Amplifiers for the Moran Process*

Andreas Galanis[1], Andreas Göbel[1], Leslie Ann Goldberg[1], John Lapinskas[1], and David Richerby[1]

1 **Department of Computer Science, University of Oxford Parks Road, Oxford, United Kingdom, OX1 3QD**

Abstract

The Moran process, as studied by Lieberman, Hauert and Nowak, is a randomised algorithm modelling the spread of genetic mutations in populations. The algorithm runs on an underlying graph where individuals correspond to vertices. Initially, one vertex (chosen uniformly at random) possesses a mutation, with fitness $r > 1$. All other individuals have fitness 1. During each step of the algorithm, an individual is chosen with probability proportional to its fitness, and its state (mutant or non-mutant) is passed on to an out-neighbour which is chosen uniformly at random. If the underlying graph is strongly connected then the algorithm will eventually reach *fixation*, in which all individuals are mutants, or *extinction*, in which no individuals are mutants. An infinite family of directed graphs is said to be *strongly amplifying* if, for every $r > 1$, the extinction probability tends to 0 as the number of vertices increases. Strong amplification is a rather surprising property — it means that in such graphs, the fixation probability of a uniformly-placed initial mutant tends to 1 even though the initial mutant only has a fixed selective advantage of $r > 1$ (independently of n). The name "strongly amplifying" comes from the fact that this selective advantage is "amplified". Strong amplifiers have received quite a bit of attention, and Lieberman et al. proposed two potentially strongly-amplifying families — superstars and metafunnels. Heuristic arguments have been published, arguing that there are infinite families of superstars that are strongly amplifying. The same has been claimed for metafunnels. We give the first rigorous proof that there is an infinite family of directed graphs that is strongly amplifying. We call the graphs in the family "megastars". When the algorithm is run on an n-vertex graph in this family, starting with a uniformly-chosen mutant, the extinction probability is roughly $n^{-1/2}$ (up to logarithmic factors). We prove that all infinite families of superstars and metafunnels have larger extinction probabilities (as a function of n). Finally, we prove that our analysis of megastars is fairly tight — there is no infinite family of megastars such that the Moran algorithm gives a smaller extinction probability (up to logarithmic factors). Also, we provide a counter-example which clarifies the literature concerning the isothermal theorem of Lieberman et al. A full version [11] containing detailed proofs is available at `http://arxiv.org/abs/1512.05632`. Theorem-numbering here matches the full version.

1998 ACM Subject Classification F.2 [Analysis of Algorithms] G.3 [Mathematics of Computing: Probability and Statistics]: Markov processes, Probabilistic algorithms, Stochastic processes

Keywords and phrases Moran process, randomised algorithm on graphs, evolutionary dynamics.

* The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013) ERC grant agreement no. 334828. The paper reflects only the authors' views and not the views of the ERC or the European Commission. The European Union is not liable for any use that may be made of the information contained therein.

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

This paper is about a randomised algorithm called the Moran process. This algorithm was introduced in biology [16, 20] to model the spread of genetic mutations in populations. Similar algorithms have been used to model the spread of epidemic diseases, the behaviour of voters, the spread of ideas in social networks, strategic interaction in evolutionary game theory, the emergence of monopolies, and cascading failures in power grids and transport networks [2, 3, 12, 15, 17].

There has been past work about analysing the expected convergence time of the algorithm [7, 8]. In fact, the fast-convergence result of [7] implies that when the algorithm is run on an undirected graph, and the "fitness" of the initial mutation is some constant $r > 1$, there is an FPRAS for the "fixation probability", which is the probability that a randomly-introduced initial mutation spreads throughout the whole graph.

This paper answers an even more basic question, originally raised in [16], about the long-term behaviour of the algorithm when it is run on directed graphs. In particular, the question is whether there even exists an infinite family of (directed) graphs such that, when the algorithm is run on an n-vertex graph in this family, the fixation probability is $1 - o(1)$, as a function of n. A heuristic argument that this is the case was given in [16], but a counter-example to the argument (and to the hypothesized bound on the fixation probability) was given in [6]. A further heuristic argument (with a revised bound) was given in [14]. Here we give the first rigorous proof that there is indeed a family of "amplifiers" with fixation probability $1 - o(1)$. Before describing this, and the other results of this paper, we describe the model.

The Moran algorithm has a parameter r which is the fitness of "mutants". All non-mutants have fitness 1. The algorithm runs on a directed graph. In the initial state, one vertex is chosen uniformly at random to become a mutant. After this, the algorithm runs in discrete steps as follows. At each step, a vertex is selected at random, with probability proportional to its fitness. Suppose that this is vertex v. Next, an out-neighbour w of v is selected uniformly at random. Finally, the state of vertex v (mutant or non-mutant) is copied to vertex w.

If the graph is finite and strongly connected then with probability 1, the process will either reach the state where there are only mutants (known as *fixation*) or it will reach the state where there are only non-mutants (*extinction*). In this paper, we are interested in the probability that fixation is reached, as a function of the mutant fitness r, given the topology of the underlying graph. If $r < 1$ then the single initial mutant has lower fitness than the non-mutants that occupy every other vertex in the initial configuration, so the mutation is overwhelmingly likely to go extinct. If $r = 1$, an easy symmetry argument shows that the fixation probability is $\frac{1}{n}$ in any strongly connected graph on n vertices [7, Lemma 1].[1] Because of this, we restrict attention to the case $r > 1$. Perhaps surprisingly, a single advantageous mutant can have a very high probability of reaching fixation, despite being heavily outnumbered in the initial configuration.

A directed graph is said to be *regular* if there is some positive integer d so that the in-degree and out-degree of every vertex is d. In a strongly connected regular graph on n vertices, the fixation probability of a mutant with fitness $r > 1$ when the Moran algorithm is run is given by

$$\rho_{\mathrm{reg}}(r, n) = (1 - \tfrac{1}{r})/(1 - \tfrac{1}{r^n}), \tag{1}$$

[1] The result is stated in [7] for undirected graphs but the proof goes through unaltered for strongly connected directed graphs.

so the extinction probability of such a mutant is given by

$$\zeta_{\mathrm{reg}}(r,n) = (\tfrac{1}{r} - \tfrac{1}{r^n})/(1 - \tfrac{1}{r^n}). \tag{2}$$

Thus, in the limit, as n tends to ∞, the extinction probability tends to $1/r$. To see why (1) and (2) hold, note that, for every configuration of mutants, the number of edges from mutants to non-mutants is the same as the number of edges from non-mutants to mutants. Suppose that the sum of the individuals' fitnesses is W and consider an edge (u, v). If u is a mutant in the current state, it is selected to reproduce with probability r/W, and, if this happens, the offspring is placed at v with probability $1/d$. Similarly, if u is not a mutant, reproduction happens along (u, v) with probability $1/(dW)$. So, in any state, the number of mutants is r times as likely to increase at the next step of the process as it is to decrease. If we observe the number of mutants every time it changes, the resulting stochastic process is a random walk on the integers, that starts at 1, absorbs at 0 and n, increases with probability $\frac{r}{r+1}$ and decreases with probability $\frac{1}{r+1}$. It is well known that this walk absorbs at n with probability (1) and at 0 with probability (2). In particular, the undirected n-vertex complete graph is regular. Thus, by (2), its extinction probability tends to $1/r$.

When the Moran process is run on non-regular graphs the extinction probability may be quite a bit lower than $1/r$. Consider the undirected $(n + 1)$-vertex "star" graph, which consists of single centre vertex that is connected by edges to each of n leaves. In the limit as $n \to \infty$, the n-leaf star has extinction probability $\frac{1}{r^2}$ [5, 16]. Informally, the reason that the extinction probability is so small is that the initial mutant is likely to be placed in a leaf, and, at each step, a mutation at a leaf is relatively unlikely to be overwritten.

Lieberman et al. [16] refer to graphs which have smaller extinction probability than (2) (and therefore have larger fixation probability than (1)) as *amplifiers*. The terminology comes from the fact that the selective advantage of the mutant is being "amplified" in such graphs.

The purpose of this paper is to explore the long-term behaviour of the Moran process by quantifying how good amplifiers can be. For this, it helps to have some more formal definitions.

▶ **Definition 1.** Consider a function $\zeta(r, n) \colon \mathbb{R}_{>1} \times \mathbb{Z}_{\geq 1} \to \mathbb{R}_{\geq 0}$. An infinite family Υ of directed graphs is said to be *up-to-ζ fixating* if, for every $r > 1$, there is an n_0 (depending on r) so that, for every graph $G \in \Upsilon$ with $n \geq n_0$ vertices, the following is true: When the Moran process is run on G, starting from a uniformly-random initial mutant, the extinction probability is at most $\zeta(r, n)$.

Equation (2) demonstrates that the infinite family of strongly-connected regular graphs is up-to-ζ_{reg} fixating and since $\zeta_{\mathrm{reg}} \leq 1/r$, this family is also up-to-$1/r$ fixating. Informally, an infinite family of graphs is said to be *amplifying* if it is up-to-ζ fixating for a function $\zeta(r, n)$ which is "smaller" than $\zeta_{\mathrm{reg}}(r, n)$. Here is the formal definition.

▶ **Definition 2.** An infinite family of directed graphs is *amplifying* if it is up-to-ζ fixating for a function $\zeta(r, n)$ which, for every $r > 1$, satisfies $\lim_{n\to\infty} \zeta(r, n) < 1/r$.

The infinite family of graphs containing all undirected stars (which can be viewed as directed graphs with edges in both directions) is up-to-$\zeta(r, n)$ fixating for a function $\zeta(r, n)$ satisfying $\lim_{n\to\infty} \zeta(r, n) = 1/r^2$, so this family of graphs is amplifying.

Lieberman et al. [16] were interested in infinite families of digraphs for which the extinction probability tends to 0, prompting the following definition.

▶ **Definition 3.** An infinite family of directed graphs is *strongly amplifying* if it is up-to-ζ fixating for a function $\zeta(r, n)$ which, for every $r > 1$, satisfies $\lim_{n\to\infty} \zeta(r, n) = 0$.

Note that the infinite family of undirected stars is not strongly amplifying since the extinction probability of stars tends to $1/r^2$ rather than to 0.

Prior to this paper, there was no (rigorous) proof that a strongly amplifying family of digraphs exists (though there were heuristic arguments, as we explain later). Proving rigorously that there is an infinite family of directed graphs that is strongly amplifying for the Moran algorithm is one of our main contributions.

Lieberman et al. [16] produced good intuition about strong amplification and defined two infinite families of graphs — superstars and metafunnels — from which it turns out that strongly amplifying families can be constructed. It is extremely difficult to analyse the Moran process on these families, due mostly to the complexity of the graphs, and the difficulty of dealing with issues of dependence and concentration. Thus, all previous arguments have been heuristic. For completeness, we discuss these heuristic arguments in Section 6.

In this paper, we define a new family of digraphs called megastars. The definition of megastars is heavily influenced by the superstars of Lieberman et al. Our main theorem is the following.

▶ **Theorem 4.** *There exists an infinite family of megastars that is strongly amplifying.*

Megastars are not easier to analyse than superstars or metafunnels. The reason for our focus on this class of graphs is that it turns out to be provably better amplifying than any of the previously-proposed families. We will present several theorems along these lines. Before doing so, we define the classes of graphs.

2 Metafunnels, superstars and megastars

2.1 Metafunnels

We start by defining the metafunnels of [16]. Let k, ℓ and m be positive integers. The (k,ℓ,m)-*metafunnel* is the directed graph $\mathcal{G}_{k,\ell,m}$ defined as follows. (See Figure 1.)

The vertex set $V(\mathcal{G}_{k,\ell,m})$ is the union of $k+1$ disjoint sets $V_0,\dots,V_k$. The set V_0 contains the single vertex v^* which is called the *centre vertex*. For $i\in[k]$, V_i is the union of ℓ disjoint sets $V_{i,1},\dots,V_{i,\ell}$, each of which has size m^i. The edge set of $\mathcal{G}_{k,\ell,m}$ is

$$(V_0\times V_k)\cup(V_1\times V_0)\cup\bigcup_{i\in[k-1]}\bigcup_{j\in[\ell]}(V_{i+1,j}\times V_{i,j})\,.$$

Lieberman et al. refer to metafunnels with $\ell=1$ as "funnels".

2.2 Superstars

We next define the superstars of [16]. Let k, ℓ and m be positive integers. The (k,ℓ,m)-*superstar* is the directed graph $\mathcal{S}_{k,\ell,m}$ defined as follows. (See Figure 2.) The vertex set $V(\mathcal{S}_{k,\ell,m})$ of $\mathcal{S}_{k,\ell,m}$ is the disjoint union of ℓ size-m sets $R_1,\dots,R_\ell$ (called *reservoirs*) together with $k\ell$ vertices $v_{1,1},v_{1,2},\dots,v_{\ell,k}$ and a single centre vertex v^*. The edge set of $\mathcal{S}_{k,\ell,m}$ is given by

$$E(\mathcal{S}_{k,\ell,m})=\bigcup_{i=1}^{\ell}\Big((\{v^*\}\times R_i)\cup(R_i\times\{v_{i,1}\})\cup\{(v_{i,j},v_{i,j+1})\mid j\in[k-1]\}\cup\{(v_{i,k},v^*)\}\Big).$$

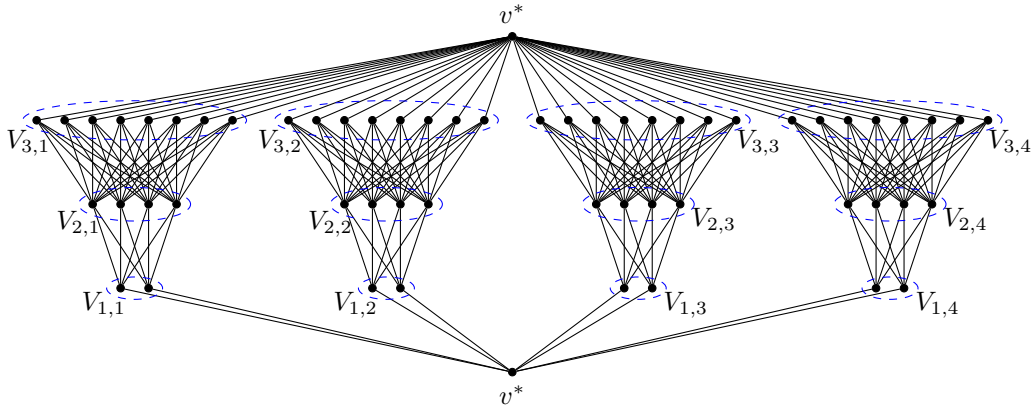

Figure 1 The metafunnel $\mathcal{G}_{3,4,2}$. All edges are directed downwards in the diagram and the centre vertex v^* is shown twice, once at the top and once at the bottom of the diagram. There are $\ell = 4$ copies of the basic unit, each of which consists of $k = 3$ levels $V_{1,j}$, $V_{2,j}$ and $V_{3,j}$, with $|V_{i,j}| = m^i = 2^i$.

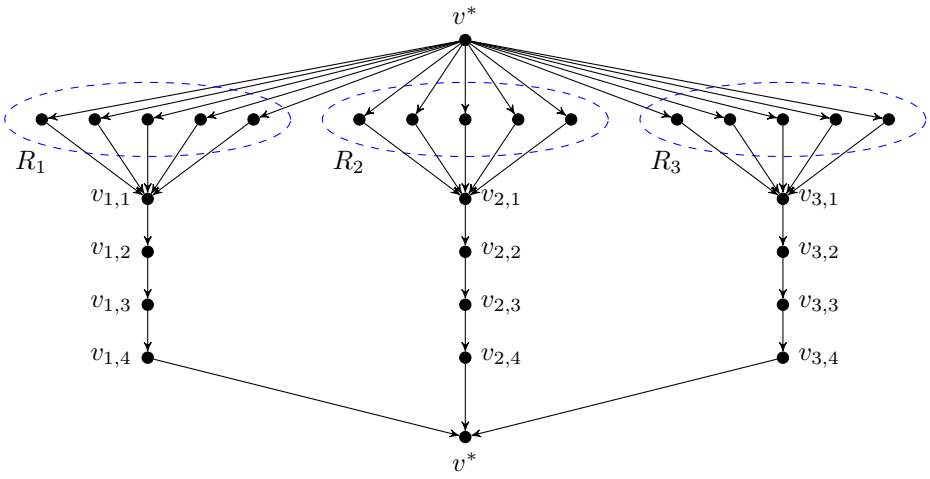

Figure 2 The superstar $\mathcal{S}_{4,3,5}$, with $\ell = 3$ reservoirs R_1, R_2 and R_3, each of size $m = 5$, connected by a path with $k = 4$ vertices to v^*. The centre vertex v^* is shown twice, at both the top and bottom of the diagram.

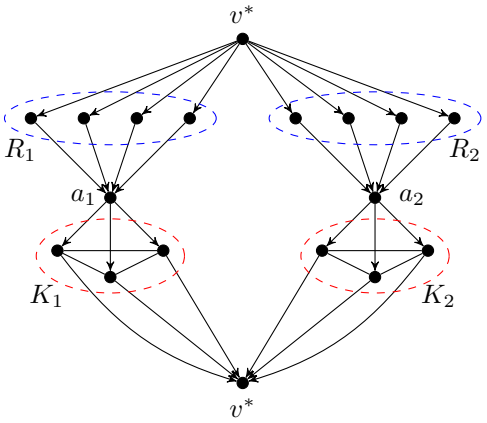

Figure 3 The megastar $\mathcal{M}_{3,2,4}$, with $\ell = 2$ reservoirs R_1 and R_2, each of size $m = 4$. Each reservoir R_i is attached, via the feeder vertex a_i to a clique of size $k = 3$. The centre vertex v^* is shown twice, once at the top and once at the bottom of the diagram. The edges within the cliques K_1 and K_2 are bidirectional.

2.3 Megastars

Finally, we define the new class of megastars, which turn out to be provably-better amplifiers than either metafunnels or superstars. The intuition behind the design of this class of graphs is that the path $v_{i,1}v_{i,2}\dots v_{i,k}$ linking the i'th reservoir R_i of a superstar to the centre vertex v^* is good for amplifying but that a clique is even better.

Let k, ℓ and m be positive integers. The (k, ℓ, m)-megastar is the directed graph $\mathcal{M}_{k,\ell,m}$ defined as follows. (See Figure 3.) The vertex set $V(\mathcal{M}_{k,\ell,m})$ of $\mathcal{M}_{k,\ell,m}$ is the disjoint union of ℓ sets $R_1, \dots, R_\ell$ of size m, called reservoirs, ℓ sets $K_1, \dots, K_\ell$ of size k, called cliques, ℓ "feeder vertices" $a_1, \dots, a_\ell$ and a single centre vertex v^*. The edge set of $\mathcal{M}_{k,\ell,m}$ consists of the following edges:

- an edge from v^* to every vertex in $R_1 \cup \dots \cup R_\ell$,
- for each $i \in [\ell]$, an edge from each vertex in R_i to a_i,
- for each $i \in [\ell]$, an edge from a_i to each vertex in K_i,
- for each $i \in [\ell]$, edges in both directions between every pair of distinct vertices in K_i,
- an edge from every vertex in $K_1 \cup \dots \cup K_\ell$ to v^*.

3 Our results

Our main result is that there is an infinite family of megastars that is strongly amplifying, so we start by defining this family. Although megastars are parameterised by three parameters, k, ℓ and m, the megastars in the family that we consider have a single parameter ℓ, so we define k and m to be functions of ℓ.

▶ **Definition 5.** Let $m(\ell) = \ell$ and $k(\ell) = \lceil (\log \ell)^{23} \rceil$. Let $\Upsilon_{\mathcal{M}} = \{\mathcal{M}_{k(\ell),\ell,m(\ell)} \mid \ell \in \mathbb{Z}, \ell \geq 2\}$.

Our main result can then be stated as follows.

▶ **Theorem 6.** *Let $\zeta_{\mathcal{M}}(r, n) = (\log n)^{23} n^{-1/2}$. The family $\Upsilon_{\mathcal{M}}$ is up-to-$\zeta_{\mathcal{M}}$ fixating.*

▶ **Corollary 7.** *The family $\Upsilon_{\mathcal{M}}$ is strongly amplifying.*

The proof of Theorem 6 requires a complicated analysis, accounting for dependencies and concentration. The theorem, as stated here, follows directly from Theorem 75 of the full version.

The reason that we studied megastars rather than the previously-introduced superstars and metafunnels is that megastars turn out to be provably better amplifying than any of the previously-proposed families. To demonstrate this, we prove the following theorem about superstars.

▶ **Theorem 8.** *Suppose that $\zeta(r,n)$ is any function such that, for any $r > 1$, we have $\lim_{n\to\infty} \zeta(r,n)(n\log n)^{1/3} = 0$. Then there is no infinite family of superstars that is up-to-ζ fixating.*

The function $\zeta_{\mathcal{M}}(r,n)$ from Theorem 6 certainly satisfies $\lim_{n\to\infty} \zeta_{\mathcal{M}}(r,n)(n\log n)^{1/3} = 0$, so Theorem 8 shows that there is no infinite family of superstars that is up-to-$\zeta_{\mathcal{M}}$ fixating. More mundanely, it shows, for example, that if $\zeta(r,n) = n^{-1/3}(\log n)^{-1}$, then no infinite family of superstars is up-to-ζ fixating. Theorem 8 is a direct consequence of Theorem 29 of the full version.

Taken together, Theorems 6 and 8 show that superstars are worse amplifiers than megastars. We next show that metafunnels are *substantially* worse. We start with the following simple-to-state theorem.

▶ **Theorem 9.** *Fix any $\delta > 0$ and let $\zeta(r,n) = n^{-\delta}$. Then there is no infinite family of metafunnels that is up-to-ζ fixating.*

In fact, Theorem 9 can be strengthened by an exponential amount.

▶ **Theorem 10.** *Fix any $\epsilon < 1/2$ and let $\zeta(r,n) = n^{-1/(\log n)^{\epsilon}}$. Then there is no infinite family of metafunnels that is up-to-ζ fixating.*

Theorems 9 and 10 are a direct consequence of Theorem 47 in the full version. In fact, Theorem 47 provides even tighter bounds, as we will see in Section 5, though these are more difficult to state.

The theorems that we have already described (Theorem 6, Theorem 8 and Theorem 10) are the main contributions of the paper. Together, they show that there is a family of megastars that is strongly amplifying, and that there are no families of superstars or metafunnels that amplify as well. For completeness, we present a theorem showing that the analysis of Theorem 6 is fairly tight, in the sense that there are no infinite families of megastars that amplify substantially better than $\Upsilon_{\mathcal{M}}$ — in particular, our bound on extinction probability can only be improved by factors of $\log(n)$. It cannot be improved more substantially.

▶ **Theorem 11.** *Let $\zeta(r,n) = n^{-1/2}/(52r^2)$. There is no infinite family of megastars that is up-to-ζ fixating.*

Theorem 11 follows from Theorem 119 in the full version, which is straightforward. We conclude the paper with a digression which perhaps clarifies the literature. It is stated, and seems to be commonly believed, that an evolutionary graph (a weighted version of the Moran process — see Section 8 of the full version for details) is "isothermal" if and only if the fixation probability of a mutant placed uniformly at random is $\rho_{\text{reg}}(r,n)$. This belief seems to have come from an informal statement of the "isothermal theorem" in the main body of [16] (the formal statement in the supplementary material of [16] is correct, however) and it has spread, for example, as Theorem 1 of [22]. We clear this up by proving the following proposition, which says that there is a counter-example.

▶ Proposition 12. There is an evolutionary graph that is not isothermal, but has fixation probability $\rho_{\text{reg}}(r, n)$.

4 Proof techniques

As we have seen, it is easy to study the Moran process on a d-regular graph by considering the transition matrix of the corresponding Markov chain (which looks like a one-dimensional random walk). Highly symmetric graphs such as undirected stars can also be handled in a straightforward matter, by directly analysing the transition matrix. Superstars, metafunnels and megastars are more complicated, and the number of mutant-configurations is exponential, so instead we resort to dividing the process into phases, as is typical in the study of randomised algorithms and stochastic processes.

An essential and common trick in the area of stochastic processes (for example, in work on the voter model) is moving to continuous time. Instead of directly studying the discrete-time Moran process, one could consider the following natural continuous-time model which was studied in [8]: Given a set of mutants at time t, each vertex waits an amount of time before reproducing. For each vertex, the period of time before the next reproduction is chosen according to the exponential distribution with parameter equal to the vertex's fitness, independently of the other vertices. If the first vertex to reproduce is v at time $t + \tau$ then, as in the standard, discrete-time version of the process, one of its out-neighbours w is chosen uniformly at random, the individual at w is replaced by a copy of the one at v, and the time at which w will next reproduce is exponentially distributed with parameter given by its new fitness. The discrete-time process is recovered by taking the sequence of configurations each time a vertex reproduces. Thus, the fixation probability of the discrete-time process is *exactly the same* as the fixation probability of the continuous-time process. So moving to the continuous-time model causes no harm. As [8] explains, analysis can be easier in the continuous-time model because certain natural stochastic domination techniques apply in the continuous-time setting but not in the discrete-time setting.

It turns out that moving to the model of [8] does not suffice for our purposes. A major problem in our proofs is dealing with dependencies. In order to make this feasible, we instead study a continuous-time model (see "the clock process" in Section 3.1 of the full version) in which every edge of the underlying graph G is equipped with two Poisson processes, one of which is called a *mutant clock* and the other of which is called a *non-mutant clock*. The clock process is a stochastic process in which all of these clocks run independently. The continuous-time Moran process as defined in [8] can be recovered as a function of the times at which these clocks trigger.

Having all of these clocks available still does not give us the flexibility that we need. We say that a vertex u "spawns a mutant" in the Moran process if, at some point in time, u is a mutant, and it is selected for reproduction. We wish to be able to discuss events such as the event that the vertex u does not spawn a mutant until it has already been a mutant for some particular amount of time. In order to express such events in a clean way, making all conditioning explicit, we define additional stochastic processes called "star-clocks" (see Section 3.3 of the full version). All of the star-clocks run independently in the star-clock process.

In Section 3.4 of the full version we provide a coupling of the star-clock process with the Moran process. The coupling is valid in the sense that the two projections are correct — the projection onto the Moran process runs according to the correct distribution and so does the projection onto the star-clock process. The point of the coupling is that the different

star-clocks can be viewed as having their own "local" times. In particular, there is a star-clock $M^*_{(u,v)}$ which controls reproductions from vertex u onto vertex v *during the time that* u *is a mutant*. Similarly, there is a star-clock $N^*_{(u,v)}$ which controls reproductions from vertex u onto vertex v during the time that u is a non-mutant. The coupling enables us to focus on relevant parts of the stochastic process, making all conditioning explicit.

The processes that we have described so far are all that we need to derive our upper bound on the fixation probability of superstars (Section 4 of the full version). This is the easiest of our main results.

Analysing the Moran process on metafunnels is more difficult. By design, the initial mutant x_0 is likely to be placed in the "top of a funnel" (in the set V_k). In the analysis, it is useful to be able to create independence by considering a "strain" of mutants which contains all of the descendants of a particular mutant spawned by x_0. Like the Moran process itself, a strain can be viewed as a stochastic process depending on the triggering of the clocks. In order to facilitate the proof, we define a general notion of "mutant process" (Section 3.2 of the full version) — so the Moran process is one example of a mutant process, and a strain is another. The analysis of the Moran process on metafunnels involves both of these and also a third mutant process which is essentially the bottom level of a strain (called its head). Strains and heads-of-strains share some common properties, and they are analysed together as "colonies" in Section 5.4.1 of the full version. The analysis of the metafunnel is the technically most difficult of our results so we discuss it further in the next section.

Fortunately, the analysis of the megastar in Section 6 of the full version does not require three different types of mutant processes — it only requires one. The process that is considered is not the Moran process itself. Instead, it is a modification of the Moran process called the megastar process. The megastar process is similar to the Moran process except that the feeder vertices are forced to be non-mutants, except when their corresponding cliques are completely full or completely empty. It is easy to show (see the proof of Theorem 75 of the full version) that the fixation probability of the Moran process is at least as high as the fixation probability of the megastar process. However, the megastar process is somewhat easier to analyse because the cliques evolve somewhat independently. The proof of the key lemma (Lemma 77 of the full version) is fairly long but it is not conceptually difficult. The point is to prove that, with high probability, the cliques fill up and cause fixation.

5 Sketch of the analysis of metafunnels

The proofs of this paper are fairly technical, so the full version is 98 pages. In order to give a flavour, we give a brief sketch of the proof of Theorem 10 which is our most difficult result. We use $n = 1 + \ell \sum_{i=1}^{k} m^i$ to denote the number of vertices of a (k, ℓ, m)-metafunnel. We use X_t to denote the set of mutants at time t and x_0 to denote the initial mutant so $X_0 = \{x_0\}$. We prove the following theorem, which is slightly stronger than Theorem 10, and implies it.

▶ **Theorem 47.** *Let $r > 1$. Then there is a constant $c_r > 0$, depending on r, such that the following holds for all $k, \ell, m \in \mathbb{Z}_{\geq 1}$ such that the (k, ℓ, m)-metafunnel $\mathcal{G}_{k,\ell,m}$ has $n \geq 3$ vertices. Suppose that the initial state X_0 of the Moran process with fitness r is chosen uniformly at random from all singleton subsets of $V(\mathcal{G}_{k,\ell,m})$. The probability that the Moran process goes extinct is at least $e^{-\sqrt{\log r \cdot \log n}} (\log n)^{-c_r}$.*

Here is the sketch of the proof of Theorem 47. If $k = 1$ then $\mathcal{G}_{k,\ell,m}$ is a star and has extinction probability roughly $1/r^2$ so Theorem 47 follows easily. So for most of the proof (and the rest of this sketch) we assume $k \geq 2$. To prove the theorem, we divide the parameter space into two regimes.

In the (straightforward) first regime, $m \leq r^{\sqrt{\log_r n}}$. Since m is small, V_k is not too large compared to $V_0 \cup \cdots \cup V_{k-1}$. Thus, it is fairly likely that x_0 is born outside V_k, and becomes a non-mutant ("dies") before it can spawn a single mutant.

Most of the proof focusses on the second regime, where $m \geq r^{\sqrt{\log_r n}}$ which, since $n \geq \ell m^k$, implies $k \leq \sqrt{\log_r n}$. In this regime it is likely that a uniformly-chosen initial mutant x_0 is born in V_k (Lemma 49) so we assume that this is the case in most of the proof (and the rest of this sketch). The key lemma is Lemma 74 which shows that, in this case, it is (sufficiently) likely that x_0 dies before v^* spawns a mutant.

In more detail, we define a stopping time T_{pa} which is the first time t that one of the following occurs.

(A1) $X_t = \emptyset$, or
(A2) $|X_t|$ exceeds a given threshold m^* which is a polynomial in $\log n$, or
(A3) By time t, v^* has already become a mutant more than b^* times, where b^* is about half as large as its number ℓm of in-neighbours, or
(A4) t exceeds some threshold t_{max} which is (very) exponentially large in n.

The subscript "pa" is for "pseudo-absorption time" because (A1) implies that the Moran process absorbs by going extinct and (A2) is a prerequisite for absorbing by fixating. The proof of Lemma 74 shows that, with sufficiently high probability, (A2)–(A4) do not hold, and so the Moran process X must go extinct by time T_{pa}.

Conditioning makes it difficult to prove that (A2)–(A4) fail. To alleviate this, we divide the mutants into groups called "strains" which are easier to analyse. In particular, a strain contains all of the descendants of a particular mutant spawned by x_0. Informally, S^i is "born" at the i'th time at which x_0 spawns a mutant. It "dies" when all of the descendants of this spawn have died. It is "dangerous" if one (or more) of these descendants spawns a mutant onto v^* before T_{pa}.

Lemma 53 defines eight events $\mathcal{P}_1$–$\mathcal{P}_8$. These are defined in such a way that we can show (in the proof of Lemma 74) that if $\mathcal{P}_1$–$\mathcal{P}_8$ simultaneously occur, then (A2)–(A4) do not hold. The definitions are engineered in such a way that we can also show that it is fairly likely that they do hold simultaneously — this takes up most of the proof. Informally, the events are defined as follows.

$\mathcal{P}_1$: No star-clock $M^*_{(v^*,v)}$ triggers in the time interval $[0, 1]$.
$\mathcal{P}_2$: For some threshold $t_{x_0} < n$, the star-clock $N^*_{(v^*,x_0)}$ triggers in $[0, t_{x_0} - 2]$.
$\mathcal{P}_3$: v^* is a mutant for at most one unit of time up to time T_{pa}.
$\mathcal{P}_4$: The Moran process absorbs (either fixates or goes extinct) by time $t_{\mathsf{max}}/2$.
$\mathcal{P}_5$: Break $[0, t_{x_0}]$ into intervals of length $(\log n)^2$. During each interval, x_0 spawns at most $2r(\log n)^2$ mutants.
$\mathcal{P}_6$: Define s to be around $3rt_{x_0}$. Each of the strains $S^1, \ldots, S^s$ spawns at most $\log n$ mutants before T_{pa}.
$\mathcal{P}_7$: Each of the strains $S^1, \ldots, S^s$ dies within $(\log n)^2$ steps.
$\mathcal{P}_8$: At most $b^*/\log n$ of $S^1, \ldots, S^s$ are dangerous.

The rough sketch of Lemma 74 is as follows. $\mathcal{P}_1$ and $\mathcal{P}_3$ guarantee that v^* does not spawn a mutant until after T_{pa}. This together with $\mathcal{P}_2$ and $\mathcal{P}_3$ guarantees that the only mutants in the process before time T_{pa} are part of strains that are born before t_{x_0}. By $\mathcal{P}_5$, there are at most s such strains. By $\mathcal{P}_6$, each of these strains only has about $\log n$ mutants. Together with $\mathcal{P}_7$, this implies that (A2) does not hold at $t = T_{\mathsf{pa}}$. $\mathcal{P}_8$ and $\mathcal{P}_6$ imply that (A3) does not hold at $t = T_{\mathsf{pa}}$. Finally, $\mathcal{P}_4$ implies that (A4) does not hold at $t = T_{\mathsf{pa}}$.

The bulk of the proof involves showing (Lemma 53) that $\mathcal{P}_1$–$\mathcal{P}_8$ are sufficiently likely to simultaneously occur. Of these, $\mathcal{P}_3$–$\mathcal{P}_7$ are all so likely to occur that the probability that they

do not occur can be subtracted off using a union bound (so conditioning on the other $\mathcal{P}_i$'s is not an issue). The majority of the failure probability comes from the probability that $\mathcal{P}_2$ does not occur. This is handled in the straightforward Lemma 54 which gives a lower bound on the probability that $\mathcal{P}_1$ and $\mathcal{P}_2$ both occur. The remaining event, $\mathcal{P}_8$, is sufficiently unlikely to occur that careful conditioning is required. This is (eventually) handled in Lemma 73, which shows that it is fairly likely to occur, conditioned on the fact that both $\mathcal{P}_1$ and $\mathcal{P}_2$ occur.

In order to get a good estimate on the probability that a strain is dangerous (in $\mathcal{P}_8$), we need to consider the number of mutants spawned from the "layer" of the strain closest to the centre vertex v^*. In order to do this, we define a new mutant process called the "head" of a strain. Strains and heads of strains share some common properties, and they are analysed together as "colonies". Informally a "colony" is a mutant process Z whose mutants are in $V_1 \cup \cdots \cup V_{k-1}$ (and not in V_0 or V_k). Once a colony becomes empty, it stays empty. Since a colony is a mutant process but not necessarily a Moran process, vertices may enter and/or leave whenever a clock triggers but we say that the colony is *hit* when a vertex leaves a colony specifically because a non-mutant is spawned onto it in the underlying Moran process. We define the "spawning chain" Y^Z of a colony and show that it increases whenever the colony spawns a mutant and that it only decreases when the colony is hit. By analysing the jump chain of a spawning chain we are able to obtain the desired bounds on the probability that $\mathcal{P}_6$, $\mathcal{P}_7$ and $\mathcal{P}_8$ fail to occur.

6 Comparison with previous work

The Moran process is similar to a discrete version of directed percolation known as the contact process. There is a vast literature (e.g., [9, 10, 17, 21]) on the contact process and other related infection processes such as the voter model and susceptible-infected-susceptible (SIS) epidemic models. Often, the questions that are studied in these models are different from the question that we study here. For example, in voter systems [9] the two states (mutant/non-mutant) are often symmetric (similar to our $r = 1$ case) and the models are often studied on infinite graphs where the question is whether the process absorbs or not (both kinds of absorption, fixation and extinction, are therefore called "fixation" in some of this work). The particular details of the Moran process are very important for us because the details of the algorithm determine the long-term behaviour. For example, unlike the Moran process, in the contact process [4], the rate at which a node becomes a non-mutant is typically taken to be 1, whereas the rate at which a node becomes a mutant is proportional to the number of mutant neighbours. In the discrete-time versions of many commonly-studied models, a node is chosen randomly at each step for replacement, rather than (as in the Moran process) for reproduction. In any case, the important point for us is that the details of the algorithm are important — results do not carry over from one algorithm to the other. Therefore, we concentrate in this section on previous work about calculating the fixation probability of the Moran process itself.

Lieberman et al. [16] studied the fixation probability of the Moran process and introduced superstars and metafunnels. Intuitively, a superstar is a good amplifier because (as long as m is sufficiently large) the initial mutation is likely to be placed in a reservoir and (as long as ℓ is sufficiently large) this is unlikely to be killed quickly by the centre vertex. Moreover, the paths of a superstar are good for amplifying the selective advantage of mutants because, after the infection spreads from a reservoir vertex to the beginning of a path, it is likely to "pick up momentum" as it travels down the path, arriving at the centre vertex as a chain of $\Theta(k)$

mutants (which, taken together, are more likely to cause the centre to spread the infection than a single mutant arriving at the centre would be). As we have seen (Theorems 6 and 8) megastars are provably better for amplification than superstars. The reason for this is that a clique is substantially better than a path at doing this "amplification". Nevertheless, the amplifying properties of superstars strongly influenced our decision to study megastars.

Lieberman et al. [16, Equation (2)] claimed[2] that for sufficiently large n, the fixation probability of a superstar with parameter k tends to $1 - r^{-(k+2)}$, and that "similar results hold for the funnel and metafunnel". They provided a heuristic sketch-proof for the superstar, but not for the funnel or metafunnel. Hauert [13, Equation (5)] claims specifically that the fixation probability of funnels tends to $1 - r^{-(k+1)}$. As far as we know, no heuristic arguments have been given for funnels or metafunnels.

In any event, Díaz, Goldberg, Mertzios, Richerby, Serna and Spirakis [6] showed that the $1 - r^{-(k+2)}$ claim for superstars is incorrect for the case $k = 3$. In particular, for this case they showed that the fixation probability is at most $1 - \frac{r+1}{2r^5+r+1}$, which is less than the originally claimed value of $1 - r^{-5}$ for all $r \geq 1.42$.

Subsequently, Jamieson-Lane and Hauert [14, Equation (5)] made a more detailed but still heuristic[3] analysis of the fixation probability of superstars. They claim that for superstars with parameter k and with $\ell = m$, the fixation probability ρ_k has the following bounds for fixed $r > 1$,

$$1 - 1/(r^4(k-1)(1-\tfrac{1}{r})^2) - o(1) \leq \rho_k \leq 1 - 1/(r^4(k-1)) + o(1), \tag{3}$$

where the $o(1)$ terms tend to 0 as $\ell \to \infty$. They claim that their bounds are a good approximation as long as $k \ll \ell = m \sim \sqrt{n}$. It is not clear exactly what "$\ll$" means in this context. Certainly there are parameter regimes where $k = o(\ell)$ and $\ell = m \sim \sqrt{n}$ but nevertheless the extinction probability is much larger than the proposed upper bound $1/(r^4(k-1)(1-1/r)^2)$ from (3). For example, suppose that $\ell = m = k^{3/2}$. In this case (see Lemma 30 of the full version), the extinction probability is at least $k/(2r(m+k)) = 1/(2r(k^{1/2}+1))$ which is larger than $1/(r^4(k-1)(1-1/r)^2)$ for all sufficiently large k. Nevertheless, the bounds proposed by Jamieson-Lane and Hauert (3) seem to be close to the truth when k is very small compared to ℓ and m.

Our Corollary 34 of the full version identifies a wide class of parameters for which the extinction probability is provably at least $1/(1470r^4k)$. This is weaker than the suggested bound of Jamieson-Lane and Hauert by a factor of 1470. This constant factor is explained by the fact that our rigorous proof needs to show concentration of all random variables. We use lots of Chernoff bounds and other bounds on probabilities. In writing the proof, we optimised readability rather than optimising our constants, so our constants can presumably be improved.

There is recent work on other related aspects of the Moran process. For example, [18, 19] give fixation probability bounds on connected *undirected* graphs. [1] studies amplification with respect to adversarial or "temperature-based" placement of the initial mutation, in which the "temperature" of a vertex is proportional to the sum of all incoming edge weights. Also, [19] considers the extent to which the number of "good starts" for fixation can be bounded.

---

[2] The reader who consults [16] might wonder why "k" as written in Equation (2) of [16] has become $k+2$ here. The reason is just that we use a slightly different parameterisation from that of [16]. To allow appropriate comparison, we describe all previous work using our parameterisation.

[3] A full discussion of the argument of Jamieson-Lane and Hauert (and of the obstacles to making it a rigorous proof) are discussed in Section 9 of the full version.

References

1 B. Adlam, K. Chatterjee, and M. A. Nowak. Amplifiers of selection. *Proceedings of the Royal Society A*, 471(2181), 2015.

2 Chalee Asavathiratham, Sandip Roy, Bernard Lesieutre, and George Verghese. The influence model. *IEEE Control Systems*, 21(6):52–64, 2001.

3 Eli Berge. Dynamic monopolies of constant size. *Journal of Combinatorial Theory, Series B*, 83(2):191–200, 2001.

4 Carol Bezuidenhout and Geoffrey Grimmett. The critical contact process dies out. *Ann. Probab.*, 18(4):1462–1482, 10 1990.

5 M. Broom and J. Rychtár. 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, 2008.

6 J. Díaz, L. A. Goldberg, G. B. Mertzios, D. Richerby, M. J. Serna, and P. G. Spirakis. On the fixation probability of superstars. *Proceedings of the Royal Society A*, 469(2156):20130193, 2013.

7 J. Díaz, L. A. Goldberg, G. B. Mertzios, D. Richerby, M. J. Serna, and P. G. Spirakis. Approximating fixation probabilities in the generalised Moran process. *Algorithmica*, 69(1):78–91, 2014.

8 Josep Díaz, Leslie Ann Goldberg, David Richerby, and Maria J. Serna. Absorption time of the Moran process. *Random Structures and Algorithms*, To Appear.

9 Richard Durrett and Jeffrey E. Steif. Fixation results for threshold voter systems. *Ann. Probab.*, 21(1):232–247, 01 1993.

10 Rick Durrett. Some features of the spread of epidemics and information on random graphs. *Proceedings of the National Academy of Science*, 107(10):4491–4498, 2010.

11 Andreas Galanis, Andreas Göbel, Leslie Ann Goldberg, John Lapinskas, and David Richerby. Amplifiers for the Moran process. Preprint.

12 Herbert Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, 2000.

13 C. Hauert. Evolutionary dynamics. In A. T. Skjeltorp and A. V. Belushkin, editors, *Proceedings of the NATO Advanced Study Institute on Evolution from Cellular to Social Scales*, pages 11–44. Springer, 2008.

14 A. Jamieson-Lane and C. Hauert. Fixation probabilities on superstars, revisited and revised. *Journal of Theoretical Biology*, 382:44–56, 2015.

15 David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proc. 9th ACM International Conference on Knowledge Discovery and Data Mining*, pages 137–146. ACM, 2003.

16 E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005. Supplementary material available at `http://www.nature.com/nature/journal/v433/n7023/full/nature03204.html`.

17 Thomas M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, 1999.

18 George B. Mertzios, Sotiris E. Nikoletseas, Christoforos Raptopoulos, and Paul G. Spirakis. Natural models for evolution on networks. *Theor. Comput. Sci.*, 477:76–95, 2013.

19 George B. Mertzios and Paul G. Spirakis. Strong bounds for evolution in networks. In *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part II*, pages 669–680, 2013.

20 P. A. P. Moran. Random processes in genetics. *Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

21 Devavrat Shah. Gossip algorithms. *Found. Trends Netw.*, 3(1):1–125, January 2009.

22 P. Shakarian, P. Roos, and A. Johnson. A review of evolutionary graph theory with applications to game theory. *BioSystems*, 107(2):66–80, 2012.