

# Wasserstein estimates and convergence to equilibrium for an evolutionary biology model

Gaël Raoul

June 26, 2022

Centre de mathématiques appliquées,  
CNRS & Ecole Polytechnique, Paris, France.

# Dynamics of a kinetic model from evolutionary biology

The model

Population Genetics' approach: Gaussian approximation

A macroscopic limit to the ODE model

Exponential convergence to a steady-state

Raoul, G., 2021. Exponential convergence to a steady-state for a population genetics model with sexual reproduction and selection. *arXiv preprint arXiv:2104.06089*.

## An example of a Phenotypic trait

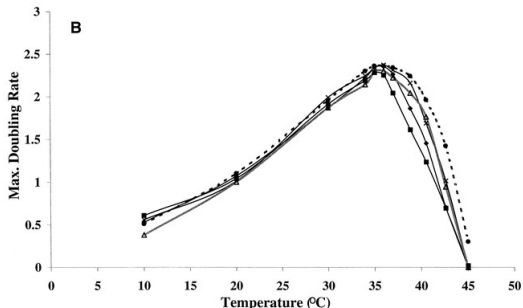


Figure: growth rate of a given E-coli strain for different temperatures<sup>1</sup>

The phenotypic trait  $x \sim 37$  is the temperature to which the strain is best adapted to.

<sup>1</sup>Bronikowski et al, *Evolution*, 2001.

# Local adaptation

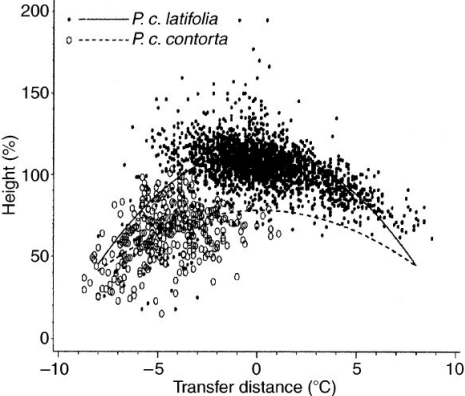
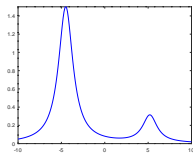
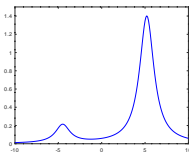
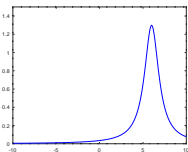
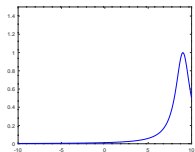
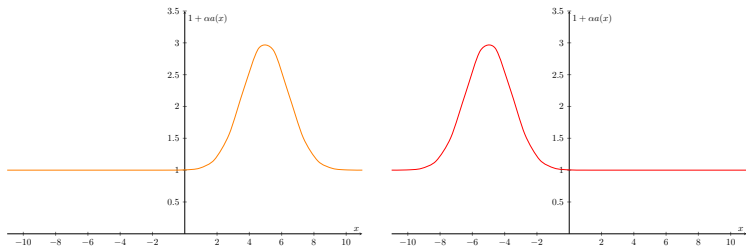


Figure: Common garden experiment on natural populations.



# Adaptation to a given fitness landscape: evolutionary rescue



# The origin of population genetics: artificial selection

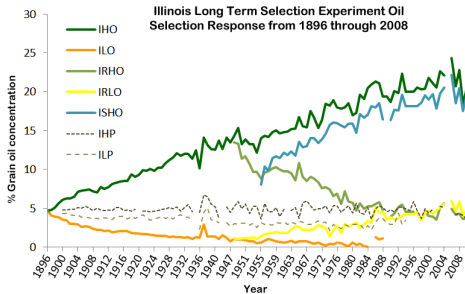
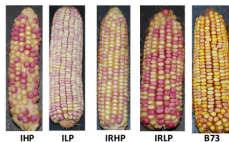
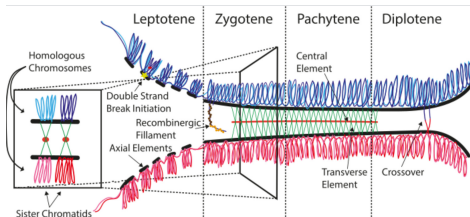
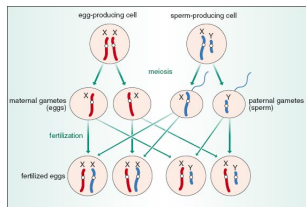


Figure: Long term artificial selection experiment.<sup>3</sup>

<sup>3</sup>S. Moose team, Department of Crop Sciences; University of Illinois; Urbana-Champaign

# Sexual reproduction and recombination



1
1
-1
1
-1
1
-1
-1
-1
1
1
1

1
-1
1
-1
-1
-1
1
-1
1
-1
1




## Infinitesimal model

Alleles  $p^i \in \{\pm 1\}$  for  $i \in \{1, \dots, N\}$ , traits

$$y_* = \frac{1}{\sqrt{N}} \sum_{i=1}^N p_*^i, \quad y'_* = \frac{1}{\sqrt{N}} \sum_{i=1}^N p_*^{i'}$$

such that  $\frac{1}{N} \sum_{i=1}^N p_*^i \sim 0$  and  $\frac{1}{N} \sum_{i=1}^N p_*^{i'} \sim 0$ . We then define the trait of the offspring by

$$y = \frac{1}{\sqrt{N}} \sum_{i=1}^N (p_*^i \delta_i + p_*^{i'} (1 - \delta_i)),$$

with  $\delta_i$  Bernouilly random variables. In the large  $N$  limit<sup>4</sup>, the law of  $y$  is

$$y \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\left(y - \frac{y_* + y'_*}{2}\right)^2}{2\sigma^2}}$$

---

<sup>4</sup>Fisher, *Earth Environ. Sci. Trans. R. Soc. Edinb.*, 1919.  
Barton, Etheridge, Véber, *Theor. popul. biol.* (2017).

## Infinitesimal model

For  $n, m \in \mathcal{P}_2(\mathbb{R})$ , we define the infinitesimal operator

$$T(n, m)(y) = \int \int \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(y - \frac{y_* + y'_*}{2}\right)^2}{2\sigma^2}} n(y_*) m(y'_*) dy * dy'_*,$$

and  $T(n) := T(n, n)$ .

### Properties:

▶  $\int T(n)(x) x dx = \int n(x) x dx,$

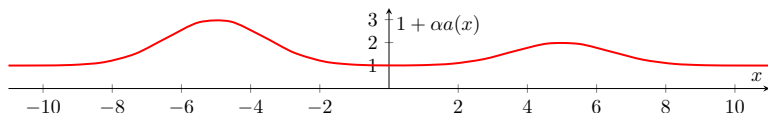
▶  $T(\Gamma_{2\sigma^2}) = \Gamma_{2\sigma^2},$  where  $\Gamma_{2\sigma^2} = \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{\left(\cdot - \frac{y_* + y'_*}{2}\right)^2}{4\sigma^2}}.$

## The model

The model we will focus on is

$$\partial_t n(t, y) =$$

$$\int \int \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\left(y - \frac{y_* + y'_*}{2}\right)^2}{2\sigma^2}} (1 + \alpha a(y_*)) n(y_*) (1 + \alpha a(y'_*)) n(y'_*) dy_* dy'_* \\ - \left(1 + \alpha \int a(x) n(t, x) dx\right)^2 n(t, y).$$



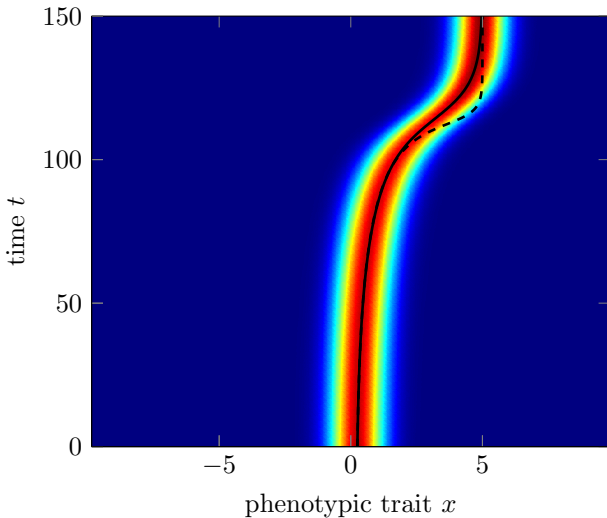
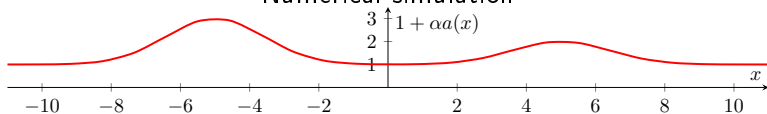
## An equivalent ecology model

The fixed population size assumption can be relaxed, as shown in the following model:

$$\begin{aligned} \partial_t f(t, x) = & \int \int \Gamma_{\sigma^2} \left( x - \frac{y_* + y'_*}{2} \right) (1 + \alpha a(y_*)) f(t, y_*) \\ & (1 + \alpha a(y'_*)) \frac{f(t, y'_*)}{\int f(t, y) dy} dy_* dy'_* \\ & - \left( \int K(y) f(t, y) dy \right) f(t, x). \end{aligned}$$

Then  $n(t, x) := \frac{f(t, x)}{\int f(t, y) dy}$  is a probability measure and satisfies our main model.

# Numerical simulation



# Dynamics of a kinetic model from evolutionary biology

The model

Population Genetics' approach: Gaussian approximation

A macroscopic limit to the ODE model

Exponential convergence to a steady-state

## References

Fisher, R.A. (1930). *The Genetical Theory of Natural Selection*. Oxford, UK: Clarendon Press.

"The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time"

MG Bulmer, *The mathematical theory of quantitative genetics*. Oxford, UK: Clarendon Press (1980).

Mirrahimi, S. and R., G. Dynamics of sexual populations structured by a space variable and a phenotypical trait. *Theoretical population biology*, **84**, 87–103 (2013).

# "Fundamental theorem of natural selection"<sup>5</sup>

---

<sup>5</sup>Fisher, R.A. (1930). *The Genetical Theory of Natural Selection*. Oxford, UK: Clarendon Press.

M. Alfaro, R. Carles, Replicator-mutator equations with quadratic fitness, *Proc. Amer. Math. Soc.* 145 (2017), 5315–5327.



## Gaussian approximation

It is classical to assume that phenotypic traits of individuals in a population are distributed normally:

$$n(t, x) \sim \Gamma_{2\sigma^2}(y - Z(t)).$$

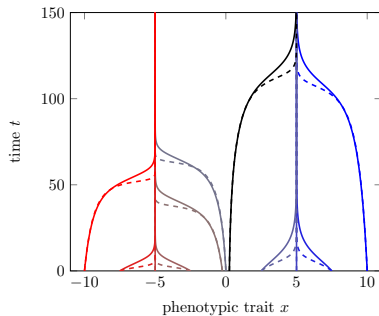
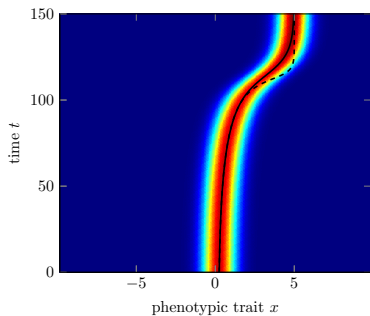
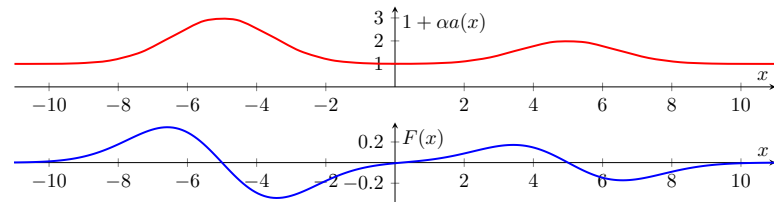
We can then estimate:

$$\begin{aligned} Z'(t) &= \frac{d}{dt} \int xn(t, x) dx \\ &= \alpha(1 + \alpha I_n(t)) \left( \int xa(x)n(t, x) dx - I_n(t) \int xn(t, x) dx \right) \\ &\sim \alpha F(Z(t)), \end{aligned}$$

where  $I_n(t) = \int a(x)n(t, x) dx$  and

$$F(Y) := \int (x - Y)a(x)\Gamma_{2\sigma^2}(x - Y) dx.$$

# Gaussian approximation



$\alpha = 1$  for the structured population model.

# Dynamics of a kinetic model from evolutionary biology

The model

Population Genetics' approach: Gaussian approximation

A macroscopic limit to the ODE model

Exponential convergence to a steady-state

## Related References

Magal, P. and Raoul, G., 2015. Dynamics of a kinetic model describing protein exchanges in a cell population. *arXiv preprint arXiv:1511.02665*.

Patout, F., 2020. The Cauchy problem for the infinitesimal model in the regime of small variance. *arXiv preprint arXiv:2001.04682*.

## Wasserstein estimate on the reproduction operator

Proposition (Tanaka-type contraction estimate)

If  $n, m \in \mathcal{P}_2(\mathbb{R})$  and  $Z_n = \int xn(x) dx = \int xm(x) dx = Z_m$ , then

$$W_2(T(n), T(m)) \leq \frac{1}{\sqrt{2}} W_2(n, m).$$

Wasserstein distance: For  $n, m \in \mathcal{P}_2(\mathbb{R})$ ,

$$W_2(n, m) = \left( \inf_{\pi \in \Pi} \int \int |x - y|^2 d\pi(x, y) \right)^{1/2},$$

where  $\Pi = \{\pi \in \mathcal{P}_2(\mathbb{R}); \pi_1 = n, \pi_2 = m\}$ .

Dual formula:

$$W_2(n, m)^2 = \max \left\{ \int \phi(y)n(y) dy + \int \psi(Y)m(Y) dY; \right. \\ \left. \phi(y) + \psi(Y) \leq |y - Y|^2 \right\}.$$

## Sketch of proof for the Tanaka-type contraction estimate

## Theorem

If  $\alpha > 0$  is small enough and under some assumptions on the initial condition,

$$\forall t \in [-C \ln \alpha / \alpha, \infty), \quad W_2(n(t, \cdot), \Gamma_{2\sigma^2}(\cdot - Z_n(t))) \leq C\alpha, \quad (1)$$

and

$$\forall t \in [0, \infty), \quad |Z_n(t) - Y(t/\alpha)| \leq \frac{C}{-\ln \alpha}, \quad (2)$$

where  $Y$  is the solution of  $Y' = F(Y)$  with initial data  $Z^0 = \int x n^0(x) dx$ .

Slow-fast structure:

- ▶ Slow manifold: Gaussian distributions  $\{\Gamma_{2\sigma^2}(\cdot - Z), Z \in \mathbb{R}\}$ .
- ▶ Fast dynamics:  $W_2(n(t, \cdot), \Gamma_{2\sigma^2}(\cdot - Z_n(t)))$
- ▶ Slow dynamics:  $Z_n(t)$ , with speed  $\sim \alpha$

## Sketch of proof

**Step 1:**  $n(t, \cdot)$  is close to a Gaussian distribution.

$$\begin{aligned}n(t, x) &= n(0, x) e^{-\int_0^t (1 + \alpha I_n(s))^2 ds} \\ &+ \int_0^t (1 + \alpha I_n(s))^2 T\left(\frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}\right)(x) e^{-\int_s^t (1 + \alpha I_n(\tau))^2 d\tau} ds,\end{aligned}$$

while

$$\begin{aligned}\Gamma_{2\sigma^2}(x - Z_n(t)) &= \Gamma_{2\sigma^2}(x - Z_n(t)) e^{-\int_0^t (1 + \alpha I_n(s))^2 ds} \\ &+ \int_0^t (1 + \alpha I_n(s))^2 \Gamma_{2\sigma^2}(\cdot - Z_n(t))(x) e^{-\int_s^t (1 + \alpha I_n(\tau))^2 d\tau} ds \\ &= \Gamma_{2\sigma^2}(x - Z_n(t)) e^{-\int_0^t (1 + \alpha I_n(s))^2 ds} \\ &+ \int_0^t (1 + \alpha I_n(s))^2 T(\Gamma_{2\sigma^2}(\cdot - Z_n(t)))(x) e^{-\int_s^t (1 + \alpha I_n(\tau))^2 d\tau} ds.\end{aligned}$$



## Sketch of proof

$$\begin{aligned} & W_2(n(t, x), \Gamma_{2\sigma^2}(x - Z_n(t))) \\ &= W_2(n(0, x), \Gamma_{2\sigma^2}(x - Z_n(t))) e^{-\int_0^t (1 + \alpha I_n(s))^2 ds} \\ &+ \int_0^t (1 + \alpha I_n(s))^2 e^{-\int_s^t (1 + \alpha I_n(\tau))^2 d\tau} \\ &W_2\left(T\left(\frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}\right), T(\Gamma_{2\sigma^2}(\cdot - Z_n(t)))\right) ds, \end{aligned}$$

and

$$\begin{aligned} & W_2\left(T\left(\frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}\right), T(\Gamma_{2\sigma^2}(\cdot - Z_n(t)))\right) \\ &\leq \frac{1}{\sqrt{2}} W_2\left(\frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}, \Gamma_{2\sigma^2}(\cdot - Z_n(t))\right) \\ &+ \mathcal{O}(\alpha)(1 + |t - s|). \end{aligned}$$

## Sketch of proof

Step 2:  $Z_n(t) := \int xn(t, x) dx$  satisfies  $Z'_n(t) \sim F(Z_n(t))$

$$\begin{aligned} Z'_n(t) &= \alpha l_n(t)(1 + \alpha l_n(t)) \left[ \int y \frac{a(y)n(t, y)}{l_n(t)} dy - Z_n(t) \right] \\ &= \alpha l_n(t)(1 + \alpha l_n(t)) \left[ \int y \frac{a(y)\Gamma_{2\sigma^2}(y - Z_n(t))}{l_n(t)} dy - Z_n(t) \right] \\ &\quad + \mathcal{O}(\alpha) W_2(n(t, \cdot - \Gamma_{2\sigma^2}(\cdot - Z_n(t))), \end{aligned}$$

and then

$$Z'_n(t) = \alpha F(Z_n(t)) + \mathcal{O}(\alpha),$$

which yields

$$|Z_n(t) - Y(t/\alpha)| \leq C\alpha te^t.$$

# Dynamics of a kinetic model from evolutionary biology

The model

Population Genetics' approach: Gaussian approximation

A macroscopic limit to the ODE model

Exponential convergence to a steady-state

## References and alternative analysis ideas

T. Bourgeron, V. Calvez, J. Garnier, T. Lepoutre, Existence of recombination-selection equilibria for sexual populations. *arXiv preprint arXiv:1703.09078* (2017).

V. Calvez, J. Garnier, F. Patout, Asymptotic analysis of a quantitative genetics model with nonlinear integral operator. *Journal de l'École polytechnique–Mathématiques*, **6**:537–579 (2019).

Calvez, V., Lepoutre, T. and Poyato, D., 2021. Ergodicity of the Fisher infinitesimal model with quadratic selection. *arXiv preprint arXiv:2107.00383*.

# Recall

## Slow-fast structure:

- ▶ Slow manifold: Gaussian distributions  $\{\Gamma_{2\sigma^2}(\cdot - Z), Z \in \mathbb{R}\}$ .
- ▶ Fast dynamics:  $W_2(n(t, \cdot), \Gamma_{2\sigma^2}(\cdot - Z_n(t)))$
- ▶ Slow dynamics:  $Z_n(t)$ , with speed  $\sim \alpha$

If the fast dynamics has a stable equilibrium, do we have a unique steady-state of the full model, when  $\alpha > 0$  is small ?

## Pushing the analysis further

For  $n$ ,  $m$  solutions, we estimate two quantities (slow and fast):

$$|Z_n(t) - Z_m(t)|,$$

$$\begin{aligned} w(t) &:= W_2(n(t, \cdot - Z_n(t)), m(t, \cdot - Z_m(t))) \\ &= \min_{\xi} W_2(n(t, \cdot - \xi), m(t, \cdot)). \end{aligned}$$

To estimate the second quantity, we consider

$$\begin{aligned} n(t, x) &= n(0, x) e^{-\int_0^t (1 + \alpha I_n(s))^2 ds} \\ &+ \int_0^t (1 + \alpha I_n(s))^2 T \left( \frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)} \right) (x) e^{-\int_s^t (1 + \alpha I_n(\tau))^2 d\tau} ds, \end{aligned}$$

and need to estimate

$$\begin{aligned} &W_2 \left( \frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}, \frac{(1 + \alpha a)m(s, \cdot)}{1 + \alpha I_m(s)} \right) \\ &\leq \mathcal{O}(\alpha) W_2(n(t, \cdot), m(t, \cdot)) + C |Z_n(t) - Z_m(t)|. \end{aligned}$$

$$W_2 \left( \frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}, \frac{(1 + \alpha a)m(s, \cdot)}{1 + \alpha I_m(s)} \right) \leq \mathcal{O}(\alpha) W_2(n(t, \cdot), m(t, \cdot)) + C |Z_n(t) - Z_m(t)|.$$

Monge-Kantorovich:

$$\left| \int f(x)n(s, x) dx - \int f(x)m(s, x) dx \right| \leq \|f'\|_\infty W_1(n(s, \cdot), m(s, \cdot)).$$

## Ideas from non-convex granular media equations

Let  $n, m \in \mathcal{P}_2(\mathbb{R})$  solutions of the following model:

$$\partial_t n - \Delta_x n = \nabla_x \cdot (n \nabla_x W *_x n + \nabla_x V),$$

where  $*$  is a convolution and  $W$  a long-range interaction potential. If  $W$  and  $V$  are convex, solutions do contract for the  $W_2$  distance:

$$\frac{d}{dt} W_2(n(t, \cdot), m(t, \cdot)) \leq -\kappa W_2(n(t, \cdot), m(t, \cdot)).$$

If  $V$  is slightly non-convex close to the origin, this property persists<sup>6</sup>

---

<sup>6</sup>J. A. Carrillo, R. J. McCann, C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, **179**(2):217–263 (2006).

F. Bolley, I. Gentil, A. Guillin, Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations. *J. Funct. Anal.*, **263**(8):2430–2457 (2012).





## Ideas from non-convex granular media equations

The analysis relies on the convexity of the energy

$$E(n(t, \cdot)) = \int n(t, x) \log(n(t, x)) dx \\ + \frac{1}{2} \int \int W(x - y) n(t, x) n(t, y) dx dy + \int V(x) n(t, x) dx,$$

along geodesics. The first part of this energy is always convex.

- ▶ **Tails:** Far from the center of mass of the distribution,  $W * n(t, \cdot) + V$  is convex.
- ▶ **Core:** Close to the center of mass,  $n(t, \cdot)$  is bounded from below, so that  $\int n(t, x) \log(n(t, x)) dx$  is strictly convex.

The convexity of the energy implies the contraction estimate.

## Wasserstein estimate with a lower bound assumption

$$\begin{aligned} W_2 \left( \frac{(1 + \alpha a)n(s, \cdot)}{1 + \alpha I_n(s)}, \frac{(1 + \alpha a)m(s, \cdot)}{1 + \alpha I_m(s)} \right) \\ \leq \mathcal{O}(\alpha) W_2(n(t, \cdot), m(t, \cdot)) + C |Z_n(t) - Z_m(t)|. \end{aligned}$$

- **Tails:** Since  $\int a(x)n(t, x) dx \neq \int b(x)n(t, x) dx$ , large distance couplings appear. We take advantage of exponential tails estimates:

$$\forall x \geq R', \quad \partial_x n(t, x) \leq n(t, x).$$

- **core:** If  $a$  is compactly supported,  $n(t, \cdot)$  is bounded away from 0 around  $\text{supp } n(t, \cdot)$ . Then,

$$\begin{aligned} W_2 \left( \left( 1 - \alpha + \frac{\alpha a}{\int a(x) n(x) dx} \right) n, \left( 1 - \alpha + \frac{\alpha a}{\int a(x) m(x) dx} \right) m \right) \\ \leq \mathcal{O}(\alpha) W_2(n, m) + C |Z_n(t) - Z_m(t)|, \end{aligned}$$



## Convergence

$$\begin{cases} \frac{d}{dt} w'(t) \leq - \left( \frac{\sqrt{2}-1}{\sqrt{2}} + C\sqrt{\alpha} \right) w(t) + C\alpha |Z_m(t) - Z_n(t)|, \\ \frac{d}{dt} |Z_m(t) - Z_n(t)| \leq C\alpha w(t) + \alpha \frac{F'(\bar{Z})}{2} |Z_m(t) - Z_n(t)|. \end{cases}$$

Then,

$$\begin{aligned} & \frac{d}{dt} (\sqrt{\alpha} w(t) + |Z_n \circ \varphi_n(t) - Z_m \circ \varphi_m(t)|) \\ & \leq \left[ -\sqrt{\alpha} \frac{\sqrt{2}-1}{\sqrt{2}} + C\alpha \right] w(t) \\ & \quad + \left[ \alpha \frac{F'(\bar{Z})}{2} + \sqrt{\alpha} C\alpha \right] |Z_m(t) - Z_n(t)| \\ & \leq \alpha \frac{F'(\bar{Z})}{3} (\sqrt{\alpha} w(t) + |Z_n \circ \varphi_n(t) - Z_m \circ \varphi_m(t)|). \end{aligned}$$

## Convergence

We have shown that

$$(\sqrt{\alpha} w(t) + |Z_n \circ \varphi_n(t) - Z_m \circ \varphi_m(t)|) \leq C e^{\frac{F'(\bar{Z})}{3} \alpha t}.$$

Then, in particular,

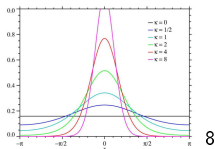
$$W_2(n(t, \cdot), m(t, \cdot)) \leq C e^{\frac{F'(\bar{Z})}{3} \alpha t}.$$

If we show that there exists a steady-state  $\bar{n}$  close to  $\Gamma_{2\sigma^2}(\cdot - \bar{Z})$ , and in particular,

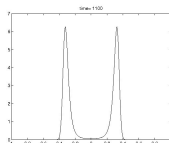
$$W_2(n(t, \cdot), \bar{n}) \leq C e^{\frac{F'(\bar{Z})}{3} \alpha t}.$$

## Perspectives

- ▶ (With B. Perthame and L. Kanzler) Intermediate distances between  $W_1$  and  $W_2$ , inspired by Eberle<sup>7</sup> and Hairer-Mattingly.
- ▶ Consider other fitness landscapes, still with a small parameter,



- ▶ Considering other aggregation questions with a small parameter,



---

<sup>7</sup>A. Eberle, PTRF 2016.

<sup>8</sup>Degond et al, Séminaire Laurent Schwartz.

Thank you for your attention.