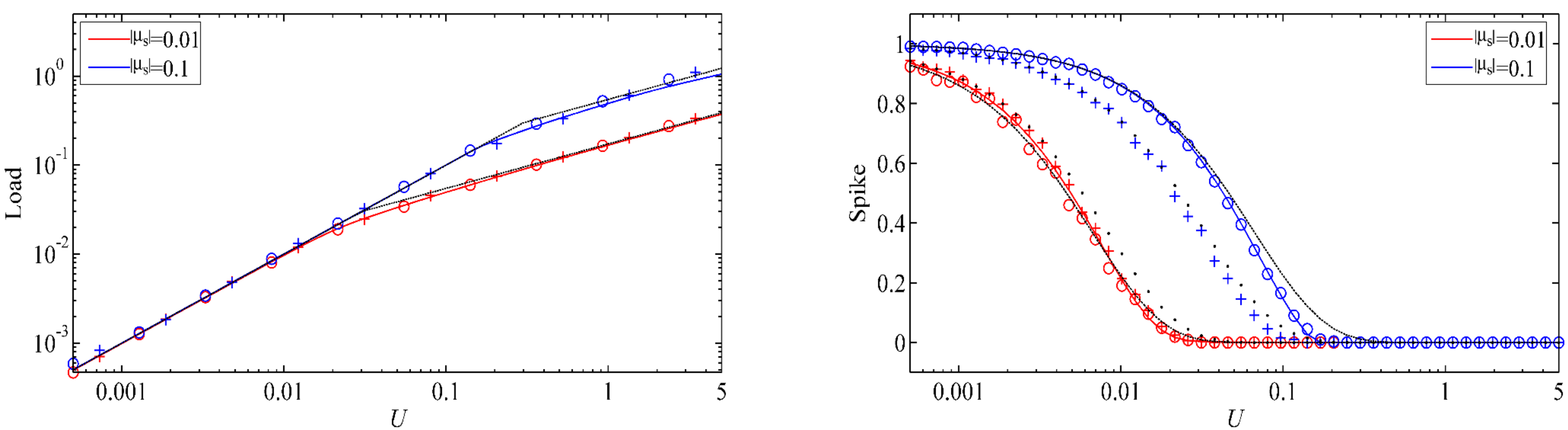
Amaury Lambert (UPMC, Collège de France)

Noemie Harmand (CEFE, Montpellier)

Luis Miguel Chevin (CEFE, Montpellier)



$\bar{m}_t$
V_t
x 10^{-3}
t
0
-0.05
-0.1
-0.15
-0.2
-0.25
0 100 200 300 400 500
0 1 2 3 4 5



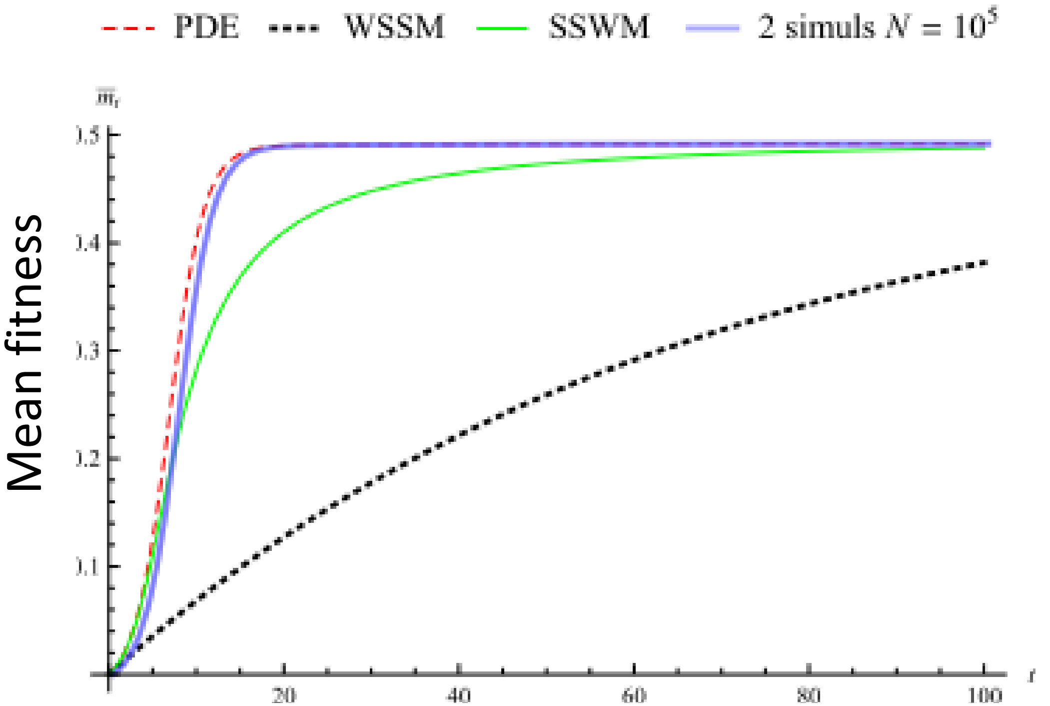
Load
Spike
U
$|\mu_s|$=0.01
$|\mu_s|$=0.1
10^0
10^{-1}
10^{-2}
10^{-3}
0.001 0.01 0.1 1 5
0 0.2 0.4 0.6 0.8 1

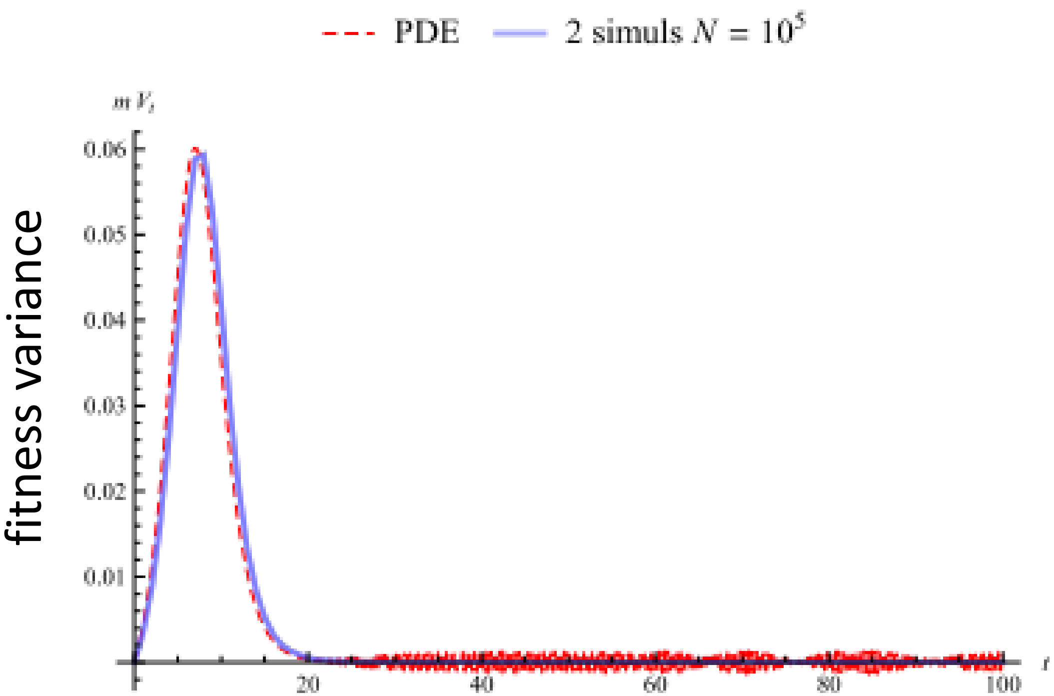
House of Cards Model

$C_*(z)$: arbitrary

(here gamma DFE)

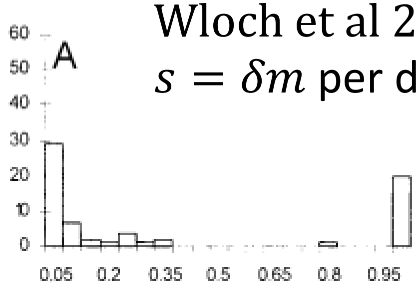
$\omega(z) = -z$





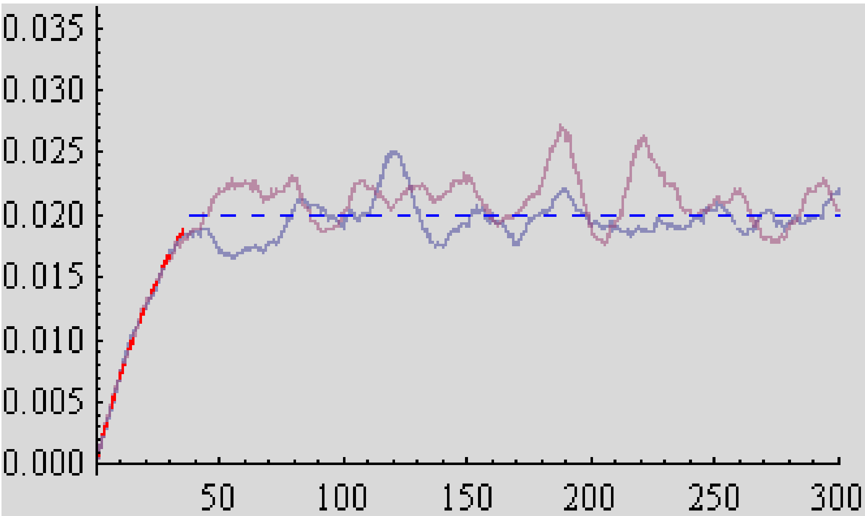
Wloch et al 2001, random single mutations in yeast

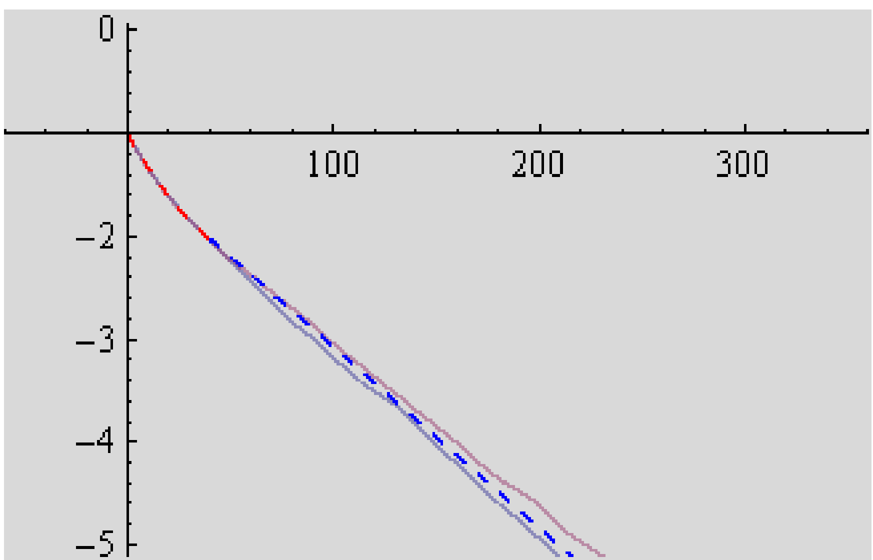
$s = \delta m$ per division in (haploid effect in YPD)

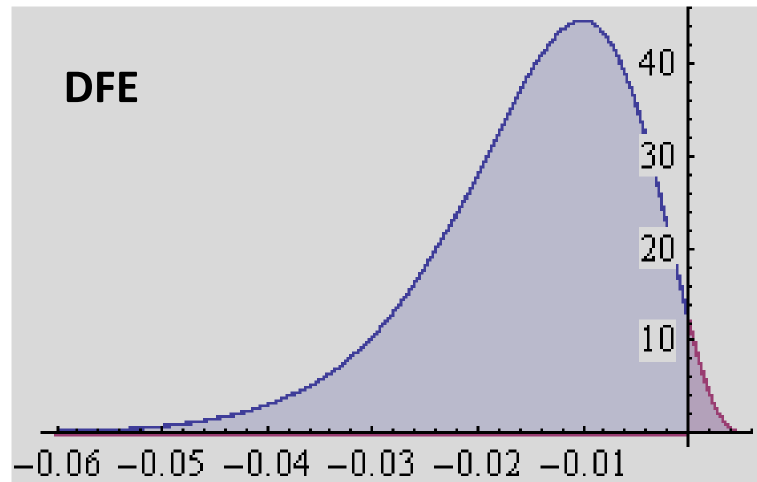


$$U = 0.66 \times 10^{-3}\ (div^{-1})$$
$$\bar{s} = 0.086\ (div^{-1})$$
$$n = 1/CV^2$$

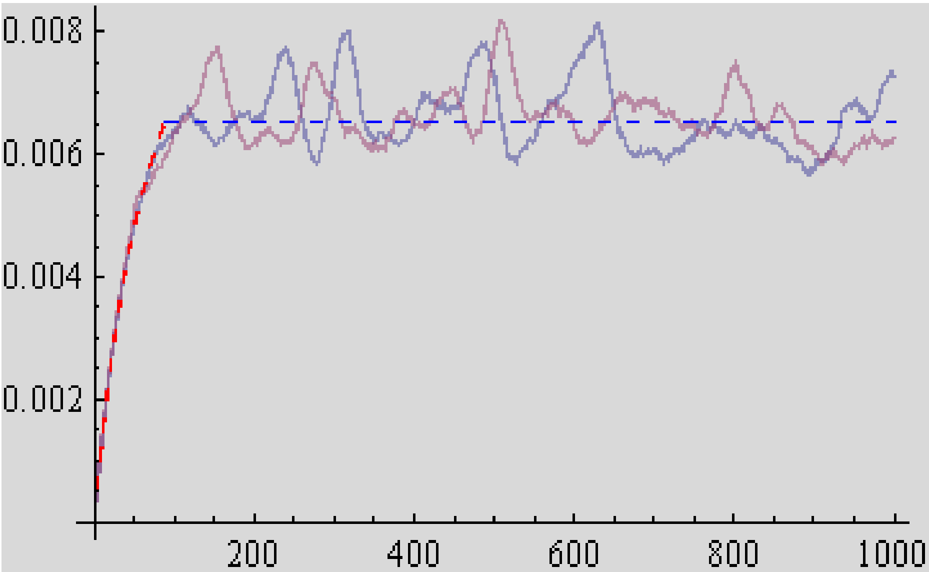
$N = 10^4, U = 500\lambda, \theta = 4, s_0 = \lambda, \lambda = 0.005$

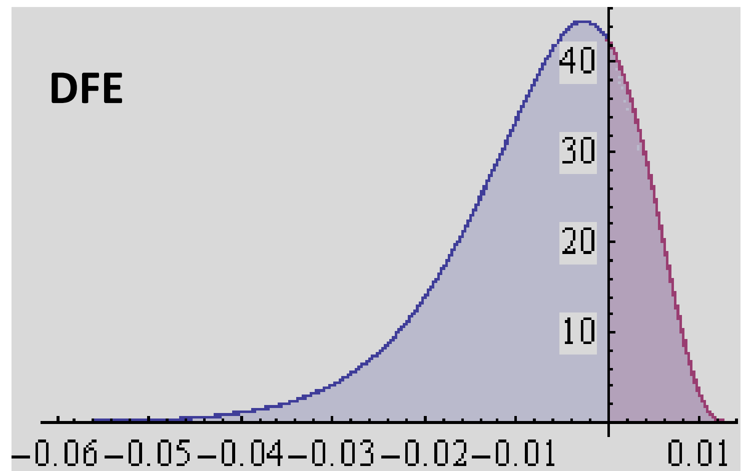


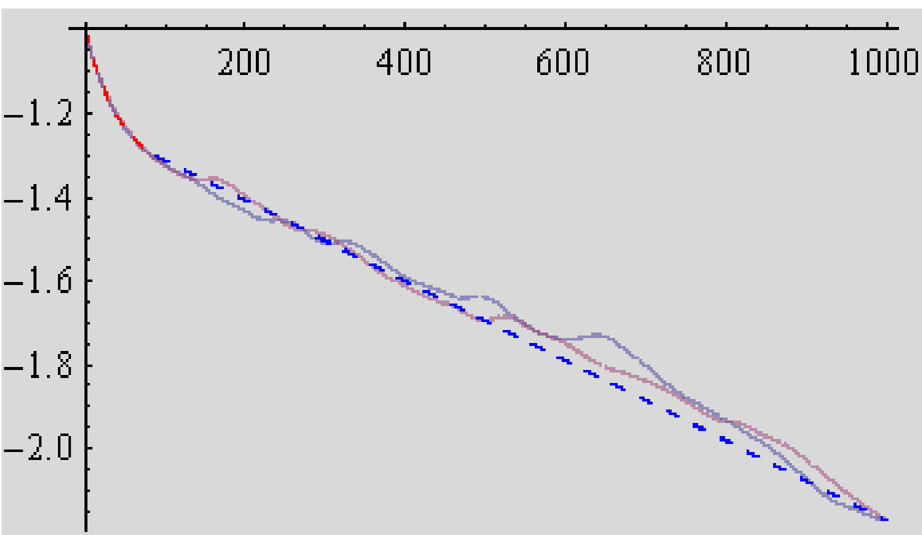




$$N = 10^5, U = 200\lambda, \theta = 4, s_0 = 2.5\lambda, \lambda = 0.005$$







$N = 10^5, U = 200\lambda, \theta = 4, s_0 = 3\lambda, \lambda = 0.01$