Fixation probabilities on superstars, revisited and revised

Alastair Jamieson-Lane[a,*], Christoph Hauert[a]

[a]*Mathematics Department, University of British Columbia, 1984 Mathematics Road, Vancouver BC V6T 1Z2, Canada*

Abstract

Population structures can be crucial determinants of evolutionary processes. For the Moran process on graphs certain structures suppress selective pressure, while others amplify it (Lieberman *et al.* 2005 *Nature* **433** 312-316). Evolutionary amplifiers suppress random drift and enhance selection. Recently, some results for the most powerful known evolutionary amplifier, the superstar, have been invalidated by a counter example (Díaz *et al.* 2013 *Proc. R. Soc. A* **469** 20130193). Here we correct the original proof and derive improved upper and lower bounds, which indicate that the fixation probability remains close to $1 - 1/(r^4 H)$ for population size $N \to \infty$ and structural parameter $H \gg 1$. This correction resolves the differences between the two aforementioned papers. We also confirm that in the limit $N, H \to \infty$ superstars remain capable of eliminating random drift and hence of providing arbitrarily strong selective advantages to any beneficial mutation. In addition, we investigate the robustness of amplification in superstars and find that it appears to be a fragile phenomenon with respect to changes in the selection or mutation processes.

Keywords: evolution, Moran process, evolutionary graph theory, structured populations

1. Introduction

Populations evolve according to the principles of natural selection and random drift. The balance between the two competing processes is determined by numerous factors, including both population size and structure (1; 2; 3; 4). The most malignant tumour is unlikely to cause harm if it arises in the outermost layer of skin and is easily brushed aside, and the most imperative model for climate change has limited influence until it has worked its way from a researchers desk, through the literature into policy making and public awareness. Position matters.

*Corresponding author
Email addresses: aja107@math.ubc.ca (Alastair Jamieson-Lane),
christoph.hauert@math.ubc.ca (Christoph Hauert)

One of the simplest and most influential models of stochastic evolutionary processes in finite populations is the Moran process (5; 6). It is based on an unstructured (or well-mixed) population of constant size N, where each individual is classified either as a resident (wild type) or mutant. Each type is assigned a constant fitness, which determines its propensity to reproduce. The fitness of wild types is normalized to 1 and mutants have fitness r. An advantageous mutant has $r > 1$, a disadvantageous mutant has $r < 1$ and a neutral mutant is indistinguishable in terms of fitness, $r = 1$. In every time step, an individual is randomly selected for reproduction with a probability proportional to its fitness and produces a clonal offspring that replaces an individual, selected uniformly at random, in the population. This process is repeated until eventually the population has reached one of the homogenous states of all residents, if the mutant went extinct, or all mutants, if the mutant successfully took over the entire population (5; 6; 7). In both cases, the population has reached fixation. In the absence of mutation, the two homogenous states are absorbing.

The Moran process models evolutionary dynamics based on selection and random drift in finite populations: an advantageous mutant has a higher probability, but no guarantee, to reach fixation and, similarly, an inferior mutant is more likely to be eliminated, but not with certainty. The fixation probability of either type is analytically accessible for any given initial configuration. Of particular interest is the fixation probability of a single mutant, ρ_M, that arises in an otherwise homogenous population of wild types:

$$\rho_M = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}. \tag{1}$$

In the neutral limit, $r \to 1$, all individuals in the population are equally likely to end up as the single common ancestor, leading to a fixation probability of $1/N$.

The original Moran process ignores population structures – but this is easily addressed by arranging individuals of a population on a graph, such that each node refers to one individual and the links to other nodes define its neighbourhood. (8) and (9) conjectured that the fixation probability of a mutant in this Moran process on graphs remains unaffected by population structures. (7) proved that this is indeed true for a broad class of structures and, in particular, holds for simple structures such as lattices or regular networks. At the same time, this classification indicated that fixation probabilities, ρ, may differ for some structures by tilting the balance between selection and random drift. Evolutionary suppressors enhance random drift and suppress selection ($\rho < \rho_M$ for $r > 1$ and $\rho > \rho_M$ for $r < 1$), whereas evolutionary amplifiers exhibit the intriguing property to enhance selection and suppress random drift ($\rho > \rho_M$ for $r > 1$ and $\rho < \rho_M$ for $r < 1$).

In recent years, various aspects of the Moran process on graphs have been explored, including effects of population structures on fixation probabilities (1; 10; 11), or fixation times (12; 13), as well as computational techniques (14; 15). However, the most intriguing result remains that even perfect evolu-

tionary amplification appears to be possible: "The superstar. . . [has] the amazing property that for large [population sizes] N, the fixation probability of an advantageous mutant converges to one, while the fixation probability of a disadvantageous mutant converges to zero." (7).

More recently (16) provided a sophisticated and elaborate counter example that contradicted the fixation probabilities reported in (7). Here we identify the problem in the original proof, correct it and report new upper and lower bounds on the fixation probability for superstars. Moreover, for any $r > 1$, a graph can be constructed such that ρ is arbitrarily close to 1, thus confirming the existence of perfect evolutionary amplification.

2. Model

2.1. Moran process on graphs

Population structure can be represented by assigning individuals to nodes on a graph with links representing each individuals' neighbourhood. The Moran process on graphs follows the same procedure as the original Moran process except for the crucial difference that the offspring does not replace a random member of the entire population but rather replaces a *neighbour* of the reproducing individual, selected uniformly at random (Fig. 1). On directed graphs, the offspring replaces a downstream neighbour by selecting one outgoing link uniformly at random. As before, the population has reached fixation once either one of the absorbing, homogenous states is reached. For any number of mutants, m $(0 < m < N)$, the fixation probabilities of residents and mutants are both non-zero on strongly connected graphs, i.e. graphs where every node can be reached from any other node through a series of moves between nodes that are connected by links (for directed graphs, only moves in the direction of the link are permitted). If a graph is not strongly connected, then the structure may prevent the spreading or elimination of a mutant type regardless of its fitness and hence the fixation probability for either or both types can be zero.

For the Moran process on graphs, the fixation probabilities are the same as in unstructured populations, c.f. Eq. (1), provided that the graph is a circulation (7). For circulations the sum of weights of all outgoing links is equal to the sum of weights of all incoming links for every node. This means that every node has the same impact on the environment as the environment has on the node.

A graph is an evolutionary suppressor if the fixation probability of an advantageous mutant is less than for the original Moran process, $\rho < \rho_M$. The simplest example is a linear chain: a graph with a single root node, which connects to one end of a (directed) chain of nodes (17). Any mutation that does not occur at the root has no chance of reaching fixation. However, if the mutation occurs in the root node it eventually takes over with certainty. Assuming that mutations arise spontaneously and are equally likely in any location, the resulting fixation probability is simply $1/N$, regardless of the mutant's fitness r. The linear chain is an example of a graph that is not strongly connected,
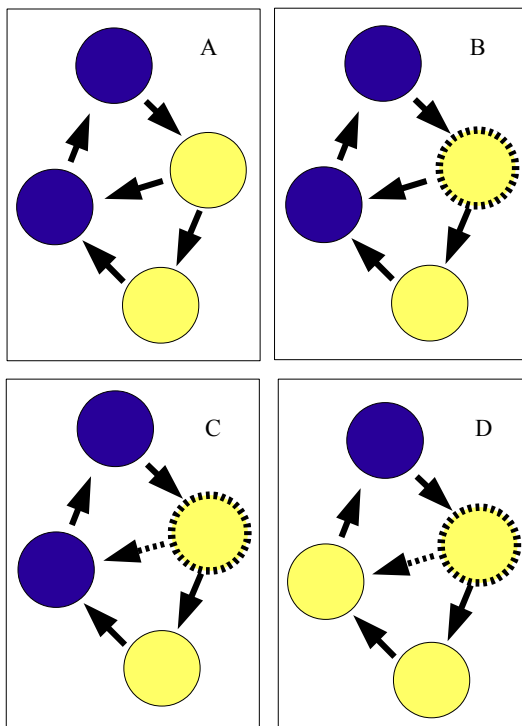
Figure 1: Moran process on a graph. (A) graph structure and distribution of residents (blue) and mutants (yellow). (B) selection: an individual (dashed outline) is selected to reproduce with a probability proportional to its fitness. (C) replacement: a downstream neighbour (dashed arrow) is randomly selected for replacement. (D) reproduction: the neighbour is replaced by the clonal offspring of the upstream reproducing individual.

because the root node cannot be reached from any node in the chain. Evolutionary suppressors are often found when high fidelity copying is of paramount importance, such as in slowing down the somatic evolution of cancer (17; 18).

In contrast, an evolutionary amplifier is a graph, which increases the fixation probability of advantageous mutants as compared to the original Moran process, $\rho > \rho_M$. The simplest evolutionary amplifier is the star graph: a single root node is connected to a reservoir of peripheral leaf nodes through bi-directional links. The fixation probability of a single mutant for $N \gg 1$ is (7; 10)

$$\rho_0 \approx \frac{1 - \frac{1}{r^2}}{1 - \frac{1}{r^{2N}}}. \tag{2}$$

On the star, a mutant with fitness r has roughly the same fixation probability as a mutant with fitness r^2 in an unstructured population. Thus, the fixation probability of beneficial mutations ($r > 1$) is enhanced, but for deleterious

4

mutants ($r < 1$) it is reduced. Note that the fixation probability depends on where the single mutant arises. If the mutant is located in the root node then, for $N \gg 1$, it is almost certainly replaced in the next time step because one of the $N-1$ reservoir nodes is selected for reproduction. However, if mutants arise at random, then for $N \gg 1$ they almost surely arise in the reservoir and the fixation probability is as specified in Eq. (2). Evolutionary amplifiers would seem to provide promising structures for tasks where strong selection is advantageous, such as in the adaptive immune system or in collaboration networks.

2.2. Superstars

The two most prominent features of the star graph are the large reservoir where changes occur on a slow time scale, and the bottleneck caused by the hub or root, where changes occur quickly. In particular, the bottleneck introduces a second level for selection to act upon- a mutant needing to reproduce in both the leaves and the hub before it successfully increases its population in the leaves. This basic insight can be exploited to increase evolutionary amplification by elongating the bottleneck and providing further levels where selection can act. Superstars act as a more extreme version of the basic star, and have been proposed as a way to increase evolutionary amplification further (7). The superstar consists of a single root node surrounded by B branches (Fig. 2). Each branch consists of a large reservoir of L nodes feeding into one end of a linear, directed chain of length H, the stem. The last stem node in each branch feeds into the root node, which then connects to all reservoir nodes in every branch. The population size is thus given by $N = B(L + H) + 1$. Nodes are classified based on their locations on the graph. This classification is designed to simplify discussions but does not affect the rate of reproduction of the individual occupying the node. (7) report the fixation probability for superstars with $L, B \gg H$ as

$$\rho_H \approx \frac{1 - \dfrac{1}{r^k}}{1 - \dfrac{1}{r^{kN}}}, \tag{3}$$

where $k = H + 2$ is a structural parameter and indicates the number of moves needed to reach any reservoir node from any other reservoir node. This is the number of levels selection can act upon. Consequently it is argued that a single mutant that arises in the reservoir of a superstar with fitness r has approximately the same fixation probability as a mutant with fitness r^k in an unstructured population. This result would then imply that by increasing the length of the stem, the fixation probability, ρ_H, of any advantageous mutant, $r > 1$, could be brought arbitrarily close to 1, indicating arbitrarily strong amplification or perfect selection.

Recently (16) provided a counter-example demonstrating that the fixation probability in Eq. (3) is too optimistic in the particular case of $H = 3$ and thus invalidated the proof in (7). In addition, (16) provide substantial simulation
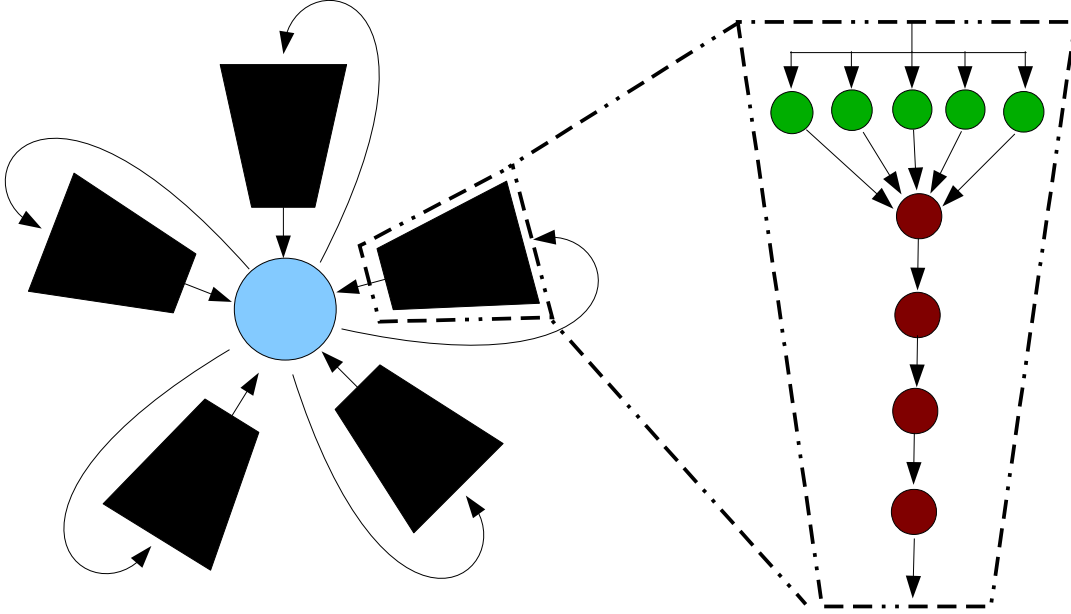
Figure 2: The superstar consists of three distinct types of nodes: the root node (pale blue), the reservoir nodes (green) and the stem nodes(dark red). The reservoir nodes connect to the start of the stem, the end of the stem connects to the root node and the root node connects to all reservoir nodes in each branch. The depicted superstar has $B = 5$ branches each with $L = 5$ reservoir nodes and a stem of length $H = 4$, which yields a total population size of $N = B(L + H) + 1 = 46$.

based evidence indicating that Eq. (3) also fails for higher values of H. For the counter-example they show that in the limit $N \to \infty$:

$$\rho_3 < 1 - \frac{1 + r}{2r^5 + r + 1}. \qquad (4)$$

This upper bound reflects the probability that a mutant in a reservoir creates a second mutant in any reservoir before getting replaced by resident offspring. Clearly, the fixation probability according to Eq. (3) grows faster with increasing r than Eq. (4) and for $r > 1.42$, results in a contradiction. It turns out that the original proof (7) was based on an optimistic assumption concerning the amplification along the stem. Taking correlations in the dynamics along the stems into account we obtain new bounds on the fixation probability. More specifically, for $L = B$ we find in the limit $B \to \infty$

$$1 - \frac{1}{r^4(H-1)(1 - \frac{1}{r})^2} \le \rho_H \le 1 - \frac{1}{1 + r^4 H}. \qquad (5)$$

Fixation thus tends to certainty for $H \to \infty$, as suggested by (7), while no longer violating the upper bound identified by (16) for $H = 3$.

6

In the following we borrow a number of valuable concepts and techniques from both articles, adding and extending where necessary. Exact bounds on the error terms for finite populations are provided in the appendices.

3. Proof

For the proof of the fixation probability on superstars, we follow the tradition of (7, 16) and consider superstars with many branches, B, and large reservoirs, L. More specifically, we study the dynamics of a single branch in detail, and use this to determine the much slower dynamics of changes in the reservoirs. For any given stem length, $H > 2$, the following arguments become exact in the limit $B, L \to \infty$. In practice, we obtain good approximations for $H \ll B, L$. In an effort to reduce notation and increase clarity we assume in the following that $H \ll B = L \approx \sqrt{N}$. However, in the full proof (see appendix) we only require that B and L scale with N but not that they are of the same size.

If mutations arise spontaneously and with equal probability at any node, then the initial mutant almost certainly arises in a reservoir node, because reservoir nodes vastly outnumber nodes of all other types. This marks the starting point for the remainder of our proof (for details see Appendix A).

On occasion, we need to refer to the total fitness of our superstar population at a given time, F_t, with $N < F_t < rN$. However, all instances of F_t cancel throughout the proof and hence we do not need to keep track of its exact value. Moreover, various necessary approximations introduce different error terms that are accounted for in full detail in the appendices. All error terms tend to zero as $B, L \to \infty$. It is sufficient to assume that $B = L$ as we take this limit, however other relations are also possible. The exact restrictions on how we can take these limits can be found in Appendix E, but for now let us simply assume that limits are taken simultaneous, with some suitable relation binding B and L togeather.

3.1. Timescales

Different nodes get updated at different rates. More precisely, any given node is updated if one of its upstream neighbours reproduces and the node of interest is chosen for replacement. Here we follow the convention that all nodes have weighted out-degree of 1, and that weight is evenly distributed among outgoing links (each link has weight $1/d$ for a node with d outgoing links). Thus, nodes with high (weighted) in-degrees, tend to be short lived, while nodes with low (weighted) in-degrees, tend to be long lived.

Assuming $r \ll N$, every node is selected for reproduction approximately with probability $1/N$. The root node has an in-degree of B and all its upstream neighbours have out-degrees of 1, hence it updates with a probability close to $B/N \approx 1/\sqrt{N}$. Recall that we assume $H \ll B = L \approx \sqrt{N}$. Similarly, reservoir nodes are replaced with probability of approximately $1/N^2$, the first stem node with probability on the order of $1/\sqrt{N}$, and all other stem nodes with probability of approximately $1/N$. For $N \gg 1$, this results in three different timescales: the

slowest for reservoir nodes that get replaced, on average, only once in N^2 time steps; an intermediate timescale for the stem nodes (with the exception of the first node), which get replaced once in N time steps; and a fast timescale for the root node as well as the first stem node in each branch, which update once in $\sqrt{N}$ time steps, respectively.

For $N \gg 1$, it is possible to separate the three timescales and analyze the dynamics of the different types of nodes individually. More specifically, this allows us to focus on the intermediate timescale associated with the dynamics in the stem, while treating the state of the slowly updating reservoir nodes as constant and the fast updating nodes as random variables. In the following, we derive the evolutionary dynamics for the top, middle and bottom of the stem in a single branch. The results determine the slow dynamics of reservoir nodes and describes the early stages of the invasion process, when mutants are rare among the reservoir nodes. This allows us to derive upper and lower bounds on the fixation probabilities.

3.2. Top of stem

The first node of the stem gets replaced on the fast time scale, which allows us to treat its state as a random variable. Initially, out of all L upstream neighbours in the reservoir of the corresponding branch, only one is a mutant. Hence, at any given time step, the top node is occupied by a mutant with probability close to r/L. This mutant reproduces with a probability r/F_t and hence the probability that a mutant is placed in the second stem node is approximately $r^2/(F_t L)$ in each time step (for error terms, see Appendix A).

3.3. Middle of stem

The structure of the stem causes the state of any given stem node to be highly correlated with its neighbours, both upstream and downstream. More specifically, if a mutant reproduces it is highly likely to end up replacing its own offspring. This correlation had been neglected in (7). In the absence of correlations, whenever a mutant in the stem reproduces, it almost certainly replaces a resident since residents are more common. Along a stem of length H this results in an overall amplification of r^H. However, due to correlations, if a mutant in the stem reproduces, it *likely* replaces its former offspring and hence diminishes the resulting amplification.

Simulations nicely illustrate the characteristic features of the dynamics along the body of the stem: clusters of mutants begin at the top of the stem, then grow and move along the stem. In the following, we refer to these clusters as *trains*. A train moves forward and increases in length whenever the front mutant reproduces, which happens at a rate r/F_t, but shrinks whenever a resident reproduces and replaces the back end of the train, which occurs at a rate $1/F_t$, see Fig. 3. Thus, as the train moves along the stem, the train length for beneficial mutants increases, on average.

Note that for small superstars with a single node in the stem body, which corresponds to $H = 2$ (or $k = 4$), the two stem nodes are indeed uncorrelated.
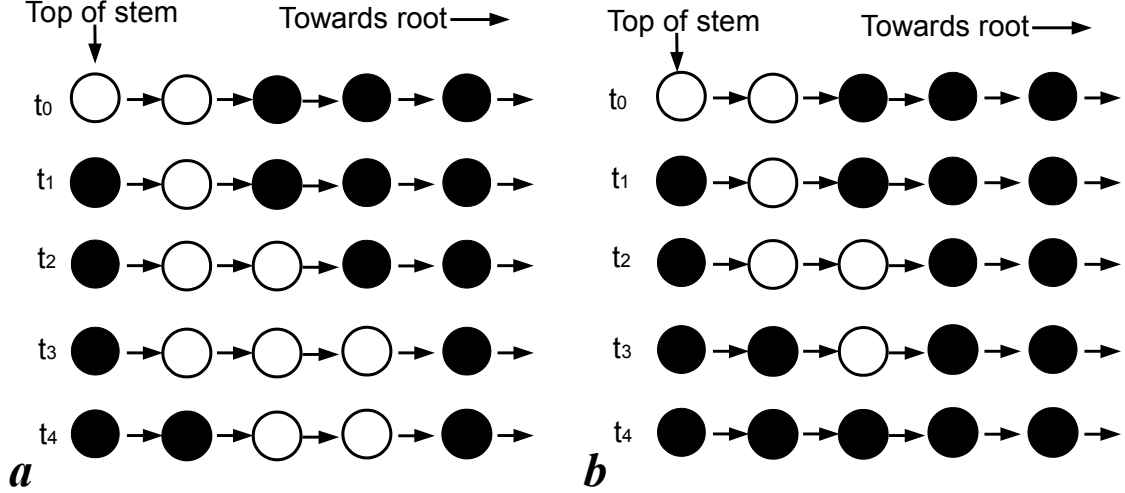
Figure 3: Two possible histories of a train of mutants (white) proceeding along a stem filled with residents (black). **a** We begin with two mutants (t_0). The top node is quickly replaced by a resident (on the fast time scale) (t_1). Some time later the remaining mutant reproduces (t_2), and then the new top node reproduces again (t_3). Finally we lose a single mutant from the back of out train (t_4). This general growing pattern applies whenever $r > 1$. **b** We begin with two mutants (t_0), and immediately lose the back mutant of our train (t_1). The front of the train reproduces, creating a second mutant (t_2), but both fall prey to bad fortune (or low fitness) and are removed (t_3, t_4). This behaviour is likely when $r < 1$, but even for beneficial mutations many trains do not reach the end of the stem.

However, for $H > 2$ this assumption breaks down and results in an overestimation of the fixation probabilities as pointed out by (16).

In order to link the stem dynamics to the slow timescale of reservoir nodes, we need to know the expected train length, T, by the time the train first reaches the root end of the stem. The history of any given train can be represented as a random sequence of increments and decrements with a bias that increments are r times more likely. Essentially, we need to sum over all possible sequences of increments and decrements given an initial train length of 1 and weigh the resulting train length with the probability of the respective sequence. All sequences where trains disappear (zero train length) are omitted (zero weight). We can count these sequences that result in extinction using the reflection principle (see Appendix B). This method yields the expected train length:

$$T = \left(\frac{r}{1+r}\right)^{H-2} \sum_{z=1}^{H-1} (H-z)\left(\frac{1}{1+r}\right)^{z-1}\left[\binom{H-4+z}{z-1} - \binom{H-4+z}{z-2}\right].$$

(6)

9

For $H \geq 2, r > 1$ simple bounds for T exist (see Appendix B.1):

$$(H-1)\left(1-\frac{1}{r}\right)^2 \leq T \leq H. \tag{7}$$

This indicates that for $r > 1$ the train length T grows approximately linearly with increasing stem length H.

3.4. Bottom of stem

Whenever a train reaches the root end of the stem, its mutants compete with the resident nodes from the other branches to occupy the root node. Since the root node is updated on the fast timescale we can again treat its state as a random variable. Thus, once a train has reached the root end of the stem, the root node is a mutant with probability close to r/B. Thus, as long as the train sits at the root end of the stem, the probability in any given time step that the root node is a mutant, reproduces and creates a second mutant in any reservoir is $r^2/(F_t B)$. However, the train is simultaneously eroded from behind, with train mutants being replaced by residents with probability $1/F_t$. Thus, the train remains at the root end for TF_t time steps, on average. Put together, this means that any given train succeeds in producing a second mutant in any reservoir with a probability close to $r^2 T/B$ (for detailed error bounds see Appendix C).

3.5. Slow dynamics in reservoirs

At any given time step, the probability of losing the initial mutant in the reservoir is $1/(F_t B L)$. Based on the dynamics in the stem, we derive the per time step probability that a second mutant is generated in the reservoir of any branch as the product of the probability that a train is generated and the probability that the train succeeds in producing a second mutant, which yields approximately $r^4 T/(F_t B L)$. Thus, the probability to eventually go from one to two mutants in the reservoirs, as opposed to losing the initial mutant, is close to

$$\frac{r^4 T}{1 + r^4 T}. \tag{8}$$

Since T can be made arbitrarily large (by increasing the stem length H, see Eq. (7)), the transition from one to two mutants becomes almost certain and, similarly, the probability of losing the initial mutant becomes vanishingly small.

3.6. Upper bound on fixation probability

To find an upper bound on the fixation probability, ρ_H, we note that before our mutant can reach fixation, the superstar must first transition from a state with one mutant in a reservoir to a state with two mutants in the reservoirs. Thus, an upper bound on this transition probability serves as an upper bound on the mutant fixation probability. Moreover, the upper bound can be made

independent of T by assuming that all trains have the maximum possible train length. Thus, in the limit of large B and L we find

$$\rho_H \leq 1 - \frac{1}{1 + r^4 T} \leq 1 - \frac{1}{1 + r^4 H}. \tag{9}$$

For any given H, r we can find T explicitly using Eq. (6). In particular, we note that for $H = 3$, we find $T = 2r/(r+1)$, thus recovering the upper bound identified in (16).

3.7. Lower bound on fixation probability

We find a lower bound on the fixation probability by approximating the dynamics of the system with a random walk. This random walk has a forward bias given by Eq. (8) as long as mutants are rare, and we assume no forward bias otherwise. Because even for larger numbers of mutants the forward bias persists (but there is no simple way to quantify the bias) we obtain a lower bound of the fixation probability, ρ_H.

For any finite number of steps, a sufficiently strong initial bias would suffice to ensure that the random walk eventually reaches fixation with high probability. However, the limit $N \to \infty$ also requires an arbitrarily large number of forward steps. In order to resolve the interplay between these two limiting behaviours we set up a martingale and apply the optional stopping theorem (19, p. 210) (see Appendix D.2 for details). In the limit of large B and L we find:

$$\rho_H \geq \hat{\rho}_H = 1 - \frac{1}{r^4 T} \geq 1 - \frac{1}{r^4 (H-1)(1-\frac{1}{r})^2}. \tag{10}$$

Once again we note that for any given H, r we can find T explicitly, Eq. (6). Combined with the upper bound, Eq. (9), this means that ρ_H must exist in a narrow window, Eq. (5).

4. Robustness

Unfortunately, it turns out that the intriguing feature of superstars to act as evolutionary amplifiers holds only under very specific conditions. Here we discuss the most important requirements.

4.1. Selection & Sequence

The original Moran process is formulated as a fecundity based birth-death process, that is, fitness affects the rate of birth (reproduction) whereas death (replacement) occurs uniformly at random. Alternatively, fitness could just as well affect survival such that birth events occur uniformly at random but death events might, for example, occur with probability inversely proportional to fitness. Similarly, the sequence of events could be reversed such that first an individual dies and then the remaining individuals compete to repopulate the vacant site. This yields a total of four distinct scenarios: *Bd*, *bD*, *dB* and *Db*,

11

where capital letters refer to the fitness dependent selection step. The original Moran process corresponds to Bd and the fixation probability is given in Eq. (1). In unstructured populations the four dynamical scenarios result in only marginal differences in fixation probabilities. However, they can have crucial effects on the evolutionary outcome in structured populations (20; 21; 22). (23) examine all four cases for both complete graphs and star graphs, showing that stars act as evolutionary suppressors in both the dB and Db cases, and are significantly less effective in the bD case compared to the original Bd case. Similar results apply to superstars:

bD *updates:.* For the birth-death process with selection on survival, mutants only gain any advantage whenever the root node reproduces. Whenever any other node reproduces, there is only a single downstream node, and thus no opportunity for competition, rendering any fitness advantage irrelevant. This lack of advantage in the stem leads to an expected train length of 1, regardless of stem length or mutant's fitness. The chance of launching a successful train in a given time step is $1/(NLB)$ and the chance of replacing the original mutant is $1/(NBLr - N(r-1))$. This results in a bias of approximately $\frac{r}{1+r}$ for the initial mutant to eventually create a second reservoir mutant – the same bias as for the original Moran process. Thus, we might expect fixation probabilities similar to the original Moran process on BL nodes, and certainly nowhere close to the amplification observed for Bd updates.

Db *updates:.* For the death-birth process with selection on survival, the prospects of mutants drop even further. The probability to successfully place even a single offspring in the top of the stem is only $r/(L+r)$. Note that for death-birth processes the top of the stem no longer changes on the fast timescale and hence trains start at the top instead of the second node. As the train propagates along the stem it tends to grow because the mutant at the back of the train is less likely to die than the resident in front of it, leading to the same train dynamics observed for the Bd process. Upon reaching the end of the stem the train competes with the other branches for control over the root node and succeeds with probability near rT/B (over the lifetime of the train). Once a mutant occupies the root, it is predestined to have many offspring – in each time step a reservoir node dies with high probability and gets replaced by an offspring of the mutant in the root node, whereas the probability is low that the root node is replaced. More specifically, we expect $rN/(1+r)$ reservoir nodes to become mutants before the root node is replaced. At that point it is reasonable to assume that mutants reach fixation with high probability. We conjecture that the probability of mutant fixation on the superstar is close to the probability of a mutant eventually being placed in the root node. Thus, we expect a fixation probability close to r^2T/BL. This result is significantly *less* than the almost certain fixation in the original Moran process (c.f. Eq. (1)). The result does, however, match well with the $1/N$ scaling found for the fixation on stars (23).

dB *updates:.* The final case is the death-birth process with selection on reproduction. Once again the probability of placing a mutant offspring in the stem

12

before losing the reservoir mutant is near r/L, but now without further benefits along the stem. Consequently, trains that do reach the root still have an expected length of 1. Thus, each train has a probability of roughly r/B for claiming the root node, which then produces $N/2$ mutants in the reservoirs, on average – enough to suggest fixation with high probability, but less than for Db. Thus, we conjecture fixation probabilities near r^2/N – the worst outcome of the four scenarios. Once again, we note the significant penalty as compared to the original Moran process as well as the similarities to the $1/N$ scaling for stars (23).

4.2. Mutations

Even though we did not explicitly model the process of mutation, we implicitly assumed that mutations are rare and arise spontaneously in any node selected uniformly at random. For the superstar this means that most mutations arise in a reservoir node – simply because the overwhelming majority of nodes are reservoir nodes.

An alternative and equally natural assumption is that mutations arise during reproduction events. Such a change does not affect the fixation probabilities in the original Moran process. However, in highly heterogenous population structures crucial differences in the fixation probabilities can arise because mutants preferentially arise in certain locations (24). For superstars, when using the Bd or bD update rules, mutants most likely arise at the top of a stem. This is an unfortunate position because the mutant is highly likely replaced before reproducing even once – extinction is almost certain. In contrast, for Db and dB, mutants again most likely arise among the reservoir nodes – but for those updates superstars do not act as evolutionary amplifiers.

Even though the dynamical properties of superstars are intriguing, the list of caveats demonstrates that the evolutionary amplification is highly sensitive to the details of the model – maybe this is the reason that superstar-like structures have not been reported in nature.

5. Conclusion

Superstars represent the most prominent representatives of evolutionary amplifiers – structures that are capable of increasing selection and suppressing random drift. For $r > 1$ and in the limit of large N we have derived upper and lower bounds for the fixation probability, ρ_H:

$$1 - \frac{1}{r^4 T} \le \rho_H \le 1 - \frac{1}{1 + r^4 T}, \tag{11}$$

where $(H-1)(1-1/r)^2 \le T \le H$.

Even though fixation probabilities can be made arbitrarily close to 1 on large superstars and sufficiently large H, the fixation probability remains bounded away from 1 for any finite graph. As a concrete example, consider $r = 2$ and $H = 50$, which yeilds $T \approx 13.25$ and $0.995283 \le \rho_{50} \le 0.995306$ in the limit of

large N. Similarly, a sizeable, finite superstar with $B = L = 5000$ ($N \approx 2.5 \cdot 10^5$) yields $0.985323 \leq \rho_{50}^N \leq 0.995375$, which includes all error terms (see Appendix E). In contrast, the fixation probability for a similarly sized isothermal graph (e.g. a lattice, complete or random regular graph) is just short of 0.5.

The upper bound for ρ_H in Eq. (11) results in a contradiction with the originally reported fixation probability, Eq. (3), for sufficiently large r. For the specific case of $H = 3$ the discrepancy was pointed out in (16). At the same time, the lower bound for ρ_H in Eq. (11) confirms that superstars are indeed capable of providing an arbitrarily strong evolutionary advantage to any beneficial mutation, as suggested in (7). Using symmetry arguments, it also follows that for $r < 1$ the fixation probability can be made arbitrarily small, as required for a perfect evolutionary amplifier (see Appendix F).

In the case $H = 2$ (or $k = 4$ in (7)) we obtain an expected train length of $T = 1$ and recover the original bias, $r^4/(1 + r^4)$. Discrepancies arise only for $H \geq 3$ (or $k \geq 5$) but those cases were not included in the simulations in (7). For $H = 3$, we obtain $T = 2r/(1+r)$, which results in a bias of $2r^5/(1+r+2r^5)$ and recovers the upper bound reported by (16). Extending the technique in (16) to higher values of H numerically, we find that the upper bounds found match Eq. (11).

An appropriately skeptical reader might ask why the theory presented here should be trusted over those previously presented in the literature – after all, both claim to offer rigorous proof. First, we note the agreement between predictions made here, and both (7) and (16) for the appropriate values of H. Second, we identify correlations between neighbouring stem nodes as the cause for the discrepancies between the two previous papers. Finally, we invite readers to scrutinize the proof offered here most thoroughly. Superstars have already presented unexpected subtleties, and as always, we need caution and vigilance to discern between scientific selection and random drift.

References

[1] Antal, T., Redner, S. & Sood, V., 2006 Evolutionary dynamics on degree-heterogeneous graphs. *Physical Review Letters* **96**, 188104.

[2] Bürger, R. & Lande, R., 1994 On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics* **138**.

[3] Nowak, M. A. & May, R. M., 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829.

[4] Fu, F. & Nowak, M. A., 2013 Global migration can lead to stronger spatial selection than local migration. *Journal of Statistical Physics* **151**, 637–653. (doi:10.1007/s10955-012-0631-6).

[5] Moran, P. A. P., 1962 *The Statistical Processes of Evolutionary Theory*. Clarendon Press, Oxford.

[6] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D., 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650.

[7] Lieberman, E., Hauert, C. & Nowak, M. A., 2005 Evolutionary dynamics on graphs. *Nature* **433**, 312–316. ISSN 1476-4687.

[8] Maruyama, T., 1970 Effective number of alleles in a subdivided population. *Theoretical Population Biology* **1**, 273–306.

[9] Slatkin, M., 1981 Fixation probabilities and fixation times in a subdivided population. *Evolution* **35**, 477–488.

[10] Broom, M. & Rychtář, J., 2008 An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A* **464**, 2609–2627. (doi:10.1098/rspa.2008.0058).

[11] Burton Voorhees, A. M., 2013 Fixation probabilities for simple digraphs. *Proc. R. Soc. A* **469**.

[12] Joshua L. Payne, M. J. E., 2009 Evolutionary dynamics on scale-free interaction networks. *IEEE, transactions on evolutionary computation* .

[13] Frean, M., Rainey, P. & Traulsen, A., 2013 The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B* **280**, 20130211. (doi:10.1098/rspb.2013.0211).

[14] Shakarian, P. & Roos, P., 2011 Fast and deterministic computation of fixation probability in evolutionary graphs. In *Computational Intelligence and Bioinformatics / Modelling, Identification, and Simulation* (eds. T. Znati, Z. Bar-Joseph & L. Rothrock). Calgary AB: ACTA Press.

[15] Fu, F., Nowak, M. A. & Hauert, C., 2009 Evolutionary dynamics on graphs: Efficient method for weak selection. *Physical Review E* **79**, 046707.

[16] Diaz, J., Mertzios, L. A. G. A. G. B., Richerby, D., Serna, M. & Spirakis, P. G., 2013 On the fixation probability of superstars. *Proc. R. Soc. A* **469**.

[17] Nowak, M. A., Michor, F. & Iwasa, Y., 2003 The linear process of somatic evolution. *Proceedings of the National Academy of Sciences USA* **100**, 14966–14969.

[18] Michor, F., Iwasa, Y. & Nowak, M. A., 2004 Dynamics of cancer progression. *Nature Reviews Cancer* **4**, 197–205.

[19] Klenke, A., 2006 *Probability Theory: A comprehensice Course*. Springer.

[20] Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A., 2006 A simple rule for the evolution of cooperation on graphs. *Nature* **441**, 502–505.

[21] Ohtsuki, H. & Nowak, M. A., 2006 The replicator equation on graphs. *Journal of Theoretical Biology* **243**, 86–97.

[22] Zukewich, J., Kurella, V., Doebeli, M. & Hauert, C., 2013 Consolidating birth-death and death-birth processes in structured populations. *PLoS One* **8**, e54639.

[23] Frean, M. R. & Baxter, G. J., 2008 Death-birth ordering and suppression of fitness in networks. Working paper: http://homepages.mcs.vuw.ac.nz/~marcus/manuscripts/FreanBaxterJTB.pdf.

[24] Maciejewski W, H. C., Fu F, 2014 Evolutionary game dynamics in populations with heterogenous structure. *PLoS Computational Biology* (doi: 10.1371/journal.pcbi.1003567).

[25] Koralov, L. B. & Sinai, Y. G., 2007 *Theory of Probability and Random Process*. Springer, 2nd edition.

Appendix A. Initial conditions

If mutations arise spontaneously and with equal probability in any node then the initial mutant arises in a reservoir node with probability $BL/(BL+1+HB)$. This probability can be made arbitrarily close to one, for suitably large L or B. The mutant arises in a stem or root node with probability

$$\epsilon_0 = \frac{1+HB}{BL+1+HB}. \tag{A.1}$$

Thus, the final fixation probability, ρ_H (see Eq. (11)), needs to be multiplied by $1-\epsilon_0$ to account for this possibility. Because ϵ_0 is only used to derive the lower bound on ρ_H, assuming extinction of all mutants not arising in a reservoir node preserves inequalities by effectively ignoring the small possibility that an invading mutant placed in stem nodes or the root node could still reach fixation.

The first node of the stem gets replaced on the fast time scale, which allows us to treat its state as a random variable. However, at early stages of invasion, only one of the L upstream neighbours is a mutant. Hence, at any given time step, the top node is occupied by a mutant with probability

$$\frac{r}{L-1+r} = \frac{r}{L}(1-\epsilon_1) \tag{A.2}$$

$$\epsilon_1 = \frac{r-1}{L+r-1} \tag{A.3}$$

This mutant reproduces with a probability r/F_t and hence the probability that a mutant is placed in the second stem node is $r^2/(F_t(L+r-1))$ in each time step.

16

However, we need to account for the possibility that the initial mutant in the reservoir is replaced *before* the first node in that stem. On a given time step, the chance that the reservoir mutant is replaced by a resident is less than $1/(F_tBL)$. Conversely, the probability of the first node in the chain being replaced exceeds L/F_t. Thus the chance that the initial mutant gets replaced before its offspring replaces the first node in the chain is

$$\epsilon_2 < 1/(1 + BL^2). \tag{A.4}$$

For our proof we assume that the first node in any chain can be treated as a random variable. The above error term accounts for any slight discrepancy caused by our initial conditions.

Appendix B. Expected train length T.

Mutants placed in the main body of the stem (excluding the first node, which updates on a fast timescale) propagate down the stem in trains. Trains grow at one end as mutants reproduce, and shrink at the other end as mutants are replaced by residents (see Fig. 3). At any time, t, the state of a train is given by two integers: A_t and Z_t. Here A_t refers to the position of the mutant at the front of the train, and Z_t refers to the position of the resident directly behind the train. The current length of the train is thus given by $A_t - Z_t$. Because in most time steps no change occurs in this particular stem, we consider a condensed process, which only accounts for events that change the state of the train. This means that A_t increases with probability $r/(1+r)$ while Z_t increases with probability $1/(1+r)$. Thus, for beneficial mutants the length tends to increase as the train progresses down the stem. If at any time $Z_t \geq A_t$ the train has vanished and the stem is cleared of mutants. In this case, we say that the train, which "arrives" at the end of the stem, has length zero.

In order to determine the expected train length, T, we consider the above process on a grid, where the horizontal axis represents the position of the front, A_t, and the vertical axis the back of the train, Z_t. Each point on the grid and below the diagonal, $A_t = Z_t$, represents a possible configuration of a train in the stem, see Fig. B.4. All other points represent invalid configurations, which we refer to as ghost states. For each train, the initial configuration is $(A_0, Z_0) = (2, 1)$, that is, the second stem node is a mutant, while the state of the first stem node is a resident (due to replacement on the fast time scale, leading to a slight underestimate).

Each train produces a trajectory or path on the grid that originates in (A_0, Z_0) and ends at time τ once the train has reached its destination: the bottom of the stem, where $A_\tau = H$. However, if at any point in time $A_t \leq Z_t$ then this represents an *invalid* path because the train has vanished. Every invalid path touches or crosses the diagonal $A_t = Z_t$ at least once. For a *valid* path $A_t > Z_t$ must hold at all times. The expected train length, T, is the weighted average over all paths, with invalid paths being considered as having length zero. The number of valid paths can be calculated using the reflection
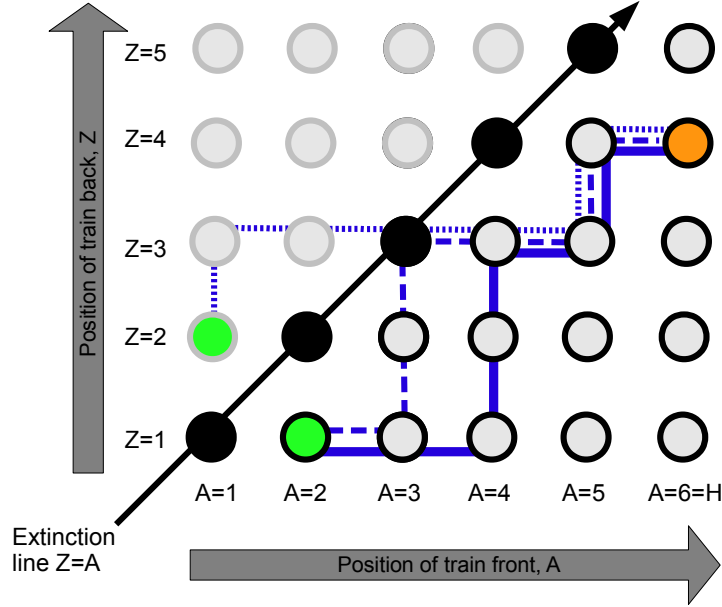
17

Figure B.4: Grid showing collection of possible train states. Permitted states (black outline), ghost states (grey outline) and extinction states (black fill) as well as a number of possible paths from our initial state (black outline, green fill) to a sample end state (black outline orange fill). Depicted are a permitted path (continuous), an invalid path (leading to extinction, long dash) and the associated "ghost path" from the reflection of our initial state to the sample end state (fine dash).

principle (25, p. 88), which states that for every *invalid* path a ghost path exists, starting from (Z_0, A_0). The trajectory of a ghost path is the reflection of the corresponding invalid path along the diagonal $A_t = Z_t$, up to the point where the invalid path touches or crosses $A_t = Z_t$ for the first time. From then on the ghost path and the remainder of the invalid path coincide, see Fig. B.4. In order to calculate the expected train length, T, we consider the train lengths based on *all* paths and subtract all *invalid* paths, to obtain the train length based on *valid* paths only. The number of ghost paths corresponds to all paths starting from (Z_0, A_0), the reflection of (A_0, Z_0) and hence the name of the method. Having counted the number of paths we then weigh the corresponding train length by the probability of each possible path and obtain:

$$T = (1 - \alpha)^{H-2} \sum_{z=1}^{H-1} (H - z)\alpha^{z-1} \left[\binom{H - 3 + z - 1}{z - 1} - \binom{H - 3 + z - 1}{z - 2} \right]$$

(B.1)

with $\alpha = 1/(1 + r)$. We assume beneficial mutations, $r > 1$, such that $0 < \alpha < 1/2$. All paths require $H - 2$ steps that increase A_t from the starting point at 2

to the end point at H, which occurs with probability $1 - \alpha$ for each step. The combinatorial sum then accounts for all possibilities and probabilities that Z_t is increased along valid paths. In particular, the index variable z indicates the position of the tail of the train and hence $H - z$ specifies the train length. The tail starts at 1 and, for any valid path, reaches at most $H - 1$. Because we are interested in the length of the train at the moment of arrival, the final step must be an increment of A_t. In particular, it follows that $z = H - 1$ has zero valid paths – a reassuring result as we know that no train could possibly have length one at the moment of its arrival. Note that we have used the convention that $\binom{n}{k} = 0$ for $k < 0$, which applies only if the tail remains at $Z_t = 1$ and admits only a single valid path.

Appendix B.1. Simplifying T

We now resort to algebraic manipulation. Various binomial coefficient identities are used throughout:

$$T = (1-\alpha)^{H-2} \sum_{z=1}^{H-1} (H-z)\alpha^{z-1} \left[\binom{H+z-4}{z-1} - \binom{H+z-4}{z-2} \right]$$

using Pascal's rule

$$= (1-\alpha)^{H-2} \sum_{z=1}^{H-1} (H-z)\alpha^{z-1} \left[\binom{H+z-3}{z-1} - 2\binom{H+z-4}{z-2} \right]$$

splitting sum

$$= (1-\alpha)^{H-2} \sum_{z=1}^{H-1} (H-z)\alpha^{z-1} \binom{H+z-3}{z-1} - 2\alpha(1-\alpha)^{H-2} \sum_{z=1}^{H-1} (H-z)\alpha^{z-2} \binom{H+z-4}{z-2}$$

changing second summation to obtain lower bound

$$\geq (1-\alpha)^{H-2} \sum_{z=1}^{H-1} (H-z)\alpha^{z-1} \binom{H+z-3}{z-1} - 2\alpha(1-\alpha)^{H-2} \sum_{z=2}^{H} (H+1-z)\alpha^{z-2} \binom{H+z-4}{z-2}$$

merging sums and relabelling indices

$$= (1-2\alpha)(1-\alpha)^{H-2} \sum_{z=0}^{H-2} (H-1-z)\alpha^z \binom{H+z-2}{z}$$

expanding factor

$$= (1-2\alpha)(1-\alpha)^{H-2} \sum_{z=0}^{H-2} (2(H-1) - (H-1) - z)\alpha^z \binom{H+z-2}{z}$$

19

using the combinatorial identity $(n+k)\binom{n+k-1}{k} = (n+k)\binom{n+k-1}{n-1} = n\binom{n+k}{n} = n\binom{n+k}{k}$

$$= (H-1)(1-2\alpha)(1-\alpha)^{H-2}\sum_{z=0}^{H-2}\alpha^z\left[2\binom{H+z-2}{z} - \binom{H+z-1}{z}\right]$$

extending the sum to ∞ can only decrease the lower bound because $2\binom{n+k}{k} - \binom{n+k+1}{k} = \binom{n+k}{k} - \binom{n+k}{k-1} \le 0$ for $k > n$ and $(1-2\alpha) > 0$

$$= (H-1)(1-2\alpha)(1-\alpha)^{H-2}\sum_{z=0}^{\infty}\alpha^z\left[2\binom{H+z-2}{z} - \binom{H+z-1}{z}\right]$$

using $(1-\alpha)^{-n-1} = \sum_{k=0}^{\infty}\alpha^k\binom{n+k}{k}$ since $|\alpha| < 1$

$$= (H-1)(1-2\alpha)(1-\alpha)^{H-2}\left[2(1-\alpha)^{1-H} - (1-\alpha)^{-H}\right] = (H-1)\left(\frac{1-2\alpha}{1-\alpha}\right)^2.$$

And so we have

$$T \ge (H-1)\left(1 - \frac{1}{r}\right)^2. \tag{B.2}$$

Hence, for $r > 1$, the expected train length, T, can be made arbitrarily long by choosing a suitably long chain, H.

Appendix B.2. Train collisions

The above derivation of the expected train length neglects the possibility that two trains may collide and merge, which introduces a source of error. When two trains collide, the first finds itself being erased at greater than the expected rate, effectively reducing its train length. Therefore, collisions *decrease* the expected train length T – despite the fact that merging trains may lead to longer overall lengths – hence Eq. (B.1) overestimates the expected train length. An upper bound on our true T is found by assuming that trains never interfere with one another, and simply using Eq. (B.1). Conversely, a lower bound for T is obtained by assuming that the second train completely eradicates the first train. In order to find this lower bound we need to determine the probability that a train collision occurs. An upper bound on this collision probability is given by the probability that a second train is generated while another train is still occupying the stem. This can be formulated in terms of a negative binomial distribution where the generation of a new train counts as a "success" while a decrease in length of the existing train in the stem counts as a "failure". In each time step a new train is generated with probability $r^2/(F_t(L+r-1))$ whereas the probability that the existing train length decreases, i.e. the resident directly behind the train reproduces, is at least $1/F_t$ (exactly $1/F_t$ along the stem, but greater for the first stem node). After H failure events we know that the stem must be cleared and contain only residents. Therefore, train collisions occur

at most with the probability that a new train is generated prior to H failure events:

$$P(\text{no } 2^{nd} \text{ train}) = \left(1 - \frac{r^2}{L + r - 1 + r^2}\right)^H > 1 - H\frac{r^2}{L + r - 1 + r^2}. \quad \text{(B.3)}$$

The inequality in Eq. (B.3) results from expanding and then truncating the power term. This is permitted, because when expanding the alternating sum for sufficiently large L, the absolute value of subsequent terms is strictly decreasing. Thus, the chance that a second train is launched while another one still occupies the stem is at most

$$\epsilon_3 = \frac{Hr^2}{L + r - 1 + r^2} \quad \text{(B.4)}$$

and becomes small for $L \gg Hr^2$. Thus, the true expected train length lies between T and $T(1 - \epsilon_3)$.

Appendix C. Interaction of trains with root node

Here we calculate the probability that a train of mutants, which arrives at the base of the stem with an initial length l, succeeds in taking over the root node and placing a new mutant in one of the reservoirs. We adapt the technique used in (16), considering a finite state Markov process with two absorbing states: either a new mutant is placed in one reservoir, or the mutant train has disappeared. All other states represent a particular train length with the root node occupied by either resident or mutant.

For each state, the probability of eventually succeeding is $p_{i\updownarrow}$, where $i \leq l$ indicates the current train length and $\updownarrow$ indicates whether the root node is occupied by a mutant, $\uparrow$, or resident, $\downarrow$. Clearly $p_{0\downarrow} = 0$ because the train has disappeared, the root is a resident and hence an absorbing state has been reached. Similarly, $p_{0\uparrow} = r/(B + r)$, denotes the probability that the mutant in the root node reproduces before being replaced by the offspring of residents in any of the B branches. By examining all possible transitions we obtain:

$$(B + r)p_{i\uparrow} = (B - 1)p_{i\downarrow} + r + p_{i-1\uparrow} \quad \text{(C.1)}$$

$$(r + 1)p_{i\downarrow} = rp_{i\uparrow} + p_{i-1\downarrow} \quad \text{(C.2)}$$

If the train is i mutants long, and the root is a mutant, a "success" occurs with relative weight r, whereas the root mutant is lost, with relative weight $B - 1$ (because there are $B - 1$ other branches), or our train may erode, with relative weight 1, leaving the root node unchanged. If the root node is a resident, then the only possible actions are the replacement of the root node, or eroding the train from behind, with relatively probabilities r and 1 respectively. Finally, we have coefficients on the left hand side to normalize over all possible courses of action. Written as a matrix equation this gives:

$$\begin{bmatrix} B + r & 1 - B \\ -r & r + 1 \end{bmatrix} \begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

which yields

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \frac{1}{B+2r+r^2} \begin{bmatrix} r+1 & B-1 \\ r & r+B \end{bmatrix} \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right). \qquad \text{(C.3)}$$

We now calculate both upper and lower bounds on the expected success probability upon arrival, $\mathbb{E}(p_{l\downarrow})$. Let us start with the upper bound.

By neglecting terms in the normalization term and increasing several of the matrix entries, we find an upper bound on the R.H.S. of Eq. (C.3). This works because $p_{i\uparrow}, p_{i\downarrow} \geq 0$, and thus we are making the R.H.S. strictly more positive. It also significantly simplifies our equation to

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} < \frac{1}{B+2r+1} \begin{bmatrix} r+1 & B+r \\ r+1 & r+B \end{bmatrix} \left(\begin{bmatrix} r \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

where the inequality applies element-wise. Substituting $p_{0\uparrow} = r/(B+r), p_{0\downarrow} = 0$ into the above gives

$$p_{1\updownarrow} < \frac{r^2+r}{B+2r+1} \left(1 + \frac{1}{B+r} \right)$$

Using the fact that the upper bounds for $p_{i\uparrow}$ and $p_{i\downarrow}$ are equal, along with the fact that the denominator of the fraction is equal to the row sum of the transition matrix gives:

$$p_{i\updownarrow} < \frac{r^2+r}{B+2r+1} + \frac{B+2r+1}{B+2r+1} p_{i-1\updownarrow}.$$

By induction we find

$$p_{i\updownarrow} < \frac{r^2+r}{B+2r+1} \left(i + \frac{1}{B+r} \right).$$

This yields an upper bound for $p_{i\uparrow}$, which we then use to calculate a tighter bound for $p_{i\downarrow}$. From the last line of Eq. (C.3) we derive:

$$p_{i\downarrow} < \frac{r}{B}(r + p_{i-1\uparrow}) + p_{i-1\downarrow}$$

which leads to

$$p_{i\downarrow} < \sum_{n=1}^{n=i} \frac{r}{B}(r + p_{n-1\uparrow}) \leq \frac{ir^2}{B} \left(1 + \frac{r+1}{(B+2r+1)(B+r)} + \frac{H-1}{2} \frac{r+1}{B+2r+1} \right).$$

Thus we end up with $p_{i\downarrow} < ir^2(1 + \epsilon_{4+})/B$, where

$$\epsilon_{4+} = \frac{1+r}{(B+2r+1)(B+r)} + \frac{(H-1)(r+1)}{2B+4r+2} \qquad \text{(C.4)}$$

22

is our error term. This error term can be made arbitrarily small for sufficiently large B.

For the lower bound, we must deal with the possibility that a small number of other branches, δ, also contain mutants. Because the above train collision argument (Appendix B.2) is based on at most a single mutant existing in the reservoir of each branch, we wish to only consider reproductive events, which place new mutants into branches that currently contain no mutants and neglect the rest (which we can do, as this calculation desires a *lower* bound). Further, we must contend with the fact that trains from other mutants may compete for control of the root node. Although generically we do not expect to encounter other trains, we calculate our lower bound as if all other mutant occupied branches have mutants at the base of their stems at all times. This arrangement, while unrealistic, describes the situation which minimizes the success probability of a given train, and is thus useful for finding lower bounds. The following equations are written under the assumption of this worst case scenario (worst from the perspective of the train we are focusing on):

$$\begin{bmatrix} B+r+\delta(r-1) & 1-B-\delta(r-1) \\ -r & r+1 \end{bmatrix} \begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \left(\begin{bmatrix} r\frac{B-\delta}{B} \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right)$$

which can be rewritten as

$$\begin{bmatrix} p_{i\uparrow} \\ p_{i\downarrow} \end{bmatrix} = \frac{1}{B+\gamma(r-1)+2r+r^2} \begin{bmatrix} r+1 & B+\delta(r-1)-1 \\ r & r+B+\delta(r-1) \end{bmatrix} \left(\begin{bmatrix} r\frac{B-\delta}{B} \\ 0 \end{bmatrix} + \begin{bmatrix} p_{i-1\uparrow} \\ p_{i-1\downarrow} \end{bmatrix} \right).$$

By ignoring the positive effects of $p_{i-1\uparrow}$ on $p_{i\downarrow}$ we form the inequality

$$p_{i\downarrow} > \frac{1}{B+\delta(r-1)+2r+r^2} \left(r^2 \frac{B-\delta}{B} + (r+B+\delta(r-1))p_{i-1\downarrow} \right)$$

for $i \geq 1$. By induction this leads to

$$p_{i\downarrow} > \frac{B-\delta}{B} \frac{r^2}{B+\delta(r-1)+2r+r^2} \sum_{n=0}^{i-1} \left(\frac{B+r+\delta(r-1)}{B+\delta(r-1)+r^2+2r} \right)^n$$

$$= \frac{B-\delta}{B} \frac{r^2}{B+\delta(r-1)+2r+r^2} \frac{1-\left(\dfrac{B+\delta(r-1)+r}{B+\delta(r-1)+r^2+2r} \right)^i}{1-\dfrac{B+\delta(r-1)+r}{B+\delta(r-1)+r^2+2r}}$$

and finally to

$$p_{i\downarrow} > \frac{B-\delta}{B} \frac{r^2}{r+r^2} \left(1 - \left(1 - \frac{r^2+r}{B+\delta(r-1)+r^2+2r} \right)^i \right).$$

As long as $i(r^2+r)/(B+\delta(r-1)+r^2+2r) < 1$, the series expansion of the inner bracket gives an alternating sequence with monotone decreasing absolute terms.

23

Therefore, we can truncate the series after three terms while still preserving the inequality because the sum of the first three terms is greater than any subsequent sum. This leads to:

$$p_{i\downarrow} > \frac{B-\delta}{B} \frac{ir^2}{B+\delta(r-1)+r^2+2r} \left(1 - \frac{i-1}{2} \frac{r^2+r}{B+\delta(r-1)+r^2+2r}\right)$$

If we rearrange the above, remembering that $1/(1+x) < 1-x$ whenever $x > -1$, we find

$$p_{i\downarrow} > \frac{ir^2}{B} \left(1 - \frac{\delta}{B} - \frac{\delta(r-1)+r^2+2r}{B} - \frac{H-1}{2}\frac{r^2+r}{B} + O(B^{-2})\right).$$

We are free to drop the very small positive terms at the end, and find the result $p_{i\downarrow} > (1 - \epsilon_{4-})ir^2/B$, where the error term

$$\epsilon_{4-} = \frac{2\delta r + r^2 + 3r + H(r^2+r)}{2B} \tag{C.5}$$

can be made small whenever $\delta, H \ll B$. Thus $\mathbb{E}(p_{l\downarrow}) > Tr^2 B^{-1}(1 - \epsilon_{4-})$. Armed with the constrains $(1-\epsilon_{4-})lr^2/B < p_{l\downarrow} < (1+\epsilon_{4+})lr^2/B$, we note that $\mathbb{E}(p_{l\downarrow}) \approx \mathbb{E}(lr^2/B) = Tr^2/B$.

Appendix D. Bounds on fixation probabilities

Appendix D.1. Upper bound

Starting with a single mutant in one reservoir initially, we first must have two before we can have BL reservoir mutants. Thus, the probability of transitioning from 1 to 2 mutants in the reservoir serves as a straight forward upper bound for the mutants fixation probability. To further simplify the upper bound, we make several optimistic (from the mutants point of view) assumptions: (i) the original mutant appears in a reservoir node (ignoring ϵ_0, see Eq. (A.1)); (ii) We slightly increase the train launch probability, dividing by L rather than $L + r - 1$ (ignoring ϵ_1, see Eq. (A.3)) (iii) no detrimental effects based on our initial conditions (ignoring ϵ_2, see Eq. (A.4)); (iv) no train collisions (ignoring ϵ_3, see Appendix B.2); and finally (v) we use the upper bound for the probability that a train succeeds in producing another reservoir mutant ($(1 + \epsilon_{4+})Tr^2/B$, see Eq. (C.4)).

A single mutant in any reservoir produces a new train with a probability of at most $r^2/(F_t L)$ per time step. Subsequently, each train succeeds in placing another mutant in any reservoir with a probability of at most $(1 + \epsilon_{4+})Tr^2/B$. At the same time, the root node has a probability of at least

$$\frac{B-1}{B+r-1}\frac{1}{F_t BL}$$

24

to remove the mutant node from the reservoir. Thus, the chance of the mutant producing a successful train before being erased by the root node is *at most* ρ_{H+} with

$$\rho_H \le \rho_{H+} = \frac{Tr^4(1 + \epsilon_{4+})}{Tr^4(1 + \epsilon_{4+}) + \frac{B-1}{B+r-1}} \approx 1 - \frac{1}{Tr^4 + 1}. \qquad \text{(D.1)}$$

The approximation in Eq. (D.1) becomes exact in the limit of large B. Moreover, because $T < H$, we can replace T by H in Eq. (D.1) to obtain a simpler and more generous upper bound.

Appendix D.2. Lower bound

On the slow timescale the dynamics of the reservoir can be approximated by a random walk X_t on the number of mutants in the reservoir. Consider the random walk on the integers from 0 to BL with forward bias γ for $X_t < \delta \ll B$ and no bias for $X_t \ge \delta$. Because the chance of any particular reservoir mutant replacing any particular reservoir resident is always higher than the converse, we can be sure that some forward bias persists for all X_t. Unfortunately, because the analytic arguments in the previous sections fail for many reservoir mutants, we must make the conservative assumption that no bias applies in this region. The fixation probability of this random walk acts as a lower bound on the fixation probability of the true process. In the following, we assume that H and r are fixed and that $B, L \gg H, \delta$.

In order to determine the fixation probability of the random walk X_t, we construct a martingale, $Q(X_t)$. A martingale is a function of a random variable such that the expected value of the martingale in the next time step is equal to the current value:

$$\mathbb{E}(Q(X_{t+1})|Q(X_t)) = Q(X_t). \qquad \text{(D.2)}$$

For $Q(X_t)$ to be a martingale, we require:

$$Q(k) = \begin{cases} \dfrac{\gamma}{1+\gamma}Q(k+1) + \dfrac{1}{1+\gamma}Q(k-1) & 0 < k < \delta & \text{(forward bias)} \\[2mm] \dfrac{1}{2}Q(k+1) + \dfrac{1}{2}Q(k-1) & \delta \le k < BL & \text{(no bias).} \end{cases}$$
$$\text{(D.3)}$$

These constraints admit the solution $Q(k) = \gamma^{-k}$ for $k < \delta$, and $Q(k) = Ak + D$ for $k \ge \delta$. For $Q(k)$ to satisfy the martingale conditions as needed, we demand $\delta \in \mathbb{N}$. The constants A, D are determined by connecting the solutions for the two regions. In particular,

$$\gamma^{-\delta} = Q(\delta) = A\delta + D$$

must hold such that $Q(\delta)$ is well defined and

$$2(A\delta + D) = \gamma^{-\delta+1} + A\delta + A + D$$

to satisfy the martingale property at δ. Thus,

$$A = \gamma^{-\delta}(1 - \gamma)$$
$$D = \gamma^{-\delta}(1 - \delta(1 - \gamma))$$

which yields

$$Q(0) = 1$$
$$Q(1) = \gamma^{-1}$$
$$Q(BL) = \gamma^{-\delta} + \gamma^{-\delta}(1 - \gamma)(BL - \delta).$$

Let τ be the first time we reach one end of our random walk. Because $Q(k)$ is bounded for all relevant values of k we are able to invoke the optional stopping theorem (19, p. 210), which renders

$$Q(1) = Q(X_0) = \mathbb{E}(Q(X_\tau)) = Q(0)P(0) + Q(BL)P(BL),$$

where $P(0)$ and $P(BL)$ represent the probabilities of reaching either end of our random walk. Using $P(0) = 1 - P(BL)$ we find

$$P(BL) = \frac{Q(1) - Q(0)}{Q(BL) - Q(0)} = \frac{1 - \gamma^{-1}}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1 - \gamma)(BL - \delta)}. \tag{D.4}$$

In order to keep error terms small, we must select δ such that $\gamma^\delta \gg BL$ and $\delta \ll B, L$. As long as B and L are large and sufficiently similar, this is possible provided $\gamma > 1$ (see below). Thus, for any choice of H, and for any $r > 1$ we can select B and L such that

$$P(BL) = \frac{Q(1) - Q(0)}{Q(BL) - Q(0)} = \frac{1 - \gamma^{-1}}{1 + \epsilon_5} \tag{D.5}$$

with

$$\epsilon_5 = \gamma^{-\delta}\left((\gamma - 1)(BL - \delta) - 1\right) \ll 1.$$

In order to find an upper bound on ϵ_5, we require a lower bound on the forward bias γ. In particular, we would like to show that $\gamma > 1$ and hence that ϵ_5 can be made small. This can be seen by taking the lower bound on the production rate of successful trains, $(1 - \epsilon_1)(1 - \epsilon_3)(1 - \epsilon_{4-})Tr^4/(BLF_t)$, and comparing to our upper bound on the removal probability for reservoir mutants, $1/(BLF_t)$. This represents the eventual forward bias after the top of the stem has been replaced at least once. To account for the possibility of mutant loss before the top of the stem has been replaced we must consider ϵ_2, which acts as an additive penalty (because it only applies once per reservoir mutant). This yields

$$\gamma \geq r^4 T(1 - \epsilon_1)(1 - \epsilon_3)(1 - \epsilon_{4-}) - \epsilon_2. \tag{D.6}$$

In the limit of large B and L all error terms tend to zero. Thus, to show $\gamma > 1$, it is sufficient to show that $r^4 T > 1$. Recalling that $A_t - Z_t$ represents the length of a train at time t (see Appendix B), and noting that it is submartingale (the expected future value is greater than the current value) whenever $r > 1$, we can easily show that $T = \mathbb{E}(A_\tau - Z_\tau) \geq A_0 - Z_0 = 1$, and thus, in the limit, $\gamma \geq r^4 > 1$. Thus ϵ_5 can indeed be made arbitrarily small.

Substituting Eq. (D.6), the lower bound of γ into Eq. (D.5) yields a lower bound on $P(BL)$, which in turn provides a lower bound on the fixation probability, see Eq. (10). To simplify the lower bound substitute the lower bound for T, $(H-1)(1-1/r)^2$, into Eq. (D.6) and obtain a looser bound on fixation probability.

Appendix E. Bringing it all together

Collating Eq. (D.5) and Eq. (D.1), we find:

$$\frac{1-\epsilon_0}{1+\epsilon_5}\left(1 - \frac{1}{r^4 T(1-\epsilon_1)(1-\epsilon_3)(1-\epsilon_{4-}) - \epsilon_2}\right) \leq \rho_H \leq 1 - \frac{B-1}{(B+r-1)Tr^4(1+\epsilon_{4+}) + B - 1}$$
$$\text{(E.1)}$$

with the train length, Eq. (B.1)

$$T = (1-\alpha)^{H-2}\sum_{z=1}^{H-1}(H-z)\alpha^{z-1}\left[\binom{H-3+z-1}{z-1} - \binom{H-3+z-1}{z-2}\right],$$
$$(H-1)(1-r^{-1})^2 \leq T \leq H,$$

the chance that the initial mutant is not in the reservoir, Eq. (A.1)

$$\epsilon_0 = \frac{1+HB}{BL+1+HB},$$

Simplifying approximation in train launch probability, Eq. (A.3)

$$\epsilon_1 = \frac{r-1}{L+r-1}$$

the chance that the initial mutant is removed before it reproduces, Eq. (A.4)

$$\epsilon_2 = \frac{1}{1+BL^2},$$

the chance of train collisions, Eq. (B.4)

$$\epsilon_3 = H\frac{r^2}{L+r-1+r^2},$$

27

the lower bound for train success, Eq. (C.5)

$$\epsilon_{4-} = \frac{2\delta r + r^2 + 3r + H(r^2 + r)}{2B},$$

the upper bound for train success, Eq. (C.4)

$$\epsilon_{4+} = \frac{(r+1)}{B + 2r + 1} \left(\frac{1}{B+1} + \frac{H-1}{2} \right),$$

and finally the Martingale error term, Eq. (C.4)

$$\epsilon_5 = \gamma^{-\delta}[(\gamma - 1)(BL - \delta) - 1].$$

Because many of the error terms are dependent on both B and L it makes sense to take the limit of both simultaneously. This will force all error terms to zero as long as $B, L \gg H, \delta$ and $\epsilon_5 \to 0$. In particular, in the limit $B \to \infty$, with $\sqrt{B} - 1 < \delta \leq \sqrt{B}$ and $L = B$ all error terms tend to zero. Other variations, such as $L = B^\beta$, $\beta > 0$ are possible, although it is suspected that relations of the form $L = \beta^B$ would prove problematic, as we would then need to reconcile the bounds $\delta \ll B$ and $\gamma^\delta \gg B\beta^B$, a problem we do not run into for $L = B^\beta$.

Simpler and looser upper and lower bounds independent of T are obtained by substituting upper and lower bounds for T (respectively) into Eq. (E.1).

Appendix F. Deleterious mutations, $r < 1$

As a final remark, we note that the above arguments are based on the assumption of a beneficial mutation, $r > 1$. In addition to promoting beneficial mutations, an evolutionary amplifier must also suppress the fixation of deleterious mutations, $r < 1$. Here we argue that a deleterious mutant indeed disappears almost surely on sufficiently large superstars.

Consider a single mutant with fitness 1, in a population of residents with fitness $1/r$. Note that we can rescale fitness without changing the dynamics of the system, since fitness is never used in an absolute sense, only relative fitness matters. We next observe that all calculations performed previously with respect to a rare mutant with a fitness advantage would now apply to a resident, if it were to become rare – that is, if residents were rare, we would expect trains of residents to propagate down the stem (incrementing with probability $1/(r+1) > 1/2$ and shrinking with probability $\alpha = r/(1+r) < 1/2$). The same martingale argument that we previously used to find a lower bound on mutant fixation probability can now be used to obtain a lower bound on resident fixation probability. This time we use the random walk X_t to track the number of residents in our reservoir nodes. Thus $X_0 = BL - 1$. This leads to a formula

for the fixation probability of *residents*

$$1 - \rho_H \geq P(BL) = \frac{Q(0) - Q(BL - 1)}{Q(0) - Q(BL)}$$

$$= \frac{1 - \gamma^{-\delta} - \gamma^{-\delta}(1-\gamma)(BL - 1 - \delta)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1-\gamma)(BL - \delta)}$$

$$= 1 - \frac{\gamma^{-\delta}(\gamma - 1)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1-\gamma)(BL - \delta)}$$

and hence

$$\rho_H \leq \frac{\gamma^{-\delta}(\gamma - 1)}{1 - \gamma^{-\delta} - \gamma^{-\delta}(1-\gamma)(BL - \delta)}$$

$$\leq \gamma^{1-\delta}. \tag{F.1}$$

In the above we require γ to be the bias in favour of the *resident*. In order to calculate it we need to find the expected train lengths of resident trains. Structurally the train equation is the same as previously, but with r replaced by $1/r$ (the fitness of residents). Because $1/r$ is now greater than one, all train length arguments that previously depended on $r > 1$ for beneficial mutations can now be applied to the resident with $1/r > 1$. Thus, we can show that the expected length of resident trains is large. Because $\gamma \approx r^{-4}T$, large T leads directly to a large γ (in the residents favour), and $\gamma^{-\delta}BL \ll 1$. The Martingale argument will significantly underestimate the fixation probability for the resident and thus overestimate the probability of mutant fixation. In order to derive exact bounds on the new bias in favour of residents, γ, we must apply error terms similar to $\epsilon_2, \epsilon_3, \epsilon_{4-}$, to obtain lower bounds on the effectiveness of "resident trains". No error term equivalent to ϵ_0 arises because the possibility that the initial mutant is not placed in a reservoir node only reduces the mutant fixation probability. Thus, the fixation probability of a deleterious mutant is bounded by Eq. (F.1) and can be made arbitrarily small for sufficiently big H, L and B.