SOME MATHEMATICAL MODELS OF EVOLUTION

II: Expanding Populations

Alison Etheridge
University of Oxford

Thanks

Félix Foutel-Rodier, Sorbonne Université/Collège de France

Sarah Penington, University of Bath

Wright-Fisher model with selection

Two types a, A Population size N, discrete generations

- Each individual produces (effectively infinite) number of gametes;
- A type a produces $(1-s)$ times as many gametes as a type A;
- Sample N offspring uniformly at random from pool of gametes.

If proportion of a-alleles in parental population is p, then the probability that an offspring is type a is

$$\frac{(1-s)p}{(1-s)p+(1-p)} = \frac{(1-s)p}{1-sp} \approx p - sp(1-p).$$

Wright Fisher diffusion with selection

For large populations, approximate dynamics by

$$dp = -sp(1-p)dt + \sqrt{\frac{1}{N_e}p(1-p)}dW_t$$

$\mathbb{P}$[deleterious mutation arising in single individual fixes] $\approx e^{-2N_e s}$

Selection more effective when population size is bigger

Wright Fisher diffusion with selection

For large populations, approximate dynamics by

$$dp = -sp(1-p)dt + \sqrt{\frac{1}{N_e}p(1-p)}dW_t$$

$\mathbb{P}$[deleterious mutation arising in single individual fixes] $\approx e^{-2N_e s}$

Selection more effective when population size is bigger

less effective at the front of an expanding population?

Space

. . . there can be advantages to living life on the edge, which might compensate for deleterious mutations

Accumulation of deleterious mutations (no space)

Assumptions:

- asexual population;
- constant size N;
- individuals accumulate mutations;
- all mutations are deleterious.

An individual that has accumulated k deleterious mutations has *relative* fitness $(1-s)^k$.

Notation:
$X_k(t) =$ proportion of individuals at time t that carry exactly k deleterious mutations.

Haigh's model

Wright-Fisher style dynamics

- Population evolves in discrete generations.
- Each offspring (independently) chooses parent (in a weighted way) from previous generation.
- All mutations are passed from parent to offspring.
- In addition, each offspring accumulates Poisson number of new mutations.

A more mathematical formulation

Frequency profile in generation t: $\mathbf{x}(t) = (x_k(t))_{k=0,1,\dots} \in \mathcal{P}(\mathbb{N}_0)$

$Nx_k(t)$ individuals in generation t carry k mutations.
Number of mutations inherited from parent:

$$\mathbb{P}[H = k] = \frac{(1-s)^k x_k(t)}{Z(t)},$$

where $Z(t) = \sum_{k=0}^{\infty} (1-s)^k x_k(t)$.

Number of new mutations: $J \sim \texttt{Poiss}(\mu)$.

Let $K_1, \dots, K_N$ be independent copies of $H + J$.

$$\mathbf{X}_k(t+1) = \frac{1}{N} \sharp\{i : K_i = k\}.$$

Infinite population limit: Poisson frequency profile

Suppose

$$x_k(t) = \frac{\theta^k}{k!} e^{-\theta}, \qquad k \in \mathbb{N}_0; \qquad Z(t) = e^{-\theta s}.$$

$$\mathbb{P}[H = k] = \frac{(1-s)^k x_k(t)}{Z(t)} = \frac{(1-s)^k \theta^k e^{-\theta(1-s)}}{k!}.$$

Then $H + J \sim \texttt{Poiss}\big((1-s)\theta + \mu\big)$.

As $N \to \infty$

$$X_k(t+1) \to \frac{\big((1-s)\theta + \mu\big)^k e^{-\big((1-s)\theta+\mu\big)}}{k!}.$$

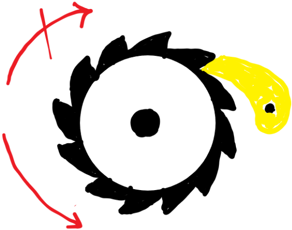
Choose $\theta = \mu/s$, $\mathbf{x}(t+1) = \mathbf{x}(t)$.

Adding noise: Muller's ratchet

For $N \to \infty$, proportion of population with no mutations $x_0(t) \to e^{-\mu/s} = e^{-\theta}$ as $t \to \infty$.

For $N < \infty$, after finite random time $N x_0(T_0) = 0$. *The ratchet clicks.*

$$N x_1(T_1) = 0, N x_2(T_2) = 0 \ldots$$



Population becomes inexorably less fit.

Introducing space

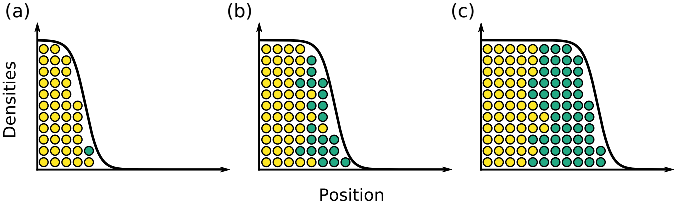
- Population expanding its range.

Introducing space

- Population expanding its range.
- Mutations arising at the front, can reach high frequency, even if they confer no selective advantage.

Introducing space

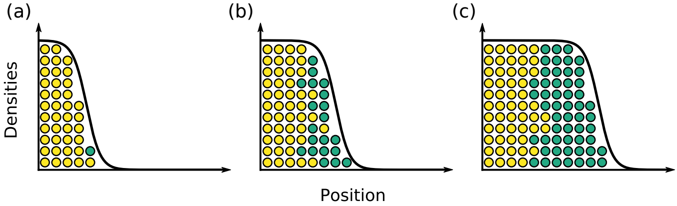
- Population expanding its range.
- Mutations arising at the front, can reach high frequency, even if they confer no selective advantage.



Gene surfing;

Introducing space

- Population expanding its range.
- Mutations arising at the front, can reach high frequency, even if they confer no selective advantage.



Gene surfing; Deleterious surfers $\rightsquigarrow$ Expansion load

What if life is not so nice on the edge?

What if there is an Allee effect?

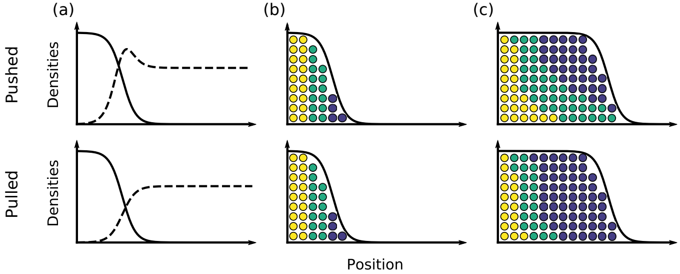
- Maximum per capita growth rate achieved at intermediate density.

What if there is an Allee effect?

- Maximum per capita growth rate achieved at intermediate density.
- No Allee effect $\rightsquigarrow$ pulled wave;
- Strong Allee effect $\rightsquigarrow$ pushed wave,

What if there is an Allee effect?

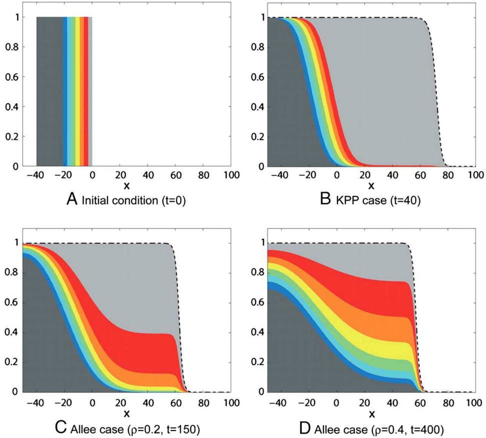
- Maximum per capita growth rate achieved at intermediate density.
- No Allee effect $\rightsquigarrow$ pulled wave;
- Strong Allee effect $\rightsquigarrow$ pushed wave,



Allee effect deterministic case

Roques et al. PNAS (2012).

$$\partial_t u = \partial_{xx} u + u(1-u)(u-\rho), \qquad \rho \in (0, 1/2).$$



A Initial condition (t=0)

B KPP case (t=40)

C Allee case (ρ=0.2, t=150)

D Allee case (ρ=0.4, t=400)

An individual based model

$n_{i,k}(t)$ number of individuals in deme $i \in \mathbb{Z}$ carrying k mutations.

$N_i(t) = \sum_{k=0}^{\infty} n_{i,k}(t)$, total population size in deme i.

An individual based model

$n_{i,k}(t)$ number of individuals in deme $i \in \mathbb{Z}$ carrying k mutations.

$N_i(t) = \sum_{k=0}^{\infty} n_{i,k}(t)$, total population size in deme i.

An individual in deme i and carrying k mutations

- gives birth to a new individual at rate $\lambda_k(N_i)$;
- dies at rate $\delta(N_i)$;
- migrates to $i \pm 1$ at rate m.

An individual based model

$n_{i,k}(t)$ number of individuals in deme $i \in \mathbb{Z}$ carrying k mutations.

$N_i(t) = \sum_{k=0}^{\infty} n_{i,k}(t)$, total population size in deme i.

An individual in deme i and carrying k mutations

- gives birth to a new individual at rate $\lambda_k(N_i)$;
- dies at rate $\delta(N_i)$;
- migrates to $i \pm 1$ at rate m.

$$\lambda_k(N_i) = r(1-s)^k(B\frac{N_i}{N}+1), \quad \delta(N_i) = r(B\frac{N_i}{N}+1)\frac{N_i}{N}.$$

An individual based model

$n_{i,k}(t)$ number of individuals in deme $i \in \mathbb{Z}$ carrying k mutations.

$N_i(t) = \sum_{k=0}^{\infty} n_{i,k}(t)$, total population size in deme i.

An individual in deme i and carrying k mutations

- gives birth to a new individual at rate $\lambda_k(N_i)$;
- dies at rate $\delta(N_i)$;
- migrates to $i \pm 1$ at rate m.

$$\lambda_k(N_i) = r(1-s)^k (B\frac{N_i}{N} + 1), \quad \delta(N_i) = r(B\frac{N_i}{N} + 1)\frac{N_i}{N}.$$

With probability μ, offspring has one more mutation than parent.

The role of B

$$\lambda_k(N_i) = r(1-s)^k(B\frac{N_i}{N}+1), \quad \delta(N_i) = r(B\frac{N_i}{N}+1)\frac{N_i}{N}.$$

The role of B

$$\lambda_k(N_i) = r(1-s)^k(B\frac{N_i}{N} + 1), \quad \delta(N_i) = r(B\frac{N_i}{N} + 1)\frac{N_i}{N}.$$

$B = 0$, stochastic version of usual logistic growth, with ‘carrying capacity’ N.

The role of B

$$\lambda_k(N_i) = r(1-s)^k(B\frac{N_i}{N}+1), \quad \delta(N_i) = r(B\frac{N_i}{N}+1)\frac{N_i}{N}.$$

$B = 0$, stochastic version of usual logistic growth, with ‘carrying capacity’ N.

$n \mapsto \lambda_0(n) - \delta(n)$ maximised at

$$n = \max\left\{0, \frac{N(B-1)}{2B}\right\}.$$

The role of B

$$\lambda_k(N_i) = r(1-s)^k\Big(B\frac{N_i}{N}+1\Big), \quad \delta(N_i) = r\Big(B\frac{N_i}{N}+1\Big)\frac{N_i}{N}.$$

$B = 0$, stochastic version of usual logistic growth, with 'carrying capacity' N.

$n \mapsto \lambda_0(n) - \delta(n)$ maximised at

$$n = \max\Big\{0, \frac{N(B-1)}{2B}\Big\}.$$

For $B \leq 1$, no Allee effect; $B > 1$ Allee effect.
Increasing B increases strength of Allee effect.

Large population scalings

Scale population density N, diffusive scaling of spatial motion, *no scaling* of reproduction rates.

For $i \in \mathbb{Z}$ and $x = i/L$ set

$$\forall k \geq 0, \; u_k^N(x,t) = \frac{n_{i,k}^N(t)}{N};$$

Large population scalings

Scale population density N, diffusive scaling of spatial motion, *no scaling* of reproduction rates.

For $i \in \mathbb{Z}$ and $x = i/L$ set

$$\forall k \geq 0,\ u_k^N(x,t) = \frac{n_{i,k}^N(t)}{N};$$

$$U = \sum_{k \geq 0} u_k, \quad u_{-1} \equiv 0,$$

$$\frac{\partial u_k}{\partial t} = m\frac{\partial^2 u_k}{\partial x^2} + r(BU+1)\big(u_k((1-\mu)(1-s)^k - U) \\ + \mu(1-s)^{k-1} u_{k-1}.$$

Weak selection and mutation

We suppose that $s, \mu \ll 1$.

Limiting model becomes:

$$U = \sum_{k \geq 0} u_k, \quad u_{-1} \equiv 0,$$

$$\frac{\partial u_k}{\partial t} = m \frac{\partial^2 u_k}{\partial x^2} + r(BU+1)\big[u_k(1 - ks - U) \\ + \mu(u_{k-1} - u_k)\big].$$

Weak selection and mutation

We suppose that $s, \mu \ll 1$.

Limiting model becomes:

$$U = \sum_{k \geq 0} u_k, \quad u_{-1} \equiv 0,$$

$$\frac{\partial u_k}{\partial t} = m \frac{\partial^2 u_k}{\partial x^2} + r(BU+1)\big[u_k(1 - ks - U) \\ + \mu(u_{k-1} - u_k)\big].$$

If $u_k(x,0)$ independent of x, then so is $u_k(x,t)$ for all t.
For each $k_0 \geq 0$, there corresponds a steady state solution of the form

$$u_k^* = (1 - \mu - k_0 s) e^{-\theta} \frac{\theta^{k-k_0}}{(k-k_0)!}, \qquad k \geq k_0$$

A travelling wave solution

If initial genetic composition is $\texttt{Poiss}(\theta)$ for all $x \in \mathbb{R}$, with $\theta = \mu/s$, then it remains so. Then,

$$\frac{\partial U}{\partial t} = m\frac{\partial^2 U}{\partial x^2} + rU(BU+1)(1-\mu-U).$$

A travelling wave solution

If initial genetic composition is $\mathtt{Poiss}(\theta)$ for all $x \in \mathbb{R}$, with $\theta = \mu/s$, then it remains so. Then,

$$\frac{\partial U}{\partial t} = m\frac{\partial^2 U}{\partial x^2} + rU(BU+1)(1-\mu-U).$$

Hadeler & Rothe (1975): travelling wave solution for all speeds $c \geq c_0$, where c_0 is given by

$$c_0 = \begin{cases} 2\sqrt{mr(1-\mu)} & \text{if } B \leq \frac{2}{1-\mu} \text{ (pulled)} \\ \\ \sqrt{\frac{mr}{2B}}(B(1-\mu)+2) & \text{if } B \geq \frac{2}{1-\mu} \text{ (pushed).} \end{cases}$$

A travelling wave solution

If initial genetic composition is $\texttt{Poiss}(\theta)$ for all $x \in \mathbb{R}$, with $\theta = \mu/s$, then it remains so. Then,

$$\frac{\partial U}{\partial t} = m\frac{\partial^2 U}{\partial x^2} + rU(BU+1)(1-\mu-U).$$

Hadeler & Rothe (1975): travelling wave solution for all speeds $c \geq c_0$, where c_0 is given by

$$c_0 = \begin{cases} 2\sqrt{mr(1-\mu)} & \text{if } B \leq \frac{2}{1-\mu} \text{ (pulled)} \\ \sqrt{\frac{mr}{2B}}(B(1-\mu)+2) & \text{if } B \geq \frac{2}{1-\mu} \text{ (pushed).} \end{cases}$$

If $\widehat{U}$ wave profile of a travelling wave with speed c, then

$$\forall k \geq 0, \forall x \in \mathbb{R}, \forall t \geq 0,\ u_k(t,x) = e^{-\theta}\frac{\theta^k}{k!}\widehat{U}(x-ct)$$

is a travelling wave solution to our system.

A first look at recovery of fitness

The ‘population’ travelling wave above connects the stable limit

$$u_k^* = (1 - \mu - k_0 s)e^{-\theta}\frac{\theta^{k-k_0}}{(k-k_0)!}$$

with $k_0 = 0$ to the trivial null equilibrium.

A first look at recovery of fitness

The ‘population’ travelling wave above connects the stable limit

$$u_k^* = (1-\mu-k_0 s)e^{-\theta}\frac{\theta^{k-k_0}}{(k-k_0)!}$$

with $k_0 = 0$ to the trivial null equilibrium.

Consider the system with initial condition

$$\forall x \in \mathbb{R}, \forall k \geq 0,\ u_k(0,x) = \begin{cases} (1-\mu-s)e^{-\theta}\frac{\theta^{k-1}}{(k-1)!} & \text{if } x > x_0 \\ (1-\mu)e^{-\theta}\frac{\theta^k}{k!} & \text{if } x \leq x_0, \end{cases}$$

for some $x_0 \in \mathbb{R}$.

A first look at recovery of fitness

The 'population' travelling wave above connects the stable limit

$$u_k^* = (1 - \mu - k_0 s) e^{-\theta} \frac{\theta^{k-k_0}}{(k-k_0)!}$$

with $k_0 = 0$ to the trivial null equilibrium.

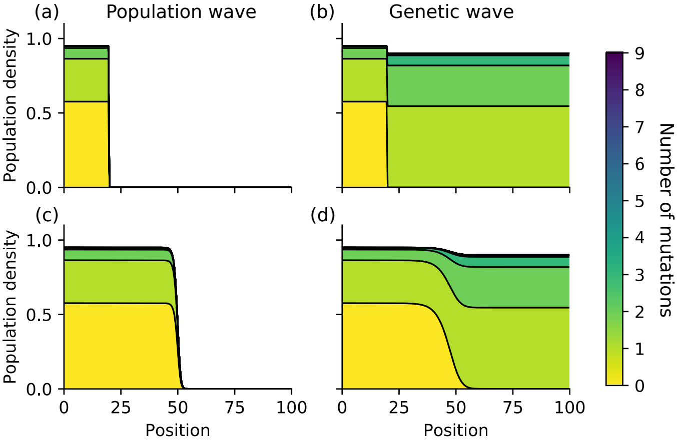
Consider the system with initial condition

$$\forall x \in \mathbb{R}, \forall k \geq 0,\ u_k(0,x) = \begin{cases} (1-\mu-s)e^{-\theta}\frac{\theta^{k-1}}{(k-1)!} & \text{if } x > x_0 \\ (1-\mu)e^{-\theta}\frac{\theta^k}{k!} & \text{if } x \leq x_0, \end{cases}$$

for some $x_0 \in \mathbb{R}$.

Individuals without mutations are able to invade the region where 'the ratchet has clicked'. The 'wave of expansion' is always pulled. It has speed $\mathcal{O}(\sqrt{s})$.

Population waves and genetic waves



Adding noise

We expect our system of deterministic equations to be replaced by

$$U = \sum_{k\geq 0} u_k, \quad u_{-1} \equiv 0,$$

$$\partial_t u_k = m\partial_{xx} u_k + r(BU+1)\big(u_k(1-ks-U) + \mu(u_{k-1}-u_k)\big)$$

$$+ \sqrt{\frac{r}{N} u_k (BU+1)(1-ks+U)}\dot{W}_k,$$

where $(\dot{W}_k)_{k\geq 0}$ are independent space-time white noises.

Adding noise

We expect our system of deterministic equations to be replaced by

$$U = \sum_{k \geq 0} u_k, \quad u_{-1} \equiv 0,$$

$$\partial_t u_k = m \partial_{xx} u_k + r(BU+1)\big(u_k(1-ks-U) + \mu(u_{k-1}-u_k)\big)$$

$$+ \sqrt{\frac{r}{N} u_k (BU+1)(1-ks+U)}\dot{W}_k,$$

where $(\dot{W}_k)_{k \geq 0}$ are independent space-time white noises.

We have not analysed this system but instead simulated our individual based model.

Defining a click

Write

$$n_k^{\texttt{max}}(t) = \max\{i \in \mathbb{Z} : n_{i,k}(t) > 0\}$$

(location of rightmost individual carrying k mutations at time t).

Defining a click

Write

$$n_k^{\mathtt{max}}(t) = \max\{i \in \mathbb{Z} : n_{i,k}(t) > 0\}$$

(location of rightmost individual carrying k mutations at time t).

$$T_1 = \inf \Big\{ t \geq 0 : \exists s \geq t, n_1^{\mathtt{max}}(s) - n_0^{\mathtt{max}}(s) > d \text{ and } \forall r \in [t,s], n_1^{\mathtt{max}}(r) > n_0^{\mathtt{max}}(r) \Big\}.$$

T_1 is the first moment when individuals with one mutation get ahead of individuals with no mutations *and* will get d demes ahead before being caught up.

Defining a click

Write

$$n_k^{\texttt{max}}(t) = \max\{i \in \mathbb{Z} : n_{i,k}(t) > 0\}$$

(location of rightmost individual carrying k mutations at time t).

$$T_1 = \inf\Big\{t \geq 0 : \exists s \geq t, n_1^{\texttt{max}}(s) - n_0^{\texttt{max}}(s) > d \\ \text{and } \forall r \in [t,s], n_1^{\texttt{max}}(r) > n_0^{\texttt{max}}(r)\Big\}.$$

T_1 is the first moment when individuals with one mutation get ahead of individuals with no mutations *and* will get d demes ahead before being caught up.

We set $d = 30$. Once $n_1^{\texttt{max}} - n_0^{\texttt{max}} > d$ it is unlikely that the inner wave catches up before entire habitat colonised.

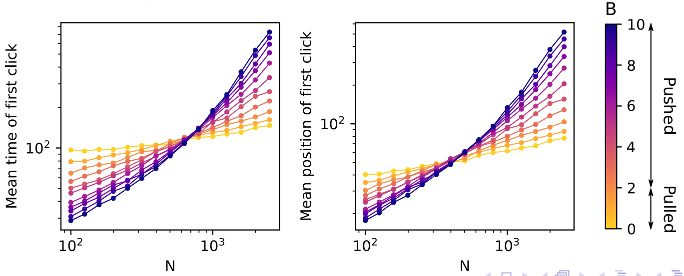
Click rate and the Allee effect

$$\partial_t u_k = m\partial_{xx}u_k + r(BU+1)\big(u_k(1-ks-U) + \mu(u_{k-1}-u_k)\big) + \sqrt{\frac{r}{N}u_k(BU+1)(1-ks+U)}\dot{W}_k,$$

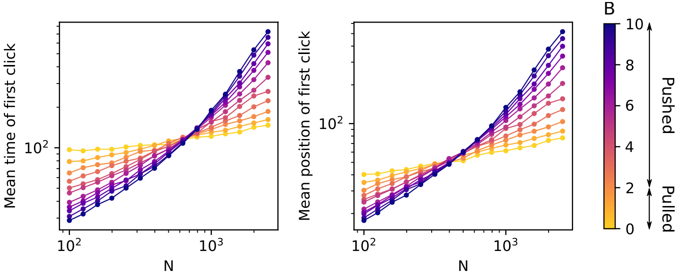
Changing B has antagonistic effects on click rate:

- increases the strength of the Allee effect (slows down clicks);
- increases the strength of drift (speeds up clicks).

Which prevails depends on N

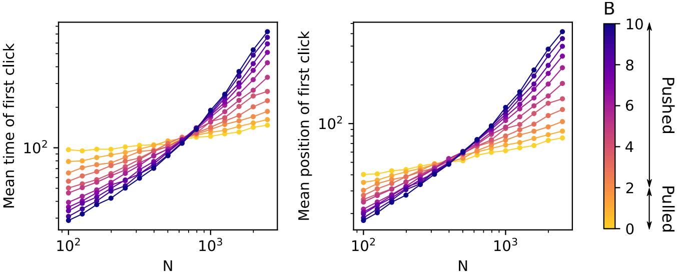


Click rate and the Allee effect



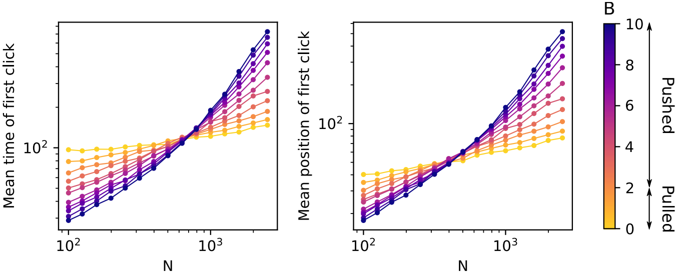
- Fix B, click rate decreases as N increases;

Click rate and the Allee effect



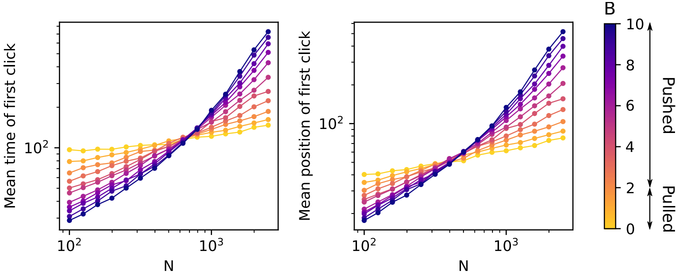
- Fix B, click rate decreases as N increases;
- Small N, genetic drift prevails, gene flow less efficient at restoring diversity in the front;

Click rate and the Allee effect



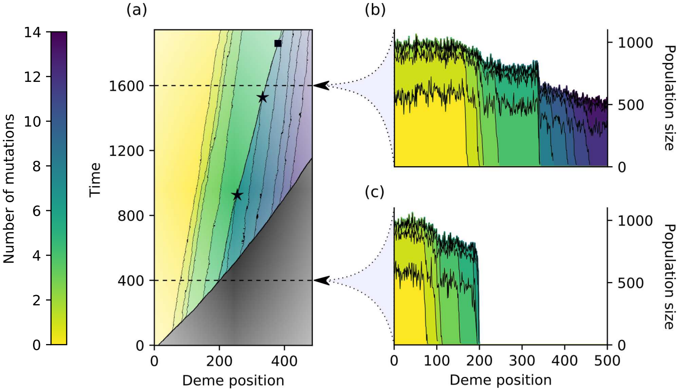
- Fix B, click rate decreases as N increases;
- Small N, genetic drift prevails, gene flow less efficient at restoring diversity in the front;
- Large N, gene flow prevails;

Click rate and the Allee effect



- Fix B, click rate decreases as N increases;
- Small N, genetic drift prevails, gene flow less efficient at restoring diversity in the front;
- Large N, gene flow prevails;
- N has more impact on T_1 for pushed waves, almost no clicks for large N.

How do mutations accumulate?



Two dimensions

At times $t = 1, 2, \ldots, 150$, record least number of mutations in each newly colonised deme.

Two dimensions

At times $t = 1, 2, \ldots, 150$, record least number of mutations in each newly colonised deme.

$n_{i,j;k}(t) :=$ number of individuals in deme (i, j) with k mutations.
$N_{i,j} :=$ total number of individuals in deme (i, j).

Two dimensions

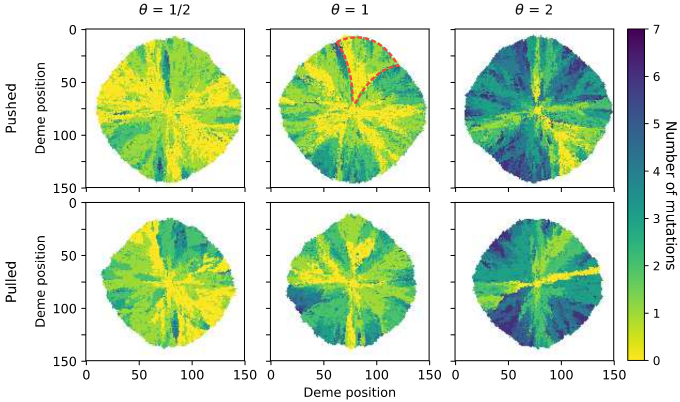
At times $t = 1, 2, \ldots, 150$, record least number of mutations in each newly colonised deme.

$n_{i,j;k}(t) :=$ number of individuals in deme (i, j) with k mutations.
$N_{i,j} :=$ total number of individuals in deme (i, j).

$$t_{i,j}^{\texttt{col}} = \inf \Big\{ t \in \{1, 2, \ldots, 150\} : N_{i,j}(t) > 0 \text{ and } N_{i,j}(t-1) = 0 \Big\};$$

$$k_{i,j}^{\texttt{col}} = \inf \Big\{ k \geq 0 : n_{i,j;k}(t_{i,j}^{\texttt{col}}) > 0 \Big\}.$$

Two dimensions



Genealogies of pulled and pushed waves?

Stochastic Fisher-KPP equation

$$\partial_t u = \partial_{xx} u + u(1-u) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

Stochastic analogue of example of Roques et al. (2012)

$$\partial_t u = \partial_{xx} u + u(1-u)(u-\rho) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

$\rho \in (0, 1/2)$

Genealogies of pulled and pushed waves?

Stochastic Fisher-KPP equation

$$\partial_t u = \partial_{xx} u + u(1-u) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

Genealogy of sample from close to the wavefront dominated by rare events in which individual far in the front produces large family; on suitable timescales Bolthausen-Sznitman coalescent

Stochastic analogue of example of Roques et al. (2012)

$$\partial_t u = \partial_{xx} u + u(1-u)(u-\rho) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

$\rho \in (0, 1/2)$

Genealogies of pulled and pushed waves?

Stochastic Fisher-KPP equation

$$\partial_t u = \partial_{xx} u + u(1-u) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

Genealogy of sample from close to the wavefront dominated by rare events in which individual far in the front produces large family; on suitable timescales Bolthausen-Sznitman coalescent

Stochastic analogue of example of Roques et al. (2012)

$$\partial_t u = \partial_{xx} u + u(1-u)(u-\rho) + \sqrt{\frac{1}{\rho_e} u(1-u)}\dot{W}$$

$\rho \in (0, 1/2)$
For a discrete space, individual based analogue (a spatial Moran model), the genealogy of sample from close to the wavefront on suitable timescales given by Kingman coalescent (Penington 2000).