



HAL
open science

Effects of demographic stochasticity and life-history strategies on times and probabilities to fixation

Diasla Abu Awad, Camille Coron

► **To cite this version:**

Diasla Abu Awad, Camille Coron. Effects of demographic stochasticity and life-history strategies on times and probabilities to fixation. *Heredity*, 2018, 121 (4), pp.374 - 386. 10.1038/s41437-018-0118-6 . hal-02621517

HAL Id: hal-02621517

<https://hal.inrae.fr/hal-02621517>

Submitted on 5 Jan 2024

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.

1 Effects of demographic stochasticity and life-history strategies
2 on times and probabilities to fixation: an individual-based
3 model

4 Diala Abu Awad*, Camille Coron†

5 October 21, 2017

6 **Abstract**

7 Previous works has suggested that the harmonic mean population size can sum-
8 marize the consequences of demographic fluctuations on the genetic frequencies of
9 populations. We test this hypothesis by studying a model in which the demography
10 and genetic composition of the population are both determined by the behavior of
11 the individuals within the population. We propose an effective population size that
12 allows us to compare our model with the classical Wright-Fisher diffusion both for
13 neutral alleles and those under selection. We find that using our approximation for
14 the effective population size, the Wright-Fisher diffusion provides good results for the
15 times to absorption and probabilities of fixation of a given neutral allele and in cases
16 where selection is not too strong. However, the times and laws to fixation are not
17 always well predicted due to large fluctuations in population size caused by small
18 growth rates or strong competition between individuals, that cannot be captured by
19 the constant population size approximation. The discrepancy between our model and
20 the Wright-Fisher diffusion is accentuated in the presence of demo-genetic feed-back.
21 Our results imply that the Wright-Fisher diffusion is not appropriate when studying
22 probabilities and times to fixation in long-lived species with low reproductive rates.

23 **Keywords:** Individual-based models; demo-genetic feedback; demographic stochastic-
24 ity; life-history traits; fixation time; self-fertilization.

*INRA, UMR 1334 AGAP, Montpellier, France

†Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France

25 1 Introduction

26 Adaptive and non adaptive evolution is characterized by the dynamics of allele frequencies
27 and their eventual loss or fixation. For more than half a century, the diffusion limit of the
28 Wright-Fisher model ([15, 37]), introduced by [22, 23], has provided one of the key tools in
29 population genetics for predicting the dynamics of allelic frequencies. Due to simple and
30 strong analytical results obtained for this general model ([24]), it has been extended to
31 take into account populations with more general and complicated behaviors such as non-
32 random mating and structured populations (see for example [2, 1, 34]). The Wright-Fisher
33 model makes two simplifying assumptions: (1) all individuals reproduce and die at the
34 same time (discrete non-overlapping generations), and (2) population size is fixed, which
35 has led to the concept of “effective population size”, denoted N_e (and discussed below).
36 However, population size tends to vary stochastically, notably since births and deaths
37 can be independent events: reproduction by an individual is not necessarily immediately
38 followed by its death (see for instance [5]), and the speeds at which reproduction and death
39 occur representing different life-history strategies (*i.e.* r/K strategies). The ubiquity of
40 stochastic demographic phenomena, such as extinction, rapid expansions and bottlenecks
41 on a macroscopic scale, or independent births and deaths on a microscopic scale, requires
42 a better understanding of their interaction with allele frequency dynamics (and notably
43 with allele fixation).

44 In existing models studying allele dynamics, N_e is a central notion which aims at
45 bringing any population as “close” as possible (the definition of closeness being dependent
46 on the indicators of interest) to a classical Wright-Fisher diffusion. In particular for
47 populations with a deterministically varying population size, this parameter is defined
48 as the harmonic mean of the population size (as shown in [38, 25], for instance). In
49 the presence of selection, [31] explored the impact of macroscopic demographic events
50 (introduced by the use of a non-constant deterministic population size) on the probability
51 of fixation of alleles. They found that the harmonic mean sufficed in reflecting the change
52 in fixation probabilities of fluctuating populations as long as selection was not too strong.

53 On the other hand, [21, 20] showed that the harmonic mean size is sometimes an inadequate
54 definition of the effective population size when population size varies stochastically and
55 the authors proposed a new definition for N_e (the heterozygosity effective size). The
56 harmonic mean seems therefore insufficient in capturing the effects of stochastic events on
57 a more microscopic level even in models where the deaths and births of individuals are
58 not considered explicitly, the general effects of these processes being averaged to reflect
59 the behavior of the entire population.

60 Recently, individual-based models examining the interaction between population size
61 dynamics on the microscopic scale and probabilities of fixation have been developed ([5, 6,
62 32]). However, the feedback of genetics on demography is not considered nor modeled in
63 these diffusions, whereas it can have a major impact on population viability, notably when
64 selection parameters are not small, as can be observed in models of evolutionary rescue ([29,
65 18]), where this feedback is a central aspect. In [32], the authors explored the consequences
66 of different life-history strategies and proposed an individual-based model with “quasi-
67 neutral” selection so that the impact of population genetics on population demography can
68 be neglected and found that they could not define an appropriate N_e for which a classical
69 neutral Wright-Fisher diffusion would give the same mean time to absorption and fixation
70 probability as their model. Mean times to fixation of neutral alleles, and eventually
71 the distribution of these times, in the Wright-Fisher diffusion depend on the population’s
72 N_e ([25]) and are thus expected to be affected by a population’s demographic dynamics
73 (notably due to macroscopic events such as bottlenecks, expansions and extinctions, as
74 can be deduced from works on coalescent theory [19]). On the contrary, the fixation
75 probability of a neutral allele is always expected to be equal to its initial frequency. That
76 [32]’s results for quasi-neutrality are better described by a Wright-Fisher diffusion with
77 selection (Figure 4 in their paper) thus raises three questions: *i*) How should fitness be
78 defined in individual-based models in order to render them, if possible, comparable to a
79 Wright-Fisher framework? *ii*) What role do life-history strategies play in the probabilities
80 and times to fixation? and *iii*) If genotypes under selection present different demographic

81 behaviors (*i.e.* growth rate), how is the ensuing change in population size likely to influence
82 the probabilities and times to fixation?

83 In this article we propose an individual-based model in order to study the absorption
84 times and fixation probabilities in a demo-genetic context, which we then compare to a
85 Wright-Fisher diffusion. In this probabilistic model both the demography and genetics
86 of a given population are defined through the dynamics of each individual within the
87 population. The behavior of each individual is stochastic, and dependent on demographic
88 parameters that can be estimated ([27]). More precisely, we consider a population of
89 diploid individuals experiencing weak selection at a single bi-allelic locus. As population
90 size is directly determined by frequent birth and death events, it changes stochastically
91 with time, and can also depend on the population's genetic composition. The originality
92 of our approach and model lies in four main features: (1) We consider linked stochastic
93 dynamics of both the population size and its genetic composition. (2) The life-history
94 strategy of a population is a natural behavior of the model and depends directly on
95 the demographic parameters (as in *e.g.* [32]), being in no way forced. (3) Extinction
96 occurs in finite time, which notably impacts fixation times. (4) We consider a sexually
97 reproducing diploid population, with general dominance relationships between alleles, and
98 possibility of self-fertilization (previous models considered haploid individuals, [5, 6, 32]).
99 The obtained model can also be seen as a generalization of the Wright-Fisher diffusion,
100 since this diffusion can be obtained when letting some parameters of the model (namely the
101 growth and competition rates) go to $+\infty$. We compare the laws of the time to absorption
102 (fixation or loss of a given allele) and the probability of fixation for our model to the
103 classical Wright-Fisher diffusion (presented in [2] for populations with self-fertilization).
104 These results are obtained by simulating trajectories of diffusion processes, and we find
105 notably that

106 (i) There are parameter sets for which the laws of the time to absorption for our demo-
107 genetic model and for the classical population genetics model (Wright-Fisher dif-
108 fusion calibrated with an appropriate effective size) are very close, notably when

109 there is a high population growth rate and high death rate (due to competition for
110 resources). Population genetics models therefore provide very good predictions for
111 species with r -strategies (high reproductive output and short life-span).

112 (ii) Laws of time to absorption can be very different when taking into account population
113 size variability, notably in rapid expansion and diminution contexts (Section 3.3),
114 as well as when population size fluctuations are highly stochastic, which is the case
115 for populations with low reproductive rates and low death rates (K -strategies). In
116 particular, we find that due to the fluctuations in population size, the frequencies of
117 small and large absorption times of rare alleles are underestimated in the classical
118 Wright-Fisher diffusion model (*i.e.* there is a greater variance in times to absorption
119 than predicted by a fixed population size).

120 (iii) The demographic consequences of taking the feedback of genetics on demography
121 can impact the probabilities and times to fixation in a way that can not be fully
122 captured by the proposed effective population size if population growth rates are
123 low.

124 **2 Model**

125 We consider a population of diploid individuals, characterized by their genotype at a sin-
126 gle bi-allelic locus with alleles A and a . The population is modeled by a 3-dimensional
127 stochastic birth-and-death process (detailed below) giving the respective numbers of in-
128 dividuals with genotype AA , Aa and aa . Contrary to previous models where population
129 size is a parameter, here it is a random variable. The dynamics of population size are
130 stochastic, and population extinction occurs with probability 1. Below we detail the re-
131 scaled diffusion approximation, highlighting the main differences between our model and
132 the diffusion approximation proposed by [24].

133 2.1 Rescaled birth-and-death process

134 The Wright-Fisher diffusion is obtained by considering the dynamics of the proportion
 135 of a given allele when re-scaling a discrete time population model with constant effective
 136 population size N_e and non-overlapping generations. In this article, population dynamics
 137 are determined by individual-based demographic parameters, therefore inducing variable
 138 population size. We introduce a scaling parameter $K \in \{1, 2, \dots\}$ that will go to infinity
 139 (as in [16, 7, 9, 10]), modeling an infinite size approximation. The population is made up
 140 of three types of individuals (AA , Aa and aa , represented by 1, 2 and 3 respectively), the
 141 number of individuals of each type being of order K . At each time t the population is
 142 represented by a vector

$$(\mathbf{Z}_t^K)_{t \geq 0} = (Z_t^{1,K}, Z_t^{2,K}, Z_t^{3,K})_{t \geq 0}$$

143
 144 which gives the respective number of individuals of each type, divided by K . If the pop-
 145 ulation is at a state $\mathbf{z} = (z_1, z_2, z_3)$, the birth rates $\lambda_i^K(\mathbf{z})$ for all $i \in \{1, 2, 3\}$ model sexual
 146 Mendelian reproduction either by self-fertilization (with probability α) or by random mat-
 147 ing (with probability $1 - \alpha$).

$$\begin{aligned} \lambda_1^K(\mathbf{z}) &= Kb_1^K \left[\alpha \left(z_1 + \frac{z_2}{4} \right) + (1 - \alpha) \frac{(z_1 + \frac{z_2}{2})^2}{n} \right], \\ \lambda_2^K(\mathbf{z}) &= Kb_2^K \left[\alpha \frac{z_2}{2} + (1 - \alpha) 2 \frac{(z_1 + \frac{z_2}{2})(z_3 + \frac{z_2}{2})}{n} \right], \\ \lambda_3^K(\mathbf{z}) &= Kb_3^K \left[\alpha \left(z_3 + \frac{z_2}{4} \right) + (1 - \alpha) \frac{(z_3 + \frac{z_2}{2})^2}{n} \right]. \end{aligned}$$

148

149

150 with $n = z_1 + z_2 + z_3 \neq 0$. These birth rates are naturally set to 0 when $n = 0$. Note that
 151 the parameters b_i^K that model the viability (or recruitment) of new-born individuals can
 152 depend on i , which allows for some selection at birth (as will be shown below). Individual

153 mortality can be natural or due to competition with other individuals (therefore allowing
 154 for density-dependence and limiting population size). Here we assume that death rates do
 155 not depend on genotypes, in order to focus on a small number of parameters (but see [11]
 156 for a more general model). If the population is at a state $\mathbf{z} = (z_1, z_2, z_3)$, the rate $\mu_i^K(\mathbf{z})$
 157 at which an individual with genotype i dies in the population is then given by:

$$\begin{aligned}\mu_1^K(\mathbf{z}) &= K z_1 (d^K + K(c^K z_1 + c^K z_2 + c^K z_3)), \\ \mu_2^K(\mathbf{z}) &= K z_2 (d^K + K(c^K z_1 + c^K z_2 + c^K z_3)), \\ \mu_3^K(\mathbf{z}) &= K z_3 (d^K + K(c^K z_1 + c^K z_2 + c^K z_3)).\end{aligned}$$

158

159 The demographic parameter d^K (resp. $c^K > 0$) is the intrinsic death rate (resp. the
 160 competition rate) of individuals. Population size is therefore regulated by competition,
 161 *i.e.* by density-dependence.

162 The demographic parameters b^K , d^K and c^K are scaled both by K and a parameter
 163 γ , the latter scaling the speed with which births and deaths occur, giving:

$$\begin{aligned}b_1^K &= \gamma K + \rho, \\ b_2^K &= \gamma K + \rho + h\sigma, \\ b_3^K &= \gamma K + \rho + \sigma,\end{aligned}\tag{1}$$

164 and

$$d^K = \gamma K \quad \text{and} \quad c^K = \frac{\xi}{K}.$$

165 The parameters σ and h are respectively the selection and dominance coefficients of
 166 allele a , and ρ is the population growth rate in the absence of selection.
 167 Note that in this model, we do not directly consider population size (number of individ-
 168 uals), but population mass, defined as $\mathcal{N}_t^K = Z_t^{1,K} + Z_t^{2,K} + Z_t^{3,K}$. This scaling therefore
 169 models a population with numerous and small individuals, each represented by a mass of

170 $1/K$, that reproduce frequently in such a way that both the total population mass and al-
 171 lele proportions will not be constant (and will evolve stochastically) even when the scaling
 172 parameter K goes to infinity. This is the same scaling used to obtain the Wright-Fisher
 173 diffusion process from the Wright-Fisher model, however our initial model (birth-and-
 174 death process) allows for stochastic dynamics of population mass. When K is large, the
 175 selection parameter σ has an inherent weak impact on the birth parameters b_i^K , but due
 176 to first-order compensation between birth and death events (both of order K), its impact
 177 on the growth rate is macroscopic. Therefore, it will still have an effect on the limiting
 178 population dynamics (notably by either increasing or decreasing the expected population
 179 mass, see next section).

180 2.2 Extended Hardy-Weinberg structure and limiting diffusion process

181 Let us set for all $K \geq 1$ and all $t \geq 0$,

$$Y_t^K = \alpha \frac{Z_t^{2,K}}{4} - (1 - \alpha) \frac{Z_t^{1,K} Z_t^{3,K} - (Z_t^{2,K}/2)^2}{Z_t^{1,K} + Z_t^{2,K} + Z_t^{3,K}}.$$

182 Note that in a pure random mating context ($\alpha = 0$), and if the quantity $Y_t^K = 0$, then the
 183 proportion of each genotype in the population is equal to the proportion of pairs of alleles
 184 forming this genotype, which means that the population satisfies the Hardy-Weinberg
 185 structure. More generally, Y_t^K quantifies the deviation of the population at time t from a
 186 generalized Hardy-Weinberg structure. Indeed, straightforward calculations show that, as
 187 in population genetics theory ([17], pp. 91-93), if $Y_t^K = 0$ then

$$\begin{aligned} Z_t^{1,K} &= \mathcal{N}_t^K [(1 - X_t^K)^2(1 - F) + (1 - X_t^K)F], \\ Z_t^{2,K} &= 2\mathcal{N}_t^K (1 - X_t^K)X_t^K (1 - F), \\ Z_t^{3,K} &= \mathcal{N}_t^K [(X_t^K)^2(1 - F) + X_t^K F], \end{aligned}$$

188 where X_t^K is the proportion of allele a in the population, and the coefficient of in-
 189 breeding $F = \frac{\alpha}{2-\alpha}$. We can prove following [10] that Y_t^K converges to 0 when K goes
 190 to infinity, for all t . The limiting population dynamics can then be represented at time t

191 by the couple (\mathcal{N}_t^K, X_t^K) giving the population mass and the proportion of allele a . The
 192 population process $(\mathcal{N}_t^K, X_t^K)_{t \geq 0}$ thus converges toward a bi-dimensional diffusion process
 193 $(\mathcal{N}_t, X_t)_{t \geq 0}$ whose equation can be written as:

$$d\mathcal{N}_t = \sqrt{2\gamma\mathcal{N}_t}dB_t^1 + \mathcal{N}_t \left[\rho - \xi\mathcal{N}_t + \sigma X_t \left(2h + X_t(1 - 2h) + F(1 - X_t)(1 - 2h) \right) \right] dt, \quad (2a)$$

$$dX_t = \sqrt{\frac{2\gamma X_t(1 - X_t)}{2\frac{\mathcal{N}_t}{1+F}}} dB_t^2 + \sigma X_t(1 - X_t) \left[h + X_t(1 - 2h) + F(1 - X_t - h + 2X_t h) \right] dt. \quad (2b)$$

194 where $(B_t^1, B_t^2)_{t \geq 0}$ is a bi-dimensional standard Brownian motion (stochastic component
 195 of the equation). This diffusion model can be generalized without difficulty to any finite
 196 number of alleles, as presented in [12]. Note that, without loss of generality, we can assume
 197 that the time scaling parameter γ is equal to 1/2, thus simplifying the above equations.
 198 In this case, if the stochastic quantity $\frac{\mathcal{N}_t}{1+F}$ is artificially replaced by a fixed parameter
 199 N_e , then the model given in (2b) is the Wright-Fisher diffusion with selection and self-
 200 fertilization presented in [2], where the parameter σ in our model is equal to the coefficient
 201 of selection s of [2] and N_e is the effective population mass.

202 More interestingly, this classical Wright-Fisher diffusion with selection and self-fertilization
 203 can also be directly retrieved from our model, by setting $\rho/\xi = N_e$ and letting ρ got to
 204 $+\infty$. In order to determine whether a constant effective population mass can summarize
 205 the effects of a stochastic population mass as proposed in earlier models [24, 31], in Section
 206 33.2 we define a fixed effective population mass N_e in such a way that the model in [2] is
 207 adequately calibrated.

208 **2.3 Simulating the diffusion process**

209 In [2], the authors provide explicit formulas for the probabilities of fixation as well as
 210 approximations for the times to loss or fixation of an allele. Due to the bi-dimensionality

211 of our model which largely increases the difficulty of mathematical calculations, fixation
212 probabilities as well as laws of times to fixation, loss and absorption (either loss or fixation)
213 of allele a are determined using simulations of equations (2a) and (2b). These simulations
214 are run using a script written in C++ (and available on Dryad). The stochastic elements of
215 the equations, B_t^1 and B_t^2 are obtained by successive samplings from a normal distribution
216 with mean 0 and variance dt . dt is the size of the time step and is a parameter fixed at the
217 beginning of the simulation, which we have set to 10^{-4} for a carrying capacity $\mathcal{K} = 100$ and
218 and 10^{-5} for $\mathcal{K} = 10$ and 1. Each simulation is run until the allele a is either lost or fixed
219 and 100 thousand replicas are run for each parameter set (unless otherwise mentioned)
220 from which the probability of fixation, as well as the means and laws of times to fixation,
221 loss and absorption are obtained.

222 In order to test whether deviations in times to loss or fixation from the approximations
223 provided in [2] are due to demographic stochasticity or due to the approximations made, we
224 run simulations of the Wright-Fisher Diffusion (using a fixed population mass N_e defined
225 in Section 3.2). We also run simulations to assess the effects of the feed-back between
226 selection and demography by artificially setting $\sigma = 0$ in equation (2a) only. In order to
227 evaluate the effect of the change in population size due to the fixation of an allele under
228 selection with an effect σ , we also consider the case where the carrying capacity is equal
229 to $(\rho + \sigma)/\xi$ (see Section 3.4).

230 **3 Analytical and numerical results**

231 **3.1 Demography**

232 The change in population mass given in Equation (2a) is made up of a stochastic term
233 (dependent on dB_t^1) and a deterministic one (dependent on dt). In this diffusion model
234 with selection and self-fertilization, the probability of extinction is equal to 1. The law of
235 the time to extinction depends on the ecological and genetic parameters. In the neutral
236 case where $\sigma = 0$, Equation (2) can be simplified the following way:

$$d\mathcal{N}_t = \sqrt{\mathcal{N}_t} dB_t^1 + \mathcal{N}_t [\rho - \xi \mathcal{N}_t] dt, \quad (3a)$$

$$dX_t = \sqrt{\frac{X_t(1-X_t)}{\frac{2\mathcal{N}_t}{1+F}}} dB_t^2. \quad (3b)$$

237 Here population mass is independent of its genetic composition and the deterministic
238 term of Equation (3a) cancels out when $\mathcal{N}_t = \mathcal{K}$ where

$$\mathcal{K} = \frac{\rho}{\xi} \quad (4)$$

239 is defined as the population's carrying capacity. Note that \mathcal{K} does not represent the
240 number of individuals that can be sustained in the population (since \mathcal{N}_t is scaled by K
241 which goes to infinity) but is an indicator of the amplitude of demographic stochasticity,
242 as will be shown below. When population mass is smaller (resp. larger) than \mathcal{K} , it will
243 tend to increase (resp. decrease). For a fixed value of \mathcal{K} , if ρ is large, then the population
244 mass will remain close to \mathcal{K} , whereas for small values of ρ the mass will tend to deviate
245 further from \mathcal{K} (see Figure 1). The smaller ρ the slower the population mass will come
246 back to its pseudo equilibrium \mathcal{K} ; therefore a small value of ρ can have an important
247 impact on extinction, as can be seen in Figure 1 (black lines). The role of \mathcal{K} on population
248 mass dynamics is not as straightforward since \mathcal{N}_t is implicated in both the stochastic and
249 deterministic terms (therefore both terms are increased when \mathcal{K} increases). In Figure 1
250 we also see that the effect of ρ on demographic stochasticity is weaker when \mathcal{K} is smaller.

251 **3.2 Effective population mass**

252 In the neutral case (Equation (3)), variations in population mass are modeled by a logistic
253 diffusion process (and thus are independent from the genetic composition of the popula-
254 tion) and changes in allele frequency by a Wright-Fisher diffusion with population mass
255 \mathcal{N}_t at any time t . Hence, it is natural to compare this model to the neutral Wright-Fisher
256 diffusion model of population genetics, for which the proportion X_t^{WF} of a neutral allele

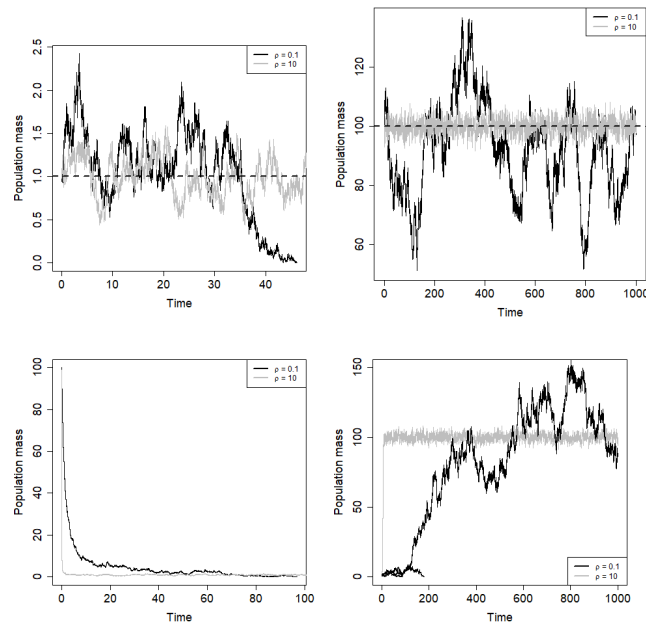


Figure 1: Top: Trajectories of the population mass ($N_t, t \geq 0$), for $N_0 = \mathcal{K}$ and $\mathcal{K} = \rho/\xi = 1$ (left), $\mathcal{K} = \rho/\xi = 100$ (right), for $\rho = 0.1$ (black) and $\rho = 10$ (grey). Bottom: Trajectories of the population mass ($N_t, t \geq 0$), for $\mathcal{K} = \rho/\xi = 1$ and $N_0 = 100$ (left), and $\mathcal{K} = \rho/\xi = 100$ and $N_0 = 1$ (right), for $\rho = 0.1$ (black) and $\rho = 10$ (grey). For $N_0 = 1$, $\rho = 0.1$ and $\mathcal{K} = 100$ (bottom-right figure), we plot 3 trajectories.

257 at all time satisfies

$$dX_t^{WF} = \sqrt{\frac{X_t^{WF}(1 - X_t^{WF})}{2N_e}} dB_t. \quad (5)$$

258 Here N_e represents the effective population mass of a self-fertilizing population (as de-
 259 scribed in [2]) and is a parameter of the Wright-Fisher diffusion model. The parameters
 260 ρ , ξ and the inbreeding coefficient F being fixed in our model, we define a fixed effective
 261 population mass N_e that allows us to compare our model with variable population mass
 262 to a Wright-Fisher diffusion.

263 In order to calibrate N_e appropriately, it is not enough for the probability of fixation to
 264 be the same in both models, as in the neutral case the fixation probability of an allele a
 265 is simply equal to its initial proportion. Therefore, we choose to calibrate N_e such that
 266 the mean absorption time (mean time to fixation of one of the two alleles) is the same in

267 both models. From Appendix A, N_e is defined as:

$$N_e = \frac{\mathbb{E}(T_{abs})}{2(1+F)\mathbb{E}\left[\int_0^{T_{abs}} \frac{1}{N_t} dt\right]}, \quad (6)$$

268 where $\mathbb{E}(V)$ represents the expectation of a stochastic variable V and T_{abs} the random
269 absorption time of the population modeled by Equation (3). Note that $\frac{\mathbb{E}(T_{abs})}{\mathbb{E}\left[\int_0^{T_{abs}} \frac{1}{N_t} dt\right]}$ is
270 not the expectation of the empirical harmonic mean of the mass till absorption, which
271 is $\mathbb{E}\left(\frac{T_{abs}}{\int_0^{T_{abs}} \frac{1}{N_t} dt}\right)$, but the ratio of two expectations (the difference between the two is
272 shown in Figure A.1, and is very important for highly fluctuating population mass). Note
273 also that with this definition, the effective population mass N_e depends on the initial
274 frequency X_0 of allele a ; this dependence is illustrated in Figure A.1. We obtain numerical
275 estimations of the quantity N_e from the simulation runs of Equation (2) with varying
276 population mass, calculating $\mathbb{E}(T_{abs})$ and $\mathbb{E}\left[\int_0^{T_{abs}} \frac{1}{N_t} dt\right]$ using all repetitions run for each
277 parameter set. In Figure 2 (left) we plot the mean times to absorption as a function of the
278 initial proportion of allele a , and for different values of ρ . This mean time to absorption
279 is given for our model with varying population mass, for the Wright-Fisher diffusion (5)
280 using the effective population mass N_e given in Equation (6), as well the theoretical result
281 provided in Equations (12) and (13) from [2]. Figure 2 therefore shows that the models
282 are indeed correctly calibrated for different values of parameters ρ , ξ and X_0 (for different
283 population densities and the effect of the inbreeding coefficient F see Figure A.2).

284 3.3 Neutral case: absorption and fixation times laws

285 Despite equal mean absorption times, the distributions of the times to absorption differ
286 between our model with stochastically varying population mass and the simulation runs
287 of the Wright-Fisher diffusion, notably when the parameters ρ and \mathcal{K} are small. This
288 is illustrated by Figure 3 in which we compare the variance of the time to absorption
289 for our demogenetic model and for the Wright-Fisher diffusion (see also Supplementary
290 Figure 1, in which the laws of these times to absorption are given), and this can be
291 understood by decomposing the absorption time into the time to loss or time to fixation

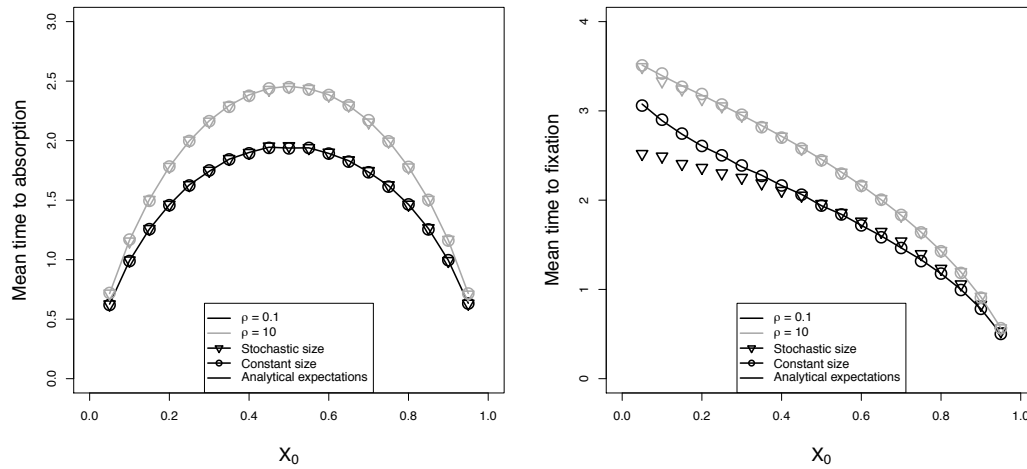


Figure 2: Mean times to absorption (left) and fixation (right) of a neutral allele ($\sigma = 0$) as a function of the initial frequency X_0 of allele a , for three cases: 1) Simulations of the stochastic diffusion process (2) (squares), 2) Simulations of the Wright-Fisher diffusion using N_e defined in Equation (6) (circles) and 3) Theoretical approximations provided by [2] using N_e (triangles). Here we considered pure random mating ($\alpha = 0$), the carrying capacity $\mathcal{K} = 1$ and the growth rate ρ equals 0.1 (black) or 10 (grey).

292 of an allele at initial frequency X_0 . Indeed we find that mean fixation times of minority
 293 alleles are lower for the model with stochastically varying population mass (Figure 2
 294 (right) and Supplementary Figure 2). This discrepancy between the results with varying
 295 and fixed sizes can be explained by the incidence of bottlenecks and extinction events,
 296 which is further accentuated by a small value of ρ . This is because a low growth rate
 297 results in a weaker impact of the deterministic forces regulating population mass (Equation
 298 (3)), further increasing demographic stochasticity. Indeed, large demographic fluctuations
 299 eventually lead to reduced population mass harmonic means, for which absorption is more
 300 rapid and fixation of minority alleles is favored (Figure 4).

301 As seen in Section 3.1, we can also consider that population mass changes drastically
 302 with time, allowing us to modeling founder effects, or drastic changes in the environment
 303 for instance. As previously, we compare the laws of the absorption time in populations with
 304 rapidly decreasing or increasing mass. Population mass trajectories are given in Section
 305 3.1 (Figure 1 (bottom)), and we start with a proportion $X = 0.1$ of a neutral allele a .
 306 We obtain that the laws of the absorption and fixation times are very different when

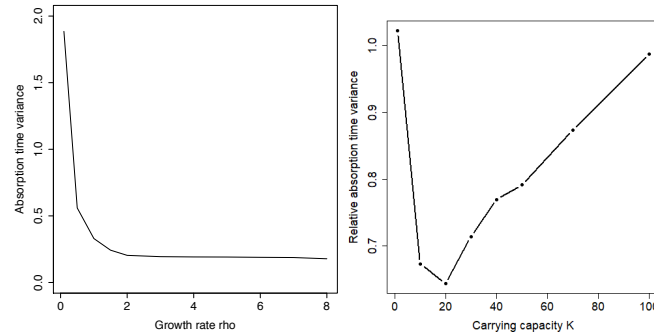


Figure 3: Variance of the absorption time in our demogenetic model and for the Wright-Fisher diffusion model, as a function of growth parameter ρ (left), and carrying capacity \mathcal{K} (right). On the left $\mathcal{K} = 1$ while on the right $\rho = 0.1$.

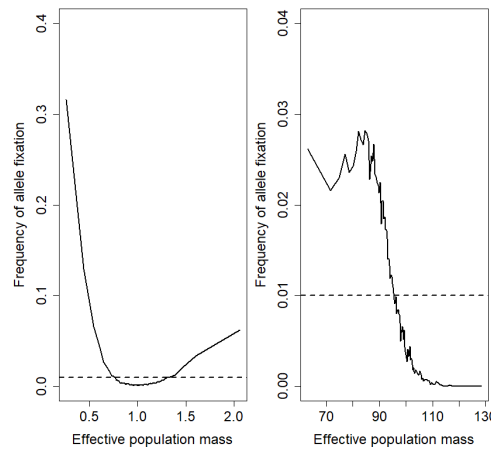


Figure 4: Fixation probability of a rare neutral allele, as a function of effective population mass. We set $X_0 = 0.01$ and $\rho = 0.1$. On the left, $\mathcal{K} = 1$ ($\xi = 0.1$), while on the right $\mathcal{K} = 100$ ($\xi = 0.001$).

307 comparing our to the Wright-Fisher diffusion model, despite the same mean absorption
 308 times (Figure 5). In particular, when population mass is kept constant, the frequency of
 309 small (and relatively large) absorption times is underestimated when the population mass
 310 increases, whereas the opposite is true when the population mass decreases.

311 3.4 Selection, demography and genetic feedback

312 In this section we introduce selection through the parameter σ in Equation (2). As men-
 313 tioned in Section 2.2, when comparing (2b) which describes the dynamics of allelic frequen-
 314 cies, to the Wright-Fisher diffusion, we find that σ has the same effect on allelic frequencies

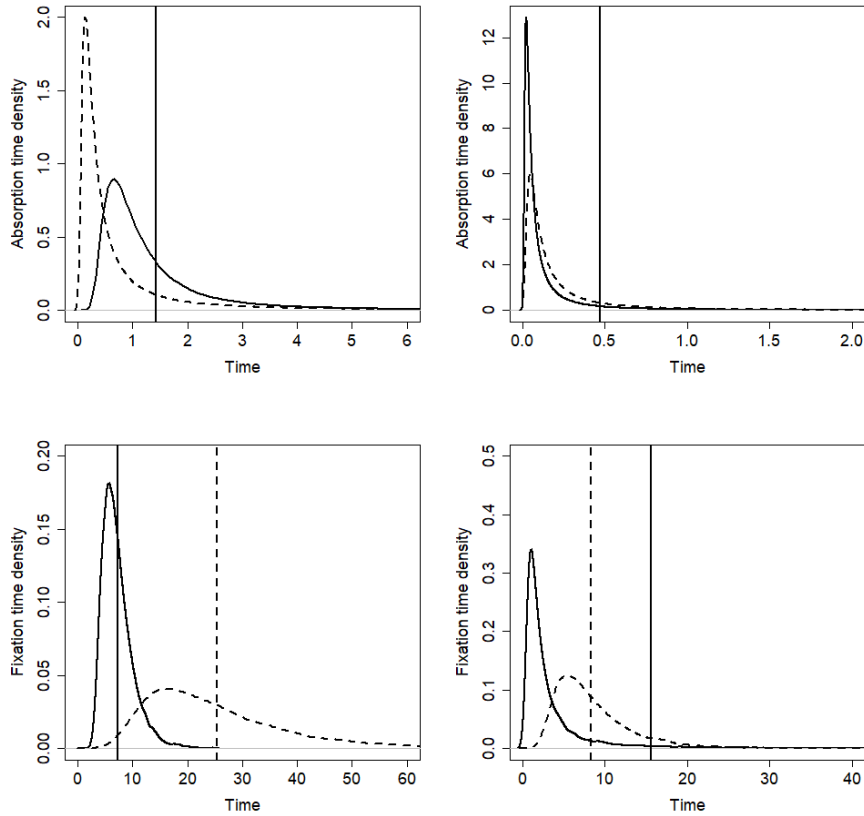


Figure 5: Absorption (top) and fixation (bottom) time density of a neutral allele with initial frequency $X_0 = 0.01$ for our model and for the classical Wright-Fisher model (dotted line). On the left (decreasing population mass), we fix $N_0 = 100$ and $\mathcal{K} = \rho/\xi = 1$, while on the right (increasing population mass), we fix $N_0 = 1$ and $\mathcal{K} = \rho/\xi = 100$, with $\rho = 0.1$.

315 as the conventionally used coefficient of selection s . The presence of σ in Equation (2a)
 316 implies that population mass and the proportion of allele a are linked through the dynam-
 317 ics of individuals that are present in the population. It is important to note that, from
 318 Equation (1), selection is in fact weak and has a negligible impact on individual birth rates
 319 (whatever value of the selection parameter $\sigma \in \mathbb{R}$). However, the proportion of a given
 320 non-neutral allele can have an important impact on the population mass dynamics. The
 321 consequences of this interaction can be quantified by the ratio σ/ξ , which is the change in
 322 the carrying capacity \mathcal{K} when the allele under selection a is fixed. Therefore, for a same
 323 \mathcal{K} before fixation but different values of ρ , similar values of σ can lead to very different

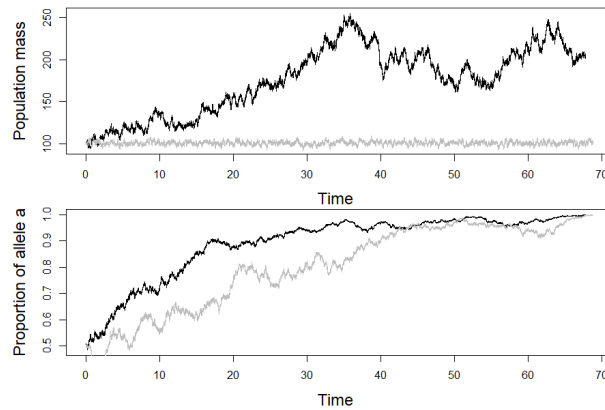


Figure 6: Population mass and proportion of allele a dynamics, for $\sigma = 0.1$, $\mathcal{K} = 100$, with $\rho = 0.1$ (black, $\xi = \rho/\mathcal{K} = 0.001$) and $\rho = 10$ (gray, $\xi = \rho/\mathcal{K} = 0.1$).

324 population mass dynamics (see Figure 6 with selection for a beneficial allele).

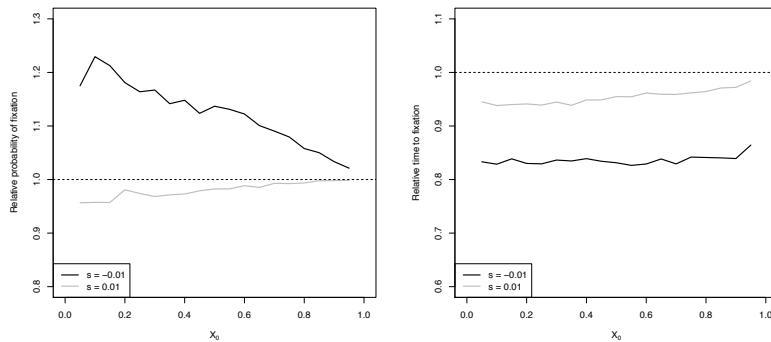


Figure 7: Relative probability of fixation (left) and relative time to fixation (right) for low growth rate ($\rho = 0.1$) compared to high growth rate ($\rho = 10$) as a function of the initial frequency X_0 of allele a with $\mathcal{K} = 100$, $h = 0.25$ and $\alpha = 0$ for $s = 0.01$ and -0.01 .

325 Due to the differences in population dynamics, probabilities and times to fixation can
 326 be affected by the growth rate, even for small values of s (Figure 7). Lower ρ results
 327 in higher probabilities of fixation of deleterious alleles, and lower relative probabilities of
 328 fixing beneficial alleles. Furthermore, times to fixation are generally lower for populations
 329 with low growth rates, independently of the coefficient of selection.

330 In order to understand and quantify the consequences of feedback of genetics on de-
 331 mography, it is natural to artificially remove all terms dependent on σ in Equation (2a),

332 hence removing any impact of changes in proportion on the dynamics of population mass.
 333 More precisely, let us for simplicity assume that $F = 0$, $h = 1/2$, and let us consider the
 334 following diffusion process $(\mathcal{N}_t^{(NF)}, X_t^{(NF)})_{t \geq 0}$ ("NF" standing for "No Feed-back"):

$$d\mathcal{N}_t^{(NF)} = \sqrt{2\gamma\mathcal{N}_t^{(NF)}} dB_t^1 + \mathcal{N}_t^{(NF)} [\rho - \xi\mathcal{N}_t^{(NF)}] dt, \quad (7)$$

$$dX_t^{(NF)} = \sqrt{\frac{2\gamma X_t^{(NF)}(1 - X_t^{(NF)})}{2\frac{\mathcal{N}_t^{(NF)}}{1+F}}} dB_t^2 + \frac{\sigma}{2} X_t^{(NF)}(1 - X_t^{(NF)}) dt. \quad (8)$$

335 For this model without feedback, we obtain that it is possible to calibrate a Wright-
 336 Fisher diffusion with selection, using Equation (6) with $\mathcal{N}_t = \mathcal{N}_t^{(NF)}$, so that the mean
 337 time to absorption and the probability of fixation are the same in both models (Figure 8).
 338 In the presence of feed-back (Equation (2)), though we generally find that for large \mathcal{K} , large
 339 ρ and/or weak selection, the proposed N_e (Equation (5)) provides a good approximation
 340 for the demographic effects on the times and probabilities of fixation, this is not the case
 341 for small values of ρ and/or \mathcal{K} . Indeed, when ρ is small there can be some discrepancies
 342 between the probability of fixation predicted by our model with feed-back and a population
 343 with constant size N_e when selection is intermediate. This can be seen in Figure 8 for
 344 $s = 0.1$ where our model with feed-back predicts a probability of up to 10% lower than
 345 the population with constant size N_e for low initial frequencies of the allele a . This
 346 difference is even greater for deleterious alleles with $s = -0.1$ (but for intermediate initial
 347 frequencies), simultaneously due to the stochastic nature of population mass and to feed-
 348 back which further contributes to decreasing the population mass in this case (Figure
 349 8). Times to fixation however are well predicted, with generally the model with feed-
 350 back being either closer to the model without feed-back and $\mathcal{K} = \rho + \sigma/\xi$ and $\mathcal{K} = \rho/\xi$
 351 depending on the initial frequency of the allele X_0 . We can see from the densities of times
 352 to absorption, fixation and loss (Supplementary Figure 3), that the laws of the times to
 353 fixation are very well captured using the constant mass model, though the times to loss
 354 are slightly underestimated. Times to fixation of a mildly deleterious allele are however

355 slightly underestimated by the simulations run with constant mass and are closer to the
 356 times to fixation of the simulations run without feed-back and $\mathcal{K} = \rho/\xi$.

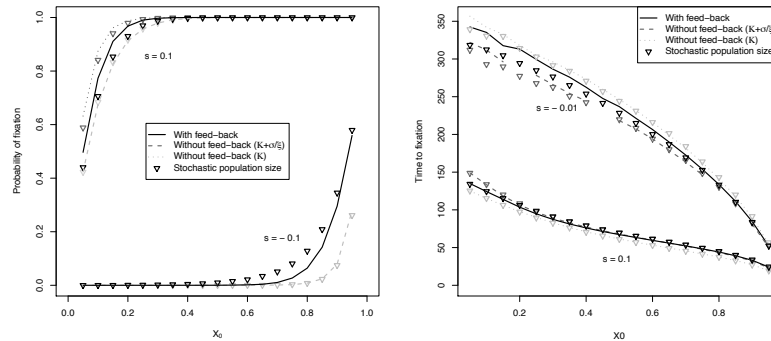


Figure 8: Left: Probability of fixation as a function of the initial frequency X_0 of allele a with $\mathcal{K} = 100$, $\rho = 0.1$ for $s = 0.1$ and -0.1 from simulations with demo-genetic feed-back (black full lines) and without demo-genetic feed-back (for $\mathcal{K} = \rho/\xi$ (dotted lines) and $\mathcal{K} = \rho + \sigma/\xi$ (dashed lines, not shown for $s = -0.1$ as $\mathcal{K} = 0$ in this case), and the corresponding results of simulations run with constant mass using the corresponding N_e . Other parameter values : $h = 0.25$ and $\alpha = 0$. Right: Time to fixation of an allele under selection as a function of its initial frequency, same parameters as the figure on the left but for $s = 0.1$ and -0.01

357 Concerning the effect of the self-fertilization rate (which are summarized in Supple-
 358 mentary Figure 4) we find that as expected from the Wright Fisher diffusion, probabilities
 359 of fixation of beneficial (respectively deleterious) alleles increase (respectively decrease)
 360 with the rate of self-fertilization α . We also find as previously predicted that the times to
 361 fixation decrease with increasing α . In all other aspects we find the same patterns as for
 362 the case without self-fertilization ($\alpha = 0$).

363 4 Discussion

364 An interesting feature of our model is that it is individual-based, in the sense that the
 365 model is characterized by simple demographic parameters that define the behavior of
 366 individuals within the population. Using these demographic parameters we are able to
 367 calculate an effective population mass N_e that allows us to predict the probabilities of
 368 fixation, as well as the times to absorption, using a Wright-Fisher diffusion and specify for

369 which parameter sets this N_e is appropriate. We generally find that for populations with
370 long-term fluctuations, induced by their intrinsic demographic parameters, the proposed
371 N_e does not fully capture the laws of times to fixation, with rare neutral alleles being
372 more frequently fixed in shorter times. We also show that, contrary to expectations,
373 despite a probability of fixation of a neutral allele being equal to its initial frequency,
374 when examining each repetition for a given parameter set separately, there is a higher
375 frequency of fixation of rare neutral alleles for populations that maintain low harmonic
376 mean masses. This result further highlights the importance of integrating demographic
377 parameters in population genetics models.

378 4.1 Interpreting demographic parameters

379 In our model the term ρ defines the speed at which individuals reproduce (hence population
380 growth) and ξ represents the competition for resources that in turn regulates population
381 mass (due to increased mortality). Thus, for a given expected population mass $\mathcal{K} = \rho/\xi$
382 a low ρ describes long-lived individuals with low death rates, whereas a high ρ describes
383 short-lived individuals with high death rates (rapid turnover). When comparing the de-
384 mographic fluctuations of two populations with different values of ρ , the short-term and
385 long-term fluctuations observed for low ρ and very rapid short-term fluctuations for high
386 ρ (Figure 1) agree with the patterns observed for long- and short-lived species respectively
387 (Figure 1.1 in [26]). For a same \mathcal{K} we estimate a lower N_e for long-lived species simultane-
388 ously due to larger population fluctuations and to the differences in population turnover
389 speeds (since for low \mathcal{K} both high ρ and low ρ populations have similar fluctuations and
390 yet we observe lower expected N_e), which implies that on the long run a population with
391 low ρ would be expected to maintain lower diversity. This prediction is supported by the
392 lower than expected times to fixation of both neutral alleles and those under selection, as
393 well as higher fixation probabilities of deleterious alleles for low ρ (see Figure 7), which
394 agrees with the observation of less efficient purifying selection in long-lived species with
395 low reproductive rates compared to that of short-lived ones with high reproductive rates

396 [33, 8]. Indeed, our results indicate that in a stable environment, the stochastic demo-
397 graphic fluctuations and the differences in the turnover speeds of species with differing r/K
398 life strategies may suffice in explaining these observations. This could explain why [33, 8]
399 found that past historical demographic disturbances were less explicative than life-history
400 strategies concerning contemporary genetic diversity.

401 4.2 Defining selection and fitness

402 One of the difficulties brought by individual-based models is how to define fitness so that it
403 remains compatible with existing population genetics models. Indeed, several definitions
404 of fitness do exist in literature (reviewed in [13, 30]), fitness generally being defined as
405 a measure of the contribution of a given entity (allele, group of alleles, individual, ...)
406 to the next generation, but the notion of generation in an individual-based model is not
407 obvious. A first way to define fitness is to focus on the Wrightian fitness (see [39]), which
408 is defined as the mean number of progeny per individual. In the logistic birth-and-death
409 model introduced in Section 2, the expected number of offspring for an individual with
410 reproduction rate b , natural death rate d and competition death rate c in a population
411 with (let us say fixed for simplicity) size N is equal to $b/(d + cN)$. Obviously, when
412 a population is at its demographic equilibrium $N = (b - d)/c$ where births and deaths
413 compensate, the fitness of each individual is equal to 1. In this framework the effect
414 of a non-neutral allele or genotype (*i.e.* its coefficient of selection) can be defined as
415 $b'/(d' + c'N) - b/(d + cN) = 0$ if $b'/b = c'/c = d'/d$ (where b' , c' and d' respectively
416 represent the new genotype's birth competition and death rate). However, as shown by
417 the results obtained for “quasi-neutral” selection in [32], where genotypes with the same
418 Wrightian fitness but different values of b were considered, this definition is not sufficient
419 in a continuous time frame. Hence a second way to consider fitness is to focus on the
420 Malthusian fitness, which is defined as the growth rate of the population size. With this
421 definition, fitness for our logistic birth-and-death model can be defined by the quantity
422 $[b/(d + cN)] \times (b + d + cN) = W \times V$ where W is the Wrightian fitness and V measures

423 the speed of reproduction and death of individuals. This second definition of fitness is a
424 more appropriate definition of fitness when studying differences in life-history strategies,
425 as done in [32]. For both of these definitions, fitness is a quantity that is not inherent to
426 the individual but depends on one side on demographic parameters and on the other side
427 on both the population size and, in a non neutral framework, its genetic composition. This
428 releases the exponential growth hypothesis naturally emerging from a concept of constant
429 individual absolute fitness ([29]).

430 In this present work, we have chosen to take into account only the Wrightian fitness so
431 as to first explore the consequences of demographic stochasticity in a model with the same
432 genetic properties as the Wright-Fisher diffusion. Our main conclusion is that, depending
433 on the life-history strategy of a population, the Wright-Fisher diffusion is not always able
434 to capture the trajectories of allelic frequencies. Future work on defining an expression
435 for the coefficient of selection in which the speed of reproduction and death V is also
436 included may provide a better bridge between individual-centered models and the more
437 mathematically manipulable Wright-Fisher diffusion.

438 **4.3 Implications for empirical works**

439 Various methods have been developed to estimate the effective size of populations (see [35]
440 and references therein) with the aim of understanding their past and, in some cases, pre-
441 dicting their future evolution. However, contemporary genetic data can be greatly affected
442 by historical events and so N_e is a parameter that is very population dependent ([36]).
443 Furthermore, from an experimental point of view, the intricacy of population dynamics
444 and population genetics requires the definition of theoretical models whose parameters can
445 be estimated using laboratory experiments for a better understanding of their respective
446 behaviors (reviewed in [28]). Here we provide another definition for N_e that is a result of
447 both the demographic parameters of a population and, in the case of selection, its genetic
448 properties. We find, that contrary to previous works the effects of demographic fluctua-
449 tions can not always be summarized using the mean harmonic population size as proposed

450 in [14, 24, 31]. Using the harmonic mean size is valid only when population fluctuations
451 are sufficiently fast compared to the coalescent times [35], hence for populations with a
452 large growth rate ρ and high death rates due to competition (parameter ξ), which repre-
453 sent short-lived species with high reproductive rates. This remains true even for strong
454 fluctuations in population size when the carrying capacity \mathcal{K} is low. However, for long
455 lived species times to fixation cannot be summarized by N_e , this being greatly due to near
456 extinction events, often ignored in deterministic models (see Chapter 1 in [26]), that can
457 contribute to lower times to fixation. Thus depending on life-history and population size,
458 the Wright-Fisher diffusion is more or less appropriate in predicting population evolution.
459 Though maintained genetic polymorphism is often used as a proxy for adaptive potential,
460 one can also argue that the speed at which an advantageous allele goes to fixation is also
461 important, especially in the face of environmental change ([18] REF). According to our
462 model, long lived species will have a tendency to have lower probabilities of fixation of
463 advantageous alleles, but this may be compensated by the speed at which this fixation
464 occurs compared to that observed in short-lived species.

465 Previous works on integrating stochasticity into demographic models have done so by
466 introducing a demographic variance, meant to reflect the differences between individuals
467 in their survival and reproduction, into deterministic models (see for example [26]). How-
468 ever, as [26] point out, empirical measures of demographic variance may be difficult to
469 obtain, all the more so in the ubiquitous presence of environmental stochasticity. One of
470 the properties of our proposed models is that inter-individual variance occurs naturally,
471 depending on the death and birth rates, and very few parameters are required in order
472 for this variance to be ensured. Indeed, statistical methods using time series have been
473 developed so as to estimate parameters compatible with our model ([3, 4]). Because of the
474 hypotheses we have made concerning birth, death and competition, our model represents
475 a logistic population growth model with extinction. In such a setting, [4] have shown that
476 death and birth rates can be estimated separately and so be used as parameters for our
477 model and compare it to empirical data, either from natural or experimental populations.

478 If our model does indeed agree with empirical observations, a natural next step would
479 be to extend this model so as to consider multiple loci, either neutral or under selection,
480 with possible mutation, so as to provide predictions in a more general genetic setting
481 all the while incorporating intrinsic demographic behaviors which we may be of a great
482 importance in shaping species diversity and evolvability.

483 **Acknowledgements:** *This work was partially funded by the Chair "Modélisation*
484 *Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MNHN-F.X., and was*
485 *also supported by the Mission for Interdisciplinarity at CNRS and by a public grant as*
486 *part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx*
487 *LMH. Diala Abu Awad was funded by the Agence National de la Recherche (ANR SEAD*
488 *- ANR-13-ADAP-0011).*

489 A Definition of N_e and strength of selection

Let us consider our diffusion model introduced in Equation (2)

$$dN_t = \sqrt{N_t} dB_t^1 + N_t[\rho - \xi N_t + \rho\sigma X_t(2h + X_t(1 - 2h) + F(1 - X_t)(1 - 2h))]dt,$$

$$dX_t = \sqrt{\frac{X_t(1 - X_t)}{2\frac{N_t}{(1+F)}}} dB_t^2 + \rho\sigma X_t(1 - X_t)[h + X_t(1 - 2h) + F(1 - X_t - h + 2X_th)]dt.$$

490 Now let us define as in [10] the time change $(\tau_t, t \geq 0)$ such that for all $t > 0$,

$$\int_0^{\tau_t} \frac{(1 + F)N_e}{2N_s} ds = t \quad (9)$$

491 for a given real number N_e , and let us define the time changed diffusion process $(\tilde{N}_t, \tilde{X}_t)_{t \geq 0} =$

492 $(N(\tau_t), X(\tau_t))_{t \geq 0}$.

493 Then $(\tilde{N}_t, \tilde{X}_t)_{t \geq 0}$ satisfies the diffusion equation:

$$d\tilde{N}_t = \sqrt{\frac{2\tilde{N}_t^2}{(1 + F)N_e}} dB_t^1$$

$$+ \frac{2\tilde{N}_t^2}{(1 + F)N_e} \left[\rho - \xi\tilde{N}_t + \sigma\tilde{X}_t \left(2h + \tilde{X}_t(1 - 2h) + F(1 - \tilde{X}_t)(1 - 2h) \right) \right] dt,$$

$$d\tilde{X}_t = \sqrt{\frac{\tilde{X}_t(1 - \tilde{X}_t)}{N_e}} dB_t^2$$

$$+ \sigma \frac{2\tilde{N}_t}{(1 + F)N_e} \tilde{X}_t(1 - \tilde{X}_t) \left[h + \tilde{X}_t(1 - 2h) + F(1 - \tilde{X}_t - h + 2\tilde{X}_th) \right] dt.$$

From this equation and Equation (9) we can, in a neutral case, provide a definition of the effective population mass in our model, defined as the effective population mass of a Wright-Fisher diffusion whose mean absorption time is the same than for our diffusion model with stochastically varying mass. Indeed from Equation (9)

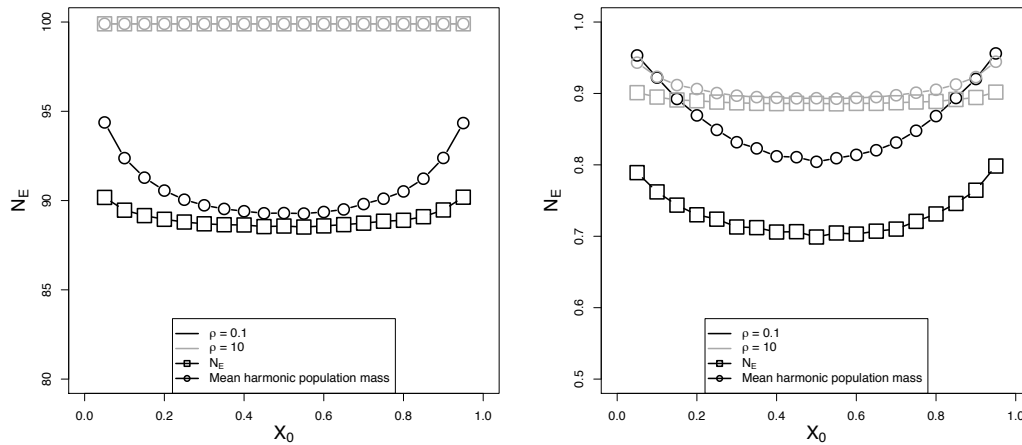
$$\mathbb{E}(T_{abs}) = \mathbb{E} \left(\int_0^{T_{abs}} \frac{(1 + F)N_e}{2N_s} ds \right), \quad \text{which gives}$$

494

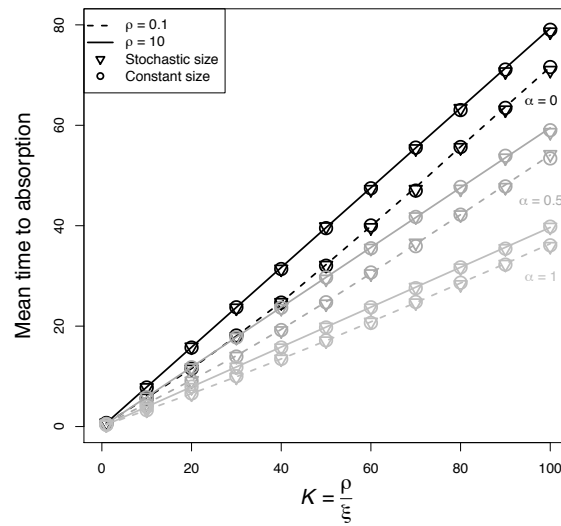
$$N_e = \frac{\mathbb{E}(T_{abs})}{\mathbb{E}\left[\int_0^{T_{abs}} \frac{(1+F)}{2N_s} ds\right]} = \frac{\mathbb{E}(T_{abs})}{\frac{4}{2-\alpha} \mathbb{E}\left[\int_0^{T_{abs}} \frac{1}{N_s} ds\right]}.$$

495 Note that using the more widely used harmonic mean of population mass so as to
 496 describe N_e results in over-estimations fo N_e (Figure A.1).

497 Note also that in the non-neutral case this change of time to obtain a Wright-Fisher
 498 diffusion with selection is not possible. Indeed, in this case the time-changed diffusion
 499 giving the proportion \tilde{X}_t of allele a follows a haploid Wright-Fisher diffusion with effective
 500 population mass equal to N_e but with selection coefficient equal to $\sigma \frac{2\tilde{N}_t}{(1+F)N_e}$ at time t .
 501 Population demography can therefore be seen and defined as a changing environment,
 502 though this environment is in this case itself influenced by the feedback of genetics. In
 503 this case a natural approximation is to take $s = \sigma$, as shown in Section 3.4.



A.1: Comparing the Effective population mass proposed in equation 6 to the Mean harmonic population mass obtained from simulations run as a function of the initial frequency X_0 of a neutral allele a ($\sigma = 0$) for two values of ρ (0.1 in black and 10 in gray) and two values of \mathcal{K} , on the left $\mathcal{K} = 1$ and on the right 100.



A.2: Mean times to absorption of a neutral allele ($\sigma = 0$) with random mating ($\alpha = 0$), partial selfing ($\alpha = 0.5$) and strict selfing ($\alpha = 1$) and different values of the growth rate ρ , as a function of the ratio $\frac{\rho}{\sigma^2} = \mathcal{K}$ (Equation 4) for three cases: 1) Simulations of the stochastic diffusion process (2), 2) Simulations of the Wright-Fisher diffusion using N_e defined in Equation (6) and 3) Theoretical approximations provided by [2] using N_e , represented by the lines (dashed for $\rho = 0.1$ and full for $\rho = 10$).

504 References

- 505 [1] T. Bataillon and M. Kirkpatrick. Inbreeding depression due to mildly deleterious mu-
506 tations in finite populations: size does matter. GENETICAL RESEARCH, 75(1):75–
507 81, FEB 2000.
- 508 [2] A. Caballero and W. Hill. Effects of partial inbreeding on fixation rates and variation
509 of mutant-genes. Genetics, 131:493–507, 1992.
- 510 [3] F. Campillo and F. L. Gland. Mle for partially observed diffusions: direct maximiza-
511 tion vs. the em algorithm. Stochastic Processes and their Applications, 33(2):245 –
512 274, 1989.
- 513 [4] F. Campillo, M. Joannides, and I. Larramendy-Valverde. Parameter identification for
514 a stochastic logistic growth model with extinction. Communications in Statistics -
515 Simulation and Computation, 0(0):1–17, 2017.
- 516 [5] N. Champagnat, R. Ferrière, and S. Méléard. Unifying evolutionary dynamics: From
517 individual stochastic processes to macroscopic models. Theor. Popul. Biol., 69:297–
518 321, 2006.
- 519 [6] N. Champagnat and A. Lambert. Evolution of discrete populations and the canonical
520 diffusion of adaptive dynamics. Ann. Appl. Probab., 17(1):102–155, 2007.
- 521 [7] N. Champagnat and S. Méléard. Invasion and adaptive evolution for individual-based
522 spatially structured populations. J. Math. Biol., 55:147–188, 2007.
- 523 [8] J. Chen, S. Glémin, and M. Lascoux. Genetic diversity and the efficacy of puri-
524 fying selection across plant and animal species. Molecular Biology and Evolution,
525 34(6):1417–1428, 2017.
- 526 [9] P. Collet, S. Méléard, and J. A. J. Metz. A rigorous model study of the adaptive
527 dynamics of Mendelian diploids. J. Math. Biol., 67(3):569–607, 2013.

- 528 [10] C. Coron. Slow-fast stochastic diffusion dynamics and quasi-stationary distributions
529 for diploid populations. J. Math. Biol., Published Online, 2013.
- 530 [11] C. Coron. A model for mendelian populations demogenetics. To appear in ESAIM:
531 Proceedings and Surveys, 2015.
- 532 [12] C. Coron, S. Méléard, and D. Villemonais. Perpetual integrals convergence and ex-
533 tinctions in population dynamics. Prépublication, 2017.
- 534 [13] T. Day and S. P. Otto. Fitness. John Wiley & Sons, Ltd, 2001.
- 535 [14] W. Ewens. The probability of survival of a new mutant in a fluctuating environment.
536 Heredity, 22:438–443, 1967.
- 537 [15] S. Fisher, Ronald A. The Genetical Theory of Natural Selection. Oxford University
538 Press, Oxford, 1930.
- 539 [16] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regu-
540 lated population and macroscopic approximations. Ann. Appl. Probab., 14(4):1880–
541 1919, 2004.
- 542 [17] J. Gillespie. Population Genetics: A Concise Guide. Population Genetics. Johns
543 Hopkins University Press, 1998.
- 544 [18] S. Glemin and J. Ronfort. Adaptation and maladaptation in selfing and outcrossing
545 species: new mutations versus standing variation. Evolution, 67(1):225–240, 2013.
- 546 [19] G. Greenbaum. Revisiting the time until fixation of a neutral mutant in a finite
547 population a coalescent theory approach. Journal of theoretical biology, 380:98–102,
548 2014.
- 549 [20] M. Iizuka. Effective population size of a population with stochastically varying size.
550 Journal of Mathematical Biology, 61(3):359–375, 2010.
- 551 [21] M. Iizuka, H. Tachida, and H. Matsuda. A neutral model with fluctuating population
552 size and its effective size. Genetics, 161(1):381–388, 2002.

- 553 [22] M. Kimura. Some problems of stochastic processes in genetics. Ann. Math. Statist.,
554 28(4):882–901, 12 1957.
- 555 [23] M. Kimura. On the probability of fixation of mutant genes in a population. Genetics,
556 47:713–719, 1962.
- 557 [24] M. Kimura. Stochastic processes in population genetics, with special reference to
558 distribution of gene frequencies and probability of gene fixation. In Mathematical
559 Topics in Population Genetics, pages 178–209. Springer Berlin Heidelberg, 1970.
- 560 [25] M. Kimura and T. Ohta. The average number of generations until fixation of a mutant
561 gene in a finite population. Genetics, 61(3):763–771, 1969.
- 562 [26] R. Lande, S. Engen, and B.-E. Saether. Stochastic Population Dynamics in Ecology and Conservation (Ox
563 Oxford University Press, USA, illustrated edition edition, 2003.
- 564 [27] P. A. P. Moran. The estimation of the parameters of a birth and death process.
565 Journal of the Royal Statistical Society. Series B (Methodological), 15(2):241–245,
566 1953.
- 567 [28] L. Mueller. Fitness, demography, and population dynamics in laboratory experi-
568 ments. In Experimental evolution: concepts, methods, and applications of selection
569 experiments, chapter 9. Berkeley, Calif, University of California Press, 2009.
- 570 [29] A. Orr and R. Unckless. Population extinction and the genetics of adaptation. The
571 American Naturalist, 172(2):160–169, 2008.
- 572 [30] H. A. Orr. Fitness and its role in evolutionary genetics. Nat Rev Genet., 10(8):531–
573 539, 2009.
- 574 [31] S. Otto and M. Whitlock. The probability of fixation in populations of changing size.
575 Genetics, 146(2):723–733, 1997.
- 576 [32] T. L. Parsons, C. Quince, and J. B. Plotkin. Some consequences of demographic
577 stochasticity in population genetics. Genetics, 1354:1345–1354, 2010.

- 578 [33] J. Romiguier, P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil,
579 Y. Chiari, R. Dernat, L. Duret, N. Faivre, E. Loire, J. M. Lourenco, B. Nabholz,
580 C. Roux, G. Tsagkogeorga, A. A.-T. Weber, L. A. Weinert, K. Belkhir, N. Bierne,
581 G. S., and N. Galtier. Comparative population genomics in animals uncovers the
582 determinants of genetic diversity. Nature, 515:261–263, 2014.
- 583 [34] D. Roze and F. Rousset. Joint effects of self-fertilization and population structure on
584 mutation load, inbreeding depression and heterosis. Genetics, 167:1001–1015, 2004.
- 585 [35] P. Sjödin, I. Kaj, S. Krone, M. Lascoux, and M. Nordborg. On the meaning and
586 existence of an effective population size. Genetics, 169(2):1061–1070, 2005.
- 587 [36] J. Wang. Estimation of effective population sizes from data on genetic markers. Philos
588 Trans R Soc Lond B Biol Sci, 360(1459):1395–409, 2005.
- 589 [37] S. Wright. Evolution in mendelian populations. Genetics, 16(2):97–159, 1931.
- 590 [38] S. Wright. Size of population and breeding structure in relation to evolution. Science,
591 87:430–431, 1938.
- 592 [39] B. Wu, C. S. Gokhale, M. van Veelen, L. Wang, and A. Traulsen. Interpretations
593 arising from wrightian and malthusian fitness under strong frequency dependent se-
594 lection. Ecol Evol., 3(5):12761280, 2013.