Some mathematical models from population genetics

5: Muller's ratchet and the rate of adaptation

Alison Etheridge

University of Oxford

joint work with Peter Pfaffelhuber (Vienna), Anton Wakolbinger (Frankfurt)

Charles Cuthbertson (Oxford), Feng Yu (Bristol)

Muller's ratchet

Asexually reproducing population. All mutations deleterious.

Chromosomes passed down as indivisible blocks.

Muller’s ratchet

Asexually reproducing population. All mutations deleterious.
Chromosomes passed down as indivisible blocks.

Number of deleterious mutations accumulated along any ancestral line can only increase.

Muller’s ratchet

Asexually reproducing population. All mutations deleterious. Chromosomes passed down as indivisible blocks.

Number of deleterious mutations accumulated along any ancestral line can only increase.

When everyone in the current ‘best’ class has accumulated at least one additional bad mutation the ratchet clicks.

Muller's ratchet

Asexually reproducing population. All mutations deleterious. Chromosomes passed down as indivisible blocks.

Number of deleterious mutations accumulated along any ancestral line can only increase.

When everyone in the current 'best' class has accumulated at least one additional bad mutation the ratchet clicks.

How many generations will it take for an asexually reproducing population to lose its best class?

Haigh's model

Wright-Fisher model:
Individuals in $(t+1)$st generation select parent at random from generation t.

Probability individual which has accumulated k mutations is selected as parent proportional to *relative fitness* $(1-s)^k$.

Number of mutations carried by offspring then $k+J$, where $J \sim \texttt{Poiss}(\lambda)$ (independent).

Type frequencies: $\mathbf{x}(t) = (x_k(t))_{k=0,1,\dots}$
$Nx_k(t) =$ No. of individuals carrying exactly k mutations at time t.

Type frequencies: $\mathbf{x}(t) = (x_k(t))_{k=0,1,\ldots}$
$Nx_k(t) =$ No. of individuals carrying exactly k mutations at time t.
H an $\mathbb{N}_0$-valued random variable.

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

Type frequencies: $\mathbf{x}(t) = (x_k(t))_{k=0,1,\dots}$

$Nx_k(t) =$ No. of individuals carrying exactly k mutations at time t.

H an $\mathbb{N}_0$-valued random variable.

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

$J \sim \texttt{Poiss}(\lambda)$ independent of H.

Type frequencies: $\mathbf{x}(t) = (x_k(t))_{k=0,1,\dots}$
$Nx_k(t) =$ No. of individuals carrying exactly k mutations at time t.
H an $\mathbb{N}_0$-valued random variable.

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

$J \sim \texttt{Poiss}(\lambda)$ independent of H.
$K_1, K_2, \dots, K_N$ independent copies of $H + J$.
Random type frequencies in next generation are

$$X_k(t+1) = \frac{1}{N}\#\{i : K_i = k\}.$$

Infinite populations

As $N \to \infty$, LLN$\Rightarrow$

$$\mathbf{x}(t+1) = \mathbb{E}_{\mathbf{x}(t)}[X(t+1)].$$

Infinite populations

As $N \to \infty$, LLN$\Rightarrow$

$$\mathbf{x}(t+1) = \mathbb{E}_{\mathbf{x}(t)}[X(t+1)].$$

Suppose $\mathbf{x}(t) \sim \mathtt{Poiss}(\alpha)$.

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

Then $H \sim \mathtt{Poiss}(\alpha(1-s))$, $J \sim \mathtt{Poiss}(\lambda)$, so

$$H + J \sim \mathtt{Poiss}(\alpha(1-s) + \lambda).$$

Infinite populations

As $N \to \infty$, LLN$\Rightarrow$

$$\mathbf{x}(t+1) = \mathbb{E}_{\mathbf{x}(t)}[X(t+1)].$$

Suppose $\mathbf{x}(t) \sim \texttt{Poiss}(\alpha)$.

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

Then $H \sim \texttt{Poiss}(\alpha(1-s))$, $J \sim \texttt{Poiss}(\lambda)$, so

$$H + J \sim \texttt{Poiss}(\alpha(1-s) + \lambda).$$

Poisson weights $\mapsto$ Poisson weights.

For every initial condition with $x_0 > 0$, the solution to the deterministic dynamics converges as $t \to \infty$ to the stationary point

$$\pi := \texttt{Poiss}(\lambda/s).$$

Back to finite populations

Write $k*$ for the number of mutations carried by individuals in fittest class.

Back to finite populations

Write $k*$ for the number of mutations carried by individuals in fittest class.

$$\mathbf{Y} := (Y_k)_{k=0,1,\ldots} := (X_{k*+k})_{k=0,1,2\ldots}$$

forms a recurrent Markov chain.

Back to finite populations

Write $k*$ for the number of mutations carried by individuals in fittest class.

$$\mathbf{Y} := (Y_k)_{k=0,1,\ldots} := (X_{k*+k})_{k=0,1,2\ldots}$$

forms a recurrent Markov chain.

Condition on $\mathbf{Y}(t) = \mathbf{y}(t)$. Size of new best class, $y_0(t+1) \sim \texttt{Binom}(N, p_0(t))$, with $p_0(t)$ probability of sampling parent from best class *and* not acquiring any additional mutations:

$$p_0(t) = \frac{y_0(t)}{W(t)} e^{-\lambda}, \quad W(t) = \sum_{i=0}^{\infty} y_i(t)(1-s)^i.$$

Back to finite populations

Write $k*$ for the number of mutations carried by individuals in fittest class.

$$\mathbf{Y} := (Y_k)_{k=0,1,\ldots} := (X_{k*+k})_{k=0,1,2\ldots}$$

forms a recurrent Markov chain.
Condition on $\mathbf{Y}(t) = \mathbf{y}(t)$. Size of new best class, $y_0(t+1) \sim \mathtt{Binom}(N, p_0(t))$, with $p_0(t)$ probability of sampling parent from best class *and* not acquiring any additional mutations:

$$p_0(t) = \frac{y_0(t)}{W(t)} e^{-\lambda}, \quad W(t) = \sum_{i=0}^{\infty} y_i(t)(1-s)^i.$$

Evolution of best class determined by $W(t)$, the *mean fitness* in the population.

Elements of Haigh’s analysis

- Immediately after a click, the type frequencies are

$$\tilde{\pi} := \frac{1}{1 - \pi_0}(\pi_1, \pi_2, \ldots).$$

Elements of Haigh's analysis

- Immediately after a click, the type frequencies are

$$\tilde{\pi} := \frac{1}{1-\pi_0}(\pi_1, \pi_2, \ldots).$$

- **Phase one:** deterministic dynamical system dominates, decaying exponentially fast towards its equilibrium

Elements of Haigh's analysis

- Immediately after a click, the type frequencies are

$$\tilde{\pi} := \frac{1}{1-\pi_0}(\pi_1, \pi_2, \ldots).$$

- **Phase one:** deterministic dynamical system dominates, decaying exponentially fast towards its equilibrium

- **Phase two:** the 'bulk' of the population changes only slowly. Mean fitness assumed constant and then No. of individuals in best class approximated by Galton-Watson branching process with Poisson offspring distribution.

The Fleming Viot diffusion

Relevant for applications are: large N, small s, small λ.

The Fleming Viot diffusion

Relevant for applications are: large N, small s, small λ.
Dynamics then captured by

$$dX_k = \left(\sum_j s(j-k)X_jX_k + \lambda(X_{k-1} - X_k) \right) dt + \sum_{j \neq k} \sqrt{\frac{1}{N} X_j X_k} dW_{jk},$$

$k = 0, 1, 2, \ldots$ where $X_{-1} = 0$ and $(W_{jk})_{j>k}$ array of independent Brownian motions, $W_{kj} := -W_{jk}$.

The Fleming Viot diffusion

Relevant for applications are: large N, small s, small λ.

Dynamics then captured by

$$dX_k = \left(\sum_j s(j-k) X_j X_k + \lambda (X_{k-1} - X_k) \right) dt + \sum_{j \neq k} \sqrt{\frac{1}{N} X_j X_k} dW_{jk},$$

$k = 0, 1, 2, \ldots$ where $X_{-1} = 0$ and $(W_{jk})_{j>k}$ array of independent Brownian motions, $W_{kj} := -W_{jk}$.

As before $Y_k = X_{k*+k}$,

$$dY_0 = s(M_1(\mathbf{Y}) - \lambda) Y_0(t) dt + \sqrt{\frac{1}{N} Y_0 (1 - Y_0)} dW_0, \quad M_1(\mathbf{Y}) = \sum_j j Y_j.$$

Infinite population limit

$$dx_k = (s(M_1(\mathbf{x}) - k) - \lambda)x_k + \lambda x_{k-1})\,dt, \quad k = 0, 1, 2, \ldots$$

with $x_{-1} = 0$.

Infinite population limit

$$dx_k = \left(s(M_1(\mathbf{x}) - k) - \lambda)x_k + \lambda x_{k-1}\right) dt, \quad k = 0, 1, 2, \ldots$$

with $x_{-1} = 0$.

Transform into system of equations for cumulants:

$$\log \sum_{k=0}^{\infty} x_k e^{-\xi k} = \sum_{k=1}^{\infty} \kappa_k \frac{(-\xi)^k}{k!}.$$

Assume $x_0 > 0$ and set $\kappa_0 = -\log x_0$. Then

$$\dot{\kappa}_k = -s\kappa_{k+1} + \lambda, \qquad k = 0, 1, 2, \ldots$$

System can be solved. In particular,

$$\kappa_1(t) = \sum_{k=0}^{\infty} k x_k(t) = -\frac{\partial}{\partial \xi} \log \sum_{k=0}^{\infty} x_k(0) e^{-\xi k} \Bigg|_{\xi = st} + \frac{\lambda}{s}(1 - e^{-st}).$$

System can be solved. In particular,

$$\kappa_1(t) = \sum_{k=0}^{\infty} k x_k(t) = -\frac{\partial}{\partial \xi} \log \sum_{k=0}^{\infty} x_k(0) e^{-\xi k} \Bigg|_{\xi = st} + \frac{\lambda}{s}(1 - e^{-st}).$$

Notice the exponential decay towards the equilibrium vaue of λ/s. The time

$$\tau = \frac{\log(\lambda/s)}{s}$$

corresponds exactly to the end of Haigh's phase one.

Approximations

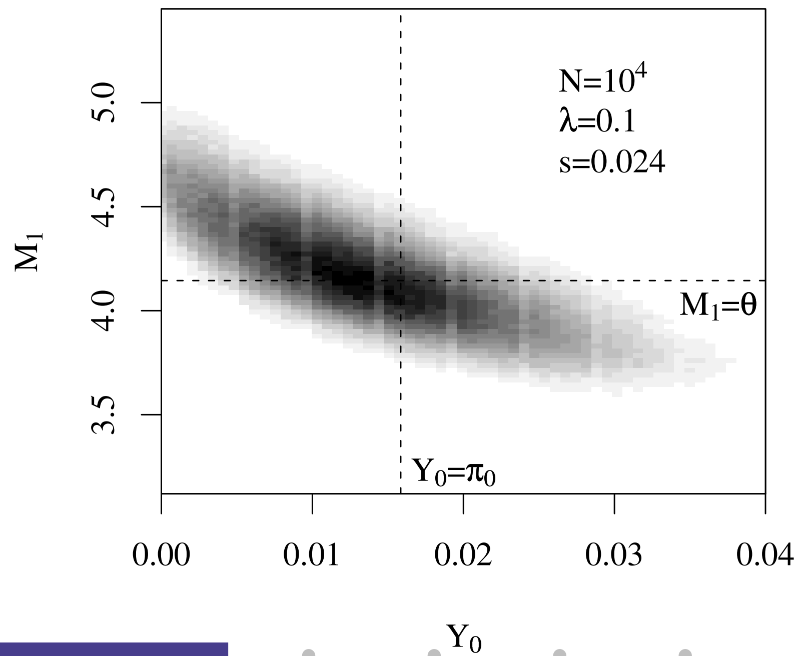
$$dY_0 = s(M_1(\mathbf{Y}) - \lambda)Y_0(t)dt + \sqrt{\frac{1}{N}Y_0(1-Y_0)}dW_0.$$

Cannot solve for $M_1(\mathbf{Y})$. Instead seek a *good approximation* of M_1 given Y_0.

Approximations

$$dY_0 = s(M_1(\mathbf{Y}) - \lambda)Y_0(t)dt + \sqrt{\frac{1}{N}Y_0(1-Y_0)}dW_0.$$

Cannot solve for $M_1(\mathbf{Y})$. Instead seek a *good approximation* of M_1 given Y_0. Simulations suggest a good fit to a linear relationship between Y_0 and M_1.



Extending Haigh's approach

Haigh assumes at click times π_0 distributed evenly over other classes. Suppose now that this holds in between click times too: given Y_0 approximate state of system by *the PPA* (Poisson Profile Approximation)

$$\Pi(Y_0) = \left(Y_0, \frac{1 - Y_0}{1 - \pi_0}(\pi_1, \pi_2, \ldots) \right).$$

Extending Haigh's approach

Haigh assumes at click times π_0 distributed evenly over other classes. Suppose now that this holds in between click times too: given Y_0 approximate state of system by *the PPA* (Poisson Profile Approximation)

$$\Pi(Y_0) = \left(Y_0, \frac{1 - Y_0}{1 - \pi_0}(\pi_1, \pi_2, \ldots) \right).$$

Estimate M_1 not from PPA but from *relaxed* PPA obtained by evolving PPA according to the deterministic dynamical system for time $A\tau := A \log(\lambda/s)/s$. This gives

$$M_1 = \theta + \frac{\eta}{e^\eta - 1}\left(1 - \frac{Y_0}{\pi_0}\right), \quad \eta = (\lambda/s)^{1-A}.$$

Three one dimensional diffusions

Substituting in the one-dimensional diffusion approximation for Y_0 gives:

$$A \text{ small}, \qquad dY_0 = \lambda(\pi_0 - Y_0)Y_0 dt + \sqrt{\frac{1}{N} Y_0} dW,$$

$$A = 1, \qquad dY_0 = 0.58 s \left(1 - \frac{Y_0}{\pi_0}\right) Y_0 dt + \sqrt{\frac{1}{N} Y_0}\, dW,$$

$$A \text{ large}, \qquad dY_0 = s \left(1 - \frac{Y_0}{\pi_0}\right) Y_0 dt + \sqrt{\frac{1}{N} Y_0}\, dW,$$

A rescaling

$$Z(t) = \frac{1}{\pi_0} Y_0(N\pi_0 t).$$

Set

$$\gamma = \frac{N\lambda}{Ns\log(N\lambda)}.$$

A rescaling

$$Z(t) = \frac{1}{\pi_0} Y_0(N\pi_0 t).$$

Set

$$\gamma = \frac{N\lambda}{Ns\log(N\lambda)}.$$

A small, $\quad dZ = (N\lambda)^{1-2\gamma}(1-Z)Zdt + \sqrt{Z}dW,$

$A = 1,$ $\quad dZ = 0.58\frac{1}{\gamma\log(N\lambda)}(N\lambda)^{1-\gamma}(1-Z)Zdt + \sqrt{Z}\,dW,$

A large, $\quad dZ = \frac{1}{\gamma\log(N\lambda)}(N\lambda)^{1-\gamma}(1-Z)Zdt + \sqrt{Z}\,dW.$

Implications

A small $\equiv$ fast clicking:

$$dZ = (N\lambda)^{1-2\gamma}(1-Z)Z dt + \sqrt{Z} dW.$$

The ratchet never clicks for $\gamma < 1/2$*.*

Implications

A small $\equiv$ fast clicking:

$$dZ = (N\lambda)^{1-2\gamma}(1-Z)Z\,dt + \sqrt{Z}\,dW.$$

The ratchet never clicks for $\gamma < 1/2$.

$A = 1$ (A large) $\equiv$ moderate clicking (slow clicking):

$$dZ = 0.58(1)\frac{1}{\gamma\log(N\lambda)}(N\lambda)^{1-\gamma}(1-Z)Z\,dt + \sqrt{Z}\,dW.$$

Implications

A small $\equiv$ fast clicking:

$$dZ = (N\lambda)^{1-2\gamma}(1-Z)Z dt + \sqrt{Z} dW.$$

The ratchet never clicks for $\gamma < 1/2$.

$A = 1$ (A large) $\equiv$ moderate clicking (slow clicking):

$$dZ = 0.58(1)\frac{1}{\gamma \log(N\lambda)}(N\lambda)^{1-\gamma}(1-Z)Z dt + \sqrt{Z}\, dW.$$

In order for $0.58\frac{1}{\gamma \log(N\lambda)}(N\lambda)^{1-\gamma}$ to be > 5,

γ	0.3	0.4	0.5	0.55	0.6	0.7	0.8	0.9
$N\lambda \geq$	20	10^2	$9 \cdot 10^2$	$4 \cdot 10^3$	$2 \cdot 10^4$	$4 \cdot 10^6$	$2 \cdot 10^{11}$	$8 \cdot 10^{26}$

Rule of thumb

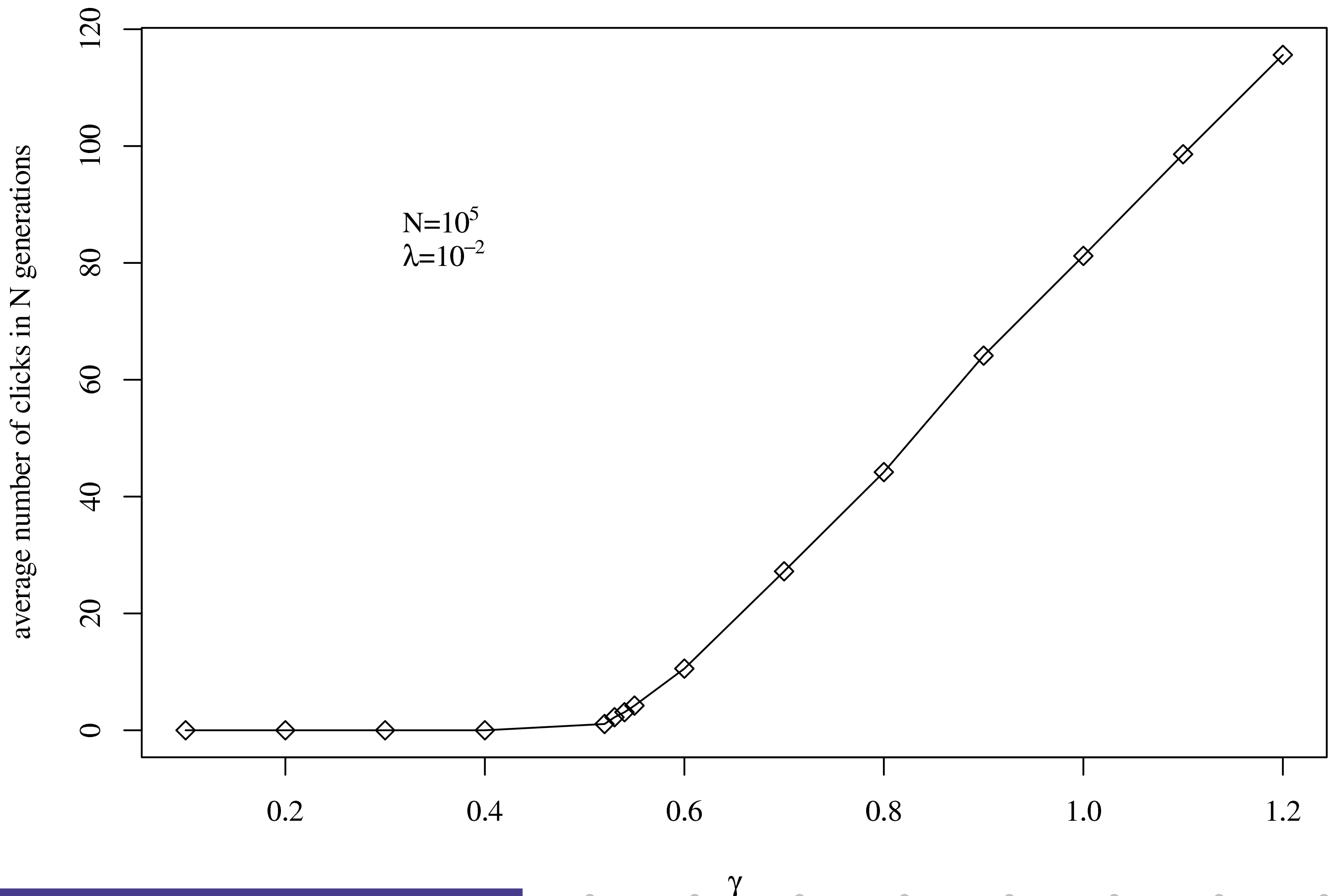
For biologically realistic parameters, transition from no clicks to moderate clicks (on evolutionary timescale) around $\gamma = 0.5$.

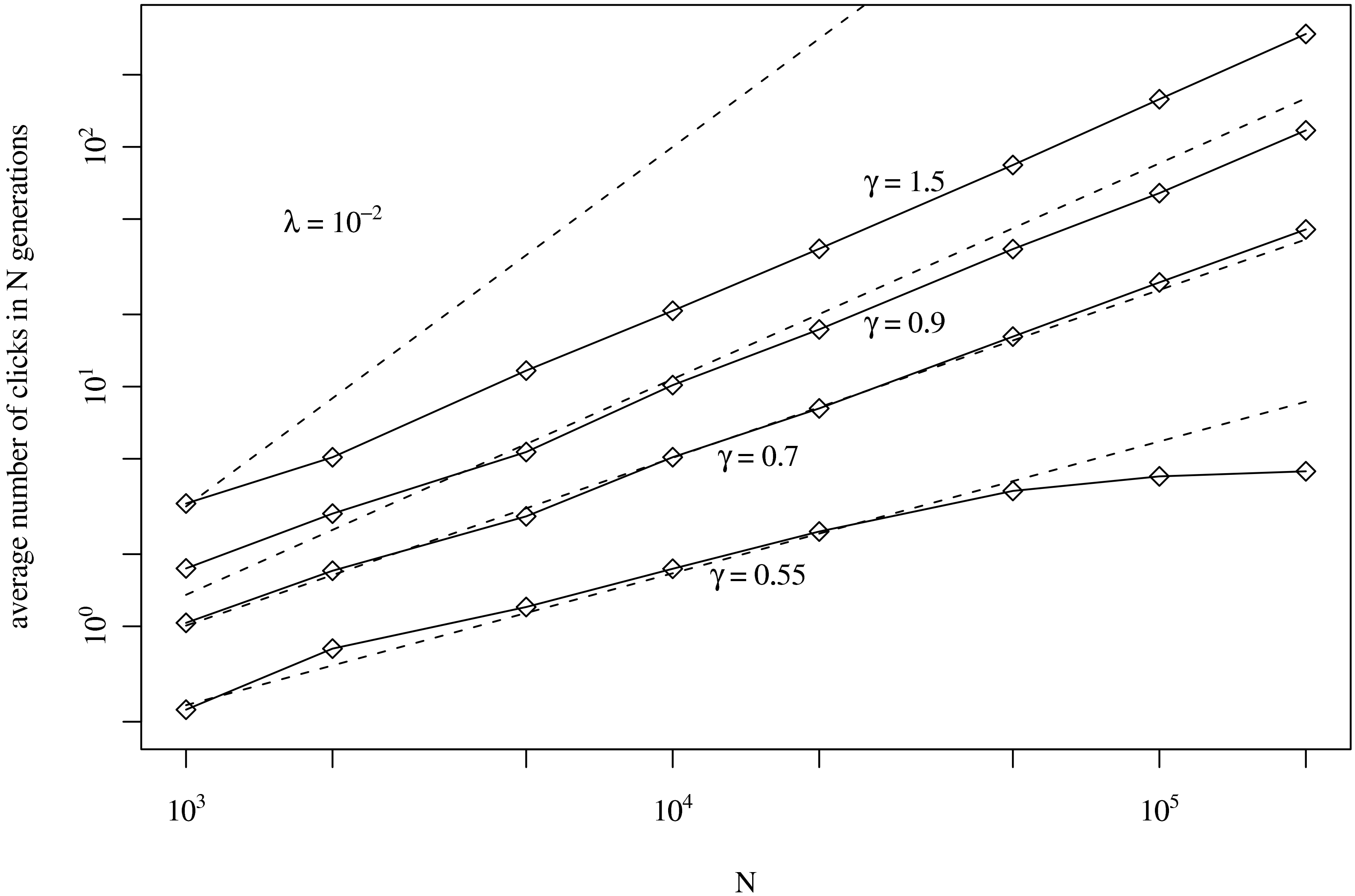
Rule of thumb

For biologically realistic parameters, transition from no clicks to moderate clicks (on evolutionary timescale) around $\gamma = 0.5$.

The rate of the ratchet is of the order $N^{\gamma-1}\lambda^{\gamma}$ *for* $\gamma \in (1/2, 1)$*, whereas it is exponentially slow in* $(N\lambda)^{1-\gamma}$ *for* $\gamma < 1/2$*.*

Simulations




average number of clicks in N generations
λ = 10^−2
γ = 1.5
γ = 0.9
γ = 0.7
γ = 0.55
10^0
10^1
10^2
10^3
10^4
10^5
N

Purely beneficial mutations

We saw before that the probability of an isolated selected sweep $\approx 2\sigma$.

Purely beneficial mutations

We saw before that the probability of an isolated selected sweep $\approx 2\sigma$.

What about overlapping sweeps?

Purely beneficial mutations

We saw before that the probability of an isolated selected sweep $\approx 2\sigma$.

What about overlapping sweeps?

Interference reduces the chance of fixation as beneficial mutations compete with one another.

Purely beneficial mutations

We saw before that the probability of an isolated selected sweep $\approx 2\sigma$.

What about overlapping sweeps?

Interference reduces the chance of fixation as beneficial mutations compete with one another.

Is there a limit to the rate of adaptation?

A mixture of mutations

All mutations equal 'strength' so ith individual's fitness characterized by *net* number, X_i, of beneficial mutations.

A mixture of mutations

All mutations equal ‘strength’ so ith individual’s fitness characterized by *net* number, X_i, of beneficial mutations.

- *Mutation:* For each individual i a mutation event occurs at rate μ. With probability $1-q$, X_i changes to $X_i - 1$ and with probability q, X_i changes to $X_i + 1$.
- *Selection:* For each pair of individuals (i, j), at rate $\frac{\sigma}{N}(X_i - X_j)^+$, individual i replaces individual j.
- *Resampling:* For each pair of individuals (i, j), at rate $\frac{1}{N}$, individual i replaces individual j.

A mixture of mutations

All mutations equal ‘strength’ so ith individual’s fitness characterized by *net* number, X_i, of beneficial mutations.

- *Mutation:* For each individual i a mutation event occurs at rate μ. With probability $1-q$, X_i changes to $X_i - 1$ and with probability q, X_i changes to $X_i + 1$.
- *Selection:* For each pair of individuals (i, j), at rate $\frac{\sigma}{N}(X_i - X_j)^+$, individual i replaces individual j.
- *Resampling:* For each pair of individuals (i, j), at rate $\frac{1}{N}$, individual i replaces individual j.

Technical modification: suppress mutations which would make the ‘width’ of the population $> L \equiv N^{1/4}$.

- $\ldots, P_0(t), \ldots, P_k(t), \ldots$ - Proportion of individuals with k mutations at time t
- k_{max} (k_{min}) - type of the fittest (least fit) individual

- $\ldots, P_0(t), \ldots, P_k(t), \ldots$ - Proportion of individuals with k mutations at time t
- k_{max} (k_{min}) - type of the fittest (least fit) individual

$$\begin{aligned} dP_k &= \bar{\mu}_k(P)\,dt + \sigma \sum_{l \in \mathbb{Z}} (k-l) P_k P_l\,dt + dM_k^P \\ &= [\bar{\mu}_k(P) + \sigma(k - m(P))P_k]\,dt + dM_k^P \end{aligned}$$

$$\bar{\mu}_k(P) \approx \mu\,(qP_{k-1} - P_k + (1-q)P_{k+1})$$

- $\ldots, P_0(t), \ldots, P_k(t), \ldots$ - Proportion of individuals with k mutations at time t
- k_{max} (k_{min}) - type of the fittest (least fit) individual

$$\begin{aligned} dP_k &= \bar{\mu}_k(P)\,dt + \sigma \sum_{l \in \mathbb{Z}} (k-l) P_k P_l\,dt + dM_k^P \\ &= \left[\bar{\mu}_k(P) + \sigma(k - m(P))P_k\right] dt + dM_k^P \end{aligned}$$

$$\bar{\mu}_k(P) \approx \mu\left(qP_{k-1} - P_k + (1-q)P_{k+1}\right)$$

M^P is a martingale with

$$\left[M_k^P\right](t) \leq \frac{2\mu}{N} t + \frac{1}{N} \int_0^t \sum_{l \in \mathbb{Z}} (2 + \sigma(k-l)^+ + \sigma(l-k)^+) P_k(s) P_l(s)\,ds$$

Moments

Mean fitness $m(P) = \sum_k kP_k$ satisfies

$$\begin{aligned} dm(P) &= (\bar{\mu}(P) + \sigma c_2(P))\, dt + dM^{P,m} \\ \bar{\mu}(P) &\approx \mu(2q - 1) \end{aligned}$$

Moments

Mean fitness $m(P) = \sum_k kP_k$ satisfies

$$\begin{aligned} dm(P) &= (\bar{\mu}(P) + \sigma c_2(P))\, dt + dM^{P,m} \\ \bar{\mu}(P) &\approx \mu(2q-1) \end{aligned}$$

Ignoring mutation terms, centred moments $c_k = \sum_k (k - m(p))^n P_k$ satisfy

$$\begin{aligned} dc_2 &\approx \sigma c_3 && dt + small\ noise\ terms \\ dc_3 &\approx \sigma(c_4 - 3c_2c_2) && dt + small\ noise\ terms \\ dc_4 &\approx \sigma(c_5 - 4c_3c_2) && dt + small\ noise\ terms \\ dc_5 &\approx \sigma(c_6 - 5c_4c_2) && dt + small\ noise\ terms \end{aligned}$$

Heuristics

Centred process has a stationary distribution.

Heuristics

Centred process has a stationary distribution.

Approximate by a deterministic distribution.

Heuristics

Centred process has a stationary distribution.

Approximate by a deterministic distribution.

Equations for centred moments give:

$$
\begin{aligned}
0 &\approx \sigma c_3 \, dt \\
0 &\approx \sigma(c_4 - 3c_2c_2) \\
0 &\approx \sigma(c_5 - 4c_3c_2) \\
0 &\approx \sigma(c_6 - 5c_4c_2) \ldots
\end{aligned}
$$

Heuristics

Centred process has a stationary distribution.

Approximate by a deterministic distribution.

Equations for centred moments give:

$$\begin{aligned} 0 &\approx \sigma c_3 \, dt \\ 0 &\approx \sigma(c_4 - 3c_2c_2) \\ 0 &\approx \sigma(c_5 - 4c_3c_2) \\ 0 &\approx \sigma(c_6 - 5c_4c_2) \ldots \end{aligned}$$

Stationary distribution approximately Gaussian.

- Suppose stationary distribution Gaussian with mean $m(P)$, variance b^2.

- Suppose stationary distribution Gaussian with mean $m(P)$, variance b^2.

Front is at K where

$$\frac{1}{\sqrt{2\pi}b}e^{-K^2/2b^2} = \frac{1}{N} \Rightarrow K \approx b\sqrt{2\log N}$$

- Suppose stationary distribution Gaussian with mean $m(P)$, variance b^2.

Front is at K where

$$\frac{1}{\sqrt{2\pi}b}e^{-K^2/2b^2} = \frac{1}{N} \Rightarrow K \approx b\sqrt{2\log N}$$

If there is a single individual at $m(P)+K$ at time $t=0$, how long until there is an individual at $m(P)+K+1$?

- Suppose stationary distribution Gaussian with mean $m(P)$, variance b^2.

Front is at K where

$$\frac{1}{\sqrt{2\pi}b} e^{-K^2/2b^2} = \frac{1}{N} \Rightarrow K \approx b\sqrt{2\log N}$$

If there is a single individual at $m(P)+K$ at time $t=0$, how long until there is an individual at $m(P)+K+1$?

Ignoring beneficial mutations occurring to individuals at $m(P)+K-1$, until the front advances

$$Z(t) \approx e^{(\sigma K-(1-q)\mu)t}.$$

More heuristics

Front advances as soon as individual of $Z(t)$ accumulates a beneficial mutation.

More heuristics

Front advances as soon as individual of $Z(t)$ accumulates a beneficial mutation. The probability that the front doesn't advance by time t is

$$\exp\left\{-q\mu\int_0^t Z(s)\,ds\right\} \approx \exp\left\{-\frac{q\mu}{\sigma K-(1-q)\mu}(e^{(\sigma K-(1-q)\mu)t}-1)\right\}$$

More heuristics

Front advances as soon as individual of $Z(t)$ accumulates a beneficial mutation. The probability that the front doesn't advance by time t is

$$\exp\left\{-q\mu\int_0^t Z(s)\,ds\right\} \approx \exp\left\{-\frac{q\mu}{\sigma K-(1-q)\mu}\left(e^{(\sigma K-(1-q)\mu)t}-1\right)\right\}$$

Average time for front to advance by one is

$$\approx \frac{1}{\sigma K-(1-q)\mu}\log(\sigma K-(1-q)\mu).$$

More heuristics

Front advances as soon as individual of $Z(t)$ accumulates a beneficial mutation. The probability that the front doesn't advance by time t is

$$\exp\left\{-q\mu \int_0^t Z(s)\, ds\right\} \approx \exp\left\{-\frac{q\mu}{\sigma K - (1-q)\mu}(e^{(\sigma K-(1-q)\mu)t} - 1)\right\}$$

Average time for front to advance by one is

$$\approx \frac{1}{\sigma K - (1-q)\mu} \log(\sigma K - (1-q)\mu).$$

But wave speed is

$$\approx \mu(2q-1) + \sigma c_2(P) \approx \mu(2q-1) + \sigma b^2 \approx \mu(2q-1) + \frac{\sigma K^2}{2\log N}.$$

Conclusion

Consistency condition:

$$\frac{\sigma K - (1-q)\mu}{\log(\sigma K - (1-q)\mu)} = \mu(2q-1) + \frac{\sigma K^2}{2\log N}.$$

Conclusion

Consistency condition:

$$\frac{\sigma K - (1-q)\mu}{\log(\sigma K - (1-q)\mu)} = \mu(2q-1) + \frac{\sigma K^2}{2\log N}.$$

For large K, this approximately reduces to

$$K \log(\sigma K) = 2\log N.$$

Conclusion

Consistency condition:

$$\frac{\sigma K - (1-q)\mu}{\log(\sigma K - (1-q)\mu)} = \mu(2q-1) + \frac{\sigma K^2}{2\log N}.$$

For large K, this approximately reduces to

$$K \log(\sigma K) = 2 \log N.$$

- If $K = \log N$, then $LHS > RHS$;
- $K = \log^{1-\delta} N$, $\delta > 0$, then $LHS < RHS$.

Conclusion

Consistency condition:

$$\frac{\sigma K - (1-q)\mu}{\log(\sigma K - (1-q)\mu)} = \mu(2q-1) + \frac{\sigma K^2}{2\log N}.$$

For large K, this approximately reduces to

$$K\log(\sigma K) = 2\log N.$$

- If $K = \log N$, then $LHS > RHS$;
- $K = \log^{1-\delta} N$, $\delta > 0$, then $LHS < RHS$.

So K between $\log N$ and any fractional power of $\log N \Rightarrow$ rate of adaptation, of order between $\log N$ and any fractional power of $\log N$.

Rigorous result

Theorem.
If $q > 0$, then for any $\beta > 0$, there exists a positive constant $c_{\mu,\sigma,q}$ such that

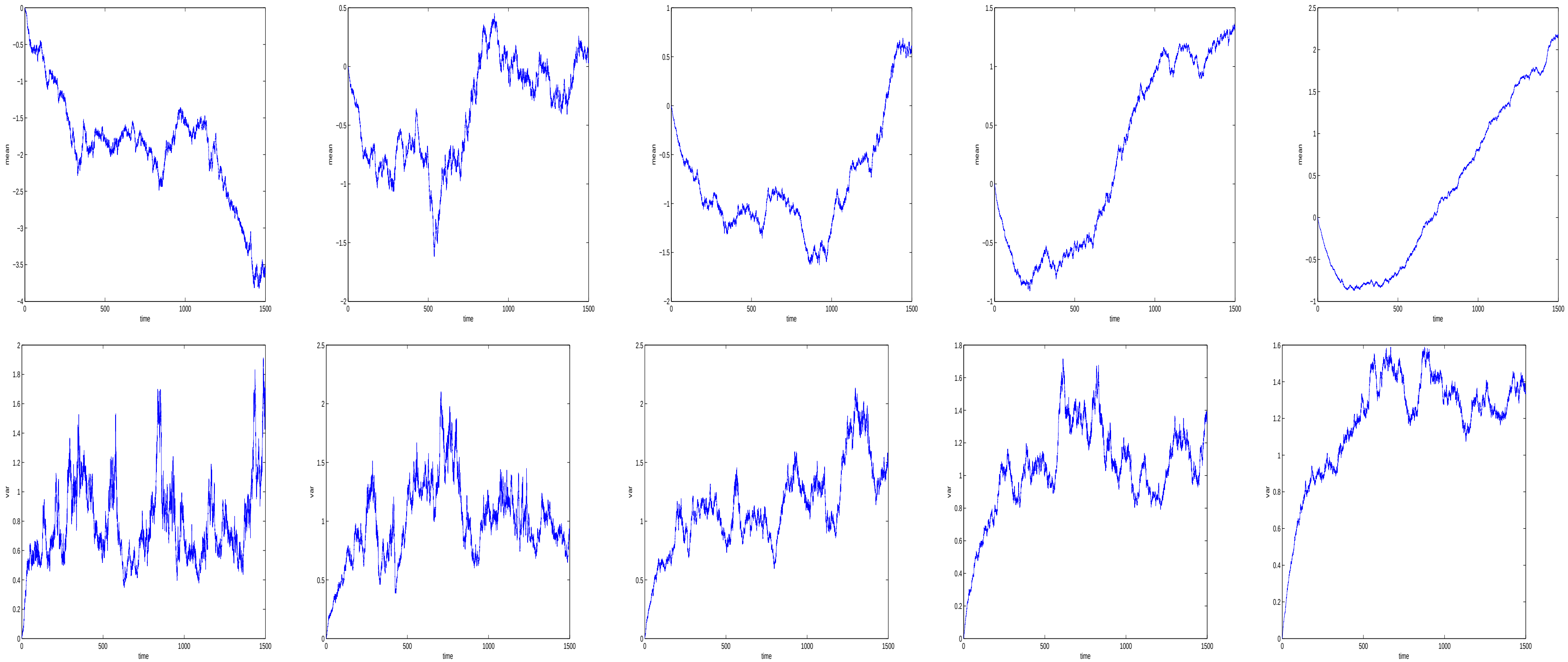
$$\mathbb{E}^{\pi}[m(1)] \geq \mathbb{E}^{\pi}[c_2] \geq c_{\mu,\sigma,q} \log^{1-\beta} N$$

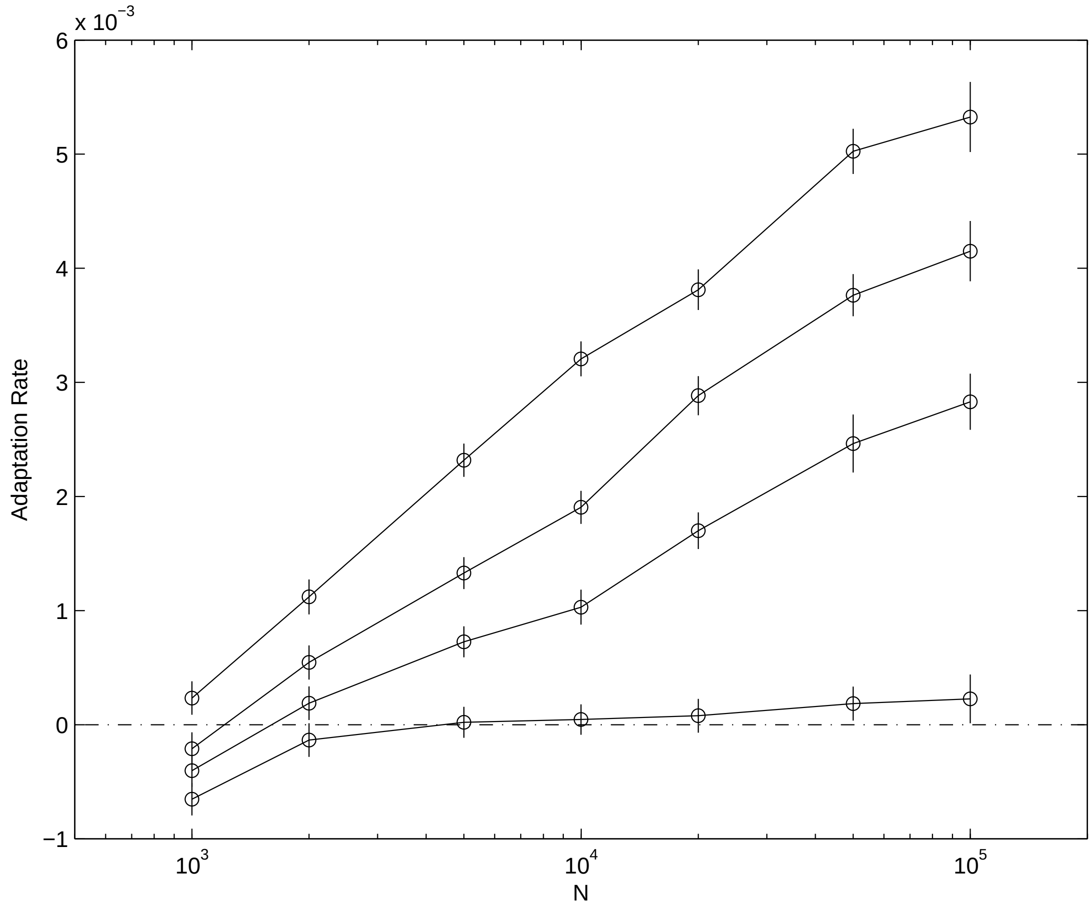
if N is sufficiently large.

Simulations

With $\mu = 0.01$, $q = 0.01$, $\sigma = 0.01$, $N = 1000, 2000, 5000, 10000, 30000$.

First row: mean; second row: variance.





Adaptation rate against population size.

From top to bottom, $q = 4\%, 2\%, 1\%, 0.2\%$, $\mu = 0.01$, $\sigma = 0.01$.