Strong Amplifiers of Natural Selection: Proofs

Andreas Pavlogiannis[1], Josef Tkadlec[1], Krishnendu Chatterjee[1], Martin A. Nowak[2]

[1]IST Austria, A-3400 Klosterneuburg, Austria
[2]Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, Cambridge, MA 02138, USA

Abstract

We consider the modified Moran process on graphs to study the spread of genetic and cultural mutations on structured populations. An initial mutant arises either spontaneously (aka *uniform initialization*), or during reproduction (aka *temperature initialization*) in a population of n individuals, and has a fixed fitness advantage $r > 1$ over the residents of the population. The fixation probability is the probability that the mutant takes over the entire population. Graphs that ensure fixation probability of 1 in the limit of infinite populations are called *strong amplifiers*. Previously, only a few examples of strong amplifiers were known for uniform initialization, whereas no strong amplifiers were known for temperature initialization.

In this work, we study necessary and sufficient conditions for strong amplification, and prove negative and positive results. We show that for temperature initialization, graphs that are unweighted and/or self-loop-free have fixation probability upper-bounded by $1 - 1/f(r)$, where $f(r)$ is a function linear in r. Similarly, we show that for uniform initialization, bounded-degree graphs that are unweighted and/or self-loop-free have fixation probability upper-bounded by $1 - 1/g(r, c)$, where c is the degree bound and $g(r, c)$ a function linear in r. Our main positive result complements these negative results, and is as follows: every family of undirected graphs with (i) self loops and (ii) diameter bounded by $n^{1-\epsilon}$, for some fixed $\epsilon > 0$, can be assigned weights that makes it a strong amplifier, both for uniform and temperature initialization.

Contents

1 Introduction

The Moran process. Evolutionary dynamics study the change of population over time under the effect of natural selection and random drift [28]. The Moran process [27] is an elegant stochastic model for the rigorous study of how mutations spread in a population. Initially, a population of n individuals, called the residents, exists in a homogeneous state, and a random individual becomes mutant. The mutants are associated with a fitness advantage $r \geq 1$, whereas the residents have fitness normalized to 1. The Moran process is a discrete-time stochastic process, described as follows. In every step, a single individual is chosen for reproduction with probability proportional to its fitness. This individual produces a single offspring (a copy of itself), which replaces another individual chosen uniformly at random from the population. The main quantity of interest is the *fixation probability* $\rho(n, r)$, defined as the probability that the single invading mutant will eventually take over the population. As typically r is small (i.e., $r = 1+\epsilon$, for some small $\epsilon > 0$) and n is large, we study the fixation probability at the limit of large populations, i.e., $\rho(r) = \lim_{n\to\infty} \rho(n, r)$. It is known that $\rho(r) = 1 - r^{-1}$.

The Moran process on graphs. The standard Moran process takes place on *well-mixed* populations where the reproducing individual can replace any other in the population. However, natural populations have spatial structure, where each individual has a specific set of neighbors, and mutation spread must respect this structure. Evolutionary graph theory represents spatial structure as a (generally weighted, directed) graph, where each individual occupies a vertex of the graph, and edges define interactions between neighbors [22]. The Moran process on graphs is similar to the standard Moran process, with the exception that the offspring replaces a neighbor of the reproducing individual. The well-mixed population is represented by the complete graph K_n. If the graph is strongly connected, the Moran process is guaranteed to reach a homogeneous state where mutants either fixate or go extinct.

Mutant initialization. The asymmetry introduced by the population structure makes the fixation probability depend on the placement of the initial mutant. In *uniform initialization*, the initial mutant arises *spontaneously*, i.e., uniformly at random on each vertex. In *temperature initialization*, the initial mutant arises*during reproduction* i.e., on each vertex with probability proportional to the rate that the vertex is replaced by offspring from its neighbors. Hence our interested is on the fixation probability $\rho(G_n^{\mathsf{w}}, r, Z)$ for a weighted graph G_n^{w} of n vertices and under initialization $Z \in \{\mathsf{U}, \mathsf{T}\}$, denoting uniform and temperature initialization, respectively.

Amplifiers of selection. Population structure affects the fixation probability of mutants. An infinite family of graphs $(G_n^{\mathsf{w}})_n$ is *amplifying* for initialization Z if $\lim_{n\to\infty} \rho(G_n^{\mathsf{w}}, r, Z) > 1 - r^{-1}$, Intuitively, the fitness advantage of mutants is being "amplified" by the structure compared to the well-mixed population. *Strong amplifying* families have $\lim_{n\to\infty} \rho(G_n^{\mathsf{w}}, r, Z) = 1$, and hence ensure the fixation of mutants. On the other hand, *bounded amplifiers* have $\lim_{n\to\infty} \rho(G_n^{\mathsf{w}}, r, Z) \leq 1 - 1/f(r)$, where f is a linear function, and hence provide limited amplification at best.

Existing results. The Moran process on graphs was introduced in [22], where several amplifying and strongly amplifying families were presented. Under uniform initialization, the canonical example is the family of undirected Star graphs, with fixation probability $1 - r^{-2}$, making it a *quadratic uniform amplifier* [22, 5, 26]. Among directed graphs, strongly amplifying families are known to exist: (i) Superstars and Metafunnels were already introduced in [22], where their strong amplifying properties were outlined, and (ii) more recently, the family of Megastars was rigorously proved to be a strong amplifying family [13]. Megastars were subsequently shown to be optimal (up to logarithmic factors) wrt the rate that fixation probability converges to 1 as a function of n [15]. Among undirected graphs, the family of Stars was the best amplifying family know for a long time, and the existence of strong amplifiers was open. Recently, undirected strong amplifiers were presented independently in [15] and [14].

Under temperature initialization, the landscape is more scarce. None of the uniform amplifiers mentioned in the previous paragraph is a temperature amplifier. It turns out that on all those structures the mutants go extinct with high probability when the initial placement is according to temperature. Recently, the Looping Star family was introduced in [1] and was shown to be a quadratic amplifier under both initialization schemes. Crucially, Looping Stars contain self-loops and weights. To our knowledge, no other temperature amplifier has been known.

Our contributions. In this work, we study necessary and sufficient conditions for strong amplifiers, and prove negative and positive results.

1. Our negative results are as follows. For temperature initialization, we show that graphs which are unweighted and/or self-loop-free have fixation probability upper-bounded by $1 - 1/f(r)$, where $f(r)$ is a function linear in r. Hence, without both weights and self-loops, there are only bounded temperature amplifiers. Similarly, we show that for uniform initialization, bounded-degree graphs that are unweighted and/or self-loop-free have fixation probability upper-bounded by $1 - 1/g(r, c)$, where c is the degree bound and $g(r, c)$ a function linear in r. Hence, without both weights and self-loops, bounded-degree graph families are only bounded uniform amplifiers.
2. Our positive result complements these negative results and is as follows. We show that every family of undirected graphs with (i) self loops and (ii) diameter bounded by $n^{1-\epsilon}$, for some fixed $\epsilon > 0$, can be assigned weights that makes the family a strong amplifier, both for uniform and temperature initialization. Moreover, the weight construction requires $O(n)$ time.

Our proof techniques rely on the analysis of Markov chains, the Cauchy-Schwarz inequality, concentration bounds, stochastic domination and coupling arguments. The weight construction in our positive result is straightforward, however proving the amplification properties of the resulting structure is more involved.

1.1 Other Related Work

Strong amplifiers were already introduced in [22], however it was later shown that the fixation probability on Superstars is weaker than originally stated, and hence the heuristic argument for strong amplification cannot be made formal [7]. In [13], it was shown that the fixation probability on Superstars as appeared in [22] is indeed too optimistic, by proving an upper bound on the rate that the probability can tend to 1 as a function of n. A revised analysis of Superstars appeared in [17]. The work of [30] introduced the Metastars as a family of unweighted undirected graphs with better amplification properties than Stars, for specific values of the fitness advantage r. Other aspects of the Moran process on graphs have also been studied in the literature. In [24], the authors studied undirected suppressors of selection, which are graphs that suppress the selective advantage of mutants, as opposed to amplifying it. Recently, a family of strong suppressors was presented [14]. The work of [25] studies selective amplifiers, a notion that characterizes the number of initial vertices that guarantee mutant fixation. Randomly structured populations were shown to have no effect on fixation probability in [2]. Besides the fixation probability, the absorption time of the Moran process is crucial for characterizing the rate of evolution [11] and has been studied on various graphs [9]. Finally, computational aspects of computing the fixation probability on graphs were studied in [8], where the problem was shown to admit a fully polynomial randomized approximation scheme, later improved in [6].

2 Organization

The organization of this document is as follows: Before presenting our proofs we present the detailed description of our model and the results in Section 2. We then present the formal notation (Section 3), the proofs of our negative results (Section 4) and the proofs of our positive results (Section 5).

3 Model and Summary of Results

3.1 Model

The birth-death Moran process. The *Moran process* considers a population of n individuals, which undergoes reproduction and death, and each individual is either a resident or a mutant [27]. The residents and the mutants have constant fitness 1 and r, respectively. The Moran process is a discrete-time stochastic process defined as follows: in the initial step, a single mutant is introduced into a homogeneous resident population. At each step, an individual is chosen randomly for reproduction with probability proportional to its fitness; another individual is chosen uniformly at random for death and is replaced by a new individual of the same type as the reproducing individual. Eventually, this Markovian process ends when all individuals become of one of the two types. The probability of the event that all individuals become mutants is called the *fixation* probability.

The Moran process on graphs. In general, the Moran process takes place on a population structure, which is represented as a graph. The vertices of the graph represent individuals and edges represent interactions between individuals [22, 28]. Formally, let $G_n = (V_n, E_n, W_n)$ be a weighted, directed graph, where $V_n = \{1, 2, \dots, n\}$ is the vertex set , E_n is the Boolean edge matrix, and W_n is a stochastic weight matrix. An edge is a pair of vertices (i, j) which is indicated by $E_n[i, j] = 1$ and denotes that there is an interaction from i to j (whereas we have $E_n[i, j] = 0$ if there is no interaction from i to j). The stochastic weight matrix W_n assigns weights to interactions, i.e., $W_n[i, j]$ is positive iff $E_n[i, j] = 1$, and for all i we have $\sum_j W_n[i, j] = 1$. For a vertex i, we denote by $\mathsf{In}(i) = \{j \mid E_n[j, i] = 1\}$ (resp., $\mathsf{Out}(i) = \{j \mid E_n[i, j] = 1\}$) the set of vertices that have incoming (resp., outgoing) interaction or edge to (resp., from) i. Similarly to the Moran process, at each step an individual is chosen randomly for reproduction with probability proportional to its fitness. An edge originating from the reproducing vertex is selected randomly with probability equal to its weight. The terminal vertex of the chosen edge takes on the type of the vertex at the origin of the edge. In other words, the stochastic matrix W_n is the weight matrix that represents the choice probability of the edges. We only consider graphs which are *connected*, i.e., every pair of vertices is connected by a path. This is a sufficient condition to ensure that in the long run, the Moran process reaches a homogeneous state (i.e., the population consists entirely of individuals of a single type). See Figure 1 for an illustration. The well-mixed population is represented by a complete graph where all edges have equal weight of $1/n$.

Classification of graphs. We consider the following classification of graphs:

1. *Directed vs undirected graphs.* A graph $G_n = (V_n, E_n, W_n)$ is called *undirected* if for all $1 \leq i, j \leq n$ we have $E_n[i, j] = E_n[j, i]$. In other words, there is an edge from i to j iff there is an edge from j to i, which represents symmetric interaction. If a graph is not undirected, then it is called a *directed* graph.

2. *Self-loop free graphs.* A graph $G_n = (V_n, E_n, W_n)$ is called a *self-loop free* graph iff for all $1 \leq i \leq n$ we have $E_n[i, i] = W_n[i, i] = 0$.

3. *Weighted vs unweighted graphs.* A graph $G_n = (V_n, E_n, W_n)$ is called an *unweighted* graph if for all $1 \leq i \leq n$ we have
$$W_n[i, j] = \begin{cases} \frac{1}{|\mathsf{Out}(i)|} & j \in \mathsf{Out}(i); \\ 0 & j \notin \mathsf{Out}(i) \end{cases}$$
In other words, in unweighted graphs for every vertex the edges are choosen uniformly at random. Note that for unweighted graphs the weight matrix is not relevant, and can be specified simply by the graph structure (V_n, E_n). In the sequel, we will represent unweighted graphs as $G_n = (V_n, E_n)$.

4. *Bounded degree graphs.* The degree of a graph $G_n = (V_n, E_n, W_n)$, denoted $\deg(G_n)$, is $\max\{\mathsf{In}(i), \mathsf{Out}(i) \mid 1 \leq i \leq n\}$, i.e., the maximum in-degree or out-degree. For a family of graphs $(G_n)_{n>0}$ we say that the family has bounded degree, if there exists a constant c such that the degree of all graphs in the family is at most c, i.e., for all n we have $\deg(G_n) \leq c$.

Initialization of the mutant. The fixation probability is affected by many different factors [29]. In a well-mixed population, the fixation probability depends on the population size n and the relative fitness advantage r of mutants [23, 28]. For the Moran process on graphs, the fixation probability also depends on the population structure, which breaks the symmetry and homogeneity of the well-mixed population [21, 20, 10, 22, 5, 12, 31, 16]. Finally, for general population structures, the fixation probability typically depends on the initial location of the mutant [3, 4], unlike the well-mixed population where the probability

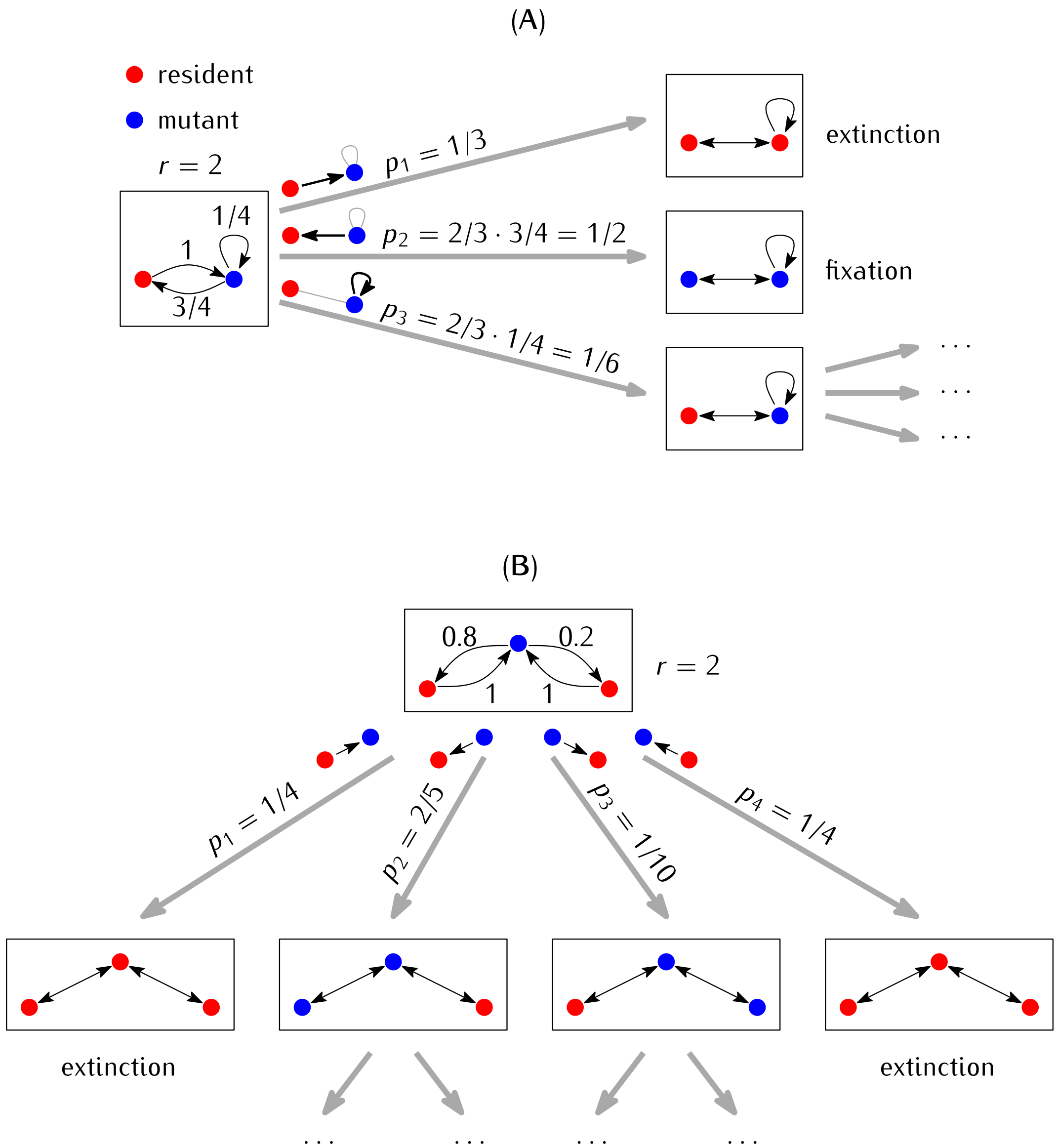

Figure 1: Illustration of one step of the Moran process on a weighted graph with self-loops. Residents are depicted as red vertices, and mutants as blue vertices. As a concrete example, we consider the relative fitness of the mutants is $r = 2$. In Figure 1(A), the total fitness of the population is $\mathcal{F} = 1 + 2 = 3$, and hence the probability of selecting resident (resp., mutant) for reproduction equals $1/3$ (resp., $2/3$). The mutant reproduces along an edge, and the edge is chosen randomly proportional to the edge weight. Figure 1(B) shows that different reproduction events might lead to the same outcome.

of the mutant fixing is independent of where the mutant arises [23, 28]. There are two standard ways mutants may arise in a population [22, 1]. First, mutants may arise spontaneously and with equal probability at any vertex of the population structure. In this case we consider that the mutant arise at any vertex uniformly at random and we call this *uniform initialization*. Second, mutants may be introduced through reproduction, and thus arise at a vertex with rate proportional to the incoming edge weights of the vertex. We call this *temperature initialization*. In general, uniform and temperature initialization result in different fixation probabilities.

Amplifiers, quadratic amplifiers, and strong amplifiers. Depending on the initialization, a population structure can distort fitness differences [22, 28, 5], where the well-mixed population serves as a canonical point of comparison. Intuitively, amplifiers of selection exaggerate variations in fitness by increasing (respectively decreasing) the chance of fitter (respectively weaker) mutants fixing compared to their chance of fixing in the well-mixed population. In a well-mixed population of size n, the fixation probability is

$$\frac{1-1/r}{1-(1/r)^n}.$$

Thus, in the limit of large population (i.e., as $n \to \infty$) the fixation probability in a well-mixed population is $1-1/r$. We focus on two particular classes of amplifiers that are of special interest. A family of graphs $(G_n)_{n>0}$ is a *quadratic* amplifier if in the limit of large population the fixation probability is $1-1/r^2$. Thus, a mutant with a 10% fitness advantage over the resident has approximately the same chance of fixing in quadratic amplifiers as a mutant with a 21% fitness advantage in the well-mixed population. A family of graphs $(G_n)_{n>0}$ is an *arbitrarily strong* amplifier (hereinafter called simply a strong amplifier) if for any constant $r>1$ the fixation probability approaches 1 at the limit of large population sizes, whereas when $r<1$, the fixation probability approaches 0. There is a much finer classification of amplifiers presented in [1]. We focus on quadratic amplifiers which are the most well-known among polynomial amplifiers, and strong amplifiers which represent the strongest form of amplification.

Amplifiers tend to have fixation times longer than the well mixed population. Therefore they are especially useful in situations where the rate limiting step is the discovery and evaluation of marginally advantageous mutants. An interesting direction for future work would be to consider amplifiers as well as the time-scale of evolutionary trajectories.

Existing results. We summarize the main existing results in terms of uniform and temperature initialization.

1. *Uniform initialization.* First, consider the family of Star graphs, which consist of one central vertex and $n-1$ leaf vertices, with each leaf being connected to and from the central vertex. Star graphs are unweighted, undirected, self-loop free graphs, whose degree is linear in the population size. Under uniform initialization, the family of Star graphs is a quadratic amplifier [22, 28]. A generalization of Star graphs, called Superstars [22, 28, 17, 8], are known to be strong amplifiers under uniform initialization [13]. The Superstar family consists of unweighted, self-loop free, but directed graphs where the degree is linear in the population size. Another family of directed graphs with strong amplification properties, called Megastars, was recently introduced in [13]. The Megastars are stronger amplifiers than the Superstars, as the fixation probability on the former is a approximately $1-n^{-1/2}$ (ignoring logarithmic factors), and is asymptotically optimal (again, ignoring logarithmic factors). In contrast, the fixation probability on the Superstars is approximately $1-n^{-1/2}$. In the limit of $n \to \infty$, both families approach the fixation probability 1.

2. *Temperature initialization.* While the family of Star graphs is a quadratic amplifier under uniform initialization, it is not even an amplifier under temperature initialization [1]. It was shown in [1] that by adding self-loops and weights to the edges of the Star graph, a graph family, namely the family of Looping Stars, can be constructed, which is a quadratic amplifier simultaneously under temperature and uniform initialization. Note that in contrast to Star graphs, the Looping Star graphs are weighted and also have self-loops.

Open questions. Despite several important existing results on amplifiers of selection, several basic questions have remained open:

1. *Question 1.* Does there exist a family of self-loop free graphs (weighted or unweighted) that is a quadratic amplifier under temperature initialization?
2. *Question 2.* Does there exist a family of unweighted graphs (with or without self-loops) that is a quadratic amplifier under temperature initialization?
3. *Question 3.* Does there exist a family of bounded degree self-loop free (weighted or unweighted) graphs that is a strong amplifier under uniform initialization?
4. *Question 4.* Does there exist a family of bounded degree unweighted graphs (with or without self-loops) that is a strong amplifier under uniform initialization?
5. *Question 5.* Does there exist a family of graphs that is a strong amplifier under temperature initialization? More generally, does there exist a family of graphs that is a strong amplifier both under temperature and uniform initialization?

To summarize, the open questions ask for (i) the existence of quadratic amplifiers under temperature initialization without the use of self-loops, or weights (Questions 1 and 2); (ii) the existence of strong amplifiers under uniform initialization without the use of self-loops, or weights, and while the degree of the graph is small; and (iii) the existence of strong amplifiers under temperature initialization. While the answers to Question 1 and Question 2 are positive under uniform initialization, they have remained open under temperature initialization. Questions 3 and 4 are similar to 1 and 2, but focus on uniform rather than temperature initialization. The restriction on graphs of bounded degree is natural: large degree means that some individuals must have a lot of interactions, whereas graphs of bounded degree represent simple structures. Question 5 was mentioned as an open problem in [1]. Note that under temperature initialization, even the existence of a cubic amplifier, that achieves fixation probability at least $1 - (1/r^3)$ in the limit of large population, has been open [1].

3.2 Results

In this work we present several negative as well as positive results that answer the open questions (Questions 1-5) mentioned above. We first present our negative results.

Negative results. Our main negative results are as follows:

1. Our first result (Theorem 1) shows that for any self-loop free weighted graph $G_n = (V_n, E_n, W_n)$, for any $r \geq 1$, under temperature initialization the fixation probability is at most $1 - 1/(r + 1)$. The implication of the above result is that it answers Question 1 in negative.

2. Our second result (Theorem 2) shows that for any unweighted (with or without self-loops) graph $G_n = (V_n, E_n)$, for any $r \geq 1$, under temperature initialization the fixation probability is at most $1 - 1/(4r+2)$. The implication of the above result is that it answers Question 2 in negative.

3. Our third result (Theorem 3) shows that for any bounded degree self-loop free graph (possibly weighted) $G_n = (V_n, E_n, W_n)$, for any $r \geq 1$, under uniform initialization the fixation probability is at most $1 - 1/(c + c^2 r)$, where c is the bound on the degree, i.e., $\deg(G_n) \leq c$. The implication of the above result is that it answers Question 3 in negative.

4. Our fourth result (Theorem 4) shows that for any unweighted, bounded degree graph (with or without self-loops) $G_n = (V_n, E_n)$, for any $r \geq 1$, under uniform initialization the fixation probability is at most $1 - 1/(1 + rc)$, where c is the bound on the degree, i.e., $\deg(G_n) \leq c$. The implication of the above result is that it answers Question 4 in negative.

Significance of the negative results. We now discuss the significance of the above results.

1. The first two negative results show that in order to obtain quadratic amplifiers under temperature initialization, self-loops and weights are inevitable, complementing the existing results of [1]. More importantly, it shows a sharp contrast between temperature and uniform initialization: while self-loop free, unweighted graphs (namely, Star graphs) are quadratic amplifiers under uniform initialization, no such graph families are quadratic amplifiers under temperature initialization.

2. The third and fourth results show that without using self-loops and weights, bounded degree graphs cannot be made strong amplifiers even under uniform initialization. See also Remark 2.

Positive result. Our main positive result shows the following:

1. For any constant $\epsilon > 0$, consider any connected unweighted graph $G_n = (V_n, E_n)$ of n vertices with self-loops and which has *diameter* at most $n^{1-\epsilon}$. The diameter of a connected graph is the maximum, among all pairs of vertices, of the length of the shortest path between that pair. We establish (Theorem 5) that there is a stochastic weight matrix W_n such that for any $r > 1$ the fixation probability on $G_n = (V_n, E_n, W_n)$ both under uniform and temperature initialization is at least $1 - \frac{1}{n^{\epsilon/3}}$. An immediate consequence of our result is the following: for any family of connected unweighted graphs with self-loops $(G_n = (V_n, E_n))_{n>0}$ such that the diameter of G_n is at most $n^{1-\epsilon}$, for a constant $\epsilon > 0$, one can construct a stochastic weight matrix W_n such that the resulting family $(G_n = (V_n, E_n, W_n))_{n>0}$ of weighted graphs is a strong amplifier simultaneously under uniform and temperature initialization. Thus we answer Question 5 in affirmative.

Significance of the positive result. We highlight some important aspects of the results established in this work.

1. First, note that for the fixation probability of the Moran process on graphs to be well defined, a necessary and sufficient condition is that the graph is connected. A uniformly chosen random connected unweighted graph of n vertices has diameter bounded by a constant, with high probability. Hence, within the family of connected, unweighted graphs, the family of graphs of diameter at most $O(n^{1-\epsilon})$, for any constant $0 < \epsilon < 1$, has probability measure 1. Our results establish a strong dichotomy: (a) the negative results state that without self-loops and/or without weights, *no* family of graphs can be a quadratic amplifier (even more so a strong amplifier) even for only temperature initialization; and (b) in contrast, for *almost all* families of connected graphs with self-loops, there exist weight func-

	Temperature		Uniform*	
	Loops	No Loops	Loops	No Loops
Weights	✓	$\times$	✓	$\times$
No Weights	$\times$	$\times$	$\times$	$\times$

Table 1: Summary of our results on existence of strong amplifiers for different initialization schemes (temperature initialization or uniform initialization) and graph families (presence or absence of loops and/or weights). The "✓" symbol marks that for given choice of initialization scheme and graph family, almost all graphs admit a weight function that makes them strong amplifiers. The "$\times$" symbol marks that for given choice of initialization scheme and graph family, no strong amplifiers exist (under any weight function). The asterisk signifies that the negative results under uniform initialization only hold for bounded degree graphs.

tions such that the resulting family of weighted graphs is a strong amplifier both under temperature and uniform initialization.

2. Second, with the use of self-loops and weights, even simple graph structures, such as Star graphs, Grids, and well-mixed structures (i.e., complete graphs) can be made strong amplifiers.
3. Third, our positive result is constructive, rather than existential. In other words, we not only show the existence of strong amplifiers, but present a construction of them.

Our results are summarized in Table 1.

Remark 1. *Edges with zero weight.* Note that edges can be effectively removed by being assigning zero weight (however, no weight assignment can create edges that don't exist.) Therefore, when our construction works for some graph, it also works for a graph that contains some additional edges. In particular, our construction easily works for complete graphs. The construction can also be extended to a scenario in which we insist that each edge is assigned a positive (non-zero) weight.

4 Preliminaries: Formal Notation

4.1 The Moran Process on Weighted Structured Populations

We consider a population of n individuals on a graph $G_n = (V_n, E_n, W_n)$. Each individual of the population is either a *resident*, or a *mutant*. Mutants are associated with a *reproductive rate* (or *fitness*) r, whereas the reproductive rate of residents is normalized to 1. Typically we consider the case where $r > 1$, i.e., mutants are *advantageous*, whereas when $r < 1$ we call the mutants *disadvantageous*. We now introduce the formal notation related to the process.

Configuration. A *configuration* of G_n is a subset $S \subseteq V$ which specifies the vertices of G_n that are occupied by mutants and thus the remaining vertices $V \setminus S$ are occupied by residents. We denote by $\mathsf{F}(S) = r \cdot |S| + n - |S|$ the total fitness of the population in configuration S, where $|S|$ is the number of mutants in S.

The Moran process. The birth-detah Moran process on G_n is a discrete-time Markovian random process. We denote by X_i the random variable for a configuration at time step i, and $\mathsf{F}(\mathsf{X}_i)$ and $|\mathsf{X}_i|$ denote the total fitness and the number of mutants of the corresponding configuration, respectively. The probability distribution for the next configuration X_{i+1} at time $i + 1$ is determined by the following two events in succession:

Birth: One individual is chosen at random to reproduce, with probability proportional to its fitness. That is, the probability to reproduce is $r/\mathsf{F}(\mathsf{X}_i)$ for a mutant, and $1/\mathsf{F}(\mathsf{X}_i)$ for a resident. Let u be the vertex occupied by the reproducing individual.

Death: A neighboring vertex $v \in \mathsf{Out}(u)$ is chosen randomly with probability $W_n[u,v]$. The individual occupying v dies, and the reproducing individual places a copy of its own on v. Hence, if $u \in \mathsf{X}_i$, then $\mathsf{X}_{i+1} = \mathsf{X}_i \cup \{v\}$, otherwise $\mathsf{X}_{i+1} = \mathsf{X}_i \setminus \{v\}$.

The above process is known as the *birth-death* Moran process, where the death event is conditioned on the birth event, and the dying individual is a neighbor of the reproducing one.

Probability measure. Given a graph G_n and the fitness r, the birth-death Moran process defines a probability measure on sequences of configurations, which we denote as $\mathbb{P}^{G_n,r}[\cdot]$. If the initial configuration is $\{u\}$, then we define the probability measure as $\mathbb{P}_u^{G_n,r}[\cdot]$, and if the graph and fitness r is clear from the context, then we drop the superscript.

Fixation event. The fixation event, denoted $\mathcal{E}$, represents that all vertices are mutants, i.e., $\mathsf{X}_i = V$ for some i. In particular, $\mathbb{P}_u^{G_n,r}[\mathcal{E}]$ denotes the fixation probability in G_n for fitness r of the mutant, when the initial mutant is placed on vertex u. We will denote this fixation probability as $\rho(G_n, r, u) = \mathbb{P}_u^{G_n,r}[\mathcal{E}]$.

4.2 Initialization and Fixation Probabilities

We will consider three types of initialization, namely, (a) uniform initialization, where the mutant arises at vertices with uniform probability, (b) temperature initialization, where the mutant arises at vertices proportional to the temperature, and (c) convex combination of the above two.

Temperature. For a weighted graph $G_n = (V_n, E_n, W_n)$, the temperature of a vertex u, denoted $\mathsf{T}(u)$, is $\sum_{v \in \mathsf{In}(u)} W_n[v,u]$, i.e., the sum of the incoming weights. Note that $\sum_{u \in V_n} \mathsf{T}(u) = n$, and a graph is *isothermal* iff $\mathsf{T}(u) = 1$ for all vertices u.

Fixation probabilities. We now define the fixation probabilities under different initialization.

1. *Uniform initialization.* The fixation probability under uniform initialization is
$$\rho(G_n, r, \mathsf{U}) = \sum_{u \in V_n} \frac{1}{n} \cdot \rho(G_n, r, u).$$
2. *Temperature initialization.* The fixation probability under temperature initialization is
$$\rho(G_n, r, \mathsf{T}) = \sum_{u \in V_n} \frac{\mathsf{T}(u)}{n} \cdot \rho(G_n, r, u).$$
3. *Convex initialization.* In *η-convex initialization*, where $\eta \in [0,1]$, the initial mutant arises with probability $(1-\eta)$ via uniform initialization, and with probability η via temperature initialization. The fixation probability is then
$$\rho(G_n, r, \eta) = (1-\eta) \cdot \rho(G_n, r, \mathsf{U}) + \eta \cdot \rho(G_n, r, \mathsf{T}).$$

4.3 Strong Amplifier Graph Families

A *family* of graphs $\mathcal{G}$ is an infinite sequence of weighted graphs $\mathcal{G} = (G_n)_{n \in \mathbb{N}^+}$.

- *Strong amplifiers.* A family of graphs $\mathcal{G}$ is a *strong uniform amplifier* (resp. *strong temperature amplifier*, *strong convex amplifier*) if for every fixed $r_1 > 1$ and $r_2 < 1$ we have that

$$\liminf_{n\to\infty} \rho(G_n, r_1, Z) = 1 \qquad \text{and} \qquad \limsup_{n\to\infty} \rho(G_n, r_2, Z) = 0 \ ;$$

where $Z = \mathsf{U}$ (resp., $Z = \mathsf{T}$, $Z = \eta$).

Intuitively, strong amplifiers ensures (a) fixation of advantageous mutants with probability 1 and (b) extinction of disadvantageous mutants with probability 1. In other words, strong amplifiers represent the strongest form of amplifiers possible.

5 Negative Results

In the current section we present our negative results, which show the nonexistence of strong amplifiers in the absence of either self-loops or weights. In our proofs, we consider weighted graph $G_n = (V_n, E_n, W_n)$, and for notational simplicity we drop the subscripts from vertices, edges and weights, i.e., we write $G_n = (V, E, W)$. We also consider that G_n is connected and $n \geq 2$. Throughout this section we will use a technical lemma, which we present below. Given a configuration $\mathsf{X}_i = \{u\}$ with one mutant, let x and y be the probability that in the next configuration the mutants increase and go extinct, respectively. The following lemma bounds the fixation probability $\rho(G_n, r, u)$ as a function of x and y.

Lemma 1. *Consider a vertex u and the initial configuration $\mathsf{X}_0 = \{u\}$ where the initial mutant arises at vertex u. For any configuration $\mathsf{X}_i = \{u\}$, let*

$$x = \mathbb{P}^{G_n,r}[|\mathsf{X}_{i+1}| = 2 \mid \mathsf{X}_i = \{u\}] \qquad \textit{and} \quad y = \mathbb{P}^{G_n,r}[|\mathsf{X}_{i+1}| = 0 \mid \mathsf{X}_i = \{u\}] \ .$$

be the probability that the number of mutants increases (or decreases) in a single step. Then the fixation probability from u is at most $x/(x+y)$, i.e.,

$$\rho(G_n, r, u) \leq \frac{x}{x+y} = 1 - \frac{y}{x+y} \ .$$

Proof. We upperbound the fixation probability $\rho(G_n, r, u)$ starting from u by the probability that a configuration X_t is reached with $|\mathsf{X}_t| = 2$. Note that to reach fixation the Moran process must first reach a configuration with at least two mutants. We now analyze the probability to reach at least two mutants. This is represented by a three-state one dimensional random walk, where two states are absorbing, one absorbing state represents a configuration with two mutants, and the other absorbing state represents the extinction of the mutants, and the bias towards the absorbing state representing two mutants is x/y. See Figure 2 for an illustration. Using the formulas for absorption probability in one-dimensional three-state Markov chains (see, e.g., [18], [28, Section 6.3]), we have the probability that a configuration with two mutants is reached is

$$\frac{1-(x/y)^{-1}}{1-(x/y)^{-2}} = \frac{1}{1+(x/y)^{-1}} = \frac{x}{x+y} \ .$$

Hence it follows that $\rho(G_n, r, u) \leq 1 - \frac{y}{x+y}$. $\square$

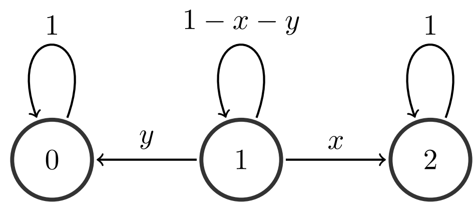

Figure 2: Illustration of the Markov chain of Lemma 1.

5.1 Negative Result 1

We now prove our negative result 1.

Theorem 1. *For all self-loop free graphs G_n and for every $r \geq 1$ we have $\rho(G_n, r, \mathsf{T}) \leq 1 - 1/(r+1)$.*

Proof. Since G_n is self-loop free, for all u we have $W[u,u] = 0$. Hence $\mathsf{T}(u) = \sum_{v \in \mathsf{In}(u)\setminus\{u\}} W[v,u]$. Consider the case where the initial mutant is placed on vertex u, i.e, $\mathsf{X}_0 = \{u\}$. For any configuration $\mathsf{X}_i = \{u\}$, we have the following:

$$x = \mathbb{P}^{G_n,r}[|\mathsf{X}_{i+1}| = 2 \mid \mathsf{X}_i = \{u\}] = \frac{r}{\mathsf{F}(\mathsf{X}_i)}$$

$$y = \mathbb{P}^{G_n,r}[|\mathsf{X}_{i+1}| = 0 \mid \mathsf{X}_i = \{u\}] = \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \sum_{v \in \mathsf{In}(u)\setminus\{u\}} W[v,u] = \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \mathsf{T}(u)\ .$$

Thus $x/y = r/\mathsf{T}(u)$. Hence by Lemma 1 we have

$$\rho(G_n, r, u) \leq 1 - \frac{\mathsf{T}(u)}{\mathsf{T}(u) + r}\ .$$

Summing over all u, we obtain

$$\rho(G_n, r, \mathsf{T}) = \sum_u \frac{\mathsf{T}(u)}{n} \cdot \rho(G_n, r, u) \leq \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \left(1 - \frac{\mathsf{T}(u)}{\mathsf{T}(u)+r}\right) = 1 - \frac{1}{n} \cdot \sum_u \frac{\mathsf{T}(u)^2}{\mathsf{T}(u)+r}\ ; \quad (1)$$

since $\sum_u \mathsf{T}(u) = n$. Using the Cauchy-Schwarz inequality, we obtain

$$\sum_u \frac{\mathsf{T}(u)^2}{\mathsf{T}(u)+r} \geq \frac{\left(\sum_u \mathsf{T}(u)\right)^2}{\sum_u (\mathsf{T}(u)+r)} = \frac{n^2}{n + n \cdot r} = \frac{n}{r+1}\ ;$$

and thus Eq. (1) becomes

$$\rho(G_n, r, \mathsf{T}) \leq 1 - \frac{1}{n} \cdot \frac{n}{r+1} = 1 - \frac{1}{r+1}$$

as desired. □

We thus arrive at the following corollary.

Corollary 1. *There exists no self-loop free family of graphs which is a strong temperature amplifier.*

5.2 Negative Result 2

We now prove our negative result 2.

Theorem 2. *For all unweighted graphs G_n and for every $r \geq 1$ we have $\rho(G_n, r, \mathsf{T}) \leq 1 - 1/(4r+2)$.*

Proof. For every vertex $u \in V$, let

$$\mathsf{T}'(u) = \sum_{v \in \mathsf{In}(u) \setminus \{u\}} \frac{1}{|\mathsf{Out}(v)|} \ .$$

We establish two inequalities related to T'. Since G_n is unweighted, we have

$$\mathsf{T}(u) = \sum_{v \in \mathsf{In}(u)} \frac{1}{|\mathsf{Out}(v)|} \geq \mathsf{T}'(u) \ .$$

For a vertex u, let $\mathsf{sl}(u) = 1$ if u has a self-loop and $\mathsf{sl}(u) = 0$ otherwise. Since G_n is connected, each vertex u has at least one neighbor other than itself. Thus for every vertex u with $\mathsf{sl}(u) = 1$ we have that $|\mathsf{Out}(u)| \geq 2$. Hence

$$\sum_u \mathsf{T}'(u) = \sum_u \left(\sum_{v \in \mathsf{In}(u)} \frac{1}{|\mathsf{Out}(v)|} - \mathsf{sl}(u) \frac{1}{|\mathsf{Out}(u)|} \right) = \sum_u \left(\sum_{v \in \mathsf{In}(u)} \frac{1}{|\mathsf{Out}(v)|} \right) - \sum_{u : \mathsf{sl}(u) = 1} \left(\frac{1}{|\mathsf{Out}(u)|} \right)$$

$$\geq \sum_u \mathsf{T}(u) - \sum_u \frac{1}{2} = n - \frac{n}{2} = \frac{n}{2}. \quad (2)$$

Similarly to the proof of Theorem 1, the fixation probability given that a mutant is initially placed on vertex u is at most

$$\rho(G_n, r, u) \leq 1 - \frac{\mathsf{T}'(u)}{\mathsf{T}'(u) + r}$$

Summing over all u, we obtain

$$\rho(G_n, r, \mathsf{T}) = \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \rho(G_n, r, u) \leq \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \left(1 - \frac{\mathsf{T}'(u)}{\mathsf{T}'(u) + r}\right) \leq 1 - \frac{1}{n} \cdot \sum_u \frac{\mathsf{T}'(u)^2}{\mathsf{T}'(u) + r} \ ; \quad (3)$$

since $\sum_u \mathsf{T}(u) = n$ and $\mathsf{T}(u) \geq \mathsf{T}'(u)$.

Using the Cauchy-Schwarz inequality we get

$$\sum_u \frac{\mathsf{T}'(u)^2}{\mathsf{T}'(u) + r} \geq \frac{(\sum_u \mathsf{T}'(u))^2}{\sum_u (\mathsf{T}'(u) + r)} = \frac{x^2}{x + n \cdot r},$$

where $x = \sum_u \mathsf{T}'(u)$. Note that the function $f(x) = \frac{x^2}{x + n \cdot r}$ is increasing in x for $x > 0$ and any $r, n > 0$. Since $x > n/2$, the right-hand side is minimized for $x = n/2$, that is

$$\sum_u \frac{\mathsf{T}'(u)^2}{\mathsf{T}'(u) + r} \geq \frac{(n/2)^2}{n/2 + n \cdot r} = \frac{n}{4r + 2}.$$

Thus Eq. (3) becomes

$$\rho(G_n, r, \mathsf{T}) \leq 1 - \frac{1}{n} \cdot \frac{n}{4r+2} = 1 - \frac{1}{4r+2}$$

as desired.

□

We thus arrive at the following corollary.

Corollary 2. *There exists no unweighted family of graphs which is a strong temperature amplifier.*

5.3 Negative Result 3

We now prove our negative result 3.

Theorem 3. *For all self-loop free graphs G_n with $c = \deg(G_n)$, and for every $r \geq 1$ we have $\rho(G_n, r, \mathsf{U}) \leq 1 - 1/(c + r \cdot c^2)$.*

Proof. Let $G_n = (V, E, W)$ and $\gamma = 1/c$. For a vertex u, denote by $\mathsf{Out}^{\gamma}(u) = \{v \in \mathsf{Out}(u) \ : \ W[u, v] \geq \gamma\}$. Observe that since $\deg(G_n) = c$, every vertex u has an outgoing edge of weight at least $1/c$, and thus $\mathsf{Out}^{\gamma}(u) \neq \emptyset$ for all $u \in V$. Let $V^h = \bigcup_u \mathsf{Out}^{\gamma}(u)$. Intuitively, the set V^h contains "hot" vertices, since each vertex $u \in V^h$ is replaced frequently (with rate at least γ) by at least one neighbor v.

Bound on size of V^h. We first obtain a bound on the size of V^h. Consider a vertex $u \in V$ and a vertex $v \in \mathsf{Out}^{\gamma}(u)$ (i.e., $v \in V^h$). For every vertex $w \in \mathsf{In}(v)$ such that $v \in \mathsf{Out}^{\gamma}(w)$ we can count $v \in V^h$ and to avoide multiple counting, we consider for each count of v a contribution of $\frac{1}{|\{w \in \mathsf{In}(v) \colon \ v \in \mathsf{Out}^{\gamma}(w)\}|}$, which is at least $\frac{1}{c}$ due to the degree bound. Hence we have

$$|V^h| = \sum_{u \in V} \sum_{v \in \mathsf{Out}^{\gamma}(u)} \frac{1}{|\{w \in \mathsf{In}(v) \ : \ v \in \mathsf{Out}^{\gamma}(w)\}|} \geq \sum_{u \in V} \sum_{v \in \mathsf{Out}^{\gamma}(u)} \frac{1}{c} \geq \sum_{u \in V} \frac{1}{c} = \frac{n}{c} \ ;$$

where the last inequality follows from the fact that $\mathsf{Out}^{\gamma}(u) \neq \emptyset$ for all $u \in V$. Hence the probability that the initial mutant is a vertex in V^h has probability at least $1/c$ according to the uniform initialization.

Bound on probability. Consider that the initial mutant is a vertex $u \in V^h$. Consider any configuration $\mathsf{X}_i = \{u\}$, we have the following:

$$x = \mathbb{P}^{G_n, r}[|\mathsf{X}_{i+1}| = 2 \mid \mathsf{X}_i = \{u\}] = \frac{r}{\mathsf{F}(\mathsf{X}_i)}$$

$$y = \mathbb{P}^{G_n, r}[|\mathsf{X}_{i+1}| = 0 \mid \mathsf{X}_i = \{u\}] = \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \sum_{(v,u) \in E} W[v, u] \geq \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \sum_{v : u \in \mathsf{Out}^{\gamma}(v)} \gamma \geq \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \gamma \ .$$

Thus $x/y \leq r/\gamma$. Hence by Lemma 1 we have

$$\rho(G_n, r, u) \leq \frac{r \cdot c}{1 + r \cdot c} \ .$$

Finally, we have

$$\rho(G_n, r, \mathsf{U}) = \sum_{u \in V^h} \frac{1}{n} \cdot \rho(G_n, r, u) + \sum_{u \in V \setminus V^h} \frac{1}{n} \cdot \rho(G_n, r, u)$$

$$\leq \frac{1}{c} \cdot \frac{r \cdot c}{1 + r \cdot c} + \frac{c-1}{c} \cdot 1 = 1 - \frac{1}{c} \cdot \left(1 - \frac{r \cdot c}{1 + r \cdot c}\right) = 1 - \frac{1}{c + r \cdot c^2} \; .$$

The desired result follows. □

We thus arrive at the following corollary.

Corollary 3. *There exists no self-loop free, bounded-degree family of graphs which is a strong uniform amplifier.*

5.4 Negative Result 4

We now prove our negative result 4.

Theorem 4. *For all unweighted graphs G_n with $c = \deg(G_n)$, and for every $r \geq 1$ we have $\rho(G_n, r, \mathsf{U}) \leq 1 - 1/(1 + r \cdot c)$.*

Proof. Let $G_n = (V, E, W)$ and consider that $\mathsf{X}_0 = u$ for some $u \in V$. Consider any configuration $\mathsf{X}_i = \{u\}$, we have the following:

$$x = \mathbb{P}^{G_n, r}[|\mathsf{X}_{i+1}| = 2 \mid \mathsf{X}_i = \{u\}] \leq \frac{r}{\mathsf{F}(\mathsf{X}_i)} \; .$$

$$y = \mathbb{P}^{G_n, r}[|\mathsf{X}_{i+1}| = 0 \mid \mathsf{X}_i = \{u\}] = \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \sum_{v \in \mathsf{In}(u) \setminus \{u\}} W[v, u] \geq \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \frac{1}{c} \; .$$

Thus $x/y \leq r \cdot c$. By Lemma 1 we have

$$\rho(G_n, r, u) \leq \frac{r \cdot c}{1 + r \cdot c}.$$

Finally, we have

$$\rho(G_n, r, \mathsf{U}) = \frac{1}{n} \cdot \sum_{u} \rho(G_n, r, u) \leq \frac{r \cdot c}{1 + r \cdot c} = 1 - \frac{1}{1 + r \cdot c} \; .$$

The desired result follows. □

We thus arrive at the following corollary.

Corollary 4. *There exists no unweighted, bounded-degree family of graphs which is a strong uniform amplifier.*

Remark 2. Theorems 3 and 4 establish the nonexistence of strong amplification with bounded degree graphs. A relevant result can be found in [24], which establishes an upperbound of the fixation probability of mutants under uniform initialization on unweighted, undirected graphs. If the bounded degree restriction is relaxed to bounded average degree, then recent results show that strong amplifiers (called *sparse incubators*) exist [15].

6 Positive Result

In the previous section we showed that self-loops and weights are necessary for the existence of strong amplifiers. In this section we present our positive result, namely that every family of undirected graphs with self-loops and whose diameter is not "too large" can be made a strong amplifier by using appropriate weight functions. Our result relies on several novel conceptual steps, therefore the proof is structured in three parts.

1. First, we introduce some formal notation that will help with the exposition of the ideas that follow.
2. Second, we describe an algorithm which takes as input an undirected graph $G_n = (V_n, E_n)$ of n vertices, and constructs a weight matrix W_n to obtain the weighted graph $G_n^{\mathsf{w}} = (V_n, E_n, W_n)$.
3. Lastly, we prove that G_n^{w} is a strong amplifier both for uniform and temperature initialization.

Before presenting the details we introduce some notation to be used in this section.

6.1 Undirected Graphs and Notation

We first present some additional notation required for the exposition of the results of this section.

Undirected graphs. Our input is an unweighted undirected graph $G_n = (V_n, E_n)$ with self loops. For ease of notation, we drop the subscript n and refer to the graph $G = (V, E)$ instead. Since G is undirected, for all vertices u we have $\mathsf{In}(u) = \mathsf{Out}(u)$, and we denote by $\mathsf{Nh}(u) = \mathsf{In}(u) = \mathsf{Out}(u)$ the set of neighbors of vertex u. Hence, $v \in \mathsf{Nh}(u)$ iff $u \in \mathsf{Nh}(v)$. Moreover, since G has self-loops, we have $u \in \mathsf{Nh}(u)$. Also we consider that G is connected, i.e., for every pair of vertices u, v, there is a path from u to v.

Symmetric weight function. So far we have used a stochastic weight matrix W, where for every u we have $\sum_v W[u, v] = 1$. In this section, we will consider a weight function $\mathsf{w} : E \to \mathbb{R}_{\geq 0}$, and given a vertex $u \in V$ we denote by $\mathsf{w}(u) = \sum_{v \in \mathsf{Nh}(u)} \mathsf{w}(u, v)$. Our construction will not only assign weights, but also ensure symmetry. In other words, we we construct *symmetric* weights such that for all u, v we have $\mathsf{w}(u, v) = \mathsf{w}(v, u)$. Given such a weight function w, the corresponding stochastic weight matrix W is defined as $W[u, v] = \mathsf{w}(u, v)/\mathsf{w}(u)$ for all pairs of vertices u, v. Given a unweighted graph G and weight function w, we denote by G^{w} the corresponding weighted graph.

Vertex-induced subgraphs. Given a set of vertices $X \subseteq V$, we denote by $G^{\mathsf{w}}[X] = (X, E[X], \mathsf{w}[X])$ the subgraph of G induced by X, where $E[X] = E \cap (X \times X)$, and the weight function $\mathsf{w}[X] : E[X] \to \mathbb{R}_{\geq 0}$ defined as

$$\mathsf{w}[X](u, v) = \begin{cases} \mathsf{w}(u, u) + \sum_{(u,w) \in E \setminus E[X]} \mathsf{w}(u, w) & \text{if } u = v \\ \mathsf{w}(u, v) & \text{otherwise} \end{cases}$$

In words, the weights on the edges of u to vertices that do not belong to X are added to the self-loop weight of u. Since the sum of all weights does not change, we have $\mathsf{w}[X](u) = \mathsf{w}(u)$ for all u. The temperature of u in $G[X]$ is

$$\mathsf{T}[X](u) = \sum_{v \in \mathsf{Nh}(u) \cap X} \frac{\mathsf{w}[X](v, u)}{\mathsf{w}[X](v)} .$$

6.2 Algorithm for Weight Assignment on G

We start with the construction of the weight function w on G. Since we consider arbitrary input graphs, w is constructed by an algorithm. The time complexity of the algorithm is $O(n \cdot \log n)$. Since our focus is on the properties of the resulting weighted graph, we do not explicitly analyze the time complexity.

Steps of the construction. Consider a connected graph G with diameter $\mathrm{diam}(G) \leq n^{1-\varepsilon}$, where $\varepsilon > 0$ is a constant independent of n. We construct a weight function w such that whp an initial mutant arising under uniform or temperature initialization, eventually fixates on G^{w}. The weight assignment consists of the following conceptual steps.

1. *Spanning tree construction and partition.* First, we construct a *spanning tree* $\mathcal{T}_n^x$ of G rooted on some arbitrary vertex x. In words, a spanning tree of an undirected graph is a connected subgraph that is a tree and includes all of the vertices of the graph. Then we partition the tree into a number of component trees of appropriate sizes.
2. *Hub construction.* Second, we construct the *hub* of G, which consists of the vertices x_i that are roots of the component trees, together with all vertices in the paths that connect each x_i to the root x of $\mathcal{T}_n^x$. All vertices that do not belong to the hub belong to the *branches* of G.
3. *Weight assignment.* Finally, we assign weights to the edges of G, such that the following properties hold:
   (a) The hub is an isothermal graph, and evolves exponentially faster than the branches.
   (b) All edges between vertices in different branches are effectively cut-out (by being assigned weight 0).

In the following we describe the above steps formally.

Spanning tree $\mathcal{T}_n^x$ construction and partition. Given the graph G, we first construct a spanning tree using the standard breadth-first-search (BFS) algorithm. Let $\mathcal{T}_n^x$ be such a spanning tree of G, rooted at some arbitrary vertex x. We now construct the partitioning as follows: We choose a constant $c = 2\varepsilon/3$, and pick a set $S \subset V$ such that

1. $|S| \leq n^c$, and
2. the removal of S splits $\mathcal{T}_n^x$ into k trees $T_{n_1}^{x_1}, \ldots, T_{n_k}^{x_k}$, each $T_{n_i}^{x_i}$ rooted at vertex x_i and of size n_i, with the property that $n_i \leq n^{1-c}$ for all $1 \leq i \leq k$.

The set S is constructed by a simple bottom-up traversal of $\mathcal{T}_n^x$ in which we keep track of the size $\mathsf{size}(u)$ of the subtree marked by the current vertex u and the vertices already in S. Once $\mathsf{size}(u) > n^{1-c}$, we add u to S and proceed as before. Since every time we add a vertex u to S we have $\mathsf{size}(u) > n^{1-c}$, it follows that $|S| \leq n^c$. Additionally, the subtree rooted in every child of u has size at most n^{1-c}, otherwise that child of u would have been chosen to be included in S instead of u.

Hub construction: hub $\mathcal{H}$. Given the set of vertices S constructed during the spanning tree partitioning, we construct the set of vertices $\mathcal{H} \subset V$ called the *hub*, as follows:

1. We choose a constant $\gamma = \varepsilon/3$.
2. For every vertex $u \in S$, we add in $\mathcal{H}$ every vertex v that lies in the unique simple path $P_u : x \rightsquigarrow u$ between the root x of $\mathcal{T}_n^x$ and u (including x and u). Since $\mathrm{diam}(G) \leq n^{1-\varepsilon}$ and $|S| \leq n^c$, we have that $|\mathcal{H}| \leq n^{1-\varepsilon+c} \leq n^{1-\gamma}$.

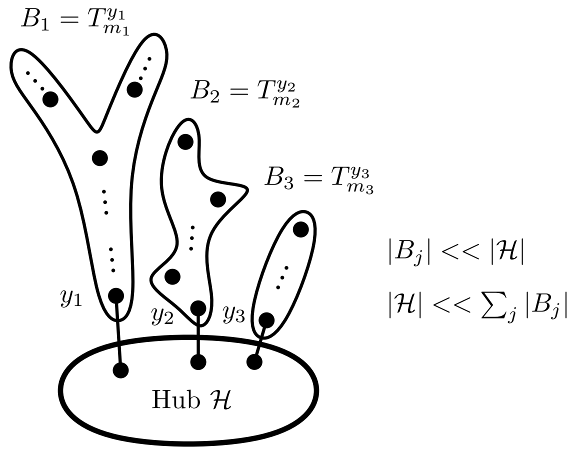

Figure 3: Illustration of the hub $\mathcal{H}$ and the branches $T^{y_j}_{m_j}$.

3. We add $n^{1-\gamma} - |\mathcal{H}|$ extra vertices to $\mathcal{H}$, such that in the end, the vertices of $\mathcal{H}$ form a connected subtree of $\mathcal{T}^x_n$ (rooted in x). This is simply done by choosing a vertex $u \in \mathcal{H}$ and a neighbor v of u with $v \notin \mathcal{H}$, and adding v to $\mathcal{H}$, until $\mathcal{H}$ contains $n^{1-\gamma}$ vertices.

Branches $B_j = T^{y_j}_{m_j}$. The hub $\mathcal{H}$ defines a number of trees $B_j = T^{y_j}_{m_j}$, where each tree is rooted at a vertex $y_j \notin \mathcal{H}$ adjacent to $\mathcal{H}$, and has m_j vertices. We will refer to these trees as *branches*(see Figure 3).

Proposition 1. *Note that by construction, we have $m_j \leq n^{1-2/3\cdot\varepsilon}$ for every j, and $|\mathcal{H}| = n^{1-\varepsilon/3}$, and $\sum_j m_j = n - n^{1-\varepsilon/3}$.*

Notation. To make the exposition of the ideas clear, we rely on the following notation.

1. *Parent* $\mathsf{par}(u)$ *and ancestors* $\mathsf{anc}(u)$. Given a vertex $u \neq x$, we denote by $\mathsf{par}(u)$ the parent of u in $\mathcal{T}^x_n$ and by $\mathsf{anc}(u)$ the set of ancestors of u.
2. *Children* $\mathsf{chl}(u)$ *and descendants* $\mathsf{des}(u)$. Given a vertex u that is not a leaf in $\mathcal{T}^x_n$, we denote by $\mathsf{chl}(u)$ the children of u in $\mathcal{T}^x_n$ that do not belong to the hub $\mathcal{H}$, and by $\mathsf{des}(u)$ the set of descendants of u in $\mathcal{T}^x_n$ that do not belong to the hub $\mathcal{H}$.

Frontier, distance, and branches. We present few notions required for the weight assignment:

1. *Frontier* $\mathcal{F}$. Given the hub $\mathcal{H}$, the *frontier* of $\mathcal{H}$ is the set of vertices $\mathcal{F} \subseteq \mathcal{H}$ defined as
$$\mathcal{F} = \bigcup_{u \in V \setminus \mathcal{H}} \mathsf{Nh}(u) \cap \mathcal{H} \ .$$
In words, $\mathcal{F}$ contains all vertices of $\mathcal{H}$ that have a neighbor not in $\mathcal{H}$.
2. *Distance function* λ. For every vertex u, we define its *distance* $\lambda(u)$ to be the length of the shortest path $P : u \rightsquigarrow v$ in T^x_n to some vertex $v \in \mathcal{F}$ (e.g., if $u \in \mathcal{F}$, we have (i) $\lambda(u) = 0$, and (ii) for every $v \in \mathsf{Nh}(u) \setminus \mathcal{H}$ we have $\lambda(v) = 1$).
3. *Values* μ *and* ν. For every vertex $u \in \mathcal{H}$, we define $\mathsf{deg}(u) = |(\mathsf{Nh}(u) \cap \mathcal{H}) \setminus \{u\}|$ i.e., $\mathsf{deg}(u)$ is the number of neighbors of u that belong to the hub (excluding u itself). Let
$$\mu = \max_{u \in \mathcal{F}} |\mathsf{chl}(u)| \qquad \text{and} \qquad \nu = \max_{u \in \mathcal{H}} \mathsf{deg}(u) \ .$$

Weight assignment. We are now ready to define the weight function $\mathsf{w} : E \to \mathbb{R}_{\geq 0}$.

1. For every edge (u, v) such that $u \neq v$ and $u, v \notin \mathcal{H}$ and u and v are not neighbors in $\mathcal{T}_n^x$, we assign $\mathsf{w}(u, v) = 0$.
2. For every vertex $u \in \mathcal{F}$ we assign $\mathsf{w}(u, u) = (\mu - |\mathsf{chl}(u)|) \cdot 2^{-n} + \nu - \mathsf{deg}(u)$.
3. For every vertex $u \in \mathcal{H} \setminus \mathcal{F}$ we assign $\mathsf{w}(u, u) = \mu \cdot 2^{-n} + \nu - \mathsf{deg}(u)$.
4. For every vertex $u \notin \mathcal{H}$ we assign $\mathsf{w}(u, u) = n^{-2\cdot\lambda(u)}$.
5. For every edge $(u, v) \in E$ such that $u \neq v$ and $u, v \in \mathcal{H}$ we assign $\mathsf{w}(u, v) = 1$.
6. For every remaining edge $(u, v) \in E$ such that $u = \mathsf{par}(v)$ we assign $\mathsf{w}(u, v) = 2^{-n} \cdot n^{-4\cdot\lambda(u)}$.

The following lemma is straightforward from the weight assignment, and captures that every vertex in the hub has the same weight.

Lemma 2. *For every vertex $u \in \mathcal{H}$ we have $\mathsf{w}(u) = \sum_{v\in\mathsf{Nh}(u)} \mathsf{w}(u, v) = \mu \cdot 2^{-n} + \nu$.*

Proof. Consider any vertex $u \in \mathcal{H} \setminus \mathcal{F}$. We have

$$\begin{aligned}\mathsf{w}(u) =& \mathsf{w}(u, u) + \sum_{v\in\mathsf{Nh}(u)\setminus\{u\}} \mathsf{w}(u, v)\\ =& \mu \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \sum_{v\in\mathsf{Nh}(u)\setminus\{u\}} 1\\ =& \mu \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \mathsf{deg}(u)\\ =& \mu \cdot 2^{-n} + \nu \end{aligned} \tag{4}$$

Similarly, consider any $u \in \mathcal{F}$. We have

$$\begin{aligned}\mathsf{w}(u) =& \mathsf{w}(u, u) + \sum_{v\in(\mathsf{Nh}(u)\cap\mathcal{H})\setminus\{u\}} \mathsf{w}(u, v) + \sum_{v\in\mathsf{chl}(u)} \mathsf{w}(u, v)\\ =& (\mu - |\mathsf{chl}(u)|) \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \sum_{v\in(\mathsf{Nh}(u)\cap\mathcal{H})\setminus\{u\}} 1 + \sum_{v\in\mathsf{chl}(u)} 2^{-n}\\ =& \mu \cdot 2^{-n} - |\mathsf{chl}(u)| \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \mathsf{deg}(u) + |\mathsf{chl}(u)| \cdot 2^{-n}\\ =& \mu \cdot 2^{-n} + \nu \end{aligned} \tag{5}$$

□

6.3 Analysis of the Fixation Probability

In this section we present detailed analysis of the fixation probability and we start with the outline of the proof.

6.3.1 Outline of the proof

The fixation of new mutants is guaranteed by showing that each of the following four stages happens with high probability.

(A) In stage 1 we consider the event $\mathcal{E}_1$ that a mutant arises in one of the branches (i.e., outside the hub $\mathcal{H}$). We show that event $\mathcal{E}_1$ happens whp.

(B) In stage 2 we consider the event $\mathcal{E}_2$ that a mutant occupies a vertex v of the branches which is a neighbor to the hub. We show that given event $\mathcal{E}_1$ the event $\mathcal{E}_2$ happens whp.

(C) In stage 3 we consider the event $\mathcal{E}_3$ that the mutants fixate in the hub. We show that given event $\mathcal{E}_2$ the event $\mathcal{E}_3$ happens whp.

(D) In stage 4 we consider the event $\mathcal{E}_4$ that the mutants fixate in all the branches. We show that given event $\mathcal{E}_3$ the event $\mathcal{E}_4$ happens whp.

Crux of the proof. Before the details of the proof we present the main crux of the proof. We say a vertex $v \notin \mathcal{H}$ hits the hub when it places an offspring to the hub. First, our construction ensures that the hub is isothermal. Second, our construction ensures that a mutant appearing in a branch reaches to a vertex adjacent to the hub, and hits the hub with a mutant polynomially many times. Third, our construction also ensures that the hub reaches a homogeneous configuration whp between any two hits to the hub. We now describe two crucial events.

- Consider that a mutant is adjacent to a hub of residents. Every time a mutant is introduced in the hub it has a constant probability (around $1-1/r$ for large population) of fixation since the hub is isothermal. The polynomially many hits of the hub by mutants ensure that the hub becomes mutants whp.
- In contrast consider that a resident is adjacent to a hub. Every time a resident is introduced in the hub it has exponentially small probability (around $(r-1)/(r^{|\mathcal{H}|}-1)$) of fixation.

Hence, given a hub of mutants, the probability (say, $\eta_1 = 2^{-\Omega(|\mathcal{H}|)}$) that the residents win over the hub is exponentially small. Given a hub of mutant, the probability that the hub wins over a branch B_j is also exponentially small (say, $\eta_2 = 2^{-O(|B_j|)}$). More importantly the ratio of η_1/η_2 is also exponentially small (by Proposition 1 regarding the sizes of the hub and branches). Using this property, se show that fixation the mutants reach fixation whp. We now analyze each stage in detail.

6.3.2 Analysis of Stage 1: Event $\mathcal{E}_1$

Lemma 3. *Consider the event $\mathcal{E}_1$ that the initial mutant is placed at a vertex outside the hub. Formally, the event $\mathcal{E}_1$ is that $\mathsf{X}_0 \cap \mathcal{H} = \emptyset$. The event $\mathcal{E}_1$ happens with probability at least $1 - O(n^{-\varepsilon/3})$, i.e., the event $\mathcal{E}_1$ happens whp.*

Proof. We examine the uniform and temperature initialization schemes separately.

- *(Uniform initialization):* The initial mutant is placed on a vertex $u \notin \mathcal{H}$ with probability
$$\sum_{u \notin \mathcal{H}} \frac{1}{n} = \frac{|V \setminus \mathcal{H}|}{n} = \frac{n - n^{1-\gamma}}{n} = 1 - \frac{n^{1-\gamma}}{n} = 1 - O(n^{-\varepsilon/3}) \; ;$$
since $\gamma = \varepsilon/3$.
- *(Temperature initialization):* For any vertex $u \notin \mathcal{H}$, we have
$$\sum_{v \in \mathsf{Nh}(u)\setminus\{u\}} \mathsf{w}(u,v) \leq \sum_{v \in \mathsf{Nh}(u)\setminus\{u\}} 2^{-n} = 2^{-\Omega(n)} \; ;$$

whereas since $\mathrm{diam}(G) \leq n^{1-\varepsilon}$ we have

$$\mathsf{w}(u,u) = n^{-2\cdot\lambda(u)} \geq n^{-2\cdot\mathrm{diam}(G)} \geq n^{-O(n^{1-\varepsilon})} .$$

Note that

$$n^{-O(n^{1-\varepsilon})} = 2^{-O(n^{1-\varepsilon}\cdot\log n)} >> 2^{-O(n)} .$$

Let $A = \mathsf{w}(u,u)$ and $B = \sum_{v\in\mathsf{Nh}(u)\setminus\{u\}} \mathsf{w}(u,v)$, and we have

$$\frac{\mathsf{w}(u,u)}{\mathsf{w}(u)} = \frac{A}{A+B} = 1 - \frac{B}{A+B} = 1 - \frac{2^{-\Omega(n)}}{n^{-O(n^{1-\varepsilon})} + 2^{-\Omega(n)}} = 1 - \frac{2^{-\Omega(n)}}{n^{-O(n^{1-\varepsilon})}} = 1 - 2^{-\Omega(n)} .$$

Then the desired event happens with probability at least

$$\begin{aligned}\sum_{u\notin\mathcal{H}} \mathbb{P}^{\mathsf{T}}[\mathsf{X}_0 = \{u\}] = \sum_{u\notin\mathcal{H}} \frac{\mathsf{T}(u)}{n} &= \frac{1}{n}\cdot\sum_{u\notin\mathcal{H}}\sum_{v\in\mathsf{Nh}(u)} \frac{\mathsf{w}(u,v)}{\mathsf{w}(v)} \geq \frac{1}{n}\cdot\sum_{u\notin\mathcal{H}} \frac{\mathsf{w}(u,u)}{\mathsf{w}(u)} \geq \frac{1}{n}\cdot\sum_{u\notin\mathcal{H}}\left(1-2^{-\Omega(n)}\right)\\ &= \frac{|V\setminus\mathcal{H}|}{n}\cdot\left(1-2^{-\Omega(n)}\right) = \frac{n-n^{1-\gamma}}{n}\cdot\left(1-2^{-\Omega(n)}\right) = (1-n^{-\gamma})\cdot\left(1-2^{-\Omega(n)}\right)\\ &= 1 - O(n^{-\varepsilon/3})\end{aligned}$$

since $\gamma = \varepsilon/3$. The desired result follows.

□

6.3.3 Analysis of Stage 2: Event $\mathcal{E}_2$

The following lemma states that if a mutant is placed on a vertex w outside the hub, then whp the mutant will propagate to the ancestor v of w at distance $\lambda(v) = 1$ from the hub (i.e., the parent of v belongs to the hub). This is a direct consequence of the weight assignment, which guarantees that for every vertex $u \notin \mathcal{H}$, the individual occupying u will place an offspring on the parent of u before some neighbor of u places an offspring on u, and this event happens with probability at least $1 - O(n^{-1})$.

Lemma 4. *Consider that at some time j the configuration of the Moran process on G^{w} is $\mathsf{X}_j = \{w\}$ with $w \notin \mathcal{H}$. Let $v \in \mathsf{anc}(w)$ with $\lambda(v) = 1$, i.e., v is the ancestor of w and v is adjacent to the hub. Then a subsequent configuration X_t with $v \in \mathsf{X}_t$ is reached with probability $1 - O(n^{-1})$, i.e., given event $\mathcal{E}_1$, the event $\mathcal{E}_2$ happens whp.*

Proof. Let t be the first time such that $v \in \mathsf{X}_t$ (possibly $t = \infty$, denoting that v never becomes mutant). Let s_i be the random variable such that

$$s_i = \begin{cases} |\mathsf{X}_i \cap \mathsf{anc}(w)| & \text{if } i < t \\ |\mathsf{anc}(w)| & \text{if } i \geq t \end{cases}$$

In words, s_i counts the number of mutant ancestors of u until time t. Given the current configuration X_i with $0 < s_i < |\mathsf{anc}(w)|$, let $u = \arg\min_{z\in\mathsf{X}_i\cap\mathsf{anc}(w)} \lambda(z)$. The probability that $s_{i+1} = s_i + 1$ is lowerbounded by the probability that u reproduces and places an offspring on $\mathsf{par}(u)$. Similarly, the probability that $s_{i+1} = s_i - 1$ is upperbounded by the probability that (i) $\mathsf{par}(u)$ reproduces and places an offspring on u, plus (ii) the probability that some $z \in \mathsf{des}(u) \setminus \mathsf{X}_i$ reproduces and places an offspring on $\mathsf{par}(z)$.

We now proceed to compute the above probabilities. Consider any configuration X_i, and and let z be any child of u and z' any child of z. The above probabilities crucially depend on the following quantities:

$$\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u)}; \qquad \frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u))}; \qquad \sum_{z_i \in \mathsf{des}(u)} \frac{\mathsf{w}(\mathsf{par}(z_i), z_i)}{\mathsf{w}(z_i)} .$$

Recall that

- $\mathsf{w}(u,\mathsf{par}(u)) = 2^{-n} \cdot n^{-4\cdot\lambda(\mathsf{par}(u))}$
- $\mathsf{w}(u,x) = 2^{-n} \cdot n^{-4\cdot\lambda(u)}$
- $\mathsf{w}(z,z') = 2^{-n} \cdot n^{-4\cdot\lambda(z)}$
- $\mathsf{w}(\mathsf{par}(u),\mathsf{par}(\mathsf{par}(u))) = 2^{-n} \cdot n^{-4\cdot\lambda(\mathsf{par}(\mathsf{par}(u)))}$
- $\mathsf{w}(u,u) = n^{-2\cdot\lambda(u)}$
- $\mathsf{w}(\mathsf{par}(u),\mathsf{par}(u)) = n^{-2\cdot\lambda(\mathsf{par}(u))}$
- $\mathsf{w}(z,z) = n^{-2\cdot\lambda(z)}$

Thus, we have

$$\begin{aligned}\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u)} =& \frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u,u)+\mathsf{w}(u,\mathsf{par}(u))+|\mathsf{chl}(u)|\cdot\mathsf{w}(u,x)} = \frac{2^{-n}\cdot n^{-4\cdot(\lambda(u)-1)}}{O(n^{-2\cdot\lambda(u)})}\\ =&\Omega(2^{-n}\cdot n^{-2\cdot(\lambda(u)-2)})\end{aligned} \tag{6}$$

$$\begin{aligned}\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u))} =& \frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u),\mathsf{par}(u))+\mathsf{w}(\mathsf{par}(u),\mathsf{par}(\mathsf{par}(u)))+|\mathsf{chl}(\mathsf{par}(u))|\cdot\mathsf{w}(u,\mathsf{par}(u))}\\ =&\frac{2^{-n}\cdot n^{-4\cdot(\lambda(u)-1)}}{\Omega(n^{-2\cdot(\lambda(u)-1)})} = O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)})\end{aligned} \tag{7}$$

$$\begin{aligned}\sum_{z_i\in\mathsf{des}(u)}\frac{\mathsf{w}(\mathsf{par}(z_i),z_i)}{\mathsf{w}(z_i)} =& |\mathsf{des}(u)|\cdot\frac{\mathsf{w}(u,z)}{\mathsf{w}(z,z)+\mathsf{w}(u,z)+|\mathsf{chl}(z)|\cdot\mathsf{w}(z,z')}\\ \le& |\mathsf{des}(u)|\cdot\frac{2^{-n}\cdot n^{-4\cdot\lambda(u)}}{\Omega(n^{-2\cdot(\lambda(u)+1)})} = n\cdot O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)})\\ =&O(2^{-n}\cdot n^{-2\cdot\lambda(u)+3})\end{aligned} \tag{8}$$

Thus, using Eq. (6), Eq. (7) and Eq. (8), we obtain

$$\begin{aligned}\frac{\mathbb{P}[s_{i+1}=s_i+1]}{\mathbb{P}[s_{i+1}=s_i-1]} \ge& \frac{\frac{r}{\mathsf{F}(\mathsf{X}')}\cdot\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u)}}{\frac{1}{\mathsf{F}(\mathsf{X}')}\cdot\left(\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u))}+\sum_{z_i\in\mathsf{des}(u)}\frac{\mathsf{w}(\mathsf{par}(z_i),z_i)}{\mathsf{w}(z_i)}\right)}\\ =&\frac{\Omega(2^{-n}\cdot n^{-2\cdot(\lambda(u)-2)})}{O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)})+O(2^{-n}\cdot n^{-2\cdot\lambda(u)+3})} = \Omega(n)\end{aligned} \tag{9}$$

Let $\alpha(n) = 1 - O(n^{-1})$ and consider a one-dimensional random walk $P : s'_0, s'_1, \dots$ on states $0 \leq i \leq |\mathsf{anc}(w)|$, with transition probabilities

$$\mathbb{P}[s'_{i+1} = \ell | s'_i] = \begin{cases} \alpha(n) & \text{if } 0 < s'_i < |\mathcal{H}| \text{ and } \ell = s'_i + 1 \\ 1 - \alpha(n) & \text{if } 0 < s'_i < |\mathcal{H}| \text{ and } \ell = s'_i - 1 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Using Eq. (9), we have that

$$\frac{\mathbb{P}[s'_{i+1} = s'_i + 1]}{\mathbb{P}[s'_{i+1} = s'_i - 1]} = \frac{\alpha(n)}{1 - \alpha(n)} = \Omega(n) \leq \frac{\mathbb{P}[s_{i+1} = s_i + 1]}{\mathbb{P}[s_{i+1} = s_i - 1]} \ .$$

Hence the probability that $s_\infty = |\mathsf{anc}(w)|$ is lowerbounded by the probability that $s'_\infty = |\mathsf{anc}(w)|$. The latter event occurs with probability $1 - O(n^{-1})$ (see e.g., [18], [28, Section 6.3]), as desired. ☐

6.3.4 Analysis of Stage 3: Event $\mathcal{E}_3$

We now focus on the evolution on the hub $\mathcal{H}$, and establish several useful results.

1. First, we show that $G^{\mathsf{w}}[\mathcal{H}]$ is isothermal (Lemma 5)
2. Second, the above result implies that the hub behaves as a well-mixed population. Considering advantageous mutants ($r > 1$) this implies the following (Lemma 6).
   (a) Every time a mutant hits a hub of only residents, then the mutant has at least a *constant* probability of fixating in the hub.
   (b) In contrast, every time a resident hits a hub of only mutants, then the resident has *exponentially small* probability of fixating in the hub.
3. Third, we show that an initial mutant adjacent to the hub, hits the hub a polynomial number of times (Lemma 7).
4. Finally, we show that an initial mutant adjacent to the hub ensures fixating in the hub whp (Lemma 8), i.e., we show that given event $\mathcal{E}_2$ the event $\mathcal{E}_3$ happens whp.

We start with observing that the hub is isothermal, which follows by a direct application of the definition of isothermal (sub)graphs [22].

Lemma 5. *The graph $G^{\mathsf{w}}[\mathcal{H}]$ is isothermal.*

Proof. Consider any vertex $u \in \mathcal{H} \setminus \mathcal{F}$. We have

$$\begin{aligned}
\mathsf{T}[X](u) = &\sum_{v \in \mathsf{Nh}(u) \cap \mathcal{H}} \frac{\mathsf{w}[\mathcal{H}](v,u)}{\mathsf{w}[\mathcal{H}](v)} = \frac{\mathsf{w}[\mathcal{H}](u,u)}{\mathsf{w}[\mathcal{H}](u)} + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} \frac{\mathsf{w}[\mathcal{H}](v,u)}{\mathsf{w}[\mathcal{H}](v)} \\
=&\frac{\mathsf{w}(u,u)}{\mathsf{w}(u)} + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} \frac{\mathsf{w}(v,u)}{\mathsf{w}(v)} \\
=&\frac{1}{\mu \cdot 2^{-n} + \nu} \cdot \left(\mathsf{w}(u,u) + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} 1 \right) \\
=&\frac{1}{\mu \cdot 2^{-n} + \nu} \cdot (\mu \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \mathsf{deg}(u)) \\
=&1
\end{aligned}$$

since by Lemma 2 we have $\mathsf{w}(u) = \mu \cdot 2^{-n} + \nu$. Similarly, consider any $u \in \mathcal{F}$. We have

$$\begin{aligned}
\mathsf{T}[X](u) = &\sum_{v \in \mathsf{Nh}(u) \cap \mathcal{H}} \frac{\mathsf{w}[\mathcal{H}](v,u)}{\mathsf{w}[\mathcal{H}](v)} = \frac{\mathsf{w}[\mathcal{H}](u,u)}{\mathsf{w}[\mathcal{H}](u)} + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} \frac{\mathsf{w}[\mathcal{H}](v,u)}{\mathsf{w}[\mathcal{H}](v)} \\
=&\frac{\mathsf{w}(u,u) + \sum_{v \in \mathsf{Nh}(u) \setminus \mathcal{H}} \mathsf{w}(u,v)}{\mathsf{w}(u)} + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} \frac{\mathsf{w}(v,u)}{\mathsf{w}(v)} \\
=&\frac{1}{\mu \cdot 2^{-n} + \nu} \cdot \left(\mathsf{w}(u,u) + \sum_{v \in \mathsf{Nh}(u) \setminus \mathcal{H}} 2^{-n} + \sum_{v \in (\mathsf{Nh}(u)\setminus\{u\}) \cap \mathcal{H}} 1 \right) \\
=&\frac{1}{\mu \cdot 2^{-n} + \nu} \cdot ((\mu - |\mathsf{chl}(u)|) \cdot 2^{-n} + \nu - \mathsf{deg}(u) + |\mathsf{chl}(u)| \cdot 2^{-n} + \mathsf{deg}(u)) \\
=&1
\end{aligned}$$

Thus for all $u \in \mathcal{H}$ we have $\mathsf{T}[X](u) = 1$, as desired. □

Lemma 6. *Consider that at some time j the configuration of the Moran process on G^{w} is X_j.*

1. *If $|\mathcal{H} \cap \mathsf{X}_j| \geq 1$, i.e., there is at least one mutant in the hub, then a subsequent configuration X_t with $\mathcal{H} \subseteq \mathsf{X}_t$ will be reached with probability at least $1 - r^{-1} - 2^{-\Omega(n)}$ (i.e., mutants fixate in the hub with constant probability).*
2. *If $|\mathcal{H} \setminus \mathsf{X}_j| = 1$, i.e., there is exactly one resident in the hub, then a subsequent configuration X_t with $\mathcal{H} \subseteq \mathsf{X}_t$ will be reached with probability at least $1 - 2^{-\Omega(m)}$, where $m = n^{1-\gamma}$ (i.e., mutants fixate in the hub with probability exponentially close to 1).*

Proof. Given a configuration X_i, denote by $s_i = |\mathcal{H} \cap \mathsf{X}_i|$. Let X_i be any configuration of the Moran process with $0 < s_i < |\mathsf{X}_i|$, u be the random variable that indicates the vertex that is chosen for reproduction in X_i, and X_{i+1} be the random variable that indicates the configuration of the population in the next step. By Lemma 5, the subgraph $G^{\mathsf{w}}[\mathcal{H}]$ induced by the hub $\mathcal{H}$ is isothermal, thus

$$\frac{\mathbb{P}[s_{i+1} = s_i - 1 | u \in \mathcal{H}]}{\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}]} = \frac{1}{r} \, . \tag{11}$$

Additionally,

$$\mathbb{P}[s_{i+1}=s_i-1|u\notin\mathcal{H}]\leq\sum_{\substack{v\in\mathcal{F}\\u\in\mathsf{chl}(v)}}\left(\frac{1}{\mathsf{F}(\mathsf{X}_i)}\cdot\frac{\mathsf{w}(u,v)}{\mathsf{w}(u)}\right)\leq n^{-1}\cdot\sum_{\substack{v\in\mathcal{F}\\u\in\mathsf{chl}(v)}}\frac{2^{-n}}{n^{-2}}$$
$$\leq n^{-1}\cdot n\cdot 2^{-n}\cdot n^2=O(n^2\cdot 2^{-n}) \tag{12}$$

since $1/\mathsf{F}(\mathsf{X}_i)\leq n^{-1}$, $\mathsf{w}(u,v)=2^{-n}$ and $\mathsf{w}(u,u)=n^{-2}$. Moreover, as $\mathcal{H}$ is heterogeneous, it contains at least a mutant vertex v and a resident vertex $w\in\mathsf{Nh}(v)$, and v reproduces with probability $r/\mathsf{F}(\mathsf{X}_i)\geq n^{-1}$, and replaces the individual $v\in\mathcal{H}$ with probability at least $1/\mathsf{w}(v)$. Hence we have

$$\mathbb{P}[s_{i+1}=s_i+1|u\in\mathcal{H}]\cdot\mathbb{P}[u\in\mathcal{H}]\geq\frac{1}{\mathsf{w}(u)}\cdot\frac{r}{\mathsf{F}(\mathsf{X}_i)}\geq\frac{1}{\mu\cdot 2^{-n}+\nu}\cdot n^{-1}\geq\frac{1}{n\cdot 2^{-n}+n}\cdot n^{-1}=\Omega(n^{-2}) \tag{13}$$

since by Lemma 2 we have $\mathsf{w}(v)=\mu\cdot 2^{-n}+\nu$. Using Eq. (11), Eq. (12) and Eq. (13), we have

$$\begin{aligned}\frac{\mathbb{P}[s_{i+1}=s_i-1]}{\mathbb{P}[s_{i+1}=s_i+1]}&=\frac{\mathbb{P}[s_{i+1}=s_i-1|u\in\mathcal{H}]\cdot\mathbb{P}[u\in\mathcal{H}]+\mathbb{P}[s_{i+1}=s_i-1|u\notin\mathcal{H}]\cdot\mathbb{P}[u\notin\mathcal{H}]}{\mathbb{P}[s_{i+1}=s_i+1|u\in\mathcal{H}]\cdot\mathbb{P}[u\in\mathcal{H}]+\mathbb{P}[s_{i+1}=s_i+1|u\notin\mathcal{H}]\cdot\mathbb{P}[u\notin\mathcal{H}]}\\&\leq\frac{\mathbb{P}[s_{i+1}=s_i-1|u\in\mathcal{H}]\cdot\mathbb{P}[u\in\mathcal{H}]+\mathbb{P}[s_{i+1}=s_i-1|u\notin\mathcal{H}]\cdot\mathbb{P}[u\notin\mathcal{H}]}{\mathbb{P}[s_{i+1}=s_i+1|u\in\mathcal{H}]\cdot\mathbb{P}[u\in\mathcal{H}]}\\&\leq\frac{\mathbb{P}[s_{i+1}=s_i-1|u\in\mathcal{H}]}{\mathbb{P}[s_{i+1}=s_i+1|u\in\mathcal{H}]}+O(n^2)\cdot\mathbb{P}[s_{i+1}=s_i-1|u\notin\mathcal{H}]=\frac{1}{r}+2^{-\Omega(n)}\end{aligned} \tag{14}$$

Hence, $s_j,s_{j+1},\dots$ performs a one-dimensional random walk on the states $0\leq i\leq|\mathcal{H}|$, with the ratio of transition probabilities given by Eq. (14). Let $\alpha(n)=r/(r+1+2^{-\Omega(n)})$ and consider the one-dimensional random walk $\rho: s'_j,s'_{j+1},\dots$ on states $0\leq i\leq|\mathcal{H}|$, with transition probabilities

$$\mathbb{P}[s'_{i+1}=\ell|s'_i]=\begin{cases}\alpha(n) & \text{if } 0<s'_i<|\mathcal{H}| \text{ and } \ell=s'_i+1\\ 1-\alpha(n) & \text{if } 0<s'_i<|\mathcal{H}| \text{ and } \ell=s'_i-1\\ 0 & \text{otherwise}\end{cases} \tag{15}$$

Using Eq. (14) we have that

$$\frac{\mathbb{P}[s'_{i+1}=s'_i-1]}{\mathbb{P}[s'_{i+1}=s'_i+1]}=\frac{1-\alpha(n)}{\alpha(n)}=\frac{1}{r}+2^{-\Omega(n)}\geq\frac{\mathbb{P}[s_{i+1}=s_i-1]}{\mathbb{P}[s_{i+1}=s_i+1]}\ .$$

Let ρ_1 (resp. ρ_2) be the probability that the Moran process starting on configuration X_j with $|\mathcal{H}\cap\mathsf{X}_j|\geq 1$ (resp. $|\mathcal{H}\setminus\mathsf{X}_j|=1$) will reach a configuration X_t with $\mathcal{H}\subseteq\mathsf{X}_t$. We have that ρ_1 (resp. ρ_2) is lowerbounded by the probability that ρ gets absorbed in $s'_\infty=|\mathcal{H}|$ when it starts from $s'_j=1$ (resp. $s'_j=|\mathcal{H}|-1$). Let

$$\beta=\frac{\mathbb{P}[s'_{i+1}=s'_i-1]}{\mathbb{P}[s'_{i+1}=s'_i+1]}=\frac{1}{r}+2^{-\Omega(n)}<1\ ;$$

and we have (see e.g., [18], [28, Section 6.3])

$$\rho_1\geq\frac{1-\beta}{1-\beta^{|\mathcal{H}|}}\geq 1-\beta=1-\frac{1}{r}-2^{-\Omega(n)}\ ;$$

and

$$\rho_2 \geq 1 - \frac{1-\beta^{-1}}{1-\beta^{-|\mathcal{H}|}} \geq 1 - \frac{\beta^{-1}}{\beta^{-|\mathcal{H}|}} = 1 - \beta^{|\mathcal{H}|-1} = 1 - \left(\frac{1}{r} + 2^{-\Omega(n)}\right)^{n^{1-\gamma}-1} = 1 - 2^{-\Omega(n^{1-\gamma})} ;$$

since $\beta^{-|\mathcal{H}|} > \beta^{-1}$ and thus $(\beta^{-1}-1)/(\beta^{-|\mathcal{H}|}-1) \leq \beta^{-1}/\beta^{-|\mathcal{H}|}$. The desired result follows. □

Lemma 7. *Consider that at some time j the configuration of the Moran process on G^{w} is X_j such that $v \in \mathsf{X}_j$ for some $v \notin \mathcal{H}$ that is adjacent to the hub ($\lambda(v) = 1$). Then a mutant hits the hub at least $n^{1/3}$ times with probability $1 - O(n^{-1/3})$.*

Proof. For any configuration X_i occurring after X_j, let

1. A be the event that v places an offspring on $\mathsf{par}(v)$ in X_{i+1}, and
2. B be the event that a neighbor of v places an offspring on v in X_{i+1},

and let ρ_A and ρ_B be the corresponding probabilities. Using Eq. (6), we have

$$\rho_A = \frac{r}{\mathsf{F}(\mathsf{X}_i)} \cdot \frac{\mathsf{w}(v, \mathsf{par}(v))}{\mathsf{w}(v)} = \Omega\left(n \cdot 2^{-n}\right) ; \tag{16}$$

and using Eq. (7) and Eq. (8)

$$\rho_B \leq \frac{r}{\mathsf{F}(\mathsf{X}_i)} \cdot \left(\frac{\mathsf{w}(v, \mathsf{par}(v))}{\mathsf{w}(\mathsf{par}(u))} + \sum_{z \in \mathsf{chl}(v)} \frac{\mathsf{w}(v, z)}{\mathsf{w}(z)}\right) \leq \frac{r}{n} \cdot \left(2^{-n} + O\left(n \cdot 2^{-n}\right)\right) = 2^{-\Omega(n)} . \tag{17}$$

since $\mathsf{par}(u) \in \mathcal{H}$ and by Lemma 2 we have $\mathsf{w}(\mathsf{par}(u)) \geq 1$. Let X be the random variable that counts the time required until event A occurs $n^{1/3}$ times. Then, for all $\ell \in \mathbb{N}$ we have $\mathbb{P}[X \geq \ell] \leq \mathbb{P}[X' \geq \ell]$ where X' is a random variable that follows the negative binomial distribution on $n^{1/3}$ failures with success rate $\rho_{X'} = 1 - O(n \cdot 2^{-n}) \leq \rho_A$ (using Eq. (16)). The expected value of X' is

$$\mathsf{E}[X'] = \frac{\rho_{X'} \cdot n^{1/3}}{1 - \rho_{X'}} = O\left(\frac{1 - n \cdot 2^{-n}}{n^{2/3} \cdot 2^{-n}}\right) .$$

Let $\alpha = 2^n \cdot n^{-1/3}$, and by Markov's inequality, we have

$$\mathbb{P}[X' \geq \alpha] \leq \frac{\mathsf{E}[X']}{\alpha} = \frac{O\left(\frac{1-n\cdot 2^{-n}}{n^{2/3}\cdot 2^{-n}}\right)}{2^n \cdot n^{-1/3}} = O(n^{-1/3}) .$$

Similarly, let Y be the random variable that counts the time required until event B occurs. Then, for all $\ell \in \mathbb{N}$, we have $\mathbb{P}[Y \leq \ell] \leq \mathbb{P}[Y' \leq \ell]$, where Y' is a geometrically distributed variable with rate $\rho_{Y'} = 2^{-\Omega(n)} \geq \rho_B$ (using Eq. (17)). Then

$$\mathbb{P}[Y' \leq \alpha] = 1 - (1 - \rho_{Y'})^{\alpha} = O(n^{-1/3}) ;$$

and thus

$$\mathbb{P}[Y \leq X] \leq \mathbb{P}[Y \leq \alpha] + \mathbb{P}[X \geq \alpha] \leq \mathbb{P}[Y' \leq \alpha] + \mathbb{P}[X' \geq \alpha] = O(n^{-1/3}) . \tag{18}$$

Hence, with probability at least $1 - O(n^{1/3})$, the vertex v places an offspring on $\mathsf{par}(v)$ at least $n^{1/3}$ times before it is replaced by a neighbor. The desired result follows. □

Lemma 8. *Consider that at some time j the configuration of the Moran process on G^{w} is X_j with $v \in \mathsf{X}_j$ for some $v \notin \mathcal{H}$ that is adjacent to the hub ($\lambda(v) = 1$). Then a subsequent configuration X_t with $\mathcal{H} \subseteq \mathsf{X}_t$ (mutants fixating in the hub) is reached with probability $1 - O(n^{-1/3})$, i.e., given event $\mathcal{E}_2$, the event $\mathcal{E}_3$ happens whp.*

Proof. By Lemma 7, we have that with probability at least $\Omega(n^{1/3})$, the vertex v places an offspring on $\mathsf{par}(v)$ at least $n^{1/3}$ times before it is replaced by a neighbor. Let t_i be the time that v places its i-th offspring on $\mathsf{par}(v)$, with $1 \leq i \leq n^{1/3}$. Let A_i be the event that a configuration X_t is reached, where $t \geq t_i$ and such that $\mathcal{H} \subseteq \mathsf{X}_t$. By Lemma 6, we have $\mathbb{P}[A_i] \geq 1 - r^{-1} - 2^{-\Omega(n)}$. Moreover, with probability $1 - 2^{-\Omega(n)}$, at each time t_i the hub is in a homogeneous state, i.e., either $\mathcal{H} \subseteq \mathsf{X}_{t_i}$ or $\mathcal{H} \cap \mathsf{X}_{t_i} = \emptyset$. The proof is similar to that of Lemma 9, and is based on the fact that every edge which has one end on the hub and the other outside the hub has exponentially small weight (i.e., 2^{-n}), whereas the hub $G^{\mathsf{w}}[\mathcal{H}]$ resolves to a homogeneous state in polynomial time with probability exponentially close to 1. It follows that with probability at least $p = 1 - 2^{-\Omega(n)}$, the events $\bar{A}_i$ are pairwise independent, and thus

$$\mathbb{P}[\overline{A}_1 \cap \overline{A}_2 \cdots \cap \overline{A}_{n^{1/3}}] \leq p \cdot \prod_{i=1}^{n^{1/3}} \mathbb{P}[\overline{A}_i] + (1-p) \leq \prod_{i=1}^{n^{1/3}} (1 - \mathbb{P}[A_i]) + 2^{-\Omega(n)} \leq \left(r^{-1} + 2^{-\Omega(n)}\right)^{n^{1/3}} + 2^{-\Omega(n)} . \tag{19}$$

Finally, starting from $\mathsf{X}_0 = \{u\}$, the probability that a configuration X_t is reached such that $\mathcal{H} \subseteq \mathsf{X}_t$ is lowerbounded by the probability of the events that

1. the ancestor v of u is eventually occupied by a mutant, and
2. v places at least $n^{1/3}$ offsprings to $\mathsf{par}(v) \in \mathcal{H}$ before a neighbor of v places an offspring on v, and
3. the event $\overline{A}_1 \cap \overline{A}_2 \cdots \cap \overline{A}_{n^{1/3}}$ does not occur.

Combining Lemma 4, Eq. (18) and Eq. (19), we obtain that the goal configuration X_t is reached with probability at least

$$(1 - O(n^{-1})) \cdot (1 - O(n^{-1/3})) \cdot \left(1 - \mathbb{P}[\overline{A}_1 \cap \overline{A}_2 \cdots \cap \overline{A}_{n^{1/3}}]\right) = 1 - O(n^{-1/3}) ;$$

as desired. □

6.3.5 Analysis of Stage 4: Event $\mathcal{E}_4$

In this section we present the last stage to fixation. This is established in four intermediate steps.

1. First, we consider the event of some vertex in the hub placing an offspring in one of the branches, while the hub is heterogeneous. We show that this event has exponentially small probability of occurring (Lemma 9).
2. We introduce the *modified* Moran process which favors residents when certain events occur, more than the conventional Moran process. This modification underapproximates the fixation probability of mutants, but simplifies the analysis.
3. We define a set of simple Markov chains $\mathcal{M}_j$ and show that the fixation of mutants on the j-th branch $T_{m_j}^{y_j}$ is captured by the absorption probability to a specific state of $\mathcal{M}_j$ (Lemma 11). This absorption probability is computed in Lemma 10.

4. Finally we combine the above steps in Lemma 12 to show that if the hub is occupied by mutants (i.e., given that event $\mathcal{E}_3$ holds), the mutants eventually fixate in the graph (i.e., event $\mathcal{E}_4$ holds) whp.

We start with an intermediate lemma, which states that while the hub is heterogeneous, the probability that a node from the hub places an offspring to one of the branches is exponentially small.

Lemma 9. *For any configuration X_j with $|\mathcal{H} \setminus \mathsf{X}_j| = 1$, let $t_1 \geq j$ be the first time such that $\mathcal{H} \subseteq \mathsf{X}_{t_1}$ (possibly $t_1 = \infty$), and $t_2 \geq j$ the first time in which a vertex $u \in \mathcal{F}$ places an offspring on some vertex $v \in \mathsf{Nh}(u) \setminus \mathcal{H}$. We have that $\mathbb{P}[t_2 < t_1] = 2^{-\Omega(m)}$, where $m = n^{1-\gamma}$.*

Proof. Given a configuration X_i, denote by $s_i = |\mathcal{H} \cap \mathsf{X}_i|$. Recall from the proof of Lemma 8 that $s_j, s_{j+1}, \dots$ performs a one-dimensional random walk on the states $0 \leq i \leq |\mathcal{H}|$, with the ratio of transition probabilities given by Eq. (14). Observe that in each s_i, the random walk changes state with probability at least n^{-2}, which is a lowerbound on the probability that the walk progresses to $s_{i+1} = s_i + 1$ (i.e., the mutants increase by one). Consider that the walk starts from s_j, and let H_a be the expected absorption time, H_f the expected fixation time on state $|\mathcal{H}|$, and H_e the expected extinction time on state 0 of the random walk, respectively. The unlooped variant of the random walk $\rho = s_i, s_{i+1}, \dots$ has expected absorption time $O(n)$ [19], hence the random walk $s_j, s_{j+1}, \dots$ has expected absorption time

$$H_a \leq n^2 \cdot O(n) = O(n^3) \ ;$$

and since by Lemma 6 for large enough n we have $\mathbb{P}[s_\infty = |\mathcal{H}|] \geq \mathbb{P}[s_\infty = 0]$, we have

$$H_a = \mathbb{P}[s_\infty = |\mathcal{H}|] \cdot H_f + \mathbb{P}[s_\infty = 0] \cdot H_e \implies H_f \leq 2 \cdot H_a = O(n^3) \ .$$

Let t_1' be the random variable defined as $t_1' = t_1 - j$, and we have

$$\mathsf{E}[t_1' | t_1' < \infty] = H_f = O(n^3) \ ;$$

i.e., given that a configuration X_{t_1} with $\mathcal{H} \subseteq \mathsf{X}_{t_1}$ is reached (thus $t_1 < \infty$ and $t_1' < \infty$), the expected time we have to wait after time j for this event to happen equals the expected fixation time H_f of the random walk $s_j, s_{j+1}, \dots$. Let $\alpha = 2^{\frac{n}{2}}$, and by Markov's inequality, we have

$$\mathbb{P}[t_1' > \alpha | t_1' < \infty] \leq \frac{\mathsf{E}[t_1' | t_1' < \infty]}{\alpha} = n^3 \cdot 2^{-\frac{n}{2}} \ . \tag{20}$$

Consider any configuration X_i. The probability p that a vertex $u \in \mathcal{F}$ places an offspring on some vertex $v \in \mathsf{Nh}(u) \setminus \mathcal{H}$ is at most

$$p \leq \frac{r}{\mathsf{F}(\mathsf{X}_i)} \cdot \sum_{u \in \mathcal{F}} \sum_{v \in \mathsf{Nh}(u) \setminus \mathcal{H}} \frac{\mathsf{w}(u,v)}{\mathsf{w}(u)} \leq r \cdot n^{-1} \cdot n^{1-\gamma} \cdot 2^{-n} \leq r \cdot n^2 \cdot 2^{-n} \ .$$

since $\mathsf{w}(u,v) = 2^{-n}$ and by Lemma 2 we have $\mathsf{w}(u) > 1$. Let $t_2' = t_2 - i$, and we have $\mathbb{P}[t_2' \leq \alpha] \leq \mathbb{P}[X \leq \alpha]$, where X is a geometrically distributed random variable with rate $\rho = r \cdot n^2 \cdot 2^{-n}$. Since

$\mathbb{P}[t_2 < t_1] = \mathbb{P}[t'_2 < t'_1]$, we have

$$\begin{aligned}\mathbb{P}[t_2 < t_1] =& \mathbb{P}[t'_2 < t'_1 | t'_1 < \infty] \cdot \mathbb{P}[t'_1 < \infty] + \mathbb{P}[t'_2 < t'_1 | t'_1 = \infty] \cdot \mathbb{P}[t'_1 = \infty] \\ \leq& \mathbb{P}[t'_2 < t'_1 | t'_1 < \infty] + \mathbb{P}[t'_1 = \infty] \\ \leq& \mathbb{P}[t'_2 < t'_1 | t_1 < \infty] + 2^{-\Omega(n^{1-\gamma})} \\ \leq& \mathbb{P}[t'_2 \leq \alpha | t'_1 < \infty] + \mathbb{P}[t'_1 > \alpha | t'_1 < \infty] + 2^{-\Omega(n^{1-\gamma})} \\ \leq& \mathbb{P}[t'_2 \leq \alpha | t'_1 < \infty] + n^3 \cdot 2^{-\frac{n}{2}} + 2^{-\Omega(n^{1-\gamma})} \\ \leq& \mathbb{P}[X \leq \alpha] + 2^{-\Omega(n^{1-\gamma})} \\ \leq& 1 - (1-\rho)^{\alpha} + 2^{-\Omega(n^{1-\gamma})} \\ \leq& 1 - (1 - r \cdot n^2 \cdot 2^{-n})^{2^{n/2}} + 2^{-\Omega(n^{1-\gamma})} \\ =& 2^{-\Omega(n^{1-\gamma})}\end{aligned}$$

The second inequality holds since by Lemma 6 we have $\mathbb{P}[t'_1 = \infty] = 2^{-\Omega(n^{1-\gamma})}$. The fourth inequality comes from Eq. (20). ☐

To simplify the analysis, we replace the Moran process with a *modified* Moran process, which favors the residents (hence it is conservative) and allows for rigorous derivation of the fixation probability of the mutants.

The modified Moran process. Consider the Moran process on G^{w}, and assume there exists a first time $t^* < \infty$ when a configuration X_{t^*} is reached such that $\mathcal{H} \subseteq \mathsf{X}_{t^*}$. We underapproximate the fixation probability of the Moran process starting from X_{t^*} by the fixation probability of the *modified* Moran process $\overline{\mathsf{X}}_{t^*}, \overline{\mathsf{X}}_{t^*+1}, \dots$, which behaves as follows. Recall that for every vertex y_j with $\lambda(y_j) = 1$, we denote by $T^{y_j}_{m_j}$ the subtree of $\mathcal{T}^x_n$ rooted at y_j, which has m_j vertices. Let V_i be the set of vertices of $T^{y_i}_{m_i}$, and note that by construction $m_i \leq n^{1-c}$, while there are at most n such trees. The *modified* Moran process is identical to the Moran process, except for the following modifications.

1. Initially, $\overline{\mathsf{X}}_{t^*} = \mathcal{H}$.
2. At any configuration $\overline{\mathsf{X}}_i$ with $\mathcal{H} \in \overline{\mathsf{X}}_i$, for all trees $T^{y_j}_{m_j}$, if a resident vertex $u \in V_j$ places an offspring on some vertex v with $u \neq v$, then $\overline{\mathsf{X}}_{i+1} = \overline{\mathsf{X}}_i \setminus V_j$ and $|\mathcal{H} \setminus \overline{\mathsf{X}}_{i+1}| = 1$ i.e., all vertices of $T^{y_j}_{m_j}$ become residents and the hub is invaded by a single resident.
3. If the modified process reaches a configuration $\overline{\mathsf{X}}_i$ with $\overline{\mathsf{X}}_i \cap \mathcal{H} = \emptyset$, the process instead transitions to configuration $\overline{\mathsf{X}}_i = \emptyset$, i.e., if the hub becomes resident, then all mutants go extinct.
4. At any configuration $\overline{\mathsf{X}}_i$ with $\mathcal{H} \setminus \overline{\mathsf{X}}_i \neq \emptyset$, if some vertex $u \in \mathcal{F}$ places an offspring on some vertex $v \in \mathsf{Nh}(u) \setminus \mathcal{H}$, then the process instead transitions to configuration $\overline{\mathsf{X}}_i = \emptyset$, i.e., if while the hub is heterogeneous, an offspring is placed from the hub to a vertex outside the hub, the mutants go extinct.

Note that any time a case of Item 1-Item 4 applies, the Moran and modified Moran processes transition to configurations X_i and $\overline{\mathsf{X}}_i$ respectively, with $\overline{\mathsf{X}}_i \subseteq \mathsf{X}_i$. Thus, the fixation probability of the Moran process on G^{w}_n is underapproximated by the fixation probability of the modified Moran process (i.e., we have $\mathbb{P}[\mathsf{X}_\infty = V | t^* < \infty] \geq \mathbb{P}[\overline{\mathsf{X}}_\infty = V]$). It is easy to see that Lemma 6 and Lemma 9 directly apply to the modified Moran process.

The Markov chain $\mathcal{M}_j$. Recall that $T^{y_j}_{m_j}$ refers to the j-th branch of the weighted graph G^{w}, rooted at the vertex y_j and consisting of m_j vertices. We associate $T^{y_j}_{m_j}$ with a Markov chain $\mathcal{M}_j$ of m_j+3 vertices, which captures the number of mutants in $T^{y_j}_{m_j}$, and whether the state of the hub. Intuitively, a state $0 \leq i \leq m_j$ of $\mathcal{M}_j$ represents a configuration where the hub is homogeneous and consists only of mutants, and there are i mutants in the branch $T^{y_j}_{m_j}$. The state $\mathcal{H}$ represents a configuration where the hub is heterogeneous, whereas the state $\mathcal{D}$ represents a configuration where the mutants have gone extinct in the hub, and thus the modified Moran process has terminated. We first present formally the Markov chain $\mathcal{M}_j$, and later (in Lemma 11) we couple $\mathcal{M}_j$ with the modified Moran process.

Consider any tree $T^{y_j}_{m_j}$, and let $\alpha = 1/(n^3+1)$. We define the Markov chain $\mathcal{M}_j = (\mathcal{X}_j, \delta_j)$ as follows:

1. The set of states is $\mathcal{X}_j = \{\mathcal{H}, \mathcal{D}\} \cup \{0, 1, \dots, m_j\}$
2. The transition probability matrix $\delta_j : \mathcal{X}_j \times \mathcal{X}_j \to [0,1]$ is defined as follows:
   (a) $\delta_j[i, i+1] = \alpha$ for $0 \leq i < m_j$,
   (b) $\delta_j[i, 0] = 1-\alpha$ for $1 < i < m_j$,
   (c) $\delta_j[0, \mathcal{H}] = 1-\alpha$,
   (d) $\delta_j[\mathcal{H}, 0] = 1 - 2^{-\Omega(m)}$, and $\delta_j[\mathcal{H}, \mathcal{D}] = 2^{-\Omega(m)}$, where $m = n^{1-\gamma}$,
   (e) $\delta_j[m_j, m_j] = \delta_j[\mathcal{D}, \mathcal{D}] = 1$,
   (f) $\delta_j[x, y] = 0$ for all other pairs $x, y \in \mathcal{X}_j$

See Figure 4 for an illustration. The Markov chain $\mathcal{M}_j$ has two absorbing states, $\mathcal{D}$ and m_j. We denote

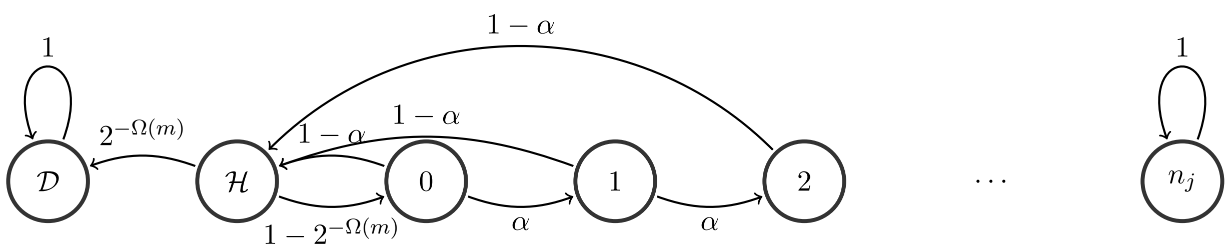

Figure 4: The Markov chain $\mathcal{M}_j$ given a tree $T^{x_j}_{n_j}$.

by ρ_j the probability that a random walk on $\mathcal{M}_j$ starting from state 0 will be absorbed in state m_j. The following lemma lowerbounds ρ_j, and comes from a straightforward analysis of $\mathcal{M}_j$.

Lemma 10. *For all Markov chains $\mathcal{M}_j$, we have $\rho_j = 1 - 2^{-\Omega(m)}$, where $m = n^{1-\gamma}$.*

Proof. Given a state $a \in \mathcal{X}_j$, we denote by x_a the probability that a random walk starting from state a will be absorbed in state m_j. Then $\rho_j = x_0$, and we have the following linear system

$$
\begin{aligned}
x_{\mathcal{H}} &= \delta[\mathcal{H}, 0] \cdot x_0 = \left(1 - 2^{\Omega(n^{1-\gamma})}\right) \cdot x_0 \\
x_i &= \delta[i, \mathcal{H}] \cdot x_{\mathcal{H}} + \delta[i, i+1] \cdot x_{i+1} = (1-\alpha) \cdot x_{\mathcal{H}} + \alpha \cdot x_{i+1} \qquad \text{for } 0 \leq i < m_j \\
x_{m_j} &= 1
\end{aligned}
$$

and thus

$$
\begin{aligned}
x_{\mathcal{H}} &= \left(1-2^{-\Omega(n^{1-\gamma})}\right)\cdot\left(x_{\mathcal{H}}\cdot(1-\alpha)\cdot\sum_{0=1}^{m_j}a^i+a^{m_j}\right)\\
\implies x_{\mathcal{H}} &= \left(1-2^{-\Omega(n^{1-\gamma})}\right)\cdot\left(x_{\mathcal{H}}\cdot\left(1-a^{m_j-1}\right)+a^{m_j}\right)\\
\implies x_{\mathcal{H}} &\left(1-\left(1-2^{-\Omega(n^{1-\gamma})}\right)\cdot\left(1-a^{m_j-1}\right)\right)=a^{m_j} \qquad (21)
\end{aligned}
$$

Note that

$$1-\left(1-2^{-\Omega(n^{1-\gamma})}\right)\cdot\left(1-a^{n_j-1}\right)\leq 2^{-\Omega(n^{1-\gamma})}+a^{n^j}\ ;$$

and from Eq. (21) we obtain

$$x_{\mathcal{H}}\geq\frac{\alpha^{n_j}}{2^{-\Omega(n^{1-\gamma})}+\alpha^{n_j}}=1-\frac{2^{-\Omega(n^{1-\gamma})}}{2^{-\Omega(n^{1-\gamma})}+\alpha^{n_j}}\geq 1-2^{-\Omega(n^{1-\gamma})}\cdot\alpha^{-n_j}=1-2^{-\Omega(n^{1-\gamma})}\cdot(n^3+1)^{n^{1-c}}=1-2^{-\Omega(n^{1-\gamma})}\ ;$$

since $a=1/(n^3+1)$ and by construction $n_j\leq n^{1-c}$ and $\gamma=\varepsilon/3<\varepsilon/2=c$. Finally, we have that $\rho_j=x_0\geq x_{\mathcal{H}}=1-2^{-\Omega(n^{1-\gamma})}$, as desired. □

Given a configuration $\overline{\mathsf{X}}_k$ of the modified Moran process, we denote by $\overline{\rho}_j(\overline{\mathsf{X}}_k)$ the probability that the process reaches a configuration $\overline{\mathsf{X}}_t$ with $\mathcal{H}\cup V_j\subseteq\overline{\mathsf{X}}_t$. The following lemma states that the probability $\overline{\rho}_j(\overline{\mathsf{X}}_\ell)$ is underapproximated by the probability ρ_j. The proof is by a coupling argument, which ensures that

1. every time the run on $\mathcal{M}_j$ is on a state $0\leq i\leq m_j$, there are at least i mutants placed on $T^{y_j}_{m_j}$, and
2. every time the modified Moran process transitions to a configuration where hub is heterogeneous (i.e., we reach a configuration X with $\mathcal{H}\setminus\mathsf{X}\neq\emptyset$), the run on $\mathcal{M}_j$ transitions to state $\mathcal{H}$.

Lemma 11. *Consider any configuration $\overline{\mathsf{X}}_\ell$ of the modified Moran process, with $\mathcal{H}\subseteq\overline{\mathsf{X}}_\ell$, and any tree $T^{y_j}_{m_j}$. We have $\overline{\rho}_j(\overline{\mathsf{X}}_\ell)\geq\rho_j$.*

Proof. The proof is by coupling the modified Moran process and the Markov chain $\mathcal{M}_j$. To do so, we let the modified Moran process execute, and use certain events of that process as the source of randomness for a run in $\mathcal{M}_j$. We describe the coupling process in high level. Intuitively, every time the run on $\mathcal{M}_j$ is on a state $0\leq i\leq m_j$, there are at least i mutants placed on $T^{y_j}_{m_j}$. Additionally, every time the modified Moran process transitions to a configuration where hub is heterogeneous (i.e., we reach a configuration X with $\mathcal{H}\setminus\mathsf{X}\neq\emptyset$), then the run on $\mathcal{M}_j$ transitions to state $\mathcal{H}$. Finally, if the modified Moran process ends on a configuration $\mathsf{X}=\emptyset$, then the run on $\mathcal{M}_j$ gets absorbed to state $\mathcal{D}$. The coupling works based on the following two facts.

1. For every state $0<i<m_j$, the ratio $\delta_j[i,i+1]/\delta_j[i,i-1]$ is upperbounded by the ratio of the probabilities of increasing the number of mutant vertices in $T^{y_j}_{m_j}$ by one, over decreasing that number by one and having the hub being invaded by a resident. Indeed, we have

$$\frac{\delta_j[i,i+1]}{\delta_j[i,i-1]}=\frac{\alpha}{1-\alpha}=\frac{1}{n^3}\ ;$$

while for every mutant vertex x of G with at last one resident neighbor, the probability that x becomes mutant in the next step of the modified Moran process over the probability that x becomes resident is at least $1/n^3$ (this ratio is at least $1/n^2$ for every resident neighbor y of x, and there are at most n such resident neighbors). The same holds for the ratio $\delta_j[0,1]/\delta_j[0,\mathcal{H}]$.

2. The probability of transitioning from state $\mathcal{H}$ to state 0 is upperbounded by the probability that once the mutant hub gets invaded by a resident the modified Moran process reaches a configuration where the hub consists of only mutants (using Lemma 6 and Lemma 9).

□

The following lemma captures the probability that the modified Moran process reaches fixation whp. That is, whp a configuration $\overline{\mathsf{X}}_i$ is reached which contains all vertices of G^{w}. The proof is based on repeated applications of Lemma 11 and Lemma 10, one for each subtree $T_{m_j}^{y_j}$.

Lemma 12. *Consider that at some time t^* the configuration of the Moran process on G^{w} is X_{t^*} with $\mathcal{H} \subseteq \mathsf{X}_{t^*}$. Then, a subsequent configuration X_t with $\mathsf{X}_t = V$ is reached with probability at least $1 - 2^{-\Omega(m)}$ where $m = n^{1-\gamma}$, i.e., given event $\mathcal{E}_3$, the event $\mathcal{E}_4$ is happens whp.*

Proof. It suffices to consider the modified Moran process on G starting from configuration $\overline{\mathsf{X}}_{t^*} = \mathcal{H}$, and showing that whp we eventually reach a configuration $\overline{\mathsf{X}}_t = V$. First note that if there exists a configuration $\overline{\mathsf{X}}_{t'}$ with $V_i \subseteq \overline{\mathsf{X}}_{t'}$ for any V_i, then for all $t'' \geq t'$ with $\overline{\mathsf{X}}_{t''} \neq \emptyset$ we have $V_i \subseteq \overline{\mathsf{X}}_{t''}$. Let $t_1 = t^*$. Since $\mathcal{H} \subseteq \overline{\mathsf{X}}_{t_1}$, by Lemma 11, with probability $\overline{\rho}_1(\overline{\mathsf{X}}_{t_1}) \geq \rho_1$ there exists a time $t_2 \geq t_1$ such that $\mathcal{H} \cup V_1 \subseteq \overline{\mathsf{X}}_{t_2}$. Inductively, given the configuration $\overline{\mathsf{X}}_{t_i}$, with probability $\overline{\rho}_i(\overline{\mathsf{X}}_{t_i}) \geq \rho_i$ there exists a time $t_{i+1} \geq t_i$ such that $\mathcal{H} \cup V_1 \cup \dots \cup V_i \subseteq \overline{\mathsf{X}}_{t_{i+1}}$. Since $V = \mathcal{H} \cup (\bigcup_{i=1}^k V_i)$, we obtain

$$\mathbb{P}[\overline{\mathsf{X}}_\infty = V] \geq \prod_{i=1}^{n} \rho_i = \prod_{i=1}^{n} \left(1 - 2^{-\Omega(n^{1-\gamma})}\right) \geq \left(1 - 2^{-\Omega(n^{1-\gamma})}\right)^n = 1 - 2^{-\Omega(m)} ;$$

as by Lemma 10 we have that $\rho_i = 1 - 2^{-\Omega(m)}$ for all i. The desired result follows. □

6.3.6 Main Positive Result

We are now ready to prove the main theorem of this section. First, combining Lemma 3, Lemma 4, Lemma 8 and Lemma 12, we obtain that if $r > 1$, then the mutants fixate G_n whp.

Lemma 13. *For any fixed $\varepsilon > 0$, for any graph G_n of n vertices and diameter $\operatorname{diam}(G_n) \leq n^{1-\varepsilon}$, there exists a weight function w such that for all $r > 1$, we have $\rho(G_n^{\mathsf{w}}, r, \mathsf{U}) = 1 - O(n^{-\varepsilon/3})$ and $\rho(G_n^{\mathsf{w}}, r, \mathsf{T}) = 1 - O(n^{-\varepsilon/3})$.*

It now remains to show that if $r < 1$, then the mutants go extinct whp. This is a direct consequence of the following lemma, which states that for any $r \geq 1$, the fixation probability of a mutant with relative fitness $1/r$ is upperbounded by one minus the fixation probability of a mutant with relative fitness r, in the same population.

Lemma 14. *For any graph G_n and any weight function w, for all $r \geq 1$, we have that $\rho(G_n^{\mathsf{w}}, 1/r, \mathsf{U}) \leq 1 - \rho(G_n^{\mathsf{w}}, r, \mathsf{U})$.*

Proof. Let σ be any irreflexive permutation of V (i.e., $\sigma(u) \neq u$ for all $u \in V$), and observe that for every vertex u, the probability that a mutant of fitness $1/r$ arising at u fixates in G_n is upperbounded by one minus the probability that a mutant of fitness r arising in $\sigma(u)$ fixates in G_n. We have

$$\begin{aligned}\rho(G_n^{\mathsf{w}}, 1/r, \mathsf{U}) =& \frac{1}{n}\sum_u \rho(G_n^{\mathsf{w}}, 1/r, u)\\ \leq& \frac{1}{n}\cdot \sum_u (1-\rho(G_n^{\mathsf{w}}, r, \sigma(u)))\\ =& 1-\frac{1}{n}\cdot\sum_{\sigma(u)}\rho(G_n^{\mathsf{w}}, r, u)\\ =& 1-\rho(G_n^{\mathsf{w}}, r, \mathsf{U})\end{aligned}$$

□

A direct consequence of the above lemma is that under uniform initialization, for any graph family where the fixation probability of advantageous mutants ($r > 1$) approaches 1, the fixation probability of disadvantageous mutants ($r < 1$) approaches zero. Since under our weight function w temperature initialization coincides with uniform initialization whp, Lemma 13 and Lemma 14 lead to the following corollary, which is our positive result.

Theorem 5. *Let $\varepsilon > 0$ and $n_0 > 0$ be any two fixed constants, and consider any sequence of unweighted, undirected graphs $(G_n)_{n>0}$ such that $\operatorname{diam}(G_n) \leq n^{1-\varepsilon}$ for all $n > n_0$. There exists a sequence of weight functions $(\mathsf{w}_n)_{n>0}$ such that the graph family $\mathcal{G} = (G_n^{\mathsf{w}_n})$ is a (i) strong uniform, (ii) strong temperature, and (iii) strong convex amplifier.*

References

[1] B. Adlam, K. Chatterjee, and M. A. Nowak. Amplifiers of selection. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2181), 2015.

[2] B. Adlam and M. A. Nowak. Universality of fixation probabilities in randomly structured populations. *Sci. Rep.*, 4, July 2014.

[3] B. Allen, C. Sample, Y. Dementieva, R. C. Medeiros, C. Paoletti, and M. A. Nowak. The molecular clock of neutral evolution can be accelerated or slowed by asymmetric spatial structure. *PLOS Computational Biology*, 11(2):1–32, 02 2015.

[4] T. Antal, S. Redner, and V. Sood. Evolutionary dynamics on degree-heterogeneous graphs. *Phys. Rev. Lett.*, 96:188104, 2006.

[5] M. Broom and J. Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 464(2098):2609–2627, Oct. 2008.

[6] K. Chatterjee, R. Ibsen-Jensen, and M. A. Nowak. Faster Monte-Carlo Algorithms for Fixation Probability of the Moran Process on Undirected Graphs. In *MFCS*, volume 83, pages 61:1–61:13, 2017.

[7] J. Díaz, L. A. Goldberg, G. B. Mertzios, D. Richerby, M. Serna, and P. G. Spirakis. On the fixation probability of superstars. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2156), 2013.

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

[9] J. Díaz, L. A. Goldberg, D. Richerby, and M. Serna. Absorption time of the moran process. *Random Structures & Algorithms*, 49(1):137–159, 2016.

[10] R. Durrett and S. A. Levin. Stochastic spatial models: a user's guide to ecological applications. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.*, 343(1305):329–350, 1994.

[11] M. Frean, P. B. Rainey, and A. Traulsen. The effect of population structure on the rate of evolution. *Proc. R. Soc. B Biol. Sci.*, 280(1762), July 2013.

[12] M. Frean, P. B. Rainey, and A. Traulsen. The effect of population structure on the rate of evolution. *Proc. R. Soc. B Biol. Sci.*, 280(1762):20130211, July 2013.

[13] A. Galanis, A. Göbel, L. A. Goldberg, J. Lapinskas, and D. Richerby. Amplifiers for the moran process. *J. ACM*, 64(1):5:1–5:90, Mar. 2017.

[14] G. Giakkoupis. Amplifiers and suppressors of selection for the moran process on undirected graphs. Technical report, 2016. https://arxiv.org/abs/1611.01585.

[15] L. A. Goldberg, J. Lapinskas, J. Lengler, F. Meier, K. Panagiotou, and P. Pfister. Asymptotically optimal amplifiers for the moran process. Technical report, 2017. https://arxiv.org/abs/1611.04209.

[16] B. Houchmandzadeh and M. Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New J. Phys.*, 13:073020, July 2011.

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

[18] J. Kemeny, D. Griffeath, J. Snell, and A. Knapp. *Denumerable Markov Chains: with a chapter of Markov Random Fields by David Griffeath*. Graduate Texts in Mathematics. Springer New York, 2012.

[19] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.

[20] S. A. Levin. Population dynamic models in heterogeneous environments. *Annu. Rev. Ecol. Syst.*, 7(1):287–310, 1976.

[21] S. A. Levin and R. T. Paine. Disturbance, patch formation, and community structure. *Proc. Natl. Acad. Sci.*, 71(7):2744–2747, 1974.

[22] E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 01 2005.

[23] T. Maruyama. A Markov process of gene frequency change in a geographically structured population. *Genetics*, 76(2):367–377, 1974.

[24] G. B. Mertzios, S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. Natural models for evolution on networks. *Theor. Comput. Sci.*, 477:76–95, 2013.

[25] G. B. Mertzios and P. G. Spirakis. Strong bounds for evolution in networks. ICALP, pages 669–680. Springer-Verlag, 2013.

[26] T. Monk, P. Green, and M. Paulin. Martingales and fixation probabilities of evolutionary graphs. *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 470(2165):20130730, 2014.

[27] P. A. P. Moran. *The statistical processes of evolutionary theory*. Oxford University Press, Oxford, England, 1962.

[28] M. Nowak. *Evolutionary Dynamics*. Harvard University Press, 2006.

[29] Z. Patwa and L. M. Wahl. The fixation probability of beneficial mutations. *J. R. Soc. Interface*, 5(28):1279–1289, Nov. 2008.

[30] A. Pavlogiannis, J. Tkadlec, K. Chatterjee, and M. A. Nowak. Amplification on undirected population structures: Comets beat stars. *Scientific Reports*, 7(1):82, 2017.

[31] M. Whitlock. Fixation probability and time in subdivided populations. *Genetics*, 779(June):767–779, 2003.