

## Appendix I : Mathematical derivations and approximations

In the following, we use a supplementary notation to simplify the mathematical derivations:  $\theta = n/2$ .

### I. General WSSM approximation to ER.

In the present model, we approximate the stochastic dynamics of each lineage by a Feller diffusion (Feller 1951) or continuous branching CB-process (Lambert 2008), with parameters  $\{r_i, \sigma_i\}$  for lineage  $i$ . We ignore any density or frequency dependence and assume that all lineages have similar stochastic reproductive variance ( $\sigma_i \approx \sigma$ ). The resulting total population size  $N_t$  (cumulating all lineage that co-segregate) can then also be approximated by a CB process, which follows the stochastic differential equation:

$$dN_t = \bar{r}_t N_t dt + \sqrt{\sigma N_t} dB_t , \quad (\text{A1})$$

where  $B_t$  is a Weiner process,  $\bar{r}_t = 1/N_t \sum_{i=1}^{N_t} r_i$  is the mean growth rate of all lineages present at time  $t$  and  $\sigma$  is the common stochastic reproductive variance of all lineages. In the WSSM regime (i.e. when  $U \gg U_c = \theta^2 \lambda$ ), we can ignore the evolutionary stochasticity introduced by mutation and drift as a first approximation. Then, the mean growth rate  $\bar{r}_t \approx \langle \bar{r}_t \rangle$  (expectation over replicates denoted by  $\langle . \rangle$ ) is approximately deterministic and given by the WSSM results in (Martin and Roques 2016) for the FGM. The probability of such a time inhomogeneous CB process to be extinct by time  $t$  is given by (Bansaye and Simatos 2015):

$$P_{\text{ext}}(t) = \exp\left(-\frac{2N_0}{\psi_t}\right) , \quad (\text{A2})$$

where  $\psi_t = \int_0^t e^{-\rho(v)} dv$  and  $\rho(u) = \int_0^u \langle \bar{r}_v \rangle dv$ . The probability of ER is the complementary probability of not being extinct over infinite time, namely  $P_R = 1 - P_{\text{ext}}(\infty)$ .

20 **II. Explicit expression for a population initially clonal**

21 Define  $r_D$  the decay rate of the initial clone, and  $r_{max}$  the maximal attainable growth rate (that of a  
 22 genotype optimal in the stressful environment). In this case, the mean fitness trajectory, relative to  
 23 the optimal genotype ( $\langle \bar{m}_t \rangle = \langle \bar{r}_t \rangle - r_{max}$ ), under the WSSM approximation, for an initially clonal  
 24 population is (eq. (12) in (Martin and Roques 2016))

$$\langle \bar{m}_t \rangle = m_0 \operatorname{sech}(\mu t)^2 - \theta \mu \tanh(\mu t), \quad (\text{A3})$$

25 where  $\mu = \sqrt{U\lambda}$  and  $\theta = n/2$ ,  $\operatorname{sech}(\cdot)$  and  $\tanh(\cdot)$  are the hyperbolic secant and tangent functions  
 26 and  $m_0$  is the fitness difference between the original clone and the optimal genotype. Now, we require  
 27 the absolute mean fitness trajectory (mean growth rate), which is obtained by noting that, by  
 28 definition,  $\langle \bar{m}_t \rangle = \langle \bar{r}_t \rangle - r_{max}$  and  $m_0 = -r_D - r_{max}$ . Denoting  $y_D = r_D/r_{max}$ ,  $\epsilon = \theta \mu/r_{max}$  we  
 29 obtain:

$$\langle \bar{r}_t \rangle = \frac{\theta \mu}{\epsilon} (1 - (y_D + 1) \operatorname{sech}(\mu t)^2 - \epsilon \tanh(\mu t)). \quad (\text{A4})$$

30 The integral  $\rho_t = \int_0^t \langle \bar{r}_v \rangle dv$  of this growth rate over time can be expressed in compact form by using  
 31 the change of variable  $\tau = \mu t$ , yielding

$$\begin{aligned} \rho_t &= \frac{\theta}{\epsilon} f(\tau) - \theta \log(h(\tau)) \\ h(\tau) &= \cosh(\tau) \\ f(\tau) &= \tau - (1 + y_D) \tanh(\tau) \end{aligned} \quad (\text{A5})$$

32 Finally, the same change of variable ( $dt = \mu d\tau$ ) can also be used to express the indefinite integral that  
 33 determines extinction probabilities, yielding the rate of ER per lineage present at the onset of stress:

$$\begin{aligned} \omega^{DN} &= -\frac{\log P_{\text{ext}}(\infty)}{N_0} = \frac{2}{\int_0^\infty e^{-\rho u} du} \\ \omega^{DN} &= 2 \mu / \left( \int_0^\infty h(\tau)^\theta \exp\left(-\frac{\theta}{\epsilon} f(\tau)\right) d\tau \right) \end{aligned} \quad (\text{A6})$$

34

35 **III. Laplace approximations for small  $\epsilon$ , for a population initially clonal**

36 Eq.(A6) is fully analytic but yields no explicit expression, which would be useful to get a more intuitive  
 37 grasp of how each parameter affects the ER probability. When the mutation rate and effects are small  
 38 enough that the load  $\theta \mu$  is small relative to the maximal growth rate  $r_{max}$  (so that we are well away  
 39 from lethal mutagenesis), then  $\epsilon \ll 1$ . This means that  $\theta/\epsilon \gg 1$  (with  $\theta \geq 1/2$ ) and the integral in  
 40 Eq.(A6) is amenable to the Laplace Approximation (as  $h(\tau)$  is monotonous over  $y \in \mathbb{R}^+$ ). This  
 41 approximation can be formulated as follows :

$$\int_0^\infty h(\tau)^\theta e^{-\theta f(\tau)/\epsilon} d\tau \underset{\epsilon \rightarrow 0}{\approx} \sqrt{\frac{2\pi\epsilon}{\theta f''(\tau_0)}} h(\tau_0)^\theta e^{-\theta f(\tau_0)/\epsilon} \quad (\text{A7})$$

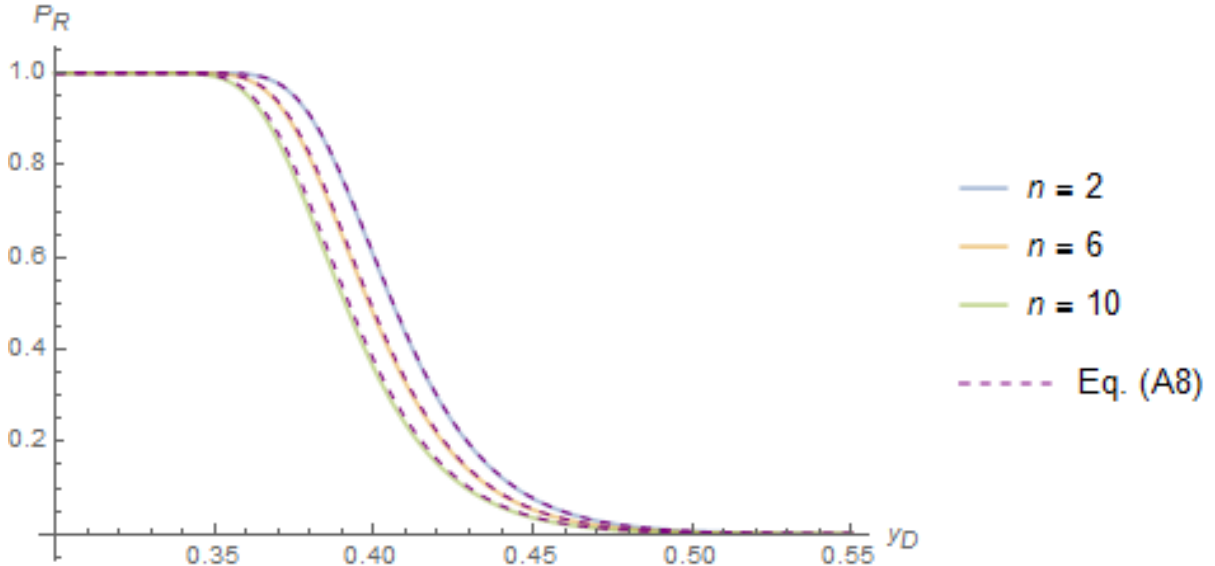
42 where  $\tau_0 = \cosh^{-1}(\sqrt{1+y_D})$  is the unique minimum of  $f(\cdot)$  over  $\tau \in \mathbb{R}^+$ . The Laplace approximation  
 43 in Eq.(A7) then yields a fully explicit expression for the rate of ER per lineage present at the onset of  
 44 stress defined in Eq.(A6). Rewriting in terms of the original parameters ( $\theta/\epsilon = r_{max}/\mu$ ), we get

$$\omega^{DN} \underset{\epsilon \rightarrow 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} \gamma(y_D)\right) \quad (\text{A8})$$

$$\gamma(y_D) = \sqrt{y_D(1+y_D)} - \cosh^{-1}(\sqrt{1+y_D}) + \epsilon \left( \left(1 + \frac{1}{2\theta}\right) \frac{\log(1+y_D)}{2} - \frac{1}{\theta} \frac{\log(y_D)}{4} \right)$$

45 The accuracy of this expression is illustrated in **Supplementary Figure 1**.

46



47

48 **Supplementary Figure 1:**  $P_R$  as a function of  $y_D$  for three values of the dimensionality (given in legend),  
 49 as computed from the 'exact' Eq.(A6) vs. Laplace Approximation Eq.(A8). Parameters are  $r_{max} = 1, \lambda =$   
 50  $5 \cdot 10^{-4}, U = 0.25, N_0 = 10^7$ .

51

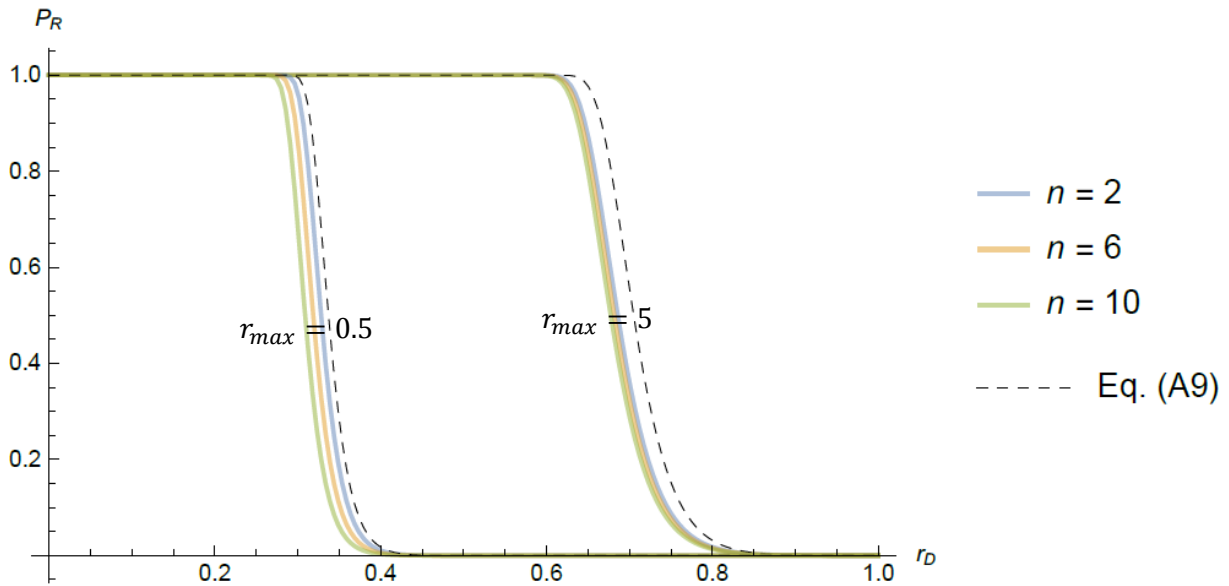
52 In fact, unless we consider narrow range of stress variation, as in **Supplementary Figure 1**, it appears  
 53 that, in the limit  $\epsilon \rightarrow 0$ ,  $P_R$  shows limited dependency on dimensionality, provided that it remain  
 54 limited ( $n$  varies by five-fold above). This is confirmed by further simplifying Eq.(A8) to produce a rough

55 but reasonably accurate approximation, in a range that is a priori of most biological relevance, say for  
 56  $y_D > 0.1$ , i.e. not-too mild a stress. In this range both  $\log(1 + y_D)/2$  and  $\log(y_D)/4$  are of similar or  
 57 smaller order than  $g(y_D) = \sqrt{y_D(1 + y_D)} - \cosh^{-1}(\sqrt{1 + y_D})$ . Therefore, as  $\epsilon \rightarrow 0$  (as we assume  
 58 here), the right hand factor in  $\gamma(y_D)$ , proportional to  $\epsilon$ , becomes negligible, relative to the left hand  
 59 term, and  $\gamma(y_D) \approx g(y_D)$ . Eq. (A8) thus simplifies to the expression given in Eq.(6) for *de novo* rescue:

$$\omega^{DN} \underset{\substack{\epsilon \rightarrow 0 \\ y_D \geq 0.1}}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} g(y_D)\right) \cdot g(y_D) = \sqrt{y_D(1 + y_D)} - \cosh^{-1}(\sqrt{1 + y_D}) \quad (A9)$$

60 This simpler approximation is less precise but still provides a good order of magnitude, when  $\epsilon$  is small  
 61 enough. **Supplementary Figure 2** illustrates its accuracy and the fact that the ER rate is indeed roughly  
 62 independent of dimensionality  $\theta$  in this parameter range.

63



64

65 **Supplementary Figure 2:** Same as **Supplementary Figure 1** (same parameters, except  $r_{max}$ , indicated  
 66 on the graph), this time as a function of  $r_D$  and compared to Eq.(A9) (dashed lines).

67 **IV. Explicit expression for a population initially at mutation-selection balance**

68 This time we assume initially mutation-selection balance, then a shift in optimum occurs without any  
 69 change in  $U$  or  $\lambda$ : adaptation is driven by both standing variance and de novo mutations. Based on the  
 70 corresponding WSSM approximation (eq. (13) in (Martin and Roques 2016) with the same approach  
 71 and parameterization as for Eq.(A4) we get

$$\langle \bar{r}_t \rangle = r_{max}(1 - e^{-2\mu t}(1 + y_D) - \epsilon), \quad (\text{A10})$$

72 which then yields the corresponding expression for  $\rho_t$  (same form as Eq.(A5)):

$$\begin{aligned} \rho_t &= \frac{\theta}{\epsilon} f(\tau) - \theta \log(h(\tau)) \\ f(\tau) &= \tau - (1 + y_D) (1 - e^{-2\tau})/2 \\ h(\tau) &= e^\tau \end{aligned} \quad (\text{A11})$$

73 The rate of ER takes a similar form as in the DN scenario (Eq.(A6)), but this time an explicit exact  
 74 expression is found:

$$\begin{aligned} \omega^{DN+SV} &= -\log P_{\text{ext}}(\infty)/N_0 = \frac{2}{\int_0^\infty e^{-\rho u} du} = 2\mu / \int_0^\infty h(\tau)^\theta e^{-\theta f(\tau)/\epsilon} d\tau \\ \omega^{DN+SV} &= 4\mu \frac{e^{-\xi} \xi^\beta}{\Gamma(\beta) - \Gamma(\beta, \xi)} \end{aligned} \quad (\text{A12})$$

75 Where  $\xi = (1 + y_D) \theta / (2 \epsilon)$  and  $\beta = (1 - \epsilon) \theta / (2 \epsilon)$ , and  $\Gamma(\cdot)$  and  $\Gamma(\cdot, \cdot)$  are respectively the  
 76 Euler's gamma function and the incomplete gamma function.

77 **V. Laplace approximations for small  $\epsilon$ , for a population initially at mutation-selection balance**

78 Although this is already explicit, a Laplace approximation can be used to get a simpler result, again  
 79 away from the lethal mutagenesis regime ( $\epsilon \ll 1$ ). The function  $f(\cdot)$  has a unique minimum at  $\tau_0 =$   
 80  $\log(\sqrt{1 + y_D})$  and the resulting Laplace approximation for the rate of ER is of the same form as  
 81 Eq.(A8):

$$\omega^{DN+SV} \underset{\epsilon \rightarrow 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} \gamma(y_D)\right), \quad (A13)$$

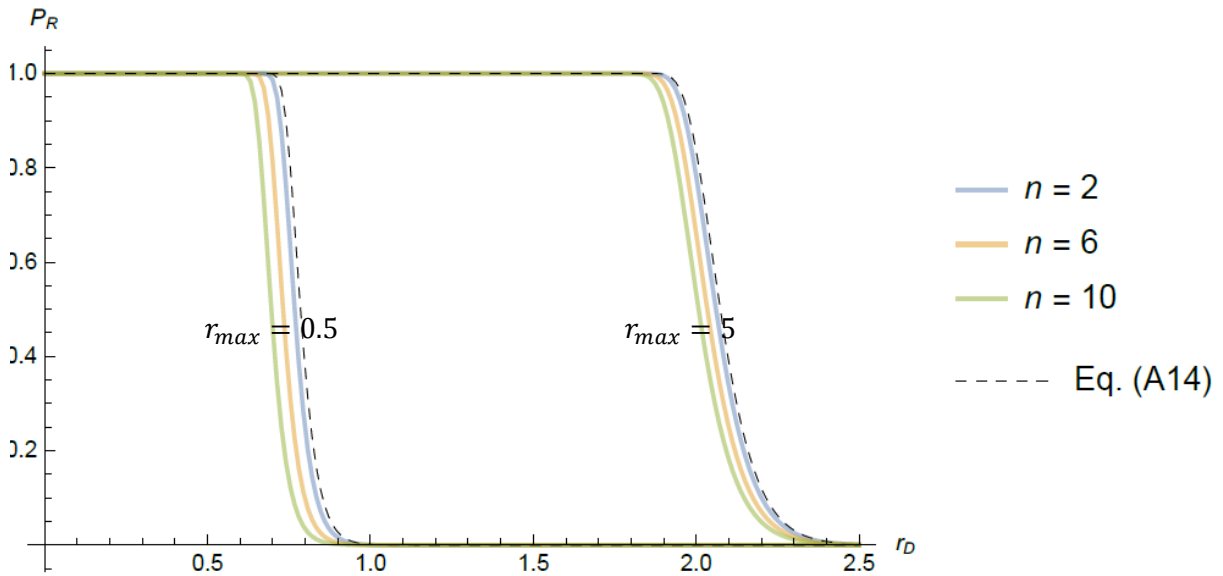
$$\gamma(y_D) = \frac{y_D - (1 - \epsilon) \log(1 + y_D)}{2}$$

82 From which, as  $\epsilon \ll 1$ , arises the same expression as Eq.(A9) with a different function  $g(\cdot)$ :

$$\omega^{DN+SV} \underset{\epsilon \rightarrow 0}{\approx} 2 \sqrt{\frac{r_{max} \mu}{\pi}} \exp\left(-\frac{r_{max}}{\mu} g(y_D)\right). \quad (A14)$$

$$g(y_D) = (y_D - \log(1 + y_D))/2$$

83 The accuracy of this approximation is illustrated in **Supplementary Figure 3** below. Note also how the  
 84 drop in ER probability occurs at higher stress levels here than with purely *de novo* mutation (in  
 85 **Supplementary Figure 2**).

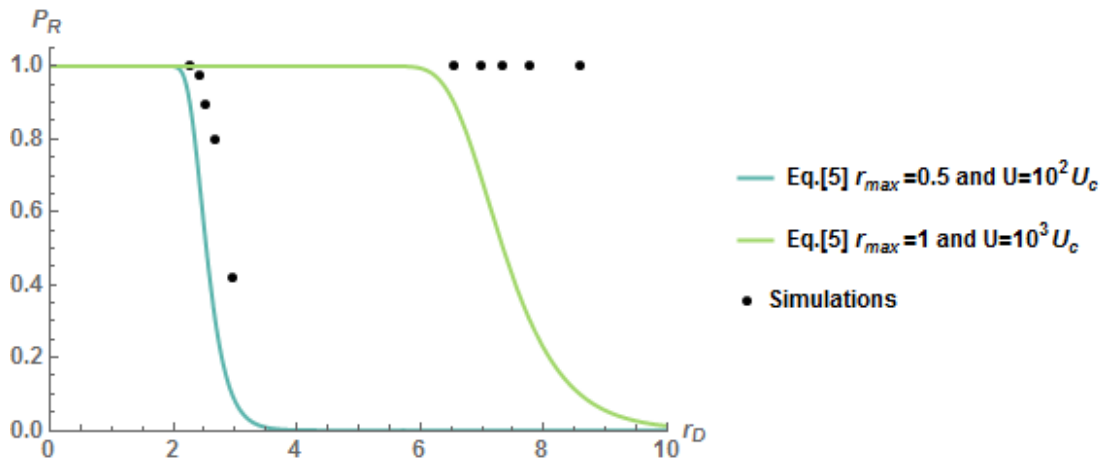


86  
 87 **Supplementary Figure 3:** Same as **Supplementary Figure 2** (same parameters) with standing variance  
 88 plus *de novo* mutation Eq.(A12), and the corresponding approximation Eq.(A14).

89  
 90 The accuracy shown in the main text between Eq.[5] (Eq.A(12)) and the simulations decrease when the  
 91 growth rates of the individuals in the population increase (in absolute value). Indeed, when growth  
 92 rates are too low or high ( $|r| \geq 1$ ), the continuous time approximation used in our model fails to  
 93 predict accurately the discrete time dynamic of the population size and therefore the probability of  
 94 rescue given by our simulations. This is illustrated in **Supplementary Figure 4** where we can see that  
 95 increasing  $r_{max}$  leads to an underestimation of the rescue probability predicted by the simulations.  
 96 Moreover, the blue line in **Supplementary Figure 4** shows that an increase in  $U$  (in comparison to the  
 97 parameters of **Figure 3** and with the same  $r_{max}$ ) also leads to an underestimation of the rescue

98 probability. Indeed, the absolute growth rates simulated are higher for a higher  $U$ , leading to the same  
99 discrepancies between continuous time and discrete time dynamics mentioned above.

100



102

103 **Supplementary figure 4:** Accuracy of the predictions of ER probabilities from both de novo and  
104 pre-existent mutants (blue: Eq.[5] *Erreur ! Source du renvoi introuvable.* (Eq.A(12))) for two  
105 values of  $r_{max}$  ( $r_{max} = 0.5$  in blue and  $r_{max} = 1$  in green) against simulations. Other  
106 parameters are  $N_0 = 10^5$ ,  $n = 4$ ,  $\lambda = 5 \cdot 10^{-3}$ ,  $U = 10^2 U_c$  for the blue line and  $U = 10^3 U_c$  for  
the green line.

107 **VI. Width of the mutation window of ER**

108 We now wish to evaluate the mutation at which ER switches from very likely to very unlikely. We know  
 109 that the rate of ER drops sharply at some upper threshold mutation rate, due to lethal mutagenesis  
 110 effects ( $\mu \theta = r_{max}$ ). This implies a maximal mutation rate  $U_{max} = r_{max}^2 / (\theta^2 \lambda)$ , above which lethal  
 111 mutagenesis leads to certain extinction of the population.

112 Here, we are interested in the parameter range, far below lethal mutagenesis, where the ER probability  
 113 increases (sharply too) with the mutation rate. We first seek the value  $\mu_*$  at which ER occurs 50% of  
 114 the time. This critical  $\mu_*$ , is the one above which ER becomes likely, so that the rescue probability is  
 115  $P_R(\mu_*) = 1/2$  and the ER rate is  $\omega_* = -\log(1 - P_R(\mu_*)) / N_0 = \log(2) / N_0$ . In this range of  
 116 parameter, we can use the approximate expressions in Eqs.(A9) and (A14) :  $\omega(\mu_*) \approx$   
 117  $2 \sqrt{r_{max} \mu_* / \pi} \exp(-r_{max} g(y_D) / \mu_*)$ . Solving for this equation yields a unique solution:

$$\mu_* \approx \frac{2 g(y_D) r_{max}}{\mathcal{W}(5.3 g(y_D) N_0^2 r_{max}^2)} \quad , \quad (A15)$$

118 where  $\mathcal{W}(\cdot)$  is Lambert's function. This rate can be approximated to a simpler form if we note that  $N_0^2$   
 119 is typically very large compared to  $g(y_D)$  and  $r_{max}$ , which are of order 1. Therefore the denominator  
 120 in Eq.(A15) is driven by the asymptotic limit of  $\mathcal{W}(\cdot)$  ( $\mathcal{W}(v) \approx \log(v / \log(v))$ , for large  $v$ ) and by the  
 121 terms in  $N_0$ . Overall, to a reasonably good approximation:

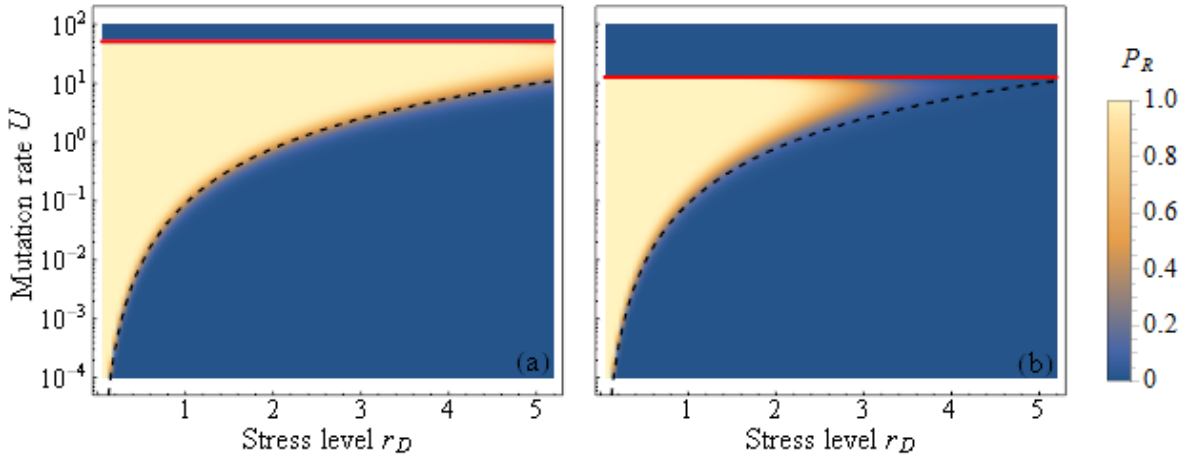
$$\mu_* \underset{N_0 \rightarrow \infty}{\approx} \frac{2 g(y_D) r_{max}}{\log\left(\frac{N_0^2}{\log(N_0^2)}\right)} \quad . \quad (A16)$$

122 Recalling that  $\mu = \sqrt{U \lambda}$ , the critical mutation rate where the ER probability is 50% is

$$U_* \approx \frac{1}{\lambda} \left( \frac{2 g(y_D) r_{max}}{\mathcal{W}(5.3 g(y_D) N_0^2 r_{max}^2)} \right)^2 \underset{N_0 \rightarrow \infty}{\approx} \frac{4 g(y_D)^2 r_{max}^2}{\lambda \log\left(\frac{N_0^2}{\log(N_0^2)}\right)^2} \quad . \quad (A17)$$

123 This means that ER is only likely within a window of mutation rate  $U_* \leq U \leq U_{max}$ . This window  
 124 narrows down as stress increases (increased decay rate and hence  $g(y_D)$ ) **Supplementary Figure 6**, or  
 125 as  $N_0$  gets smaller and its lower bound is roughly independent of dimensionality. The accuracy of this  
 126 approximation is illustrated in **Figure 4** for the DN scenario ( $g(y_D) = g_{DN}(y_D) = \sqrt{y_D(1+y_D)} -$   
 127  $\cosh^{-1}(\sqrt{1+y_D})$ ) and in **Supplementary Figure 5** below for the SV+DN scenario in the presence of  
 128 standing variance ( $g(y_D) = g_{DN+SV}(y_D) = (y_D - \log(1+y_D))/2$ ).





129

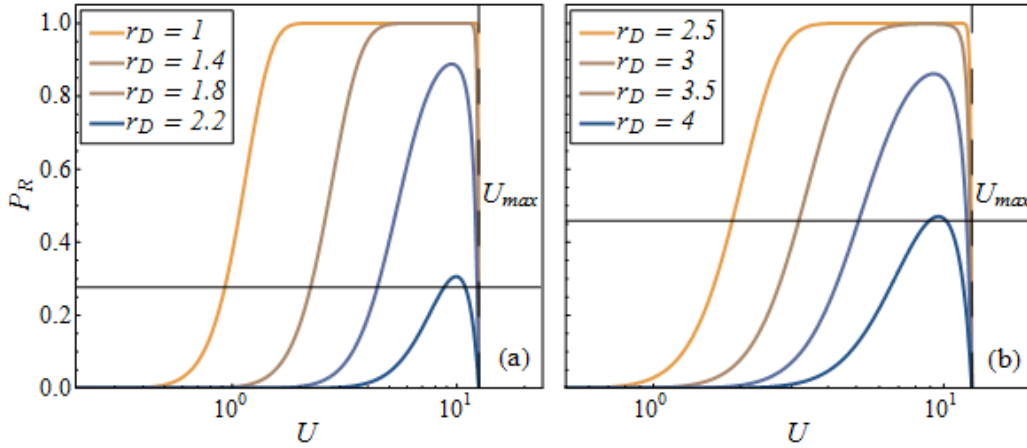
130

**Supplementary Figure 5:** same as **Figure 4** with standing variance plus *de novo* mutation.

131

132

### VII. Height of the mutation window of ER



133

134

**Supplementary Figure 6:** Decrease of the width and height of the mutation window with stress. Colored plain lines show Eq.[6] for ER from *de novo* mutations (a) or from pre-existing standing genetic variance and *de novo* mutations (b). Black lines give  $\max(P_R)$  from Eq.(A18) (for  $r_D = 2.2$  in (a) and  $r_D = 4$  in (b)). The colors and the parameters are the same as in **Figure 4** for increasing  $r_D$ .

138

139

140

141

142

143

144

**Supplementary Figure 6** shows that, for sufficiently high stress  $r_D$  or low inoculum size  $N_0$ , the ER probability cannot reach above some limited maximum  $p = \max_{\mu \in \mathbb{R}^+} (P_R)$ . **Figure 5** also shows that  $\max(P_R)$  drops sharply (from  $\max(P_R) = 1$  to  $\max(P_R) \ll 1$ ) as  $r_D$  increases or  $N_0$  decreases. The goal of this section is to compute the values of  $r_D$  where this transition occurs, namely the stress levels beyond which the population cannot avoid extinction, even if its mutation rate was to be changed (e.g. by hyper-mutators).

145

146

147

The transition observed in **Figure 5** occurs at some pair of parameters  $\{r_D^*(p), N_0^*(p)\}$  for which  $\max(P_R) = p$  for some threshold value  $p$ . The shape of the transition in **Figure 5** suggests that the values of  $r_D^*(p)$  and  $\log(N_0^*(p))$  are linearly related: so that  $\log N_0^*(p) = a + b r_D^*(p)$

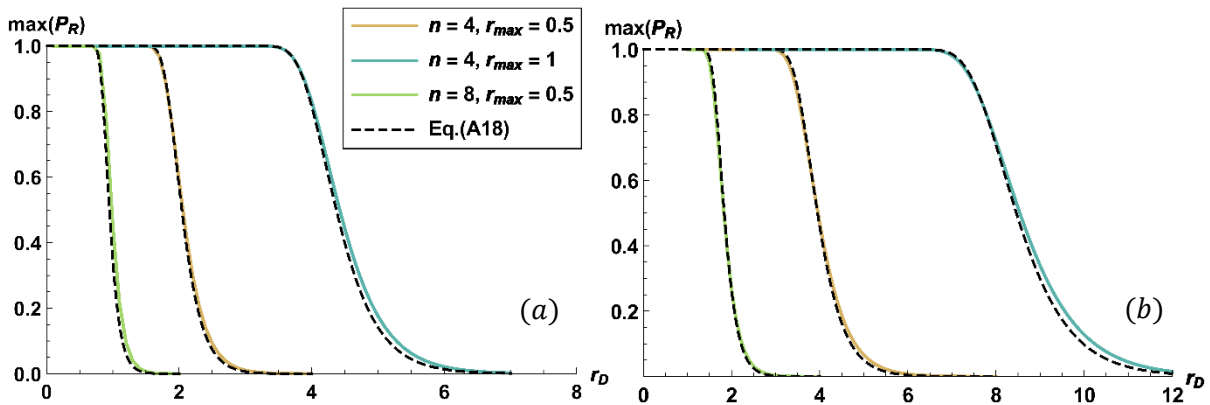
148 for some  $(a, b) > 0$ . We also know that  $P_R = 1 - \exp(-N_0 \omega)$  where  $\omega$  depends on  
 149  $r_D, r_{max}, \mu, n$  (Eq.(5)). At  $\{r_D^*, \log N_0^*\}$  we thus have  $p = \max(P_R) = 1 - \exp(-N_0^*(p) \omega^*(p))$ ,  
 150 where  $\omega^*(p)$  is the corresponding ER rate along the transition of height  $p$ . This implies  
 151 that  $\log N_0^*(p) = \log(|\log(1 - p)|) - \log \omega_*(p) = a + b r_D^*(p)$ , so that  $\omega_*(p) = \omega(r_D^*(p)) =$   
 152  $A e^{-B r_D^*(p)}$  for some  $(A, B)$ . Since  $\omega_*$  is independent of  $N_0$  and is a maximum over  $\mu$  (so  
 153 independent of  $\mu$ ), it must only depend on the remaining parameters  $\{r_D, r_{max}, n\}$ . The  
 154 coefficients  $A, B$ , which determine its relationship with  $r_D^*$ , must thus only depend on  $n$   
 155 and  $r_{max}$ . Therefore, we look for approximations of the form  $\omega_* \approx A e^{-B r_D^*}$ , for some functions  
 156  $A = A(n, r_{max})$  and  $B = B(n, r_{max})$  for both  $DN$  and  $DN + SV$  scenarios. A simple fitting procedure,  
 157 across a range of values of  $\{n, r_{max}\}$ , carried out with Matlab© curve fitting toolbox suggests that  $A \approx$   
 158  $r_{max}/n$  and  $B \approx \delta n/r_{max}$ , leading to:

$$\omega_*(p) = \frac{r_{max}}{n} \exp\left(-\delta \frac{n}{r_{max}} r_D^*(p)\right)$$

with  $\begin{cases} DN : \delta = 0.6 \\ DN + SV : \delta = 0.31 \end{cases}$  (A18)

159 From Eq.(A18) we can see that the height  $\max(P_R)$  of the mutation window increases with an increase  
 160 in the maximal growth rate  $r_{max}$  or the population size  $N_0$  and decreases with an increase in the  
 161 dimensionality  $n$  or the harshness of stress  $r_D$  (the accuracy of this fitting is illustrated in  
 162 **Supplementary figures 7**). The figure below provides a comparison between the numerical value  
 163 of  $\max(P_R)$  from Eq.[5] and the result of Eq. A(18).

164



165

166 **Supplementary Figure 7:**  $\max(P_R)$  as a function of  $r_D$  for different values of the dimensionality and  
 167 the maximal growth rate (given in legend), as computed from the 'exact' Eq.[5] vs. fitted approximation  
 168 Eq.A(18). (a) Population adapting from de novo mutation only. (b) Population adapting from both de  
 169 novo mutation standing genetic variance. In both cases  $N_0 = 10^5$ .

170

171 Solving  $\max(P_R) = p = 1 - e^{-N_0 \omega_*(p)}$  in terms of  $r_D^*(p)$  using Eq. A(18) yields the threshold stress at  
 172 which  $\max(P_R) = p$  :

$$r_D^*(p) = \frac{r_{max}}{n \delta} \log\left(\frac{N_0 r_{max}}{n |\log(1-p)|}\right) \quad (\text{A19})$$

173 where  $\delta$  is given in Eq. (A 18) depending on the scenario.

174

175 **References:**

176 Bansaye, V., and F. Simatos. 2015. On the scaling limits of Galton-Watson processes in varying  
177 environments. *Electron. J. Probab.* 20.

178 Feller, W. 1951. *Diffusion processes in genetics*. University of California Press Berkeley, Calif.

179 Lambert, A. 2008. Population Dynamics and Random Genealogies. *Stoch. Models* 24:45–163.

180 Martin, G., and L. Roques. 2016. The Non-stationary Dynamics of Fitness Distributions: Asexual Model  
181 with Epistasis and Standing Variation. *Genetics* genetics.116.187385.

182