

# Influence of viral replication mechanisms on within-host evolutionary dynamics

Claude Loverdo<sup>1</sup>, Miran Park<sup>1</sup>, Sebastian J. Schreiber<sup>2</sup>, and James O. Lloyd-Smith<sup>1,3</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

<sup>2</sup>Department of Evolution and Ecology and Center for Population Biology, University of California, Davis, California 95616

<sup>3</sup>Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

## A Appendix: Burst model with lethal mutants

### A.1 Motivation

In this article, infected cells release virions in numbers following a geometric distribution. This is a reasonable assumption for many viruses that release virions by budding, but for bursting viruses, such as lytic bacteriophages, infected cells may release a roughly fixed number of virions. As highlighted by Pearson et al. (2011), the distribution of the number of released virions can influence viral dynamics. In this appendix, we study lethal mutations when infected cells release exactly  $N$  virions.

### A.2 Generating function

Assuming that a virion of strain  $i$  has a probability  $q_i$  to infect a cell, which will then produce  $N_i$  virions, with a probability  $\mu_{ij}$  to mutate from strain  $i$  to strain  $j$ , the general generating functions are:

$$g_i^{all}(\vec{z}) = 1 - q_i + q_i \sum_j \mu_{ij} z_j^{N_i} \text{ for the all-or-none mechanism, and} \quad (1)$$

$$g_i^{ind}(\vec{z}) = 1 - q_i + q_i \left( \sum_j \mu_{ij} z_j \right)^{N_i} \text{ for the independent mechanism.} \quad (2)$$

In the particular case of one strain with lethal mutations only,

$$g_{all}(z) = 1 - q + q((1 - \mu)z^N + \mu), \quad (3)$$

$$g_{ind}(z) = 1 - q + q((1 - \mu)z + \mu)^N. \quad (4)$$

The mean number of viable offspring virions one virion generates is  $R_0 = q(1 - \mu)N$ .

### A.3 Extinction probability

#### A.3.1 General equations

Starting from one virion, using (3) and (4),

$$e_{all} = 1 - q + q((1 - \mu)e_{all}^N + \mu), \text{ and} \quad (5)$$

$$e_{ind} = 1 - q + q((1 - \mu)e_{ind} + \mu)^N. \quad (6)$$

These equations can be solved numerically for any specific set of parameters, however they have no general analytical solution.

#### A.3.2 Relation between the extinction probabilities for the all-or-none and independent mechanisms

To test whether the extinction probability for the independent mechanism with an additional initial round of mutations is equal to the extinction probability for the all-or-none mechanism (Fig.3 main text), we define:

$$e_{ind,mut} = (1 - \mu)e_{ind} + \mu. \quad (7)$$

Using (6),  $e_{ind,mut}$  can be transformed:

$$e_{ind,mut} = (1 - \mu)(1 - q + q((1 - \mu)e_{ind} + \mu)^N) + \mu = 1 - q + q((1 - \mu)e_{ind,mut}^N + \mu). \quad (8)$$

This equation is the same as equation (5), thus the extinction probability in the all-or-none case is equal to the extinction probability in the independent case with an initial round of mutations:

$$e_{all} = (1 - \mu)e_{ind} + \mu. \quad (9)$$

#### A.3.3 Bounds to the extinction probabilities

There is no general solution to the equations for the extinction probabilities, but lower and upper bounds can be found. The extinction probability at generation  $t$  is  $e(t)$ . Because an extinct viral lineage cannot come back to life, a lower bound for  $e(\infty)$  is  $e(1)$ .

For the upper bound,  $e(\infty)$  is the smallest real positive solution of  $g(z) = z$ , and as  $g(0) > 0$ , then  $g'(z = e(\infty)) < 1$ . It leads to  $Nq(1 - \mu)e_{all}^{N-1}(\infty) < 1$ , which combined with (5) leads to

$$e_{all}(\infty) \leq \text{Min} \left( 1, \frac{1 - q(1 - \mu)}{1 - 1/N} \right). \quad (10)$$

Similarly,

$$e_{ind}(\infty) \leq \text{Min} \left( 1, \frac{1 - q + \mu/(N(1 - \mu))}{1 - 1/N} \right). \quad (11)$$

The extinction probability is larger under the budding model than under the burst model: using equation (3) of main text and previous bounds,  $e_{all,budding} = 1 - q(1 - \mu) + 1/N \geq (1 - q(1 - \mu))/(1 - 1/N) \geq e_{all,burst}$ . The same can be shown for  $e_{ind}$ .

### A.3.4 When does extinction occur?

With the all-or-none mechanism,  $e_{all}(1)/e_{all}(\infty) \geq (1 - 1/N)$ . Provided that  $N$  is large, most extinctions occur during the first generation.

With the independent mechanism, the conclusion is less clear:

$$\frac{e_{ind}(1)}{e_{ind}(\infty)} \geq \frac{(1 - q + q\mu^N)(1 - 1/N)}{1 - q + \mu/(N(1 - \mu))} \geq \frac{1 - 1/N}{1 - \mu/(N(1 - q)(1 - \mu))}. \quad (12)$$

However, if  $\mu$  is small,  $N$  is large, and  $1 - q$  is not too small, most extinctions also happen during the first generation.

## A.4 Distribution of the number of virions

Counting viable virions only, it can be shown using recurrence that:

- the mean number of virions at generation  $t$  is  $\langle n(t) \rangle = R_0^t$
- the variance in this number is  $var = R_0^t(R_0^t - 1) \left( \frac{(N-1)A}{R_0-1} - 1 \right)$ , with  $A = 1$  for the all-or-none mechanism and  $A = 1 - \mu$  for the independent mechanism. The variance in the budding model is larger than that in the burst model (see equation (18)).
- the third moment is  $M_3 = R_0^t(R_0^t - 1) \left( \frac{(N-1)A^2}{(R_0-1)(R_0+1)} \left( (N-2)(R_0^t + 1) + 3(N-1) \frac{R_0(R_0^{t-1} - 1)}{R_0-1} \right) - 3(N-1)A \frac{R_0^t - 1}{R_0-1} + 2R_0^t - 1 \right)$ .

If we focus on the distribution conditioned on survival of the viral lineage, we can show that in the large  $t$  limit:

- $\langle n|alive \rangle(t) \propto R_0^t$ , and  $\langle n|alive \rangle_{ind} \simeq (1 - \mu)\langle n|alive \rangle_{all}$
- $var_{alive}(t) \propto R_0^{2t}$ , and  $var_{alive,ind} \simeq (1 - \mu)^2 var_{alive,all}$
- $M_{3,alive}(t) \propto R_0^{3t}$ , and  $M_{3,alive,ind} \simeq (1 - \mu)^3 M_{3,alive,all}$

For the first 3 moments at least, it is as if the surviving lineages had started  $v = -\log(1 - \mu)/\log(R_0)$  generations earlier with the all-or-none mechanism than with the independent mechanism ( $R_0^{-v} = 1 - \mu$ ). However, the distribution of virions with the all-or-none mechanism is non-zero only for multiples of  $N$ , contrary to the independent mechanism.

## A.5 Lethal mutants, burst model: conclusion

Compared to the budding model, the variance and extinction probability are smaller in the burst model. However, the comparison between the all-or-none and independent mechanisms gives the same results for the burst and budding models.  $e_{all} = (1 - \mu)e_{ind} + \mu$ , and, at least up to the third moment and in the large  $t$  limit, the distribution of the number of virions conditioned on survival of the viral lineage is as if the independent had started  $v = -\log(1 - \mu)/\log(R_0)$  generations later than the all-or-none.

## B Appendix: lethal mutants, budding model

In this appendix, we give the details of the calculations leading to the results on lethal mutants with the budding model discussed in the main text.

We assume that there is one strain, which undergoes lethal mutations with probability  $\mu$ , leading to the generating functions:

$$g_{all}(z) = 1 - q + q \left( \frac{1 - \mu}{1 + N(1 - z)} + \mu \right), \text{ and} \quad (13)$$

$$g_{ind}(z) = 1 - q + \frac{q}{1 + N(1 - \mu)(1 - z)}. \quad (14)$$

### B.1 Extinction probability

#### B.1.1 Overall extinction probability

Extinction probabilities are solution of  $g(e) = e$ :

$$e(\infty) = 1 - \frac{R_0 - 1}{AN}, \quad (15)$$

with  $A_{all} = 1$  and  $A_{ind} = 1 - \mu$ . We can check that  $e_{ind}(1 - \mu) + \mu = e_{all}$ .

#### B.1.2 How fast does extinction occur?

From the generating functions (13) and (14) taken at  $z = 0$ ,

$$e(1) = 1 - \frac{R_0}{1 + NA}. \quad (16)$$

We combine this equation with equation (15):

$$\frac{e(1)}{e(\infty)} = \frac{NA}{1 + NA}. \quad (17)$$

Provided that  $R_0 > 1$  (which leads to  $N(1 - \mu) > 1$  and  $N > 1$ ), at least half of the extinctions occur during the first generation, and often a much larger proportion.

## B.2 Distribution of the number of virions

Using recurrence, it can be shown that:

- The mean number of virions at generation  $t$  is  $\langle n(t) \rangle = R_0^t$ . If conditioned on survival,  $\langle n_{alive}(t) \rangle = R_0^t/e(t)$ . In the limit of large  $t$ , when  $e(t) \rightarrow e(\infty)$ , we obtain  $\langle n_{alive,ind}(t + v) \rangle \simeq \langle n_{alive,all}(t) \rangle$ , with  $v = -\log(1 - \mu)/\log(R_0)$ .
- The variance is

$$var = R_0^t(R_0^t - 1) \left( \frac{2NA}{R_0 - 1} - 1 \right). \quad (18)$$

If conditioned on viral survival,  $var_{alive} = \frac{R_0^t}{1 - e} \left( 2NA \frac{R_0^t - 1}{R_0 - 1} + 1 - \frac{R_0^t}{1 - e} \right)$ , which in the large  $t$  limit is  $\simeq \left( \frac{R_0^t NA}{R_0 - 1} \right)^2$ . Again,  $var_{alive,ind}(t + v) \simeq var_{alive,all}(t)$ .

It can also be shown that the distribution of the number of virions ( $n = 0$  excluded) is a geometric distribution for any generation  $t > 0$ .

## C Appendix: Adaptive evolution

A virion of strain  $i$  successfully infects a cell with probability  $q_i$ , which then produces  $N_i$  new virions. The mean reproductive number without mutations is  $R_i = q_i N_i$ . Under the all-or-none mechanism, these new virions are all mutants with probability  $\mu_i$ . Under the independent mechanism, each of these new virions may be a mutant with probability  $\mu_i$ . The generating functions (1) and (2) in the main text become:

$$g_{1,all}(z_1, z_2) = 1 - q_1 + q_1 \left( \frac{1 - \mu_1}{1 + N_1(1 - z_1)} + \frac{\mu_1}{1 + N_1(1 - z_2)} \right), \text{ and} \quad (19)$$

$$g_{1,ind}(z_1, z_2) = 1 - q_1 + \frac{q_1}{1 + N_1(1 - (1 - \mu_1)z_1 - \mu_1 z_2)}. \quad (20)$$

Similar equations can be written for  $g_2$ , the generating function starting from one virion of strain 2.

### C.1 Survival probabilities

#### C.1.1 No general relation between $s_{all}$ and $s_{ind}$

From the argument in the main text (Fig.3 main text), we guess that in some cases  $s_{all,1} = s_{ind,mut,1}$  with the latter defined as  $(1 - \mu_1)s_{ind,1} + \mu_1 s_{ind,2}$ , as verified for the lethal mutants ( $s_2 = 0$ ). Here, using (20):

$$s_{ind,mut,1} = (1 - \mu_1)q_1 + \mu_1 q_2 - \frac{(1 - \mu_1)q_1}{1 + N_1 s_{ind,mut,1}} - \frac{\mu_1 q_2}{1 + N_2 s_{ind,mut,2}}, \quad (21)$$

to compare with (19):

$$s_{all,1} = q_1 - q_1 \left( \frac{1 - \mu_1}{1 + N_1 s_{all,1}} + \frac{\mu_1}{1 + N_1 s_{all,2}} \right). \quad (22)$$

Thus, there is no straightforward relation between  $s_{all}$  and  $s_{ind}$ , except in the special cases of neutral mutations ( $q_1 = q_2$  and  $N_1 = N_2$ ), or lethal mutations ( $s_2 = 0$ ).

#### C.1.2 Analytical solutions for survival probabilities

Using (19) and (20), the survival probabilities starting from a virion of strain  $i$  are solutions of:

$$s_{all,1} = f_{all,1}(s_{all,1}, s_{all,2}) = q_1 - q_1 \left( \frac{1 - \mu_1}{1 + N_1 s_{all,1}} + \frac{\mu_1}{1 + N_1 s_{all,2}} \right) \quad (23)$$

and the symmetric equation (with  $1 \leftrightarrow 2$ ) for the all-or-none mechanism, and

$$s_{ind,1} = f_{ind,1}(s_{ind,1}, s_{ind,2}) = q_1 - \frac{q_1}{1 + N_1((1 - \mu_1)s_{ind,1} + \mu_1 s_{ind,2})} \quad (24)$$

and the symmetric equation (with  $1 \leftrightarrow 2$ , i.e. with indices 1 and 2 interchanged) for the independent mechanism. These systems can be solved by computer algebra systems, but the expressions obtained are very large. Since it is neither easy to find which solution is real and in  $(0, 1]$ , nor to gain insights from the results, we use approximations as detailed below.

### C.1.3 Iterative approximation of the survival probabilities

**Survival probability without mutations** If we neglect mutations, it is as if there were only one strain:

$$s_i^{(0)} = \max \left\{ 0, q_i - \frac{1}{N_i} \right\}. \quad (25)$$

We recover the result from Pearson et al. (2011): when the initial virion successfully infects a cell (probability  $q$ ), it has a probability  $1/R_0 = 1/(qN)$  of extinction (a classic result for geometric distributions (Harris, 1963)).

**Definitions**  $(s_1^*, s_2^*)$  is solution of  $s_1 = f_1(s_1, s_2)$  and  $s_2 = f_2(s_1, s_2)$ . We reformulate these equations as  $s_1 = h_1(s_2)$  and  $s_2 = h_2(s_1)$ . With the all-or-none mechanism:

$$s_{1,all} = h_{1,all}(s_2) = \frac{1}{2} \left( -\frac{1}{N_1} + q_1 - \frac{q_1 \mu_1}{1 + N_1 s_2} + \sqrt{\left( \frac{1}{N_1} - q_1 + \frac{q_1 \mu_1}{1 + N_1 s_2} \right)^2 + \frac{4q_1 \mu_1 s_2}{1 + N_1 s_2}} \right). \quad (26)$$

With the independent mechanism:

$$s_{1,ind} = h_{1,ind}(s_2) = \frac{1}{2} \left( q_1 - \frac{1 + N_1 \mu_1 s_2}{N_1(1 - \mu_1)} + \sqrt{\left( \frac{1 + \mu_1 N_1 s_2}{N_1(1 - \mu_1)} - q_1 \right)^2 + \frac{4q_1 \mu_1 s_2}{1 - \mu_1}} \right). \quad (27)$$

$h_2$  is similarly defined (symmetry  $1 \leftrightarrow 2$ ). In what follows, we study only one strain, but by symmetry all the conclusions are also valid for the other strain.

**Iterative process** The iterative process is shown in figure 1. In the plane  $(s_1, s_2)$ ,  $h_1$  and  $h_2$  intersect for  $(s_1^*, s_2^*)$ , the solution of the system.  $s_2^{(0)}$  is given by the intersection between the  $s_1 = s_2$  curve and  $h_2$ . Indeed, the  $0^{th}$  order does not consider mutations (25), so it is equivalent to having mutational exchange with a strain with the same fitness. Then,  $h_1$  taken at  $s_2 = s_2^{(0)}$  gives  $s_1^{(1)}$ ; then  $h_2$  taken at  $s_1 = s_1^{(1)}$  gives  $s_2^{(2)}$ , and so on.

**Convergence of the iterative approximations** Does this iterative process converge to the solutions  $(s_1^*, s_2^*)$ ? To answer this question, we study properties of  $f$  to deduce properties of  $h$  and ultimately properties of  $h_1(h_2)$ , which we use to prove convergence. All  $f$  and  $g$  functions are  $C^\infty$  for  $s \in [0, 1]$ .

It can be easily checked that  $\forall (s_2, s) \in ([0, 1], [0, 1])$ ,  $f_1(s_2, s) \in [0, 1)$ ,  $\partial f_1 / \partial s_1 > 0$ ,  $\partial f_1 / \partial s_2 > 0$ ,  $\partial^2 f_1 / \partial s_1^2 < 0$ ,  $\partial^2 f_1 / \partial s_2^2 < 0$ ,  $\partial^2 f_1 / (\partial s_1 \partial s_2) \leq 0$ . We additionally need to prove that for  $s_1 \in [h_1(0), h_1(1)]$  (relevant range as  $s_1$  cannot be outside it during the iterative process),  $\partial f_1 / \partial s_1 < 1$ .

$$\frac{\partial f_{1,all}}{\partial s_1} = \frac{(1 - \mu_1) q_1 N_1}{(1 + N_1 s_1)^2}, \text{ and} \quad (28)$$

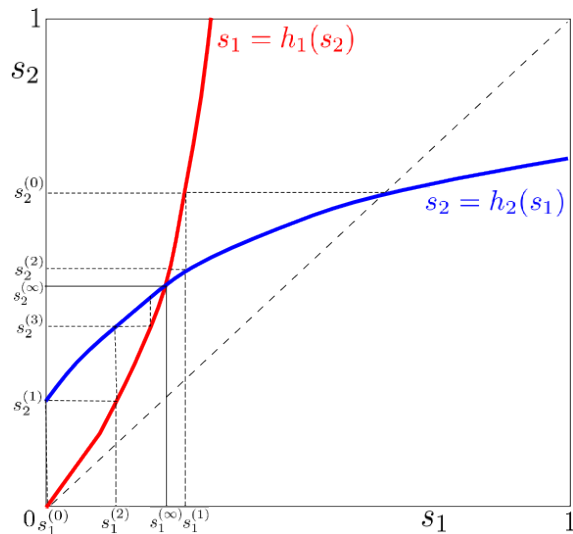


Figure 1: Iterative process for the approximation of the survival probability. Here, the strain 1 is bound to extinction without mutations ( $s_1^{(0)} = 0$ ).

$$\frac{\partial f_{1,ind}}{\partial s_1} = \frac{(1 - \mu_1)q_1 N_1}{(1 + N_1((1 - \mu_1)s_1 + \mu_1 s_2))^2}. \quad (29)$$

If  $(1 - \mu_1)q_1 N_1 < 1$ , then  $\partial f_1 / \partial s_1 < 1$  for any value of  $s_1$  and  $s_2$ . Else, if  $(1 - \mu_1)q_1 N_1 \geq 1$ , we can check that  $\left. \frac{\partial f_1}{\partial s_1} \right|_{s_1=h_1(0)} < 1$ . As  $\partial^2 f_1 / \partial s_1^2 < 0$ , this is enough to prove that  $\partial f_1 / \partial s_1 < 1$  for any  $s_1$  in  $[h_1(0), h_1(1)]$ .

Now, we use the properties of  $f$  to infer the properties of  $h$ . It is clear that  $h_1([0, 1]) \subset [0, 1)$ . As  $s_1 = h_1(s_2)$  and  $s_1 = f_1(s_1, s_2)$ , differentiating these equations with respect to  $s_2$  leads to  $dh_1/ds_2 = \alpha = \partial_{s_2} f_1 / (1 - \partial_{s_1} f_1)$ . We have shown previously that  $\partial_{s_2} f_1 > 0$  and  $\partial_{s_1} f_1 < 1$  (for the pertinent range of  $s_1$ ). As a consequence,  $\frac{dh_1}{ds_2} > 0$ .

With the second derivative we eventually obtain:

$$\frac{d^2 h_1}{ds_2^2} = \frac{\frac{\partial^2 f_1}{\partial s_2^2} + 2\alpha \frac{\partial^2 f_1}{\partial s_2 \partial s_1} + \alpha^2 \frac{\partial^2 f_1}{\partial s_1^2}}{1 - \frac{\partial f_1}{\partial s_1}}. \quad (30)$$

$\alpha$  is positive,  $1 - \frac{\partial f_1}{\partial s_1} > 0$  for  $s_1^{(0)}$ , and the second derivatives of  $f_1$  are negative or zero. As a consequence,  $\frac{d^2 h_1}{ds_2^2} < 0$ .

We now use the properties of  $h$  to deduce the properties of  $h_1(h_2)$ .

As  $h_i([0, 1]) \subset [0, 1)$ ,  $h_1(h_2([0, 1])) \subset [0, 1)$ .

As for the derivative,  $(h_1(h_2))' = h_2' \cdot h_1'(h_2)$ , thus as  $h_1$  and  $h_2$  have positive derivatives,  $h_1(h_2)$  has also a positive derivative.

For the second derivative,  $(h_1(h_2))'' = h_2'' \cdot h_1'(h_2) + (h_2')^2 \cdot h_1''(h_2)$ : the derivatives of  $h_i$  are positive, and the second derivatives are negative, so the second derivative of  $h_1(h_2)$  is negative.

The iterative approximation leads to  $s_1^{(j+2)} = h_1(h_2(s_1^{(j)}))$ . If  $h_1(h_2(0)) > 0$ , as  $h_1(h_2(1)) < 1$ , the derivative is positive and the second derivative is negative,

there is at one fixed point in  $(0, 1)$ , and starting the iteration from any  $s_1 \in (0, 1)$  converges to it. The same reasoning can be made for  $s_2$ .

If  $h_1(h_2(0)) = 0$ , the procedure for  $s_1^{(2k)}$  and  $s_2^{(2k+1)}$  starts on the fixed point  $(0, 0)$ , and stays to this point. In this case, the procedure should be done for  $s_1^{(2k+1)}$  and  $s_2^{(2k)}$  only.

#### C.1.4 Approximation of the survival probabilities in the limit of small mutation rates

Another approximation is to take both mutation rates small. We derive this approximation in the general case of  $n$  different strains. The generating map is  $G(z) = (G_1(z), \dots, G_n(z))$  where  $z = (z_1, \dots, z_n)$ . We define  $F_i = G_i(1 - s) - (1 - s_i)$ , where  $1 - s = (1 - s_1, \dots, 1 - s_n)$ , and obtain:

$$F_i^{all}(s, \mu) = s_i + q_i \left( \frac{1 - \sum_{j \neq i} \mu_{ij}}{1 + N_i s_i} + \sum_{j \neq i} \frac{\mu_{ij}}{1 + N_i s_j} - 1 \right) \quad \text{and} \quad (31)$$

$$F_i^{ind}(s, \mu) = s_i + q_i \left( \frac{1}{1 + N_i \left( (1 - \sum_{j \neq i} \mu_{ij}) s_i + \sum_{j \neq i} \mu_{ij} s_j \right)} - 1 \right). \quad (32)$$

Let  $\hat{s} = \hat{s}(\mu)$  denote the survival probabilities which satisfy  $F_i(\hat{s}, \mu) = 0$  for all  $i = 1, 2, \dots, n$ . To write down a first order approximation of  $\hat{s}_i$  with respect to  $\mu$ , we begin by implicitly differentiating  $F_i(\hat{s}, \mu) = 0$  with respect to  $\mu_{ij}$  for  $j \neq i$ .

$$\frac{\partial F_i}{\partial \mu_{ij}} + \sum_{k=1}^n \frac{\partial F_i}{\partial s_k} \frac{\partial \hat{s}_k}{\partial \mu_{ij}} = 0 \quad (33)$$

and then evaluate these expressions at  $\mu = 0$  and  $\hat{s}(0)$ . Since  $\frac{\partial F_i}{\partial s_j}$  equals zero at  $\mu = 0$  for all  $j \neq i$ , equation (33) simplifies to

$$\frac{\partial F_i}{\partial \mu_{ij}} + \frac{\partial F_i}{\partial s_i} \frac{\partial \hat{s}_i}{\partial \mu_{ij}} = 0 \quad (34)$$

at  $\mu = 0$  and  $s = \hat{s}(0)$ . Solving for the derivatives evaluated at  $\mu = 0$  and  $\hat{s} = \hat{s}(0)$  yields

$$\frac{\partial F_i^{all}}{\partial \mu_{ij}} = q_i \left( \frac{1}{1 + N_i \hat{s}_j} - \frac{1}{1 + N_i \hat{s}_i} \right), \quad (35)$$

$$\frac{\partial F_i^{ind}}{\partial \mu_{ij}} = q_i N_i \frac{(\hat{s}_i - \hat{s}_j)}{(1 + N_i \hat{s}_i)^2}, \quad \text{and} \quad (36)$$

$$\frac{\partial F_i}{\partial s_i} = 1 - \frac{q_i N_i}{(1 + N_i \hat{s}_i)^2}. \quad (37)$$

Next, we implicitly differentiate  $F_i(\hat{s}, \mu) = 0$  with respect to  $\mu_{jk}$  for  $j \neq i$  and  $k \neq j$ :

$$\frac{\partial F_i}{\partial \mu_{jk}} + \sum_{l=1}^n \frac{\partial F_i}{\partial s_l} \frac{\partial \hat{s}_l}{\partial \mu_{jk}} = 0, \quad (38)$$



and evaluate these expressions at  $\mu = 0$  and  $\hat{s}(0)$ . Since  $\frac{\partial F_i}{\partial s_i}$  equals zero at  $\mu = 0$  for all  $j \neq i$  and  $\frac{\partial F_i}{\partial \mu_{jk}} = 0$ , equation (38) simplifies to

$$\frac{\partial F_i}{\partial s_i} \frac{\partial \hat{s}_i}{\partial \mu_{jk}} = 0 \quad (39)$$

at  $\mu = 0$  and  $s = \hat{s}(0)$ .

Since  $1 + N_i \hat{s}_i(0) = 1$  if  $q_i N_i < 1$  and  $q_i N_i$  otherwise, it follows from equation (37) that  $\frac{\partial F_i}{\partial s_i} > 0$  at  $\mu = 0$ . Hence, (39) implies that  $\frac{\partial \hat{s}_i}{\partial \mu_{jk}} = 0$  for  $j \neq i$  and  $k \neq j$ . On the other hand, solving (34) yields

$$\frac{\partial \hat{s}_i}{\partial \mu_{ij}}(0) = -\frac{\frac{\partial F_i}{\partial \mu_{ij}}}{\frac{\partial F_i}{\partial s_i}} \quad (40)$$

where

$$\hat{s}_i = \hat{s}_i(0, 0) = (q_i - 1/N_i)^+ \quad (41)$$

and  $x^+ = \max\{0, x\}$ . Hence, the first order Taylor approximation of  $\hat{s}_i(\mu)$  is:

$$(q_i - 1/N_i)^+ + \sum_{j \neq i} \mu_{ij} \frac{q_i \left( \frac{1}{1+(q_i N_i - 1)^+} - \frac{1}{1+N_i(q_j - 1/N_j)^+} \right)}{1 - \frac{q_i N_i}{(1+(q_i N_i - 1)^+)^2}} \quad (42)$$

for the all-or-none mechanism, and:

$$(q_i - 1/N_i)^+ + \sum_{j \neq i} \mu_{ij} \frac{q_i N_i ((q_j - 1/N_j)^+ - (q_i - 1/N_i)^+)}{((q_i N_i - 1)^+ + 1)^2 - q_i N_i} \quad (43)$$

for the independent mechanism.

In the special case of evolutionary escape, say  $\hat{s}_1(0) = 0$  and  $\hat{s}_2(0) > 0$  and  $n = 2$ , these first order approximations simplify to

$$\hat{s}_1^{all} \approx \mu_1 \frac{q_1 N_1}{1 - N_1 q_1} \frac{q_2 - 1/N_2}{1 + N_1(q_2 - 1/N_2)}, \quad (44)$$

$$\hat{s}_2^{all} \approx q_2 - \frac{1}{N_2} - \mu_2 q_2, \quad (45)$$

$$\hat{s}_1^{ind} \approx \mu_1 \frac{q_1 N_1}{1 - N_1 q_1} \left( q_2 - \frac{1}{N_2} \right) \text{ and} \quad (46)$$

$$\hat{s}_2^{ind} \approx q_2 - \frac{1}{N_2} - \frac{\mu_2}{N_2}. \quad (47)$$

### C.1.5 Comparison of the approximations

It can be shown that the iterative approximation  $s_1^{(1)}$  taken in the limit  $\mu_1$  small results in the same expressions as the Taylor expansion when  $\mu$  is small. When  $R_1$  is not too close to 1, iterative  $s_i^{(1)}$  and small  $\mu$  approximations work almost equally well (Fig. 2). However,  $s_i^{(1)}$  is more suitable on a larger range of parameters.

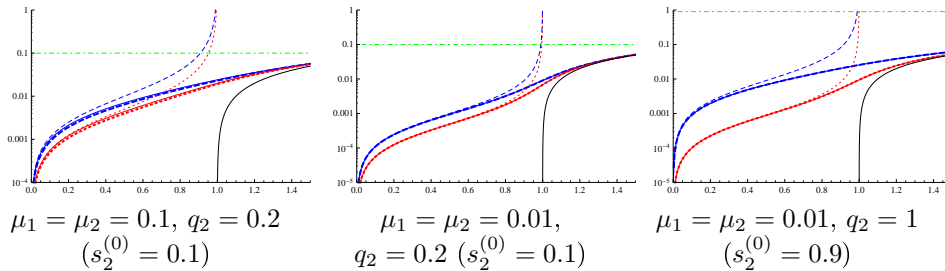


Figure 2: Two strains, evolutionary escape. Survival probabilities as a function of  $R_1 = N_1 q_1$ , with  $q_1$  variable and  $N_1$  fixed. All-or-none (red, dotted lines) and independent (blue, dashed lines).  $s_2^{(0)}$  (green dot-dashed line),  $s_1^{(0)}$  (black solid line). Exact expression for  $s_1$  (thick dotted or dashed lines), approximation  $s_1^{(1)}$  (solid lines), further approximation with  $\mu_1$  small (equations 44 and 46) (thin dotted or dashed lines).  $N_1 = N_2 = 10$ .

### C.1.6 The independent mechanism results in more survival of the viral lineage than the all-or-none mechanism

Consider the case of  $n$  strains. Let  $c_i(z)$  be the generating function for the number of virions released by a cell infected with a virion of type  $i$ . To avoid degenerate cases, we assume that  $c_i(0) > 0$  (i.e. there is a positive probability of no offspring) and  $\frac{d^k c_i}{dz^k}(0) > 0$  for some  $k \geq 2$  (i.e. there is a positive probability of at least two offspring). Let  $\mu_{ij}$  be the probability that a virion of type  $i$  produces a virion of type  $j$ . Note that  $\sum_j \mu_{ij} = 1$ . For the case of all-or-none-mutations, the generating map  $G_{all}(z) = (G_{all,1}(z), \dots, G_{all,n}(z))$  where  $z = (z_1, \dots, z_n)$  is given by

$$G_{all,i}(z) = 1 - q_i + q_i \sum_j \mu_{ij} c_j(z_j). \quad (48)$$

In the case of independent-mutations, the generating map  $G_{ind}(z)$  is given by

$$G_{ind,i}(z) = 1 - q_i + q_i c_i \left( \sum_j \mu_{ij} z_j \right). \quad (49)$$

Since  $c_i$  are generating maps, they are given by power series with all positive coefficients and, consequently, they are convex functions. For  $x, y \in \mathbb{R}^n$ , we write  $x \geq y$  if  $x_i \geq y_i$  for all  $i$ ,  $x > y$  if  $x \geq y$  and  $x_i > y_i$  for some  $i$ , and  $x \gg y$  if  $x_i > y_i$  for all  $i$ . By Jensen's inequality  $G_{all,i}(z) \geq G_{ind,i}(z)$  for all  $z \geq 0$  and  $i$ . This inequality is strict provided that  $z_i \neq z_j$  for some  $i, j$ . Furthermore, the fact that these generating maps correspond to power series with positive coefficients implies that  $G_{all}$  and  $G_{ind}$  are monotone maps i.e.  $G(x) \geq G(y)$  if  $x \geq y$ . Moreover, if the mutation matrix  $(\mu_{ij})$  is primitive, then  $G^n(x) \gg G^n(y)$  whenever  $x > y$ . Here,  $G^t$  denotes composing the map  $G$  with itself  $t$  times.

Let  $e_{all,i}$  and  $e_{ind,i}$  be the probabilities of extinction for a population initiated with one individual of type  $i$  with all-or-none-mutations or independent-mutations, respectively. Let  $e_{all} = (e_{all,1}, \dots, e_{all,n})$  and  $e_{ind} = (e_{ind,1}, \dots, e_{ind,n})$ . Standard branching process theory (Harris, 1963) implies that  $\lim_{t \rightarrow \infty} G_{all}^t(0) =$

$e_{all}$  and  $\lim_{t \rightarrow \infty} G_{ind}^t(0) = e_{ind}$ . We claim that  $G_{all}^t(0) \geq G_{ind}^t(0)$  for all  $t$ . Clearly  $G_{ind}(0) = G_{all}(0)$ . Assume  $G_{all}^t(0) \geq G_{ind}^t(0)$ . Then, by the aforementioned properties of the generating maps,

$$\begin{aligned} G_{all}^{t+1}(0) &= G_{all}(G_{all}^t(0)) \\ &\geq G_{all}(G_{ind}^t(0)) \\ &\geq G_{ind}(G_{ind}^t(0)) = G_{ind}^{t+1}(0). \end{aligned}$$

Hence, induction implies the claim and we have shown that  $e_{all} \geq e_{ind}$ . When the  $c_i$  are different and the mutation matrix is primitive, it can also be shown that this inequality is strict as  $G_{all}^2(0) \gg G_{ind}^2(0)$ .

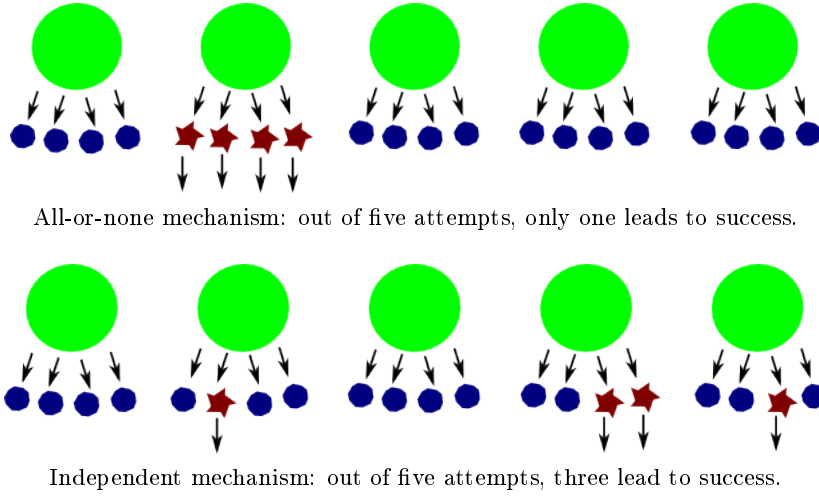


Figure 3: For a given mean mutation rate, the independent mechanism is more likely to lead to survival than the all-or-none mechanism. Consider the following simplified situation. In five separate instances, a very unfit virus (depicted by a blue polygon) infects a host cell (represented by large green disks), producing each time 4 new virions. With a probability of 20%, it mutates, and this mutant (depicted by a red star) has a very high survival probability. Clustered mutations lead to less overall survival.

An heuristic explanation for the independent mechanism leading to more survival is detailed in figure 3.

## C.2 Distribution of the number of virions

### C.2.1 Mean number of virions

We define  $\langle n_i^{(j)}(t) \rangle$  the mean number of virions  $i$  at generation  $t$ , starting from a particle  $j$  at generation 0;  $\alpha_i = q_i N_i (1 - \mu_i)$  and  $\beta_i = q_i N_i \mu_i$  the mean number of virions produced in one cycle by a virion of strain  $i$ , of the same strain ( $\alpha$ ), and of the other strain ( $\beta$ ). This leads to:

$$N(t) = \begin{pmatrix} \langle n_1^{(j)}(t) \rangle \\ \langle n_2^{(j)}(t) \rangle \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_2 \\ \beta_1 & \alpha_2 \end{pmatrix}^t \begin{pmatrix} \delta_{j,1} \\ \delta_{j,2} \end{pmatrix}. \quad (50)$$

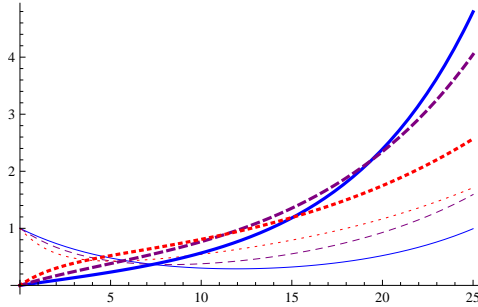


Figure 4: Two strains, influence of the mutation rate on the mean number of virions.  $\langle n_1(t) \rangle$  (thin lines) and  $\langle n_2(t) \rangle$  (thick lines) as a function of generation  $t$ , from (53). Here the dynamics start at generation  $t=0$  with one viral particle of strain 1.  $\mu_1 = \mu_2 = \mu$ ,  $\mu = 0.05$  (blue, solid lines),  $\mu = 0.1$  (purple, dashed lines),  $\mu = 0.2$  (red, dotted lines).  $q_1 N_1 = 0.9$ ,  $q_2 N_2 = 1.2$ .

Defining  $r = \sqrt{(\alpha_1 - \alpha_2)^2 + 4\beta_1\beta_2}$ , the eigenvalues are:

$$\lambda_{\pm} = \frac{\alpha_1 + \alpha_2 \pm r}{2}, \text{ and their associated eigenvectors} \quad (51)$$

$$\vec{V}_{\pm} = \begin{pmatrix} \frac{\alpha_1 - \alpha_2 \pm r}{2\beta_1} \\ 1 \end{pmatrix}. \quad (52)$$

Starting with one particle of strain 1,  $(1, 0) = c_+ \vec{V}_+ + c_- \vec{V}_-$ , with  $c_{\pm} = \pm\beta_1/r$ . At generation  $t$ ,  $N(t) = c_+ \vec{V}_+ \lambda_+^t + c_- \vec{V}_- \lambda_-^t$ , which can also be written:

$$\begin{pmatrix} \langle n_1(t) \rangle \\ \langle n_2(t) \rangle \end{pmatrix} = \frac{1}{\lambda_+ - \lambda_-} \begin{pmatrix} ((\lambda_+ - \alpha_2)\lambda_+^t - (\lambda_- - \alpha_2)\lambda_-^t) \\ \beta_1 (\lambda_+^t - \lambda_-^t) \end{pmatrix}. \quad (53)$$

If strain 1 is less fit than strain 2, a larger mutation rate means that a beneficial mutation happens earlier, leading to a faster adaptation. However, it also means that the long-term mutation-selection balance is shifted towards more mutations, resulting in a slower long-term growth (Fig. 4).

### C.2.2 Variance of the number of virions

Defining  $var(i, n_j, t)$  as the variance of the number of virions of type  $j$ , and  $cov(i, n_k, n_j, t)$  the covariance between the number of virions of type  $k$  and  $j$ , at generation  $t$  starting from one virion of type  $i$ ;  $A_{i,all} = 1$ ,  $A_{i,ind} = 1 - \mu_i$ ,  $B_{i,all} = 1$ ,  $B_{i,ind} = \mu_i$ ,  $C_{all} = 0$ ,  $C_{ind} = 1$ ,

$$V(t) = \begin{pmatrix} var(1, n_1, t) \\ cov(1, n_1, n_2, t) \\ var(1, n_2, t) \end{pmatrix}, \quad (54)$$

$$M = \begin{pmatrix} \alpha_1^2 & 2\alpha_1\beta_2 & \beta_2^2 \\ \alpha_1\beta_1 & \alpha_1\alpha_2 + \beta_1\beta_2 & \alpha_2\beta_2 \\ \beta_1^2 & 2\alpha_2\beta_1 & \alpha_2^2 \end{pmatrix}, \text{ and} \quad (55)$$

$$G = \begin{pmatrix} \alpha_1(2N_1A_1 + 1 - \alpha_1) & \beta_2(2N_2B_2 + 1 - \beta_2) \\ \alpha_1\beta_1 \left( \frac{2CN_1}{\alpha_1 + \beta_1} - 1 \right) & \alpha_2\beta_2 \left( \frac{2CN_2}{\alpha_2 + \beta_2} - 1 \right) \\ \beta_1(2N_1B_1 + 1 - \beta_1) & \alpha_2(2N_2A_2 + 1 - \alpha_2) \end{pmatrix}, \quad (56)$$

it can be shown using the derivatives of the generating function that:

$$V(t+1) = MV(t) + GN(t), \text{ leading to} \quad (57)$$

$$V(t) = \sum_{i=0}^{t-1} M^{t-1-i} GN(i) \quad (58)$$

since  $V(0) = 0$ . For  $M$ , eigenvalues are:

$$\lambda_0 = \alpha_1\alpha_2 - \beta_1\beta_2, \quad (59)$$

$$\lambda_1 = \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + 2\beta_1\beta_2 - \alpha_1r - \alpha_2r) = \lambda_-^2, \text{ and} \quad (60)$$

$$\lambda_2 = \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + 2\beta_1\beta_2 + \alpha_1r + \alpha_2r) = \lambda_+^2. \quad (61)$$

$\lambda_2$  is the largest eigenvalue. Indeed,  $\lambda_2 \geq \lambda_1$  and  $\lambda_2 \geq \lambda_0$  (except if  $\alpha = \alpha_2 = 0$ ; or  $\alpha = \alpha_2$  with  $\beta_1 = 0$  or  $\beta_2 = 0$ : these regimes are not relevant to our study).  $\lambda_2$  is the square of  $\lambda_+$ , the largest eigenvalue for the mean number of virions: in the long time limit, the variances and covariances scale as the square of the mean number of virions. As in the lethal mutant case, the variance is relatively large (the standard deviation scales as the mean). The eigenvectors are:

$$v_0 = \begin{pmatrix} -2\beta_2 \\ \alpha_1 - \alpha_2 \\ 2\beta_1 \end{pmatrix}, \quad (62)$$

$$v_1 = \frac{1}{2\beta_1^2(\alpha_1 + \alpha_2 - r)} \begin{pmatrix} \alpha_1r(-\alpha_1 + \alpha_2 + r) - \beta_1\beta_2(\alpha_1 + \alpha_2 + r) \\ \beta_1(2\beta_1\beta_2 - \alpha_1(-\alpha_1 + \alpha_2 + r)) \\ \beta_1^2(\alpha_1 + \alpha_2 - r) \end{pmatrix}, \text{ and} \quad (63)$$

$$v_2 = \frac{1}{4\beta_1^2(\alpha_1\alpha_2 - \beta_1\beta_2)} \begin{pmatrix} \alpha_1r(\alpha_1 - \alpha_2 + r) - \beta_1\beta_2(\alpha_1 + \alpha_2 - r) \\ \beta_1(2\beta_1\beta_2 - \alpha_1(-\alpha_1 + \alpha_2 - r)) \\ \beta_1^2(\alpha_1 + \alpha_2 + r) \end{pmatrix}. \quad (64)$$

Projecting on the eigenvectors eventually leads to:

$$V(t) = \frac{\beta_1}{(\lambda_+ - \lambda_-)r^2} \times \left[ v_0 \left\{ \left( \frac{\lambda_0^t - \lambda_+^t}{\lambda_0 - \lambda_+} - \frac{\lambda_0^t - \lambda_-^t}{\lambda_0 - \lambda_-} \right) f_0' + \left( \frac{\alpha_1 - \alpha_2 + r}{2} \frac{\lambda_0^t - \lambda_+^t}{\lambda_0 - \lambda_+} - \frac{\alpha_1 - \alpha_2 - r}{2} \frac{\lambda_0^t - \lambda_-^t}{\lambda_0 - \lambda_-} \right) f_0'' \right\} \right. \\ + v_1 \left\{ \left( \frac{\lambda_1^t - \lambda_+^t}{\lambda_1 - \lambda_+} - \frac{\lambda_1^t - \lambda_-^t}{\lambda_1 - \lambda_-} \right) f_1' + \left( \frac{\alpha_1 - \alpha_2 + r}{2} \frac{\lambda_1^t - \lambda_+^t}{\lambda_1 - \lambda_+} - \frac{\alpha_1 - \alpha_2 - r}{2} \frac{\lambda_1^t - \lambda_-^t}{\lambda_1 - \lambda_-} \right) f_1'' \right\} \\ \left. + v_2 \left\{ \left( \frac{\lambda_2^t - \lambda_+^t}{\lambda_2 - \lambda_+} - \frac{\lambda_2^t - \lambda_-^t}{\lambda_2 - \lambda_-} \right) f_2' + \left( \frac{\alpha_1 - \alpha_2 + r}{2} \frac{\lambda_2^t - \lambda_+^t}{\lambda_2 - \lambda_+} - \frac{\alpha_1 - \alpha_2 - r}{2} \frac{\lambda_2^t - \lambda_-^t}{\lambda_2 - \lambda_-} \right) f_2'' \right\} \right], \quad (65)$$

with:

$$f_0' = -\beta_1G_{12} + (\alpha_1 - \alpha_2)G_{22} + \beta_2G_{32} \quad (66)$$

$$f_0'' = \frac{1}{\beta_1} (-\beta_1 G_{11} + (\alpha_1 - \alpha_2)G_{21} + \beta_2 G_{31}) \quad (67)$$

$$f_1' = 2\beta_1 \left( \beta_1 G_{12} - G_{22}(\alpha_1 - \alpha_2 + r) + G_{32} \left( -\beta_2 + \frac{r(\alpha_1 - \alpha_2 + r)}{2\beta_1} \right) \right) \quad (68)$$

$$f_1'' = 2 \left( \beta_1 G_{11} - G_{21}(\alpha_1 - \alpha_2 + r) + G_{31} \left( -\beta_2 + \frac{r(\alpha_1 - \alpha_2 + r)}{2\beta_1} \right) \right) \quad (69)$$

$$f_2' = \beta_1 \left( \beta_1 G_{12}(\alpha_1 + \alpha_2 - r) - 2\beta_2 G_{22} \left( 2\beta_1 + \frac{\alpha_1}{\beta_2}(\alpha_1 - \alpha_2 - r) \right) - G_{32} \left( \frac{\alpha_1}{\beta_1} r(\alpha_1 - \alpha_2 - r) + \beta_2(\alpha_1 + \alpha_2 + r) \right) \right) \quad (70)$$

$$f_2'' = \beta_1 G_{11}(\alpha_1 + \alpha_2 - r) - 2\beta_2 G_{21} \left( 2\beta_1 + \frac{\alpha_1}{\beta_2}(\alpha_1 - \alpha_2 - r) \right) - G_{31} \left( \frac{\alpha_1}{\beta_1} r(\alpha_1 - \alpha_2 - r) + \beta_2(\alpha_1 + \alpha_2 + r) \right) \quad (71)$$

**Large  $t$  limit** We now study this result in the large  $t$  limit.  $\lambda_2$  is larger than  $\lambda_0$  and  $\lambda_1$ .  $\lambda_- \leq \lambda_+$ , and  $\lambda_+^2 = \lambda_2$ . We focus here on the case where virus may survive, i.e.  $\lambda_+ > 1$ . Then  $\lambda_2 = \lambda_+^2 > \lambda_+$ . So now we simplify previous results in the large  $t$  limit, keeping only the leading terms in  $\lambda_2^t$ .

$$V(t) \simeq \frac{\beta_1 \lambda_2^t v_2}{r^2 (\lambda_2 - \lambda_-)(\lambda_2 - \lambda_+)} (2f_2' + (\lambda_2 - \alpha_2)f_2''). \quad (72)$$

In this limit, the ratio between the variances/covariances in the all-or-none *vs.* independent is:

$$\frac{V^{all}}{V^{ind}} \simeq \frac{2f_2'^{all} + (\lambda_2 - \alpha_2)f_2''^{all}}{2f_2'^{ind} + (\lambda_2 - \alpha_2)f_2''^{ind}}. \quad (73)$$

We now study this ratio to show that in the large  $t$  limit,  $var_{i,all} > var_{i,ind}$ .

First,  $\lambda_2 \geq \alpha_2$ . Indeed,  $\lambda_2 = \lambda_+^2$ ,  $\lambda_+ = (\alpha_1 + \alpha_2 + r)/2$ . If  $\alpha_1 \leq \alpha_2$ ,  $r \geq \alpha_2 - \alpha_1$ , leading to  $\lambda_+ \geq \alpha_2$ . Then  $\lambda_2 \geq \lambda_+ \alpha_2 \geq \alpha_2$ , because we are interested in the situation where the virus may survive, i.e.  $\lambda_+ > 1$ . If  $\alpha_1 \geq \alpha_2$ ,  $r \geq \alpha_1 - \alpha_2$ , leading to  $\lambda_+ \geq \alpha_1 \geq \alpha_2$ . Then  $\lambda_2 \geq \lambda_+ \alpha_2 \geq \alpha_2$ .

It can be easily checked that  $G_{1j}^{all} \geq G_{1j}^{ind} \geq 0$ ,  $G_{3j}^{all} \geq G_{3j}^{ind} \geq 0$ , and  $G_{2j}^{all} \leq 0 \leq G_{2j}^{ind}$ .

The coefficients of  $G_{1j}$  in  $f_2'$  and  $f_2''$  are of the sign of  $\alpha_1 + \alpha_2 - r$ .  $r = \sqrt{((1 - \mu_1)R_1 - (1 - \mu_2)R_2)^2 + 4\mu_1\mu_2R_1R_2}$ . Assuming  $\mu_1 \leq 0.5$ ,  $4\mu_1\mu_2R_1R_2 \leq 4(1 - \mu_1)(1 - \mu_2)R_1R_2$ , leading to  $r \leq ((1 - \mu_1)R_1 + (1 - \mu_2)R_2)$ , thus  $\alpha_1 + \alpha_2 - r \geq 0$ . The coefficients of  $G_{1j}$  in  $f_2'$  and  $f_2''$  are positive, and as  $G_{1j}^{all} \geq G_{1j}^{ind}$ , so we conclude that these terms are larger with the all-or-none mechanism than with independent mechanism.

The coefficients of  $G_{2j}$  in  $f_2'$  and  $f_2''$  are opposite of the sign of  $2\beta_1\beta_2 + \alpha_1(\alpha_1 - \alpha_2 - r)$ . Fixing  $a = \mu_2/\mu_1$ , this expression is positive for  $\mu_1$  between 0 and  $\mu_c = (3R_1 + R_2(1 + a) - \sqrt{(3R_1 + R_2(1 + a))^2 - 8R_1(R_1 + R_2)})/(4R_1)$ . For  $a = 1$  (i.e.  $\mu_1 = \mu_2$ ),  $\mu_c = 1/2$ .  $\partial\mu_c/\partial a < 0$ : The higher  $\mu_2$  compared to  $\mu_1$ , the smaller  $\mu_c$ . As  $G_{2j}^{all} \leq 0 \leq G_{2j}^{ind}$ ,  $\mu_1 < \mu_c$  ensures that these terms are larger with the all-or-none mechanism than with the independent mechanism.

The coefficients of  $G_{3j}$  in  $f_2'$  and  $f_2''$  are opposite of the sign of  $\alpha_1 r(\alpha_1 \alpha_2 r) + \beta_1 \beta_2(\alpha_1 + \alpha_2 + r)$ . This expression is negative for  $0 < \mu_1 < 1/(1 + \mu_2/\mu_1)$ . As  $G_{3j}^{all} \geq G_{3j}^{ind}$ , these terms are larger with the all-or-none mechanism than with the independent mechanism.

**Conclusion on the variance** We have proved that in the long time limit the variance is larger in the all-or-none than in the independent case, provided that the mutation rate is small enough (e.g. when  $\mu_1 = \mu_2$ , a sufficient condition is  $\mu_i \leq 0.5$ ). We have not proven that the variance is larger with the all-or-none mechanism than with the independent mechanism for any generation  $t$ , only for the long time limit, but numerically, we see no instance of the contrary.

### C.3 Extension to multi-step evolutionary trajectories

We have discussed a simplified model where one mutation is enough for adaptation. In many real systems, however, several mutations are needed for a virus to adapt to some challenge (Shih et al., 2007; Bloom et al., 2010). Consider a network of genotypes where each node is a strain and each edge represents one mutational step. For low mutation rates, the survival probability of fit strains (i.e. those with  $R = qN > 1$ ) is relatively unchanged by mutations. Then the survival probabilities of their unfit ( $R < 1$ ) neighbors can be approximated for low mutation rates by adding the contributions to survival resulting from the mutations to each of the neighboring fit strains (see Iwasa et al. (2004) and Weissman et al. (2009) for a more detailed discussion of this approach). Then, survival probabilities of the next-nearest neighbors are considered, and so on.

Heuristic arguments help us to focus on which steps in a multi-step trajectory will be most influenced by replication mechanism. It is clear from arguments presented above, and in figure 3, that the survival probability of the viral lineage for a mutational step from an unfit strain to a fit strain differs markedly between the all-or-none and independent mechanisms. In contrast, survival of fit strains is not strongly affected by mutations and consequently is insensitive to the replication mechanism. Steps between unfit strains will also be relatively insensitive to replication mechanism, by the following logic. Consider a strain  $i$  with several neighbors, none of them fit, and among which strain  $j$  is closest (i.e. the smallest number of mutational steps from) a fit strain. For sufficiently low mutation rates, the dominant contribution to  $s_i$  is from mutations to  $j$ . Then the ratio between the survival probabilities under the two mechanisms,  $s_i^{ind}/s_i^{all}$ , tends to  $1 + N_i s_j$ . Because  $j$  is also an unfit strain, its survival probability is of the order of  $\mu$  or smaller, so  $N_i s_j$  is  $\leq N_i \mu$ , which in this limit of small mutation rates is much less than 1. Thus  $s_i^{ind}/s_i^{all}$  is not very different from one. This can be understood using a variation of the argument shown in figure 3: if the mutant leads to survival only 20% of the time, there is no more clustering of escape mutants in the all-or-none mechanism, so the all-or-none and independent mechanisms have approximately the same probability of escape. Combining these arguments to consider multi-step trajectories in general, we see that the independent mechanism will still lead to more viral survival than the all-or-none mechanism, but the ratio of the survival probabilities will be dominated by the steps from unfit to fit strains. The two-strain case presented above provides the essential building block to explore more complex evolutionary trajectories.

## References

Bloom, J. D., L. I. Gong, and D. Baltimore, 2010. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* (New

- York, N.Y.) 328:1272–5.
- Harris, T. E., 1963. The theory of branching processes. Dover phoenix editions.
- Iwasa, Y., F. Michor, and M. A. Nowak, 2004. Evolutionary dynamics of invasion and escape. *Journal of Theoretical Biology* 226:205–214.
- Pearson, J. E., P. Krapivsky, and A. S. Perelson, 2011. Stochastic Theory of Early Viral Infection: Continuous versus Burst Production of Virions. *PLoS Computational Biology* 7:e1001058.
- Shih, A. C.-C., T.-C. Hsiao, M.-S. Ho, and W.-H. Li, 2007. Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104:6283–8.
- Weissman, D. B., M. M. Desai, D. S. Fisher, and M. W. Feldman, 2009. The rate at which asexual populations cross fitness valleys. *Theoretical population biology* 75:286–300.