HAL open science

A stochastic model for a single click of Muller's ratchet

D. Waxman, L. Loewe

▶ **To cite this version:**

D. Waxman, L. Loewe. A stochastic model for a single click of Muller's ratchet. Journal of Theoretical Biology, 2010, 264 (4), pp.1120. 10.1016/j.jtbi.2010.03.014 . hal-00594146

HAL Id: hal-00594146

https://hal.science/hal-00594146v1

Submitted on 19 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Author's Accepted Manuscript

A stochastic model for a single click of Muller's ratchet

D. Waxman, L. Loewe

PII: S0022-5193(10)00137-2
DOI: doi:10.1016/j.jtbi.2010.03.014
Reference: YJTBI 5915

To appear in: *Journal of Theoretical Biology*

www.elsevier.com/locate/yjtbi

Received date: 11 July 2009
Revised date: 9 March 2010
Accepted date: 9 March 2010

Cite this article as: D. Waxman and L. Loewe, A stochastic model for a single click of Muller's ratchet, *Journal of Theoretical Biology*, doi:10.1016/j.jtbi.2010.03.014

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

9/3/2010

A Stochastic Model for a Single Click of Muller's Ratchet

D. Waxman[1] and L. Loewe[2]

[1]Centre for the Study of Evolution, School of Life Sciences,

University of Sussex, Brighton BN1 9QG, Sussex UK

[2]Centre for Systems Biology at Edinburgh, School of Biological Sciences,

University of Edinburgh, Edinburgh EH9 3JD, Scotland UK

Accepted manuscript

Running Head: A single click of Muller's ratchet.

[1]Corresponding Author

E-mail address: D.Waxman@sussex.ac.uk

E-mail address: Laurence.Loewe@evolutionary-research.net

ACCEPTED MANUSCRIPT

Abstract

This work presents a new approach to Muller's ratchet, where Haigh's model is approximately mapped into a simpler model that describes the behaviour of a population after a click of the ratchet, i.e., after loss of what was the fittest class. This new model predicts the distribution of times to the next click of the ratchet and is equivalent to a Wright-Fisher model for a population of haploid asexual individuals with one locus and two alleles. Within this model, the fittest members of a population correspond to carriers of one allele, while all other individuals have suboptimal fitness and are represented as carriers of the other allele. In this way, all suboptimal fitness individuals are amalgamated into a single 'mutant' class.

The approach presented here has some limitations and the potential for improvement. However, it does lead to results for the rate of the ratchet that, over a wide range of parameters, are accurate within one order of magnitude of simulation results. This contrasts with existing approaches, which are designed for only one or the other of the two different parameter regimes known for the ratchet and are more accurate only in the parameter regime they were designed for.

Numerical results are presented for the mean time between clicks of the ratchet for (i) the Wright-Fisher model, (ii) a diffusion approximation of

Accepted manuscript

ACCEPTED MANUSCRIPT

this model and (iii) individually based simulations of a full model. The diffusion approximation is validated over a wide range of parameters by its close agreement with the Wright-Fisher model.

The present work predicts that: (a) the time between clicks of the ratchet is insensitive to the value of the selection coefficient when the genomic mutation rate is large compared with the selection coefficient against a deleterious mutation, (b) the time interval between clicks of the ratchet has, approximately, an exponential distribution (or its discrete analogue). It is thus possible to determine the variance in times between clicks, given the expected time between clicks. Evidence for both (a) and (b) is seen in simulations.

Accepted manuscript

Key words: Muller's ratchet; random genetic drift; asexual population; theoretical population genetics.

ACCEPTED MANUSCRIPT

1 Introduction

Muller's ratchet is a random process that occurs in a finite population of asexual individuals. It corresponds to the repeated irreversible loss of the fittest class of individuals and results in degradation of the fitness of a population over time. It can arise from the combined effects of recurrent deleterious mutations, the effective absence of beneficial (back) mutations, the absence of recombination, and random genetic drift (Muller 1964; Felsenstein 1974).

In a finite population there is inherent randomness in the contribution of different individuals to the next generation. When the set of individuals with the smallest number of mutations (the fittest class) makes no contribution to the next generation, it is lost from the population and one click of Muller's ratchet is said to have taken place. The next fittest class then becomes the fittest class. In the effective absence of both beneficial mutations and recombination, each click of Muller's rachet is irreversible. Clicks of the ratchet randomly occur, and lead to the mean fitness of the population decreasing over time. This process naturally leads to the expectation that asexual species persist for shorter times than broadly comparable recombining sexual species, in which parents of suboptimal fitness may produce offspring of higher fitness than either parent (Felsenstein 1974; Maynard Smith 1978).

Muller's ratchet is believed to be relevant to a variety of replicators

Accepted manuscript

ACCEPTED MANUSCRIPT

that include: asexually reproducing organisms (Loewe and Lamatsch 2008); purely selfing organisms (Heller and Maynard Smith 1978; Loewe and Cutter 2008); non recombining regions of the genome such as the Y chromosome (Rice 1994; Gordo and Charlesworth 2000a, b; Bachtrog 2008); endosymbionts (Moran 1996; Rispe and Moran 2000), and mitochondria (Bergstrom and Pritchard 1998; Loewe 2006).

In order to assess the relevance of Muller's ratchet, to any given biological system, it is necessary to determine the rate it predicts for the accumulation of deleterious mutations. There have been many approaches that attempt to do this (see e.g. Söderberg and Berg 2007, Rouzine et. al. 2008, Etheridge et. al. 2008 and references in Loewe 2006). Models that can predict the operation of Muller's ratchet with reasonable accuracy, over a range of parameter values, are typically associated with one of two parameter regimes associated with different behaviours of the ratchet (Loewe and Hill 2010). One parameter regime covers the case where, at deterministic mutation selection balance (see Haigh 1978), the expected number of individuals in the fittest class is larger than unity, Muller's ratchet operates in a relatively *slow regime* that is characterised by a large mean time between clicks and also an appreciable level of variation in the time between clicks. The accumulation of mutations in a population occurs when the relatively improbable event that all individuals in the fittest class fail to contribute to the next generation.

Accepted manuscript

ACCEPTED MANUSCRIPT

Under the slow regime, variants of Stephan et al's diffusion approximation of Haigh's stochastic treatment (of Haigh's model) produce the most accurate predictions for the ratchet's rate (Stephan et. al. 1993; Stephan and Kim 2002; Gordo and Charlesworth 2000a, b; Etheridge et. al. 2008). The other parameter regime occurs when the expected number of individuals in the fittest class falls below unity. Muller's ratchet then operates in a relatively *fast regime* that is characterised by a quasi-continuous accumulation of mutations – which has been described as quasi-deterministic. Here clicks are expected to happen with near deterministic necessity before the system can equilibrate. In this regime the speed of the ratchet can be reasonably described by the models which deal with the lack of an approximate equilibrium between mutation and selection. This includes the work of Gessler (1995), who employed simulation and analytical approaches to the problem and estimated how the distribution of mutations within a population differs from a Poisson distribution, and the more recent work of Rouzine et. al. (2008), who modelled the ratchet as a semi-deterministic combination of deterministic and stochastic processes associated with a travelling wave. No approximation, however, has yet been produced that accurately spans these two distinct regimes of operation.

In the present work we present an approach to Muller's ratchet that differs from previous approaches. A single click of the ratchet is treated as

Accepted manuscript

ACCEPTED MANUSCRIPT

being describable by a simple Wright-Fisher model of random genetic drift (Fisher 1922; Wright 1931). This entails truncating Haigh's model of Muller's ratchet (Haigh 1978) and mapping it onto a Wright-Fisher model where just two mutational classes are used to represent all different types of individuals in a population. One of the mutational classes consists of the fittest individuals in a population; the other class is an amalgamation of all suboptimal fitness individuals. The resulting Wright-Fisher model of the ratchet takes the same form in all parameter regimes. It predicts the rate of the ratchet and allows many statistical properties to be numerically calculated, for a wide range of parameter values, without the need for simulation and hence without the associated statistical errors. This approach, however, does have some limitations and does have the potential for improvement – as we point out below. We view the approach presented here as a possible starting point for a uniformly valid model for the rate of Muller's ratchet.

The structure of this paper is as follows. We first introduce a somewhat general model for Muller's ratchet for an asexual haploid population. Next, a truncated version of this model is considered and parameters in the truncated model are explicitly related to parameters in Haigh's (1978) model of Muller's ratchet. In the remaining part of the paper the stochastic dynamics of the truncated model are analysed and properties of the time between clicks of Muller's ratchet are obtained in three different ways: (i) from the

Accepted manuscript

ACCEPTED MANUSCRIPT

numerical solution of properties of a potentially large matrix associated with the Wright-Fisher model, (ii) from a diffusion approximation of the Wright-Fisher model, and (iii) from individual-based simulations of a full model of Muller's ratchet. These three approaches are compared with each another and with the previous approaches of Stephan and Kim (2002), Gessler (1995) and Rouzine et al (2008).

The model presented in this work provides an estimate of the probability distribution for time intervals between clicks of the ratchet, and hence allows the determination of statistics of time intervals. Furthermore, under a wide range of parameter combinations, the model predicts the rate of accumulation of deleterious mutations from Muller's ratchet that lie within one order of magnitude of simulation results. These results include parameter combinations from the fast and slow regimes of operation of the ratchet described above. This work thus raises the possibility that a single theoretical approach may exist for predicting the rate of Muller's ratchet, that accurately covers both parameter regimes.

Accepted manuscript

2 General Model

Consider a population of haploid asexual individuals with discrete generations, labelled $t = 0, 1, 2, \ldots$. The census point in a generation is taken to

ACCEPTED MANUSCRIPT

be the *adult* stage, after all selection has occurred and immediately prior to reproduction. Many other treatments, including Haigh's treatment (Haigh 1978), implicitly census immediately after birth of offspring, before viability selection has occurred (we discuss this further in Section 4 and in Appendix A).

The lifecycle of the population is

Adults

$\downarrow$ asexual reproduction with mutation,

$\downarrow$ followed by the death of all adults

Juveniles

$\downarrow$ viability selection and

$\downarrow$ number regulation

Adults.

New mutations can occur at reproduction and we assume an individual can carry a maximum number of Q mutations. All mutations are assumed to deleteriously affect viability but have no effect on fertility, hence the expected number of offspring of an adult is independent of their mutational load.

In a finite population, where N adults are maintained each generation, the regulation of their number is assumed to occur by a non selective ecological

Accepted manuscript

ACCEPTED MANUSCRIPT

thinning process that occurs after viability selection has acted. It effectively corresponds to randomly picking N individuals from the population, without replacement, to constitute the adults of the next generation. The chance an individuals is picked is independent of the number of mutations they carry.

The viability of a newly born individual is taken to be determined solely by the number of the mutations they carry (Haigh 1978). This feature allows us to collect the population into mutational classes. We use $X_k(t)$ to represent the proportion or relative frequency (henceforth abbreviated to *frequency*) of all *adults* carrying k mutations in generation t, where $k = 0, 1, 2, ..., Q$. Since the $X_k(t)$ are frequencies they sum to unity at all times: $\sum_{k=0}^{Q} X_k(t) = 1$.

For a very large (i.e., effectively infinite) population, the changes of $X_k(t)$ over time are essentially deterministic in character. In such a case we can write

$$X_k(t+1) = B_k(\mathbf{X}(t)), \qquad k = 0, 1, 2, ..., Q \tag{1}$$

where $\mathbf{X}(t)$ denotes the vector $\mathbf{X}(t) = (X_0(t), X_1(t), X_2(t), ...X_Q(t))$ and the function $B_k(\mathbf{X})$ incorporates the frequency changes of k-mutant adults that are caused by selection and mutation.

To derive the precise form of Eq. (1), let v_j denote the relative viability of an individual with j mutations and let $M_{k,j}^{(Q)}$ be the probability that an offspring has k mutations, given its parent had j mutations. Then

Accepted manuscript

ACCEPTED MANUSCRIPT

the frequency of k-mutant offspring produced by adults of generation t is $\sum_{j=0}^{Q} M_{k,j}^{(Q)} X_j(t)$. After selection, the frequency of these individuals is proportional to $v_k \sum_{j=0}^{Q} M_{k,j}^{(Q)} X_j(t)$. It follows that $B_k(\mathbf{X})$ is given by

$$B_k(\mathbf{X}) = \frac{v_k \sum_{j=0}^{Q} M_{k,j}^{(Q)} X_j}{\sum_{i,j=0}^{Q} v_i M_{k,j}^{(Q)} X_j}, \qquad k = 0, 1, 2, ..., Q. \tag{2}$$

We note that while Eq. (1) describes the change of frequencies over time of different mutational classes, it has the same mathematical form as a *different model*, namely a *one locus haploid asexual population* with Q different alleles. In this one locus model, alleles are denoted as A_k, i.e., labelled by k where $k = 0, 1, ..., Q$. The quantity $X_k(t)$ then has the interpretation as the frequency of the k'th allele in generation t, while the v_k characterise the viability of different alleles and the $M_{k,j}^{(Q)}$ contain their probabilities of mutation.

3 Truncated Model: $Q = 1$

The analysis presented in this work is based upon a truncated model which is a simplification of the somewhat general model presented in the previous section. We shall proceed by introducing the truncated model and first describe the deterministic behaviour it exhibits when population size is effectively infinite. We then investigate the stochastic behaviour the model exhibits when

Accepted manuscript

ACCEPTED MANUSCRIPT

population size is finite.

The truncated model follows from the model of Section 2 by restricting the number of mutational classes to two. We achieve this by taking $Q = 1$ in Eqs. (1) and (2). Since there are *two* mutational classes, the truncated model is mathematically equivalent to an effective model for asexual haploid individuals with one locus and two alleles, namely A_0 and A_1.

The two mutational classes within the truncated model are as follows.

(i) The set of all adults that are mutation free constitute one mutational class. More generally, we take this mutational class to consist of all individuals in a population with the smallest number of mutations and hence the highest fitness and shall refer to it as the fittest class. We represent members of this class, within the effective one locus model, as carriers of the A_0 allele.

(ii) The set of all adults carrying any non-zero number of mutations make up the other mutational class. More generally, we take this mutation class to consist of all suboptimal fitness individuals in the population, i.e., all individuals carrying more mutations than members of the fittest class and shall refer to it as the suboptimal class. We represent its members, within the effective one locus model, as carriers of the A_1 allele. In this way all individuals, who carry more mutations than members of the fittest class, are amalgamated, into a single class.

The frequency of the fittest class in generation t is $X_0(t)$ while that of

Accepted manuscript

ACCEPTED MANUSCRIPT

the suboptimal class is $1 - X_0(t)$. The absence of back mutations means that in the truncated model, mutation is only in one direction: from the fittest class to the suboptimal class. Let μ be the probability that an offspring of a mutation-free (i.e., fittest) adult has a non-zero number of mutations. Mutation is then characterised in the truncated model by the mutation rates $M_{k,j}^{(Q)}$ of Eq. (2) being replaced by a truncated set of mutation rates, $M_{k,j}^{(1)}$, which have the form $M_{0,0}^{(1)} = 1 - \mu$, $M_{0,1}^{(1)} = 0$, $M_{1,0}^{(1)} = \mu$ and $M_{1,1}^{(1)} = 1$. The relative viability of individuals in the fittest class is $v_0 = 1$ while individuals in the suboptimal class have a relative viability of $v_1 = 1 - \sigma$. The parameter σ can be thought of as an effective selection coefficient associated with deleterious mutations in the truncated model; σ lies in the range 0 to 1. We shall make the fittest class the focus of our attention and hence shall adopt a more intuitive notation by writing $X_{\text{fittest}}(t)$ for $X_0(t)$:

$$X_{\text{fittest}}(t) = X_0(t) = \begin{cases} \text{frequency of the fittest} \\ \text{adults in generation } t. \end{cases} \tag{3}$$

In the next Section we relate the parameters in the truncated model to the parameters in Haigh's model of the ratchet (Haigh 1978).

Using the forms given above for $M_{k,j}^{(1)}$ and v_j in the truncated model, the terms of $B_k(\mathbf{X})$ (Eq. (2)), with $Q = 1$, are $B_0(\mathbf{X}) \equiv B_0(X_0, X_1)$ and

Accepted manuscript

ACCEPTED MANUSCRIPT

$B_1(\mathbf{X}) = 1 - B_0(\mathbf{X})$ and these can be explicitly found. Because frequencies of the two mutational classes add to unity, we can eliminate X_1 in $B_0(X_0, X_1)$ and write this just as a function of X_0, namely $B_0(X_0, 1 - X_0)$. We thus define $b(x) = B_0(x, 1 - x)$ and

$$b(x) = \frac{(1-\mu)\,x}{1 - \sigma + \sigma(1-\mu)x}. \tag{4}$$

It then follows from Eqs. (1), (3) and (4) that in an effectively infinite population, the frequency of the fittest class, $X_{\text{fittest}}(t)$, obeys the deterministic equation $X_{\text{fittest}}(t+1) = b(X_{\text{fittest}}(t))$. When $\mu < \sigma$ this equation has the equilibrium solution

$$\hat{X}_{\text{fittest}} = \frac{\sigma - \mu}{\sigma(1-\mu)}. \tag{5}$$

In a large (effectively infinite) population, any non-zero frequency of the fittest class will approach the polymorphic equilibrium of Eq.(5).

As we next show, when relating μ and σ to parameters in Haigh's Model, the condition $\mu < \sigma$ automatically emerges.

Accepted manuscript

4 Relating Parameters

In the standard model of Muller's ratchet (Haigh 1978), there is no upper limit to the number of mutations that an individual can carry, hence

ACCEPTED MANUSCRIPT

the parameter Q (introduced in Section 2) is infinite: $Q = \infty$. The mutations in Haigh's Model have a Poisson distribution given by $M_{k,j}^{(\infty)} = U^{(k-j)}e^{-U}/(k-j)!$ where U is the genomic mutation rate (the expected number of new deleterious mutations in a newly born offspring). Furthermore, the relative viability of individuals with k mutations is given by $(1-s)^k$ where the selection coefficient s lies in the range 0 to 1.

A plausible way to relate parameters in the truncated model (which has $Q = 1$) and Haigh's model (which has $Q = \infty$) is to ensure that:

(i) the probability of the fittest class undergoing mutation is the same in each model

(ii) the equilibrium frequency of the fittest class, in an infinite population, is the same in each model.

We note that the probability of mutation of the fittest class is μ in the truncated model, and is $1 - \exp(-U)$ in Haigh's model. We also note that in the truncated model, the frequency of the fittest class in adults, in an infinite equilibrium population, is given by Eq. (5), i.e., $(\sigma - \mu) \,/\, [\sigma(1-\mu)]$. In Appendix A we show that in Haigh's model, the frequency of the fittest class *in adults*, in an infinite equilibrium population, is $\exp(-(1-s)U/s)$.

Accepted manuscript

ACCEPTED MANUSCRIPT

Combining these results leads to

$$\mu = 1 - \exp(-U) \tag{6}$$

$$\sigma = \frac{1 - \exp(-U)}{1 - \exp(-U/s)}. \tag{7}$$

There are several points to note about these forms for μ and σ.

(a) When the genomic mutation rate, U, is small ($U \ll 1$), the parameter μ is also small and approximately equal to U. However, since U or U/s need not necessarily be small, the parameters μ and σ are not always small, and in some cases they may be close to unity.

(b) When the ratio U/s is large ($U/s \gg 1$), we have $\sigma \simeq 1 - \exp(-U)$ and in this regime, σ changes appreciably with U, but is quite insensitive to changes in s.

(c) From Eqs. (6) and (7), it follows that $\mu/\sigma = 1 - \exp(-U/s)$, hence the above method of relating parameters in the two models automatically yields $\mu/\sigma < 1$. This is the condition required for the truncated model to have a polymorphic equilibrium when population size is effectively infinite. If the condition $\mu/\sigma < 1$ did not hold then the analogue of an error catastrophe (Maynard Smith 1989, pp 20-23) would occur in the truncated model, where the fittest class would not be present at equilibrium in an effectively infinite

Accepted manuscript

ACCEPTED MANUSCRIPT

population.

There is an intuitive explanation why $\mu/\sigma < 1$: In Haigh's model, the quantity s is a measure of how deleterious a single mutation is. However there can be an indefinitely large number of deleterious mutations in Haigh's model, hence selection against deleterious mutations can become arbitrarily large and can always exceed the "strength" of mutation (U). It is thus natural that in the mapping from Haigh's model to the truncated model, the feature that 'selection is stronger than mutation' directly manifests itself.

5 Time of Loss of the Fittest Class

We now consider population number to be regulated at a finite value, via unspecified ecological processes, so that N adults are present in the population each generation. We assume that the number of offspring that are born and survive viability selection are sufficiently large ($\gg N$), that both selection and mutation are essentially deterministic processes (with negligible variation around expected values). Given this, the principal place in the lifecycle where randomness occurs is the ecological thinning of the population to N individuals.

Thinning corresponds to sampling a population without replacement. However, given that a sample (of size N) is drawn from the much larger

Accepted manuscript

ACCEPTED MANUSCRIPT

number of individuals that survive selection, thinning is well approximated by sampling with replacement. This means, in particular, that when the population size is finite, the truncated model can be treated as a Wright-Fisher model (Fisher 1922; Wright 1931) of haploid asexual individuals with one locus and two alleles. We shall perform a statistical analysis within the framework of such a model. This is equivalent to following the fates of a very large number of replicate populations and determining the probability distribution for the random time at which the frequency of the fittest class of adults vanishes.

We note that in any replicate population, the frequency of adults in the fittest class in generation t, namely $X_{\text{fittest}}(t)$, is a random variable that takes one of the particular values

$$x_n = n/N, \qquad n = 0, 1, 2, ..., N. \tag{8}$$

In different populations, $X_{\text{fittest}}(t)$ may have different values. If $X_{\text{fittest}}(t)$ equals x_n in a particular replicate population then there are n fittest individuals in that population and $N - n$ individuals who carry more mutations than individuals in the fittest class.

When the fittest class is lost, its frequency, $X_{\text{fittest}}(t)$, achieves the value zero, and a click of the ratchet is said to have taken place. Our approach

Accepted manuscript

ACCEPTED MANUSCRIPT

to determining the distribution of times between clicks of the ratchet is non-standard since we describe the behaviour of a population immediately after a click of the ratchet. That is, immediately after the time where loss of (what was) the fittest class has occurred, and offspring of some suboptimal adults have just been promoted to the fittest class.

We measure time so that a click of the ratchet occurred immediately prior to time $t = 0$. Thus in our approach, all replicate populations possess *one or more* individuals in the (newly promoted) fittest class at time $t = 0$. There are then two processes that occur to the distribution (over all replicate populations) of the frequency of the fittest class.

The first process is *relaxation* over time of the distribution of the frequency of the fittest class. Immediately after a click of the ratchet, the distribution of the frequency of the fittest class will not have an equilibrium form. We assume, however, that this distribution relaxes to a characteristic form, and during the time of establishment of this characteristic form, a negligible fraction of all replicate populations lose the fittest class. This is a simplifying assumption that neglects the time of relaxation of the distribution of the frequency of the fittest class, compared with the mean time it takes to lose the fittest class. The characteristic form of the distribution of the frequency of the fittest class that is achieved after relaxation is termed a *quasistationary distribution* (Darroch and Seneta 1965). Formally,

Accepted manuscript

ACCEPTED MANUSCRIPT

the quasistationary distribution can be defined as the equilibrium probability distribution of the frequency of the fittest class, when loss of the fittest class cannot occur.

The next process that occurs is *loss* of the fittest class. As time proceeds, loss of the fittest class occurs in some of the replicate populations, and this is signalled by the occurrence (and growth) of a spike in the distribution of the frequency of the fittest class at the value zero (i.e., at $X_{\text{fittest}} = 0$). The overall distribution of X_{fittest} then has a spike at $X_{\text{fittest}} = 0$ and the distribution of *non-zero* values of X_{fittest} has a quasistationary *shape*, albeit with a reduced amplitude. Figure 1 illustrates this situation.

Figure 1

In many regards, the quasistationary distribution plays the role of an equilibrium distribution of all non-zero values X_{fittest}.

In Appendix B, we give details of the stochastic loss of the fittest class in terms of a Wright-Fisher model. To determine the expected time of loss of the fittest class (and the relaxation time), we introduce a column vector $\mathbf{v}$ and a matrix $\mathbf{w}$ that originate in the Wright-Fisher model underlying the problem. The vector $\mathbf{v}$ contains the probabilities of transitions of non-zero values of X_{fittest} to *loss* of the fittest class (i.e., to X_{fittest} becoming zero);

Accepted manuscript

ACCEPTED MANUSCRIPT

the matrix $\mathbf{w}$ contains the probabilities of transitions between all different non-zero values of X_{fittest} (see Appendix B for details). The elements of $\mathbf{v}$ and $\mathbf{w}$ are given by

$$\left.\begin{aligned} v_n &= [1 - b(x_n)]^N \\ w_{n,m} &= \binom{N}{n} [b(x_m)]^n [1 - b(x_m)]^{N-n} \end{aligned}\right\} n, m = 1, 2, ..., N \tag{9}$$

where $b(x)$ is given in Eq. (4) and $\binom{N}{n} = \frac{N!}{(N-n)!n!}$ denotes a binomial coefficient.

The quasistationary distribution is a column vector, which we write as $\hat{\mathbf{q}}$. It contains the probabilities of all non-zero values of X_{fittest}. The quasistationary distribution is straightforward to find, using numerical packages, since it is the right eigenvector of the matrix $\mathbf{w}$ associated with the largest eigenvalue; it is normalised to unity: $\sum_{n=1}^{N} \hat{q}_n = 1$.

The random time of loss of the fittest class, which is identified with the time of a click of the ratchet, is written T_{click}. This has a distribution which is shown in Appendix B to be approximately that of a geometric random variable. Such a distribution is the discrete analogue of an exponential distribution.

With a prime, $'$, denoting transpose of a matrix, and with $t = 1, 2, 3,$

Accepted manuscript

ACCEPTED MANUSCRIPT

..., the distribution of T_{click}, is approximately given by $\text{Prob}\{T_{\text{click}} = t\} = \mathbf{v}'\hat{\mathbf{q}}\,(1 - \mathbf{v}'\hat{\mathbf{q}})^{t-1}$. Such a distribution is fully characterised by its expected value, which is the mean time of loss of the fittest class, and this is given by

$$\mathrm{E}[T_{\text{click}}] = (\mathbf{v}'\hat{\mathbf{q}})^{-1}\,. \tag{10}$$

This result is intuitively reasonable. Loss of the fittest class occurs at a characteristic rate of $\mathbf{v}'\hat{\mathbf{q}}$ which is determined by transitions of X_{fittest} to the value zero, which are governed by the vector $\mathbf{v}$ of Eq. (9), and these transitions occur from *the* characteristic distribution of the problem, namely the quasistationary distribution, $\hat{\mathbf{q}}$.

The mean value of the relaxation time is given by $T_{\text{relax}} = 1/\ln(\lambda_1/\lambda_2)$ where λ_1 and λ_2 are eigenvalues of the matrix $\mathbf{w}$ of largest and next largest size. The eigenvalue λ_1 is always positive (see Appendix B) and λ_2 is assumed real and positive, and found to be so in all cases investigated.

6 Diffusion Approximation

The expression given in Eq. (10) for $\mathrm{E}[T_{\text{click}}]$ is straightforward to evaluate using standard numerical methods. There are, however, practical limits on the largest population size that can be considered (determined by computer

Accepted manuscript

ACCEPTED MANUSCRIPT

memory size). We present results from a diffusion approximation, since solving a diffusion equation is often more tractable than dealing with very large matrices. The diffusion approximation has been found to be of great utility in problems associated with random genetic drift (Kimura 1964).

In Appendix C we provide details of a diffusion analysis that captures all of the features present in Figure 1, namely a description of all values of the frequency of the fittest class, including the value zero, by ensuring that probability is conserved for all times (cf. McKane and Waxman 2007).

Under a diffusion approximation it is assumed that $1/N$ is a small parameter and both frequencies and time are approximated as continuous quantities. This results in the quasistationary distribution, $\hat{\mathbf{q}}$, being replaced by a *probability density* of a *continuous* random variable – which represents the frequency of the fittest class. We write this probability density as $q(x)$.

A diffusion approximation leads to the time of loss of the fittest class, T_{click}, having, approximately, an exponential distribution. The quantity $(\mathbf{v}'\hat{\mathbf{q}})^{-1}$, appearing in Eq. (10), is replaced in a diffusion approximation by $2N\,[q(0)]^{-1}$ and the expected time of loss of the fittest class is

$$\mathrm{E}[T_{\text{click}}] \simeq 2N\,[q(0)]^{-1}\,. \tag{11}$$

The diffusion analysis given in Appendix C yields an equation for $q(x)$

Accepted manuscript

ACCEPTED MANUSCRIPT

that is numerically solved for $q(0)$. Table 1 indicates that the diffusion results for $\mathrm{E}[T_{\mathrm{click}}]$ typically differ by only a few percent from those of the Wright-Fisher model, with one exception: $N = 100$, $U = 1$ and $s = 0.5$. For these parameter values, gene frequency changes/generation are not small, and a diffusion approximation would not be expected to hold. However even in this case, Wright-Fisher and diffusion results differ only by a factor ~ 3 in the mean time of loss of the fittest class.

7 Individual-based simulations

In addition to the analysis already presented, we have also carried out individual-based simulations, employing methods described elsewhere (Loewe 2006). These were computed by Simulator005, Project 1 of evolution@home, a global computing system for evolutionary biology (Loewe 2002, 2007). The simulated model is equivalent to the model of Haigh (1978), with several extensions. In short, individuals have multiplicative fitness, discrete generations and the number of offspring produced in their lifetime is drawn from a Poisson distribution. The expected number of offspring is determined by the product of (i) the maximal reproductive capacity, (ii) an individual's multiplicative fitness and (iii) a density dependent thinning factor, which depends on the fitnesses of all individuals of the previous generation and a fixed carrying

Accepted manuscript

ACCEPTED MANUSCRIPT

capacity of the habitat, which equals the expected value of population size. Additionally, each mutation causes a fixed reduction in viability, by a factor $1 - s$, and the number of new mutations/generation are Poisson distributed, with expected value U. A click of the ratchet is recorded at the time when all individuals in the population have more deleterious mutations than the last recorded best class. The first click of a population is not used in click-time estimates, since the population may not be in a typical state at the time of click. The click times reported here are effective click times, which we define as (click time)/(click size), where "click size" denotes the number of mutational classes that are skipped at a click. For most clicks, click size is 1, but for high mutation rates and low selection coefficients, larger click sizes do occur. During a simulation, properties of all ratchet clicks are recorded as a stream, which is also simultaneously evaluated to compute some descriptive statistics, such as the arithmetic mean of all effective click times, or the mean of various phases of a click. This mean is then treated as one "single-run-result". To obtain the values reported here, we considered the set of all such single-run-results that were available in the evolution@home database for that parameter combination. For each of these sets we report the arithmetic mean and the approximate 2.5% and 97.5% percentiles in Table 1. Naturally, the percentiles for this distribution of averages are expected to be narrower than the corresponding percentiles for the distribution of individual

Accepted manuscript

ACCEPTED MANUSCRIPT

click times themselves.

8 Comparison of Results

In Table 1 we give results for the mean time between clicks of the ratchet, under: (i) the Wright-Fisher model, (ii) a diffusion analysis of this model, (iii) individual based simulations of a full model of Muller's ratchet and (iv) previous results. The results are illustrated in Figure 2, which compares the performance of the various approaches for a number of selected parameter combinations, and provides an overview of the results in Table 1.

Table 1

Figure 2

In order to compare the predictions of the model introduced here with simulation results, the error of magnitude (EoM) is calculated. The EoM is a measure of agreement of quantities that cannot become negative, but can span many orders of magnitude (Loewe 2007). An EoM of $+1$ (-1) indicates that the model predictions for a click time are 10 times faster (10 times slower) than those observed in simulations.

Accepted manuscript

ACCEPTED MANUSCRIPT

The methods presented here are the first to generate reasonably accurate predictions across the slow *and* the fast regimes of Muller's ratchet (which correspond to parameter regions that are separated by the point where the expected number of individuals in the fittest class is unity). In Table 1 most of the mean click times obtained from our model are within one order of magnitude of simulated values ($|EoM| < 1$). The expected click times that overestimate the simulation click times belong to both Stephan's regime (data sets 1, 5, 17, 19, 21, 22, 29) and Gessler's regime (data sets 2, 3, 6, 8, 9, 11, 12, 15, 18, 21, 24, 27, 32, 33, 35, 36).

We note that the approximations presented here provide more than just point estimates; they allow estimation of arbitrary percentiles. These often overlap with the percentiles of the set of mean effective click times observed in simulations, even though the latter percentiles are technically different from the analytical approximations, which give broader percentiles for the individual click times. (The simulations give the percentiles of the *average* click time, as observed in a simulation with *many* clicks.) This leads to the expectation that the simulated percentiles in Table 1 should be smaller than the analytically percentiles, which is indeed the case.

One prediction of Appendices B and C is that the click time-intervals of Muller's ratchet are approximately exponentially distributed. To our knowledge this prediction is new and we approximately test it using our simulation

Accepted manuscript

ACCEPTED MANUSCRIPT

results. To this end, we compare two independent approximate estimates of 95% percentiles of the mean click time. One estimate *predicts* percentiles from an exponential distribution based on the mean click time observed in our simulation results and the number of observations. The other estimate *observes* percentiles from repeated simulations, where each simulation contributes one observation of mean click time. We assume that errors associated with estimating the mean follow a normal distribution, thus 95% percentiles can be approximated by twice the standard error SE. We derive SE predictions from the average number of observed clicks / simulation (see Table 1) and an estimate of the mean, which equals the standard deviation under the assumption of an exponential distribution. We thus have

$$SE[T_{\text{click}}] \approx E[T_{\text{click}}]/\sqrt{clicks/runs} \tag{12}$$

where the corresponding quantities are reported in Table 1. The predicted approximate 95% quantiles are often very close to results from simulations. Both values are usually correct within a factor of two, the only exceptions are the lower bounds in datasets 1 and 28, and the lower bounds in datasets 16 and 34 are borderline. Inaccuracies are probably due to noise and the fact that not all runs observed an equal number of clicks, rendering Eq. (12) only approximate. Previous observations of Muller's ratchet confirm our results.

Accepted manuscript

ACCEPTED MANUSCRIPT

In a small meta-analysis we estimated the coefficient of variation (CV = stdev/mean) from reported means, standard errors and numbers of observed clicks for previous simulation studies (Gordo & Charlesworth 2001, Gessler 1995, Stephan et al. 1993, Haigh 1978 and Charlesworth & Charlesworth 1997). Assuming that such an analysis is possible, we find that out of 165 reported parameter combinations that we analysed, 27 (or 16%) have CV > 2 and 11 (or 7%) have CV < 0.5. The resulting percentages are not substantially different from those in our results and again might say more about inherent noise and the difficulties of the estimation procedure than about true deviations. As a result, from this work we can conclude that the variance of the click time of Muller's ratchet is now reasonably well understood for a majority of parameter combinations and much easier to predict from a known mean click time than the mean itself. More detailed analyses of the minority of outliers might eventually reveal additional insights into the ratchet mechanism.

Accepted manuscript

9 Discussion

In this work, we have presented a new approach to computing the speed of Muller's ratchet, by relating it to a one locus Wright-Fisher model for a haploid asexual population. Direct analysis of this model reduces to employing

ACCEPTED MANUSCRIPT

numerical linear algebra to solve a matrix equation. An approximate approach uses a diffusion approximation, based on the probability conserving diffusion approach of McKane and Waxman (2007). Both approaches perform reasonably well over a broad range of parameters, as can be seen from comparisons with individual-based simulations of a somewhat full model of Muller's ratchet obtained from the evolution@home global computing system (Loewe 2006, Loewe 2007). Our theoretical results show that approximately, the time intervals between clicks of Muller's ratchet are exponentially distributed. This finding is confirmed with reasonable accuracy by our simulations as well as a small meta-analysis of previous simulation results, and makes it possible to estimate the percentiles of click times of Muller's ratchet from the expected time between clicks, $E[T_{\text{click}}]$. We can compute $P_x[T_{\text{click}}]$, the x-th percentile of click time, from the known quantile function of the exponential distribution:

$$P_x[T_{\text{click}}] = E[T_{\text{click}}] \times \log\left(\frac{1}{1-x}\right). \tag{13}$$

Since predicting the mean speed of Muller's ratchet is notoriously difficult, it would be very surprising if the approach presented here is capable of perfect predictions. Indeed, given the difficulties of predicting click times, which can vary over *many* orders of magnitude, it is surprising how often the approach presented agrees with simulations within one order of magnitude.

Accepted manuscript

ACCEPTED MANUSCRIPT

There is generally very good agreement between the Wright-Fisher and its diffusion approximation. This makes it clear that the Wright-Fisher model is robust to the approximation of frequencies and time as continuous quantities, which is made under a diffusion approximation. Also, in the diffusion analysis the click time arises from the quantity $2N/q(0)$, where $q(x)$ is the quasistationary distribution. Even in situations with very large times between clicks, (small $q(0)$), the close comparison of Wright-Fisher and diffusion results indicates that numerical errors are not a significant problem. We can go further and say that the high level of agreement of Wright-Fisher and diffusion results validate the diffusion approximation over a broad range of parameters.

The one locus model introduced here is defined in a parameter space of the mutation rate, μ, and the selection coefficient, σ. The predictions, however, relate to a multilocus model whose parameter space is the genomic mutation rate, U, and the selection coefficient of deleterious mutations, s. The detailed predictions rely on the mapping between (μ, σ) and (U, s). The mapping implicitly assumes a mutation-selection balance that is not obviously attainable in the fast (quasi-deterministic) regime of operation of the ratchet. There is thus the suspicion that different – perhaps more realistic – mappings between (μ, σ) and (U, s) might improve the accuracy of the model. This suspicion is somewhat strengthened by the observation that estimates in the fast regime

Accepted manuscript

ACCEPTED MANUSCRIPT

are typically larger than the simulation results, even though the results are correct within an order of magnitude. There may thus be room for further development in the mapping between the different models. However, some features suggested by the mapping do appear to be qualitatively correct. In particular, the property of Eqs. (6) and (7), that when the ratio of genomic mutation rate to selection coefficient is large ($U/s \gg 1$), the selection coefficient of the one locus model, σ, is approximately equal to $1 - \exp(-U)$, i.e., insensitive to changes in s. This behaviour suggests that in such a regime, the time between clicks of the ratchet is insensitive to the value of s and depends primarily on U. Such a feature can be seen to roughly hold in Table 1 when $U/s \gtrsim 10$.

There is also room for further improvement in our simplifying approximation, of neglecting the time of relaxation of the distribution of the frequency of the fittest class. This simplifying approximation is equivalent to the approximation of the distribution of the frequency of the fittest class, immediately after a click of the ratchet, by the quasistationary distribution. It may be possible to determine a more realistic form for this initial distribution, and hence obtain a more accurate distribution of times between clicks of the ratchet.

An approximate method is useful if its computational complexity remains within reasonable bounds. The solution of the Wright-Fisher model depends

Accepted manuscript

ACCEPTED MANUSCRIPT

on numerically solving a matrix that grows with the square of population size. The resulting computing demands do place some limits the applicability of this method, however with modern packages (such as Matlab), calculation for the largest population size that we have considered here ($10,000$ individuals) takes approximately 30 minutes on a standard desktop computer. The complexity of simulations also grows with population size and simulations require many runs to accurately estimate the distribution of click times. The diffusion approximation may be a convenient half-way house, since it agrees closely with the Wright-Fisher results and can be rapidly computed for arbitrary values of N, U and s. For example, the time required for calculation of the diffusion result for data set 36 (the final row of the table) was around 20 seconds.

To summarise, we believe three advances have been made in this work. First, we have presented a simple model for Muller's ratchet that allows computation of the mean time between clicks, with reasonable accuracy, over both the slow stochastic and the fast quasi-deterministic regimes of operation of the ratchet. This makes it plausible that a single theoretical approach can be developed to determine the rate of the ratchet, to high precision, over all regimes of operation, which is an important goal in the theoretical development of integrated linkage theories in the population genetics of mutations (Loewe and Hill 2010). Second, this work allows the approximate estima-

Accepted manuscript

ACCEPTED MANUSCRIPT

tion of percentiles of the distribution of click times of Muller's ratchet, based on the mean time between clicks. Lastly, the distribution of time intervals between clicks is approximately exponential. These last two points could facilitate better comparisons between analytical predictions and simulations.

The work presented here may be a reasonable starting point towards further theoretical understanding of the rate of Muller's ratchet.

Accepted manuscript

ACCEPTED MANUSCRIPT

ACKNOWLEDGEMENTS

We thank the participants of evolution@home, in particular Rechenkraft.net, for contributing the CPU-time required to compute the simulation results presented here.

We thank the anonymous reviewers for constructive comments that have significantly improved the manuscript.

The Centre for Systems Biology at Edinburgh is a Centre for Integrative Systems Biology (CISB) funded by BBSRC and EPSRC, reference BB/D019621/1.

Accepted manuscript

ACCEPTED MANUSCRIPT

Appendix A: Relative Frequency of the Fittest Class in Adults

In this Appendix we establish a result for Haigh's model of Muller's ratchet (Haigh 1978) that is made use of in this work. The result states that when population size is infinite, so there are no stochastic effects, the relative frequency of the fittest class *in adults at equilibrium is* $\exp(-(1-s)U/s)$, where s is the selection coefficient against a deleterious mutation and U is the genomic mutation rate.

To begin, let $X_k(t)$ denote the frequency of adults with k $(= 0, 1, 2, ...)$ mutations in generation t. For an infinite population, Haigh's model is summarised by Eqs. (1) and (2) of the main text in which $Q = \infty$, the distribution of mutations is $M_{k,j}^{(\infty)} = U^{(k-j)}e^{-U}/(k-j)!$, and the viability of individuals with k mutations is $(1-s)^k$.

A simple way to proceed is to start with a mutation-free population, and iterate Eq. (1) for $X_k(t)$. We thus take $X_0(0) = 1$ and set $X_k(0) = 0$ for $k \neq 0$. After one iteration we find $X_k(1) = [(1-s)U]^k e^{-(1-s)U}/k!$ and this suggests that $X_k(t)$ may have the form of a Poisson distribution for general t. Accordingly, we conjecture that $X_k(t) = [\alpha(t)]^k \exp[-\alpha(t)]/k!$ and find that $X_k(t+1)$ has, indeed, the same form as $X_k(t)$ with $\alpha(t+1) = (1-s)[U+\alpha(t)]$. It follows that $X_k(t)$ has the Poisson form conjectured, when starting from

Accepted manuscript

ACCEPTED MANUSCRIPT

a mutation-free population (which has $\alpha(0) = 0$) and the solution for $\alpha(t)$ is $\alpha(t) = (1-s)U\left[1-(1-s)^t\right]/s$. The equilibrium relative frequencies of adults are thus $X_k(\infty) = \left[\alpha(\infty)\right]^k \exp\left[-\alpha(\infty)\right]/k!$ where $\alpha(\infty) = (1-s)U/s$. It may be verified that these frequencies constitute a stationary solution of Eq. (1). The relative frequency of the fittest class in adults at equilibrium, $X_0(\infty)$, is thus $\exp(-(1-s)U/s)$.

For completeness, we shall derive a result that appears in Haigh's (1978) analysis. We note that the equilibrium relative frequency of the fittest class in juveniles (i.e., in newly born individuals who have not been subject to selection, and carry no mutations) is derived from the distribution of adults, by subjecting it to a round of mutation. The equilibrium relative frequency of fittest juveniles is thus given by $\sum_{j=0}^{\infty} M_{0,j}^{(\infty)} X_j(\infty) \equiv \exp(-U) X_0(\infty) = \exp(-U/s)$. This is the standard result presented in the literature (Haigh 1978).

Accepted manuscript

ACCEPTED MANUSCRIPT

Appendix B: Wright-Fisher Model

In this Appendix we determine properties of the model used in this work to calculate the rate of Muller's ratchet. The model is equivalent to a Wright-Fisher model (Fisher 1922; Wright 1931) for a population of haploid asexual individuals with one locus and two alleles. Carriers of the A_0 allele correspond to individuals in the fittest class, while carriers of the A_1 allele have more mutations than individuals in the fittest class and correspond to an amalgamation of all suboptimal fitness individuals. An assumption underlying this work is that such a Wright-Fisher model describes the behaviour of the fittest class of individuals after one click of the ratchet, i.e., immediately after loss from the population of (what was) the fittest class.

The total number of adults in a population is taken to be ecologically regulated, so that in every generation, the number of adults takes the finite fixed value N. The relative frequency of adults in the fittest class, in generation t, is written $X_{\text{fittest}}(t)$.

Accepted manuscript

Stochastic dynamics

The random dynamical behaviour of $X_{\text{fittest}}(t)$ is given in a Wright-Fisher model by $X_{\text{fittest}}(t+1) = R(t)/N$ where $t = 0, 1, 2, ...$ and $R(t)$, for different t, are independent binomial random variables. If, in generation t, the frequency of the fittest class has the particular value x_n, then $R(t)$ has the binomial

ACCEPTED MANUSCRIPT

distribution $\mathrm{Binom}(N, b(x_n))$, where $b(x)$ is given in Eq. (4). In this case the expected value of $X_{\mathrm{fittest}}(t+1)$ is $b(x_n)$ and its variance is $b(x_n)\left[1-b(x_n)\right]/N$.

A Wright-Fisher model determines the statistics associated with an infinite number of replicate populations. We write the probability that $X_{\mathrm{fittest}}(t)$ has the value x_n as $F_n(t)$. The quantity $F_0(t)$ is thus the probability that a population has no individuals in the fittest class by generation t. Equivalently, we can say that $F_0(t)$ is the probability that the fittest class has been lost from a population by generation t.

The dynamical "Wright-Fisher" rule connecting the distribution of the frequency of the fittest class in different generations is $F_n(t+1) = \sum_{m=0}^{N} W_{n,m} F_m(t)$ where $W_{n,m}$ are elements of a matrix of transition probabilities. The $W_{n,m}$ are given by $W_{n,m} = \binom{N}{n}\left[b\left(x_m\right)\right]^n\left[1-b(x_m)\right]^{N-n}$ where $\binom{N}{n} = \frac{N!}{(N-n)!n!}$ is a binomial coefficient and $b(x)$ is given in Eq. (4). The dynamical rule can be written as the matrix equation

$$\mathbf{F}(t+1) = \mathbf{W}\mathbf{F}(t) \tag{14}$$

Accepted manuscript

ACCEPTED MANUSCRIPT

where $F(t)$ is a column vector with $N+1$ elements and W is the matrix

$$\mathbf{W} = \begin{pmatrix} W_{0,0} & W_{0,1} & \cdots & W_{0,N} \\ W_{1,0} & W_{1,1} & \cdots & \vdots \\ \vdots & & & \vdots \\ W_{N,0} & \cdots & \cdots & W_{N,N} \end{pmatrix}. \tag{15}$$

A property of Eq. (14) that we shall shortly use is

$$\sum_{n=0}^{N} F_n(t) = 1. \tag{16}$$

This result holds for all t, and follows from

$$\sum_{n=0}^{N} W_{n,m} = 1 \tag{17}$$

(which can be proved using the binomial theorem). Equation (16) is a statement that the total probability of all possibilities (the occurrence of all possible values of $X_{\text{fittest}}(t)$) is maintained by Eq. (14) and takes the value of unity in all generations.

Properties of the dynamics

It is convenient to go to a dynamical description where the probability of loss of the fittest class by generation t, namely $F_0(t)$, is treated separately to

Accepted manuscript

ACCEPTED MANUSCRIPT

the probabilities of other values of $X_{\text{fittest}}(t)$ (cf. Waxman 2008). To obtain this description, we write the the distribution $F(t)$ in the form

$$\mathbf{F}(t) = \begin{pmatrix} F_0(t) \\ \\ \mathbf{f}(t) \end{pmatrix}. \tag{18}$$

Here the column vector $f(t)$ contains the probabilities of all non-zero values of $X_{\text{fittest}}(t)$. The elements of $f(t)$ are written as $f_n(t)$, with n running from 1 to N (i.e., n does not take the value zero).

In what follows, we use a prime, $'$, to denote the transpose of a matrix, and write a column vector of N ones as

$$\mathbf{L} = (\underbrace{1, 1, ..., 1}_{N})'. \tag{19}$$

We then use Eq. (16) to write $F_0(t) = 1 - \sum_{n=1}^{N} f_n(t)$ and express this as

$$F_0(t) = 1 - \mathbf{L}'\mathbf{f}(t). \tag{20}$$

Attention can now be focussed on $f(t)$ which underlies the behaviour of $F_0(t)$.

Proceeding, it may be verified that the matrix W can be written in the

Accepted manuscript

ACCEPTED MANUSCRIPT

form

$$\mathbf{W} = \begin{pmatrix} 1 & \mathbf{v}' \\ \mathbf{0} & \mathbf{w} \end{pmatrix}. \tag{21}$$

Here 0 is a column vector with N vanishing components, v is a column vector with N non-zero components given by $v_m = W_{0,m}$ where m runs from 1 to N (and does not take the value zero). The quantity w is an $N \times N$ matrix with elements $w_{n,m} = W_{n,m}$, where n and m run from 1 to N (and also do not take the value zero).

Equations (14), (18) and (21) yield

$$\mathbf{f}(t+1) = \mathbf{w}\mathbf{f}(t) \tag{22}$$

$$F_0(t+1) = F_0(t) + \mathbf{v}'\mathbf{f}(t) \tag{23}$$

Equation (22) has the solution

$$\mathbf{f}(t) = \mathbf{w}^t\mathbf{f}(0). \tag{24}$$

The principal long time behaviour of $f(t)$ is exponential decay, corresponding to an exponential decrease of the probability of obtaining any non-zero frequency of the fittest class. To see this, consider the probability of

Accepted manuscript

ACCEPTED MANUSCRIPT

obtaining any non-zero value of the frequency of the fittest class, namely $\sum_{n=1}^{N} f_n(t) = L'f(t) = L'w^t f(0)$. Using a spectral decomposition of the matrix w (see, e.g., Eq. (G11) in Bulmer 1994), the leading large time behaviour of $L'w^t f(0)$ is determined by the eigenvalue of w of largest magnitude. We write this eigenvalue as λ_1 and, since all elements of w are strictly positive, λ_1 is positive by the Perron Frobenius theorem (see e.g., Strang 1988). It follows that the leading large time behaviour of $\sum_{n=1}^{N} f_n(t)$ is proportional to λ_1^t which can be written as $\exp\left[-t\log\left(\lambda_1^{-1}\right)\right]$. Hence $\sum_{n=1}^{N} f_n(t)$ decays exponentially on a timescale of $1/\log(1/\lambda_1)$. This result, combined with Eq. (20), has the direct implication that at large times $1 - F_0(t)$ is proportional to λ_1^t. This last result can be written as $\exp(-t/\tau)$ where $\tau = 1/\log(1/\lambda_1)$ and defines the characteristic timescale τ over which $1 - F_0(t)$ vanishes.

Using Eq. (20), we express Eq. (24) as

$$\mathbf{f}(t) = [1 - F_0(t)]\,\mathbf{q}(t) \tag{25}$$

where

$$\mathbf{q}(t) = \frac{\mathbf{w}^t\mathbf{f}(0)}{\mathbf{L}'\mathbf{w}^t\mathbf{f}(0)}. \tag{26}$$

For large t, the quantity $q(t)$ is guaranteed to approach the right eigenvector of the matrix w associated with the eigenvalue of largest magnitude (see,

Accepted manuscript

ACCEPTED MANUSCRIPT

e.g., Eq. (G11) in Bulmer 1994). We write this eigenvector as $\hat{q}$ and

$$\hat{\mathbf{q}} = \lim_{t\to\infty} \frac{\mathbf{w}^t \mathbf{f}(0)}{\mathbf{L}'\mathbf{w}^t \mathbf{f}(0)} \tag{27}$$

This vector is normalised to unity in the sense $L'\hat{q} \equiv \sum_{n=1}^{N} \hat{q}_n = 1$.

The vector $\hat{q}$, which is a probability distribution, is unambiguously defined (it is the right eigenvector of w associated with the largest eigenvalue) and is straightforward to numerically determine. In the genetic drift literature the distribution $\hat{q}$ has been referred to as a quasistationary distribution (Darroch and Seneta 1965) or in a diffusion approximation, as an asymptotic conditional distribution (Ewens 1965) and in a different context as a conditional probability distribution (Nisbet and Gurney 1982). Here the word "conditional" refers to conditioning on the non absence (or loss) of the fittest class.

The largest eigenvalue of the matrix w, namely λ_1, can be expressed in terms of $\hat{q}$. At large t, we replace $q(t)$ in Eq. (25) by $\hat{q}$, yielding $f(t) \simeq [1 - F_0(t)]\,\hat{q}$. Using this result in Eq. (23) yields $F_0(t+1) \simeq F_0(t) + [1 - F_0(t)]\, v'\hat{q}$ which can be written $1 - F_0(t+1) \simeq (1 - \mathbf{v}'\hat{\mathbf{q}})\,[1 - F_0(t)]$. This has the solution $1 - F_0(t) \propto (1 - \mathbf{v}'\hat{\mathbf{q}})^t$ and since, at long times, $1 - F_0(t) \propto \lambda_1^t$ it follows, from comparison of the time dependence of these two expressions

Accepted manuscript

ACCEPTED MANUSCRIPT

for $1 - F_0(t)$, that the largest eigenvalue of w is

$$\lambda_1 = 1 - \mathbf{v}'\hat{\mathbf{q}}. \tag{28}$$

Relaxation time

The characteristic time it takes $q(t)$ (of Eq. (26)) to relax to the quasi-stationary distribution, $\hat{q}$, is given by

$$T_{\text{relax}} = 1/\ln(\lambda_1/\lambda_2) \tag{29}$$

where λ_1 and λ_2 are largest and next largest eigenvalues of the matrix w (λ_2 is assumed real and positive, and found so in all cases considered). This form for T_{relax} follows from a spectral decomposition (Bulmer 1994) of Eq. (26). That is, we take the distribution of $X_{\text{fittest}}(0)$ to have the form $f(0) = a\hat{q} + b\hat{r} + ...$ where a and b are constants and $\hat{r}$ is the right eigenvector of the matrix w associated with eigenvalue λ_2. It follows that numerator on the right side of Eq. (26) is $\lambda_1^t a\hat{q} + \lambda_2^t b\hat{r} + ...$, while the denominator is $\lambda_1^t a + \lambda_2^t c + ...$ where c is another constant. Thus at long times, $q(t) \simeq \hat{q} \times [1 - (\lambda_2/\lambda_1)^t c/a] + (\lambda_2/\lambda_1)^t \hat{r} b/a$. Since $(\lambda_2/\lambda_1)^t$ can be written as $\exp(-t/T_{\text{relax}})$, this means that at long times, $q(t)$ deviates from $\hat{q}$ by

Accepted manuscript

ACCEPTED MANUSCRIPT

an amount proportional to $\exp(-t/T_{\text{relax}})$ and T_{relax} has the natural interpretation as the time over which relaxation occurs to the quasistationary distribution.

There is also a characteristic timescale in Haigh's Model of Muller's ratchet, corresponding to the time, after a click of the ratchet, that it takes the expected numbers of individuals in different mutational classes to come close to their deterministic equilibrium values (Haigh 1978). However this timescale is distinct from T_{relax}, which is a relaxation time of the distribution of the frequency of just the fittest class; the timescale of Haigh is associated with the approximate approach to equilibrium of many different mutational classes.

Time to loss

The distribution of non-zero values of $X_{\text{fittest}}(0)$ is $f(0)$; this distribution describes all non-zero values of the frequency of the fittest class immediately after a click of the ratchet. We shall take $f(0)$ to coincide with the quasistationary distribution, $\hat{q}$. This, we emphasise, is a simplifying assumption that neglects the time of relaxation of $q(t)$ of Eq. (26) to $\hat{q}$. Writing the random time of loss of the fittest class as T_{click}, the simplifying assumption leads to

Accepted manuscript

ACCEPTED MANUSCRIPT

T_{click} having taking the value t with probability

$$\text{Prob}\,\{T_{\text{click}} = t\} = (1-\lambda_1)\lambda_1^{t-1}, \qquad t = 1, 2, ... \tag{30}$$

which is a geometric distribution. It follows from Eq. (30) that the expected time to loss of the fittest class is

$$E[T_{\text{click}}] = \sum_{t=1}^{\infty} t\,\text{Prob}\,\{T_{\text{click}} = t\} = (1-\lambda_1)^{-1}$$

$$= (\mathbf{v}'\hat{\mathbf{q}})^{-1}\,. \tag{31}$$

To establish Eq. (30), we note that since $F_0(t)$ is the probability that the fittest class contains no individuals by generation t, it means $F_0(t) = \text{Prob}\,\{T_{\text{click}} \leq t\}$. Hence for $t = 1$, 2, ..., the probability that T_{click} takes the particular value t is $\text{Prob}\,\{T_{\text{click}} = t\} = F_0(t) - F_0(t-1)$. Using Eqs. (23) and (24) yields $\text{Prob}\,\{T_{\text{click}} = t\} = v'w^{t-1}f(0)$. Under our simplifying assumption, that $f(0) = \hat{q}$, we have that $w^{t-1}f(0) = \lambda_1^{t-1}\hat{q}$ and since, by Eq. (28) $v'\hat{q} = 1-\lambda_1$, we have $v'w^{t-1}f(0) = (1-\lambda_1)\lambda_1^{t-1}$, and hence obtain Eq. (30).

Accepted manuscript

ACCEPTED MANUSCRIPT

Appendix C: Diffusion Approximation

In this Appendix we use a diffusion approximation of the one locus Wright-Fisher model used in this work to describe Muller's ratchet. This Appendix may be omitted by readers who do not require an in-depth knowledge of the technical details of this approximation.

Under a diffusion approximation, frequencies are treated as continuous quantities, and this results in the distribution of the frequency of the fittest class being a probability density. Our analysis establishes the quasistationary distribution (see Appendix B) which is the characteristic description of non-zero values of the frequency of the fittest class.

Equation for quasistationary distribution

Let $\phi \equiv \phi(x,t)$ denote the probability density, in a diffusion approximation, of the frequency of the fittest class; this is an approximation of the distribution $F(t)$ appearing in Eq. (14). The equation obeyed by $\phi(x,t)$ is the forward Kolmogorov diffusion equation (Kimura 1964) for a haploid asexual population of size N, namely $\partial\phi/\partial t = \partial^2 \left[(2N)^{-1} x(1-x)\phi\right]/\partial x^2 - \partial\left[\beta(x)\phi\right]/\partial x$ with $\beta(x) = b(x) - x$ where $b(x)$ is given in Eq. (4). Thus

$$\beta(x) = \frac{x\left[(\sigma-\mu) - \sigma(1-\mu)x\right]}{1-\sigma+\sigma(1-\mu)x}. \tag{32}$$

Accepted manuscript

ACCEPTED MANUSCRIPT

The associated probability current density is $j(x,t) = -\partial\left[(2N)^{-1}x(1-x)\phi\right]/\partial x + \alpha(x)\phi$.

We seek a complete solution of the diffusion equation, similar to that obtained by McKane and Waxman (2007), where probability is conserved at all times. We shall thus explicitly include a term associated with loss of the fittest class. Following the simplifying approximation made in Appendix B, we ignore the assumed short time of relaxation of the solution. Thus with $\delta(x)$ a Dirac delta function and $\Theta(x)$ a Heaviside step function, we directly look for a solution of the form

$$\phi(x,t) = F_0(t)\delta(x) + [1 - F_0(t)]\, q(x)\Theta(x). \tag{33}$$

In this equation $q(x)$ is the quasistationary distribution and $F_0(t)$ is the probability of loss of the fittest class from a population by time t. It follows that $F_0(t)\delta(x)$ is the probability density associated with loss of the fittest class. The other part of the solution, namely $[1 - F_0(t)]\, q(x)\Theta(x)$, describes all non zero values of the frequency of the fittest class. This part of the solution maintains its shape – i.e., its dependence on x, apart from a time-dependent amplitude of $1 - F_0(t)$. The explicit presence of the step function, $\Theta(x)$, in Eq. (33), incorporates the feature that $q(x)$ is zero for $x < 0$ but need not continuously vanish at $x = 0$ (cf. McKane and Waxman 2007).

Accepted manuscript

ACCEPTED MANUSCRIPT

Distributions related to Eq. (33), that exclude the singular delta function term in Eq. (33), have been considered elsewhere (Ewens 1965, Nisbet and Gurney 1982), in different contexts. As the following arguments show, the singular delta function term in Eq. (33) is essential to obtain the correct equation for $q(x)$ and we know of no way, within a correct diffusion analysis, to avoid inclusion of this term.

We shall determine the probability of loss of the fittest class by time t, that is to say $F_0(t)$. As will become evident, this requires knowledge of the quasistationary distribution, $q(x)$, which is the solution of the diffusion equation, Eq. (33), subject to the following conditions:

(a) All replicate populations possess the fittest class at time $t = 0$. Hence $F_0(0) = 0$.

(b) No probability flows out of the $x = 1$ boundary, hence $q(x)$ has vanishing probability current density at $x = 1$.

(c) The zero frequency class is not included in the term $[1 - F_0(t)]\, q(x)\Theta(x)$. Hence $q(x)$ does not contain a delta function at $x = 0$.

(d) The quasistationary distribution, $q(x)$, is normalised to unity: $\int_0^1 q(x)dx = 1$, thereby ensuring that probability is conserved for all times (cf. McKane and Waxman 2007).

In the following calculations, properties of delta functions and step func-

Accepted manuscript

ACCEPTED MANUSCRIPT

tions that we shall repeatedly use are $x\delta(x) = 0$ and $d\Theta(x)/dx = \delta(x)$.

We proceed by substituting Eq. (33) into the diffusion equation with the result

$$\frac{dF_0(t)}{dt}\left[\delta(x) - q(x)\Theta(x)\right]$$

$$= \left[1 - F_0(t)\right]\Theta(x)\left(\frac{1}{2N}\frac{d^2}{dx^2}\left[x(1-x)q(x)\right] - \frac{d}{dx}\left[\beta(x)q(x)\right]\right)$$

$$+ \left[1 - F_0(t)\right]\left[\frac{1}{2N}\frac{d}{dx}\left[x(1-x)q(x)\right]\right]_{x=0}\delta(x). \tag{34}$$

The coefficients of the delta functions on the left and right must balance, and since $\left[(2N)^{-1}\, d\left[x(1-x)q(x)\right]/dx\right]_{x=0} = q(0)/(2N)$ we obtain

$$\frac{dF_0(t)}{dt} = \left[1 - F_0(t)\right]\frac{q(0)}{2N}. \tag{35}$$

The solution of this equation, satisfying condition (a) above (vanishing at $t = 0$) is

$$F_0(t) = 1 - \exp\left(-\frac{q(0)t}{2N}\right). \tag{36}$$

This indicates that the time of loss of the fittest class, following from Eq. (33), is an exponentially distributed random variable. Thus, under a diffusion

Accepted manuscript

ACCEPTED MANUSCRIPT

approximation, the mean time between clicks of the ratchet is $\mathrm{E}[T_{\mathrm{click}}] = 2N/q(0)$. The quantity $(2N)^{-1}q(0)$ is the diffusion approximation of the Wright-Fisher result of Eq. (31), namely $v'\hat{q}$. The diffusion result for $\mathrm{E}[T_{\mathrm{click}}]$ requires knowledge of $q(0)$, which follows from knowledge of $q(x)$.

The equation for $q(x)$ follows by equating the terms on both sides of Eq. (34) that are coefficients of the step functions, and then using Eq. (35) to eliminate $dF_0(t)/dt$. This leads to

$$-\frac{1}{2N}\frac{d^2}{dx^2}\left[x(1-x)q(x)\right] + \frac{d}{dx}\left[\beta(x)q(x)\right] = \frac{q(0)}{2N}q(x) \tag{37}$$

which is a nonlinear eigenvalue equation, since the eigenvalue is proportional to $q(0)$.

Equation (37) fully determines the quasistationary solution describing the problem, when subject to the conditions (b), (c) and (d) given above. We note that the normalisation condition $\int_0^1 q(x)dx = 1$ is required in this problem since Eq. (37) is explicitly nonlinear and this feature makes it unlike a standard (linear) eigenvalue problem, where the eigenvalue is independent of normalisation of the solution.

Numerical solution for $q(x)$

We have found that a successful way of proceeding is to derive and solve

Accepted manuscript

ACCEPTED MANUSCRIPT

an integral equation for $q(x)$. To this end, we introduce

$$H(x) = 2N \int_0^x y^{-1}(1-y)^{-1}\beta(y)dy = a\ln(1+cx) + b\ln(1-x) \tag{38}$$

where we have used Eq. (32) and defined

$$a = 2N\frac{1-\mu}{1-\sigma\mu}, \quad b = 2N\frac{\mu(1-\sigma)}{1-\sigma\mu}, \quad c = \frac{\sigma(1-\mu)}{1-\sigma}. \tag{39}$$

Using $H(x)$ allows Eq. (37) to be written as

$d^2\left[x(1-x)q(x)\right]/dx^2 - d\left[(dH(x)/dx)\,x(1-x)q(x)\right]/dx = -q(0)q(x)$.

We then carry out four steps.

(i) We integrate the above equation from an arbitrary x to $x = 1$ and use vanishing of the probability current density at $x = 1$. This results in $d\left[x(1-x)q(x)e^{-H(x)}\right]/dx = q(0)e^{-H(x)}\int_x^1 q(y)dy$.

(ii) We integrate the last equation from $x = 0$ to an arbitrary x, and obtain $q(x) = q(0)\,(1-x)^{-1}\,x^{-1}e^{H(x)}\int_0^x dye^{-H(y)}\int_y^1 q(z)dz$.

(iii) We use normalisation of $q(x)$, i.e., $\int_0^1 q(x)dx = 1$, to write $\int_y^1 q(z)dz = 1 - \int_0^y q(z)dz$. This allows us to write

$$q(x) = q(0)\left(r(x) - \int_0^1 G(x,y)q(y)dy\right) \tag{40}$$

Accepted manuscript

ACCEPTED MANUSCRIPT

where

$$r(x) = (1-x)^{-1}\, x^{-1} e^{H(x)} \int_0^x e^{-H(z)} dz \tag{41}$$

$$G(x,y) = \Theta(x-y)\,(1-x)^{-1}\, x^{-1} e^{H(x)} \int_y^x e^{-H(z)} dz. \tag{42}$$

(iv) We eliminate $q(0)$ in Eq. (40), by again using normalisation of $q(x)$. The result is an integral equation for $q(x)$ that incorporates Eq. (37) and conditions (b), (c) and (d) above:

$$q(x) = \frac{r(x) - \int_0^1 G(x,y) q(y) dy}{\int_0^1 r(y) dy - \int_0^1 dy \int_0^1 dz G(y,z) q(z)}. \tag{43}$$

We have discretised this equation and then iterated it from an arbitrary initial choice of $q(x)$ and have found that in all cases considered the iteration converges. The solutions for $q(x)$ lead to a result for $\mathrm{E}[T_{\text{click}}] = 2N/q(0)$ that is typically very close to the Wright-Fisher result, $\mathrm{E}[T_{\text{click}}] = (v'\hat{q})^{-1}$.

While a possible approximation of Eq. (43) is $q(x) \simeq r(x)/\int_0^1 r(y) dy$, this is found to be valid in only about 20% of the parameter combinations appearing in Table 1 and is not pursued further here.

Accepted manuscript

ACCEPTED MANUSCRIPT

References

Bachtrog D. 2008 The temporal dynamics of processes underlying Y chromosome degeneration. Genetics 179: 1513-1525.

Bergstrom, C. T. and Pritchard, J. 1998 Germline bottlenecks and the evolutionary maintenance of the mitochondrial genome. Genetics 149:2135-2146.

Bulmer, M. Theoretical Evolutionary Ecology. Sinauer Associates, Sunderland, Massachusetts (1994).

Darroch, J. N. and Seneta, E. 1965 On Quasi-Stationary Distributions in Absorbing Discrete-Time Finite Markov Chains. J. Appl. Prob. 2: 88-100.

Etheridge, A., Pfaffelhuber P. and Wakolbinger A. 2008. How often does the ratchet click? Facts, heuristics, asymptotics. Trends in Stochastic Analysis. Part IV 13 Cambridge University Press

Ewens,W. J. 1965 The adequacy of the diffusion approximation to certain distributions in genetics. Biometrics 21: 386-394.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics 78:737-756.

Fisher, R. A. 1922 On the dominance ratio. Proc. R. Soc. Edinb. 42:321-341.

Gessler, D. D. G. 1995. The constraints of finite size in asexual populations and the rate of the ratchet. Genet. Res. 66:241–253.

Gordo, I. and Charlesworth B. 2000a The degeneration of asexual haploid

Accepted manuscript

ACCEPTED MANUSCRIPT

populations and the speed of Muller's ratchet. Genetics 154:1379-1387.

Gordo, I. and Charlesworth B. 2000b On the speed of Muller's ratchet. Genetics 156:2137-2140.

Haigh, J. 1978 The accumulation of deleterious genes in a population – Muller's ratchet. Theor. Pop. Biol. 14: 251-267.

Haigh, J. 2002 Probability Models, Springer, London (2002).

Heller, R, and Maynard Smith J. 1978. Does Muller's ratchet work with selfing? Genet. Res. 32:289-293.

Kimura M. 1964. Diffusion models in Population Genetics. J. Appl. Prob. 1:177-232

Loewe, L. 2002. Global computing for bioinformatics. Briefings in Bioinformatics 3:377-388.

Loewe, L., 2006 Quantifying the genomic decay paradox due to Muller's ratchet in human mitochondrial DNA. Genet. Res. 87:133-159.

Loewe, L. 2007. Evolution@home: observations on participant choice, work unit variation and low-effort global computing. Software Practice & Experience 37: 1289-1318.

Loewe, L. and Cutter, A. 2008. On the potential for extinction by Muller's Ratchet in Caenorhabditis elegans", BMC Evolutionary Biology 8: 125.

Loewe, L., and Lamatsch, D. K. 2008. Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (Poecilia formosa). BMC

Accepted manuscript

ACCEPTED MANUSCRIPT

Evolutionary Biology 8:88.

Loewe, L. and Hill, W. G. 2010. The population genetics of mutations: good, bad and indifferent. Phil Trans R Soc B 365: 1153-1167.

Maynard Smith, J., 1978. The Evolution of Sex. Cambridge University Press, Cambridge.

Maynard Smith, J., 1989. Evolutionary Genetics. Oxford University Press, Oxford.

Muller, H. J. 1964. The relation of recombination to mutational advance. Mutat. Res. 1: 2-9.

McKane, A. J. and Waxman, D. 2007 Singular solutions of the diffusion equation of population genetics. J. Theor. Biol. 247: 849-858.

Moran, N. A. 1996 Accelerated evolution and Muller's ratchet in endosymbiont bacteria 1996 PNAS 93: 2873-2878

Nisbet R. M. and Gurney W. S. C. 1982 Modelling Fluctuating Populations, John Wiley, New York

Rice, W. R., 1994 Degeneration of a nonrecombining chromosome. Science 263: 230-232

Rispe, C. and N. A. Moran 2000 Accumulation of deleterious mutations in endosymbionts: Muller's ratchet with two levels of selection. Am. Nat. 156: 425-441.

Rouzine, I. M., Brunet, E. Wilke, C.O. 2008. The traveling-wave approach

Accepted manuscript

ACCEPTED MANUSCRIPT

to asexual evolution: Muller's ratchet and speed of adaptation. Theoretical population biology 73: 24-46

Söderberg R. J., Berg O. G, 2007.Mutational interference and the progression of Muller's ratchet when mutations have a broad range of effects. Genetics 177: 971-986.

Stephan, W., Chao, L. and Smale J. G. 1993 The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. Genet. Res. Camb. 61: 225-231.

Stephan W., Kim Y. 2002. Recent applications of diffusion theory to population genetics. In Modern developments in theoretical population genetics Edited by: Slatkin M, Veuille M. Oxford: Oxford University Press 72-93.

Strang, G 1988 Linear Algebra and Its Applications. San Diego: Harcourt Brace.

Wright, S. 1931 Evolution in Mendelian populations. Genetics, 16:97-159.

Accepted manuscript

ACCEPTED MANUSCRIPT

Figure 1 Caption

The Figure illustrates the distribution of the frequency of the fittest class, $X_{\text{fittest}}(t)$, over replicate populations. The Figure was calculated from a Wright-Fisher model describing two mutational classes, namely the fittest class of adults and the class encompassing all suboptimal adults. The spike at $x = 0$ represents populations where the fittest class has been lost by generation t. As time progresses, the spike slowly grows in height, at the expense of the remainder of the distribution, which describes non-zero values of the frequency of the fittest class, and has a near quasistationary form. The height of the part of the distribution describing non-zero values of the frequency slowly decreases over time, but apart from this, maintains its shape.

Figure 2 Caption

The Figure illustrates the dependency of the time between clicks of Muller's ratchet on mutation rate, selection coefficient and population size. For numerical details see Table 1. The first three bars (for each parameter value) represent, respectively, the 2.5%, 50% and 97.5% percentiles of the click time-intervals estimated from the Wright-Fisher model using Eq. (13). The fourth bar gives the estimate from simulations, including a vertical bar to indicate the 2.5% and 97.5% percentiles of the mean click time intervals, as estimated

Accepted manuscript

ACCEPTED MANUSCRIPT

from the simulations specified in Table 1. We employ the results of Stefan and Kim (2002) and Gessler (1995) in a complementary manner to determine the fifth bar: when $N \exp(-U/s) > 1$ we use the result of Stefan and Kim (2002), and when $N \exp(-U/s) < 1$ we use the result of Gessler (1995). The sixth bar gives the result of Rouzine et al. (2008).

Table 1 Caption

This Table gives the speed of the accumulation of mutations by Muller's ratchet, as computed from various different approaches, namely (i) a numerical treatment of a Wright-Fisher model for an asexual population with a single haploid locus with two alleles, (ii) a diffusion analysis of this model and (iii) individually-based simulations of a full model of Muller's ratchet. The table also summarises key results from other approaches.

The parameters N ($= N_e =$ effective population size), U (genomic deleterious mutation rate/individual/generation) and s (deleterious selection coefficient) determine all other columns, which contain results of the various approaches. Combinations of input parameters that lead to $N \exp(-U/s) > 1$ (cf. Haigh 1978), are marked with a "+".

The column headings are as follows.

$\mathrm{E}[T_{\mathrm{click}}]$ gives the arithmetic mean of the effective number of generations between individual clicks that are expected under a Wright-Fisher/diffusion

Accepted manuscript

ACCEPTED MANUSCRIPT

model or as observed in simulations.

2.5% P and 97.5% P denote the lower and upper percentiles P of the distribution of click times. Note that percentiles refer to the distribution of individual click times in the Wright-Fisher and Diffusion columns, as computed using Eq. (13), but in the Simulation column, the percentiles refer to the distribution of mean click times, where each mean is an average over all clicks observed during a single run.

T_{relax} gives the mean time required for the distribution of replicate populations to relax to a quasistationary distribution.

"Simulations": the combined number of observed clicks; the number of simulation runs in which these clicks were observed are given in a separate column. The number of clicks per simulation was not constant but usually similar.

"Expect E±2SE": specify, approximately, the percentiles that would be expected in the two previous columns based on the observed $\mathrm{E}[T_{\text{click}}]$ and our theoretical result that the distribution of click times is approximately exponential.

SGmethod and Rmethod give the predicted click times in generations as computed by previous approaches, where (S) refers to Stephan's diffusion approximation, (G) to Gessler's simulation/analytic method and (R) to the traveling wave approximation of Rouzine et al. (2008, their equation 36).

Accepted manuscript

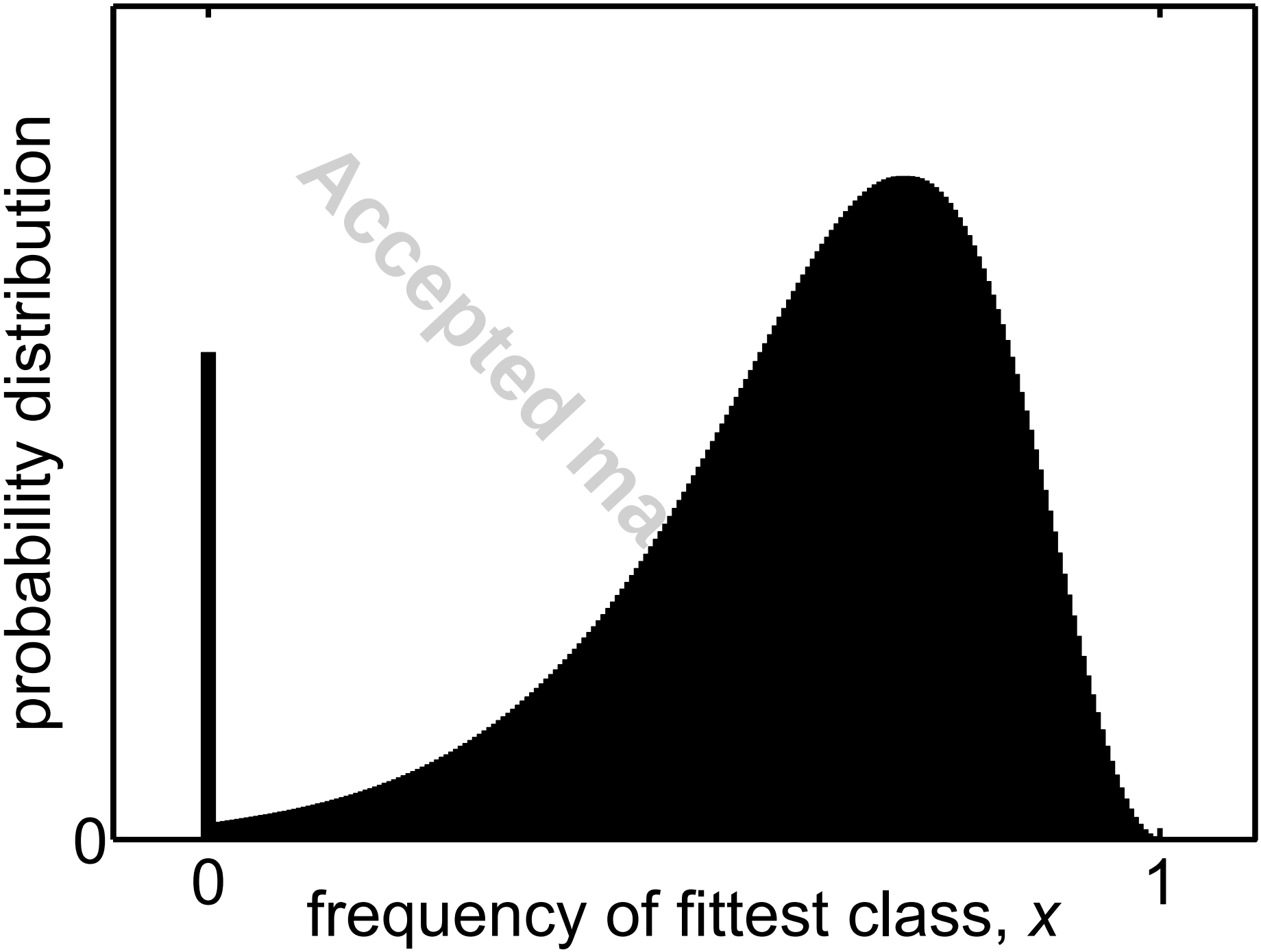
ACCEPTED MANUSCRIPT

We present the results of Stefan and Kim (2002) and Gessler (1995) in a strictly complementary manner; when $N\exp(-U/s) > 1$ we give the result of Stefan and Kim (2002), and when $N\exp(-U/s) < 1$ we give the result of Gessler (1995). The notation 'Err' indicates numerical errors that prevented the automated computation of results in our implementation.

"Error": the error of magnitude (EoM) between the diffusion approximation of the Wright-Fisher model and the mean click time as observed in simulations. The quantity EoM is defined by $\log_{10}(\mathrm{E}[T_{\mathrm{click}}]_{\mathrm{diffusion}}/\,\mathrm{E}[T_{\mathrm{click}}]_{\mathrm{simulation}})$. Each order of magnitude in discrepancies between predictions and observations is highlighted by one star (*).

Accepted manuscript

Figure 1

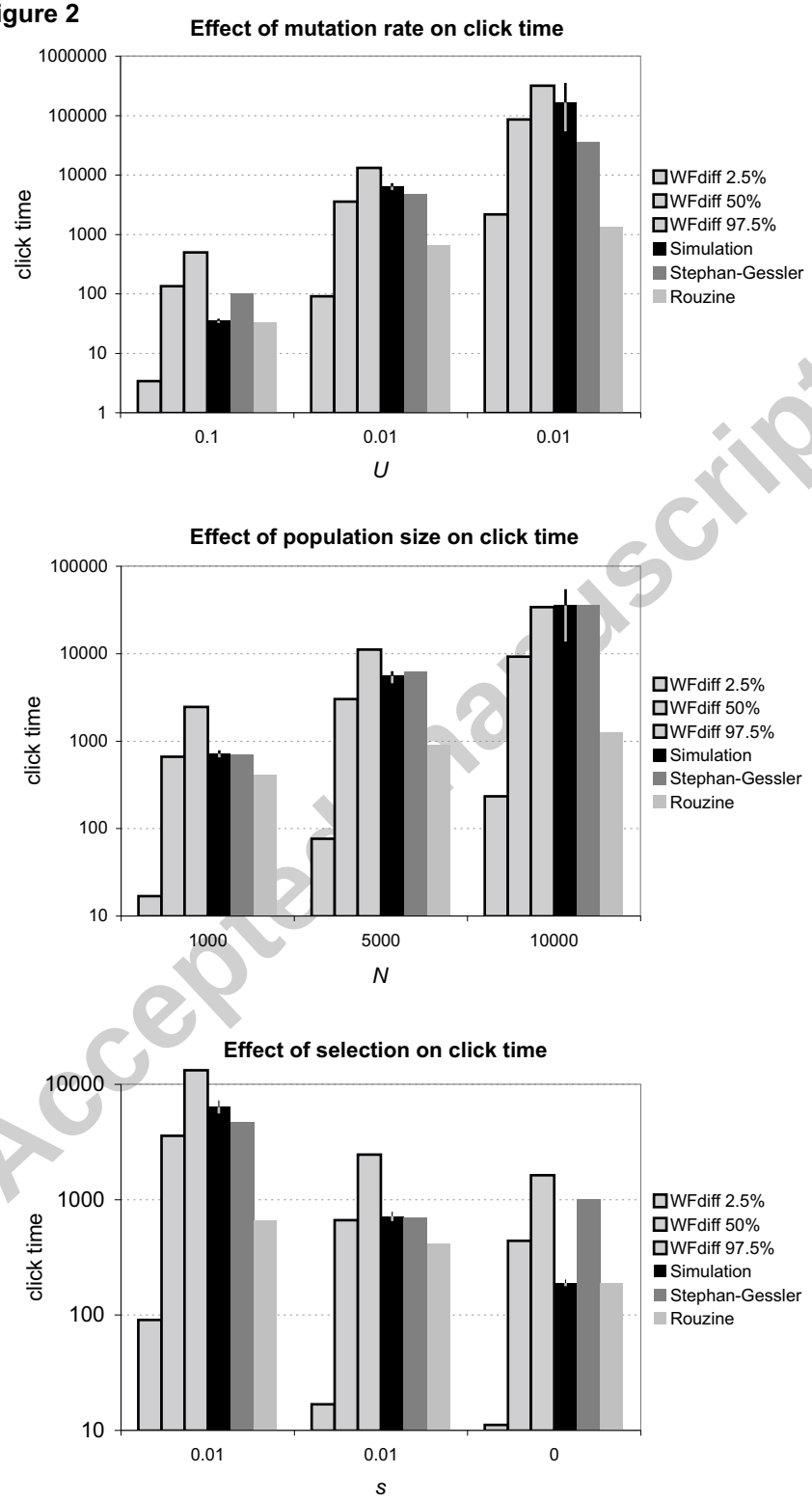


Accepted ma

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure 2



Accepted Manuscript

Table1

	Ratchet parameters			Wright Fisher				Diffusion			Error	Simulation							
data set	N	U	s	E[Tclick]	2.5% P	97.5% P	Trelax	E[Tclick]	2.5% P	97.5% P	EoM (Diff/Sim)	E[Tclick]	2.5% P	97.5% P	Expect E-2SE	Expect E+2SE	clicks/runs	E[Tclick] /SGmethod	E[Tclick] /Rmethod
1	100+	1	0.5	**567734**	14374	2094302	**4**	**1753961**	44406	6470151	3.51 ***	**546**	17	1033	251	840	757/55	**360/S**	12.67
2	100	1	0.1	**17**	0	63	**6**	**17**	0	63	0.75	**3**	2	4	2	4	4061/60	**3.3/G**	3.30
3	100	1	0.005	**17**	0	63	**6**	**17**	0	63	1.06 *	**1**	1	2	1	2	15360/40	**1.2/G**	1.31
4	100+	0.1	0.1	**526**	13	1940	**19**	**539**	14	1988	-0.08	**643**	575	737	538	747	1200/8	**472/S**	65.54
5	100+	0.1	0.05	**72**	2	266	**18**	**72**	2	266	0.04	**66**	53	74	56	76	8591/51	**70/S**	41.30
6	100	0.1	0.005	**46**	1	170	**16**	**46**	1	170	0.48	**15**	14	16	14	17	9811/24	**25/G**	15.51
7	100+	0.05	0.1	**16539**	419	61010	**19**	**18261**	462	67362	-0.02	**19290**	11240	29020	8942	29638	139/10	**3557/S**	133.14
8	100	0.05	0.01	**69**	2	255	**23**	**69**	2	255	0.23	**41**	36	45	36	45	13083/40	**33/G**	40.02
9	100	0.05	0.001	**66**	2	243	**23**	**66**	2	243	0.42	**25**	23	27	23	27	16689/33	**32/G**	23.86
10	1000+	0.1	0.05	**1226**	31	4523	**69**	**1138**	29	4198	-0.60	**4568**	3737	5342	3963	5173	40423/177	**3595/S**	126.73
11	1000	0.1	0.01	**135**	3	497	**52**	**135**	3	498	0.59	**35**	32	38	31	39	13785/43	**100/G**	33.04
12	1000	0.1	0.001	**135**	3	496	**52**	**135**	3	498	0.96	**15**	14	16	14	16	17757/33	**16/G**	14.72
13	1000+	0.01	0.01	**3695**	94	13630	**192**	**3587**	91	13234	-0.25	**6359**	5574	7249	5494	7224	39361/182	**4724/S**	655.42
14	1000+	0.01	0.005	**673**	17	2482	**174**	**665**	17	2454	-0.03	**717**	652	785	645	789	15779/40	**701/S**	413.04
15	1000	0.01	0.001	**442**	11	1632	**162**	**441**	11	1626	0.37	**189**	178	203	172	206	25054/50	**1000/G**	189.94
16	1000+	0.005	0.01	**89160**	2257	328901	**205**	**86342**	2186	318506	-0.28	**164300**	53960	353600	121960	206640	9035/150	**35567/S**	1331.35
17	1000+	0.005	0.001	**657**	17	2425	**231**	**654**	17	2414	0.19	**421**	395	449	381	462	17280/40	**1166/S**	400.17
18	1000	0.005	0.0001	**650**	16	2398	**230**	**647**	16	2388	0.42	**246**	228	267	223	269	12304/27	**435/G**	238.58
19	5000+	0.5	0.1	**230**	6	850	**64**	**226**	6	835	0.13	**166**	97	198	136	196	12238/101	**280/S**	32.77
20	5000	0.5	0.05	**145**	4	535	**57**	**146**	4	539	0.98	**15**	11	18	13	18	7622/48	**Err/G**	13.79
21	5000	0.5	0.01	**145**	4	534	**56**	**145**	4	537	1.49 *	**5**	4	5	4	5	14972/38	**5/G**	4.82
22	5000+	0.05	0.01	**456**	12	1680	**166**	**455**	12	1678	0.28	**240**	216	260	211	268	178368/647	**256/S**	144.60
23	5000	0.05	0.005	**415**	11	1532	**162**	**415**	11	1532	0.67	**89**	84	95	81	97	265516/577	**Err/G**	81.66
24	5000	0.05	0.001	**415**	11	1531	**162**	**415**	11	1531	1.04 *	**38**	36	40	35	41	64485/123	**50/G**	38.28
25	5000+	0.01	0.005	**3049**	77	11246	**439**	**3025**	77	11158	-0.26	**5565**	4581	6276	4858	6272	35959/145	**6145/S**	893.31
26	5000	0.01	0.001	**939**	24	3465	**360**	**939**	24	3463	0.51	**292**	273	308	265	318	72576/148	**Err/G**	280.67
27	5000	0.01	0.0005	**939**	24	3464	**360**	**938**	24	3462	0.65	**208**	195	222	190	227	48959/98	**500/G**	207.98
28	5000+	0.005	0.005	**192123**	4864	708719	**550**	**190172**	4815	701522	-0.64	**833700**	42530	1442000	369817	1297583	323/25	**239722/S**	2255.71
29	5000+	0.005	0.001	**1386**	35	5111	**514**	**1384**	35	5107	0.25	**786**	729	838	713	859	52112/112	**1275/S**	676.04
30	5000	0.005	0.0005	**1349**	34	4976	**510**	**1348**	34	4973	0.45	**483**	447	515	438	527	49394/104	**Err/G**	475.26
31	10000+	0.5	0.1	**411**	10	1516	**95**	**396**	10	1460	-0.31	**805**	448	1252	518	1091	20666/655	**1460/S**	42.85
32	10000	0.5	0.05	**205**	5	756	**80**	**207**	5	763	1.04 *	**19**	15	22	16	22	7603/48	**6.7/G**	16.17
33	10000	0.5	0.001	**204**	5	753	**80**	**206**	5	761	1.83 *	**3**	3	3	3	3	18291/32	**2.6/G**	2.82
34	10000+	0.01	0.005	**9393**	238	34650	**667**	**9249**	234	34117	-0.59	**35580**	13700	54440	28423	42737	4251/43	**35955/S**	1267.33
35	10000	0.01	0.001	**1313**	33	4843	**508**	**1312**	33	4840	0.57	**353**	332	370	321	385	78183/163	**333/G**	330.44
36	10000	0.01	0.0001	**1312**	33	4840	**508**	**1312**	33	4838	0.94	**150**	141	160	136	163	22340/44	**192/G**	147.24

ACCEPTED MANUSCRIPT