# A role of graphs in evolutionary processes

by

Josef Tkadlec

January, 2020

*A thesis presented to the*
*Graduate School*
*of the*
*Institute of Science and Technology Austria, Klosterneuburg, Austria*
*in partial fulfillment of the requirements*
*for the degree of*
*Doctor of Philosophy*

IST AUSTRIA
*Institute of Science and Technology*

The thesis of Josef Tkadlec, titled *A role of graphs in evolutionary processes*, is approved by:

**Supervisor**: Krishnendu Chatterjee, IST Austria, Klosterneuburg, Austria

Signature: ____________________________

**Committee Member**: Krzysztof Pietrzak, IST Austria, Klosterneuburg, Austria

Signature: ____________________________

**Committee Member**: Martin A. Nowak, Harvard University, Cambridge, US

Signature: ____________________________

**Defense Chair**: Herbert Edelsbrunner, IST Austria, Klosterneuburg, Austria

Signature: ____________________________

signed page is on file

IST Austria Thesis, ISSN: 2663-337X

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: ______________________

Josef Tkadlec
January, 2020
signed page is on file

# Abstract

In this thesis we study certain mathematical aspects of evolution. The two primary forces that drive an evolutionary process are mutation and selection. Mutation generates new variants in a population. Selection chooses among the variants depending on the reproductive rates of individuals. Evolutionary processes are intrinsically random – a new mutation that is initially present in the population at low frequency can go extinct, even if it confers a reproductive advantage. The overall rate of evolution is largely determined by two quantities: the probability that an invading advantageous mutation spreads through the population (called *fixation probability*) and the time until it does so (called *fixation time*). Both those quantities crucially depend not only on the strength of the invading mutation but also on the population structure. In this thesis, we aim to understand how the underlying population structure affects the overall rate of evolution. Specifically, we study population structures that increase the fixation probability of advantageous mutants (called *amplifiers of selection*). Broadly speaking, our results are of three different types: We present various strong amplifiers, we identify regimes under which only limited amplification is feasible, and we propose population structures that provide different tradeoffs between high fixation probability and short fixation time.

# Acknowledgments

First, I would like to thank my collaborators: My advisor Krishnendu Chatterjee, for navigating me through my PhD and for always being available to help; my colleague Andreas Pavlogiannis, for all the whiteboard discussions and the effort he put in our joint papers; and my external supervisor Martin Nowak, for stimulating discussions about research outside math and inner workings of academia. Second, I would like to thank IST for providing a great environment to do research in a field of my choice and for allowing me to learn some wuzzler along the way. Last but not least, I would like to thank my family and friends for continuous unconditional support.

# About the Author

Josef "Pepa" Tkadlec completed his BSc in mathematics at the Charles University in Prague and joined IST in September 2014. His main line of research focuses on understanding discrete-time stochastic processes on graphs but he is also interested in other combinatorial problems, often with geometric or game-theoretical flavor. He has been deeply involved in the world of mathematics competitions for high school students: First participating himself at the International Mathematical Olympiad (IMO), later helping to run competitions in central Europe (Czech-Polish-Slovak match, Náboj), proposing problems for international competitions (e.g. IMO 2012/p5 and IMO 2016/p6), and lecturing at math camps around the world (AwesomeMath in the US, African Math Initiative in Kenya, Maths Beyond Limits in Poland). He likes ultimate frisbee, outdoor puzzle hunts, and musical instruments of any kind.

# List of Publications

Chapters 2, 3, and 4 of this thesis are closely based on the following joint work with A. Pavlogiannis, K. Chatterjee and M. A. Nowak.

1. Pavlogiannis, A., Tkadlec, J., Chatterjee, K., & Nowak, M. A. (2018). Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory. **Communications biology**, 1(1), 71.
2. Tkadlec, J., Pavlogiannis, A., Chatterjee, K., & Nowak, M. A. (2019). Population structure determines the tradeoff between fixation probability and fixation time. **Communications biology**, 2.
3. Tkadlec, J., Pavlogiannis, A., Chatterjee, K., & Nowak, M. A. (2019). Limits on amplifiers of natural selection under death-Birth updating. arXiv preprint arXiv:1906.02785.

# Table of Contents

# 1 Introduction

In this thesis we study certain mathematical aspects of evolution. That is, we study processes that act on populations of reproducing individuals. The two primary forces that drive these evolutionary processes are mutation and selection. Mutation generates new variants in a population. Selection chooses among the variants depending on the reproductive rates of individuals. A reproductive rate is also called fitness. Evolutionary processes are intrinsically random. A new mutation that is initially present in the population at low frequency can go extinct, even if it confers a reproductive advantage.

The key quantities of evolutionary dynamics which affect the overall rate of evolution are then [1; 2; 3; 4; 5]: (a) the mutation rate $\mu$, which is the rate at which new mutations are generated; (b) the fixation probability $\rho$, which is the probability that a new mutation takes over the whole population; and (c) the fixation time $\tau$, which is the time until the mutation fixates in the population, once generated. In the regime of low mutation rate, the population spends most of the time in a homogeneous state in which all individuals are of the same type and the fixation time is no longer so relevant.

A classical evolutionary process for studying the fate of a newly generated mutation is the discrete-time Moran Birth-death process [6]. Given a population of $N$ individuals, at each time step, an individual is chosen for reproduction proportionally to its fitness (reproductive rate). It produces an offspring (a copy of itself) and the offspring replaces a random individual (see Figure 1.1a). Hence the total population size remains constant. In the case of a well-mixed population, each offspring is equally likely to replace any other individual. The fixation probability of a single new mutant with relative fitness $r > 1$ invading a homogeneous population of residents with normalized fitness 1 is equal to $\rho = (1 - 1/r)/(1 - 1/r^N)$. Thus, for large $N$ we have $\rho \approx 1 - 1/r$ [7; 3]. The fixation time is of the order of $N \log N$ steps [8; 9].

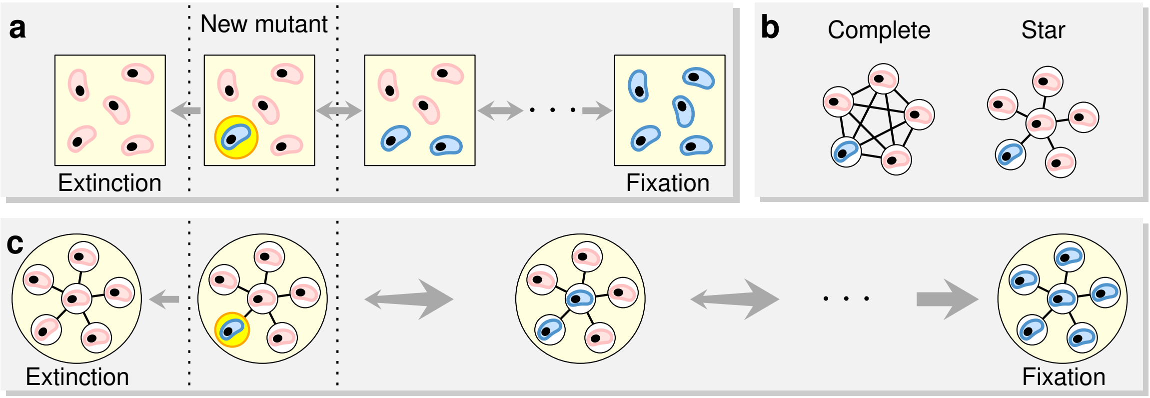

Figure 1.1: **Moran process on graphs. a,** A new mutant (blue) appears in a population of finite size. The lineage of the new mutant can either become extinct or reach fixation. The Moran process is a birth-death process; in any one time step one new offspring is generated and one individual dies. **b,** All fixed spatial structures can be described by graphs. The classical, well-mixed population corresponds to a complete graph. The star graph is a well-studied example of extreme heterogeneity, where one individual, the center, is connected to all others, but each leaf is only connected to the center. **c,** Population structure influences both the fixation probability and the fixation time. An advantageous mutant introduced at a random vertex of a star graph is more likely to fixate than on a complete graph (the arrows pointing to the right are thicker), but the (average) fixation time on the star graph is much longer than on the complete graph (the arrows are longer).

A population structure has profound effects on fixation probability and fixation time of an invading mutant [10; 11; 12; 13; 14; 15; 16; 17; 18]. The mathematical framework for studying these effects is called Evolutionary graph theory. In Evolutionary graph theory, the structure of a population is represented by a graph (network) [7; 19; 20; 21; 22; 23]: each individual occupies a node (vertex) and the edges (links) represent the connections to neighboring sites where a reproducing individual can place an offspring. The edges can have different weights representing different interaction rates between pairs of sites. Moran Birth-death process can be run on any graph. The well-mixed population is given by the complete graph $K_N$ where each individual is connected to each other individual (Figure 1.1b). Graphs can also represent deme structured populations, where islands are represented by complete graphs and connections of different weights exist between islands. Graphs can also represent spatial lattices or asymmetric structures.

A well-studied example is the star graph $S_N$, which has one central vertex and $N-1$

surrounding vertices each connected to the central vertex (Figure 1.1b). For the star graph, the fixation probability tends to approximately $1 - 1/r^2$ for $r > 1$ and large $N$ [7; 24]. Hence, if a mutant has 10 % fitness advantage, which means $r = 1.1$, the star graph amplifies its advantage to 21 %. The fixation time is of the order of $N^2 \log N$ steps [25; 26].

Our goal is to understand how the key quantities of fixation probability and fixation time depend on the structure of the underlying graph. Specifically, we are interested in two types of questions:

1. What are the quantities for some natural fixed graph $G$, or for a sequence of structurally similar graphs of increasing population size $N$? E.g.: What is the fixation probability on a 1D-lattice of size $N = 10$? How does the fixation time on a square grid scale when the size grows large?

2. What is the possible range of values that the quantities attain, over all graphs from some natural class? E.g.: How high can a fixation probability be on a regular graph? How slow are the slowest undirected graphs?

In addressing (1), the ideal outcome is obtaining a concise formula for the quantity in question. If that is not possible, as is often the case, we might still be able to compute the quantity numerically [27], or to approximate it, e.g. using a simulation-based algorithm [28]. In addressing (2), the ideal outcome is typically an asymptotically tight bound obtained via some probabilistic tools [29; 30; 31; 32].

We note that the key quantities are crucially affected not only by the population structure but also by other features, such as by the choice of the underlying stochastc process [33], by the initialization scheme describing where the invading mutant appears [34], and obviously by its fitness $r$. By default, we consider Moran Birth-death process and a mutant with fixed fitness advantage $r > 1$ initialized at a node selected uniformly at random. However, we also discuss different update rules (e.g. Moran death-Birth process), different initialization schemes (e.g. temperature initialization) and mutants that are either neutral ($r = 1$) or disadvantageous ($r < 1$).

Finally, we note that random processes of similar flavor have been independently studied accross various fields, each time representing a different phenomenon [35]. In social science, nodes of different colors can represent people with different opinions who interact along edges of a social network [36]. Similar models of agent-based dynamics have also appeared in statistical physics to study interacting particle systems [37], in epidemiology to study spread of infectious diseases [38, Chapter 7], or in the study of evolution of cooperation [39].

## 1.1 Preliminaries and notation

Here we introduce the key concepts and terminology. Specifically, we give an overview of different types of population structures, different versions of the evolutionary dynamics, and the quantities of interest. For reader's convenience, we summarize the key notation in Table 1.1 on page 7. Moreover, each of the following chapters includes a self-contained description of the precise model and notation.

**Population structure, degree, temperature.** The population structure is represented by a connected graph $G_N$ whose $N$ nodes represent *individuals* and whose edges represent possible interactions. A *well-mixed* population is represented by a complete graph $K_N$.

In general, the edges could be directed (one-way), they could be weighted, and they could include self-loops. Formally, for a pair of nodes $u$, $v$, the weight of an edge $(u, v)$ is denoted by $w_{u,v}$. If the nodes $u$, $v$ are not connected by an edge then we set $w_{u,v} = 0$. In the special case of *undirected* graphs, each edge is two-way and has the same weight in both directions, that is, $w_{u,v} = w_{v,u}$ for all pairs of nodes $u$, $v$. In the special case of *unweighted* graphs, each edge has weight 1. In the special case of *self-loop free* graphs, nodes do not have self-loops, that is, $w_{u,u} = 0$ for each node $u$. In the most general case of directed graphs with weighted edges and self-loops, two nodes $u$, $v$ could be interacting in both directions with different weights $w_{u,v} \neq w_{v,u}$.

Given a node $u$ of a graph, its *outdegree* $\mathsf{Out}(u)$ is the sum $\mathsf{Out}(u) = \sum_v w_{u,v}$ of the weights of the outgoing edges and, similarly, its *indegree* $\mathsf{In}(u)$ is the sum $\mathsf{In}(u) = \sum_v w_{v,u}$ of the weights of the incoming edges. When both notions coincide, which is the case e.g. for undirected graphs, we call this the *degree* of the node. When the graph is unweighted, the degree is simply the number of neighbors of the node.

A temperature of a node is a notion that corresponds to the rate at which the node is typically replaced by its neighbors. Formally, a *temperature* $\mathsf{T}(v)$ of a node $v$ is defined by

$$\mathsf{T}(v) = \sum_{u} \frac{w(u,v)}{\sum_{v' \in N(u)} w(u,v')}.$$

**Evolutionary dynamics, update rules.** Initially, each node is of one of several possible *types* and each type is associated with a non-negative real number called a *fitness*. Typically there are only two types: *residents* with normalized fitness 1 and *mutants* with relative fitness $r$. A *configuration* is a mapping from nodes of $G_N$ into types.

The configuration updates randomly in discrete time-steps according to one of several possible versions of a Moran process. Here we list two of them. For notation purposes, let $f(v)$ be the fitness of the individual at node $v$.

1. Moran Birth-death (Bd) updating: The individuals compete for reproduction and the winner replaces a random neighbor.
   - (Birth) A random individual is selected for reproduction with probability proportional to its fitness. That is, an individual $u$ is selected with probability $f(u)/\sum_v f(v)$.
   - (migration) An outgoing edge from $u$ is selected with probability proportional to its weight. That is, an edge $(u,v)$ is selected with probability $w_{u,v}/\sum_{v'} w_{u,v'}$.
   - (replacement) Node $v$ becomes the same type as node $u$.
2. Moran death-Birth (dB) updating: The individuals randomly die and the neighbors compete to fill the resulting gap.
   - (death) A random individual is selected for death uniformly at random. That is, an individual $v$ is selected with probability $1/N$.
   - (competition) An incoming edge to $v$ is selected with probability proportional to its weight and the fitness of the type at its other end. That is, an edge $(u,v)$ is selected with probability $f(u) \cdot w_{u,v}/(\sum_{u'} f(u') \cdot w_{u',v})$.
   - (replacement) Node $v$ becomes the same type as node $u$.

We note that in both abbreviations the letter “B” is upper case and “d” is lower case to signify that the fitness plays a role in the “Birth” step and not in the “death” step.

There are other versions of Moran process – for example, in Moran Db process, first the death happens with probability inversely proportional to the fitness and then the gap is filled ignoring the fitness of the reproducing individual [40]. It is also possible to combine the update rules. For example, given $\delta \in [0,1]$, in *$\delta$-dB updating* each step independently follows a dB updating with probability $\delta$ and Bd updating with probability $1-\delta$ [33]. In general, different versions of the underlying stochastic process can give rise to qualitatively different behavior [41].

**Initialization schemes and key quantities.** We are interested in the fate of a newly generated mutation appearing in a population of indistinguishable residents. The fate depends on the node at which the mutant appears. Given a (connected) graph $G$, a node $v$ and $r > 0$, we denote by $\mathrm{fp}(G, r, v)$ the fixation probability of a single mutant with relative fitness $r$ initialized at a node $v$ of graph $G$, under Moran Bd process. Typically, we consider that the first mutant appears at a random node according to some probability distribution over nodes. Specifically, we denote by $\mathrm{fp}(G, r) = \frac{1}{N}\sum_v \mathrm{fp}(G, r, v)$ the fixation probability under *uniform initialization* and by $\mathrm{fp}^{\mathsf{T}}(G, r) = \frac{1}{N}\sum_v \mathsf{T}(v) \cdot \mathrm{fp}(G, r, v)$ the fixation probability under *temperature initialization*.

Regarding the duration of the process, we denote by $T(G, r, v)$ the random variable counting the number of steps until the process reaches one of the absorbing states (either all mutants or all residents), starting with a single initial mutant with fitness $r$ at node $v$. Typically we are interested not in the full distribution but only in its expectation $\mathrm{AT}(G, r, v) = \mathbb{E}[T(G, r, v)]$ or even in the (mean, expected) *absorption time* $\mathrm{AT}(G, r) = \frac{1}{N}\sum \mathrm{AT}(G, r, v)$. Similarly, we define the (mean, expected) conditional *fixation time* $\mathrm{CT}(G, r)$ where we average over only those evolutionary trajectories that eventually reach fixation. Unless stated otherwise, we work with uniform initialization and Moran Bd updating.

**Amplifiers and suppressors of selection, superamplifiers.** Given $r > 1$, some population structures $G_N$ increase the fixation probability of advantageous mutants compared to the complete graph $K_N$ whereas others decrease it. This gives rise to a classification of graphs into so-called amplifiers of selection and suppressors of selection.

Formally, given a graph $G_N$ and $r > 1$, we say that $G_N$ is an *$r$-amplifier* if $\mathrm{fp}(G_N, r) > \mathrm{fp}(K_N, r)$, where $K_N$ is a complete graph of the same size representing a well-mixed

| Notation | Description |
|---|---|
| $G$ or $G_N$ | connected graph (network) with $N$ nodes |
| $u, v$ | nodes, each hosting one individual: a mutant or a resident |
| $r$ | relative fitness (reproductive rate) of any mutant (residents have fitness 1) |
| $\mathrm{fp}(G, r, v)$ | fixation probability: probability that in the Moran Birth-death process starting with a single mutant with relative fitness $r$ at a node $v$ of graph $G$, the mutants spread to all nodes of the graph (the mutants "fixate") |
| $\mathrm{fp}(G, r)$ | (average) fixation probability when the initial mutant appears at a node selected uniformly at random |
| $\mathsf{T}(v)$ | temperature of a node $v$: a measure of how frequently the node is replaced by its neighbors when all individuals are residents |
| $\mathrm{fp}^{\mathsf{T}}(G, r)$ | fixation probability when the initial mutant appears at a node selected randomly according to the node's temperature |
| $\mathrm{fp}^{\mathrm{dB}}(G, r)$ | (average) fixation probability under Moran death-Birth process |
| $T(G, r, v)$ | a random variable counting the number of steps until the process reaches either mutant fixation or mutant extinction |
| $\mathrm{AT}(G, r, v)$ | Expectation of $T(G, r, v)$ |
| $\mathrm{AT}(G, r)$ | (unconditional) Absorption time: $\frac{1}{N}\,\mathrm{AT}(G, r, v)$ |
| $\mathrm{CT}(G, r, v)$ | Expectation of $T(G, r, v)$, conditioned on the final state being fixation |
| $\mathrm{CT}(G, r)$ | (conditional) Fixation time: $\frac{1}{N}\,\mathrm{CT}(G, r, v)$ |

Table 1.1: Summary of the notation.

population. When the inequality is reversed, $\mathrm{fp}(G_N, r) < \mathrm{fp}(K_N, r)$, we call $G_N$ an $r$-*suppressor.* If $G$ is an $r$-amplifier for all $r > 1$, we say it is *universal.* (Some authors further require that a universal amplifier has to satisfy $\mathrm{fp}(G_N, r) < \mathrm{fp}(K_N, r)$ for any $r < 1$, that is, the population structure not only favours advantageous mutants but also disfavours disadvantageous mutants.)

We remark that universal amplifiers can be further classified based on their strength [34]. The strongest possible form of amplification is called superamplification. It is meaningful only in the limit of increasing population size. Formally, a sequence of graphs $(G_N)_{N=1}^{\infty}$ of increasing size is called a *superamplifier* (or *arbitrarily strong amplifier*) if it guarantees fixation of any advantageous mutant in the limit of large $N$, that is, $\mathrm{fp}(G_N, r) \to 1$ as $N \to \infty$ for any $r > 1$.

**Notation for asymptotic behavior.** To talk about asymptotic behavior (in the limit of large population size $N$), we use standard mathematical notations $o(\cdot)$, $O(\cdot)$, and $\Theta(\cdot)$ that denote "asymptotically strictly smaller than", "asymptotically less than or equal to", and "asymptotically equal to" (up to a constant factor), respectively. For example, we will write $\frac{1}{N} = o(1)$ (since $\frac{1}{N}$ is much smaller than 1, for large $N$) or $\frac{1}{2}N(N+1) = \Theta(N^2)$. For detailed treatment see [42, Section 1.3].

## 1.2 Related work

Here we give a brief summary of the existing literature, with the focus on results that are directly relevant to the role of population structure in Moran Birth-death process on graphs. A substantial portion of the research done in the field in the past decade has gone into computing the key quantities for fixed graphs and graph families and into identifying population structures that exhibit some sort of extreme behavior for advantageous mutants (when $r > 1$).

### 1.2.1 Computing the key quantities

**Exact formulas.** Moran Bd process on a complete graph is well understood as it can be mapped into a random walk on $[0, N]$ with constant forward bias $r$. The fixation probability of a single mutant is given by $\mathrm{fp}(K_N, r) = (1 - 1/r)/(1 - 1/r^N)$. The formula exhibits a threshold behavior for large $N$. When $r > 1$, this tends to a positive constant $1 - 1/r$; however when $r > 1$ it tends to zero roughly as $1/r^{N-1}$. (When $r = 1$ we clearly get $\mathrm{fp}(K_N, r) = 1/N$ by symmetry.) The absorption time and fixation time are also known. Using a Coupon Collector-like argument we have $\mathrm{AT}(K_N, r) \approx \frac{r+1}{r} \cdot N \log N$ and $\mathrm{CT}(K_N, r) \approx \frac{r+1}{r-1} \cdot N \log N$ when $r > 1$.

In fact, the mapping into a random walk with constant bias shows that the fixation probability is the same for all regular graphs. This is called the Isothermal theorem [7]. However, the time changes: For a Ring graph $R_N$ (aka Cycle, aka 1D lattice) the absorption and fixation time are $\Theta(N^2)$ when $r > 1$ [8] and for the Grid graph (aka 2D square lattice) the correct asymptotics is not known since the waiting time to gain a mutant depends not only on the number of current mutants but also on their relative position in the grid.

Another well understood family of graphs are the Stars $S_N$ consisting of one central node and $N - 1$ leaf nodes, all connected to the center. When $r > 1$ we have $\mathrm{fp}(S_N, r) \to_{N\to\infty} 1 - 1/r^2$ (see [7; 19]) and $\mathrm{AT}(S_N, r) = \Theta(N^2 \log N)$ (see [25; 26]). Using martingales, the fixation probability can be explicitly computed for complete bipartite graphs [43] and for so-called bithermal graphs [44].

**Numerical computation.** The problem of computing the fixation probabilities and absorption or fixation times can be approached numerically. Given two types competing on a graph $G_N$, the resulting dynamical system is a Markov chain with $2^N$ states (since there are $2^N$ possible subsets of nodes that could be occupied by mutants). Thus the problem of computing fixation probabilities, absorption times, or fixation times reduces to solving a system of $2^N$ linear equations [27]. Such approach is feasible for graphs of small population size $N \le 20$ [45; 46; 26] and for graphs that are highly regular since then the size of the system of equations can be greatly reduced (e.g. to $2N$ equations for $S_N$) [47; 48]. We remark that Grid graphs are not susceptible to this approach.

**Approximation algorithms.** Another approach to find the quantities is to use simulation-based approximation algorithms. When the underlying graph is undirected, there exists a FPRAS to compute the quantities [28]. It relies on the fact that the absorption time on an undirected graph is polynomial, hence simulations of the process finish quickly. A subsequent work improved the algorithm by simulating only the steps when the configuration changes, by improving the bound on the absorption time, and by terminating certain runs early [49; 50]. We remark that without further modifications, such a simulation-based approach does not extend to directed graphs since the absorption time on some directed graphs is exponentially long [8].

### 1.2.2 Highest and lowest probability and time

**Probability: Amplifiers of selection.** In terms of probability, the main quest in the field has been a hunt for amplifiers of selection – as simple as possible and as strong as possible.

The first directed superamplifiers, called Superstars and Metafunnels, were proposed in the seminal paper [7], supported by a heuristic argument. Later, it turned out that the argument can not be made formal, even though the structures probably do work [51; 52]. However, directed structures called Megastars, inspired by Superstars, were proved to be superamplifiers. Moreover, as $N \to \infty$, the extinction probability of Megastars decays as $1/N^{1/2}$ which is optimal up to logarithmic factors [53].

All those structures heavily rely on directed edges. Among a more restricted family of undirected graphs, Star graphs have long been the strongest known amplifiers for any fixed $r > 1$ and any fixed $N \geq 2$. However, now we know stronger undirected amplifiers for certain combinations of $r$ and $N$ [47] and even undirected superamplifiers [54; 55]. Moreover, the extinction probability on those superamplifiers decays as $1/N^{1/3}$ which is again optimal up to logarithmic factors.

Although those structures are superamplifiers under uniform initialization, they all cease to work under temperature initialization. In fact, superamplifiers can not exist under temperature initialization unless the underlying graph has both weighted edges and self-loops. Conversely, among graphs that have both weighted edges and self-loops, many superamplifiers exist [48].

**Probability: Suppressors of selection.** Among directed graphs, there are plenty suppressors of selection – for instance, for a directed path $P_N$ on $N$ nodes we have $\mathrm{fp}(P_N, r) = 1/N$ for any $r > 1$, hence the fixation probability of mutants tends to 0 no matter how high their fitness advantage. Among undirected graphs, first $r$-suppressors have been constructed for $r \in (1, 4/3)$ (see [56; 57]). Later, Giakkoupis constructed the first super-suppressors [54] with fixation probability decaying as roughly $\mathcal{O}(1/N^{1/4})$. Current champions achieve roughly $\mathcal{O}(1/N^{1/2})$ (see [50]).

**Time.** As alluded above, for undirected graphs both the absorption and the fixation time can be exponential [8]. For undirected graphs, the absorption time is $\mathcal{O}(N^{3+\varepsilon})$ for any $\varepsilon > 0$ [50] and it is $\Omega(N \log N)$ provided that the graph is not a strong suppressor [26]. In fact, the complete graph $K_N$ is the fastest known graph in terms of both absorption and fixation time among undirected graphs, for any fixed $r > 1$ and $N \geq 2$. However, certain directed graphs do achieve shorter fixation time [26].

**Moran dB updating.** When it comes to dB Moran process, rather than having a fairly complete picture we have a few known results here and there [33]. The Isothermal theorem no longer holds but there is a formula for the fixation probability on the complete graph [41] and on the star graph [19; 40]. An exhaustive search for amplifiers among small undirected unweighted graphs has not found any [58]. The first weighted $r$-amplifiers have been constructed for $r \in (1, \phi)$ where $\phi = \frac{1}{2}(1 + \sqrt{5})$ is the golden ratio [59]. On the negative side, it is known that among self-loop free graphs, no superamplifiers and no universal amplifiers exist under dB updating [60].

## 1.3 Our contribution

Our contribution comes in four parts. Each part is treated separately in a different chapter. Each chapter is self-contained. Here we give a broad overview of the main result presented in each chapter and the idea behind its proof.

### Chapter 2: Weighted loopy superamplifiers [48].

*Result.* We present a dichotomy result on the existence of superamplifiers under uniform and temperature initialization. Our result highlights the importance of weighted edges and self-loops in the underlying graph structure. On the positive side, we show that superamplifiers are abundant among graphs with both weighted edges and self-loops. On the negative side, we show that no temp-superamplifiers exist among graphs that lack either weighted edges or self-loops (or both), and that unif-superamplifiers can exist only when the maximum degree of the graphs grows unbounded.

*Idea behind the proof.* The main idea behind the negative results is to look at the evolutionary trajectories in which the initial mutant dies before reproducing even once. We show that in most scenarios the total mass of those trajectories is non-negligible. The first exceptional scenario is when the graph contains both weighted edges and self-loops. The other one is when it contains vertices of unbounded degree and we operate under uniform initialization.

The main idea behind the positive result is to show that for almost all connected graphs we can assign weights to its edges and self-loops such that we obtain a strong amplifier under both initializations. In more detail, we partition the graph into a central part called a *hub* and many small perimeter parts called *branches* and assign the weights such that, with high probability, the initial mutant appears at a branch, travels to a hub, spreads through the hub, and eventually spreads through all the branches too.

### Chapter 3: Time-probability tradeoff [26].

*Result.* We present two results related to tradeoff between high fixation probability and short fixation time. First, motivated by the exhaustive search through all graphs of small size $N \leq 9$, we present population structures, called *$\alpha$-Balanced bipartite graphs*, that are roughly as good as Stars in terms of fixation probability and close to as good as Complete graphs in terms of fixation time. Second, we show that no amplifiers can have asymptotically shorter absorption time than the Complete graphs.

*Idea behind the proof.* The main idea behind the $\alpha$-Balanced bipartite graphs is to speed up the Star without harming the fixation probability. This is achieved by considering Complete bipartite graphs $K_{N,N^{1-\alpha}}$ for some $\alpha > 0$. The hard part is to argue about the fixation time. We manage to formalize the following intuition: As with the Star, the critical events are those in which a node from a smaller part reproduces on a node in a larger part. On $\alpha$-Balanced bipartite graphs, those critical events occur at a higher rate than they would on a Star (by a factor of roughly $N^{1-\alpha}$) so one would expect the fixation time to be faster by that factor.

### Chapter 4: No strong dB-amplifiers [60].

*Result.* We present a negative result on the existence of strong amplifiers under pure death-Birth updating and $\delta$-death-Birth updating. Namely, we show that any amplifiers under death-Birth updating are necessarily *transient* and *bounded*, and that any amplifiers under $\delta$-death-Birth updating are necessarily *transient* and at most *linear*.

*Idea behind the proof.* The main idea behind the proofs for the (pure) death-Birth updating on a generic graph is to account for those evolutionary trajectories in which the initial mutant dies before reproducing even once. When the accounting is done smartly, Jensen's inequality yields a tight upper bound on the fixation probability in terms of the average degree $d$ of the graph. Transience then follows from the fact that the Complete graph has strictly higher average degree than any other graph (directed or undirected). The method extends to $\delta$-death-Birth updating.

**Chapter 5: Selection reactors.**

*Result.* We present a simple and natural underlying population structure that we call a *Selection reactor*. The edges and self-loops of the Selection reactor can be assigned weights in many ways. For certain weight assignment, we present a simple proof that the result is a superamplifier. For other weight assignment, we prove that the result is a superamplifier with absorption time comparable to that of the Star graph.

*Idea behind the proof.* The main idea behind the proofs is to build on the general positive result presented in Chapter 2 in the special case when the underlying graph structure is given by the Selection reactor. In the first proof, we exploit the fact that on a Selection reactor, the general partitioning into a hub and branches can be done in such a way that every branch consists of a single vertex. In the second proof, we use a weight assignment that makes the hub interact with the branches more often than they would do in the weight assignment used for the general construction. We show that even with this more frequent interaction, the resulting Selection rectors are still superamplifiers and that the higher interaction rate substantially shortens the absorption time.

## 1.4 Relevance and implications of our results

Our negative results show that the phenomenon of superamplification is highly sensitive to seemingly secondary low-level details of the evolutionary process. In Chapter 4 we show that superamplification is impossible under Moran dB updating or under any mixed $\delta$-dB updating. This leaves only (pure) Moran Bd updating. In Chapter 2 we show that superamplification is impossible under adversarial or temperature initialization – though it is possible under uniform initialization. However, in Chapter 2 we also show that even under uniform initialization, superamplification is possible only when the degree of the underlying graphs grows unbounded.

Our positive results all relate to the rate of evolution at different mutation-rate regimes. In the limit of extremely low mutation rate, the overall rate of evolution does not depend on the fixation time but only on the fixation probability. In this regime, the plethora of superamplifiers identified in Chapter 2 are relevant. With higher mutation rate, the fixation time enters the picture and so do the Selection reactors from Chapter 5. With an even higher mutation rate, it is important to optimize the tradeoff between high fixation probability and low fixation time. Different tradeoffs are achieved by the $\alpha$-Balanced bipartite graphs from Chapter 3.

## 1.5 Further directions

We conclude this chapter with an outlook at possible future research. We focus on three broad research directions and list several open questions in each of them.

**Amplification under temperature and adversarial initialization.** The first direction is to investigate whether amplifiers exist under various mutant initialization schemes. Under Bd updating, fixation probability tends to be higher when the initial mutant is placed at a node with low degree [20]. It is then not that surprising that graphs that have many nodes with low degree (such as stars) tend to be amplifiers under uniform initialization. When mutations occur during reproduction events rather than spontaneously, a more relevant notion is that of a temperature initialization. Under temperature initialization, the initial mutant typically appears at a node with high turnover rate and such nodes tend to have high degree so we would generally expect the fixation probability to be lower and amplification to be harder to achieve. Empirically, this seems to be the case and in fact no self-loop free amplifiers under temperature initialization are known.

**Open Problem 1** (Temperature vs. uniform initialization)**.** Does there exist a self-loop free graph $G_N$ (possibly directed and/or weighted) and $r > 1$ such that $\mathrm{fp}^{\mathsf{T}}(G_N, r) > \mathrm{fp}(G_N, r)$?

**Open Problem 2** (Temperature amplifiers)**.** Does there exist a self-loop free graph $G_N$ (possibly directed and/or weighted) and $r > 1$ such that $\mathrm{fp}^{\mathsf{T}}(G_N, r) > \mathrm{fp}(K_N, r)$?

An even stricter condition is to require that a population structure amplifies starting from any initial mutant node [34]. No such structures are known.

**Open Problem 3** (Adversarial amplifiers)**.** Does there exist a graph $G_N$ (possibly directed and/or weighted) and $r > 1$ such that for all nodes $v$ we have $\mathrm{fp}(G_N, r, v) > \mathrm{fp}(K_N, r)$?

A complementary question would be to ask whether there exists a graph $G_N$ such that $\mathrm{fp}(G_N, r, v) < \mathrm{fp}(K_N, r)$ for all nodes $v$. We remark that [54, Lemma 6] implies that the answer to this complementary question is negative for undirected unweighted graphs.

**Extreme points of the time-probability tradeoff.** The second direction is to better understand the extreme points of the time-probability tradeoff. When it comes to short time, Theorem 6 implies that $\mathrm{AT}(A_N, r) = \Omega(N \log N)$ for any amplifier $A_N$. Even though we have $\mathrm{AT}(K_N, r) = \Theta(N \log N)$, the constants do not match. Specifically, the following two questions are open:

**Open Problem 4** (Faster than $K_N$)**.** Does there exist a $G_N$ and $r > 1$ such that $\mathrm{AT}(G_N, r) < \mathrm{AT}(K_N, r)$?

**Open Problem 5** (Asymptotically faster than $K_N$)**.** Does there exist a family $\{G_N\}$ of graphs of increasing size and $r > 1$ such that $\mathrm{AT}(G_N, r) = o(N \log N)$?

Of course, by Theorem 6, such a family would have to be a super-suppressor.

At the other end of the spectrum, there is a gap between the fastest known superamplifiers ($\widetilde{\mathcal{O}}(N^2)$, given by Theorem 15) and the lower bound on time for any superamplifiers ($\Omega(N \log N)$, given by Theorem 6).

**Open Problem 6** (Asymptotically fastest superamplifiers)**.** How fast are the asymptotically fastest superamplifiers?

**Amplification under dB and $\delta$-dB updating.** Finally, the third direction is to investigate the dB updating and the mixed $\delta$-dB updating. Here there are many interesting questions, of which we list three.

First, even though $r$-universal amplifiers are impossible under dB updating, transient amplifiers have recently been constructed [59]. However, their edges have (rather extreme) weights.

**Open Problem 7** (Unweighted dB $r$-amplifiers)**.** Does there exist an unweighted graph $G_N$ and $r > 1$ such that $\mathrm{fp}^{\mathrm{dB}}(G_N, r) > \mathrm{fp}^{\mathrm{dB}}(K_N, r)$?

Second, dB updating generally seems to be less conductive to amplification than Bd updating.

**Open Problem 8** (Monotonicity in $\delta$)**.** Does $\mathrm{fp}^{\mathrm{dB}}(G_N, r) < \mathrm{fp}(G_N, r)$ hold for every graph $G_N$ and every $r > 1$? If so, is $\mathrm{fp}^{\delta}(G_N, r)$ a decreasing function of $\delta$, for any fixed graph $G$ and any fixed $r > 1$?

Third, contrary to Bd updating, there are no superamplifiers under dB updating. Hence the following question is meaningful.

**Open Problem 9** (Strongest $r$-amplifier under dB updating)**.** For fixed $r > 1$, find $M(r) = \sup_G\{\mathrm{fp}^{\mathrm{dB}}(G, r)\}$.

Theorem 11 implies that $M(r) \leq 1 - 1/(r + 1)$ and equality is attained for $r = 1$ and $K_2$. On the other hand, large complete graphs show that, $M(r) \geq 1 - 1/r$ which is relatively tight for $r$ large (see Figure 4.6).

# 2 Weighted & loopy superamplifiers

In general, the fixation probability depends not only on the graph, but also on the initial placement of the invading mutants [61; 34]. The two most natural cases are the following. First, mutation is independent of reproduction and occurs at all locations at a constant rate per unit time. Thus, mutants arise with equal probability in each location. This is called uniform initialization. Second, mutation happens during reproduction. In this case, mutants are more likely to occur in locations that have a higher turnover. This is called temperature initialization. Our approach also allows us to study any combination of the two cases: some mutants arise spontaneously while others occur during reproduction.

For a wide class of population structures [7], which include symmetric ones [62], the fixation probability is the same as for the well-mixed population. A population structure is an *amplifier* if it exaggerates the fitness difference between the invading mutant and the resident when compared to the well-mixed population [7; 34; 41]. A population structure is a *superamplifier* (in this section also called *strong amplifier*) if it ensures a fixation probability arbitrarily close to one for any advantageous mutant, $r > 1$. Superamplifiers can only exist in the limit of large population size.

Numerical studies [58] suggest that for spontaneously arising mutants and small population size, many unweighted graphs amplify for some values of $r$. But for large population size, randomly constructed, unweighted graphs do not amplify [63]. Moreover, proven amplifiers for all values of $r$ are rare. For spontaneously arising mutants (uniform initialization): (i) the Star has fixation probability approximately $1 - 1/r^2$ in the limit of large $N$, and is thus an amplifier [7; 43; 24]; (ii) the Superstar (introduced in [7], see also [52]) and the Incubator (introduced in [55; 54]), which are graphs with unbounded degree, are superamplifiers. The mathematical proofs of these assertions are intricate [53].

For mutants that arise during reproduction (temperature initialization), neither the

Star nor the Superstar amplify [34]. The Star can be modified with self-loops and edge weights to obtain the Looping Star, which has fixation probability $1 - 1/r^2$ in the limit of large $N$ both for mutants that arise during reproduction and for mutants that arise spontaneously. The Looping Star is the only known amplifier for both uniform and temperature initialization [34], but it is not a superamplifier. In fact, no superamplifiers for temperature initialization had been known.

In this work we resolve several open questions regarding superamplification under uniform and temperature initialization. First, we show that there exists a vast variety of graphs with self-loops and weighted edges that are superamplifiers for both uniform and temperature initialization. Moreover, many of those superamplifiers are structurally simple, therefore they might be realizable in natural or laboratory setting. Second, we show that both self-loops and weighted edges are key features of superamplification. Namely, we show that without either self-loops or weighted edges, no graph is a superamplifier under temperature initialization, and no simple graph is a superamplifier under uniform initialization.

## 2.1 Model

**The Moran Process on Weighted Structured Populations.** We consider a population of $n$ individuals on a weighted, directed graph $G_n = (V_n, E_n, W_n)$, where $V_n = \{1, 2, \ldots, n\}$ 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) $i$. 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 $\mathsf{X}_i$ 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 $\mathsf{X}_{i+1}$ 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 $\mathbb{P}^{G_n,r}_u[\cdot]$, 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, $\mathbb{P}^{G_n,r}_u[\mathcal{E}]$ 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 $\mathrm{fp}(G_n, r, u) = \mathbb{P}^{G_n,r}_u[\mathcal{E}]$.

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

**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 (aka "superamplifiers").* 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
$$\liminf_{n \to \infty} \mathrm{fp}(G_n, r_1, Z) = 1 \qquad \text{and} \qquad \limsup_{n \to \infty} \mathrm{fp}(G_n, r_2, Z) = 0 \ ;$$
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.

**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 $\deg(G_n) \leq c$.

**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?

## 2.2 Overview of theoretical results

In this work we present several negative and one positive result 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.

**Discussion 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 [34]. 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 1.

**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.

**Discussion of the positive result.** We highlight some important aspects of the positive result 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 functions such that the resulting family of weighted graphs is a strong amplifier both under temperature and uniform initialization.

2. Second, our positive result is constructive, rather than existential.

   Specifically, given an underlying graph with sublinear diameter we first specify certain subset of vertices that we call a *hub*. The remaining vertices are then split by the hub into a number of so-called *branches*. The construction guarantees that the combined population size of all the branches is much larger than that of the hub. Therefore, with high probability, the first mutant arises on a branch.

   The weights of all edges are then defined so that each of the following steps happens with high probability (see Fig. 2.1). First, the mutants spread on the branch until they reach a vertex that is connected to the hub. Second, the mutants repeatedly invade the hub and eventually fixate there. Third, one by one the mutants spread from the hub to all branches and fixate.

   Intuitively, the weight assignment creates a sense of global flow in the branches, directed towards the hub. This guarantees that the first two steps happen with high probability. For the third step, we show that once the mutants fixate in the hub, they are extremely likely to resist all resident invasion attempts and instead they will invade and take over the branches one by one thereby fixating on the whole graph.

3. Third, we show that with the use of self-loops and weights, even simple graph structures, such as Star graphs and Grids can be turned into strong amplifiers (see Fig. 2.4).

4. Fourth, note that in using weights, edges can be effectively removed by assigning to their weight a small value. However, edges cannot be created. Thus, for complete

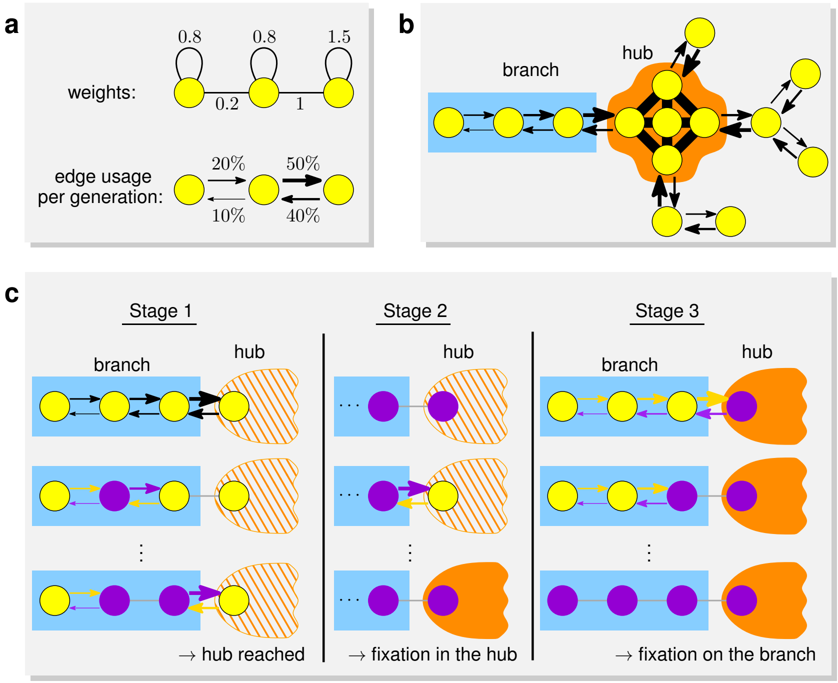

Figure 2.1: **Details of steps to fixation.** **a,** Assigning different weights to edges and self-loops changes the frequency with which each edge is used in each direction. Thicker arrows indicate edges that are used more frequently. **b,** Our weight assignment creates a global sense of flow in the branches, directed towards the hub. The hub itself is almost isothermal and evolves fast. **c,** Three stages to fixation, illustrated on a single branch and the connecting vertex in the hub. After fixating on the hub at the end of Stage 2 (hub becomes dark orange), mutants spread to all the branches and fixate on the whole graph.

> graphs, desired sub-graphs can be created easily using weights. Our positive result states that for almost all graphs, one can use weights to create sub-graphs which are strong amplifiers both under uniform and temperature initialization.

Our results are summarized in Table 2.1.

## 2.3 Proofs of the 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 $G_n$

| | Temperature | | Uniform$^\star$ | |
|---|---|---|---|---|
| | Loops | No Loops | Loops | No Loops |
| Weights | ✓ | × | ✓ | × |
| No Weights | × | × | × | × |

Table 2.1: 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 "×" 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.

is connected and $n \geq 2$. Throughout this section we will use a simple 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 $\mathrm{fp}(G_n, r, u)$ 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

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

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.,

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

*Proof.* Let's focus on the first event that changes the number of mutants. The probability that this event decreases the number of mutants equals $\frac{y}{x+y}$. In such case, the mutants have gone extinct, hence the extinction probability is at least $\frac{y}{x+y}$ and the fixation probability is at most $\frac{x}{x+y}$. $\square$

### 2.3.1 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*

$$\mathrm{fp}(G_n, r, \mathsf{T}) \leq 1 - 1/(r+1).$$

*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

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

Summing over all $u$, we obtain

$$\mathrm{fp}(G_n, r, \mathsf{T}) = \sum_u \frac{\mathsf{T}(u)}{n} \cdot \mathrm{fp}(G_n, r, u) \leq \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \left(1 - \frac{\mathsf{T}(u)}{\mathsf{T}(u)+r}\right) = 1 - \frac{1}{n} \cdot \sum_u \frac{\mathsf{T}(u)^2}{\mathsf{T}(u)+r}\ ; \tag{2.1}$$

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. (2.1) becomes

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

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.

### 2.3.2 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* $\mathrm{fp}(G_n, r, \mathsf{T}) \leq 1 - 1/(4r+2)$.

*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 $\mathsf{T}'$. 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

$$\begin{aligned}
\sum_u \mathsf{T}'(u) =& \sum_u \left( \sum_{v \in \mathsf{In}(u)} \frac{1}{|\mathsf{Out}(v)|} - \mathsf{sl}(u) \frac{1}{|\mathsf{Out}(u)|} \right) \\
=& \sum_u \left( \sum_{v \in \mathsf{In}(u)} \frac{1}{|\mathsf{Out}(v)|} \right) - \sum_{u : \mathsf{sl}(u)=1} \left( \frac{1}{|\mathsf{Out}(u)|} \right) \\
\geq& \sum_u \mathsf{T}(u) - \sum_u \frac{1}{2} = n - \frac{n}{2} = \frac{n}{2}. \qquad (2.2)
\end{aligned}$$

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

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

Summing over all $u$, we obtain

$$\begin{aligned}
\mathrm{fp}(G_n, r, \mathsf{T}) = \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \mathrm{fp}(G_n, r, u) &\leq \frac{1}{n} \cdot \sum_u \mathsf{T}(u) \cdot \left(1 - \frac{\mathsf{T}'(u)}{\mathsf{T}'(u) + r}\right) \\
&\leq 1 - \frac{1}{n} \cdot \sum_u \frac{\mathsf{T}'(u)^2}{\mathsf{T}'(u) + r} \ ; \qquad (2.3)
\end{aligned}$$

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

Using the Cauchy-Schwarz inequality we get

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

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. (2.3) becomes

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

as desired. □

We thus arrive at the following corollary.

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

### 2.3.3 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* $\mathrm{fp}(G_n, r, \mathsf{U}) \leq 1 - 1/(c + r\cdot c^2)$.

*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 $\gamma$) 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 $\frac{1}{|\{w\in\mathsf{In}(v):\ v\in\mathsf{Out}^{\gamma}(w)\}|}$, 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

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

Finally, we have

$$\begin{aligned}\mathrm{fp}(G_n,r,\mathsf{U}) &= \sum_{u\in V^h}\frac{1}{n}\cdot\mathrm{fp}(G_n,r,u) + \sum_{u\in V\setminus V^h}\frac{1}{n}\cdot\mathrm{fp}(G_n,r,u)\\ &\leq \frac{1}{c}\cdot\frac{r\cdot c}{1+r\cdot c}+\frac{c-1}{c}\cdot 1 = 1-\frac{1}{c}\cdot\left(1-\frac{r\cdot c}{1+r\cdot c}\right) = 1-\frac{1}{c+r\cdot c^2}\ .\end{aligned}$$

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.

### 2.3.4 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* $\mathrm{fp}(G_n,r,\mathsf{U})\leq 1-1/(1+r\cdot c)$.

*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

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

Finally, we have

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

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* 1. Theorems 3 and 4 establish the nonexistence of strong amplification with bounded degree graphs. A relevant result can be found in [56], 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 [64].

## 2.4 Proof of the 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 $G_n^{\mathsf{w}}$ is a strong amplifier both for uniform and temperature initialization.

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

### 2.4.1 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 $\mathsf{w}$, 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 $\mathsf{w}$, we denote by $G^{\mathsf{w}}$ 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

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

### 2.4.2 Algorithm for weight assignment on $G$

We start with the construction of the weight function $\mathsf{w}$ on $G$. Since we consider arbitrary input graphs, $\mathsf{w}$ 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 $\mathsf{w}$ such that whp an initial mutant arising under uniform or temperature initialization, eventually fixates on $G^{\mathsf{w}}$. 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 $T_{n_1}^{x_1}, \dots, T_{n_k}^{x_k}$, 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 $|\mathcal{H}| \leq n^{1-\varepsilon+c} \leq n^{1-\gamma}$.
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 $\mathcal{T}_n^x$ (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.

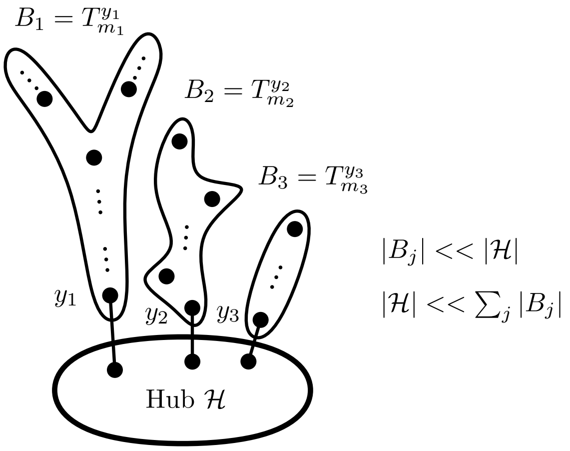

Figure 2.2: Illustration of the hub $\mathcal{H}$ and the branches $T^{y_j}_{m_j}$.

**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 Fig. 2.2).

**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* $\lambda$. For every vertex $u$, we define its *distance* $\lambda(u)$ to be the length of the shortest path $P : u \rightsquigarrow v$ in $T_n^x$ 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* $\mu$ *and* $\nu$. 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

$$\begin{aligned}
\mathsf{w}(u) =& \mathsf{w}(u, u) + \sum_{v \in \mathsf{Nh}(u) \setminus \{u\}} \mathsf{w}(u, v) \\
=& \mu \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \sum_{v \in \mathsf{Nh}(u) \setminus \{u\}} 1 \\
=& \mu \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \mathsf{deg}(u) \\
=& \mu \cdot 2^{-n} + \nu
\end{aligned} \tag{2.4}$$

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

$$\begin{aligned}
\mathsf{w}(u) =& \mathsf{w}(u,u) + \sum_{v \in (\mathsf{Nh}(u) \cap \mathcal{H}) \setminus \{u\}} \mathsf{w}(u,v) + \sum_{v \in \mathsf{chl}(u)} \mathsf{w}(u,v) \\
=& (\mu - |\mathsf{chl}(u)|) \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \sum_{v \in (\mathsf{Nh}(u) \cap \mathcal{H}) \setminus \{u\}} 1 + \sum_{v \in \mathsf{chl}(u)} 2^{-n} \\
=& \mu \cdot 2^{-n} - |\mathsf{chl}(u)| \cdot 2^{-n} + \nu - \mathsf{deg}(u) + \mathsf{deg}(u) + |\mathsf{chl}(u)| \cdot 2^{-n} \\
=& \mu \cdot 2^{-n} + \nu \qquad (2.5)
\end{aligned}$$

□

### 2.4.3 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.

**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 $\eta_1/\eta_2$ 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.

### 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 $\operatorname{diam}(G) \leq n^{1-\varepsilon}$ we have

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

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

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

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

□

### 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 $G^{\mathsf{w}}$ is $\mathsf{X}_j = \{w\}$ with $w \not\in \mathcal{H}$. 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 $\mathsf{X}_t$ 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 $\mathsf{X}_i$ 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 $\mathsf{X}_i$, 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

$$\begin{aligned}\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u)} =&\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u,u)+\mathsf{w}(u,\mathsf{par}(u))+|\mathsf{chl}(u)|\cdot\mathsf{w}(u,x)} = \frac{2^{-n}\cdot n^{-4\cdot(\lambda(u)-1)}}{O(n^{-2\cdot\lambda(u)})}\\ =&\Omega(2^{-n}\cdot n^{-2\cdot(\lambda(u)-2)}),\end{aligned} \tag{2.6}$$

$$\begin{aligned}\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u))} =&\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u),\mathsf{par}(u))+\mathsf{w}(\mathsf{par}(u),\mathsf{par}(\mathsf{par}(u)))+|\mathsf{chl}(\mathsf{par}(u))|\cdot\mathsf{w}(u,\mathsf{par}(u))}\\ =&\frac{2^{-n}\cdot n^{-4\cdot(\lambda(u)-1)}}{\Omega(n^{-2\cdot(\lambda(u)-1)})} = O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)}),\end{aligned} \tag{2.7}$$

$$\begin{aligned}\sum_{z_i\in\mathsf{des}(u)}\frac{\mathsf{w}(\mathsf{par}(z_i),z_i)}{\mathsf{w}(z_i)} =&|\mathsf{des}(u)|\cdot\frac{\mathsf{w}(u,z)}{\mathsf{w}(z,z)+\mathsf{w}(u,z)+|\mathsf{chl}(z)|\cdot\mathsf{w}(z,z')}\\ \le&|\mathsf{des}(u)|\cdot\frac{2^{-n}\cdot n^{-4\cdot\lambda(u)}}{\Omega(n^{-2\cdot(\lambda(u)+1)})} = n\cdot O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)})\\ =&O(2^{-n}\cdot n^{-2\cdot\lambda(u)+3}).\end{aligned} \tag{2.8}$$

Thus, using Eq. (2.6), Eq. (2.7) and Eq. (2.8), we obtain

$$\begin{aligned}\frac{\mathbb{P}[s_{i+1}=s_i+1]}{\mathbb{P}[s_{i+1}=s_i-1]} \ge&\frac{\frac{r}{\mathsf{F}(\mathsf{X}')}\cdot\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(u)}}{\frac{1}{\mathsf{F}(\mathsf{X}')}\cdot\left(\frac{\mathsf{w}(u,\mathsf{par}(u))}{\mathsf{w}(\mathsf{par}(u))}+\sum_{z_i\in\mathsf{des}(u)}\frac{\mathsf{w}(\mathsf{par}(z_i),z_i)}{\mathsf{w}(z_i)}\right)}\\ =&\frac{\Omega(2^{-n}\cdot n^{-2\cdot(\lambda(u)-2)})}{O(2^{-n}\cdot n^{-2\cdot(\lambda(u)-1)})+O(2^{-n}\cdot n^{-2\cdot\lambda(u)+3})} = \Omega(n)\end{aligned} \tag{2.9}$$

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

$$\mathbb{P}[s'_{i+1}=\ell|s'_i] = \begin{cases}\alpha(n) & \text{if } 0<s'_i<|\mathcal{H}| \text{ and } \ell=s'_i+1\\ 1-\alpha(n) & \text{if } 0<s'_i<|\mathcal{H}| \text{ and } \ell=s'_i-1\\ 0 & \text{otherwise}\end{cases} \tag{2.10}$$

Using Eq. (2.9), we have that

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

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., [65], [3, Section 6.3]), as desired. □

### 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 [7].

**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 $G^{\mathsf{w}}$ is $\mathsf{X}_j$.

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 $\mathsf{X}_t$ 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 $\mathsf{X}_t$ 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 $\mathsf{X}_i$, denote by $s_i = |\mathcal{H} \cap \mathsf{X}_i|$. Let $\mathsf{X}_i$ 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 $\mathsf{X}_i$, and $\mathsf{X}_{i+1}$ 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

$$\frac{\mathbb{P}[s_{i+1} = s_i - 1 | u \in \mathcal{H}]}{\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}]} = \frac{1}{r} \,. \tag{2.11}$$

Additionally,

$$\begin{aligned}
\mathbb{P}[s_{i+1} = s_i - 1 | u \notin \mathcal{H}] \leq \sum_{\substack{v \in \mathcal{F} \\ u \in \mathsf{chl}(v)}} \left( \frac{1}{\mathsf{F}(\mathsf{X}_i)} \cdot \frac{\mathsf{w}(u,v)}{\mathsf{w}(u)} \right) \leq n^{-1} \cdot \sum_{\substack{v \in \mathcal{F} \\ u \in \mathsf{chl}(v)}} \frac{2^{-n}}{n^{-2}} \\
\leq n^{-1} \cdot n \cdot 2^{-n} \cdot n^2 = O(n^2 \cdot 2^{-n})
\end{aligned} \tag{2.12}$$

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

$$\begin{aligned}\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}] \cdot \mathbb{P}[u \in \mathcal{H}] \geq \frac{1}{\mathsf{w}(u)} \cdot \frac{r}{\mathsf{F}(\mathsf{X}_i)} \geq \frac{1}{\mu \cdot 2^{-n} + \nu} \cdot n^{-1} \geq \frac{1}{n \cdot 2^{-n} + n} \cdot n^{-1} \\ = \Omega(n^{-2}),\end{aligned} \tag{2.13}$$

since by Lemma 2 we have $\mathsf{w}(v) = \mu \cdot 2^{-n} + \nu$. Using Eq. (2.11), Eq. (2.12) and Eq. (2.13), we have

$$\begin{aligned}\frac{\mathbb{P}[s_{i+1} = s_i - 1]}{\mathbb{P}[s_{i+1} = s_i + 1]} &= \frac{\mathbb{P}[s_{i+1} = s_i - 1 | u \in \mathcal{H}] \cdot \mathbb{P}[u \in \mathcal{H}] + \mathbb{P}[s_{i+1} = s_i - 1 | u \notin \mathcal{H}] \cdot \mathbb{P}[u \notin \mathcal{H}]}{\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}] \cdot \mathbb{P}[u \in \mathcal{H}] + \mathbb{P}[s_{i+1} = s_i + 1 | u \notin \mathcal{H}] \cdot \mathbb{P}[u \notin \mathcal{H}]} \\ &\leq \frac{\mathbb{P}[s_{i+1} = s_i - 1 | u \in \mathcal{H}] \cdot \mathbb{P}[u \in \mathcal{H}] + \mathbb{P}[s_{i+1} = s_i - 1 | u \notin \mathcal{H}] \cdot \mathbb{P}[u \notin \mathcal{H}]}{\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}] \cdot \mathbb{P}[u \in \mathcal{H}]} \\ &\leq \frac{\mathbb{P}[s_{i+1} = s_i - 1 | u \in \mathcal{H}]}{\mathbb{P}[s_{i+1} = s_i + 1 | u \in \mathcal{H}]} + O(n^2) \cdot \mathbb{P}[s_{i+1} = s_i - 1 | u \notin \mathcal{H}] = \frac{1}{r} + 2^{-\Omega(n)}\end{aligned} \tag{2.14}$$

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. (2.14). 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

$$\mathbb{P}[s'_{i+1} = \ell | s'_i] = \begin{cases} \alpha(n) & \text{if } 0 < s'_i < |\mathcal{H}| \text{ and } \ell = s'_i + 1 \\ 1 - \alpha(n) & \text{if } 0 < s'_i < |\mathcal{H}| \text{ and } \ell = s'_i - 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

Using Eq. (2.14) 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 $\rho_1$ (resp. $\rho_2$) be the probability that the Moran process starting on configuration $\mathsf{X}_j$ with $|\mathcal{H} \cap \mathsf{X}_j| \geq 1$ (resp. $|\mathcal{H} \setminus \mathsf{X}_j| = 1$) will reach a configuration $\mathsf{X}_t$ with $\mathcal{H} \subseteq \mathsf{X}_t$. We have that $\rho_1$ (resp. $\rho_2$) is lowerbounded by the probability that $\rho$ 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., [65], [3, 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 $G^{\mathsf{w}}$ is $\mathsf{X}_j$ 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 $\mathsf{X}_i$ occurring after $\mathsf{X}_j$, let

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

and let $\rho_A$ and $\rho_B$ be the corresponding probabilities. Using Eq. (2.6), we have

$$\rho_A = \frac{r}{\mathsf{F}(\mathsf{X}_i)} \cdot \frac{\mathsf{w}(v, \mathsf{par}(v))}{\mathsf{w}(v)} = \Omega\left(n \cdot 2^{-n}\right) \ ; \tag{2.16}$$

and using Eq. (2.7) and Eq. (2.8)

$$\rho_B \leq \frac{r}{\mathsf{F}(\mathsf{X}_i)} \cdot \left( \frac{\mathsf{w}(v, \mathsf{par}(v))}{\mathsf{w}(\mathsf{par}(u))} + \sum_{z \in \mathsf{chl}(v)} \frac{\mathsf{w}(v,z)}{\mathsf{w}(z)} \right) \leq \frac{r}{n} \cdot \left(2^{-n} + O\left(n \cdot 2^{-n}\right)\right) = 2^{-\Omega(n)} \ . \tag{2.17}$$

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. (2.16)). The expected value of $X'$ is

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

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

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

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. (2.17)). Then

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

and thus

$$\mathbb{P}[Y \leq X] \leq \mathbb{P}[Y \leq \alpha] + \mathbb{P}[X \geq \alpha] \leq \mathbb{P}[Y' \leq \alpha] + \mathbb{P}[X' \geq \alpha] = O(n^{-1/3}) \ . \tag{2.18}$$

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 $G^{\mathsf{w}}$ is $\mathsf{X}_j$ 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 $\mathsf{X}_t$ 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 $\mathsf{X}_t$ 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

$$\begin{aligned}\mathbb{P}[\overline{A}_1 \cap \overline{A}_2 \cdots \cap \overline{A}_{n^{1/3}}] &\leq p \cdot \prod_{i=1}^{n^{1/3}} \mathbb{P}[\overline{A}_i] + (1-p) \leq \prod_{i=1}^{n^{1/3}} (1 - \mathbb{P}[A_i]) + 2^{-\Omega(n)} \\ &\leq \left(r^{-1} + 2^{-\Omega(n)}\right)^{n^{1/3}} + 2^{-\Omega(n)}.\end{aligned} \tag{2.19}$$

Finally, starting from $\mathsf{X}_0 = \{u\}$, the probability that a configuration $\mathsf{X}_t$ 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. (2.18) and Eq. (2.19), we obtain that the goal configuration $\mathsf{X}_t$ 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. □

**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 $T^{y_j}_{m_j}$ 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 $\mathsf{X}_j$ 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 $\mathsf{X}_i$, 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. (2.14). 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 $O(n)$ [29], 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

$$\mathbb{E}[t'_1 | t'_1 < \infty] = H_f = O(n^3) \ ;$$

i.e., given that a configuration $\mathsf{X}_{t_1}$ 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

$$\mathbb{P}[t'_1 > \alpha | t'_1 < \infty] \leq \frac{\mathbb{E}[t'_1 | t'_1 < \infty]}{\alpha} = n^3 \cdot 2^{-\frac{n}{2}} \ . \tag{2.20}$$

Consider any configuration $\mathsf{X}_i$. 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. (2.20). $\square$

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 $G^{\mathsf{w}}$, and assume there exists a first time $t^* < \infty$ when a configuration $\mathsf{X}_{t^*}$ is reached such that $\mathcal{H} \subseteq \mathsf{X}_{t^*}$. We underapproximate the fixation probability of the Moran process starting from $\mathsf{X}_{t^*}$ 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 $T_{m_j}^{y_j}$ the subtree of $\mathcal{T}_n^x$ rooted at $y_j$, which has $m_j$ vertices. Let $V_i$ be the set of vertices of $T_{m_i}^{y_i}$, 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 $T_{m_j}^{y_j}$, 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 $T_{m_j}^{y_j}$ 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 $\mathsf{X}_i$ and $\overline{\mathsf{X}}_i$ respectively, with $\overline{\mathsf{X}}_i \subseteq \mathsf{X}_i$. Thus, the fixation probability of the Moran process on $G^{\mathsf{w}}_n$ 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 $G^{\mathsf{w}}$, 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 Fig. 2.3 for an illustration. The Markov chain $\mathcal{M}_j$ has two absorbing states, $\mathcal{D}$ and $m_j$. We denote by $\rho_j$ 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 $\rho_j$, and comes from a straightforward analysis of $\mathcal{M}_j$.

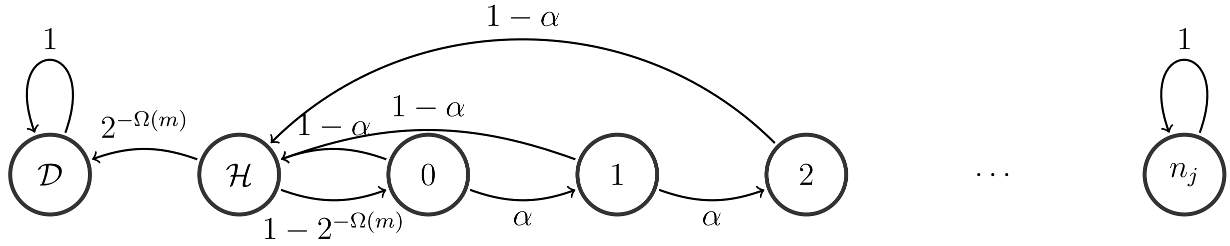

Figure 2.3: The Markov chain $\mathcal{M}_j$ given a tree $T^{x_j}_{n_j}$.

**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

$$
\begin{aligned}
x_{\mathcal{H}} &= \left(1 - 2^{-\Omega(n^{1-\gamma})}\right) \cdot \left(x_{\mathcal{H}} \cdot (1-\alpha) \cdot \sum_{0=1}^{m_j} a^i + a^{m_j}\right) \\
\Longrightarrow x_{\mathcal{H}} &= \left(1 - 2^{-\Omega(n^{1-\gamma})}\right) \cdot \left(x_{\mathcal{H}} \cdot \left(1 - a^{m_j - 1}\right) + a^{m_j}\right) \\
\Longrightarrow x_{\mathcal{H}} & \left(1 - \left(1 - 2^{-\Omega(n^{1-\gamma})}\right) \cdot \left(1 - a^{m_j - 1}\right)\right) = a^{m_j} \qquad (2.21)
\end{aligned}
$$

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. (2.21) we obtain

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

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 $\rho_j$. 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 $\mathsf{X}$ 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 $\mathsf{X}$ 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 $G^{\mathsf{w}}$. The proof is based on repeated applications of Lemma 11 and Lemma 10, one for each subtree $T^{y_j}_{m_j}$.

**Lemma 12.** Consider that at some time $t^*$ the configuration of the Moran process on $G^{\mathsf{w}}$ is $\mathsf{X}_{t^*}$ with $\mathcal{H} \subseteq \mathsf{X}_{t^*}$. Then, a subsequent configuration $\mathsf{X}_t$ 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. □

### 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 $\mathrm{diam}(G_n) \leq n^{1-\varepsilon}$, there exists a weight function $\mathsf{w}$ such that for all $r > 1$, we have $\mathrm{fp}(G_n^{\mathsf{w}}, r, \mathsf{U}) = 1 - O(n^{-\varepsilon/3})$ and $\mathrm{fp}(G_n^{\mathsf{w}}, r, \mathsf{T}) = 1 - O(n^{-\varepsilon/3})$.

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 $\mathsf{w}$, for all $r \geq 1$, we have that $\mathrm{fp}(G_n^{\mathsf{w}}, 1/r, \mathsf{U}) \leq 1 - \mathrm{fp}(G_n^{\mathsf{w}}, r, \mathsf{U})$.

*Proof.* Let $\sigma$ 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

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

□

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 $\mathsf{w}$ 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.*

## 2.5 Numerical and simulation results

Here we present related numerical and simulation results. Specifically, we illustrate that even simple graph structures, such as Star graphs and Grids can be turned into strong amplifiers.

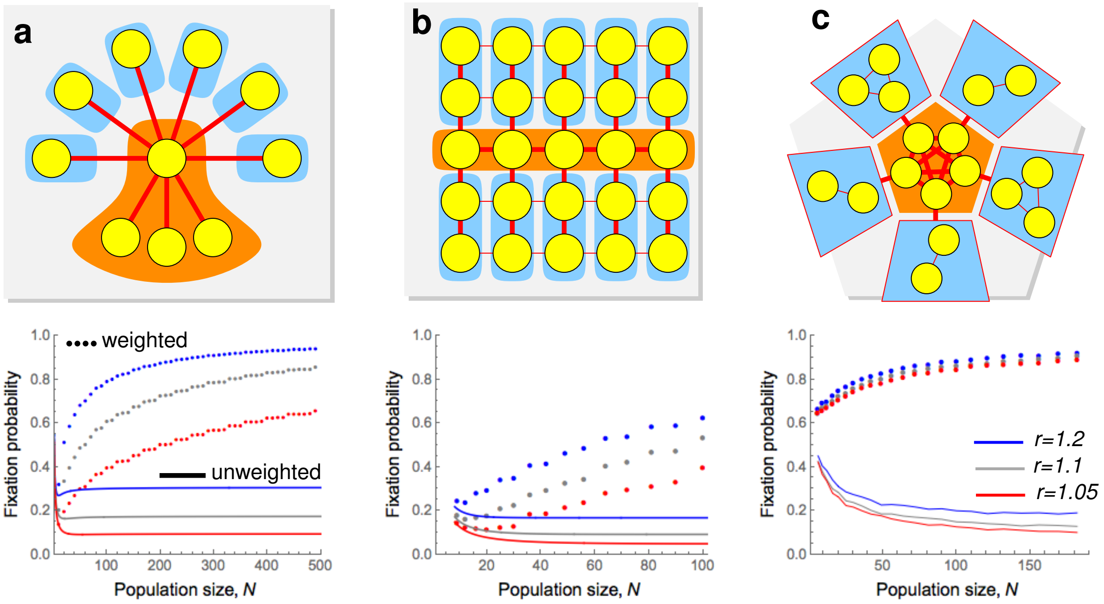

Figure 2.4: **Almost any topology can be turned into a strong amplifier.** We illustrate our positive result on the topology of a Star (**a**), a Grid (**b**), and a Sunflower (**c**). The hub is shown in orange, the branches in blue. Thin edges are assigned negligibly small (or zero) weights. For each graph, we compare the fixation probability of the unweighted version (lines) and the weighted version (dots) as a function of the population size, $N$, For unweighted graphs, we plot the maximum of the fixation probabilities for uniform and temperature initialization. For weighted graphs we plot the minimum of the two. A Sunflower graph consists of a well-mixed population of size $n$ in the center surrounded by $n$ petals which are local well-mixed populations of size $n-1$ or $n-2$ each. Each petal is connected to a unique vertex of the center. Our construction assigns negligibly small weights to edges within petals. The values plotted are obtained by simulating the process 10 000 times.

# 3 Time-probability tradeoff

Several population structures have been identified that alter the fixation probability of advantageous mutants. Structures that decrease the fixation probability are known as *suppressors of selection* and those that increase it are known as *amplifiers of selection* [7; 34; 58; 47]. However, amplification is usually achieved at the cost of increasing fixation time compared to the well-mixed population [13; 17; 28; 66]. For example, the star graph has higher fixation probability but also longer fixation time as compared to the well-mixed population. There also exist *superamplifiers* (also known as arbitrarily strong amplifiers of natural selection) that guarantee fixation of advantageous mutants in the limit of large population size [53; 54; 55; 48]. But those structures tend to require even longer fixation times.

We can refer to population structures that decrease the fixation time with respect to the well-mixed population as *accelerators.* Both the fixation probability and the fixation time play an important role in the speed of evolution. Ideally, we prefer a population structure that is both an amplifier and an accelerator, but all known amplifiers achieve amplification at the cost of deceleration. In fact, this slowdown can be so prominent that it outweighs the amplification and leads to longer evolutionary timescales [17].

Here we show that absorption time on any amplifier is asymptotically at least as large as both the absorption and the fixation time on the well-mixed population. Given this negative result, we proceed to study the tradeoff between fixation probability and time more closely. We have computed fixation probabilities and fixation times for a large class of graphs. While within this class, the well-mixed population is optimal with respect to fixation time, and the star graph is favorable with respect to fixation probability, there is a very interesting tradeoff curve between fixation probability and fixation time. In other words, there exist population structures which provide different tradeoffs between high fixation probability and short fixation time. As our main analytical results, we present population structures that asymptotically achieve fixation probability equal to that of star graphs and fixation time similar to that of well-mixed populations. Thus, we achieve amplification with negligible deceleration. Finally, while the above analytical results are established for large population sizes, we also study evolutionary processes on population structures of small or intermediate size by numerical simulation. Specifically, we consider the effective rate of evolution as proposed by Frean, Rainey, and Traulsen [17]. Generally speaking, the well-mixed population has a high effective rate of evolution if the mutation rate is high, while the star graph has a high effective rate of evolution if the mutation rate is very low. We show that for a wide range of intermediate mutation rates, our new structures achieve higher effective rate of evolution than both the well-mixed population and the star graph.

## 3.1 Model

**Moran process on graphs.** Moran Birth-death process is a discrete-time stochastic (random) process that models evolutionary dynamics in a spatially structured population. The population structure is represented by a connected graph $G$, possibly with weighted edges and/or self-loops. At all times, each vertex of the graph is occupied by a single individual that is of one of two types: either a *resident* or a *mutant*. The individuals of one type are considered indistinguishable. Moreover, residents are assigned (normalized) fitness 1 while the mutants have fitness $r$. Here we consider advantageous mutants ($r > 1$). In one step of the process, an individual is selected for reproduction randomly and proportionally to its fitness. This individual produces an offspring that is a copy of itself. This offspring then selects one of the adjacent edges proportionally to the edge weight and travels along that edge to replace the individual at its other endpoint. (If the selected edge happened to be a self-loop then the offspring replaces the parent and nothing changes.) These steps continue until the population becomes homogeneous: either all individuals are mutants (*fixation* occurred) or they are all residents (*extinction* occurred). The well-mixed population is modelled by an unweighted complete graph (without self-loops).

**Initialization scheme.** We study the situation of a single mutant invading a population of residents. This initial mutant can appear either spontaneously or during reproduction.

In the first case, called *uniform initialization*, the mutant is placed at a vertex chosen uniformly at random. In the second case, called *temperature initialization*, we perform one step of the Moran process in a population that consists entirely of residents and place the mutant at the vertex that the offspring migrates to. Formally, the mutant is placed at a random vertex, proportionally to the temperature (or turnover rate) of that vertex. Here temperature $\mathsf{T}(v)$ of a vertex $v$ is defined by

$$\mathsf{T}(v) = \sum_{u \in N(v)} \frac{w(u,v)}{\sum_{v' \in N(u)} w(u,v')},$$

where $w(u,v)$ is the weight of edge between $u$ and $v$ and $N(v)$ is the set of *neighbors* of $v$, that is vertices connected to $v$ by an edge.

**Fixation probability and time.** Given a graph $G$ and relative fitness advantage $r$, let $\mathrm{fp}(G,r)$ be the fixation probability of a single mutant under uniform initialization, and let $\mathrm{fp}_{\mathsf{T}}(G,r)$ stand for temperature initialization. Similarly, we define $\mathrm{CT}(G,r)$ (or $\mathrm{CT}_{\mathsf{T}}(G,r)$) to be the fixation time, that is the expected number of steps of the Moran process until the mutants reach fixation (conditioning on them doing so). Likewise we define $\mathrm{ET}(G,r)$ (or $\mathrm{ET}_{\mathsf{T}}(G,r)$) to be the extinction time and $\mathrm{AT}(G,r)$ (or $\mathrm{AT}_{\mathsf{T}}(G,r)$) to be the (unconditional) absorption time.

**Amplifiers and superamplifiers.** A graph $G_N$ with $N$ vertices is called an *amplifier* if it increases the fixation probability of any advantageous mutant, as compared to the Complete graph (that is, $\mathrm{fp}(G_N,r) > \mathrm{fp}(K_N,r)$ for any $r > 1$). A sequence of graphs $(G_N)_{N=1}^{\infty}$ of increasing size is called a *superamplifier* (or *arbiterarily strong amplifier*) if, in the limit of large $N$, it guarantees the fixation of any advantageous mutant (that is, $\mathrm{fp}(G_N,r) \to 1$ as $N \to \infty$).

**Graphs.** We introduce and study the following graphs.

**Complete graph.** Complete graph $K_N$ on $N$ vertices models a well-mixed population. This case is well understood. In particular, the fixation probability satisfies

$$\mathrm{fp}(K_N,r) = \mathrm{fp}_{\mathsf{T}}(K_N,r) = \frac{1-1/r}{1-1/r^N} \to 1-1/r$$

for $r > 1$ as $N \to \infty$ and the (unconditional) absorption time is of the order of $\Theta(N \log N)$ (see [8]). In fact, using a standard difference method one can derive that, for $r > 1$, we have $\mathrm{AT}(K_N, r) \approx \frac{r+1}{r} \cdot N \log N$ and $\mathrm{CT}(K_N, r) \approx \frac{r+1}{r-1} \cdot N \log N$. For reference purposes we present those proofs in Section 3.3.4.

**Star graph.** Star graph $S_N$ consists of one central vertex connected to each of the remaining $N-1$ vertices on the periphery. For large $N$, it is known that $\mathrm{fp}(S_N, r) \to 1 - 1/r^2$ and that the absorption and fixation time are of the order of at most $O(N^2 \log N)$ and $O(N^3)$, respectively [40]. In fact, as a corollary of our results on $\varepsilon$-Balanced bipartite graph, we show that both the absorption time and the fixation time are of the order of $\Theta(N^2 \log N)$. The bottom line is that, under uniform initialization, the Star graph amplifies the fixation probability but at the cost of substantially increasing the fixation time.

**$\varepsilon$-Balanced bipartite graph.** For uniform initialization we present a family of graphs that, in the limit of large population size, achieve the fixation probability of the Star graph and the fixation time almost as good as the Complete graph. The graphs are complete bipartite graphs with both parts large but one part asymptotically larger than the other one. Formally, given $N$ and $\varepsilon \in (0, 1]$, the $\varepsilon$-Balanced bipartite graph $B_{N,\varepsilon}$ is a complete bipartite graph with parts of size $N^{1-\varepsilon}$ and $N$. That is, there are $N^{1-\varepsilon}$ vertices in one part, $N$ vertices in the other part, and all edges that connect vertices in different parts. The case $\varepsilon = 1$ corresponds to a Star graph.

**Weighted bipartite graphs.** For temperature initialization, the Star graph and the $\varepsilon$-Balanced bipartite graphs fail to amplify. We present another family of weighted graphs with self-loops that, in the limit of large population size, provide fixation probability $1 - 1/r^2$ (the same as Star graph under uniform initialization) and the fixation time almost as good as the Complete graph. The graphs are obtained by adding self-loops of relatively large weight to all vertices in the larger part of an $\varepsilon$-Balanced bipartite graph. Formally, given $N$ and $\varepsilon \in (0, 1)$, the Weighted bipartite graph $W_{N,\varepsilon}$ is a complete bipartite graphs with one (smaller) part of size $N^{1-\varepsilon}$, one (larger) part of size $N$, and every vertex of the larger part having a self-loop of such a weight $w$ that $N^{-\varepsilon/2} = \frac{N^{1-\varepsilon}}{w+N^{1-\varepsilon}}$. The case $\varepsilon = 1$ is closely related to a Looping Star [34].

## 3.2 Overview of theoretical results

Here we present the summary of our theoretical results and discuss their significance. They are all related to the tradeoff between fixation probability and fixation time, under both uniform and temperature initialization.

First, we prove that no amplifier is asymptotically faster than the Complete graph in terms of absorption time (recall that $\mathrm{AT}(K_N, r) = \Theta(N \log N)$, see Section 3.3.4).

**Theorem 6.** *Fix $r > 1$. Let $G$ be any graph with $N \geq 2$ vertices and let $p = \mathrm{fp}(G, r)$ be the fixation probability of a single mutant under uniform initialization. Then*

$$\mathrm{AT}(G, r) \geq \frac{p}{r} \cdot N \cdot H_{N-1},$$

*where $H_{N-1} = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{N-1} \geq \log N$. In particular, $\mathrm{AT}(G, r) \geq \frac{p}{r} \cdot N \log N$ for an arbitrary graph $G$ and $\mathrm{AT}(A, r) \geq \frac{r-1}{r^2} \cdot N \log N$ for an arbitrary amplifier $A$.*

Second, we give tight results for the fixation time on Bipartite graphs. In particular, we prove that under uniform initialization, certain $\varepsilon$-Balanced bipartite graphs $B_{N,\varepsilon}$ asymptotically achieve the fixation probability of the Star graph and the fixation time almost as good as the Complete graph. The analysis of fixation probability is relatively straightforward. For fixation time, we provide tight lower and upper bounds. We first present the lower bound that is proportional to $N^{1+\varepsilon} \log N$. For the upper bound we then distinguish two cases: If the size of the smaller part is small, that is $N^{1-\varepsilon} = o(\sqrt{N})$, then the argument is simpler and we get a matching upper bound. If the size of the smaller part is relatively close to $N$, the upper bound has an additional factor of $N^{\varepsilon}$. As a consequence, we can prove the following theorem.

**Theorem 7.** *Fix $\varepsilon \in (0, 1]$ and $r > 1$. Let $B_{N,\varepsilon}$ be the $\varepsilon$-Balanced bipartite graph. Then*

- $\mathrm{fp}(B_{N,\varepsilon}, r) \to 1 - 1/r^2$.
    - *(small center) If $\varepsilon \in (0.5, 1)$ then there exist constants $c_1$, $c_2$ such that*

    $$c_1 \cdot N^{1+\varepsilon} \log N \leq \mathrm{AT}(B_{N,\varepsilon}, r) \leq c_2 \cdot N^{1+\varepsilon} \log N.$$

    - *(large center) If $\varepsilon \in (0, 0.5)$ then there exist constants $c_1$, $c_2$ such that*

    $$c_1 \cdot N^{1+\varepsilon} \log N \leq \mathrm{AT}(B_{N,\varepsilon}, r) \leq c_2 \cdot N^{1+2\varepsilon} \log N.$$

*Moreover, the fixation time* $\mathrm{CT}(B_{N,\varepsilon}, r)$ *satisfies the same inequalities.*

As an immediate corollary, we obtain that for any fixed $r > 1$, both the absorption and the fixation time on a Star graph ($\alpha = 1$) are of the order of $\Theta(N^2 \log N)$. This is in alignment with earlier results [40; 25].

Third, we prove that under temperature initialization, analogous results can be achieved using Weighted bipartite graphs $W_{N,\varepsilon}$.

**Theorem 8.** *Fix* $\varepsilon \in (0, 1]$ *and* $r > 1$. *Let* $W_{N,\varepsilon}$ *be the Weighted bipartite graph. Then*

- $\mathrm{fp}(W_{N,\varepsilon}, r) \to 1 - 1/r^2$.
- *There exist constants* $c_1$, $c_2$ *such that*

$$c_1 \cdot N^{1+\varepsilon} \log N \le \mathrm{AT}(B_{N,\varepsilon}, r) \le c_2 \cdot N^{1+\frac{3}{2}\varepsilon} \log N.$$

*Moreover, the fixation time* $\mathrm{CT}(W_{N,\varepsilon}, r)$ *satisfies the same inequalities.*

We note that the same upper bounds on expected time in Theorems 7 and 8 apply regardless of the initial configuration of mutants and residents. Thus we also get strong concentration results: For example, splitting the time into $\log N$ stages of length $3 \cdot \mathrm{AT}(G, r)$ each, a repeated application of Markov inequality immediately implies that the absorption time is less than $3 \cdot \mathrm{AT}(G, r) \cdot \log N$ with probability at least $1 - 1/3^{\log N} \ge 1 - 1/N$.

Finally, for reference purposes we compute the absorption, fixation, and extinction times of a single advantageous mutant ($r > 1$) on a Complete graph, using the standard difference method. We remark that for the absorption time we also identify the non-leading terms, up to an error of the order of $o(1)$.

**Theorem 9.** *Fix* $r > 1$ *and let* $K_N$ *be the Complete graph on* $N$ *vertices. Then*

$$\begin{aligned}
\mathrm{AT}(K_N, r) &= (N-1)H_{N-1} \cdot \frac{r+1}{r} + (N-1) \cdot \log(1 - 1/r) - \frac{1}{r(r-1)} + o(1), \\
\mathrm{CT}(K_N, r) &= (N-1)H_{N-1} \cdot \frac{r+1}{r-1} + (N-1) \cdot \frac{r+1}{r-1} \log(1 - 1/r) + o(N), \\
\mathrm{ET}(K_N, r) &= (N-1) \cdot \log\left(\frac{r}{r-1}\right) + o(N).
\end{aligned}$$

*In particular for* $r = 1 + s$, $s > 0$ *small, we have* $\mathrm{AT}(K_N, r) \approx 2 \cdot N \log N$, $\mathrm{CT}(K_N, r) \approx \frac{2}{s} \cdot N \log N$, *and* $\mathrm{ET}(K_N, r) \approx \frac{1}{s} \cdot N$.

## 3.3 Proofs

### 3.3.1 Lower bound on absorption time

Here we show that for $r > 1$ no family of graphs with fixation probability bounded away from zero can have asymptotically smaller absorption time than the Complete graphs. Specifically, no amplifiers can absorb asymptotically faster than the Complete graphs. Recall that for the Complete graph on $N$ vertices, both the fixation time and the absorption time is of the order of $\Theta(N \log N)$ (see Section 3.3.4).

**Theorem 6.** *Fix $r > 1$. Let $G$ be any graph with $N \geq 2$ vertices and let $p = \mathrm{fp}(G, r)$ be the fixation probability of a single mutant under uniform initialization. Then*

$$\mathrm{AT}(G, r) \geq \frac{p}{r} \cdot N \cdot H_{N-1},$$

*where $H_{N-1} = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{N-1} \geq \log N$. In particular, $\mathrm{AT}(G, r) \geq \frac{p}{r} \cdot N \log N$ for an arbitrary graph $G$ and $\mathrm{AT}(A, r) \geq \frac{r-1}{r^2} \cdot N \log N$ for an arbitrary amplifier $A$.*

*Proof.* Consider a modified Moran process $M'$ that is identical with the standard Moran process, except that if the mutation goes extinct then in the next step we again initialize a single mutant uniformly at random and continue the process. Clearly, the modified process $M'$ always terminates with the mutants fixating and its expected fixation time is given by $T'(G, r) = \frac{1}{p} \cdot \mathrm{AT}(G, r)$.

Given any subset $X$ of the vertices, let $p^X$ be the probability to gain a mutant in a single step from a configuration consisting of mutants at vertices of $X$ and residents elsewhere. To gain a mutant, one of the $|X|$ mutants has to be selected for reproduction and then the offspring has to replace a resident. The probability of the first event alone equals $\frac{r|X|}{N+(r-1)|X|}$, hence we get an upper bound

$$p^X \leq \frac{r|X|}{N + (r-1)|X|} \leq \frac{r|X|}{N} \equiv p_{|X|}$$

that doesn't depend on $X$ but only on $|X|$.

Finally, fix $k \in \{1, 2, \ldots, N-1\}$ and observe that any evolutionary trajectory in $M'$ has to, at some point, reach a state with $k$ mutants and gain another mutant from there. Hence, in expectation, the evolutionary trace spends at least $\frac{1}{p_k}$ steps in states corresponding to configurations with $k$ mutants. By linearity of expectation, summing over $k$ gives

$$\mathrm{AT}(G,r) = p \cdot T'(G,r) \geq p \sum_{k=1}^{N-1} \frac{N}{r \cdot k} = \frac{p}{r} \cdot N \cdot H_{N-1}$$

as desired. □

**Remarks on the lower bound.** Several remarks are in order.

1. First, we emphasize that the proof applies to all graphs, possibly containing directed edges, weighted edges, and/or self-loops.

2. Second, we note that the same proof goes through for any initialization scheme $\mathcal{S}$ (with $p = \mathrm{fp}(G,r)$ replaced by the fixation probability $p^{\mathcal{S}}$ under that initialization scheme $\mathcal{S}$). Specifically, it applies to temperature initialization and also to schemes in which the first mutant is initialized to a fixed vertex.

3. Third, we discuss relationship between absorption time and fixation time. Note that Theorem 6 provides a lower bound on the absorption time $\mathrm{AT}(G,r)$ which is a weighted average of the fixation time $\mathrm{CT}(G,r)$ and the extinction time $\mathrm{ET}(G,r)$. Since the evolutionary trajectories leading to extinction are typically shorter than those leading to fixation, the fixation time tends to be even longer than the absorption time. In fact, the inequality $\mathrm{CT}(G,r) > \mathrm{AT}(G,r)$ holds for all undirected graphs $G$ and all values $r > 1$ that we tested. On the other hand, there do exist directed graphs for which the opposite inequality $\mathrm{CT}(G,r) < \mathrm{AT}(G,r)$ holds. As an example, consider $r = 4$ and a graph $G$ consisting of three vertices $\{u, v_1, v_2\}$ and edges $\{u \to v_1, u \to v_2, v_1 \leftrightarrow v_2\}$. Then we easily check that $\mathrm{CT}(G,4) = 3.25$ while $\mathrm{AT}(G,4) = 19.25$. In fact, in terms of fixation time, this graph $G$ is even slightly faster than the complete graph $K_3$, as we have $\mathrm{CT}(K_3,4) = 3 + \frac{4}{7} > 3.25$.

### 3.3.2 $\varepsilon$-Balanced bipartite graphs

In this subsection we analyze the $\varepsilon$-Balanced bipartite graph $B_{N,\varepsilon}$. Recall that $B_{N,\varepsilon}$ consists of $c = N^{1-\varepsilon}$ vertices in the (smaller) center and $N$ vertices in the outside part, each two vertices from different parts connected by an edge.

We prove the following theorem.

**Theorem 7.** *Fix $\varepsilon \in (0,1]$ and $r > 1$. Let $B_{N,\varepsilon}$ be the $\varepsilon$-Balanced bipartite graph. Then*

- $\mathrm{fp}(B_{N,\varepsilon}, r) \to 1 - 1/r^2$.
    - *(small center) If $\varepsilon \in (0.5, 1)$ then there exist constants $c_1$, $c_2$ such that*

$$c_1 \cdot N^{1+\varepsilon} \log N \leq \mathrm{AT}(B_{N,\varepsilon}, r) \leq c_2 \cdot N^{1+\varepsilon} \log N.$$

    - *(large center) If $\varepsilon \in (0, 0.5)$ then there exist constants $c_1$, $c_2$ such that*

$$c_1 \cdot N^{1+\varepsilon} \log N \leq \mathrm{AT}(B_{N,\varepsilon}, r) \leq c_2 \cdot N^{1+2\varepsilon} \log N.$$

*Moreover, the fixation time $\mathrm{CT}(B_{N,\varepsilon}, r)$ satisfies the same inequalities.*

### Martingales background

First, we recall the following facts about martingales (see [43]). Fix $r > 1$. Given a complete bipartite graph with $v$ vertices at the outside part and $c$ vertices in the center, the state (configuration) space can be parametrized by the number $0 \leq i \leq v$ of mutants in the outside part and the number $0 \leq j \leq c$ of mutants in the center. For each state $(i, j)$, let $\mathrm{fp}(i, j)$ be the fixation probability starting from that state. There is a formula for $\mathrm{fp}(i, j)$ which can be computed as follows: Let

$$h_v = \frac{v + cr}{vr^2 + cr}, \qquad h_c = \frac{c + vr}{cr^2 + vr}$$

and for every state $(i, j)$ define a potential function $\phi(i, j) = h_v^i \cdot h_c^j$. (Note that $\phi(i+1, j) = \phi(i, j) \cdot h_v$ and $\phi(i, j+1) = \phi(i, j) \cdot h_c$.) Then

$$\mathrm{fp}(i, j) = \frac{\phi(0,0) - \phi(i,j)}{\phi(0,0) - \phi(v,c)} = \frac{1 - \phi(i,j)}{1 - \phi(v,c)}.$$

For the rest of this section, we will be using these results for $c = N^{1-\varepsilon}$ and $v = N$.

### Fixation probability

With the martingales background, the analysis of the fixation probability is relatively straightforward.

**Lemma 15.** Fix $\varepsilon \in (0,1]$ and $r > 1$. As $N \to \infty$, we have $\mathrm{fp}(B_{N,\varepsilon}, r) \to 1 - 1/r^2$.

*Proof.* The original mutant appears at the outside part with probability $N/(N+N^{1-\varepsilon}) \to 1$. Since $\phi(1,0) = h_v \to 1/r^2$ and $\phi(v,c) = h_v^v \cdot h_c^c < h_v^v \to 0$ as $N \to \infty$, we compute

$$\mathrm{fp}(B_{N,\varepsilon}, r) = \frac{1-\phi(1,0)}{1-\phi(v,c)} \to_{N\to\infty} 1 - 1/r^2.$$

□

### Lower bound on fixation time

Next, we present the lower bounds for the absorption and fixation time. The idea is to consider the expected time $t_k$ to gain one mutant in the outside part, if there are currently $k$ mutants there. By bounding those times and summing up we obtain the following lemma.

**Lemma 16.** Fix $\varepsilon \in (0,1]$. Then

$$\mathrm{AT}(B_{N,\varepsilon}, r) = \Omega(N^{1+\varepsilon}\log N), \quad \mathrm{CT}(B_{N,\varepsilon}, r) = \Omega(N^{1+\varepsilon}\log N).$$

*Proof.* For the absorption time, we proceed as in the proof of Theorem 6, that is, we restart the process each time the mutants go extinct. The modified process $M'$ always terminates with the mutants fixating and its expected fixation time is given by $T'(B_{N,\varepsilon}, r) = \frac{1}{\mathrm{fp}(B_{N,\varepsilon},r)} \cdot \mathrm{AT}(B_{N,\varepsilon}, r)$. Consider a state with $1 \le k \le N-1$ mutants in the outside part and $0 \le j \le c$ mutants at the center. Let $F = N + c + (r-1)(j+k) > N$ be the total fitness of the population. The probability that in the next step we gain one mutant in the outside part equals

$$\frac{r\cdot j}{F} \cdot \frac{N-k}{N} \le \frac{r \cdot c}{N^2} \cdot (N-k) \equiv p_k.$$

Since $p_k$ is independent of $j$, the expected time to reach some state with $k+1$ mutants, starting in any state with $k$ mutants in the outside part, is at least

$$\frac{1}{p_k} = \frac{1}{r} \cdot \frac{N^2}{c} \cdot \frac{1}{N-k} \equiv t_k.$$

In order to fixate, we need to pass through a state with $k$ mutants in the outside part, for each $k = 1, \dots, N-1$. By linearity of expectation,

$$\begin{aligned}\mathrm{AT}(B_{N,\varepsilon}, r) &= \mathrm{fp}(B_{N,\varepsilon}, r) \cdot T'(B_{N,\varepsilon}, r) \\ &\ge \mathrm{fp}(B_{N,\varepsilon}, r) \cdot \sum_{k=1}^{N-1} \frac{1}{r} \cdot \frac{N^2}{c} \cdot \frac{1}{N-k} \to \frac{r^2-1}{r^3} \cdot N^{1+\varepsilon} \cdot \sum_{k=1}^{N-1} \frac{1}{k} = \Theta(N^{1+\varepsilon}\log N).\end{aligned}$$

For the fixation time, we perform a standard construction to obtain a different modified process $M''$ that only includes the trajectories that lead to fixation. Specifically, we remove the state $(0,0)$ (the only state $s$ with $\mathrm{fp}(s)=0$) and, for any two other states $s$ and $t$, we renormalize the transition probability $p(s\to t)$ to a new value $p''(s\to t)=p(s\to t)\cdot\frac{\mathrm{fp}(t)}{\mathrm{fp}(s)}$. It is a standard result that in this way we have constructed a Markov chain with only one absorbing state whose absorption time is equal to the fixation time of the original process, that is, $\mathrm{CT}(B_{N,\varepsilon},r)=\mathrm{AT}''(B_{N,\varepsilon},r)$. To get a lower bound for $\mathrm{AT}''(B_{N,\varepsilon},r)$, we proceed as before.

Due to the renormalization, each $p_k$ $(k=1,\dots,N-1)$ gets multiplied by a ratio of two fixation probabilities that can be upper bounded by

$$\frac{\max_j\{\mathrm{fp}(k+1,j)\}}{\min_j\{\mathrm{fp}(k,j)\}}.$$

Note that for $k\geq 1$ the denominator is at least a constant (recall that $\mathrm{fp}(1,0)\to 1-1/r^2$ for large $N$), hence the ratio can be further upper bounded by $1/c_0$ for any $c_0<1-1/r^2$ and $N\to\infty$. Hence $t''_k=1/p''_k\geq c_0/p_k$. This gives

$$\begin{aligned}\mathrm{CT}(B_{N,\varepsilon},r)=\mathrm{AT}''(B_{N,\varepsilon},r)\geq\sum_{k=1}^{N-1}t''_k\geq c_0\sum_{k=1}^{N-1}\frac{1}{r}\cdot\frac{N^2}{c}\cdot\frac{1}{N-k}=\frac{c_0}{r}\cdot N^{1+\varepsilon}\cdot\sum_{k=1}^{N-1}\frac{1}{k}\\ =\Theta(N^{1+\varepsilon}\log N)\end{aligned}$$

as desired.

□

### Upper bound: "small" center

For the upper bound, we distinguish two cases. First, we assume that $\varepsilon\in(1/2,1]$, that is $c=o(\sqrt{N})$.

The idea is to again work with the restarted process and moreover to split the set of states into *sections* as follows: section $S_i$ consists of all the states with $i$ mutants in the outside part. Then we consider a Markov chain $\mathcal{M}'$ whose nodes are the sections $S_i$. By construction, the only transitions with nonzero probability are of the form $S_i\to S_{i\pm1}$ or $S_i\to S_i$. In the following sequence of Lemmas, we provide upper bounds for the expected number of transitions from $S_{i+1}$ to $S_i$ and for the expected number of transitions within each $S_i$. Summing up, we obtain an upper bound for the fixation time in the original Markov chain.

Formally, fix $i$ and let

- $f_{\max} = \max_j\{\mathrm{fp}(i,j)\}$ be the maximum fixation probability from a state in $S_i$. Clearly, $f_{\max}$ is attained in state $(i,c)$.
- $g_{\min} = \min_j\{\mathrm{fp}(i+1,j)\}$ be the minimum fixation probability from a state in $S_{i+1}$. Clearly, $g_{\min}$ is attained in state $(i+1,0)$.
- $q = \min_j\{q_j\}$ where $q_j$ is the probability that an evolutionary trajectory starting at $(i+1,j)$ fixates at $(v,c)$ before visiting any state in $S_i$.

First, since $\varepsilon > 1/2$ we have the following:

**Lemma 17.** $h_c^c \to_{N\to\infty} 1$ and $f_{\max} < g_{\min}$ (for large enough $N$)

*Proof.* We have

$$h_c^c \approx \left(1 - \frac{r - 1/r}{N^\varepsilon}\right)^{(N^{1-\varepsilon})}$$

For $N \to \infty$ we have $N^\varepsilon \to \infty$. If the parenthesis was raised to power $N^\varepsilon$, the limit would have been $\exp(-(r-1/r))$, a constant. Since $N^{1-\varepsilon} = o(N^\varepsilon)$ for $\varepsilon > 1/2$, we have $\lim_{N\to\infty} h_c^c = 1$. Hence $h_c^c > h_v$, then $\phi(i,c) = h_c^c \cdot \phi(i,0) > h_v \cdot \phi(i,0) = \phi(i+1,0)$ and thus $f_{\max} = \mathrm{fp}(i,c) < \mathrm{fp}(i+1,0) = g_{\min}$ as desired. □

We aim to bound $q$ from below and use it to bound the expected number $X$ of transitions from (any state in) $S_{i+1}$ to (any state in) $S_i$ from above.

**Lemma 18.** $q \geq \frac{g_{\min} - f_{\max}}{1 - f_{\max}}$

*Proof.* Let's run an evolutionary trajectory from some state $(i+1,j)$ in $S_{i+1}$. The trajectory can't go extinct without hitting $S_i$. Conditioning on if the trajectory first fixates or hits $S_i$, we can write

$$g_{\min} \leq \mathrm{fp}(i+1,j) \leq q_j \cdot 1 + (1-q_j) \cdot f_{\max}$$

which rewrites as

$$q_j \geq \frac{g_{\min} - f_{\max}}{1 - f_{\max}}.$$

This is true for every $j$, hence it is true for $q = \min_j\{q_j\}$ too. □

Let $X$ be a random variable counting the transitions from any state in $S_{i+1}$ to any state in $S_i$, starting from any state.

**Lemma 19.** $\mathbb{E}[X] \leq \frac{1-q}{q} = \frac{1-g_{\min}}{g_{\min}-f_{\max}}$.

*Proof.* Any two transitions from section $S_{i+1}$ to section $S_i$ are necessarily separated by an intermediate visit to section $S_{i+1}$. Any time we are in section $S_{i+1}$, with probability at least $q$ we fixate before hitting section $S_i$ again. Hence

$$E[X] \leq q \cdot 0 + (1-q)(1+E[X]).$$

Rewriting and using the bound for $q$ we obtain

$$\mathbb{E}[X] \leq \frac{1-q}{q} = \frac{1}{q} - 1 = \frac{1-f_{\max}}{g_{\min}-f_{\max}} - 1 = \frac{1-g_{\min}}{g_{\min}-f_{\max}}.$$

□

Rewriting $g_{\min}$ and $f_{\max}$ in terms of $h_v$, $h_c$ we deduce that $\mathbb{E}[X]$ is constant.

**Lemma 20.** $\mathbb{E}[X] \leq \frac{1}{r^2-1}$ (for large enough $N$)

*Proof.* Recall that $\mathrm{fp}(i,j) = \frac{1-\phi(i,j)}{d}$ where $d = 1-\phi(v,c)$ doesn't depend on $i, j$. Plugging this in the bound from Lemma 19 we get

$$\begin{aligned}\mathbb{E}[X] \leq \frac{1-g_{\min}}{g_{\min}-f_{\max}} &= \frac{1-\frac{1-\phi(i+1,0)}{d}}{\frac{1-\phi(i+1,0)}{d}-\frac{1-\phi(i,c)}{d}} \\ &= \frac{d-(1-\phi(i+1,0))}{1-\phi(i+1,0)-(1-\phi(i,c))} < \frac{\phi(i+1,0)}{\phi(i,c)-\phi(i+1,0)}.\end{aligned}$$

Using the definition $\phi(i,j) = h_v^i h_c^j$ and dividing by $h_v^i$ this can be further rewritten as

$$\mathbb{E}[X] < \frac{\phi(i+1,0)}{\phi(i,c)-\phi(i+1,0)} = \frac{h_v}{h_c^c - h_v} \to_{N\to\infty} \frac{1/r^2}{1-1/r^2}$$

as desired. □

Let $\mathbb{E}[\mathcal{L}_i]$ be the expected number of "looping" transitions of the form $S_i \to S_i$ before a transition of the form $S_i \to S_{i\pm1}$ occurs (or the process reaches an absorbing state). The following lemma bounds $\mathbb{E}[\mathcal{L}_i]$ from above.

**Lemma 21.** For $i = 1, 2, \dots, N-1$ we have $\mathbb{E}[\mathcal{L}_i] \leq \frac{r\cdot N(N+c)}{c\cdot\min\{i,N-i\}} - 1$. Moreover, $\mathbb{E}[\mathcal{L}_0] \leq r(N+c)-1$ and $\mathbb{E}[\mathcal{L}_N] \leq r(N+c)-1$.

*Proof.* Crudely (not caring about $r$). First, let $i = 1, \dots, N-1$. We pick a vertex in the center with probability at least $\frac{1 \cdot c}{r(N+c)}$. No matter its type, there are at least $\min\{i, N-i\}$ vertices of the other type at the outside part. Hence with probability

$$p \geq \frac{1 \cdot c}{r(N+c)} \cdot \frac{\min\{i, N-i\}}{N}$$

we transition to section $S_{i\pm 1}$ in one step. As before, we get the result from $\mathbb{E}[\mathcal{L}_i] \leq \frac{1}{p} - 1$. Second, if $i = 0$ or $i = N$ and we are not in an absorbing state then there exists a vertex in the center whose type is different to the type of all vertices in the outside part. Hence $p \geq \frac{1}{r(N+c)}$ and we conclude as in the first case. □

We are ready to sum those contributions up.

**Lemma 22.** If $\varepsilon \in (1/2, 1]$, $c = N^{1-\varepsilon} = o(\sqrt{N})$ and $r > 1$ then

$$\mathrm{AT}(B_{N,\varepsilon}, r) = O\left(\frac{N^2}{c} \cdot \log N\right) = O(N^{1+\varepsilon} \log N).$$

*Proof.* As in the proof of Theorem 6 we restart the process anytime the mutants fixate. Consider the one-dimensional Markov chain $\mathcal{M}'$ whose vertices are the sections $S_i$, $i = 0, \dots, N$. Fix $i \in \{1, \dots, N-1\}$ and let $f(r) = \frac{1}{r^2-1}$. On average, there are at most $f(r)$ transitions $S_{i+1} \to S_i$. Also, on average there are at most $f(r)$ transitions $S_i \to S_{i-1}$, hence there are at most $f(r)+1$ transitions $S_{i-1} \to S_i$ for a total of at most $2f(r)+1$ transitions from outside of $S_i$ to $S_i$. Similarly, on average there are at most $f(r)$ transitions into $S_0$ and at most $f(r)+1$ transitions into $S_N$. Every time there is a transition into $S_i$, there are on average $\mathbb{E}[\mathcal{L}_i]$ transitions within $S_i$. By linearity of expectation, the total expected number of transitions is at most

$$\begin{aligned}
\mathrm{AT}(B_{N,\varepsilon}, r) &= \mathrm{fp}(B_{N,\varepsilon}, r) \cdot T'(B_{N,\varepsilon}, r) \\
&\leq \mathrm{fp}(B_{N,\varepsilon}, r) \cdot \left( (2f(r)+1) \cdot r(N+c) + \sum_{i=1}^{N-1} (2f(r)+1)(1 + \mathbb{E}[\mathcal{L}_i]) \right) \\
&= \mathrm{fp}(B_{N,\varepsilon}, r) \cdot (2f(r)+1) \cdot r(N+c) \cdot \left( 1 + \frac{N}{c} \sum_{i=1}^{N-1} \frac{1}{\min\{i, N-i\}} \right) \\
&= \Theta\left( \frac{N^2}{c} \cdot \log N \right),
\end{aligned}$$

where in the inequality we first combined the terms for $S_0$ and $S_N$. The last equality follows from the sum being $\Theta(2\log(N/2)) = \Theta(\log N)$ and from $c = o(N)$. □

### Upper bound: "large" center

Note that the argument used for small center fails for $\varepsilon \leq 1/2$ because the difference $g_{\min} - f_{\max}$ becomes zero or even negative. Indeed, for $\varepsilon = 1/2$ we have $h_c^c \to_{N\to\infty} \exp(-(r-1/r)) < 1/r^2$ and for $\varepsilon < 1/2$ the inequality is even stricter. However an analogous argument can be made to work if we split the state space into different "tilted" sections, taking $\varepsilon$ into account. The idea of the proof is that we fix $\varepsilon \in (0, 1/2)$, consider large $N$, and look at a complete bipartite graph $B_{N,\varepsilon}$. We assume that $r$ is such that there exists an integer $t$ called "tilt" satisfying $h_c^t = h_v$. This assumption guarantees that the states $(i, j+t)$ and $(i+1, j)$ are assigned exactly the same potential. We can then split the state (configuration) space into $\Theta(N)$ sections where each section is not a vertical line but a set of $c$ states that looks like a line tilted with slope $-t$ (see figure). We then proceed as before, providing an upper bound for the number of transitions across sections and within sections. The result follows by summing up.

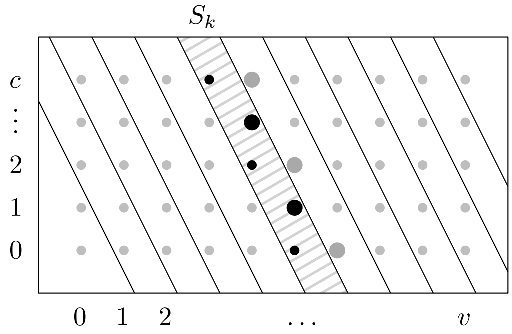

Figure 3.1: **Tilted sections of $B_{N,\varepsilon}$.** For $\varepsilon \in (0, 1/2)$ we split the state (configuration) space into $\Theta(N)$ "tilted" sections $S_k$. Here the tilt is $t = 2$. The maximum potential within $S_k$ is attained at any thick black vertex, the minimum potential within $S_{k+1}$ is attained at any thick grey vertex.

In the rest of the section we formalize this idea. First, we define the (tilted) sections. Let $s = v + \lfloor c/t \rfloor$ and for $k = 0, \dots, s$ let

$$S_k = \{(i,j) : i + \lfloor j/t \rfloor = k\}.$$

As before, we fix $k$ and define

- $f_{\max} = \max\{\mathrm{fp}(i,j) : (i,j) \in S_k\}$,
- $g_{\min} = \min\{\mathrm{fp}(i,j) : (i,j) \in S_{k+1}\}$, and

- $q = \min_{(i,j)\in S_{k+1}}\{q_{(i,j)}\}$ where $q_{(i,j)}$ is the probability that an evolutionary trajectory starting at state $(i,j)$ belonging to $S_{k+1}$ fixates at $(v,c)$ before visiting any state in $S_k$.

Clearly, $f_{\max}$ is attained for any "top" state of $S_k$ within its column (possibly not in the "top" row). Similarly, $g_{\min}$ is attained in state "bottom" state of $S_{k+1}$ (possibly not in the "bottom" row). Note that by construction, those two states are assigned potentials that differ by a factor of $h_c$.

As before, let $X$ be a random variable counting the transitions from any state in $S_{k+1}$ to any state in $S_k$, starting from any state. The following lemma bounds $\mathbb{E}[X]$ from above. Note that this time the bound is super-constant.

**Lemma 23.** $\mathbb{E}[X] \le \frac{r}{r^2-1}\cdot N^{\varepsilon}$ (for large $N$)

*Proof.* Note that Lemma 18 and Lemma 19 are valid for tilted sections too. Let $(i,j)$ be some state in $S_k$ for which the value $f_{\max}$ is attained. Then $g_{\min}$ is attained at a state whose potential is equal to $h_c\cdot\phi(i,j)$. We continue as in the proof of Lemma 20 to get

$$\mathbb{E}[X] \le \frac{1-g_{\min}}{g_{\min}-f_{\max}} = \frac{h_c\cdot\phi(i,j)-\phi(v,c)}{\phi(i,j)-h_c\cdot\phi(i,j)} < \frac{h_c}{1-h_c}.$$

Since for large $N$ we have $h_c \approx 1-(1-1/r)/N^{\varepsilon}$, the right-hand side can be approximated as

$$\frac{h_c}{1-h_c} < \frac{1}{1-h_c} = \frac{N^{\varepsilon}}{r-1/r} = \frac{r}{r^2-1}\cdot N^{\varepsilon}.$$

□

It remains to bound the expected number $\mathbb{E}[\mathcal{L}_k]$ of the "looping" transitions of the form $S_k \to S_k$ before a transition of the form $S_k \to S_{k\pm1}$ occurs. This is done as before, observing that any two states that differ only in the number of mutants in the outside part of the graph always lie in different sections. Hence Lemma 21 holds.

Now we are ready ti prove the last inequality in Theorem 7.

**Lemma 24.** If $\varepsilon \in (0,1/2)$ then

$$\mathrm{AT}(B_{N,\varepsilon},r) = O(N^{1+2\varepsilon}\log N).$$

*Proof.* As before, let $f(r) = \frac{r}{r^2-1} \cdot N^{\varepsilon}$. By linearity of expectation, the total expected number of transitions is

$$\begin{aligned}
\mathrm{AT}(B_{N,\varepsilon}, r) &\leq \mathrm{fp}(B_{N,\varepsilon}, r) \cdot \left( (2f(r)+1) \cdot r(N+c) + \sum_{i=1}^{N-1} (2f(r)+1)(1+\mathbb{E}[\mathcal{L}_i]) \right) \\
&= \mathrm{fp}(B_{N,\varepsilon}, r) \cdot (2f(r)+1) \cdot r(N+c) \cdot \left( 1 + \frac{N}{c} \sum_{i=1}^{N-1} \frac{1}{\min\{i, N-i\}} \right) \\
&= O\left( N^{\varepsilon} \cdot N \cdot \frac{N}{c} \cdot \log N \right) = O(N^{1+2\varepsilon} \log N)
\end{aligned}$$

and the result follows. □

Finally, we observe that an upper bound on $\mathrm{AT}(B_{N,\varepsilon}, r)$ immediately implies an asymptotically matching upper bound on $\mathrm{CT}(B_{N,\varepsilon}, r)$.

**Lemma 25.** Fix $r > 1$ and $\varepsilon > 0$. If $\mathrm{AT}(B_{N,\varepsilon}, r) = O(N^{\alpha} \log N)$ then $\mathrm{CT}(B_{N,\varepsilon}, r) = O(N^{\alpha} \log N)$

*Proof.* Since the absorption time is a weighted average of the fixation time and the extinction time, we have

$$\begin{aligned}
\mathrm{AT}(B_{N,\varepsilon}, r) &= \mathrm{fp}(B_{N,\varepsilon}, r) \cdot \mathrm{CT}(B_{N,\varepsilon}, r) + (1 - \mathrm{fp}(B_{N,\varepsilon}, r)) \cdot \mathrm{ET}(B_{N,\varepsilon}, r)) \\
&\geq \mathrm{fp}(B_{N,\varepsilon}, r) \cdot \mathrm{CT}(B_{N,\varepsilon}, r)
\end{aligned}$$

Since $\mathrm{fp}(B_{N,\varepsilon}, r) \to 1 - 1/r^2$ as $N \to \infty$ and $\mathrm{AT}(B_{N,\varepsilon}, r) = O(N^{\alpha} \log N)$, the result follows. □

Altogether, Lemmas 15, 16, 22, 24 and 25 prove all the statements of Theorem 7.

### 3.3.3 Weighted bipartite graphs

In this section we analyze the Weighted bipartite graphs $W_{N,\varepsilon}$ under temperature initialization. Recall that $W_{N,\varepsilon}$ is a complete bipartite graph with one (smaller) part of size $c = N^{1-\varepsilon}$, one (larger) part of size $N$, and every vertex of the larger part having a self-loop of such a weight $w$ that $N^{-\varepsilon/2} = \frac{N^{1-\varepsilon}}{w+N^{1-\varepsilon}}$.

We prove the following theorem.

**Theorem 8.** *Fix $\varepsilon \in (0,1)$ and $r > 1$. Let $W_{N,\varepsilon}$ be the Weighted bipartite graph. Then*

- $\mathrm{fp}(W_{N,\varepsilon}, r) \to 1 - 1/r^2$.
- *There exist constants* $c_1$, $c_2$ *such that*

$$c_1 \cdot N^{1+\varepsilon} \log N \le \mathrm{AT}(B_{N,\varepsilon}, r) \le c_2 \cdot N^{1+\frac{3}{2}\varepsilon} \log N.$$

*Moreover, the fixation time* $\mathrm{CT}(B_{N,\varepsilon}, r)$ *satisfies the same inequalities.*

### Martingales for Weighted bipartite graphs

First, we recall more martingales background.

Fix $r > 1$. Given integers $v$, $c$, and a real number $q \in (0, 1)$, let $W(c, v, q)$ be a Weighted complete bipartite graph with $c$ vertices in the smaller part (center) and $v$ vertices at the larger (outside) part, each of them with an extra self-loop of such weight $w$ that $q = \frac{c}{w+c}$ is the probability that when a vertex in the larger part is selected for reproduction, its offspring replaces one of the vertices in the smaller part (as opposed to replacing its parent via the self-loop). Then, as with the unweighted complete bipartite graphs, the state space can be parametrized by the number $0 \le i \le v$ of mutants in the outside part and the number $0 \le j \le c$ of mutants in the center and the fixation probabilities from all the states can be computed similarly to above, with $v$ replaced by $v \cdot q$.

Namely, let

$$h_v = \frac{qv + cr}{qvr^2 + cr}, \qquad h_c = \frac{c + qvr}{cr^2 + qvr}$$

and for every state $(i, j)$ of $i$ mutants in the outside part and $j$ mutants in the center, define a potential function $\phi(i, j) = h_v^i \cdot h_c^j$. Then we easily check the fixation probability from a state $(i, j)$ is given by

$$\mathrm{fp}(i, j) = \frac{\phi(0, 0) - \phi(i, j)}{\phi(0, 0) - \phi(v, c)} = \frac{1 - \phi(i, j)}{1 - \phi(v, c)}.$$

### Fixation probability

With the extra martingales background, the analysis of the fixation probability is again relatively straightforward.

**Lemma 26.** $\mathrm{fp}_{\mathsf{T}}(W_{N,\varepsilon}, r) \to 1 - 1/r^2$.

*Proof.* The first mutant is introduced in the center with probability proportional to $q \cdot v$ and to the outside part with probability proportional to $c + (1-q)v$. Since $q = o(1)$, it is introduced to the outside part with high probability. The fixation probability $\mathrm{fp}(1,0)$ starting from a state with a single mutant in the outside part satisfies

$$\mathrm{fp}(1,0) = \frac{1-\phi(1,0)}{1-\phi(v,c)} = \frac{1-h_v}{1-h_v^v h_c^c}.$$

Since $h_v \approx \frac{1}{r^2}$ and $h_v^v \to 0$ as $N \to \infty$, we have $\mathrm{fp}(1,0) \to 1 - 1/r^2$. ☐

**Fixation time**

The arguments for fixation time are direct translations of arguments for (unweighted) $\varepsilon$-Balanced bipartite graphs (see Section 3.3.2). For the lower bound, Lemma 16 still applies. For the upper bound, we proceed analogously.

**Lemma 27.** If $\varepsilon \in (0,1)$ and $r > 1$ then $T_{\mathsf{T}}(W_{N,\varepsilon}, r) = O(N^{1+\frac{3}{2}\varepsilon} \log N)$.

*Proof.* Fixing $k$ and considering the section $S_k$, we denote by $X$ the expected number of transitions from any state in $S_{k+1}$ to any state in $S_k$. As in Lemmas 20 and 23 we get

$$\mathbb{E}[X] \le \frac{h_c}{1-h_c} = \Theta(N^{\varepsilon/2}).$$

Lemma 21 then yields

$$T_{\mathsf{T}}(W_{N,\varepsilon}) = O\left(N^{\varepsilon/2} \cdot N \cdot \frac{N}{c} \cdot \log N\right) = O(N^{1+\frac{3}{2}\varepsilon} \log N).$$

☐

Altogether, Lemmas 16, 25, 26 and 27 prove all the statements of Theorem 8

### 3.3.4 Time on Complete graph

For reference purposes we compute the absorption, fixation, and extinction times of a single advantageous mutant ($r > 1$) on a Complete graph, using the standard difference method.

**Theorem 9.** *Fix $r > 1$ and let $K_N$ be the Complete graph on $N$ vertices. Then*

$$\begin{aligned}\mathrm{AT}(K_N, r) &= (N-1)H_{N-1} \cdot \frac{r+1}{r} + (N-1) \cdot \log(1-1/r) - \frac{1}{r(r-1)} + o(1),\\ \mathrm{CT}(K_N, r) &= (N-1)H_{N-1} \cdot \frac{r+1}{r-1} + (N-1) \cdot \frac{r+1}{r-1} \log(1-1/r) + o(N),\\ \mathrm{ET}(K_N, r) &= (N-1) \cdot \log\left(\frac{r}{r-1}\right) + o(N).\end{aligned}$$

*In particular for $r = 1+s$, $s > 0$ small, we have* $\mathrm{AT}(K_N, r) \approx 2 \cdot N \log N$, $\mathrm{CT}(K_N, r) \approx \frac{2}{s} \cdot N \log N$, *and* $\mathrm{ET}(K_N, r) \approx \frac{1}{s} \cdot N$.

*Proof.* First, we compute the absorption time, then the fixation time and finally the extinction time.

**Absorption time.** Fix $N$ and $r$ and for $k = 0, \dots, N$ let $T_k$ be the expected absorption time from a state with $k$ mutants. Clearly $T_0 = T_N = 0$ and for $k = 1, \dots, N-1$ we have

$$T_k = 1 + p(k,k)T_k + p(k,k-1)T_{k-1} + p(k,k+1)T_{k+1},$$

where $p(i,j)$ is the transition probability from a state with $i$ mutants to a state with $j$ mutants. Specifically, we have

$$p(k,k-1) = \frac{N-k}{N+(r-1)k} \cdot \frac{k}{N-1}. \qquad \text{and} \qquad p(k,k+1) = \frac{r \cdot k}{N+(r-1)k} \cdot \frac{N-k}{N-1}.$$

Plugging in those values of $p(i,j)$, the above equation can be rewritten as

$$T_{k+1} - T_k = \frac{1}{r}(T_k - T_{k-1}) - \left(\frac{N-1}{rk} + \frac{N-1}{N-k}\right).$$

Setting $\Delta_k \equiv T_k - T_{k-1}$ and $x_k = \frac{N-1}{rk} + \frac{N-1}{N-k}$ this further rewrites as

$$\Delta_{k+1} = \frac{1}{r}\Delta_k - x_k.$$

Specifically, we have $\Delta_1 = T_1 - T_0 = T_1$ and $\Delta_1 + \cdots + \Delta_N = T_N - T_0 = 0$. Let's write

$$\begin{aligned}\Delta_2 &= \frac{1}{r}\Delta_1 - x_1, && (1)\\ \Delta_3 &= \frac{1}{r}\Delta_2 - x_2, && (2)\\ \Delta_4 &= \frac{1}{r}\Delta_3 - x_3, && (3)\\ &\dots\\ \Delta_N &= \frac{1}{r}\Delta_{N-1} - x_{N-1}, && (N-1)\end{aligned}$$

We aim to express each $\Delta_k$ in terms of $\Delta_1$ only. Summing up $\frac{1}{r}(1) + (2)$ gives

$$\Delta_3 = \frac{1}{r^2}\Delta_1 - \left(x_2 + \frac{1}{r}x_1\right)$$

Similarly, summing up $\frac{1}{r^2}(1) + \frac{1}{r}(2) + (3)$ gives

$$\Delta_4 = \frac{1}{r^3}\Delta_1 - \left(x_3 + \frac{1}{r}x_2 + \frac{1}{r^2}x_1\right)$$

and similarly all the way up to

$$\Delta_N = \frac{1}{r^{N-1}}\Delta_1 - \left(x_{N-1} + \frac{1}{r}x_{N-2} + \cdots + \frac{1}{r^{N-2}}x_1\right).$$

Summing up all of them, together with an extra equation $\Delta_1 = \Delta_1$, we get

$$\begin{aligned} 0 &= \Delta_1 + \cdots + \Delta_N \\ &= \Delta_1\left(1 + \cdots + \frac{1}{r^{N-1}}\right) - \big(x_1(1 + \cdots + 1/r^{N-2}) + x_2(1 + \cdots + 1/r^{N-3}) + \cdots + x_{N-1}\cdot 1\big) \end{aligned}$$

and in turn

$$\begin{aligned} \Delta_1 &= \frac{1 - 1/r}{1 - 1/r^N} \cdot \frac{x_1(1 - 1/r^{N-1}) + x_2(1 - 1/r^{N-2}) + \cdots + x_{N-1}(1 - 1/r)}{1 - 1/r} \\ &= \frac{1}{1 - 1/r^N}\Big(\underbrace{\sum_{k=1}^{N-1} x_k}_{A} - \underbrace{\sum_{k=1}^{N-1} \frac{x_k}{r^{N-k}}}_{B}\Big). \end{aligned}$$

For $A$ we easily get $A = \frac{r+1}{r}(N-1)H_{N-1}$. For $B$, we plug in $x_k = \frac{N-1}{rk} + \frac{N-1}{N-k}$, split $B = B_1 + B_2$ and separately compute the sums using a well-known limit

$$\begin{aligned} B_2 &= (N-1)\sum_{i=1}^{N-1} \frac{1}{N-k} \cdot \frac{1}{r^{N-k}} \\ &= (N-1)\big(\log(1 - 1/r) + O(1/r^N)\big) \to (N-1)\log(1 - 1/r) + o(1) \end{aligned}$$

and an approximation

$$B_1 = \frac{1}{r}\sum_{k=1}^{N-1} \frac{N-1}{N-k}\frac{1}{r^k} = \frac{1}{r}\sum_{k=1}^{N-1}\frac{1}{r^k} + E(N) = \frac{1}{r(r-1)} + o(1)$$

whose error term

$$E(N) = \sum_{k=1}^{N-1} \frac{k-1}{N-k} \cdot \frac{1}{r^k}$$

tends to 0, because the sum $S_1$ over the first $\sqrt[3]{N}$ terms satisfies

$$S_1 \le \sqrt[3]{N} \cdot \frac{\sqrt[3]{N}}{N - \sqrt[3]{N}} \cdot \frac{1}{r} < \frac{1}{\sqrt[3]{N}} \cdot \frac{1}{r} \to 0$$

and the sum $S_2$ of the remaining terms satisfies

$$S_2 \leq (N - \sqrt[3]{N}) \cdot \frac{N}{1} \cdot \frac{1}{r^{\sqrt[3]{N}}} \to 0.$$

This conclude the proof of the absorption time.

**Fixation time.** We proceed similarly.

As before, we fix $N$ and $r$ and for $k = 1, \ldots, N$ we let $\mathrm{CT}_k$ be the expected (conditional) fixation time from a state with $k$ mutants (for $k = 0$ we define $\mathrm{CT}_0 = 0$). Then $\mathrm{CT}_N = 0$ and for $k = 1, \ldots, N-1$ we have

$$\mathrm{fp}_k \mathrm{CT}_k = \mathrm{fp}_k + p(k,k) \cdot \mathrm{fp}_k \mathrm{CT}_k + p(k,k-1) \cdot \mathrm{fp}_{k-1} \mathrm{CT}_{k-1} + p(k,k+1) \cdot \mathrm{fp}_{k+1} \mathrm{CT}_{k+1},$$

where $\mathrm{fp}_i = \frac{1-1/r^i}{1-1/r^N}$ are the fixation probabilities and $p(i,j)$ are the transition probabilities. Setting $\Delta_k \equiv \mathrm{fp}_k \mathrm{CT}_k - \mathrm{fp}_{k-1} \mathrm{CT}_{k-1}$ and $x_k \equiv \frac{\mathrm{fp}_k}{p(k,k+1)}$ this can be rewritten as

$$\Delta_{k+1} = \frac{1}{r}\Delta_k - x_k.$$

Specifically, we have $\Delta_1 = \mathrm{fp}_1 \cdot \mathrm{CT}_1 - \mathrm{fp}_0 \cdot \mathrm{CT}_0 = \mathrm{fp}_1 \cdot \mathrm{CT}_1$ and

$$\Delta_1 + \cdots + \Delta_N = \mathrm{fp}_N \cdot \mathrm{CT}_N - \mathrm{fp}_0 \cdot \mathrm{CT}_0 = 0 - 0 = 0.$$

As before, we obtain

$$(1 - 1/r^N) \cdot \Delta_1 = \underbrace{\sum_{k=1}^{N-1} x_k}_{A} - \underbrace{\sum_{k=1}^{N-1} \frac{x_k}{r^{N-k}}}_{B}.$$

This time, $p(k,k+1) = \frac{rk}{N+(r-1)k} \cdot \frac{N-k}{N-1}$ and thus

$$x_k = \frac{\mathrm{fp}_k}{p(k,k+1)} = \frac{1-1/r^k}{1-1/r^N} \cdot \left(\frac{r-1}{r} + \frac{N}{rk}\right) \cdot \frac{N-1}{N-k}$$

and

$$A = \sum_{k=1}^{N-1} x_k = \frac{N-1}{1-1/r^N} \cdot \sum_{k=1}^{N-1} (1 - 1/r^k) \cdot \left(\frac{1}{N-k} + \frac{1}{rk}\right).$$

Multiplying out the two parentheses we get

$$X \equiv \sum_{k=1}^{N-1} \frac{1}{N-k} + \frac{1}{rk} = (1 + 1/r) \cdot H_{N-1}$$

and

$$Y \equiv \sum_{k=1}^{N-1} \frac{1}{r^k (N-k)} \to 0, \quad Z \equiv \sum_{k=1}^{N-1} \frac{1}{r^k \cdot rk} \to \frac{1}{r} \log(1 - 1/r).$$

Hence

$$A = \frac{N-1}{1-1/r^N} \cdot (X+Y+Z) = (1+1/r)\cdot(N-1)H_{N-1} + \frac{1}{r}\log(1-1/r)\cdot(N-1) + o(N).$$

We proceed with $B$ analogously. This time, the only combination that survives is

$$\sum_{k=1}^{N-1} \frac{1}{r^{N-k}} \cdot \frac{1}{N-k} \to \log(1-1/r),$$

hence $B = \log(1-1/r)\cdot(N-1) + o(N)$.

In total, we get

$$\begin{aligned}(1-1/r^N)\cdot\frac{1-1/r}{1-1/r^N}\cdot \mathrm{CT}_1 = (1-1/r^N)\cdot\Delta_1 &= A+B \\ &= \frac{r+1}{r}\cdot(N-1)H_{N-1} + \frac{r+1}{r}\log(1-1/r)\cdot(N-1)+o(N)\end{aligned}$$

and finally the desired

$$\mathrm{CT}(N,r) = \mathrm{CT}_1 = \frac{r+1}{r-1}\cdot(N-1)H_{N-1} + \frac{r+1}{r-1}\log(1-1/r)\cdot(N-1) + o(N).$$

**Extinction time.** A formula for the extinction time follows easily from the absorption time and the fixation time.

It suffices to note that $\mathrm{AT}_1 = \mathrm{fp}_1 \cdot T_1 + (1-\mathrm{fp}_1)\cdot \mathrm{ET}_1$ and plug in the expressions for $\mathrm{AT}_1$ and $T_1$. The $N\log N$ term cancels out and we are left with

$$\mathrm{ET}_1 = -\log(1-1/r)\cdot(N-1) + o(N).$$

□

## 3.4 Numerical and simulation results

Here we present related numerical and simulation results. First we consider uniform initialization. We show fixation probability and fixation time for all undirected graphs on $N=8$ vertices (see Figure 3.2) and for selected graphs of size $N=100$ (see Figure 3.3). Then we do the same for temperature initialization (see Figures 3.4 and 3.5). The results are robust with respect to the selection parameter $r$ (see Figure 3.6). Next, we focus on $\varepsilon$-Balanced graphs and illustrate that as $N\to\infty$, their fixation probability tends to $1-1/r^2$ for any $\varepsilon>0$ (see Figure 3.6), whereas the fixation time seems to scale as $\Theta(N^{1+\varepsilon}\cdot\log N)$ (see Figure 3.8). Finally, we illustrate that, in terms of the effective rate of evolution, $\varepsilon$-Balanced graphs and $\varepsilon$-Weighted bipartite graphs perform well under uniform and temperature initialization, respectively.

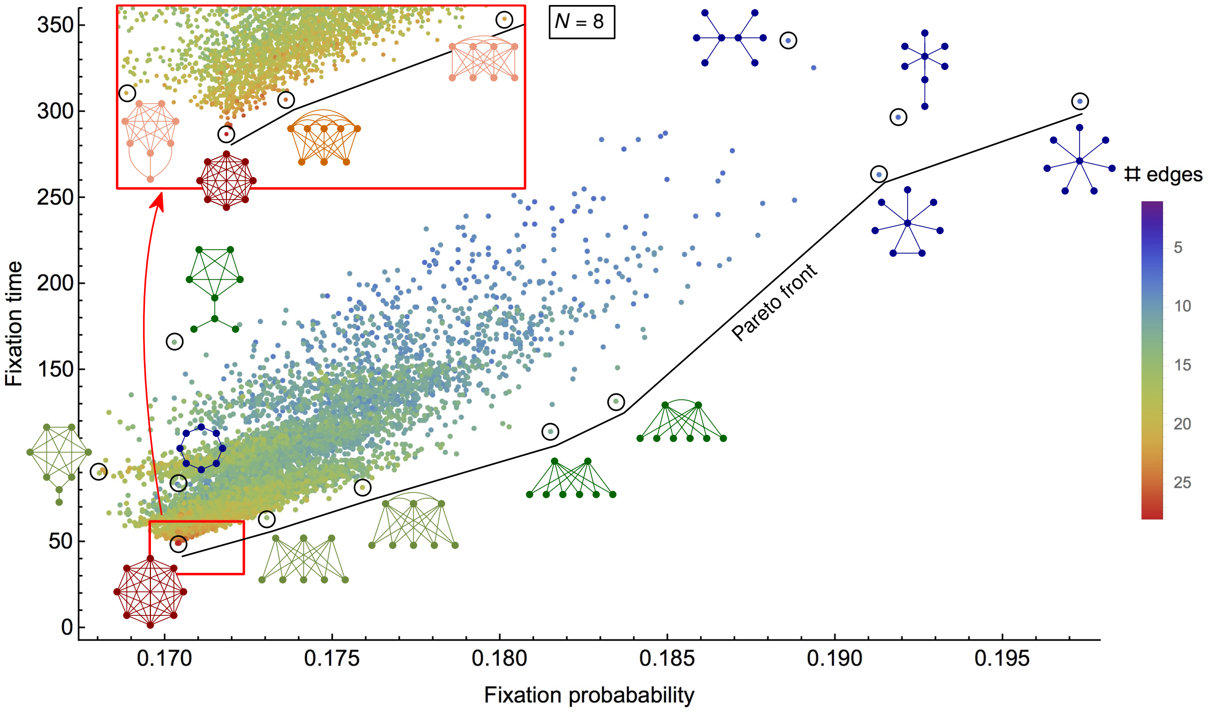

Figure 3.2: **Fixation probability and time under uniform initialization.** Numerical solutions for all 11,117 undirected connected graphs of size $N = 8$. Each graph is represented by a dot and color corresponds to the number of its edges. The $x$- and $y$-coordinates show the fixation probability and the fixation time for a single mutant with relative fitness $r = 1.1$, under uniform initialization. The graphs to the right of the complete graph are amplifiers of selection: they increase the fixation probability. Any graph below the complete graph would be an accelerator of selection: it would decrease the fixation time. Graphs close to the bottom right corner provide good trade-off between high fixation probability and short fixation time. All the values are computed by numerically solving large systems of linear equations (see e.g. [27]). See Figure 3.6 for other $r$ values.

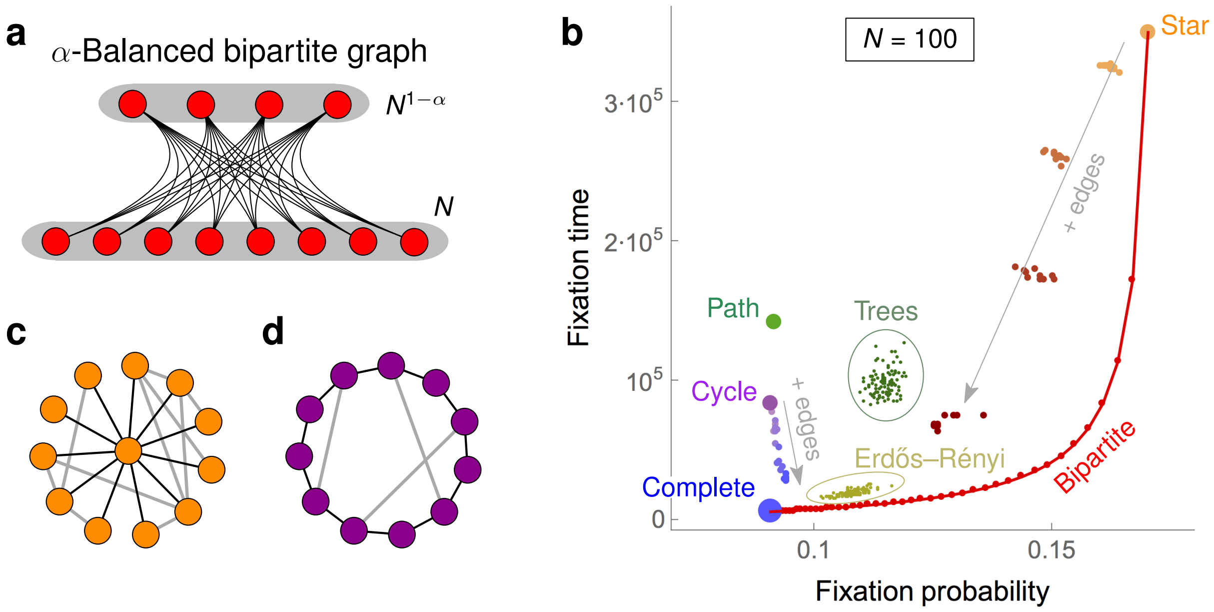

Figure 3.3: **$\varepsilon$-Balanced bipartite graphs.** **a,** An $\varepsilon$-Balanced bipartite graph $B_{N,\varepsilon}$ is a complete bipartite graph with $N$ vertices in the larger part and $N^{1-\varepsilon}$ vertices in the smaller part. Here $N = 8$ and $\varepsilon = 1/3$. We prove that for large $N$, the $\varepsilon$-Balanced bipartite graphs achieve the fixation probability of a star and, for $\varepsilon$ small, approach the fixation time of the complete graph (see Figures 3.7 and 3.8). **b,** In general, bipartite graphs provide great trade-offs between high fixation probability and short fixation time. Comparison is with selected graphs of size $N = 100$ such as Trees (100×), random Erdős–Renyi graphs (100×, $p = 0.03$), star graphs with additional 10, 30, 50, 100 random edges (10× each), and cycle graphs with additional 1, 3, 5, 10 random edges (5× each). The values were obtained by simulating the Moran process $10^5$ times. **c,** Star graph with several random edges. **d,** Cycle graph with several random edges.

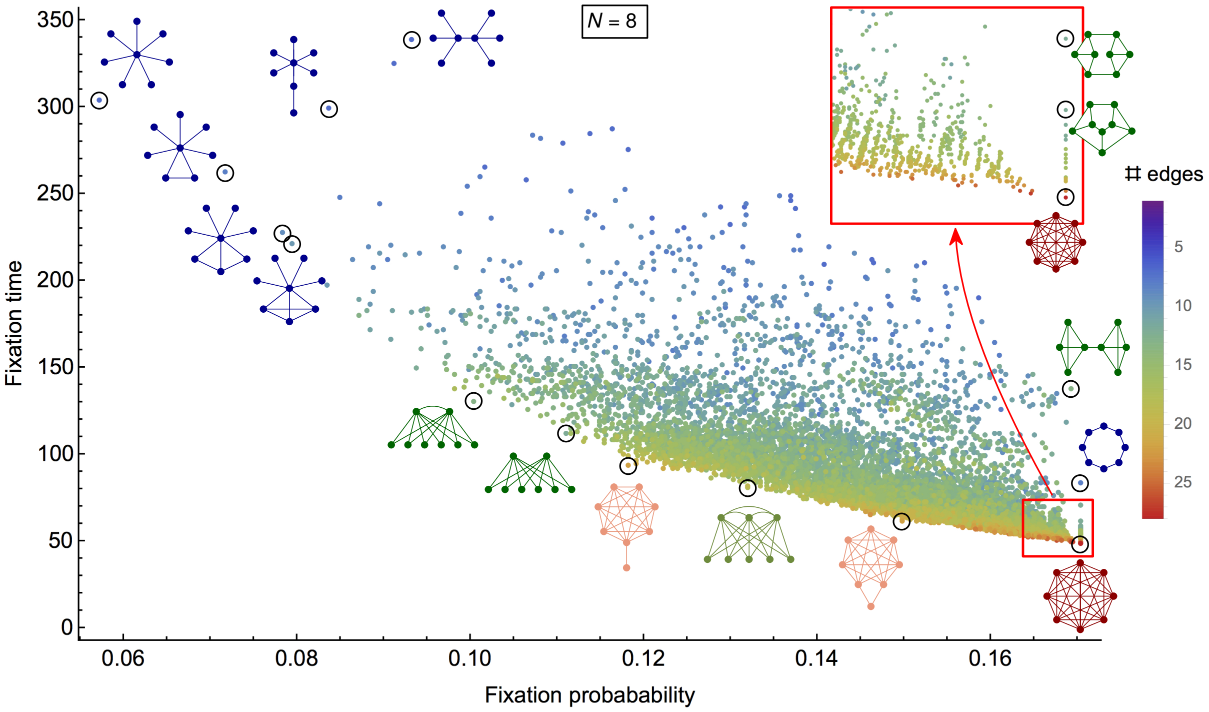

Figure 3.4: **Fixation probability and time under temperature initialization.** Numerical solutions for all undirected connected graphs of size $N = 8$, under temperature initialization ($r = 1.1$). There are no amplifiers and no (strict) accelerators. By the isothermal theorem [7], all the regular graphs achieve the same fixation probability as the complete graph.

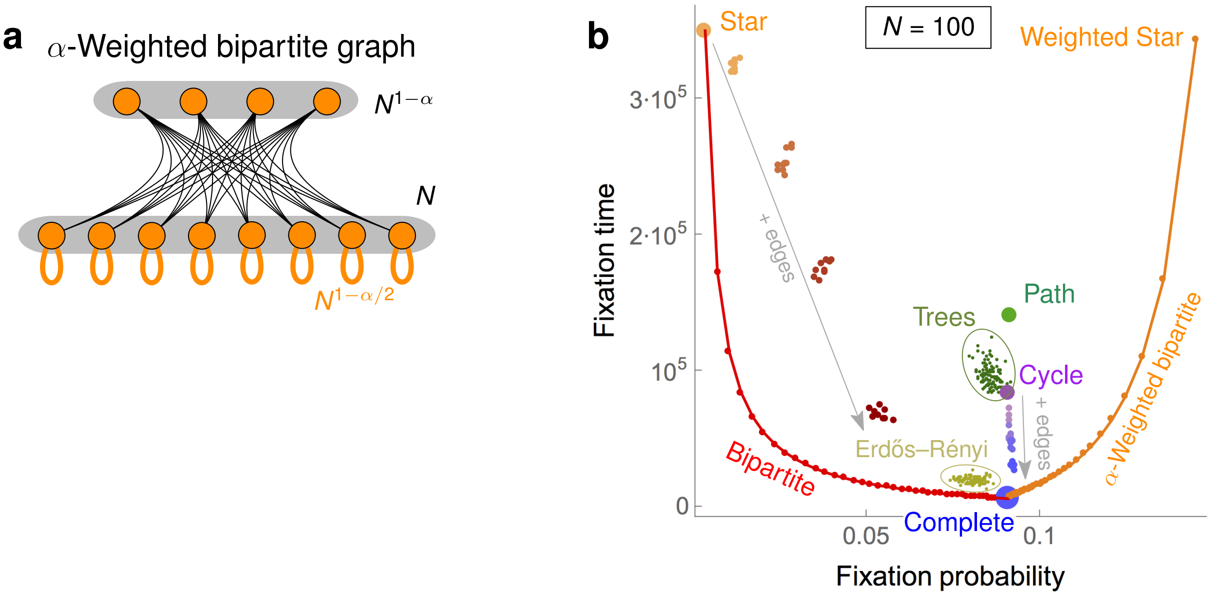

Figure 3.5: **$\varepsilon$-Weighted bipartite graphs.** **a,** An $\varepsilon$-Weighted bipartite graph $W_{N,\varepsilon}$ is obtained by adding self-loops with weight $w \doteq N^{1-\varepsilon/2}$ to all vertices in the larger part of an $\varepsilon$-Balanced bipartite graph. Here $N = 8$ and $\varepsilon = 1/3$. We prove that for large $N$, the $\varepsilon$-Weighted bipartite graphs improve the fixation probability to $1 - 1/r^2$ and, for $\varepsilon$ small, approach the fixation time of the complete graph. **b,** Computer simulations for selected graphs of size $N = 100$ (as in Figure 3.3b). It is known than among unweighted graphs, only a very limited amplification can be achieved [48]. $\varepsilon$-Weighted bipartite graphs (with self-loops of varying weight) overcome this limitation and provide trade-offs between high fixation probability and short fixation time.

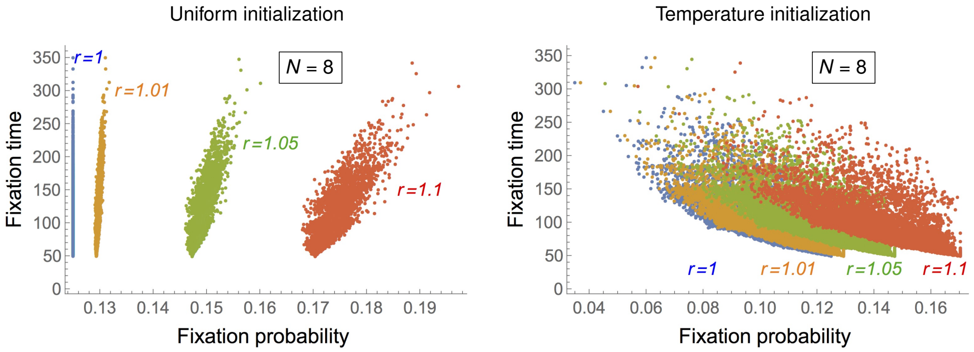

Figure 3.6: **Fixation probability and time under for other $r$ values.** Similar data as in Figures 3.2 and 3.4 for varying $r \in \{1, 1.01, 1.05, 1.1\}$, under **a** uniform or **b** temperature initialization. Under uniform initialization, the fixation probability of a neutral mutant equals $1/N$, independent of the graph structure. As $r$ approaches 1, the point cloud gets closer to a vertical line.

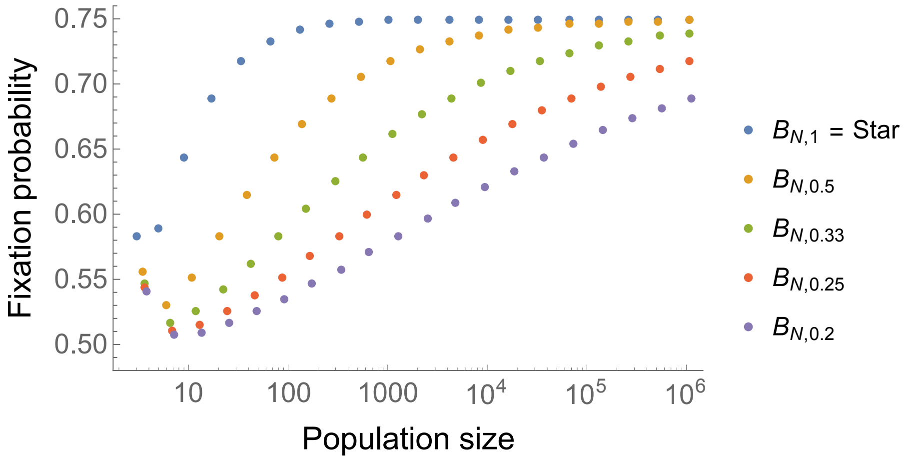

Figure 3.7: **Fixation probability on bipartite graphs $B_{N,\varepsilon}$.** We fix $r = 2$ and consider the $\varepsilon$-Balanced bipartite graphs for $\varepsilon = 1$ (i.e. a star) and $\varepsilon \in \{0.5, 0.33, 0.25, 0.2\}$. The dots are exact values of the fixation probability under uniform initialization, computed by numerically solving large systems of linear equations. The figure illustrates that the fixation probability tends to $1 - 1/r^2$ for any positive $\varepsilon > 0$.

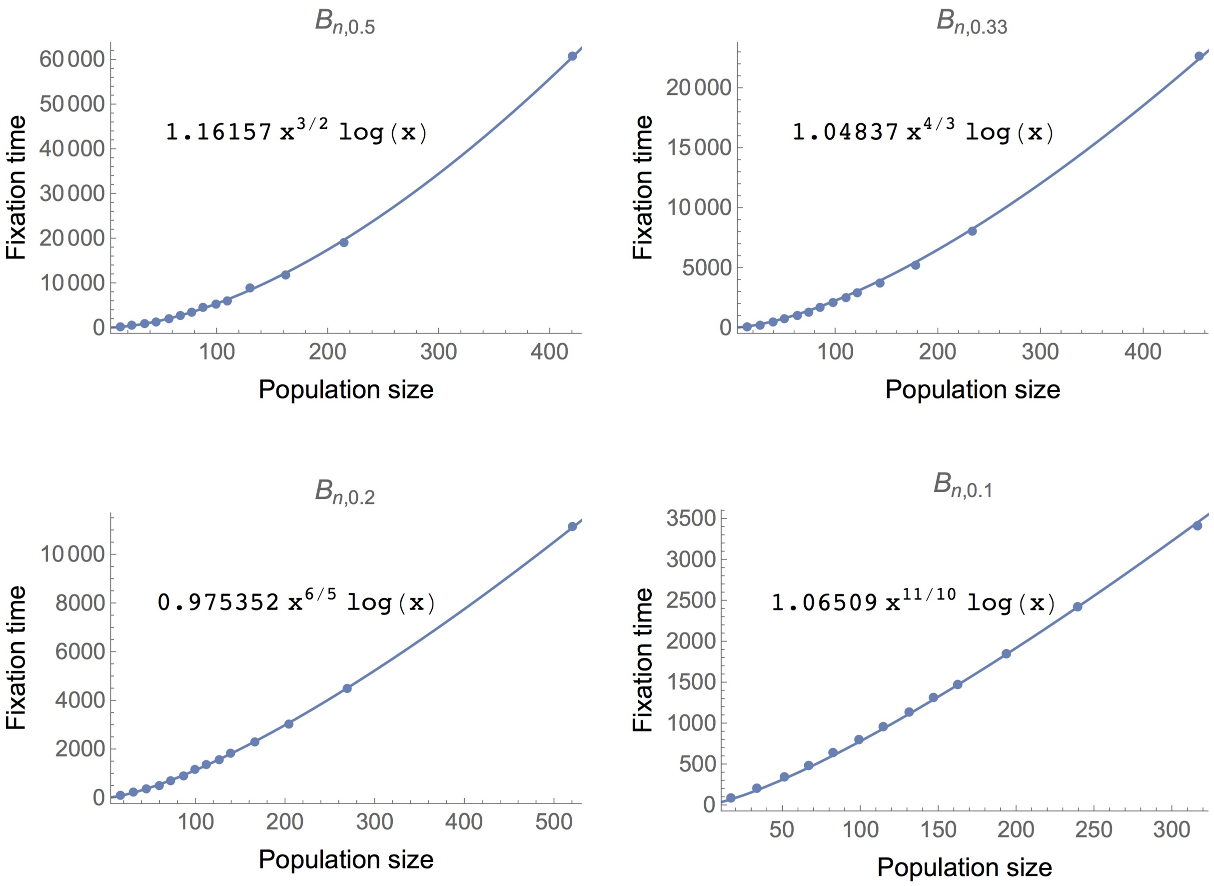

Figure 3.8: **Fixation time for $B_{n,\varepsilon}$ is proportional to $n^{1+\varepsilon}\log n$.** We fix $r = 2$ and consider the $\varepsilon$-Balanced bipartite graphs $B_{n,\varepsilon}$ for $\varepsilon \in \{0.5, 0.33, 0.2, 0.1\}$ and for $n$ up to 500. The dots are exact numerical solutions, the lines are the best fits. The figure confirms that the fixation time $T(B_{n,\varepsilon}, r)$ is proportional to $n^{1+\varepsilon}\log n$ for any $\varepsilon > 0$.

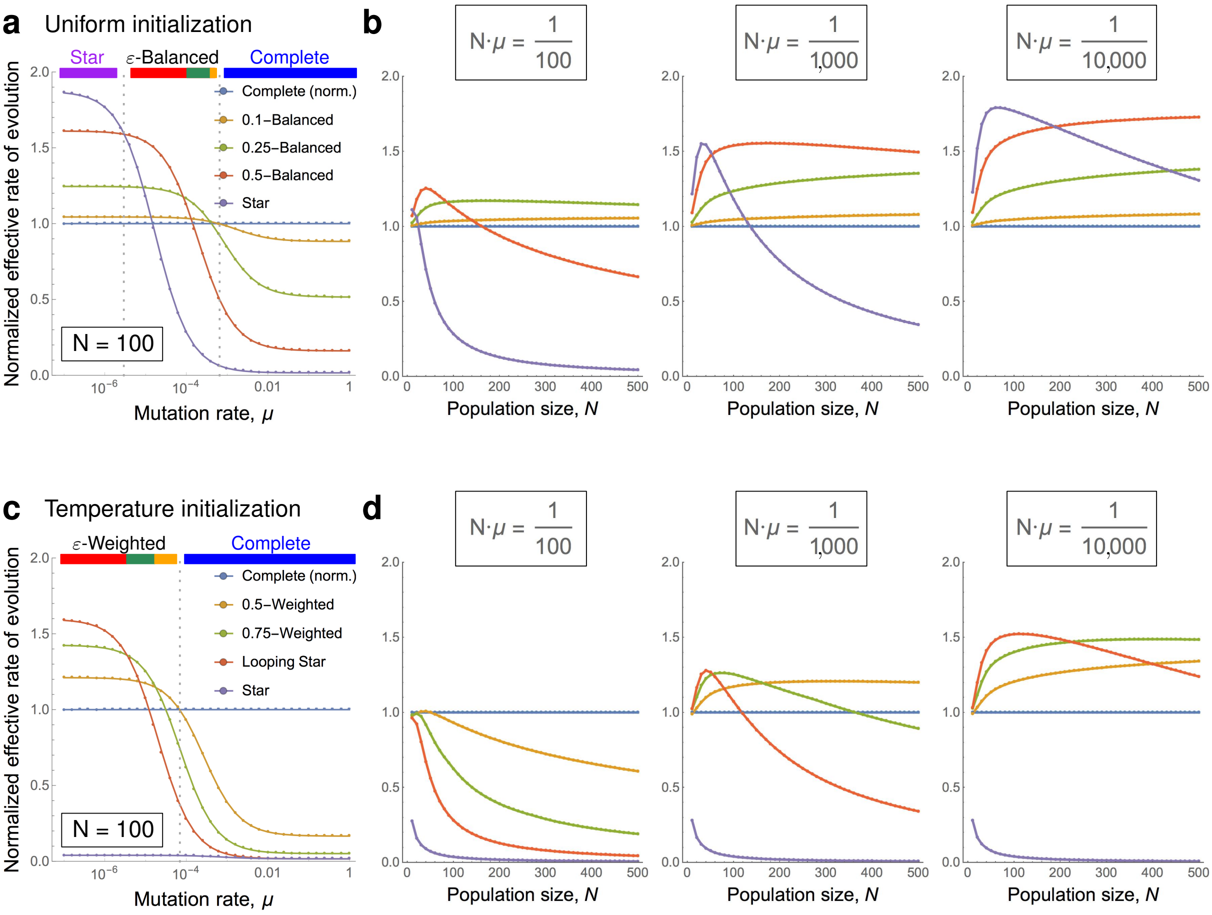

Figure 3.9: **Fig. 4. Effective rate of evolution.** The effective rate of evolution depends on the population size, $N$, the mutation rate, $\mu$, and the population structure. For uniform initialization, we compare five different population structures: the complete graph (blue), $\varepsilon$-Balanced graphs with $\varepsilon \in \{0.1, 0.25, 0.5\}$ (orange, green, red), and the star graph (purple), always showing the relative rate of evolution with respect to the complete graph. **a,** We fix $N = 100$, $r = 1.1$ and vary $\mu = 10^{-7}, \ldots, 10^{0}$. The complete graph has a higher effective rate of evolution if the mutation rate is high ($\mu > 10^{-3}$) and star is favourable if the mutation rate is low ($\mu < 3 \cdot 10^{-6}$). In the intermediate regime, suitable $\varepsilon$-Balanced graphs outperform both of them. **b,** We fix $r = 1.1$ and $N \cdot \mu \in \{10^{-2}, 10^{-3}, 10^{-4}\}$ and vary $N = 10, 20, \ldots, 500$. The star is favourable if mutations are rare ($N \cdot \mu = 10^{-4}$ and $N$ small). Otherwise, suitable $\varepsilon$-Balanced graphs are more efficient. **c, d** Analogous data for temperature initialization. This time we compare the complete graph (blue) and the star (purple) with $\varepsilon$-Weighted bipartite graphs for $\varepsilon \in \{0.25, 0.5, 1\}$ (orange, green, red). The complete graph dominates if mutations are common ($N \cdot \mu = 10^{-2}$). In other cases, $\varepsilon$-Weighted bipartite graphs are preferred. The star is not an amplifier for temperature initialization.

# 4 Amplification under death-Birth updating

Evolutionary graph theory represents population structure of size $N$ by a graph (network) $G_N$ [7; 19; 20; 21; 22; 23]: each individual occupies a vertex, and neighboring vertices mark sites of spatial proximity (see Fig. 4.1a). Mutant spread must respect the structure, in that the offspring of a reproducing individual in one vertex can only move to a neighboring vertex. The Moran process on graphs has two distinct variants:

- In the *Birth-death* Moran process, the death event is conditioned on the Birth event. That is, first an individual is chosen for reproduction and then its offspring replaces a random neighbor (see Fig. 4.1b).
- In the *death-Birth* Moran process, the Birth event is conditioned on the death event. That is, first an individual is chosen for death and then its neighbors compete to fill the vacancy with their offspring (see Fig. 4.1c).

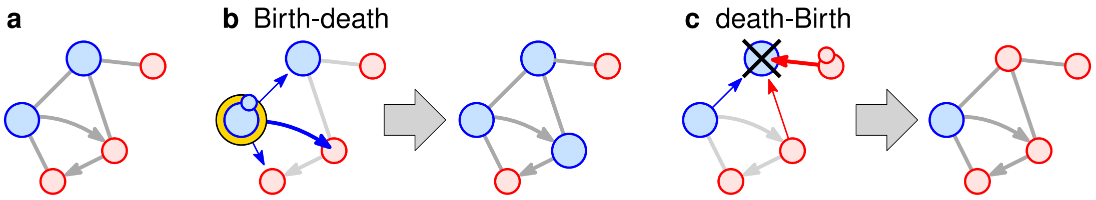

Figure 4.1: **Moran process on graphs. a,** The spatial structure is represented by a graph. Each vertex represents a site and is occupied either by a resident (red) with fitness 1 or by a mutant (blue) with relative fitness $r > 1$. Each edge can be one-way (arrow) or two-way. **b,** In each step of the Birth-death process, one individual is sampled for reproduction proportionally to fitness, and then its offspring replaces a random neighbor. **c,** In each step of the death-Birth process, a random individual dies and then it is replaced by a neighbor sampled proportionally to fitness.

The fixation probability of the invading mutant is a function of its fitness $r$, as well as the graph $G_N$. In alignment with most of the literature, we focus on advantageous mutants, where $r > 1$.

The well-mixed population of size $N$ is represented by a complete graph $K_N$. In the Birth-death Moran process, the fixation probability in the well-mixed population is $\mathrm{fp}(K_N, r) = (1-1/r)/(1-1/r^N)$ [3]. Under death-Birth updating, the fixation probability is $\mathrm{fp}^{\mathrm{dB}}(K_N, r) = (1-1/N) \cdot (1-1/r)/(1-1/r^{N-1})$ [41]. Specifically, as $N \to \infty$, both the expressions converge to $1 - 1/r$.

Amplifiers of natural selection are graphs that increase the fixation probability of the advantageous mutants compared to the well-mixed population [34; 7]. Under Birth-death updating, many amplifying families of graphs have been constructed, such as the Star graph [24; 40; 25], the Complete Bipartite graph [43] and the Comet graph [47], as well as families that guarantee fixation in the limit of large population size [7; 54; 53; 48; 55]. Extensive computer simulations on small populations have also shown that many graphs have amplifying properties [66; 58; 26]. While the above results hold for the Birth-death Moran process, no amplifiers are known for the death-Birth Moran process, and computer-assisted search has found that, under death-Birth updating, most small graphs suppress the fixation probability rather than amplifying it [58].

Here we prove two negative results on the existence of amplifiers under death-Birth updating. Our first result states that the fixation probability in any graph is bounded by $1-1/(r+1)$. Hence, even if amplifiers do exist, they can provide only limited amplification. In particular, there are no families of graphs that would guarantee fixation in the limit of large population size. Our second result states that for any graph $G_N$, there exists a threshold $r^*$ such that for all $r \geq r^*$, the fixation probability is bounded by $\mathrm{fp}^{\mathrm{dB}}(r, K_N)$. Hence, even if some graphs amplify for certain values of $r$, their amplifying property is necessarily transient, and lost when the mutant fitness advantage $r$ becomes large enough. We note that a companion work [59] identifies transient amplifiers among graphs that have weighted edges. Finally, we also study the mixed $\delta$-death-Birth Moran process, for $\delta \in [0, 1]$, under which death-Birth and Birth-death updates happen with rate $\delta$ and $1-\delta$, respectively [33]. We establish analogous negative results for mixed $\delta$-updating, for any fixed $\delta > 0$. Note that as $\delta$ vanishes ($\delta \to 0$), we approach (pure) Birth-death Moran process for which both universal and super amplifiers exist. We find that some of those amplifiers are less sensitive to variations in $\delta$ than other. In particular, certain bipartite structures achieve transient amplification for $\delta$ as big as 0.5.

## 4.1 Model

Here we present the model of Moran process on graphs.

**Population structure.** In evolutionary graph theory, a population structure is represented by a graph that has $N$ sites (nodes), some of which are connected by edges. Each site is occupied by a single individual. The edge from node $u$ to node $v$ represents that the individual at node $u$ can replace the individual at node $v$.

**Directions and weights.** The edges could be undirected (two-way) or directed (one-way) and they could be weighted. Formally, for a pair of nodes $u$, $v$, the weight of an edge $(u, v)$ is denoted by $w_{u,v}$. If the nodes $u$, $v$ are not connected then $w_{u,v} = 0$. In the special case of unweighted graphs, each edge is considered to have weight 1. In the special case of undirected graphs, each edge is two-way. In the most general case of directed graphs with weighted edges, two nodes $u$, $v$ could be interacting in both directions with different weights $w_{u,v} \neq w_{v,u}$. We don't allow self-loops, that is, $w_{u,u} = 0$ for each node $u$.

**Mutant initialization.** Initially, each site is occupied by a single resident with fitness 1. Then a single mutant with fitness $r$ appears at a certain node. This initial mutant node can be selected uniformly at random (*uniform initialization*) or with probability proportional to the turnover rate of each node (*temperature initialization*). Unless specified otherwise, we assume that the initialization is uniform and that the mutation is advantageous ($r > 1$).

**Moran dB and Bd updating.** Once a mutant has appeared, some version of Moran process takes place. Moran process is a discrete-time stochastic process. At each step, one individual is replaced by a copy of another (neighbouring) individual, hence the population size remains constant. Denote by $f(v)$ the fitness of the individual at node $v$. The two prototypical updatings are:

- Moran death-Birth (dB) updating. An individual $v$ is selected uniformly at random for death. The individuals at the neighbouring sites then compete for the vacant spot. Specifically, once $v$ is fixed, an individual $u$ is selected for placing a copy of itself on $v$ with probability proportional to $f(u) \cdot w_{u,v}$. Note that fitness of an individual doesn't play a role in the death step (thus "d" is lower case) but it does play a role in the birth step (thus "B" is upper case).

- Moran Birth-death (Bd) updating. An individual $u$ is selected for reproduction with probability proportional to its fitness $f(u)$. Then it replaces a random neighbor. Specifically, once $u$ is fixed, an individual $v$ is replaced by a copy of $u$ with probability proportional to $w_{u,v}$.

**Mixed $\delta$-dB updating.** The two regimes dB and Bd can be understood as two extreme points of a spectrum. We also consider mixed updating where some steps of the process follow the dB updating while the other ones follow Bd updating. Generally, given a $\delta \in [0,1]$, a $\delta$-dB updating is an update rule in which each step is a dB event with probability $\delta$ and a Bd event with probability $1-\delta$, independently of all the other steps. With this notation, a 1-dB updating is the same as (pure) dB updating and 0-dB updating is the same as (pure) Bd updating.

**Fixation probability.** Given a graph $G$, $r > 1$ and $\delta \in [0,1]$, we denote by $\mathrm{fp}^{\delta}(G,r)$ the fixation probability of a $\delta$-dB updating, when the first mutant is initialized uniformly at random. The complement, that is the probability that the evolutionary trajectory goes extinct, is denoted by $\mathrm{ep}^{\delta}(G,r) = 1 - \mathrm{fp}^{\delta}(G,r)$. Specifically, for $\delta = 1$ we denote the fixation (resp. extinction) probability under pure dB updating by $\mathrm{fp}^{\mathrm{dB}}(G,r)$ (resp. $\mathrm{ep}^{\mathrm{dB}}(G,r)$) and similarly for the pure Bd updating which corresponds to $\delta = 0$.

**Fixation probability on well-mixed populations.** When studying the effect of population structure on the fixation probability, our baseline is the fixation probability on a well-mixed population of the same size. A well-mixed population is modelled by a complete (unweighted) graph $K_N$, without self-loops. Under pure dB and Bd updating there are exact formulas for fixation probability [41; 7]:

$$\mathrm{fp}^{\mathrm{dB}}(K_N, r) = \frac{N-1}{N} \cdot \frac{1-\frac{1}{r}}{1-\frac{1}{r^{N-1}}} \quad \text{and} \quad \mathrm{fp}(K_N, r) = \frac{1-\frac{1}{r}}{1-\frac{1}{r^{N}}}.$$

For $\delta$-dB updating, no analogous formula is known but numerical computations for various values of $N$ and $r$ show that $\mathrm{fp}^{\delta}(K_N, r)$ is essentially indistinguishable from the linear interpolation

$$\widehat{\mathrm{fp}}^{\delta}(K_N, r) = \delta \cdot \mathrm{fp}^{\mathrm{dB}}(K_N, r) + (1-\delta) \cdot \mathrm{fp}(K_N, r)$$

between $\mathrm{fp}^{\mathrm{dB}}(K_N, r)$ and $\mathrm{fp}(K_N, r)$ (see Fig. 4.2). Therefore, in $\delta$-dB updating we use $\widehat{\mathrm{fp}}^{\delta}(K_N, r)$ as the baseline comparison.

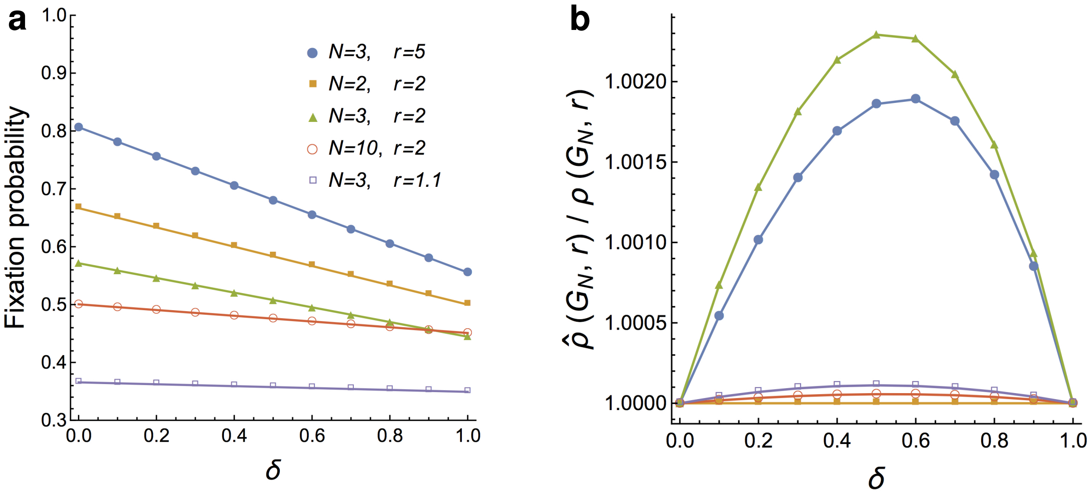

Figure 4.2: **Linear interpolation for $\delta$-dB updating.** On a complete graph $K_N$, the fixation probability $\mathrm{fp}^{\delta}(K_N, r)$ under $\delta$-dB updating is essentially indistinguishable from the linear interpolation $\widehat{\mathrm{fp}}^{\delta}(K_N, r)$ between fixation probability under pure dB and pure Bd updating. **a,** The $x$-axis shows $\delta \in [0, 1]$, the $y$-axis shows the fixation probability $\mathrm{fp}^{\delta}(K_N, r)$ (marks) and the linear interpolation $\widehat{\mathrm{fp}}^{\delta}(K_N, r)$ (lines) for several pairs $(N, r)$. The marks lie almost exactly on the lines. **b,** The ratio $\widehat{\mathrm{fp}}^{\delta}(K_N, r)/\,\mathrm{fp}^{\delta}(K_N, r)$ is well within 1 %, typically even within 0.1 % of 1. The interpolation is exact for $N = 2$.

**Amplifiers of selection.** Given $r > 1$, some population structures enhance the fixation probability of mutants, compared to the well-mixed population, whereas others decrease it. We refer to the former ones as *amplifiers of selection* and to the latter ones as *suppressors of selection*. Formally, given a graph $G_N$ with $N$ nodes and some $r > 1$, we say that $G_N$ is an *$r$-amplifier under dB updating* if $\mathrm{fp}^{\mathrm{dB}}(G_N, r) > \mathrm{fp}^{\mathrm{dB}}(K_N, r)$, where $K_N$ is a complete graph that represents a well-mixed population. If $G$ is an $r$-amplifier under dB updating for all $r > 1$, we call it *universal*. In contrast, graphs that amplify only for some range of values $r \in (1, r^{\star})$ are called *transient*. Similarly, we say that $G_N$ is an *$r$-amplifier under Bd updating* if $\mathrm{fp}(G_N, r) > \mathrm{fp}(K_N, r)$ (note that the baseline is the complete graph $K_N$ under Bd updating) and, for a fixed $\delta \in [0, 1]$, we say that $G_N$ is an *$r$-amplifier under $\delta$-dB updating* if $\mathrm{fp}^{\delta}(G_N, r) > \widehat{\mathrm{fp}}^{\delta}(K_N, r)$.

**Classification of amplifiers by strength: Implied scale of fitness.** Amplifiers can be further classified by strength [34]. We single out bounded amplifiers, linear amplifiers, quadratic amplifiers and super amplifiers. The intuition behind the classification is that, in the limit of large population size, fixation probability can often be written as $1-1/\operatorname{isf}(r)$ for a suitable function $\operatorname{isf}(r)$ of $r$. For instance, for large well-mixed population we have $\operatorname{isf}(r) = r$ (under any of dB, Bd, $\delta$-dB updating) and for large Star graphs under Bd updating we have $\operatorname{isf}(r) = r^2$. The extent to which a large population structure $G$ distorts this fixation probability can thus be classified by looking at the function $\operatorname{isf}(r)$.

Formally, given a family of graphs $\{G_N\}_{N=1}^{\infty}$ of increasing population size, the *implied scale of fitness* of the family is a function $\operatorname{isf}(r)\colon (1,\infty) \to \mathbb{R}$ such that

$$\lim \inf_{N\to\infty} \operatorname{fp}^{\mathrm{dB}}(G_N, r) = 1 - 1/\operatorname{isf}(r).$$

We say that the family is

1. an (at most) *bounded amplifier* if $\operatorname{isf}(r) \le r + c_0$ for some constant $c_0$.
2. an (at least) *linear amplifier* if $\operatorname{isf}(r) \ge c_1 r + c_0$ for some constants $c_1 > 1$, $c_0$.
3. an (at least) *quadratic amplifier* if $\operatorname{isf}(r) \ge c_2 r^2 + c_1 r + c_0$ for some constants $c_2 > 0$, $c_1$, $c_0$.
4. a *super amplifier* if $\operatorname{isf}(r) = \infty$ for all $r > 1$.

These definitions naturally carry over to Bd updating and $\delta$-dB updating. Figure 4.3 illustrates that Stars and certain bipartite graphs are quadratic amplifiers under Bd updating but they cease to amplify under dB updating.

**Remark on the regimes considered.** We intentionally restrict our attention to the following regimes:

1. $r > 1$. If $r = 1$ then $\operatorname{fp}^{\delta}(G_N, r) = 1/N$, regardless of the population structure. If $r < 1$ then $\operatorname{fp}^{\delta}(G_N, r) < 1/N \to_{N\to\infty} 0$ for any $G_N$. Thus we focus on $r > 1$.
2. Uniform initialization. For dB updating, the notions of uniform and temperature initialization coincide, since every node is, on average, selected for death and replaced equally often. Thus we focus on uniform initialization only.
3. No self-loops. For dB updating, self-loops are not biologically realistic: An individual who has just died can not replace itself. Thus we consider graphs with possibly directed and/or weighted edges but without self-loops.

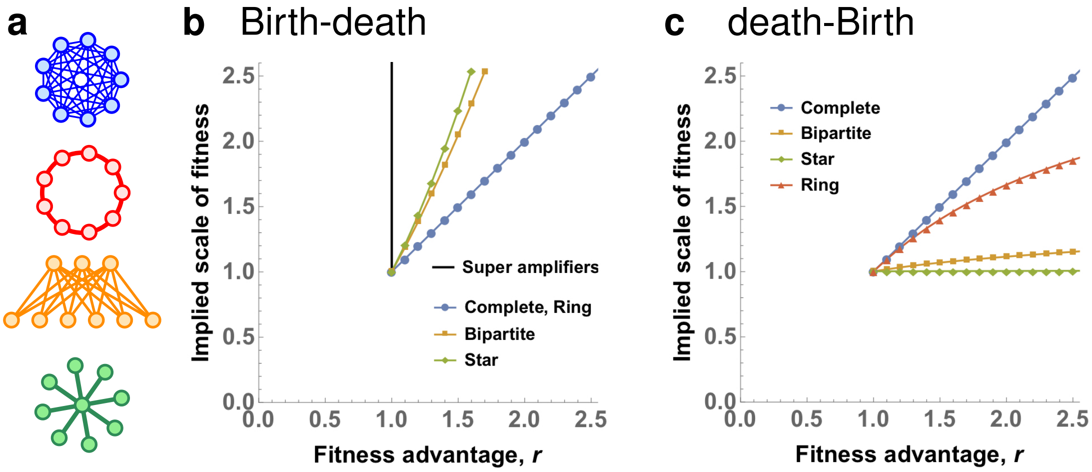

Figure 4.3: **Implied scale of fitness.** The implied scale of fitness for several graph families. **a,** Complete graphs $K_N$, Ring graphs $R_N$, complete Bipartite graphs $B_{\sqrt{N},N-\sqrt{N}}$ and Star graphs $S_N$. **b,** Under Birth-death updating, the Star graphs and the Bipartite graphs are quadratic amplifiers, whereas the Ring graphs are equivalent to Complete graphs. There also exist super amplifiers that guarantee fixation with probability 1 for any $r > 1$. (To model the limit $N \to \infty$ we show values for $N = 400$.) **c,** Under death-Birth updating, none of Bipartite graphs, Star graphs or Ring graphs amplify selection.

## 4.2 Overview of theoretical results

Here we formally state our theorems and their consequences. The first two theorems concern pure dB updating, the other two concern mixed $\delta$-dB updating.

**Theorem 10** (All dB amplifiers are bounded). *Fix $r > 1$. Then for any graph $G_N$ (possibly with directed and/or weighted edges) we have $\mathrm{fp}^{\mathrm{dB}}(G_N, r) \le 1 - \frac{1}{r+1}$.*

**Theorem 11** (All dB amplifiers are transient). *Fix a non-complete graph $G_N$ (possibly with directed and/or weighted edges). Then there exists $r^\star > 1$ such that for all $r > r^\star$ we have $\mathrm{fp}^{\mathrm{dB}}(G_N, r) < \mathrm{fp}^{\mathrm{dB}}(K_N, r)$, where $K_N$ is the complete graph on $N$ vertices. In particular, we can take $r^\star = 2N^2$.*

**Theorem 12** (All $\delta$-dB amplifiers are at most linear). *Fix $r > 1$ and $\delta \in (0, 1]$. Then for any graph $G$ (possibly with directed and/or weighted edges) we have $\mathrm{fp}^{\delta}(G, r) \le 1 - \frac{1}{(r/\delta)+1}$.*

**Theorem 13** (All $\delta$-dB amplifiers are transient). *Fix a non-complete graph $G$ on $N$ vertices (possibly with directed and/or weighted edges) and $\delta \in (0, 1]$. Then there exists*

$r^\star > 1$ *such that for all* $r > r^\star$ *we have* $\mathrm{fp}^\delta(G, r) < \widehat{\mathrm{fp}}^\delta(K_N, r)$*, where* $K_N$ *is the complete graph on* $N$ *vertices.*

**Discussion of our results.** Here we list the implications of our results.

1. Theorem 10 implies that, under dB updating, no unweighted graph is a universal amplifier and a weighted graph can only be a universal amplifier if it is a weighted version of the complete graph $K_N$.
2. Theorem 11 implies that, under dB udpating, there are no superamplifiers and no quadratic or even linear amplifiers.
3. Theorem 12 is a $\delta$-dB analogue of Theorem 10. It implies that, compared to the baseline given by a weighted average between $\mathrm{fp}^{\mathrm{dB}}(K_N, r)$ and $\mathrm{fp}(K_N, r)$, no unweighted graph is a universal amplifier and a weighted graph can only be an $r$-universal amplifier if it is a weighted version of $K_N$.
4. Theorem 13 is a $\delta$-dB analogue of Theorem 11. It implies that if $\delta > 0$ (i.e. we don't have pure Bd updating) then there are no quadratic amplifiers and no superamplifiers. For $\delta = 1$ (pure dB updating), the bound coincides with the one given in Theorem 11. For $\delta = 0$ (pure Bd updating), the bound is vacuous (in the limit $\delta \to 0$ it simplifies to $\mathrm{fp}(G, r) \leq 1$) which is in alignment with the existence of quadratic amplifiers and superamplifiers under (pure) Bd updating.

## 4.3 Proofs

Our proofs rely on Jensen's inequality. For reference purposes, we state it here. Essentially, given a convex (or concave) function $f$ and several real numbers $x_1, \dots, x_k$, Jensen's inequality bounds the (weighted) average of values $f(x_1), \dots, f(x_k)$ by the value that $f$ takes at the (weighted) average of $x_1, \dots, x_k$.

**Claim 1** (Jensen's inequality)**.** Let $a_1, \dots, a_n$ be non-negative real numbers that sum up to 1 and let $f$ be a real continuous function. Then

- If $f$ is convex then

$$\sum_{i=1}^{k} a_i \cdot f(x_i) \geq f\left(\sum_{i=1}^{k} a_i \cdot x_i\right)$$

- If $f$ is concave then

$$\sum_{i=1}^{k} a_i \cdot f(x_i) \leq f\left(\sum_{i=1}^{k} a_i \cdot x_i\right)$$

### 4.3.1 Theorems on dB updating

The key to proving our theorems on dB updating is the following lemma that gives an upper bound on the fixation probability $\mathrm{fp}^{\mathrm{dB}}(G,r)$ on an arbitrary graph (possibly with directed and/or weighted edges), in terms of the average in-degree $d$ and the relative fitness $r>1$ of the mutant. Recall that given a graph $G$ and its node $v$, the *in-degree* of $v$ is the number of nodes $u$ for which there is an edge $(u,v)$. If $G$ is undirected then the in-degree of a node is the same as the degree (the number of neighbors). For any graph $G$, the average in-degree is the same as the average out-degree (and as the average degree if $G$ is undirected).

**Lemma 28.** *Fix $r>1$ and let $G$ be a graph (possibly with directed and/or weighted edges) with average out-degree $d$. Then*

$$\mathrm{fp}^{\mathrm{dB}}(G,r) \leq \frac{d\cdot r}{d\cdot r+d+r-1}.$$

*Proof.* Denote by $u$ the initial node occupied by the mutant and recall that $\mathrm{ep}^{\mathrm{dB}}(u)$ is the extinction probability under dB updating if the initial mutant appears at $u$. Then $\mathrm{ep}^{\mathrm{dB}}(G,r)=\frac{1}{N}\sum_u \mathrm{ep}^{\mathrm{dB}}(u)$.

Denote by $\mathcal{E}^-(u)$ (resp. $\mathcal{E}^+(u)$) the event that in the next step of the dB updating the number of mutants decreases (resp. increases) and by $p^-(u)$ (resp. $p^+(u)$) the corresponding probability. Note that if neither of $\mathcal{E}^-(u)$, $\mathcal{E}^+(u)$ happens, the set of nodes occupied by the mutants stays the same, and if $\mathcal{E}^-(u)$ happens before $\mathcal{E}^+(u)$, the mutants go extinct. Therefore the extinction probability $\mathrm{ep}^{\mathrm{dB}}(u)$ starting from a configuration with a single mutant at node $u$ satisfies

$$\mathrm{ep}^{\mathrm{dB}}(u) \geq \frac{p^-(u)}{p^-(u)+p^+(u)} = \frac{1}{1+\frac{p^+(u)}{p^-(u)}}.$$

We now compute $p^-(u)$ and $p^+(u)$. The number of mutants decreases if and only if we select the single mutant for death, i.e. $p^-(u)=1/N$, for any node $u$. The number of mutants increases if and only if for death we select some node that neighbors $u$ and then we select $u$ for producing an offspring. Hence

$$p^+(u) = \sum_v p^+_{u,v},$$

where

$$p^+_{u,v} = \frac{1}{N} \cdot \frac{r \cdot w_{u,v}}{(r-1) \cdot w_{u,v} + \sum_{u'} w_{u',v}}$$

is the probability that $v$ was selected for death and then $u$ (the only mutant on $G$) was selected to place a copy of itself on $v$.

Now we bound $\mathrm{ep}^{\mathrm{Bd}}(G, r)$ in terms of $p^-(u)$ and $p^+(u)$. In the last step we use Jensen's inequality for a function $f(x) = 1/(1+x)$ which is convex on $x \in (0, \infty)$:

$$\mathrm{ep}^{\mathrm{dB}}(G, r) = \frac{1}{N} \sum_u \mathrm{ep}^{\mathrm{dB}}(u) \geq \frac{1}{N} \sum_u \frac{1}{1 + \frac{p^+(u)}{p^-(u)}} \geq \frac{1}{1 + \frac{1}{N} \sum_u \frac{p^+(u)}{p^-(u)}}.$$

Since $p^-(u) = 1/N$ for all $u$, the right-hand side simplifies and we get

$$\mathrm{ep}^{\mathrm{dB}}(G, r) \geq \frac{1}{1 + \sum_u p^+(u)}.$$

In the rest, we find a tight upper bound on $\sum_u p^+(u)$. We first rewrite each $p^+(u)$ using $p^+_{u,v}$ and interchange the sums to get

$$\sum_u p^+(u) = \sum_u \sum_v p^+_{u,v} = \sum_v \sum_u p^+_{u,v}.$$

We focus on the inner sum. Fix a node $v$ and denote by $s(v) = \sum_{u'} w_{u',v}$ the total weight of all edges incoming to $v$. Using the formula for $p^+_{u,v}$ we obtain

$$\sum_u p^+_{u,v} = \frac{1}{N} \sum_u \frac{r \cdot w_{u,v}}{(r-1) \cdot w_{u,v} + s(v)}.$$

We make three observations. First, the summation has at most $d_{in}(v)$ terms, where $d_{in}(v)$ is the number of incoming edges to $v$. Second, we have $\sum_u w_{u,v} = s(v)$. Third, for fixed $r > 0$ and any $s > 0$, the function $g(x) = \frac{r \cdot x}{(r-1)x+s}$ is concave on $x \in (0, s)$. Therefore, by another application of Jensen's inequality we can write

$$\sum_u p^+_{u,v} \leq \frac{1}{N} \cdot d_{in}(v) \cdot \frac{r \cdot \frac{s(v)}{d_{in}(v)}}{(r-1) \cdot \frac{s(v)}{d_{in}(v)} + s(v)} = \frac{1}{N} \cdot \frac{r \cdot d_{in}(v)}{r - 1 + d_{in}(v)},$$

Finally, summing up over $v$ we obtain

$$\sum_u p^+(u) = \sum_v \sum_u p^+_{u,v} \leq \frac{1}{N} \sum_v \frac{r \cdot d_{in}(v)}{r - 1 + d_{in}(v)} \leq \frac{r \cdot d}{r - 1 + d},$$

where in the last step we yet again used Jensen's inequality, this time for the function $h(x) = \frac{r\cdot x}{r-1+x}$ that is concave on $x \in (0,\infty)$, and the fact that the average in-degree of a graph is the same as its average out-degree.

We conclude by observing that this upper bound on $\sum_u p^+(u)$ yields

$$\mathrm{ep}^{\mathrm{dB}}(G,r) \geq \frac{1}{1+\sum_u p^+(u)} \geq \frac{1}{1+\frac{r\cdot d}{r-1+d}} = \frac{d+r-1}{dr+d+r-1},$$

hence

$$\mathrm{fp}^{\mathrm{dB}}(G,r) \leq 1-\mathrm{ep}^{\mathrm{dB}}(G,r) \leq \frac{d\cdot r}{d\cdot r+d+r-1}$$

as desired □

With the lemma at hand, we can prove the first two Theorems.

**Theorem 10** (All dB amplifiers are transient)**.** *Fix a non-complete graph $G_N$ (possibly with directed and/or weighted edges). Then there exists $r^\star > 1$ such that for all $r > r^\star$ we have $\mathrm{fp}^{\mathrm{dB}}(G_N, r) < \mathrm{fp}^{\mathrm{dB}}(K_N, r)$, where $K_N$ is the complete graph on $N$ vertices. In particular, we can take $r^\star = 2N^2$.*

*Proof.* Recall that

$$\mathrm{fp}^{\mathrm{dB}}(K_N) = (1-1/N)\frac{1-1/r}{1-1/r^{N-1}} \geq (1-1/N)(1-1/r),$$

hence

$$\mathrm{ep}^{\mathrm{dB}}(K_N) \leq \frac{N+r-1}{Nr}.$$

Using Lemma 28, it suffices to show that for all sufficiently large $r$ we have

$$\frac{d+r-1}{dr+d+r-1} > \frac{N+r-1}{Nr}$$

which, after clearing the denominators, is equivalent to

$$r^2\,(N-1-d) - 2r(N-1) - (d-1)(N-1) > 0.$$

Since $G$ is not complete, $d < N-1$ (a strict inequality), hence the coefficient by $r^2$ is positive and the inequality holds for all sufficiently large $r$.

In particular, it is straightforward to check that $r = 2N^2$ is large enough: If $G$ misses at least one edge then $d \leq N-1-\frac{1}{N}$ hence for $r \geq 2N^2$ the right-hand side is at least

$$(2N^2)^2 \cdot \frac{1}{N} - 4N^2(N-1) - N^2 = 3N^2 > 0.$$

□

**Theorem 11** (All dB amplifiers are bounded)**.** *Fix $r > 1$. Then for any graph $G_N$ (possibly with directed and/or weighted edges) we have* $\mathrm{fp}^{\mathrm{dB}}(G_N, r) \leq 1 - \frac{1}{r+1}$.

*Proof.* Using Lemma 28, it suffices to check that

$$\frac{d+r-1}{dr+d+r-1} \geq \frac{1}{r+1}$$

which, after clearing the denominators, is equivalent to $r(r-1) \geq 0$. The equality holds for $r = 1$. $\square$

### 4.3.2 Theorems on $\delta$-dB updating

In order to prove Theorems 13 and 12 we first use a similar technique as before to establish an analogue of Lemma 28 that applies to $\delta$-dB updating.

**Lemma 29.** Fix $r > 1$ and let $G$ be a graph (possibly with directed and/or weighted edges) with average out-degree $d$. Then

$$\mathrm{ep}^{\delta}(G, r) \geq \frac{1}{1 + \frac{dr}{d+r-1} + \frac{1-\delta}{\delta} \cdot \frac{Nr}{N+r-1}}.$$

*Proof.* Denote the initial mutant node by $u$ and, as in Lemma 28, let $p^-(u)$ (resp. $p^+(u)$) be the probability that after a single step of $\delta$-dB updating, the number of mutants in the population decreases (resp. increases).

The values $p^-(u)$ and $p^+(u)$ are weighted averages of the corresponding values under (pure) dB and Bd updating, with weights $\delta$, $1-\delta$. That is,

$$p^-(u) = \delta \cdot \frac{1}{N} \quad + \quad (1-\delta) \cdot \sum_t \frac{1}{N+r-1} \cdot \frac{w_{t,u}}{\sum_{u'} w_{t,u'}}$$

and, using the notation $p^+_{u,v}$ from Lemma 28,

$$p^+(u) = \delta \cdot \sum_v p^+_{u,v} \quad + \quad (1-\delta) \cdot \frac{r}{N+r-1}.$$

As in Lemma 28, we get

$$\mathrm{ep}^{\delta}(G, r) \geq \frac{1}{1 + \frac{1}{N} \sum_u \frac{p^+(u)}{p^-(u)}}.$$

For each fixed $u$, we bound $p^-(u)$ from below by ignoring the whole Bd-contribution. We get $p^-(u) \geq \frac{\delta}{N}$ which yields

$$\mathrm{ep}^{\delta}(G, r) \geq \frac{1}{1 + \frac{1}{\delta} \sum_u p^+(u)}$$

and it remains to bound $\sum_u p^+(u)$ from above. In $\sum_u p^+(u)$, the total Bd-contribution (summed over $u$) equals $(1-\delta)\frac{Nr}{N+r-1}$ and, as in Lemma 28, the total dB-contribution is at most $\delta \cdot \sum_u \sum_v p^+_{u,v} \le \delta \cdot \frac{rd}{r-1+d}$. In total, this yields

$$\mathrm{ep}^{\delta}(G,r) \ge \frac{1}{1+\frac{dr}{d+r-1}+\frac{1-\delta}{\delta}\cdot\frac{Nr}{N+r-1}}$$

as desired. □

Using Lemma 29 we present proofs of Theorems 3 and 4 from the main text.

**Theorem 12** (All $\delta$-dB amplifiers are transient)**.** *Fix a non-complete graph $G$ on $N$ vertices (possibly with directed and/or weighted edges) and $\delta \in (0,1]$. Then there exists $r^\star > 1$ such that for all $r > r^\star$ we have $\mathrm{fp}^{\delta}(G,r) < \widehat{\mathrm{fp}}^{\delta}(K_N,r)$, where $K_N$ is the complete graph on $N$ vertices.*

*Proof.* Let $d$ be the average in-degree of $G$. Since $G$ is not complete, we have $d < N-1$ (a strict inequality).

As in the proof of Theorem 10, recall that $\mathrm{ep}^{\mathrm{dB}}(K_N,r) \le \frac{N+r-1}{Nr}$. Moreover, $\mathrm{fp}(K_N,r) = \frac{1-1/r}{1-1/r^N} \ge 1-1/r$, hence $\mathrm{ep}^{\mathrm{Bd}}(K_N,r) \le \frac{1}{r}$. This yields

$$1-\widehat{\mathrm{fp}}^{\delta}(K_N,r) = \widehat{\mathrm{ep}}^{\delta}(K_N,r) = \delta\cdot \mathrm{ep}^{\mathrm{dB}}(K_N,r) + (1-\delta)\,\mathrm{ep}^{\mathrm{Bd}}(K_N,r) \le \frac{1}{r}+\delta\cdot\frac{r-1}{Nr}$$

and by Lemma 29 it suffices to show that for all sufficiently large $r$ we have

$$\frac{1}{1+\frac{dr}{d+r-1}+\frac{1-\delta}{\delta}\cdot\frac{Nr}{N+r-1}} \ge \frac{1}{r}+\delta\cdot\frac{r-1}{Nr},$$

Since $N$, $d$ and $\delta$ are all fixed, we can consider both sides as functions of $r$. As $r \to \infty$, the left-hand side tends to $\frac{1}{1+d+\frac{1-\delta}{\delta}N}$ while the right-hand side tends to $\frac{\delta}{N}$. In order to conclude, it suffices to show strict inequality between the respective limits:

$$\frac{1}{1+d+\frac{1-\delta}{\delta}N} > \frac{\delta}{N}.$$

After clearing the denominators, this is equivalent to $\delta(N-1-d) > 0$ which indeed holds for any $\delta > 0$ and any non-complete graph $K_N$. □

**Theorem 13** (All $\delta$-dB amplifiers are at most linear)**.** *Fix $r > 1$ and $\delta \in (0,1]$. Then for any graph $G$ (possibly with directed and/or weighted edges) we have* $\mathrm{fp}^{\delta}(G,r) \le 1-\frac{1}{(r/\delta)+1}$.

*Proof.* Since $d \leq N-1 < N$ and $r > 1$, we have

$$\frac{dr}{d+r-1} < \frac{Nr}{N+r-1},$$

hence Lemma 29 gives

$$\mathrm{ep}^{\delta}(G,r) \geq \frac{1}{1+\frac{dr}{d+r-1}+\frac{1-\delta}{\delta}\cdot\frac{Nr}{N+r-1}} > \frac{1}{1+\frac{1}{\delta}\cdot\frac{Nr}{N+r-1}} \geq \frac{\delta}{r+\delta} = \frac{1}{(r/\delta)+1},$$

where the last inequality is equivalent to $\delta \cdot r(r-1) \geq 0$ after clearing the denominators. The result follows. □

## 4.4 Numerical and simulation results

Here we present related numerical and simulation results. First, we illustrate that given a fixed population structure under $\delta$-dB updating, the fixation probability increases as $\delta$ decreases (see Figure 4.4).

Second, we note that even though universal amplification and superamplification are impossible for any $\delta > 0$ due to Theorems 12 and 13, some population structures do achieve reasonable levels of amplification for various combinations of $r$ and $\delta$. Specifically, we consider Star graphs, Bipartite graphs, and Ring graphs of fixed size $N = 10$ and $N = 100$ and show how strongly they amplify, depending on the fitness advantage $r$ of the initial mutant and on the portion $\delta$ of dB updates (see Figure 4.5). We make two observations:

1. First, when $\delta$ is small enough, both Star graphs and Bipartite graphs do amplify selection, for a certain range of $r > 1$. Interestingly, large Bipartite graphs are less sensitive to variations in $\delta$ than Star graphs, and for small $r > 1$ they maintain amplification even for $\delta$ almost as big as 0.5. On the other hand, if $\delta$ is small enough, Star graphs tend to achieve stronger amplification than Bipartite graphs.

2. Second, for any of the six population structures and any fixed $r$, increasing $\delta$ diminishes any benefit that the population structure provides to advantageous mutants. Specifically, there appears to be no regime $(r, \delta)$ where a ring graph would amplify selection.

Third, we illustrate that the upper bound on $\mathrm{fp}^{\mathrm{dB}}(G_N, r)$ given by Theorem 11 is tight for $r = 1$, due to $K_2$, and relatively tight for large $r$, due to $K_N$ with $N$ large (see Figure 4.6).

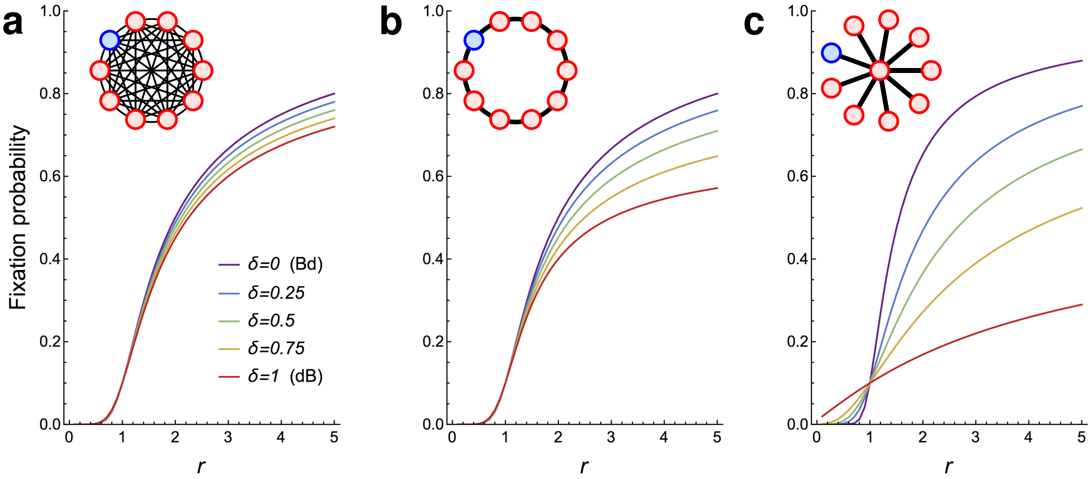

Figure 4.4: **Fixation probability under $\delta$-dB updating.** Three different graphs on $N = 10$ vertices: **a** Complete graph, **b** Ring graph, **c** Star graph. For each $\delta \in \{0, 0.25, 0.5, 0.75, 1\}$ we show the fixation probability under $\delta$-dB updating as a function of $r$. On the latter two graphs, the dependence of the fixation probability on $\delta$ is more pronounced and not roughly linear as is the case for the Complete graph. The Star graph is an amplifier under Bd updating and also a $\delta$-dB amplifier for small $\delta$ (e.g. for $\delta = 0.2$ and $r = 2$ we have $\mathrm{fp}^{\delta}(S_{10}, r) > 0.494 > 0.491 > \widehat{\mathrm{fp}}^{\delta}(K_{10}, r)$) but ceases to be an amplifier for large $\delta$ (e.g. for $\delta = 0.5$ and $r = 2$ we have $\mathrm{fp}^{\delta}(S_{10}, r) < 0.37 < 0.47 < \widehat{\mathrm{fp}}^{\delta}(K_{10}, r)$).

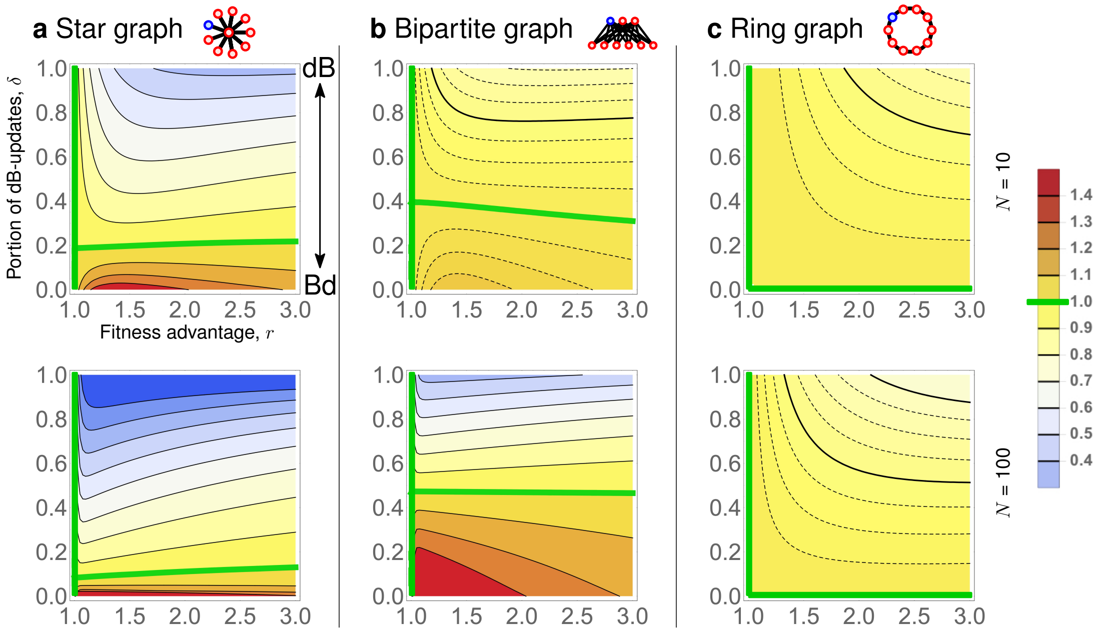

Figure 4.5: **Strength of amplification in terms of $r$ and $\delta$.** **a**, Star graphs, **b**, complete Bipartite graphs with smaller part of size $\sqrt{N}$, and **c**, Ring graphs, of size either $N = 10$ (top row) or $N = 100$ (bottom row). For each of the six graphs, we plot the ratio $\mathrm{fp}^{\delta}(G_N, r)/\widehat{\mathrm{fp}}^{\delta}(K_N, r)$ as a function of the fitness advantage $r$ ($x$-axis) and the portion of dB-updates $\delta$ ($y$-axis). Red (blue) color signifies that the population structure amplifies (suppresses) selection for the given regime $(r, \delta)$. Green curves denote regimes where the ratio equals 1. When $r = 1$, the fixation probability equals $1/N$ regardless of $\delta$ and the population structure. By Theorem 12, all $\delta$-amplifiers are transient, hence the "horizontal" green curves eventually hit the $x$-axis for $r$ large enough. Plotted values were obtained by numerically solving large systems of equations for every $r \in \{1, 1.025, \ldots, 3\}$ and $\delta \in \{0, 0.025, \ldots, 1\}$.

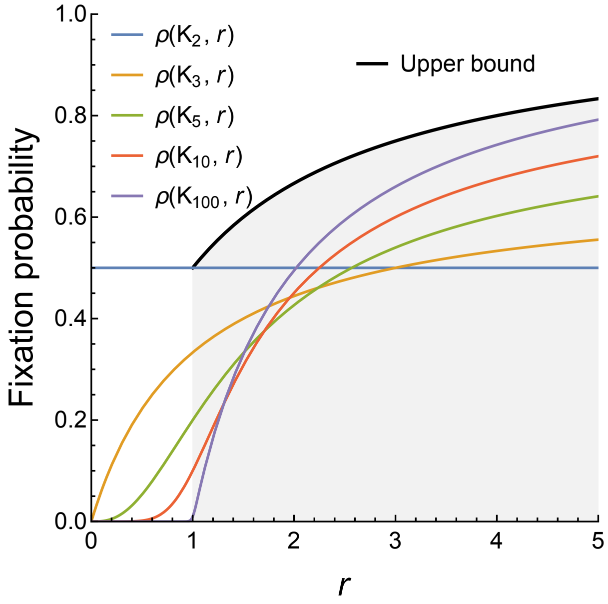

Figure 4.6: **Tightness of the upper bound.** We consider Complete graphs of sizes $N \in \{2, 3, 5, 10, 100\}$ under dB updating. The fixation probability is always below the upper bound given by Theorem 11. For $r = 1$ the bound precisely matches the fixation probability on $K_2$. For large $r$, the bound is relatively tight with respect to large Complete graphs.

# 5 Selection reactors

Here we study certain simple and natural population structures that we call *Selection reactors*. The nodes of a Selection reactor are split into two parts: a reactor-chamber and a pre-chamber (see Figure 5.1). In the reactor-chamber, the individuals directly compete. This is modelled by a complete graph. In the pre-chamber, the individuals do not compete. This is modelled by an empty graph (with self-loops). Finally, with some small probability, an offspring of a reproducing individual can migrate to the other chamber. This is modelled by a complete bipartite graph connecting each reactor-node to each pre-node.

Adjusting the weights of the self-loops and of the edges going across the chambers, we can model any combination of the two migration rates among the two chambers. Doing so, we find that the weight assignment crucially affects the fixation probability and the absorption time of a single advantageous mutant. For certain weight assignments, we present a simple proof that the resulting weighted graphs are superamplifiers. For other weight assignments, we prove that the resulting weighted graphs are superamplifiers with absorption time comparable to that of the Star graph.

## 5.1 Model

Here we formally introduce the model and our notation.

**Moran process on graphs.** Let $G$ be an undirected connected graph with $N$ nodes and edges that are possibly weighted and possibly include self-loops. Each node of $G$ is occupied by a single individual. The individuals come in two types – mutants and residents – that differ in their fitness (reproductive strength). Every mutant has relative fitness $r$ for some fixed $r > 0$ whereas the residents have normalized fitness 1. A *configuration* is a set of nodes occupied by mutants. Given a configuration $C$, Moran Birth-death process is a discrete-time stochastic process that proceeds as follows:

1. (Birth) Pick a random individual for reproduction, proportionally to their fitness. That is, if there are $i$ mutants and $N-i$ residents, each mutant is picked with probability $r/F$ and each resident is picked with probability $1/F$, where $F = i \cdot r + (N-i) \cdot 1$ is the total fitness of the population. Suppose the picked individual occupies node $x$.

2. (Migration) Pick a random edge adjacent to $x$ in $G$, proportionally to its weight. That is, edge $(x, y)$ is picked with probability $w(x, y)/w(x)$, where $w(x, y)$ is the weight of edge $(x, y)$ and $w(x)$ is the sum of weights of all edges adjacent to node $x$. Suppose the picked edge is $(x, y)$.

3. (Death) Replace the individual at node $y$ by a copy of an individual at node $x$.

When $G$ is connected, the process has two absorbing states: one where all individuals are mutants (this is called *fixation*) and one where they are all residents (*extinction*). Once one of those two states is reached, the process terminates.

**Fixation probability and time.** The main quantities of interest are the fixation probability and fixation time. Given a graph $G_N$ on $N$ nodes, a relative fitness $r$ of the mutants, and an initial configuration $C$ of nodes occupied by mutants, we denote by $\mathrm{fp}(G_N, r, C)$ the probability that Moran process reaches fixation, we call this *fixation probability*. Moreover, we denote by $T(G_N, r, C)$ the random variable counting the number of steps required until the process terminates. We use $T(G_N, r, C)$ to define two notions of time: First, we denote by $\mathrm{AT}(G_N, r, C) = \mathbb{E}[T(G_N, r, C)]$ the expected number of steps of the process; we call this (unconditional) *absorption time*. Second, we denote by $\mathrm{CT}(G_N, r, C)$ the expected number of steps among only those evolutionary trajectories that reach fixation (as opposed to reaching extinction); we call this (conditional) *fixation time*.

**Initialization scheme.** Typically, we are interested in the fixation probability and absorption time of a single mutant invading a population of residents, that is, the cases when $|C| = 1$. To that end, we define the fixation probability and time under *uniform initialization* as the average over all $N$ possible starting positions, that is:

$$\mathrm{fp}(G_N, r) = \frac{1}{N} \sum_{|C|=1} \mathrm{fp}(G_N, r, C)$$

and likewise for $\mathrm{AT}(G_N, r)$ and $\mathrm{CT}(G_N, r)$. Alternatively, we could average over the starting positions non-uniformly. One common way is a so-called *temperature initialization*, where each possible initial node $x$ is given weight $\mathsf{T}(x)$ instead of $1/N$. Here $\mathsf{T}(x)$ is the *temperature* of the node, that is, the rate at which the node is replaced by its neighbors when the whole population is residents. Formally, $\mathsf{T}(x) = \sum_y \frac{1}{N} \cdot w(y,x)/w(y)$. All our results hold for both types of initialization.

**Unweighted selection reactor.** First we describe the underlying (unweighted) graph structure of selection reactors. Given integers $N$, $n$ the unweighted selection reactor $\mathrm{USR}(N, n)$ is an unweighted graph such that:

1. There are $N$ *leaf* nodes and $n$ *hub* nodes.
2. Every two hub nodes are connected, every hub node is connected to every leaf node, and every leaf node has a self-loop.

**Firing in and out.** Due to the symmetry of selection reactors, any step of the process that changes the mutant configuration is one of the following four types:

1. A hub node replaces another hub node: We say the hub node *stays*;
2. A hub node replaces a leaf node: We say the hub node *fires out*;
3. A leaf node replaces itself due to a self loop: We say the leaf *loops*;
4. A leaf node replaces a hub node: We say the leaf *fires in*.

Moreover, we say that a fire-in event is *resident fire-in* when the individual reproducing is a resident (and similarly for mutant and/or fire-out events). Finally, we say that a set of nodes (e.g. a hub) is *heterogeneous* if it contains at least one node resident and at least one mutant. Otherwise, we call it *homogeneous*.

**Selection reactors.** Now we can formally define (weighted) selection reactors. Given integers $N$, $n$ and two real numbers $a$, $b$, the selection reactor $\mathrm{SR}(N, n, a, b)$ is a weighted graph such that:

1. There are $N$ *leaf* nodes and $n$ *hub* nodes.

2. Every two hub nodes are connected, every hub node is connected to every leaf node by a weighted edge, and every leaf node has a weighted self-loop. Moreover, the edge weights are such that:

    (a) For any leaf node, once it is selected for reproduction, it fires in the hub (as opposed to self-looping) with probability $f_{\text{in}} = 1/Na$. In other words, we expect to see one fire-in event per $a$ generations.

    (b) For any hub node, once it is selected for reproduction, it fires out to some leaf (as opposed to replacing another hub node) with probability $f_{\text{out}} = 1/nb$. In other words, we expect to see one fire-out event per $b$ generations.

It is readily checked that the properties 2b and 2a are satisfied by assigning weight $w = (n-1)/(N \cdot (nb-1)) = \Theta(1/Nb)$ to each edge connecting a hub node with a leaf node, and by assigning weight $l = n(n-1)(Na-1)/(N \cdot (nb-1)) = \Theta(na/b)$ to each self-loop, respectively.

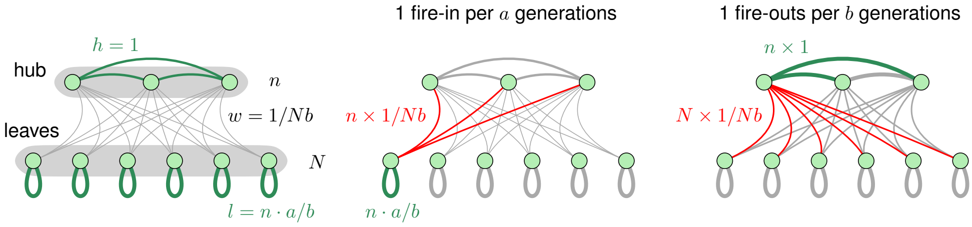

Figure 5.1: Selection reactor SR$(N, n, a, b)$ consists of $N$ leaves with heavy self-loops of weight roughly $na/b$ each, $n$ other nodes forming a well-mixed population called a hub, and $N \cdot n$ light accross edges of weight roughly $1/Nb$ each. Due to the weights, most of the time a leaf will reproduce along its self-loop and a hub node will replace another hub node but occasionally, a leaf fires in the hub or a hub node fires out to a leaf. Specifically, the weights are such that we expect to see roughly $1/a$ fire-in events and roughly $1/b$ fire-out events per generation ($N$ steps).

In what follows, we consider $N$ as the main parameter and we view $n$, $a$, $b$ as functions of $N$.

## 5.2 Overview of theoretical results

Here we present a summary of our theoretical results related to Selection reactors and superamplification.

First, we present a simple proof that, with appropriate weights, Selection reactors become superamplifiers.

**Theorem 14** (Reactors are superamplifiers)**.** *For any fixed $N$, let $n = \sqrt{N}$, $a = N^6$ and $b = N^8$. Then Selection reactors* $\mathrm{SR}(N, n, a, b)$ *are superamplifiers, under both uniform and temperature initialization.*

Second, we show that with different weights, Selection reactors become superamplifiers with absorption time asymptotically comparable to that of the Star graph.

**Theorem 15** (Reactors are fast superamplifiers)**.** *For any fixed $N$, let $n = \sqrt{N}$, $a = \log N$ and $b = \log^2 N$. Then Selection reactors* $\mathrm{SR}(N, n, a, b)$ *satisfy the following, under both uniform and temperature initialization:*

1. *(super-amplifiers)* $\mathrm{fp}(\mathrm{SR}(N, n, a, b), r) \to_{N\to\infty} 1$.
2. *(fast)* $\mathrm{AT}(\mathrm{SR}(N, n, a, b), r) = \mathcal{O}(N^{2+\varepsilon})$*, for any* $\varepsilon > 0$.

We remark that in the proof we also establish that the absorption time is $\mathcal{O}(N^{2+\varepsilon})$ not only in expectation but also with high probability.

**Discussion of the results.**

1. **Connection to Weighted & loopy superamplifiers.** We remark that our proof of Theorem 14 is based on a general construction for turning almost an arbitrary connected graph into a strong amplifier by assigning weights to edges and self-loops, see [48]. When the underlying graph is a Selection reactor, as is our case, the general proof can be greatly simplified.

2. **Tightness of analysis.** Second, we note our analysis in Theorem 15 is tight, for our proof strategy. In order for the absorption time to be $\mathcal{O}(n^{2-\varepsilon})$, the leaves would have to interact with the hub so often that the hub would frequently get interrupted before resolving to a homogeneous state. It is possible that even with this more common interaction the Selection reactors are superamplifiers but our proof strategy does not apply.

## 5.3 Proofs

Here we present formal proofs of our theoretical results.

### 5.3.1 Selection reactors as superamplifiers

The goal of this section is to give a simple proof of Theorem 14.

**Proof idea.** The intuition behind the proof is motivated by two features identified in [48]. First, due to the weights the hub evolves almost completely independently and it interacts with the leaves only rarely. Hence, with probability very close to one, it resolves to a homogeneous state between any two interactions with the leaves. Second, since the hub itself behaves like a well-mixed population, the mutants are fairly likely to win it and extremely unlikely to lose it afterwards: Indeed, for fixed $r$ and as $N$ grows large, the fixation probability $\mathrm{fp}(K_N, r)$ of a single (advantageous) mutant invading a well-mixed population of residents tends to a positive constant $1 - 1/r$, whereas the fixation probability of a single (disadvantageous) resident invading a well-mixed population of mutants $\mathrm{fp}(K_N, 1/r)$ is exponentially small in $N$.

When the hub is relatively small and the interactions between the hub and the leaves are sufficiently rare, we argue that:

1. the first mutant typically appears at a leaf;
2. the mutant repeatedly invades the hub until mutants fixate there; and
3. the mutant hub easily withstands the residents invasions while the mutants win the leaves one by one.

For the formal proof we need the following technical lemma that quantifies the intuition about the behavior of the hub between any two interactions with the leaves.

**Lemma 30.** Consider the selection reactor $\mathrm{SR}(N, \sqrt{N}, N^6, N^8)$ and a configuration where both mutants and residents are present in the hub. Then, except for an event with probability $\mathcal{O}(1/N^5)$, in the next $N^2$ steps the hub resolves to a homogeneous state without interacting with any leaf. Moreover,

1. If the hub contains exactly one mutant then this homogeneous state is “all mutants”, except for a constant probability $p_m \leq 1/r$.
2. If the hub contains exactly one resident then this homogeneous state is “all mutants”, except for an exponentially small probability $p_r \leq 1/r^{N-1}$.

*Proof.* Note that we expect to see one fire-in event per $N^6$ generations, hence a single reproduction event is a fire-in event with probability $\Theta(1/N^7)$. Similarly, it is a fire-out event with probability $\Theta(1/N^9)$. Overall, the probability that a hub interacts with the leaves in one step is $\Theta(1/N^7 + 1/N^9) = \Theta(1/N^7)$.

Consider any period of $N^2$ steps. Within this period, the hub interacts with the leaves with probability at most $N^2 \cdot \Theta(1/N^7) = \Theta(1/N^5)$.

Say that a step of the Moran process is a *hub-step* if it involves a hub node replacing another hub node. The hub is well-mixed population of size $\sqrt{N}$ so the expected number of hub-steps until it resolves to a homogeneous state is $\Theta(\sqrt{N} \cdot \log \sqrt{N})$. A hub-step happens on average once per $\Theta(\sqrt{N})$ steps (once a hub node is selected for reproduction, it fires out with probability only $o(1)$), so overall the hub is expected to resolve in $\tau = \Theta(N \log N)$ steps. By Markov inequality, the probability that the hub does *not* resolve in $3\tau$ steps is at most $1/3 < 1/e$, hence it resolves in $5 \log N \cdot 3\tau = o(N^2)$ steps with probability at least $1 - 1/N^5$.

Altogether, within the $N^2$ steps, the hub doesn't interact with the leaves and becomes homogeneous, apart from trajectories of total probability $\mathcal{O}(1/N^5)$.

Finally, when the hub evolves independently, the fixation probability of a single invading mutant is $\mathrm{fp}(K_N, r) > 1 - 1/r$ whereas the fixation probability of a single invading resident equals $\mathrm{fp}(K_N, 1/r) < 1/r^{N-1}$. □

Now we present a formal proof.

**Theorem 14** (Reactors are superamplifiers)**.** *For any fixed* $N$*, let* $n = \sqrt{N}$*,* $a = N^6$ *and* $b = N^8$*. Then Selection reactors* $\mathrm{SR}(N, n, a, b)$ *are superamplifiers, under both uniform and temperature initialization.*

*Proof.* We proceed in three phases as sketched above. Note that any specific edge connecting a leaf-node and a hub-node is $\Theta(N^8/N^6) = \Theta(N^2)$ more likely to be used for reproduction as a fire-in event than as a fire-out event.

Throughout the proof, whenever the hub becomes heterogeneous, we wait for $N^2$ steps. If, within this period, the hub does not resolve or it interacts with the leaves, we conservatively assume that the evolutionary trajectory goes extinct. By Lemma 30, in each individual case this happens with probability at most $\mathcal{O}(1/N^5) \to 0$.

First, we argue that the initial mutant typically appears at a leaf. Indeed, the probability that the first mutant appears at a hub node is $\sqrt{N}/(N+\sqrt{N}) = \mathcal{O}(1/N^{1/2}) \to 0$, both for uniform and temperature initialization. From now on, we assume that the first mutant appeared at a leaf.

Second, we argue that the mutants typically fixate on the hub. Wait for the first interaction between the initial mutant and any hub node. The interaction could be a fire-in event or a fire-out event.

- With probability $\Theta(1/N^2)$ the event is a fire-out event and the mutants go extinct right away.
- With the remaining probability the event is a fire-in event. Then we wait $N^2$ steps. Within those $N^2$ steps, the hub either gets interrupted (probability $\mathcal{O}(1/N^5)$, we conservatively assume extinction), or the residents win the hub back (probability $1/r$), or the mutants win the hub (probability $1-1/r$).

Altogether, with probability $\Theta(1/N^2 + 1/N^5) = \Theta(1/N^2)$ we go extinct and with probability at most $1/r$ nothing happens, otherwise we win the hub. We repeat this stage $X = \log_{1/r} 1/N = \Theta(\log N)$ times. The probability that "nothing happens" $X$ times in a row is at most $(1/r)^X = 1/N$. During this time, we observe a fire-out event or a hub gets interrupted while heterogeneous with probability at most $\mathcal{O}(\log N/N^2)$. Otherwise the mutants win the hub. Thus the mutants win the hub except for events with probability at most $\mathcal{O}(1/N + \log N/N^2) = \mathcal{O}(1/N) \to 0$.

Third, we argue that the mutants then typically win the leaves. Wait for the first interaction between a hub and a resident leaf. The interaction could be a fire-in event or a fire-out event.

- With probability $p^+ = \Theta(1/N^2)$ the event is a fire-out event and one leaf becomes a mutant.

- With the remaining probability it is a fire-in event. Then we wait $N^2$ steps. Within those $N^2$ steps, the hub either gets interrupted (probability $\mathcal{O}(1/N^5)$, we conservatively assume extinction) or the residents fixate on the hub (probability at most $1/r^{N-1}$)), otherwise the mutants win the hub back. Altogether, with probability $p^- = \mathcal{O}(1/N^5 + 1/r^{N-1}) = \mathcal{O}(1/N^5)$ we conservatively assume extinction, otherwise nothing happens.

Comparing the two probabilities, we see that we win a leaf before assuming extinction, except for an event with probability $p^-/(p^+ + p^-) = \mathcal{O}(N^2/N^5) = \mathcal{O}(1/N^3)$. Repeating this $N-1$ times, we obtain that once we gain the hub, we win all $N-1$ leaves that were initially resident, except for an event with probability at most

$$(N-1) \cdot \frac{p^-}{p^+ + p^-} = \mathcal{O}(1/N^2).$$

Combining all three stages, we conclude that mutants fixate on the whole graph, except for probability $\mathcal{O}(1/N^{1/2} + 1/N + 1/N^2) = \mathcal{O}(1/N^{1/2}) \to 0$. Graphs $\mathrm{SR}(N, \sqrt{N}, N^6, N^8)$ are thus superamplifiers. □

### 5.3.2 Selection reactors as fast superamplifiers

The goal of this section is to prove Theorem 15.

**Proof idea.** Here we informally describe the idea behind the proof that the selection reactors are fast super-amplifiers.

The relative sizes of the hub and leaf subpopulations and the weights are constructed such that the following conditions are satisfied:

(i) there are many more leaves than hub nodes,

(ii) leaves interact with the hub only rarely, and

(iii) any specific leaf is more likely to fire in the hub than to be fired to from any vertex in the hub.

The intuition behind the proof of super-amplification is that the following events happen with high probability:

1. By (i), the first mutant appears at a leaf.
2. By (iii), the mutant repeatedly invades the hub before being eliminated.
3. By (ii), the hub evolves mostly independently and since any invading mutant has a fitness advantage, the mutants eventually fixate on the hub.
4. Next, the mutants need to win the leaves, one by one. During this period, two forces are at play:
    - If the hub is all mutants, the next active step is typically a resident leaf firing in and making the hub heterogeneous, but occasionally a mutant from the hub fires out to a resident leaf, thereby making it mutant.
    - If the hub is heterogeneous, typically it resolves to an all-mutants state before any hub-node fires out to a leaf, however rarely it might "leak" – that is, it might fire out a resident (while in a heterogeneous state), thereby possibly turning a leaf from a mutant to a resident.

    By construction, we can relate the strength of those two forces. By (ii), hub leaks only rarely. On the other hand, by (iii), gaining a mutant leaf is "only" uncommon (not rare). In total, we gain a mutant leaf (due to some hub-node firing out while the hub is all mutants) more frequently than we lose one (due to the hub leaking).
5. One by one, all leaves become mutants and the mutants fixate.

Since the aforementioned sequence of events occurs with high probability, the selection reactors are super-amplifiers.

Moreover, we can quantify how much time each stage takes. We show that winning the hub starting from a single initial mutant in a leaf takes, on average, roughly $N^2$ steps. Moreover, we show that winning the $i$-th leaf takes, on average, roughly $N^2/i$ steps, hence all leaves are won in roughly $\sum_{i=1}^{N} N^2/i \approx N^2 \log N$ steps, on average. Therefore, the total number of steps is of the order of roughly $N^2 \log N = \mathcal{O}(N^{2+\varepsilon})$ for any $\varepsilon > 0$.

**Mathematical tools.** We recall several mathematical facts that we will rely on in our proofs.

**Lemma 31.** Let $f(n)$ be an increasing function. Then

1. If $f(n) = o(n)$ then $\lim_{n\to\infty}(1-1/n)^{f(n)} = 1$;
2. If $n = o(f(n))$ then $\lim_{n\to\infty}(1-1/n)^{f(n)} = 0$; and
3. If $f(n) = \Theta(n)$ then $\liminf_{n\to\infty}(1-1/n)^{f(n)} > 0$ and $\liminf_{n\to\infty}(1-1/n)^{f(n)} < 1$.

**Lemma 32** (Bernoulli inequality)**.** For any positive integer $n$ and any $x \in (0,1)$ we have $(1-x)^n \geq 1 - xn$.

**Lemma 33** (Markov inequality)**.** Let $X$ be a non-negative random variable with expectation $\mu$ and $c > 1$ a real constant. Then $\mathbb{P}[X \geq c \cdot \mu] \leq \frac{1}{c}$.

Finally, we recall several basic properties of a Moran process on a complete graph of size $N$ with mutants of fitness $r$: First, until the process resolves (that is, as long as it is heterogeneous), we are always $r$ times more likely to gain a mutant than to lose a mutant. Second, advantageous mutants have at least a constant fixation probability whereas disadvantageous mutants have exponentially small fixation probability. Third, for advantageous mutants both the (unconditional) absorption time and the (conditional) fixation time are of the order of $\Theta(N \log N)$. Fourth, the number of steps is short not only on average but also with high probability.

**Lemma 34.** Fix $N \in \mathbb{N}$ and $r \neq 1$. Consider the Moran Birth-death process with mutants of fitness $r$ on a complete graph $K_N$. Then

1. Suppose there are currently $i$ mutants for $1 \leq i \leq n-1$. Then $p^+(i) = r \cdot p^-(i)$, where $p^+(i)$ (resp. $p^-(i)$) is the probability that the number of mutants increases (resp. decreases) in one step. As a consequence, we expect to visit states with $i$ mutants at most a constant number of times (for each $i$).
2. $\mathrm{fp}(K_N, r) = \frac{1-1/r}{1-1/r^N}$. Specifically, $\mathrm{fp}(K_N, r) > 1 - 1/r$ when $r > 1$, and $\mathrm{fp}(K_N, r) < r^{N-1}$ when $r < 1$.
3. $\mathrm{AT}(K_N, r) = \frac{r+1}{r} \cdot N \log N + o(N \log N) = \Theta(N \log N)$ and $\mathrm{CT}(K_N, r) = \frac{r+1}{r-1} \cdot N \log N + o(N \log N) = \Theta(N \log N)$ when $r > 1$.
4. For any $a > 0$ we have $\mathbb{P}[T(K_N, r) > 2 \log a \, \mathrm{AT}(K_N, r)] < 1 - 1/a$, where $T(K_N, r)$ is a random variable counting the number of steps until the process resolves.

*Proof.* This is well known. For the last item, let $\tau = \mathrm{AT}(K_N, r)$. By Markov inequality applied to $T(K_N, r)$ we have $Pr[T(K_N, r) > 2\tau] < 1/2$. Hence the process terminates within each block of length $2\tau$ with probability at least $1/2 \geq 1/e$. The probability that the process does not terminate in any of $a$ such consecutive blocks is then at most $(1/e)^{\log a} = 1/a$ as desired. □

Now we turn to proving Theorem 15. We remark that our results apply as long as the three functions $n(N)$, $a(N)$, $b(N)$ are increasing and unbounded, and satisfy conditions:

1. $n = o(N)$: there are many more leaves than hub nodes;
2. $(\log N)^3 = o(n)$: hub is not too small;
3. $a = o(b)$: fire-ins are more common than fire-outs; and
4. $b = o(N)$: fire-ins and fire-outs occur at least once in a while.

For specificity, we set $n = \sqrt{N}$, $a = \log N$, $b = \log^2 N$ and use a shorthand notation $\mathrm{SR}(N)$ for selection reactor $\mathrm{SR}(N, \sqrt{N}, \log N, \log^2 N)$.

Recall that a *configuration*, denoted by $C$, is the set of nodes that are occupied by mutants.

First we show that the hub is interrupted by fire-ins so rarely that it is still biased towards gaining mutants by a constant factor.

**Lemma 35** (Hub with fire-ins is still biased)**.** Suppose the hub is heterogenous, that is, the number $i$ of mutants in the hub satisfies $1 \leq i \leq n-1$. Let $p^+$ (resp. $p^-$) be the probability that the number of mutants in the hub increases (resp. decreases) in one step and let $r' = \frac{1}{2}(r+1)$. Then $p^+ \geq r' \cdot p^-$ for $N$ large enough.

*Proof.* Simple counting. Let $F$ be the total fitness of the population. We can always gain a mutant in the hub due to a hub mutant replacing a hub resident. Conversely, we lose a mutant in the hub either due to a hub resident replacing a hub mutant or due to a resident leaf firing in the hub. Regardless of how many mutant leaves there are, we get:

$$p^+ \geq \frac{ri}{F} \cdot (1 - f_{\mathrm{out}}) \cdot \frac{n-i}{n-1}$$

and

$$p^- \leq \frac{n-i}{F} \cdot (1 - f_{\mathrm{out}}) \cdot \frac{i}{n-1} + \frac{N}{F} \cdot f_{\mathrm{in}} \cdot \frac{i}{n},$$

hence

$$p^+/p^- \geq \frac{r(1-f_{\text{out}})(n-i)}{(1-f_{\text{out}})(n-i)+1/a} \geq \frac{r(1-f_{\text{out}})}{(1-f_{\text{out}})+1/a} \rightarrow_{N\to\infty} r,$$

since $f_{\text{out}} \to 0$ and $a \to \infty$ as $N \to \infty$. In particular, from some $N$ on the ratio is at least $r' = (r+1)/2$. □

As a corollary, we obtain that any time the hub becomes heterogeneous, it quickly resolves to a homogeneous state and, with constant probability, that state is all-mutants.

**Lemma 36.** There exists a positive constant $c_0$ such that: For large enough $N$, if the hub is heterogeneous, it resolves to an all-mutants state within $\mathcal{O}(N \log n)$ steps with probability at least $c_0$. Moreover, if the hub is heterogeneous then it resolves (to one of the two states) within $\mathcal{O}(N \log^2 n)$ steps, with high probability.

*Proof.* For the first claim, by Lemma 35, the fixation probability for large enough $N$ is at least as high as on the complete graph with mutant fitness advantage $r'$. By Lemma 34, this fixation probability is at least $1 - 1/r'$. To bound the number of steps, note that a step of the Moran process changes a configuration on the hub with probability at least $n/F \cdot (1 - f_{\text{out}}) = \Omega(n/N)$, hence by Lemma 34, fixation happens within $\mathcal{O}(n \log n \cdot N/n) = \mathcal{O}(N \log n)$ steps with constant probability. The second claim now follows from Lemma 34. □

We deduce that starting from a configuration with at least one mutant leaf, with high probability we quickly win the hub.

**Lemma 37** (We gain the hub)**.** If the current configuration $C$ includes a leaf then, with high probability, we reach a configuration that contains all hub nodes within $\mathcal{O}(N^2 b)$ steps.

*Proof.* Let $C$ be any configuration containing at least one leaf, call that leaf $x$.

First we show that $C$ will contain $x$ for $s_0 = N^2\sqrt{ab}$ steps with high probability. Indeed, the probability that $x$ gets eliminated in a single step is at most $p_0 = n/N \cdot f_{\text{out}} \cdot 1/N = 1/N^2 b$. The probability that it survives for all $s_0$ steps is $(1-p_0)^{s_0} \geq 1 - p_0 s_0 = 1 - \sqrt{a/b} \to 1$ as desired (recall Lemma 32 and the relation $a = o(b)$).

Now we show that $x$ invades hub fairly often. The probability that in a single step $x$ gets selected for reproduction and fires in the hub is at least $p_1 \geq 1/(r(N+n)) \cdot f_{\text{in}} =$

$\Omega(1/N^2a)$. Consider $s_1 = N^2a$ steps. Then the probability that, in the next $s_1$ steps, $x$ fires-in at least once (and thus the mutants invade the hub) is at least $1-(1-p_1)^{s_1}$. By Lemma 31, this is at least some positive constant $c_1$ as $N \to \infty$.

Once $C$ contains a hub node, by Lemma 36 there is at least a constant probability $c_2$ that it will contain all hub nodes at some point in the next $\mathcal{O}(N \log n)$ steps. In total, the mutants win the hub within $s = s_1+s_2 = \mathcal{O}(N^2a)$ steps with constant positive probability $c = c_1 \cdot c_2$.

Finally, we amplify this constant probability into a high probability. Consider $\sqrt{b/a}$ stages, each consisting of $s_1 + s_2$ steps. Note that the number of stages grows unbounded ($a = o(b)$) and in total the stages take $N^2a \cdot \sqrt{b/a} = N^2\sqrt{ab}$ steps, so $x$ stays a mutant throughout all stages, with high probability. Within each stage, the mutants win the hub with at least a constant probability $c$. This implies that the mutants win the hub in one of the stages with probability at least $1-(1-c)^{b/a} \to 1$, that is, with high probability. □

Next we show that once we win the hub, it remains mostly mutants for long enough time. Specifically, we use $k = 1 + 4 \cdot \log_r N = \Theta(\log N)$ as the threshold number of residents allowed in the hub.

**Lemma 38** (Hub stays mostly mutants)**.** Set $k = 1 + 4 \cdot \log_r N = \Theta(\log N)$. If at any point a configuration $C$ contains all $n$ hub vertices, then, with high probability, the hub contains at most $k = \Theta(\log N)$ resident vertices in each of the next $N^3$ steps.

*Proof.* This follows from the behavior of the Moran process on a complete graph and some crude bounds. Fix $r' = (r+1)/2$.

Suppose that a single hub vertex has just become a resident. Denote by $p$ the probability that $k$ hub nodes become resident before the hub becomes all mutants. Recall that by Lemma 35, as long as the hub is heterogeneous, there is at least a constant bias $r' = \frac{1}{2}(r+1)$ against gaining more residents. By Lemma 34, the fixation probability with bias exactly equal to $r'$ would be $\mathrm{fp}(K_k, 1/r') < (1/r')^{k-1} = 1/N^4$. Thus $p \le 1/N^4$.

Since by Lemma 31 we have $(1-p)^{N^3} > (1-1/N^4)^{N^3} \to_{N\to\infty} 1$, the following event happens with high probability: The first $N^3$ times the hub becomes heterogeneous, it reaches a state with all mutants before reaching a state with $k$ residents. Since each such stage takes at least one step, this takes at least $N^3$ steps in total as desired. □

Now we prove that as long as the hub is mostly mutants, we are much more likely to gain a mutant leaf (due to a hub firing out while fully homogeneous) than to lose a mutant leaf (due to a hub firing out a resident while heterogeneous). We call the latter event *leaking* and consider it first.

**Lemma 39** (Leaking is rare)**.** If the hub contains at most $k = \Theta(\log N)$ residents throughout the next $\Theta(N \log n \log b)$ steps, then it leaks a resident before becoming all mutants with probability of the order of at most $\mathcal{O}(1/b)$.

*Proof.* Since there are at most $k$ residents in the hub, the probability that a single step is a resident leak is at most $p_{\text{leak}} = \frac{k}{F} \cdot f_{\text{out}} \leq \frac{k}{Nnb}$. Consider the next $s = \Theta(N \log n \log b)$ steps. With probability at least $1 - 1/b$, the hub becomes all mutants within those $s$ steps (see Lemma 36 and Lemma 34). On the other hand, by Lemma 32 the hub leaks a resident within those $s$ steps with probability at most

$$\begin{aligned} 1 - (1 - p_{\text{leak}})^s \leq 1 - (1 - p_{\text{leak}} \cdot s) = p_{\text{leak}} \cdot s &= \mathcal{O}(k \log n \log b/(nb)) \\ &= \mathcal{O}((\log N)^3/nb) = \mathcal{O}(1/b), \end{aligned}$$

where the last two steps hold provided that $b = \mathcal{O}(N)$ and $(\log N)^3 = o(n)$. In total, except for events with probability at most $1/b + 1/b = \mathcal{O}(1/b)$, the hub resolves before leaking a resident. ☐

On the other hand, as mentioned earlier, gaining a mutant leaf is more common.

**Lemma 40** (Occasionally we gain a leaf)**.** If the hub is homogeneous then it gains a mutant leaf before getting heterogeneous with probability of the order of at least $\Omega(a/b)$.

*Proof.* This is simple computation. Note that the only edges whose endpoints have different types are those that connect a resident leaf to a mutant node in the hub. Fix any such edge and denote the resident leaf by $x$ and the hub mutant by $y$. Let $F$ be the total fitness. We directly compare the probabilities that in a single step the edge is used for reproduction in the direction $x \to y$ or $y \to x$. We get

$$\mathbb{P}[x \to y] = \frac{1}{F} \cdot f_{\text{in}} \cdot \frac{1}{n} = \frac{1}{FNna} \quad \text{and} \quad \mathbb{P}[y \to x] = \frac{r}{F} \cdot f_{\text{out}} \cdot \frac{1}{N} = \frac{r}{FNnb}$$

hence

$$\frac{\mathbb{P}[y \to x]}{\mathbb{P}[y \to x] + \mathbb{P}[x \to y]} = \frac{ra}{ra + b} = \Omega(a/b).$$

Since this holds for every edge, it holds overall too. ☐

Combining the last two lemmas, we show that as long as the hub is mostly mutants, it is overwhelmingly more likely to gain a mutant leaf than to lose one.

**Lemma 41** (Gaining leaves before losing them)**.** As long as the hub contains at most $k = \Theta(\log N)$ residents, with high probability, we gain a mutant leaf before losing one.

*Proof.* By Lemma 39 the hub leaks with probability $p_{\leftarrow}$ that is of the order of at most $\mathcal{O}(1/b)$ and otherwise it resolves to all mutants. Once it resolves, by Lemma 40 we gain a mutant leaf with probability $p_{\rightarrow}$ that is of the order of at least $\Omega(a/b)$, otherwise the hub becomes heterogeneous again. From this point on, the situation repeats. Hence the probabilities of losing and gaining a mutant leaf are in ratio $p_{\leftarrow} : (1-p_{\leftarrow}) \cdot p_{\rightarrow}$. Computing the probability $p$ of losing a leaf before winning it we get

$$p = \frac{p_{\leftarrow}}{p_{\leftarrow} + (1-p_{\leftarrow}) \cdot p_{\rightarrow}} = \mathcal{O}\left(\frac{1/b}{1/b + a/b}\right) = \mathcal{O}(1/a) \rightarrow_{N \to \infty} 0,$$

hence $1 - p \rightarrow_{N\to\infty} 1$ as desired. □

Now we combine the ingredients for the proof of Theorem 15.

**Theorem 15** (Reactors are fast superamplifiers)**.** *For any fixed* $N$*, let* $n = \sqrt{N}$*,* $a = \log N$ *and* $b = \log^2 N$*. Then Selection reactors* $\mathrm{SR}(N, n, a, b)$ *satisfy the following, under both uniform and temperature initialization:*

1. *(super-amplifiers)* $\mathrm{fp}(\mathrm{SR}(N, n, a, b), r) \rightarrow_{N\to\infty} 1$.
2. *(fast)* $\mathrm{AT}(\mathrm{SR}(N, n, a, b), r) = \mathcal{O}(N^{2+\varepsilon})$*, for any* $\varepsilon > 0$.

*Proof.* Recall that the selection reactor consists of $N$ leaf nodes and $n = o(N)$ hub nodes.

First, we argue that the initial mutant appears at a leaf node with high probability, under both uniform and temperature initialization. For uniform initialization this follows immediately from $n = o(N)$. For temperature initialization, we use the fact that the leaves mostly self-loop: For any fixed leaf $x$ we have $\mathsf{T}(x) \geq \frac{1}{N+n} \cdot (1 - 1/Na)$, hence

$$\sum_{x \text{ leaf}} \mathsf{T}(x) \geq \frac{N}{N+n} \cdot (1 - 1/Na) \rightarrow_{N\to\infty} 1.$$

Since the total temperature of the population is 1, this implies that the overwhelming majority of the temperature is concentrated on the leaves, hence the initial mutant appears at a leaf vertex with high probability.

Next, Lemma 37 implies that the mutants invade the hub and win it over in $N^2 \cdot b$ steps (with high probability). After that, Lemma 38 implies that in each of the subsequent $N^3$ steps, the hub will contain at most $k = \mathcal{O}(\log N)$ residents (with high probability).

Now we consider a Markov chain $M$ with $2N+3$ states labelled $M_i$, $H_i$ for $0 \le i \le N$ and a single extra state $X$ (see Figure 5.2).

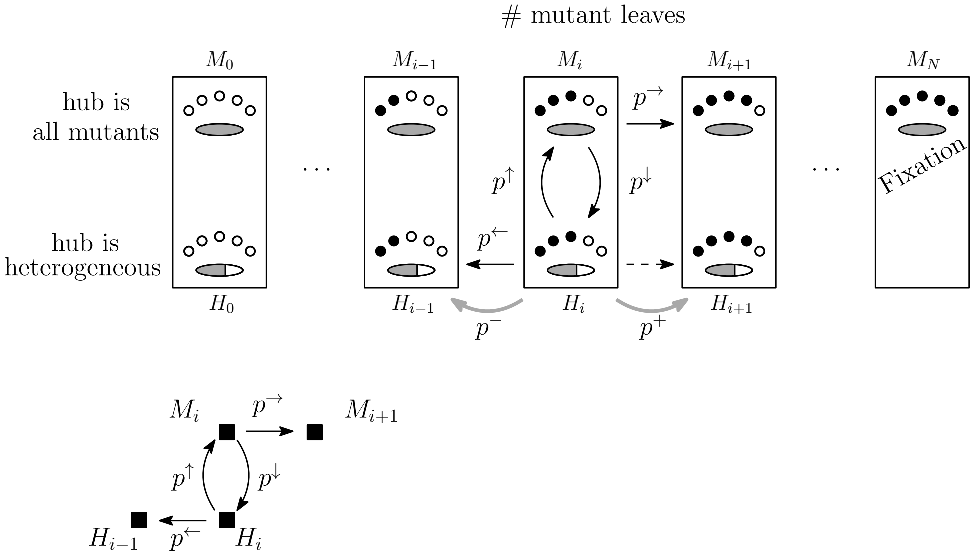

Figure 5.2: Graphical representation of the Markov chain in step 4. Denote by $M_i$ (resp. $H_i$) the states when we have $i$ mutant leaves and the hub is all mutants (resp. heterogeneous and containing at most $k = \Theta(\log N)$ residents) and assume our current state is $H_i$. Then the state keeps toggling between $H_i$ and $M_i$ until either we move right via $M_i \to M_{i+1}$ or $H_i \to H_{i+1}$ (uncommon) or left via $H_i \to H_{i-1}$ (rare). Since uncommon is more frequent than rare, we gain a mutant leaf more often than we lose it.

The states $M_i$ represent the configurations in which all $n$ hub vertices and $i$ of the leaf vertices are mutants. Specifically, $M_1$ corresponds to the current state and $M_N$ corresponds to fixation. The states $H_i$ represent configurations in which $n-1$ hub vertices and $i$ of the leaf vertices are mutants. The state $X$ represents the configurations that have more than $k$ residents in the hub (including extinction). Note that by Lemma 38, we avoid state $X$ in the next $N^3$ steps with high probability. Thus it suffices to show that those evolutionary trajectories that avoid $X$ fixate within the next $N^3$ steps with high probability. From now on we deal with such trajectories only.

Observe that when in a state $M_i$, there are only three options for the next state:

1. $M_i \to M_i$ when the configuration does not change;

2. $M_i \to M_{i+1}$ when we gain a leaf;

3. $M_i \to H_i$ when a resident fires in the hub;

and only the latter two of them change the state. Moreover, Lemma 40 implies that the probability $p_{\to}$ of moving to state $M_{i+1}$ (rather than to $H_i$) is of the order of at least $a/b$.

Similarly, when in a state $H_i$, there are four options for the next state:

1. $H_i \to H_i$ when the configuration does not change;

2. $H_i \to H_{i+1}$ when the hub leaks a mutant and we gain a mutant leaf;

3. $H_i \to H_{i-1}$ when the hub leaks a resident and we lose a mutant leaf;

4. $H_i \to M_i$ when the hub becomes all mutants before leaking;

and only the latter three of them change the state. Lemma 39 implies that the probability $p_{\leftarrow}$ of moving to state $H_{i-1}$ (rather than to $M_i$) is of the order of at most $1/b$.

A typical evolutionary trajectory starting at $H_i$ keeps toggling between $H_i$ and $M_i$ (potentially staying at each state for some time) but eventually it leaves the set $\{H_i, M_i\}$ either along $M_i \to M_{i+1}$ or along $H_i \to H_{i-1}$. Lemma 41 implies that, upon winning the hub, with high probability $p_N \to 1$, the trajectory will leave into $\{H_{i+1}, M_{i+1}\}$ (along either $M_i \to M_{i+1}$ or perhaps even $H_i \to H_{i+1}$) rather than into $\{H_{i-1}, M_{i-1}\}$ (along $H_i \to H_{i-1}$), no matter if it is currently at $H_i$ or at $M_i$.

Hence the fixation probability in the Markov chain $M$ can be bounded from below by the fixation probability on a one-dimensional Markov chain with bias $\frac{p_N}{1-p_N}$. By Lemma 34, the fixation probability for fixed bias $r^\star$ (and any $N$) is at least $1 - 1/r^\star$. As $N \to \infty$, we have $p_N \to 1$, hence

$$\mathrm{fp}(\mathrm{SR}(N), r) \geq 1 - \frac{1-p_N}{p_N} \to_{N\to\infty} 1$$

and the selection reactors are indeed super-amplifiers.

Regarding the fixation time, note that for any $i = 0, \dots, N-1$ we leave the set $\{H_i, M_i\}$ in one step with probability $p_i$ satisfying

$$p_i \geq \frac{r(n-k)}{F} \cdot f_{\mathrm{out}} \cdot \frac{N-i}{N} = \Omega((N-i)/N^2 b),$$

hence the expected number of steps to leave the set $\{H_i, M_i\}$ is of the order of at most

$$\frac{1}{p_i} \leq \frac{N^2 b}{N-i}.$$

By Lemma 34, upon leaving the set $\{H_i, M_i\}$, we are expected to revisit it only a constant number of times. Summing over $i = 0, \ldots, N-1$ we get that the expected number of steps to fixate on the whole graph after fixating on the hub is of the order of at most

$$\sum_{i=0}^{N-1} \frac{N^2 b}{N-i} = N^2 b \log N.$$

By Lemma 34, this implies that we win the leaves within $\mathcal{O}(N^2 b \log^2 N)$ steps with high probability.

Combined with the time spent winning the hub in the first place ($N^2 \cdot b$ steps, with high probability), we obtain that, with high probability, fixation happens within $\mathcal{O}(N^{2+\varepsilon})$ steps.

Finally, it remains to show that fixation happens within $\mathcal{O}(N^{2+\varepsilon})$ steps not only with high probability but also on average. This follows from two observations. First, note that the process absorbs within $\mathcal{O}(N^{2+\varepsilon})$ steps with high probability, starting from any configuration: Indeed, for configurations containing a leaf we have just proved that the process in fact fixates with high probability. Given a configuration contained in the hub, by Lemma 36 we either go extinct, or win the hub, or gain a leaf in $\mathcal{O}(N \log^2 n)$ steps (with high probability). In the first case we have just absorbed and in the latter two cases we fixate within $\mathcal{O}(N^{2+\varepsilon})$ steps as above (with high probability). Second, consider stages of length $\tau = \mathcal{O}(N^{2+\varepsilon})$ such that the process absorbs within each stage with probability at least 1/2. Then we get

$$\mathrm{AT}(\mathrm{SR}(N)) \leq \frac{1}{2}\tau + \frac{1}{4}(2\tau) + \frac{1}{8}(3\tau) + \cdots = 2\tau = \mathcal{O}(N^{2+\varepsilon})$$

where we have summed a series $\sum_{i=1}^{\infty} i/2^i = 2$. Thus $\mathrm{AT}(\mathrm{SR}(N), r) = \mathcal{O}(N^{2+\varepsilon})$. $\square$

## 5.4 Numerical and simulation results

Here we present related numerical and simulation results. First, we illustrate that Selection reactors are superamplifiers for any $r > 1$ (see Figure 5.3). Second, we show that different Selection reactors achieve different tradeoffs between high fixation probability and low fixation time (see Figure 5.4).

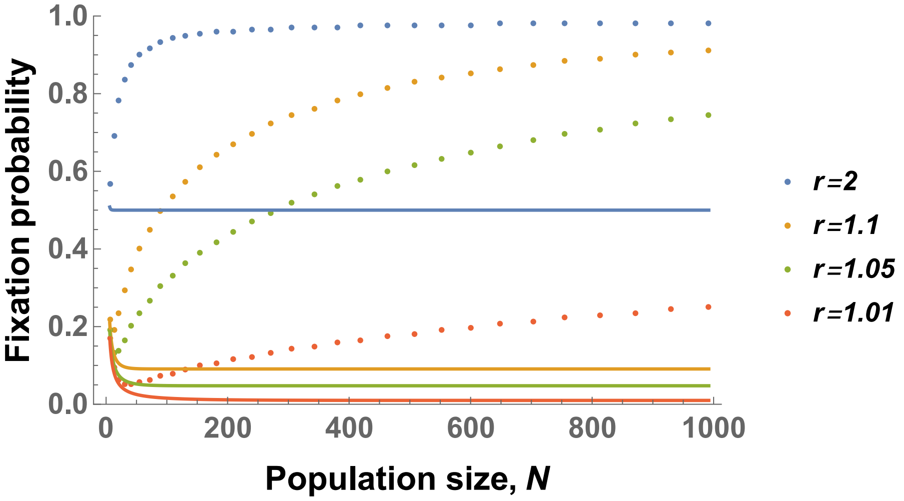

Figure 5.3: **Fixation probability on Selection reactors.** We consider Selection reactors $\mathrm{SR}'(N, \sqrt{N}, N^{1/2}, N^{3/2})$. We vary $N$ up to 1000 and consider $r \in \{2, 1.1, 1.05, 1.01\}$. The dots are exact values of the fixation probability under uniform initialization. The figure illustrates that the fixation probability tends to 1 for any $r > 1$. We remark that when $r = 1.01$ and $N = 1000$ then the Selection Reactor increases the fixation probability by a factor of 25, compared to the Complete graph.

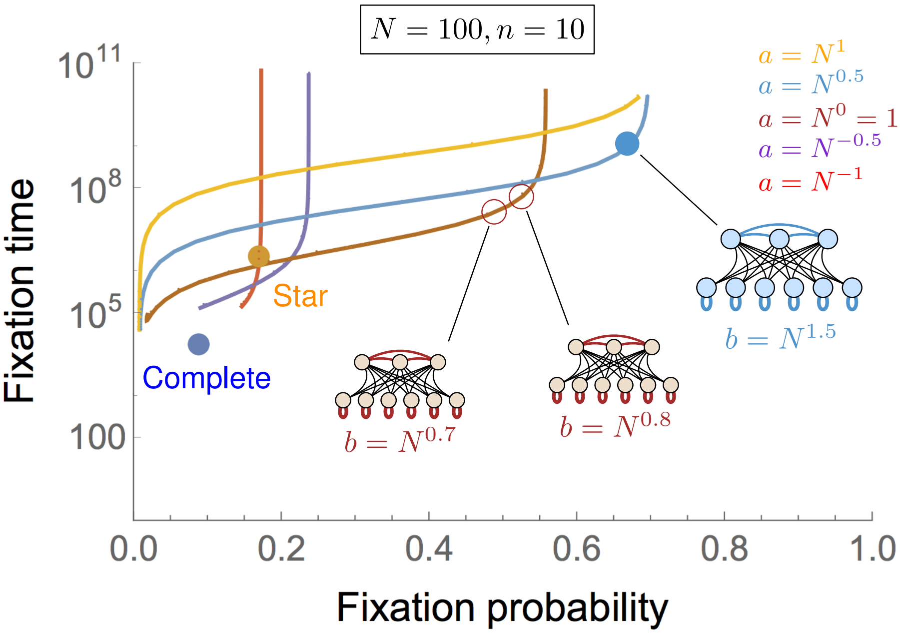

Figure 5.4: **Time-probability tradeoff for Selection reactors.** We fix $r = 1.1$, $N = 100$, $n = \sqrt{N}$ and consider Selection reactors $\mathrm{SR}(N, n, a, b)$ for five discrete values $a \in \{N^{-1}, N^{-0.5}, N^0, N^{0.5}, N^1\}$ and variable $b$. We observe that, compared to a Star graph, already for $N = 100$ certain Selection reactors substantially increase the fixation probability while other Selection reactors slightly increase the fixation probability while simultaneously decreasing the fixation time.

# Bibliography

[1] Kimura M. Evolutionary Rate at the Molecular Level. Nature. 1968;217(5129):624–626.

[2] Ewens WJ. Mathematical Population Genetics 1: Theoretical Introduction. Interdisciplinary Applied Mathematics. Springer New York; 2004.

[3] Nowak MA. Evolutionary dynamics: exploring the equations of life. Cambridge, Massachusetts: Belknap Press of Harvard University Press; 2006.

[4] Desai MM, Fisher DS, Murray AW. The speed of evolution and maintenance of variation in asexual populations. Curr Biol. 2007;17(5):385–394.

[5] McCandlish DM, Epstein CL, Plotkin JB. Formal properties of the probability of fixation: Identities, inequalities and approximations. Theor Popul Biol. 2015;99:98 – 113.

[6] Moran PAP. The statistical processes of evolutionary theory. Oxford University Press; 1962.

[7] Lieberman E, Hauert C, Nowak MA. Evolutionary dynamics on graphs. Nature. 2005;433(7023):312–316.

[8] Díaz J, Goldberg LA, Richerby D, Serna M. Absorption time of the Moran process. Random Structures & Algorithms. 2016;49(1):137–159.

[9] Altrock PM, Traulsen A. Fixation times in evolutionary games under weak selection. New Journal of Physics. 2009;11(1):013012.

[10] Slatkin M. Fixation Probabilities and Fixation Times in a Subdivided Population. Evolution. 1981;35(3):477–488.

[11] Nowak MA, May RM. Evolutionary games and spatial chaos. Nature. 1992;359(6398):826–829.

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

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

[14] Hauert C, Doebeli M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature. 2004 Apr;428(6983):643–646.

[15] Komarova NL. Spatial stochastic models for cancer initiation and progression. Bulletin of mathematical biology. 2006;68(7):1573–1599.

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

[17] Frean M, Rainey PB, Traulsen A. The effect of population structure on the rate of evolution. Proc R Soc B Biol Sci. 2013 Jul;280(1762):20130211.

[18] Komarova NL, Shahriyari L, Wodarz D. Complex role of space in the crossing of fitness valleys by asexual populations. Journal of The Royal Society Interface. 2014;11(95):20140014.

[19] Broom M, Rychtář J. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. Proc R Soc A Math Phys Eng Sci. 2008 Oct;464(2098):2609–2627. Available from: `http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.2008.0058`.

[20] Broom M, Rychtář J, Stadler BT. Evolutionary Dynamics on Graphs – the Effect of Graph Structure and Initial Placement on Mutant Spread. J Stat Theory Pract. 2011;5(3):369–381.

[21] Shakarian P, Roos P, Johnson A. A review of evolutionary graph theory with applications to game theory. Biosystems. 2012;107(2):66–80.

[22] Débarre F, Hauert C, Doebeli M. Social evolution in structured populations. Nat Commun. 2014;5.

[23] Allen B, Lippner G, Chen YT, Fotouhi B, Momeni N, Yau ST, et al. Evolutionary dynamics on any population structure. Nature. 2017 Apr;544(7649):227–230.

[24] Chalub FACC. An asymptotic expression for the fixation probability of a mutant in star graphs. Journal of Dynamics and Games. 2016;3:217–223.

[25] Askari M, Samani KA. Analytical calculation of average fixation time in evolutionary graphs. Physical Review E. 2015;92(4):042707.

[26] Tkadlec J, Pavlogiannis A, Chatterjee K, Nowak MA. Population structure determines the tradeoff between fixation probability and fixation time. Communications biology. 2019;2.

[27] Hindersin L, Möller M, Traulsen A, Bauer B. Exact numerical calculation of fixation probability and time on graphs. Biosystems. 2016;150:87–91.

[28] Díaz J, Goldberg LA, Mertzios GB, Richerby D, Serna M, Spirakis PG. Approximating fixation probabilities in the generalized Moran process. Algorithmica. 2014;69(1):78–91.

[29] Levin DA, Peres Y. Markov chains and mixing times. vol. 107. American Mathematical Soc.; 2017.

[30] Durrett R. Probability: theory and examples. vol. 49. Cambridge university press; 2019.

[31] Lindvall T. Lectures on the coupling method. Courier Corporation; 2002.

[32] Motwani R, Raghavan P. Randomized algorithms. Cambridge university press; 1995.

[33] Zukewich J, Kurella V, Doebeli M, Hauert C. Consolidating birth-death and death-birth processes in structured populations. PLoS One. 2013;8(1):e54639.

[34] Adlam B, Chatterjee K, Nowak M. Amplifiers of selection. In: Proc. R. Soc. A. vol. 471; 2015. p. 20150114.

[35] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. Physics reports. 2006;424(4-5):175–308.

[36] Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. Reviews of modern physics. 2009;81(2):591.

[37] Liggett TM. Stochastic interacting systems: contact, voter and exclusion processes. vol. 324. springer science & Business Media; 2013.

[38] Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton University Press; 2011.

[39] Szabó G, Fath G. Evolutionary games on graphs. Physics reports. 2007;446(4-6):97–216.

[40] Hadjichrysanthou C, Broom M, Rychtář J. Evolutionary games on star graphs under various updating rules. Dynamic Games and Applications. 2011;1(3):386.

[41] Kaveh K, Komarova NL, Kohandel M. The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. Royal Society open science. 2015;2(4):140465.

[42] Cormen TH. Introduction to algorithms. MIT press; 2009.

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

[44] Houchmandzadeh B, Vallade M. Exact results for fixation probability of bithermal evolutionary graphs. Biosystems. 2013;112(1):49–54.

[45] Cuesta FA, Sequeiros PG, Rojo ÁL, Benito RV. An Accurate Database of the Fixation Probabilities for All Undirected Graphs of Order 10 or Less. In: International Conference on Bioinformatics and Biomedical Engineering. Springer; 2017. p. 209–220.

[46] Möller M, Hindersin L, Traulsen A. Exploring and mapping the universe of evolutionary graphs identifies structural properties affecting fixation probability and time. Communications biology. 2019;2(1):137.

[47] Pavlogiannis A, Tkadlec J, Chatterjee K, Nowak MA. Amplification on Undirected Population Structures: Comets Beat Stars. Scientific Reports. 2017;7.

[48] Pavlogiannis A, Tkadlec J, Chatterjee K, Nowak MA. Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory. Communications Biology. 2018;1(1):71.

[49] Chatterjee K, Ibsen-Jensen R, Nowak MA. Faster Monte-Carlo Algorithms for Fixation Probability of the Moran Process on Undirected Graphs. In: MFCS. vol. 83; 2017. p. 61:1–61:13.

[50] Ann Goldberg L, Lapinskas J, Richerby D. Phase transitions of the Moran process and algorithmic consequences. Random Structures & Algorithms. 2019;Available from: `https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20890`.

[51] Díaz J, Goldberg LA, Mertzios GB, Richerby D, Serna, J M, Spirakis PG. On the fixation probability of superstars. Proc R Soc A Math Phys Eng Sci. 2013;469:20130193.

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

[53] Galanis A, Göbel A, Goldberg LA, Lapinskas J, Richerby D. Amplifiers for the Moran Process. J ACM. 2017 Mar;64(1):5:1–5:90. Available from: `http://doi.acm.org/10.1145/3019609`.

[54] Giakkoupis G. Amplifiers and Suppressors of Selection for the Moran Process on Undirected Graphs. arXiv preprint arXiv:161101585. 2016;.

[55] Goldberg LA, Lapinskas J, Lengler J, Meier F, Panagiotou K, Pfister P. Asymptotically optimal amplifiers for the Moran process. Theoretical Computer Science. 2019;758:73–93.

[56] Mertzios GB, Nikoletseas S, Raptopoulos C, Spirakis PG. Natural Models for Evolution on Networks. Theor Comput Sci. 2013;477:76–95.

[57] Mertzios GB, Spirakis PG. Strong Bounds for Evolution in Networks. ICALP. Springer-Verlag; 2013. p. 669–680.

[58] Hindersin L, Traulsen A. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. PLoS Comput Biol. 2015;11(11):e1004437.

[59] Allen B, Sample C, Jencks R, Withers J, Steinhagen P, Brizuela L, et al. Transient amplifiers of selection and reducers of fixation for death-Birth updating on graphs. arXiv preprint arXiv:190601036. 2019;.

[60] Tkadlec J, Pavlogiannis A, Chatterjee K, Nowak MA. Limits on amplifiers of natural selection under death-Birth updating. arXiv preprint arXiv:190602785. 2019;.

[61] Antal T, Redner S, Sood V. Evolutionary dynamics on degree-heterogeneous graphs. Phys Rev Lett. 2006;96:188104. Available from: `http://prl.aps.org/abstract/PRL/v96/i18/e188104`.

[62] Maruyama T. A Markov process of gene frequency change in a geographically structured population. Genetics. 1974;76(2):367–377. Available from: `http://iopscience.iop.org/1367-2630/13/7/073020http://www.genetics.org/content/76/2/367.short`.

[63] Adlam B, Nowak MA. Universality of fixation probabilities in randomly structured populations. Sci Rep. 2014 Jul;4.

[64] Goldberg LA, Lapinskas J, Lengler J, Meier F, Panagiotou K, Pfister P. Asymptotically Optimal Amplifiers for the Moran Process; 2017. Https://arxiv.org/abs/1611.04209.

[65] Kemeny JG, Griffeath DS, Snell JL, Knapp AW. Denumerable Markov Chains: with a chapter of Markov Random Fields by David Griffeath. Graduate Texts in Mathematics. Springer New York; 2012.

[66] Hindersin L, Traulsen A. Counterintuitive properties of the fixation time in network-structured populations. J R Soc Interface. 2014 Aug;11(99).