Strong Bounds for Evolution in Networks[*][†]

George B. Mertzios[1] and Paul G. Spirakis[2]

[1] School of Engineering and Computing Sciences, Durham University, UK. Email: `george.mertzios@durham.ac.uk`
[2] Department of Computer Science, University of Liverpool, UK, and University of Patras, Greece. Email: `p.spirakis@liverpool.ac.uk`

Abstract. This work studies the generalized Moran process, as introduced by Lieberman et al. [Nature, 433:312-316, 2005]. We introduce the parameterized notions of *selective amplifiers* and *selective suppressors* of evolution, i.e. of networks (graphs) with many "strong starts" and many "weak starts" for the mutant, respectively. We first prove the existence of *strong* selective amplifiers and of (quite) strong selective suppressors. Furthermore we provide strong upper bounds and almost tight lower bounds (by proving the "Thermal Theorem") for the traditional notion of fixation probability of Lieberman et al., i.e. assuming a random initial placement of the mutant.

Keywords: Evolutionary dynamics, undirected graphs, fixation probability, lower bound, Markov chain.

1 Introduction

Population and evolutionary dynamics have been extensively studied [2, 6, 7, 19, 25, 28, 29], mainly on the assumption that the evolving population is homogeneous, i.e. it has no spatial structure. One of the main models in this area is the Moran Process [23], where the initial population contains a single *mutant* with fitness $r > 0$, with all other individuals having fitness 1. At every step of this process, an individual is chosen for reproduction with probability proportional to its fitness. This individual then replaces a second individual, which is chosen uniformly at random, with a copy of itself. Such dynamics as the above have been extensively studied also in the context of strategic interaction in evolutionary game theory [14, 16–18, 27].

In a recent article, Lieberman, Hauert, and Nowak [20] (see also [24]) introduced a generalization of the Moran process, where the individuals of the population are placed on the vertices of a connected graph (which is, in general, directed) such that the edges of the graph determine competitive interaction. In the generalized Moran process, the initial population again consists of a single mutant of fitness r, placed on a vertex that is chosen uniformly at random, with each other vertex occupied by a non-mutant of fitness 1. An individual is chosen for reproduction exactly as in the standard Moran process, but now the second individual to be replaced is chosen among its neighbors in the graph uniformly at random (or according to some weights of the edges) [20, 24]. If the underlying graph is the complete graph, then this process becomes the standard Moran process on a homogeneous population [20, 24]. Several similar models describing infections and particle interactions have been also studied in the past, including the SIR and SIS epidemics [11, Chapter 21], the voter and antivoter models and the exclusion process [1, 10, 21]. However such models do not consider the issue of different fitness of the individuals.

The central question that emerges in the generalized Moran process is how the population structure affects evolutionary dynamics [20, 24]. In the present work we consider the generalized Moran process on arbitrary finite, undirected, and connected graphs. On such graphs, the generalized Moran process terminates almost surely, reaching either *fixation* of the graph (all vertices are occupied by copies of the mutant) or *extinction* of the mutants (no copy of the mutant remains). The *fixation probability* of a graph G for a mutant of fitness r, is the probability that eventually fixation is reached when the mutant is initially placed at a random vertex of G, and is

denoted by $f_r(G)$. The fixation probability can, in principle, be determined using standard Markov Chain techniques. But doing so for a general graph on n vertices requires solving a linear system of 2^n linear equations. Such a task is not computationally feasible, even numerically. As a result of this, most previous work on computing fixation probabilities in the generalized Moran process was either restricted to graphs of small size [6] or to graph classes which have a high degree of symmetry, reducing thus the size of the corresponding linear system (e.g. paths, cycles, stars, and cliques [3–5]). Experimental results on the fixation probability of random graphs derived from grids can be found in [26]. Moreover, layered directed graphs with extreme behavior due to the existence of "positive feedback loops", such as superstars, megastars, funnels, and metafunnels, have been studied in [8, 12, 20].

A recent result [9] shows how to construct fully polynomial randomized approximation schemes (FPRAS) for the probability of reaching fixation (when $r \geq 1$) or extinction (for all $r > 0$). The result of [9] uses a Monte Carlo estimator, i.e. it runs the generalized Moran process several times[3], while each run terminates in polynomial time with high probability [9]. Note that improved lower and upper bounds on the fixation probability immediately lead to a better estimator here. Until now, the only known general bounds for the fixation probability on connected undirected graphs, are that $f_r(G) \geq \frac{1}{n}$ and $f_r(G) \leq 1 - \frac{1}{n+r}$.

Lieberman et al. [20, 24] proved the *Isothermal Theorem*, stating that (in the case of undirected graphs) the fixation probability of a regular graph (i.e. of a graph with overall the same vertex degree) is equal to that of the complete graph (i.e. the homogeneous population of the standard Moran process), which equals to $(1 - \frac{1}{r})/(1 - \frac{1}{r^n})$, where n is the size of the population. Intuitively, in the Isothermal Theorem, every vertex of the graph has a *temperature* which determines how often this vertex is being replaced by other individuals during the generalized Moran process. The complete graph (or equivalently, any regular graph) serves as a benchmark for measuring the fixation probability of an arbitrary graph G: if $f_r(G)$ is larger (resp. smaller) than that of the complete graph then G is called an *amplifier* (resp. a *suppressor*) [20, 24].

Our contribution. The structure of the graph, on which the population resides, plays a crucial role in the course of evolutionary dynamics. Human societies or social networks are never homogeneous, while certain individuals in central positions may be more influential than others [24]. Motivated by this, we introduce in this paper a new notion of measuring the success of an advantageous mutant in a structured population, by counting the number of initial placements of the mutant in a graph that guarantee fixation of the graph with large probability. This provides a refinement of the notion of fixation probability. Specifically, we do not any more consider the fixation probability as the probability of reaching fixation when the mutant is placed at a random vertex, but we rather consider the probability $f_r(v)$ of reaching fixation when a mutant with fitness $r > 1$ is introduced at a specific vertex v of the graph; $f_r(v)$ is termed the *fixation probability of vertex v*. Using this notion, the fixation probability $f_r(G)$ of a graph $G = (V, E)$ with n vertices is $f_r(G) = \frac{1}{n} \sum_{v \in V} f_r(v)$.

We aim in finding graphs that have many "strong starts" (or many "weak starts") of the mutant. Thus we introduce the notions of $(h(n), g(n))$-*selective amplifiers* (resp. $(h(n), g(n))$-*selective suppressors*), which include those graphs with n vertices for which there exist at least $h(n)$ vertices v with $f_r(v) \geq 1 - \frac{c(r)}{g(n)}$ (resp. $f_r(v) \leq \frac{c(r)}{g(n)}$) for an appropriate function $c(r)$ of r. We contrast this new notion of $(h(n), g(n))$-selective amplifiers (resp. suppressors) with the notion of $g(n)$-*universal amplifiers* (resp. *suppressors*) which include those graphs G with n vertices for which $f_r(G) \geq 1 - \frac{c(r)}{g(n)}$ (resp. $f_r(G) \leq \frac{c(r)}{g(n)}$) for an appropriate function $c(r)$ of r. For a detailed presentation and a rigorous definition of these notions we refer to Section 2.

Using these new notions, we prove that there exist strong selective amplifiers, namely $(\Theta(n), n)$-selective amplifiers (called the *urchin graphs*). Furthermore we prove that there exist also quite

---

[3]For approximating the probability to reach fixation (resp. extinction), one needs a number of runs which is about the inverse of the best known lower (resp. upper) bound of the fixation probability.

strong selective suppressors, namely $(\frac{n}{\phi(n)+1}, \frac{n}{\phi(n)})$-selective suppressors (called the $\phi(n)$-*urchin graphs*) for *any* function $\phi(n) = \omega(1)$ with $\phi(n) \leq \sqrt{n}$.

Regarding the traditional measure of the fixation probability $f_r(G)$ of undirected graphs G, we provide upper and lower bounds that are much stronger than the bounds $\frac{1}{n}$ and $1 - \frac{1}{n+r}$ that were known so far [9]. More specifically, first of all we demonstrate the nonexistence of "strong" universal amplifiers by showing that for any graph G with n vertices, the fixation probability $f_r(G)$ is strictly less than $1 - \frac{c(r)}{g(n)}$, for any $g(n) = \omega(n^{\frac{3}{4}})$. This is in a wide contrast with what happens in directed graphs, as Lieberman et al. [20] provided directed graphs with arbitrarily large fixation probability (see also [24]). Motivated by our work, very recently Giakkoupis [13] and Goldberg et al. [15] slightly improved our estimate of the function $g(n)$ to $\omega(n^{\frac{1}{3}} \log^{\frac{4}{3}} n)$ [13] and to $\omega(n^{\frac{1}{3}})$ [15], respectively.

On the other hand, we provide our lower bound in the *Thermal Theorem*, which states that for any vertex v of an arbitrary undirected graph G, the fixation probability $f_r(v)$ of v is at least $(r-1)/(r + \frac{\deg v}{\deg_{\min}})$ for any $r > 1$, where $\deg v$ is the degree of v in G (i.e. the number of its neighbors) and $\deg_{\min}$ (resp. $\deg_{\max}$) is the minimum (resp. maximum) degree in G. This result extends the Isothermal Theorem for regular graphs [20]. In particular, we consider here a different notion of *temperature* for a vertex than [20]: the temperature of vertex v is $\frac{1}{\deg v}$. As it turns out, a "hot" vertex (i.e. with high temperature) affects more often its neighbors than a "cold" vertex (with low temperature). The Thermal Theorem, which takes into account the vertex v on which the mutant is introduced, provides immediately our lower bound $(r-1)/(r + \frac{\deg_{\max}}{\deg_{\min}})$ for the fixation probability $f_r(G)$ of any undirected graph G. The latter lower bound is almost tight, as it implies that $f_r(G) \geq \frac{r-1}{r+1}$ for a regular graph G, while the Isothermal Theorem implies that the fixation probability of a regular graph G tends to $\frac{r-1}{r}$ as the size of G increases. Note that our new upper/lower bounds for the fixation probability lead to better time complexity of the FPRAS proposed in [9], as the Monte Carlo technique proposed in [9] now needs to simulate the Moran process a less number of times (to estimate fixation or extinction).

Our techniques are original and of a constructive combinatorics flavor. For the class of strong selective amplifiers (the urchin graphs) we introduce a novel decomposition of the Markov chain $\mathcal{M}$ of the generalized Moran process into $n - 1$ smaller chains $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_{n-1}$, and then we decompose each $\mathcal{M}_k$ into two even smaller chains $\mathcal{M}_k^1, \mathcal{M}_k^2$. Then we exploit a new way of composing these smaller chains (and returning to the original one) that is carefully done to maintain the needed domination properties. For the proof of the lower bound in the Thermal Theorem, we first introduce a new and simpler weighted process that bounds fixation probability from below (the generalized Moran process is a special case of this new process). Then we add appropriate dummy states to its (exponentially large) Markov chain, and finally we iteratively modify the resulting chain by maintaining the needed monotonicity properties. Eventually this results to the desired lower bound of the Thermal Theorem. Finally, our proof for the non-existence of strong universal amplifiers is done by contradiction, partitioning appropriately the vertex set of the graph and discovering an appropriate independent set that leads to the contradiction.

The paper is organized as follows. Preliminaries and notation are given in Section 2. Furthermore we present our results on amplifiers and suppressors in Sections 3 and 4, respectively.

2   Preliminaries

Throughout the paper we consider only finite, connected, undirected graphs $G = (V, E)$. Our results apply to connected graphs as, otherwise, the fixation probability is necessarily zero. The edge $e \in E$ between two vertices $u, v \in V$ is denoted by $e = uv$. For a vertex subset $X \subseteq V$, we write $X + y$ and $X - y$ for $X \cup \{y\}$ and $X \setminus \{y\}$, respectively. Furthermore, throughout r denotes the fitness of the mutant, while the value r is considered to be independent of the size n of the network, i.e. we assume that r is constant. For simplicity of presentation, we call a vertex v "infected" if a

copy of the mutant is placed on v. For every vertex subset $S \subseteq V$ we denote by $f_r(S)$ the fixation probability of the set S, i.e. the probability that, starting with exactly $|S|$ copies of the mutant placed on the vertices of S, the generalized Moran process will eventually reach fixation. By the definition of the generalized Moran process $f_r(\emptyset) = 0$ and $f_r(V) = 1$, while for $S \notin \{\emptyset, V\}$,

$$
f_r(S) = \frac{\sum_{xy \in E, x \in S, y \notin S} \left(\frac{r}{\deg x} f_r(S+y) + \frac{1}{\deg y} f_r(S-x) \right)}{\sum_{xy \in E, x \in S, y \notin S} \left(\frac{r}{\deg x} + \frac{1}{\deg y} \right)} \tag{1}
$$

In the next definition we introduce the notions of *universal* and *selective* amplifiers.

Definition 1. *Let $\mathcal{G}$ be an infinite class of undirected graphs. If there exists an $n_0 \in \mathbb{N}$, an $r_0 \geq 1$, and some function $c(r)$, such that for every graph $G \in \mathcal{G}$ with $n \geq n_0$ vertices and for every $r > r_0$:*

- *$f_r(G) \geq 1 - \frac{c(r)}{g(n)}$, then $\mathcal{G}$ is a class of $g(n)$-universal amplifiers,*
- *there exists a subset S of at least $h(n)$ vertices of G, such that $f_r(v) \geq 1 - \frac{c(r)}{g(n)}$ for every vertex $v \in S$, then $\mathcal{G}$ is a class of $(h(n), g(n))$-selective amplifiers.*

Moreover, $\mathcal{G}$ is a class of strong universal *(resp.* strong selective*) amplifiers if $\mathcal{G}$ is a class of n-universal (resp. $(\Theta(n), n)$-selective) amplifiers.*

Similarly to Definition 1, we introduce the notions of *universal* and *selective* suppressors.

Definition 2. *Let $\mathcal{G}$ be an infinite class of undirected graphs. If there exist functions $c(r)$ and $n_0(r)$, such that for every $r > 1$ and for every graph $G \in \mathcal{G}$ with $n \geq n_0(r)$ vertices:*

- *$f_r(G) \leq \frac{c(r)}{g(n)}$, then $\mathcal{G}$ is a class of $g(n)$-universal suppressors,*
- *there exists a subset S of at least $h(n)$ vertices of G, such that $f_r(v) \leq \frac{c(r)}{g(n)}$ for every vertex $v \in S$, then $\mathcal{G}$ is a class of $(h(n), g(n))$-selective suppressors.*

Moreover, $\mathcal{G}$ is a class of strong universal *(resp.* strong selective*) suppressors if $\mathcal{G}$ is a class of n-universal (resp. $(\Theta(n), n)$-selective) suppressors.*

Note that $n_0 = n_0(r)$ in Definition 2, while in Definition 1 n_0 is not a function of r. The reason for this is that, since we consider the fitness value r to be constant, the size n of G needs to be sufficiently large with respect to r in order for G to act as a suppressor. Indeed, if we let r grow arbitrarily, e.g. if $r = n^2$, then for *any* graph G with n vertices the fixation probability $f_r(v)$ tends to 1 as n grows. The next lemma follows by Definitions 1 and 2.

Lemma 1. *If $\mathcal{G}$ is a class of $g(n)$-universal amplifiers (resp. suppressors), then $\mathcal{G}$ is a class of $(n - o(n), g(n))$-selective amplifiers (resp. suppressors).*

Proof. Suppose that $\mathcal{G}$ is a class of $g(n)$-universal amplifiers. That is, for every $r > r_0$ and for every graph $G = (V, E) \in \mathcal{G}$ with $n \geq n_0$ vertices, the fixation probability of G is $f_r(G) \geq 1 - \frac{c(r)}{g(n)}$, where $c(r)$ is some function that depends only on r. Let $S \subseteq V$ be the subset of vertices such that $f_r(v) \geq 1 - \frac{c'(r)}{g(n)}$ for some function $c'(r)$ that depends only on r. Then there exists an appropriate function $\phi(n, r) = \omega(1)$, i.e. $\lim_{n \to \infty} \phi(n, r) = \infty$, such that $f_r(v) \leq 1 - \frac{\phi(n, r)}{g(n)}$ for every $v \in V \setminus S$. Thus the fixation probability of G is

$$
f_r(G) \leq \frac{|S| \cdot 1 + (n - |S|) \cdot (1 - \frac{\phi(n,r)}{g(n)})}{n} = 1 - \frac{(n - |S|)}{n} \cdot \frac{\phi(n, r)}{g(n)} \tag{2}
$$

4

Now, since $f_r(G) \geq 1 - \frac{c(r)}{g(n)}$, it follows by (2) that $(n - |S|) \leq n\frac{c(r)}{\phi(n,r)}$, and thus $|S| \geq n(1 - \frac{c(r)}{\phi(n,r)}) = n - o(n)$, since $\phi(n,r) = \omega(1)$. Thus it follows by definition of the set S that $\mathcal{G}$ is a class of $(n - o(n), g(n))$-selective amplifiers.

Suppose now that $\mathcal{G}$ is a class of $g(n)$-universal suppressors. That is, for every $r > 1$ and for every graph $G = (V, E) \in \mathcal{G}$ with $n \geq n_0(r)$ vertices, the fixation probability of G is $f_r(G) \leq \frac{c(r)}{g(n)}$, where $c(r)$ is some function that depends only on r. Let $S \subseteq V$ be the subset of vertices such that $f_r(v) \leq \frac{c'(r)}{g(n)}$ for some function $c'(r)$ that depends only on r. Then there exists an appropriate function $\phi(n,r) = \omega(1)$, i.e. $\lim\limits_{n \to \infty} \phi(n,r) = \infty$, such that $f_r(v) \geq \frac{\phi(n,r)}{g(n)}$ for every $v \in V \setminus S$. Thus the fixation probability of G is

$$f_r(G) \geq \frac{|S| \cdot 0 + (n - |S|) \cdot \frac{\phi(n,r)}{g(n)}}{n} = \frac{(n - |S|)}{n} \cdot \frac{\phi(n,r)}{g(n)} \tag{3}$$

Now, since $f_r(G) \leq \frac{c(r)}{g(n)}$, it follows by (3) that $(n - |S|) \leq n\frac{c(r)}{\phi(n,r)}$, and thus $|S| \geq n(1 - \frac{c(r)}{\phi(n,r)}) = n - o(n)$, since $\phi(n,r) = \omega(1)$. Thus it follows by definition of the set S that $\mathcal{G}$ is a class of $(n - o(n), g(n))$-selective suppressors. $\square$

The most natural question that arises by Definitions 1 and 2 is whether there exists any class of strong selective amplifiers/suppressors, as well as for which functions $h(n)$ and $g(n)$ there exist classes of $g(n)$-universal amplifiers/suppressors and classes of $(h(n), g(n))$-selective amplifiers/suppressors. In Section 3 and 4 we provide our results on amplifiers and suppressors, respectively.

3   Amplifier bounds

In this section we prove that there exist no strong universal amplifiers (Section 3.1), although there exists a class of strong selective amplifiers (Section 3.2).

3.1   Non-existence of strong universal amplifiers

Theorem 1. *For any function $g(n) = \omega(n^{\frac{3}{4}})$ there exists no graph class $\mathcal{G}$ of $g(n)$-universal amplifiers for any $r > r_0 = 1$.*

Proof. The proof is done by contradiction. Let $r_0 = 1$ and $g(n) = \omega(n^{\frac{3}{4}})$. Then $g(n) = \Omega(n^{\frac{3}{4}}\phi(n))$ for some function $\phi(n) = \omega(1)$, i.e. $g(n) = \Omega(n^{1-\delta})$ where $\delta = \frac{1}{4} - \frac{\log \phi(n)}{\log n}$. Suppose that $\mathcal{G}$ is a class of $g(n)$-universal amplifiers. That is, for every graph $G = (V, E) \in \mathcal{G}$ with $n \geq n_0$ vertices, the fixation probability of G is $f_r(G) \geq 1 - \frac{c(r)}{g(n)} \geq 1 - \frac{c_0(r)}{n^{1-\delta}}$ for every $r > 1$, where $c(r), c_0(r)$ are two functions that depend only on r. We partition the vertex set V into three sets V_1, V_2, V_3 such that

$$V_1 = \{v \in V : f_r(v) \geq 1 - \frac{c_0(r)}{n^{1-\delta}}\} \tag{4}$$

$$V_2 = \{v \in V \setminus V_1 : f_r(v) \geq 1 - \frac{c_1(r)}{n^{1-2\delta}}\} \tag{5}$$

$$V_3 = \{v \in V \setminus V_1 : f_r(v) < 1 - \frac{c_1(r)}{n^{1-2\delta}}\} \tag{6}$$

where $c_1(r)$ is an appropriate function of r (to be specified below). Note that $V_1 \neq \emptyset$, since $f_r(G) \geq 1 - \frac{c_0(r)}{n^{1-\delta}}$ by assumption. Using (6), the fixation probability $f_r(G)$ of G is upper-bounded by

$$f_r(G) \leq \frac{(|V_1| + |V_2|) \cdot 1 + |V_3| \cdot (1 - \frac{c_1(r)}{n^{1-2\delta}})}{n}$$
$$= 1 - \frac{|V_3|}{n} \cdot \frac{c_1(r)}{n^{1-2\delta}} \tag{7}$$

5

Now, since $f_r(G) \geq 1 - \frac{c_0(r)}{n^{1-\delta}}$, it follows by (7) that $1 - \frac{c_0(r)}{n^{1-\delta}} \leq 1 - \frac{|V_3|}{n} \cdot \frac{c_1(r)}{n^{1-2\delta}}$, and thus

$$|V_3| \leq n^{1-\delta} \frac{c_0(r)}{c_1(r)} \tag{8}$$

For an arbitrary vertex $v \in V$, we obtain an upper bound on the probability $f_r(v)$ by assuming that fixation is reached if the process reaches at least two infected vertices, when it starts with only v being infected. Therefore

$$f_r(v) \leq \frac{r \cdot 1 + \sum_{x \in N(v)} \frac{1}{\deg x} \cdot 0}{r + \sum_{x \in N(v)} \frac{1}{\deg x}} = \frac{r}{r + \sum_{x \in N(v)} \frac{1}{\deg x}} \tag{9}$$

for every $v \in V$. It follows now by (4) and (9) that for every $v \in V_1$,

$$1 - \frac{c_0(r)}{n^{1-\delta}} \leq \frac{r}{r + \sum_{x \in N(v)} \frac{1}{\deg x}} \Leftrightarrow$$

$$\sum_{x \in N(v)} \frac{1}{\deg x} \leq \frac{r \cdot c_0(r)}{n^{1-\delta} - c_0(r)} \leq \frac{c'(r)}{n^{1-\delta}} \tag{10}$$

for an appropriate function $c'(r)$ of r. Therefore, since $\sum_{x \in N(v)} \frac{1}{\deg x} \geq \deg(v) \cdot \frac{1}{n}$, (10) implies that

$$\deg v \leq c'(r) \cdot n^{\delta} \tag{11}$$

for every $v \in V_1$. Furthermore, since $\frac{1}{\deg u} \leq \sum_{x \in N(v)} \frac{1}{\deg x}$ for every $u \in N(v)$, (10) implies that

$$\deg u \geq n^{1-\delta} \frac{1}{c'(r)} \tag{12}$$

for every $u \in N(v)$, where $v \in V_1$. Define now the value of the function $c_1(r)$ (cf. (5)-(6)) as:

$$c_1(r) = 2c_0(r) \cdot c'(r)$$

and thus (8) becomes

$$|V_3| \leq n^{1-\delta} \frac{1}{2c'(r)} \tag{13}$$

Similarly to (10), it follows by (5) and (9) that for every $v \in V_2$,

$$\sum_{u \in N(v)} \frac{1}{\deg u} \leq \frac{r \cdot c_1(r)}{n^{1-2\delta} - c_1(r)} \leq \frac{c''(r)}{n^{1-2\delta}}$$

for some function $c''(r)$ of r, and that

$$\deg v \leq c''(r) \cdot n^{2\delta} \tag{14}$$

for every $v \in V_2$.

In the remainder of the proof we will use a generic upper bound on the fixation probability of undirected graphs which was proved in [22]. To state this generic upper bound, we need to define for every vertex v the quantity $Q_v = \sum_{x \in N(v)} \frac{1}{\deg(x)}$, as well as for every edge $uv \in E$ the quantity $Q_{uv} = \sum_{x \in N(v) \setminus \{u\}} \frac{1}{\deg(x)} + \sum_{x \in N(u) \setminus \{v\}} \frac{1}{\deg(x)}$. Intuitively, Q_v is an indicator of how "strong" the neighbors of v are, compared to v. Roughly, if Q_v is larger, then v is weak compared to its neighbors. Similarly, Q_{uv} is an indicator of how "strong" a pair of neighbored vertices u, v are, compared to their neighborhood $N(u) \cup N(v) \setminus \{u, v\}$.

Let $v \in V$ such that $f_r(v)$ is maximized. The generic upper bound of [22] on the fixation probability is[4]

$$f_r(G) \leq \max_{u \in N(v)} \frac{r^2}{r^2 + rQ(v) + \frac{Q_v Q_{uv}}{2}}. \tag{15}$$

The main idea for the proof of (15) is to construct an auxiliary Markov chain $\widetilde{\mathcal{M}}$, in which the probability of reaching a specific absorbing state is at least as large as the fixation probability in the original Markov chain. To favor fixation in this auxiliary Markov chain $\widetilde{\mathcal{M}}$, fixation is assumed to be reached whenever we reach three mutants in the population. For further details on this upper bound on $f_r(G)$ we refer to Theorem 1 in [22].

In the following let $v \in V$ such that $f_r(v)$ is maximized. Then clearly $v \in V_1$. Furthermore, (15) implies that

$$f_r(G) \leq \max_{u \in N(v)} \frac{2r^2}{2r^2 + Q_v Q_{uv}}. \tag{16}$$

Let $u_0 \in N(v)$ be such that the right hand side of (16) is maximized, and thus

$$f_r(G) \leq \frac{2r^2}{2r^2 + Q_v Q_{u_0 v}} \tag{17}$$

To arrive to a contradiction to our assumption on δ, first we upper-bound the product $Q_v Q_{u_0 v}$ by a quantity proportional to $\frac{1}{n^{1-\delta}}$, and then we lower-bound $Q_v Q_{u_0 v}$ by a quantity proportional to $n^{-3\delta}$ (see inequality (19) below). Since $f_r(G) \geq 1 - \frac{c_0(r)}{n^{1-\delta}}$, it follows by (17) that

$$1 - \frac{c_0(r)}{n^{1-\delta}} \leq \frac{2r^2}{2r^2 + Q_v Q_{uv}} \Leftrightarrow$$
$$n^{1-\delta} Q_v Q_{u_0 v} \leq 2r^2 c_0(r) + c_0(r) Q_v Q_{u_0 v} \Leftrightarrow$$
$$Q_v Q_{u_0 v} \leq \frac{2r^2 c_0(r)}{n^{1-\delta} - c_0(r)} \leq \frac{c'''(r)}{n^{1-\delta}} \tag{18}$$

for an appropriate function $c'''(r)$ of r.

Since $v \in V_1$ and $u_0 \in N(v)$, it follows by (12) and (13) that u_0 has at least $n^{1-\delta} \frac{1}{2c'(r)}$ neighbors in $V_1 \cup V_2$. Thus (11) and (14) imply that

$$\sum_{x \in N(u_0) \setminus \{v\}} \frac{1}{\deg(x)} \geq \left(n^{1-\delta} \frac{1}{2c'(r)} - 1 \right) \cdot \min \left\{ \frac{1}{c'(r) \cdot n^{\delta}}, \frac{1}{c''(r) \cdot n^{2\delta}} \right\} = \Omega(n^{1-3\delta})$$

Furthermore $Q_{u_0 v} \geq \sum_{x \in N(u_0) \setminus \{v\}} \frac{1}{\deg(x)}$ by the definition of $Q_{u_0 v}$, and thus $Q_{u_0 v} = \Omega(n^{1-3\delta})$. Moreover $Q_v Q_{u_0 v} = \Omega(n^{-3\delta})$, since $Q_v = \Omega(\frac{1}{n})$. Therefore it follows by (18) that

$$\Omega(n^{-3\delta}) = Q_v Q_{u_0 v} \leq \frac{c'''(r)}{n^{1-\delta}}. \tag{19}$$

This is a contradiction, since $\delta = \frac{1}{4} - \frac{\log \phi(n)}{\log n}$ by assumption, where $\phi(n) = \omega(1)$. Therefore there exists no class $\mathcal{G}$ of $g(n)$-universal amplifiers for any $r > r_0 = 1$, where $g(n) = \omega(n^{\frac{3}{4}})$. □

The next corollary follows from Theorem 1.

Corollary 1. *There exists no infinite class $\mathcal{G}$ of undirected graphs which are strong universal amplifiers.*

---

[4]In the original statement of this upper bound (see Theorem 1 in [22]), the maximization is taken over all edges $uv \in E$. However, it follows from the proof of Theorem 1 in [22] that the maximization can actually be taken over all edges uv such that $f_r(v)$ is maximized, i.e. as stated in the inequality (15).

3.2 A class of strong selective amplifiers

In this section we present the first class $\mathcal{G} = \{G_n : n \geq 1\}$ of strong selective amplifiers, which we call the *urchin* graphs. Namely, the graph G_n has $2n$ vertices, consisting of a clique with n vertices, an independent set of n vertices, and a perfect matching between the clique and the independent set, as illustrated in Figure 1. For every graph G_n, we refer for simplicity to a vertex of the clique of G_n as a *clique vertex* of G_n, and to a vertex of the independent set of G_n as a *nose* of G_n, respectively. We prove in this section that the class $\mathcal{G}$ of urchin graphs are strong selective amplifiers. Namely, we prove that, whenever $r > r_0 = 5$, the fixation probability of any nose v of any graph G_n is $f_r(v) \geq 1 - \frac{c(r)}{n}$, where $c(r)$ is a function that depends only on the fitness r of the mutant.
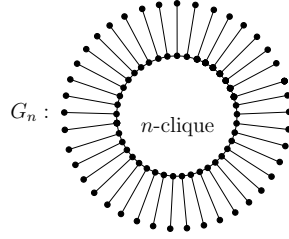

Fig. 1. The "urchin" graph G_n with $2n$ vertices.

Let v be a clique vertex (resp. a nose) and u be its adjacent nose (resp. clique vertex). Let v be infected; if u is not infected, then v is called an *isolated clique vertex* (resp. *isolated nose*), otherwise v is called a *covered clique vertex* (resp. *covered nose*). Let $k \in \{0, 1, \ldots, n\}$, $i \in \{0, 1, 2, \ldots, n-k\}$, and $x \in \{0, 1, 2, \ldots, k\}$. Denote by $Q_{i,x}^k$ the state of G_n with exactly i isolated clique vertices, x isolated noses, and $k-x$ covered noses. An example of the state $Q_{i,x}^k$ is illustrated in Figure 2(a). Furthermore, for every $k, i \in \{0, 1, \ldots, n\}$, we define the state P_i^k of G_n as follows. If $i \leq k$, then P_i^k is the state with exactly i covered noses and $k-i$ isolated noses. If $i > k$, then P_i^k is the state with exactly k covered noses and $i-k$ isolated clique vertices. Note that $Q_{i,0}^k = P_{k+i}^k$ and $Q_{0,x}^k = P_{k-x}^k$, for every $k \in \{0, 1, \ldots, n\}$, $i \in \{0, 1, 2, \ldots, n-k\}$, and $x \in \{0, 1, 2, \ldots, k\}$. Two examples of the state P_i^k, for the cases where $i \leq k$ and $i > k$, are illustrated in Figures 2(b) and 2(c), respectively.
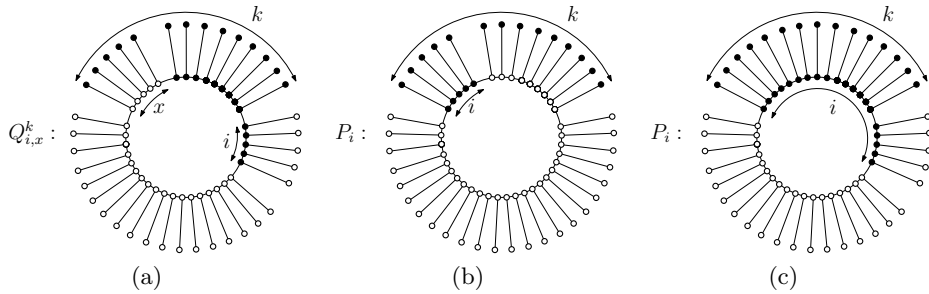

(a) (b) (c)

Fig. 2. The state (a) $Q_{i,x}^k$ and the state P_i^k, where (b) $i \leq k$, and (c) $i > k$.

Let $k \in \{1, 2, \ldots, n-1\}$. For all appropriate values of i and x, we denote by $q_{i,x}^k$ (resp. p_i^k) the probability that, starting at state $Q_{i,x}^k$ (resp. P_i^k) we eventually arrive to a state with $k+1$ infected noses before we arrive to a state with $k-1$ infected noses.

Lemma 2. *Let $k \in \{1, 2, \ldots, n-1\}$. Then $q_{i,x}^k > q_{i-1,x-1}^k$, for every $i \in \{1, 2, \ldots, n-k\}$ and every $x \in \{1, 2, \ldots, k\}$.*

Proof. Denote by $\mathcal{M}_1$ the Markov chain with starting state $Q_{i,x}^k$. Similarly, denote by $\mathcal{M}_2$ the Markov chain with starting state $Q_{i-1,x-1}^k$. Note that, initially, both Markov chains $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same number $2(k-x)+x+i$ of infected vertices. Moreover, $\mathcal{M}_1$ and $\mathcal{M}_2$ coincide initially on all their vertices except two. In particular, $\mathcal{M}_1$ has initially an isolated nose u, which is a covered nose in $\mathcal{M}_2$. Furthermore, $\mathcal{M}_1$ has initially an isolated clique vertex v, which is an uninfected clique vertex in $\mathcal{M}_2$. Denote by u' the (unique) clique vertex that is adjacent to u in G_n. Furthermore denote by v' the (unique) nose that is adjacent to v in G_n. Note that, initially, u' is uninfected in $\mathcal{M}_1$ and infected in $\mathcal{M}_2$, while v' is uninfected in both $\mathcal{M}_1$ and $\mathcal{M}_2$.

Note that at every iteration of the processes $\mathcal{M}_1$ and $\mathcal{M}_2$, one vertex w is activated and then it replaces a neighbor w' of it by an offspring of w. Thus, an equivalent way to analyze these processes is to consider that, at every iteration, one directed edge between two adjacent vertices is activated (with the appropriate probability). In order to prove that $q_{i,x}^k > q_{i-1,x-1}^k$, we simulate the progress of $\mathcal{M}_1$ by the random choices made at the corresponding steps by $\mathcal{M}_2$. In particular, we simulate the processes $\mathcal{M}_1$ and $\mathcal{M}_2$ until they reach states S_1 and S_2, respectively, such that either $S_1 = S_2$, or one of S_1 and S_2 is strictly included in the other. Furthermore, during the whole simulation of $\mathcal{M}_1$ by $\mathcal{M}_2$, before we reach such states S_1 and S_2, both $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same number of infected vertices at the corresponding iterations.

Suppose that a vertex w is activated for reproduction in $\mathcal{M}_2$ and that it places its offspring at a vertex w'. If both $w, w' \notin \{u, u', v, v'\}$ then we mimic this step in $\mathcal{M}_1$, i.e. in $\mathcal{M}_1$ we also activate w for reproduction and we place its offspring at w'. In the next three paragraphs we consider each of the remaining cases where at least one of w, w' belongs to $\{u, u', v, v'\}$. For each of these cases we specify how we simulate this step in $\mathcal{M}_1$.

If the clique vertex v (resp. u') is activated for reproduction in $\mathcal{M}_2$, then we activate u' (resp. v) in $\mathcal{M}_1$. In this case, if v (resp. u') places in $\mathcal{M}_2$ its offspring at a clique vertex $w \neq u'$ (resp. $w \neq v$), then u' (resp. v) places in $\mathcal{M}_1$ its offspring at the same clique vertex w. If v (resp. u') places in $\mathcal{M}_2$ its offspring at the clique vertex u' (resp. v), then u' (resp. v) places in $\mathcal{M}_1$ its offspring at the clique vertex v (resp. u'); in this case we arrive to two identical states in both $\mathcal{M}_1$ and $\mathcal{M}_2$. Finally, if v (resp. u') places in $\mathcal{M}_2$ its offspring at its adjacent nose v' (resp. u), then u' (resp. v) places in $\mathcal{M}_1$ its offspring at its adjacent nose u (resp. v'); in this case we arrive in $\mathcal{M}_1$ to a state, in which the infected vertices are a strict subset (resp. superset) of the infected vertices in $\mathcal{M}_2$.

If a clique vertex $w \notin \{v, u'\}$ is activated for reproduction in $\mathcal{M}_2$, and if w places in $\mathcal{M}_2$ its offspring at v (resp. u'), then w places in $\mathcal{M}_1$ its offspring at u' (resp. v). In this case, if the number of infected vertices changes in $\mathcal{M}_2$, then we arrive to the same state in both $\mathcal{M}_1$ and $\mathcal{M}_2$.

Finally, if the nose $w = v'$ (resp. $w = u$) is activated for reproduction in $\mathcal{M}_2$, then we activate the same nose also in $\mathcal{M}_1$. In this case we arrive in $\mathcal{M}_1$ to a state, in which the infected vertices are a strict subset (resp. superset) of the infected vertices in $\mathcal{M}_2$.

Note now that $q_{i,x}^k > q_{i-1,x-1}^k$ if and only if, in the above simulation of $\mathcal{M}_1$ by $\mathcal{M}_2$, the probability that we arrive to strictly more infected vertices in $\mathcal{M}_1$ than $\mathcal{M}_2$ is greater or equal to the probability that we arrive to strictly less infected vertices in $\mathcal{M}_1$ than $\mathcal{M}_2$. Furthermore, note that we arrive in $\mathcal{M}_1$ to a state with strictly more or strictly less infected vertices than in $\mathcal{M}_2$ only when one of the edges uu' or vv' is activated (in some direction) in the process $\mathcal{M}_2$. In particular, whenever this event occurs, $\mathcal{M}_1$ receives strictly more infected vertices than $\mathcal{M}_2$, if either u places its offspring at u' in $\mathcal{M}_2$, or if u' places its offspring at u in $\mathcal{M}_2$. Similarly, $\mathcal{M}_1$ receives strictly less infected vertices than $\mathcal{M}_2$, if either v places its offspring at v' in $\mathcal{M}_2$, or if v' places its offspring at v in $\mathcal{M}_2$. The ratio of these probabilities is

$$\frac{r \cdot \frac{1}{n} + r \cdot 1}{\frac{1}{n} + 1} = r > 1$$

and thus $q_{i,x}^k > q_{i-1,x-1}^k$. $\qquad\square$

Corollary 2. *Let $k \in \{1, 2, \ldots, n-1\}$, $i \in \{0, 1, \ldots, n-k\}$, and $x \in \{0, 1, \ldots, k\}$. Then $q_{i,x}^k > p_{k+i-x}^k$.*

Proof. Suppose first that $i \geq x$. Then Lemma 2 implies that $q_{i,x}^k > q_{i-1,x-1}^k > \ldots > q_{i-x,0}^k = p_{k+i-x}^k$. Suppose now that $i < x$. Then Lemma 2 implies that $q_{i,x}^k > q_{i-1,x-1}^k > \ldots > q_{0,x-i}^k = p_{k+i-x}^k$. $\square$

Now, starting from the Markov chain of the generalized Moran process, we define the Markov chain $\mathcal{M}$ by replacing any transition to a state $Q_{i,x}^k$ with a transition to state P_{k+i-x}^k. Then, for every nose v of the graph G_n, the fixation probability $f_r(v)$ of v is by Corollary 2 greater than or equal to the fixation probability of state P_0^1 in the Markov chain $\mathcal{M}$. Thus, in order to compute a lower bound on the fixation probability $f_r(v)$ of a nose v in G_n, we will compute a lower bound on the fixation probability of state P_0^1 in $\mathcal{M}$ (cf. Theorem 2).

In order to analyze $\mathcal{M}$, we decompose it first into the $n-1$ smaller Markov chains $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_{n-1}$, as follows. For every $k \in \{1, 2, \ldots, n-1\}$, the Markov chain $\mathcal{M}_k$ captures all transitions of $\mathcal{M}$ between states with k infected noses. The state graph of $\mathcal{M}_k$ is illustrated in Figure 3, where we denote by F_{k-1} (resp. F_{k+1}) an *arbitrary* state with $k-1$ (resp. $k+1$) infected noses. Moreover, we consider F_{k-1} and F_{k+1} as absorbing states of $\mathcal{M}_k$. Since we want to compute a lower bound of the fixation probability, whenever we arrive at state F_{k+1} (resp. at state F_{k-1}), we assume that we have the smallest number of infected clique vertices with $k+1$ (resp. with $k-1$) infected noses. That is, whenever $\mathcal{M}_k$ reaches state F_{k+1}, we assume that $\mathcal{M}$ has reached state P_{k+1}^{k+1} (and thus we move to the Markov chain $\mathcal{M}_{k+1}$). Similarly, whenever $\mathcal{M}_k$ reaches state F_{k-1}, we assume that $\mathcal{M}$ has reached state P_0^{k-1} (and thus we move to the Markov chain $\mathcal{M}_{k-1}$).
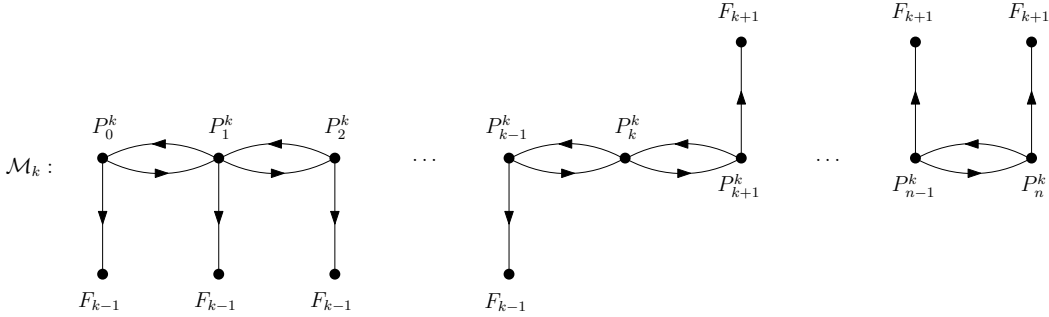

Fig. 3. The state graph of the relaxed Markov chain $\mathcal{M}_k$, where $k \in \{1, 2, \ldots, n-1\}$.

A decomposition of $\mathcal{M}_k$ into two Markov chains In order to analyze the Markov chain $\mathcal{M}_k$, where $k \in \{1, 2, \ldots, n-1\}$, we decompose it into two smaller Markov chains $\{\mathcal{M}_k^1, \mathcal{M}_k^2\}$, as they are shown in Figure 4.

In $\mathcal{M}_k^1$, we consider the state P_{k+1}^k absorbing. For every $i \in \{0, 1, \ldots, k\}$ denote by h_i^k the probability that, starting at state P_i^k in $\mathcal{M}_k^1$, we eventually reach state P_{k+1}^k before we reach state F_{k-1}, cf. Figure 4(a). In this Markov chain $\mathcal{M}_k^1$, every transition probability between two states is equal to the corresponding transition probabilities in $\mathcal{M}_k$.

In $\mathcal{M}_k^2$, we denote by s_i^k, where $i \in \{k, k+1, \ldots, n\}$, the probability that starting at state P_i^k we eventually reach state F_{k+1} before we reach state F_{k-1}, cf. Figure 4(b). In this Markov chain $\mathcal{M}_k^2$, the transition probability from state P_k^k to state P_{k+1}^k (resp. to state F_{k-1}) is equal to h_k^k (resp. $1 - h_k^k$), while all other transition probabilities between two states in $\mathcal{M}_k^2$ are the same as the corresponding transition probabilities in $\mathcal{M}_k$.

In order to prove the main result of this section, namely that the class $\mathcal{G} = \{G_n : n \geq 1\}$ of urchin graphs is a class of strong selective amplifiers (cf. Theorem 3), we first need to prove a series of technical results which can be outlined as follows.

First, we compute (cf. Lemmas 3 and 4) a lower bound on the probability h_0^k. That is, we compute a lower bound on the probability that, starting at state P_0^k in $\mathcal{M}_k^1$ (i.e. starting with
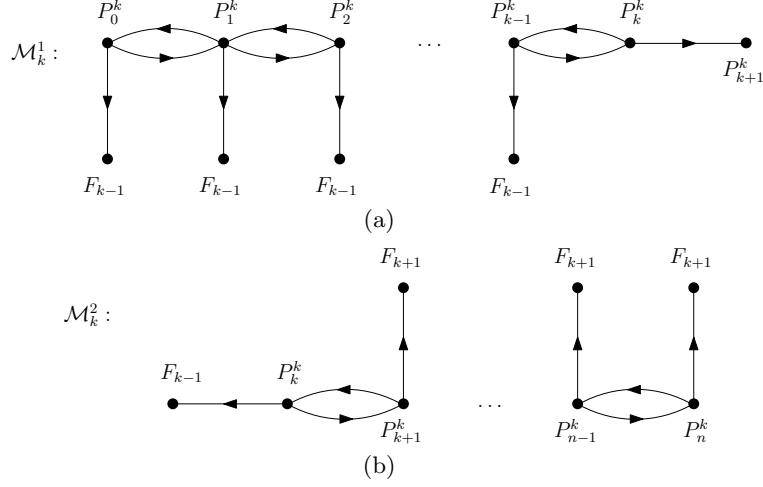
Fig. 4. The two Markov chains $\mathcal{M}_k^1$ and $\mathcal{M}_k^2$, where $k \in \{1, 2, \ldots, n-1\}$.

exactly k infected noses and all other vertices uninfected), we eventually infect at least $k+1$ clique vertices before disinfecting any infected nose. Second, we compute by Lemma 5 a lower bound on the probability s_k^k. That is, we compute a lower bound on the probability that, starting at state P_k^k in $\mathcal{M}_k^2$ (i.e. starting with k noses and their k adjacent clique vertices infected, while all other vertices are uninfected), we eventually infect one more nose before disinfecting any infected nose.

Then, using these lower bounds for h_0^k and s_k^k, we are able to compute a lower bound on the fixation probability of state P_0^1 in the Markov chain $\mathcal{M}$ (cf. Figure 5). To do so, we further relax the Markov chain $\mathcal{M}$ to another Markov chain $\mathcal{M}'$, which is equivalent to a birth-death process $\mathcal{B}_n$ (cf. Figures 6(a) and 6(b), respectively). As it turns out, the fixation probability p_1 of state P_0^1 in this birth-death process $\mathcal{B}_n$ (cf. Figure 6(b)) is a lower bound on the fixation probability of state P_0^1 in the Markov chain $\mathcal{M}$ (cf. Figure 5), which is in turn a lower bound on the fixation probability $f_r(v)$ of a nose v in the urchin graph G_n. Finally we prove in Theorem 2 that $p_1 \geq 1 - \frac{c(r)}{n}$, for some appropriate function $c(r)$ of r, and thus the same lower bound also applies to the fixation probability $f_r(v)$ of a nose v in G_n (cf. Theorem 3).

Lemma 3. *In the Markov chain $\mathcal{M}_k^1$, for any $r > 1$,*

$$h_k^k \geq 1 - \frac{2}{n(r-1)+1} = 1 - O(\frac{1}{n}) \tag{20}$$

Proof. For $i = 0$, the value of h_i^k in the Markov chain $\mathcal{M}_k^1$ is

$$h_0^k = \frac{rk \cdot h_1^k + \frac{k}{n} \cdot 0}{rk + \frac{k}{n}} = \frac{rn}{rn+1} \cdot h_1^k \tag{21}$$

and thus

$$h_1^k - h_0^k = \frac{1}{rn} h_0^k \leq \frac{1}{rn} \tag{22}$$

Furthermore, for every $i \in \{1, \ldots, k\}$, where $1 \leq k \leq n-1$, the value of h_i^k in $\mathcal{M}_k^1$ can be computed as follows.

$$h_i^k = \alpha_i^k \cdot h_{i+1}^k + \beta_i^k \cdot h_{i-1}^k + \gamma_i^k \cdot 0 \tag{23}$$

11

where

$$\alpha_i^k = \frac{r^{\frac{((k-i)n+i(n-i))}{n}}}{\Sigma_i^k}$$

$$\beta_i^k = \frac{\frac{i(n-i)}{n}}{\Sigma_i^k} \tag{24}$$

$$\gamma_i^k = \frac{\frac{k-i}{n}}{\Sigma_i^k}$$

and $\Sigma_i^k = r^{\frac{((k-i)n+i(n-i))}{n}} + \frac{i(n-i)}{n} + \frac{k-i}{n}$. Therefore (23) implies that

$$h_{i+1}^k - h_i^k = \frac{\beta_i^k}{\alpha_i^k}(h_i^k - h_{i-1}^k) + \frac{\gamma_i^k}{\alpha_i^k}h_i^k \tag{25}$$

Furthermore, (24) implies that

$$\frac{\beta_i^k}{\alpha_i^k} = \frac{1}{r} \cdot \frac{i(n-i)}{(k-i)n+i(n-i)} \leq \frac{1}{r} \tag{26}$$

$$\frac{\gamma_i^k}{\alpha_i^k} = \frac{1}{r} \cdot \frac{k-i}{(k-i)n+i(n-i)} = \frac{1}{r} \cdot \frac{1}{n + \frac{i(n-i)}{k-i}} \leq \frac{1}{rn} \tag{27}$$

Note that the inequality $\frac{\gamma_i^k}{\alpha_i^k} \leq \frac{1}{rn}$ in (27) holds also for $i = k$, since $\gamma_k^k = 0$. Therefore it follows by (25), (26), and (27) that

$$h_{i+1}^k - h_i^k \leq \frac{1}{r}(h_i^k - h_{i-1}^k) + \frac{1}{rn}h_i^k \tag{28}$$

Thus, since $h_{k+1}^k = 1$ by definition, it follows by (28) for $i = k$ that

$$1 - h_k^k \leq \frac{1}{r}(h_k^k - h_{k-1}^k) + \frac{1}{rn}h_k^k$$
$$\leq \dots$$
$$\leq \frac{1}{r^k}(h_1^k - h_0^k) + \frac{1}{rn}(h_k^k + \frac{1}{r}h_{k-1}^k + \dots + \frac{1}{r^{k-1}}h_1^k)$$

Therefore, since $h_1^k \leq h_2^k \leq \dots \leq h_k^k$ and $h_1^k - h_0^k \leq \frac{1}{rn}$ by (22), it follows that

$$1 - h_k^k \leq \frac{1}{r^{k+1}n} + \frac{1}{n(r-1)}h_k^k \leq \frac{1}{(r-1)n} + \frac{1}{n(r-1)}h_k^k$$

Therefore

$$h_k^k \geq 1 - \frac{2}{n(r-1)+1} \tag{29}$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4. *In the Markov chain $\mathcal{M}_k^1$, for any $r > 1$,*

$$h_0^k \geq 1 - \frac{k+2}{n(r-1)} \tag{30}$$

Proof. Let $1 \leq i \leq k$. Recall by (28) in the proof of Lemma 3 that

$$h_{i+1}^k - h_i^k \leq \frac{1}{r}(h_i^k - h_{i-1}^k) + \frac{1}{rn}h_i^k \leq \dots \leq \frac{1}{r^i}(h_1^k - h_0^k) + \frac{1}{n(r-1)}$$

Therefore it follows by (22) that

$$h_{i+1}^k - h_i^k \leq \frac{1}{r^{i+1}n} + \frac{1}{n(r-1)} \tag{31}$$

Summing up (31) for every $i = 1, 2, \ldots, k$, it follows that

$$1 - h_1^k \leq \frac{1}{n}\left(\frac{1}{r^2} + \ldots + \frac{1}{r^{k+1}}\right) + \frac{k}{n(r-1)}$$
$$\leq \frac{1}{n}\frac{1}{r(r-1)} + \frac{k}{n(r-1)} \leq \frac{k+1}{n(r-1)}$$

since $h_{k+1}^k = 1$, and thus

$$h_1^k \geq 1 - \frac{k+1}{n(r-1)} \tag{32}$$

Therefore it follows now by (21) that

$$h_0^k \geq \frac{rn}{rn+1}\left(1 - \frac{k+1}{n(r-1)}\right) \geq \frac{n(r-1)}{n(r-1)+1}\left(1 - \frac{k+1}{n(r-1)}\right) \geq 1 - \frac{k+2}{n(r-1)}$$

$\square$

We now provide some lower bounds for the Markov chain $\mathcal{M}_k^2$.

Lemma 5. *In the Markov chain $\mathcal{M}_k^2$, for any $r > 5$,*

$$s_k^k \geq 1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{n}{(n-k)^2} \tag{33}$$

Proof. For $i = k$, the value of s_i^k in the Markov chain $\mathcal{M}_k^2$ is

$$s_k^k = h_k^k \cdot s_{k+1}^k + (1 - h_k^k) \cdot 0 = h_k^k \cdot s_{k+1}^k \tag{34}$$

Therefore Lemma 3 implies that

$$s_k^k \geq \left(1 - \frac{2}{n(r-1)+1}\right)s_{k+1}^k \geq \left(1 - \frac{2}{n(r-1)}\right)s_{k+1}^k \tag{35}$$

and thus

$$s_{k+1}^k - s_k^k \leq \frac{2}{n(r-1)}s_{k+1}^k \leq \frac{2}{n(r-1)} \tag{36}$$

Furthermore, for every $i \in \{k+1, \ldots, n\}$, the value of s_i^k in $\mathcal{M}_k^2$ can be computed as follows.

$$s_i^k = \alpha_i^k \cdot s_{i+1}^k + \beta_i^k \cdot s_{i-1}^k + \gamma_i^k \cdot 1 \tag{37}$$

where

$$\alpha_i^k = \frac{r^{\frac{i(n-i)}{n}}}{\Sigma_i^k}$$
$$\beta_i^k = \frac{(i-k) + \frac{i(n-i)}{n}}{\Sigma_i^k} = \frac{\frac{(i-k)n+i(n-i)}{n}}{\Sigma_i^k} \tag{38}$$
$$\gamma_i^k = \frac{r^{\frac{i-k}{n}}}{\Sigma_i^k}$$

13

and $\Sigma_i^k = r\frac{i(n-i)}{n} + \frac{(i-k)n+i(n-i)}{n} + r\frac{i-k}{n}$. Therefore (37) implies that

$$s_{i+1}^k - s_i^k = \frac{\beta_i^k}{\alpha_i^k}(s_i^k - s_{i-1}^k) - \frac{\gamma_i^k}{\alpha_i^k}(1 - s_i^k) \tag{39}$$

Furthermore, (38) implies that

$$\frac{\beta_i^k}{\alpha_i^k} = \frac{1}{r} \cdot (1 + \frac{(i-k)n}{i(n-i)}) \tag{40}$$

$$\frac{\gamma_i^k}{\alpha_i^k} = \frac{i-k}{i(n-i)} \geq \frac{i-k}{i} \cdot \frac{1}{n} \tag{41}$$

We now prove that $\frac{\beta_i^k}{\alpha_i^k} \leq \frac{5}{r}$, whenever $i \leq \frac{n+k}{2}$. Suppose first that $k \leq \frac{n}{2}$. Then $i \leq \frac{n+k}{2} \leq \frac{n+\frac{n}{2}}{2}$, i.e. $i \leq \frac{3n}{4}$. Thus $\frac{1}{n-i} \leq \frac{4}{n}$, and thus (40) implies that $\frac{\beta_i^k}{\alpha_i^k} \leq \frac{1}{r} \cdot (1+4) = \frac{5}{r}$. Suppose now that $n \geq k > \frac{n}{2}$. Then also $i > \frac{n}{2}$, since $i \geq k+1$, and thus $\frac{n}{i} < 2$. Furthermore $i - k \leq n - i$, since $i \leq \frac{n+k}{2}$. Therefore $\frac{(i-k)n}{i(n-i)} = \frac{i-k}{n-i} \cdot \frac{n}{i} < 2$, and thus (40) implies that $\frac{\beta_i^k}{\alpha_i^k} < \frac{1}{r} \cdot (1+2) = \frac{3}{r}$. Summarizing, for every $k \in \{1, 2, \ldots, n-1\}$ and every $i \in \{k+1, \ldots, \frac{n+k}{2}\}$,

$$\frac{\beta_i^k}{\alpha_i^k} \leq \frac{5}{r} \tag{42}$$

Therefore it follows by (39), (41), and (42) that

$$s_{i+1}^k - s_i^k \leq \frac{5}{r}(s_i^k - s_{i-1}^k) - \frac{i-k}{in}(1 - s_i^k) \tag{43}$$

Thus, in particular

$$s_i^k - s_{i-1}^k \leq \frac{5}{r}(s_{i-1}^k - s_{i-2}^k)$$
$$\leq (\frac{5}{r})^{i-k-1}(s_{k+1}^k - s_k^k) \tag{44}$$

Now (43) and (44) imply that

$$s_{i+1}^k - s_i^k \leq (\frac{5}{r})^{i-k}(s_{k+1}^k - s_k^k) - \frac{i-k}{in}(1 - s_i^k) \tag{45}$$

Note that $s_i^k = s_k^k + (s_{k+1}^k - s_k^k) + (s_{k+2}^k - s_{k+1}^k) + \ldots + (s_i^k - s_{i-1}^k)$. Thus (44) implies that

$$s_i^k \leq s_k^k + (s_{k+1}^k - s_k^k) \cdot (1 + \frac{5}{r} + \ldots + (\frac{5}{r})^{i-k-1})$$
$$\leq s_k^k + (s_{k+1}^k - s_k^k) \cdot \frac{r}{r-5} \tag{46}$$

Therefore (45) and (46) imply that

$$s_{i+1}^k - s_i^k \leq (\frac{5}{r})^{i-k}(s_{k+1}^k - s_k^k) - \frac{i-k}{in}(1 - s_k^k - (s_{k+1}^k - s_k^k) \cdot \frac{r}{r-5}) \tag{47}$$

Note that (47) holds also for $i = k$ and that in this case it becomes an equality. Summing up (47) for every $i \in \{k, \ldots, \frac{n+k}{2}\}$, it follows that

$$s_{\frac{n+k}{2}+1}^k - s_k^k \leq \frac{r}{r-5}(s_{k+1}^k - s_k^k) - (1 - s_k^k - (s_{k+1}^k - s_k^k) \cdot \frac{r}{r-5}) \sum_{i=k+1}^{\frac{n+k}{2}} \frac{i-k}{in} \tag{48}$$

14

Note now that for any positive numbers $x, y, z, w > 0$, it holds that $\frac{x}{y} + \frac{z}{w} > \frac{x+z}{y+w}$. Therefore, for every $i \in \{k+1, \ldots, \frac{n+k}{2}\}$,

$$\frac{i-k}{in} + \frac{(\frac{n+k}{2} - i + k + 1) - k}{(\frac{n+k}{2} - i + k + 1)n} > \frac{n-k+2}{n(n+3k+2)} > \frac{n-k}{n(n+3k)}$$

Thus

$$2 \sum_{i=k+1}^{\frac{n+k}{2}} \frac{i-k}{in} > (\frac{n+k}{2} - k) \cdot \frac{n-k}{n(n+3k)} = \frac{(n-k)^2}{2n(n+3k)}$$

i.e.

$$\sum_{i=k+1}^{\frac{n+k}{2}} \frac{i-k}{in} > \frac{(n-k)^2}{4n(n+3k)} \tag{49}$$

It follows now by (48) and (49) that

$$0 \le s_{\frac{n+k}{2}+1}^{k} - s_k^k \le \frac{r}{r-5}(s_{k+1}^k - s_k^k) - \frac{(n-k)^2}{4n(n+3k)}(1 - s_k^k - (s_{k+1}^k - s_k^k) \cdot \frac{r}{r-5})$$

$$= \frac{r}{r-5}(s_{k+1}^k - s_k^k)\frac{4n(n+3k) + (n-k)^2}{4n(n+3k)} - \frac{(n-k)^2}{4n(n+3k)}(1 - s_k^k)$$

Therefore

$$(n-k)^2(1 - s_k^k) \le \frac{r}{r-5}(s_{k+1}^k - s_k^k)(4n(n+3k) + (n-k)^2)$$

and thus

$$s_k^k \ge 1 - \frac{r}{r-5}(s_{k+1}^k - s_k^k)(1 + \frac{4n(n+3k)}{(n-k)^2}) \tag{50}$$

Now (50) and (36) imply that

$$s_k^k \ge 1 - \frac{r}{r-5}\frac{2}{n(r-1)}(1 + \frac{4n(n+3k)}{(n-k)^2}) \ge 1 - \frac{r}{r-5}\frac{2}{n(r-1)} \cdot 2\frac{4n(n+3k)}{(n-k)^2}$$

$$\ge 1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{n}{(n-k)^2}$$

$\square$

The next two corollaries follow now from Lemma 5 by direct substitution.

Corollary 3. *In the Markov chain $\mathcal{M}_k^2$, for any $r > 5$ and any $k \le \frac{n}{2}$,*

$$s_k^k \ge 1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{4}{n} = 1 - O(\frac{1}{n})$$

Corollary 4. *In the Markov chain $\mathcal{M}_k^2$, for any $r > 5$ and any $k \le n - \sqrt{n \log n}$,*

$$s_k^k \ge 1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{1}{\log n} = 1 - O(\frac{1}{\log n})$$

We now present an auxiliary lemma that provides a lower bound for the probability s_k^k, for any $k \le n - 1$.

Lemma 6. *In the Markov chain $\mathcal{M}_k^2$, for any $r > 5$ and any $k \le n - 1$,*

$$s_k^k \ge \frac{1}{n}$$

15

Proof. Let $1 \leq k \leq n - 1$. Recall by (37) and (38) in the proof of Lemma 5 that for $i = k + 1$,

$$s_{k+1}^k = \frac{\alpha_{k+1}^k \cdot s_{k+2}^k + \beta_{k+1}^k \cdot s_k^k + \gamma_{k+1}^k \cdot 1}{\alpha_{k+1}^k + \beta_{k+1}^k + \gamma_{k+1}^k}$$

Therefore, since $s_{k+2}^k \geq s_{k+1}^k$, it follows that

$$s_{k+1}^k \geq \frac{\beta_{k+1}^k \cdot s_k^k + \gamma_{k+1}^k \cdot 1}{\beta_{k+1}^k + \gamma_{k+1}^k} \tag{51}$$

In particular, it follows by (51) and (38) for $i = k + 1$ that

$$s_{k+1}^k \geq \frac{(n + (k+1)(n-k-1)) \cdot s_k^k + r \cdot 1}{n + (k+1)(n-k-1) + r} \tag{52}$$

Furthermore recall by (35) in the proof of Lemma 5 that

$$s_{k+1}^k \leq \frac{n(r-1)}{n(r-1)-2} s_k^k = (1 + \frac{2}{n(r-1)-2}) s_k^k \tag{53}$$

Thus (52) and (53) imply that

$$(1 + \frac{2}{n(r-1)-2}) s_k^k \geq \frac{(n + (k+1)(n-k-1)) \cdot s_k^k + r \cdot 1}{n + (k+1)(n-k-1) + r}$$

and thus

$$(n + (k+1)(n-k-1) + r)(1 + \frac{2}{n(r-1)-2}) s_k^k \geq (n + (k+1)(n-k-1)) s_k^k + r \Leftrightarrow$$

$$(n + (k+1)(n-k-1)) \frac{2}{n(r-1)-2} s_k^k + r(1 + \frac{2}{n(r-1)-2}) s_k^k \geq r \Leftrightarrow$$

$$(2(n + (k+1)(n-k-1)) + rn(r-1)) s_k^k \geq r(n(r-1) - 2)$$

Therefore

$$s_k^k \geq \frac{r(n(r-1) - 2)}{2(n + (k+1)(n-k-1)) + rn(r-1)}$$

Note now that $(k+1)(n-k-1) < n^2$, and thus the last inequality implies that

$$s_k^k \geq \frac{r(n(r-1) - 2)}{2(n + n^2) + rn(r-1)} \geq \frac{r(r-1) - \frac{2r}{n}}{2(n+1) + r(r-1)}$$

Therefore, since $r > 5$ and $r < n$ by assumption, it follows that

$$s_k^k \geq \frac{20 - 2}{2(n+1) + 20} = \frac{9}{n + 11} > \frac{1}{n}$$

$\square$

Urchin graphs are strong selective amplifiers In this section we conclude our analysis by combining the results of Section 3.2 on the two Markov chains $\mathcal{M}_k^1$ and $\mathcal{M}_k^2$. The Markov chain $\mathcal{M}$ is illustrated in Figure 5, where the transition from state P_0^k to the states P_k^k, P_0^{k-1} is done through the Markov chain $\mathcal{M}_k^1$, and the transition from state P_k^k to the states P_{k+1}^{k+1}, P_0^{k-1} is done through the Markov chain $\mathcal{M}_k^2$, respectively.

In Figure 5, the transition probability from state P_k^k to state P_{k+1}^{k+1} (resp. P_0^{k-1}) is s_k^k (resp. $1 - s_k^k$). Recall that s_k^k is the probability that, starting at P_k^k in $\mathcal{M}_k^2$ (and thus also in $\mathcal{M}$), we reach state
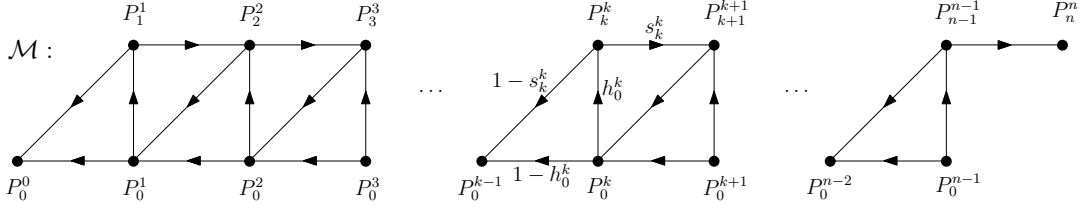
Fig. 5. The Markov chain $\mathcal{M}$, using the Markov chains $\mathcal{M}_k^1$ and $\mathcal{M}_k^2$, where $k \in \{1, 2, \ldots, n-1\}$.

F_{k+1} before we reach F_{k-1}. Furthermore, the transition probability from state P_0^k to state P_k^k is equal to the probability that, starting at P_0^k in $\mathcal{M}_k^1$, we reach P_k^k before we reach F_{k-1}. Note that this probability is larger than h_0^k. Therefore, in order to compute a lower bound of the fixation probability of a nose in G_n, we can assume that in $\mathcal{M}$ the transition probability from state P_0^k to P_k^k (resp. P_0^{k-1}) is h_0^k (resp. $1 - h_0^k$), as it is shown in Figure 5.

Note that for every $k \in \{2, \ldots, n-1\}$ the infected vertices of state P_0^k is a strict subset of the infected vertices of state P_k^k. Therefore, in order to compute a lower bound of the fixation probability of state P_0^1 in $\mathcal{M}$, we can relax $\mathcal{M}$ by changing every transition from state P_{k-1}^{k-1} to state P_k^k to a transition from state P_{k-1}^{k-1} to state P_0^k, where $k \in \{2, \ldots, n-1\}$. This relaxed Markov chain $\mathcal{M}'$ is illustrate in Figure 6(a). After eliminating the states P_k^k in $\mathcal{M}'$, where $k \in \{1, 2, \ldots, n-1\}$, we obtain the equivalent birth-death process $\mathcal{B}_n$ that is illustrated in Figure 6(b). Denote by p_1 the fixation probability of state P_0^1 in $\mathcal{B}_n$, i.e. p_1 is the probability that, starting at state P_0^1 in $\mathcal{B}_n$, we eventually arrive to state P_n^n.
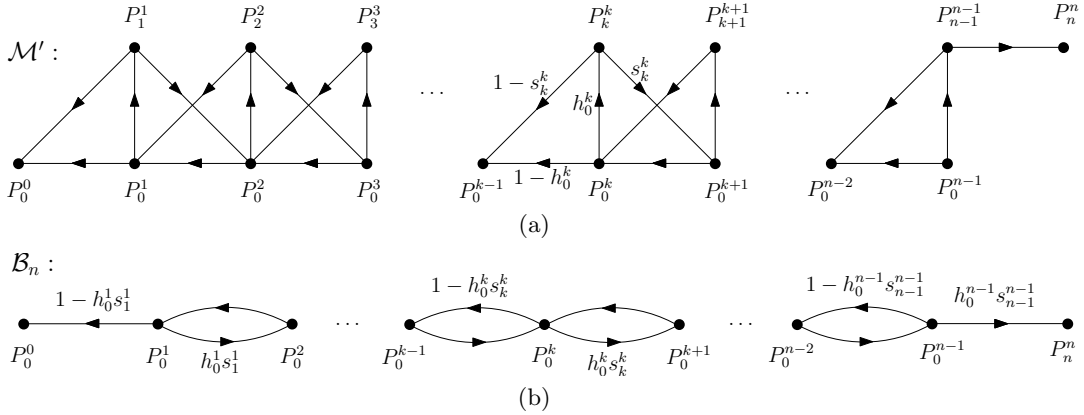

(a)

(b)

Fig. 6. (a) The relaxed Markov chain $\mathcal{M}'$ and (b) the birth-death process $\mathcal{B}_n$ that is obtained from $\mathcal{M}'$ after eliminating the states P_k^k in $\mathcal{M}'$, where $k \in \{1, 2, \ldots, n-1\}$.

Theorem 2. *For any $r > 5$ and for sufficiently large n, the fixation probability p_1 of state P_0^1 in $\mathcal{B}_n$ is $p_1 \geq 1 - \frac{c(r)}{n}$, for some appropriate function $c(r)$ of r.*

Proof. Denote by λ_k the forward bias of $\mathcal{B}_n$ at state P_0^k, i.e. $\lambda_k = \frac{h_0^k s_k^k}{1 - h_0^k s_k^k}$ is the ratio of the forward over the backward transition probability at state P_0^k. Then the fixation probability p_1 of state P_0^1 in $\mathcal{B}_n$ is

$$p_1 = \frac{1}{1 + \frac{1}{\lambda_1} + \frac{1}{\lambda_1 \lambda_2} + \frac{1}{\lambda_1 \lambda_2 \lambda_3} + \ldots + \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_{n-1}}} \tag{54}$$

17

Note now by Lemma 4 and Corollary 3 that for every $k \leq \frac{n}{2}$,

$$\lambda_k = \frac{h_0^k s_k^k}{1 - h_0^k s_k^k} \geq \frac{(1 - \frac{k+2}{n(r-1)})(1 - \frac{256r}{(r-5)(r-1)} \cdot \frac{1}{n})}{1 - (1 - \frac{k+2}{n(r-1)})(1 - \frac{256r}{(r-5)(r-1)} \cdot \frac{1}{n})}$$

$$\geq \frac{1 - \frac{256r}{(r-5)(r-1)} \cdot \frac{1}{n} - \frac{k+2}{n(r-1)}}{\frac{256r}{(r-5)(r-1)} \cdot \frac{1}{n} + \frac{k+2}{n(r-1)}} = \frac{n(r-1) - \frac{256r}{(r-5)} - (k+2)}{\frac{256r}{(r-5)} + (k+2)} \tag{55}$$

Therefore, since $k \leq \frac{n}{2}$ and $\frac{256r}{(r-5)} < \log n < \frac{n}{2} - 2$ for sufficiently large n, it follows by (55) that

$$\lambda_k > \frac{n(r-2)}{\log n + (k+2)} > \frac{n(r-2)}{2 \log n + k}$$

and thus

$$\frac{1}{\lambda_k} < \frac{2 \log n + k}{n(r-2)} \tag{56}$$

for every $k \leq \frac{n}{2}$. Furthermore, note by Lemma 4 and Corollary 4 that, whenever $\frac{n}{2} < k \leq n - \sqrt{n \log n}$,

$$\lambda_k = \frac{h_0^k s_k^k}{1 - h_0^k s_k^k} \geq \frac{(1 - \frac{k+2}{n(r-1)})(1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{1}{\log n})}{1 - (1 - \frac{k+2}{n(r-1)})(1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{1}{\log n})}$$

$$\geq \frac{1 - \frac{64r}{(r-5)(r-1)} \cdot \frac{1}{\log n} - \frac{k+2}{n(r-1)}}{\frac{64r}{(r-5)(r-1)} \cdot \frac{1}{\log n} + \frac{k+2}{n(r-1)}} = \frac{(r-1)\log n - \frac{64r}{(r-5)} - \frac{k+2}{n}\log n}{\frac{64r}{(r-5)} + \frac{k+2}{n}\log n}$$

Therefore, since $k + 2 < n$ and $\frac{64r}{(r-5)} < \frac{\log n}{r-2}$ for sufficiently large n, it follows that

$$\lambda_k > \frac{(r-1)\log n - \frac{\log n}{r-2} - \log n}{\frac{\log n}{r-2} + \log n} = r - 3$$

and thus

$$\frac{1}{\lambda_k} \leq \frac{1}{r-3} \tag{57}$$

whenever $\frac{n}{2} < k \leq n - \sqrt{n \log n}$. Moreover, note by Lemma 4 and Lemma 6 that, whenever $n - \sqrt{n \log n} < k \leq n - 1$,

$$\lambda_k = \frac{h_0^k s_k^k}{1 - h_0^k s_k^k} \geq \frac{(1 - \frac{k+2}{n(r-1)})\frac{1}{n}}{1 - (1 - \frac{k+2}{n(r-1)})\frac{1}{n}} > \frac{\frac{1}{n} - \frac{n+1}{n^2(r-1)}}{1 - \frac{1}{n} + \frac{n+1}{n^2(r-1)}} > \frac{1 - \frac{2}{r-1}}{n - 1 + \frac{2}{r-1}}$$

Note now that $1 - \frac{2}{r-1} > \frac{1}{2}$ since $r > 5$ by assumption, and thus the latter inequality implies that $\lambda_k > \frac{1/2}{n}$, i.e.

$$\frac{1}{\lambda_k} < 2n \tag{58}$$

whenever $n - \sqrt{n \log n} < k \leq n - 1$.

Since $r > 5$ by assumption, note now by (56) that $\frac{1}{\lambda_k} < \frac{3\log n}{n(r-2)} < \frac{\log n}{n}$ whenever $k \leq \log n$, and that $\frac{1}{\lambda_k} < \frac{3k}{n(r-2)} < \frac{k}{n}$ whenever $\log n < k \leq \frac{n}{2}$. Therefore, for every $k \in \{2, 3, \ldots, \log n\}$,

$$\frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \left(\frac{\log n}{n}\right)^k \leq \left(\frac{\log n}{n}\right)^2 \tag{59}$$

18

Furthermore, for every $k \in \{\log n + 1, \ldots, \frac{n}{2}\}$,

$$\frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \left(\frac{\log n}{n}\right)^{\log n} \prod_{i=\log n+1}^{k} \frac{i}{n} \tag{60}$$

$$< \left(\frac{\log n}{n}\right)^{\log n} < \left(\frac{\log n}{n}\right)^3 \tag{61}$$

Therefore, for every $k \in \{2, 3, \ldots, \frac{n}{2}\}$,

$$\sum_{k=2}^{\frac{n}{2}} \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \log n \left(\frac{\log n}{n}\right)^2 + \frac{n}{2}\left(\frac{\log n}{n}\right)^3 < \frac{1}{n} \tag{62}$$

for sufficiently large n. Note furthermore by (57) that $\frac{1}{\lambda_k} < 1$ whenever $\frac{n}{2} < k \leq n - \sqrt{n \log n}$, since $r > 5$ by assumption. Therefore, for every $k \in \{\frac{n}{2} + 1, \ldots, n - \sqrt{n \log n}\}$,

$$\frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_{2n/3}}$$

and thus it follows by (61) that

$$\sum_{k=\frac{n}{2}+1}^{n-\sqrt{n \log n}} \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \left(\frac{n}{2} - \sqrt{n \log n}\right) \left(\frac{\log n}{n}\right)^3 < \frac{\log^3 n}{2n^2} < \frac{1}{n} \tag{63}$$

for sufficiently large n. Let now $n - \sqrt{n \log n} < k \leq n - 1$. Then it follows by (58) and (60) that

$$\frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \left(\frac{\log n}{n}\right)^{\log n} \prod_{i=\log n+1}^{\frac{n}{2}} \frac{i}{n} \cdot (2n)^{k-n+\sqrt{n \log n}}$$

$$< \frac{2^{\left(\log n \log \log n + \sqrt{n \log n} + \log n \sqrt{n \log n}\right)}}{2^{n/2}} \tag{64}$$

However

$$\log n \log \log n + \sqrt{n \log n} + \log n \sqrt{n \log n} < \frac{n}{4}$$

for sufficiently large n, and thus (64) implies that

$$\frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \frac{1}{2^{n/4}}$$

for every $k \in \{n - \sqrt{n \log n} + 1, \ldots, n - 1\}$. Therefore

$$\sum_{k=n-\sqrt{n \log n}+1}^{n-1} \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \frac{n}{2^{n/4}} < \frac{1}{n} \tag{65}$$

for sufficiently large n. Thus, summing up (62), (63), and (65), it follows that

$$\sum_{k=2}^{n-1} \frac{1}{\lambda_1 \lambda_2 \lambda_3 \ldots \lambda_k} < \frac{3}{n} \tag{66}$$

For $k = 1$, (55) implies that

$$\lambda_k \geq \frac{n(r-1) - \frac{256r}{(r-5)} - 3}{\frac{256r}{(r-5)} + 3} > n\frac{(r-2)(r-5)}{259r - 15}$$

19

and thus

$$\frac{1}{\lambda_k} < \frac{259r - 15}{(r-2)(r-5)} \cdot \frac{1}{n} \qquad (67)$$

Summarizing, it follows by (54), (66), and (67) that

$$p_1 = \frac{1}{1 + \frac{1}{n} \cdot \left(3 + \frac{259r-15}{(r-2)(r-5)}\right)} \geq 1 - \frac{c(r)}{n}$$

where $c(r) = 3 + \frac{259r-15}{(r-2)(r-5)}$ is a function that inly depends on r. This completes the proof of the theorem. □

We are now ready to provide our main result in this section.

Theorem 3. *The class $\mathcal{G} = \{G_n : n \geq 1\}$ of urchin graphs is a class of strong selective amplifiers.*

Proof. Consider the urchin graph G_n, where $n \geq 1$. Let v be a nose in G_n. Then the fixation probability $f_r(v)$ of v in the generalized Moran process is greater than or equal to the fixation probability of state P_0^1 in the Markov chain $\mathcal{M}$ of Figure 5 (cf. Corollary 2 and the discussion after it in Section 3.2). Furthermore, the fixation probability of state P_0^1 in the Markov chain $\mathcal{M}$ is greater than or equal to the fixation probability p_1 of state P_0^1 in the birth-death process $\mathcal{B}_n$ in Figure 6(b). Therefore, since $p_1 \geq 1 - \frac{c(r)}{n}$ for any $r > 5$ by Theorem 2, it follows that $f_r(v) \geq 1 - \frac{c(r)}{n}$ for $r > r_0 = 5$ and sufficiently large n, where $c(r)$ is a function that depends only on r. Finally, since there exist exactly $\frac{n}{2}$ noses in G_n, it follows by Definition 2 that the class $\mathcal{G}$ of urchin graphs is a class of $(\frac{n}{2}, n)$-selective amplifiers, and thus $\mathcal{G}$ is a class of strong selective amplifiers. □

4 Suppressor bounds

In this section we prove our lower bound for the fixation probability of an arbitrary undirected graph, namely the *Thermal Theorem* (Section 4.1), which generalizes the analysis of the fixation probability of regular graphs [20]. Furthermore we present for every function $\phi(n)$, where $\phi(n) = \omega(1)$ and $\phi(n) \leq \sqrt{n}$, a class of $(\frac{n}{\phi(n)+1}, \frac{n}{\phi(n)})$-selective suppressors in Section 4.2.

4.1 The Thermal Theorem

Consider a graph $G = (V, E)$ and a fitness value $r > 1$. Denote by $\mathcal{M}_r(G)$ the generalized Moran process on G with fitness r. Then, for every subset $S \notin \{\emptyset, V\}$ of its vertices, the fixation probability $f_r(S)$ of S in $\mathcal{M}_r(G)$ is given by (1), where $f_r(\emptyset) = 0$ and $f_r(V) = 1$. That is, the fixation probabilities $f_r(S)$, where $S \notin \{\emptyset, V\}$, are the solution of the linear system (1) with boundary conditions $f_r(\emptyset) = 0$ and $f_r(V) = 1$.

Suppose that at some iteration of the generalized Moran process the set S of vertices are infected and that the edge $xy \in E$ (where $x \in S$ and $y \notin S$) is activated, i.e. either x infects y or y disinfects x. Then (1) implies that the probability that x infects y is higher if $\frac{1}{\deg x}$ is large; similarly, the probability that y disinfects x is higher if $\frac{1}{\deg y}$ is large. Therefore, in a fashion similar to [20], we call for every vertex $v \in V$ the quantity $\frac{1}{\deg v}$ the *temperature* of v: a "hot" vertex (i.e. with high temperature) affects more often its neighbors than a "cold" vertex (i.e. with low temperature).

Before we proceed, recall that $f_r(\emptyset) = 0$ and $f_r(V) = 1$. Furthermore, recall by (1) that, for every vertex subset $S \notin \{\emptyset, V\}$,

$$f_r(S) = \frac{\sum_{xy \in E, x \in S, y \notin S} \left(\frac{r}{\deg x} f_r(S+y) + \frac{1}{\deg y} f_r(S-x)\right)}{\sum_{xy \in E, x \in S, y \notin S} \left(\frac{r}{\deg x} + \frac{1}{\deg y}\right)}$$

20

Note that the summation in both the nominator and the denominator is done over all edges $xy \in E$ which have one endpoint $x \in S$ and one endpoint $y \notin S$. Therefore, for every subset $S \notin \{\emptyset, V\}$, there exists at least one such pair $x(S), y(S)$ of vertices, where $x(S) \in S$, $y(S) \notin S$, and $x(S)y(S) \in E$, such that

$$f_r(S) \geq \frac{\frac{r}{\deg x(S)} f_r(S + y(S)) + \frac{1}{\deg y(S)} f_r(S - x(S))}{\frac{r}{\deg x(S)} + \frac{1}{\deg y(S)}} \tag{68}$$

Thus, solving the linear system that is obtained from (68) by replacing inequalities with equalities, we obtain a lower bound for the fixation probabilities $f_r(S)$, where $S \notin \{\emptyset, V\}$.

In the next definition we introduce a weighted generalization of this linear system, which is a crucial tool for our analysis in obtaining the Thermal Theorem. Note that in Definition 3, as well as in the remainder of this section, for every vertex subset $S \notin \{\emptyset, V\}$ we consider the above two vertices $x(S)$ and $y(S)$ as fixed.

Definition 3 (the linear system L_0). *Let $G = (V, E)$ be an undirected graph and $r > 1$. Let every vertex $v \in V$ have weight (temperature) $d_v > 0$. The linear system L_0 on the variables $p_r(S)$, where $S \subseteq V$, is given by the following equations whenever $S \notin \{\emptyset, V\}$:*

$$p_r(S) = \frac{r d_{x(S)} p_r(S + y(S)) + d_{y(S)} p_r(S - x(S))}{r d_{x(S)} + d_{y(S)}} \tag{69}$$

with boundary conditions $p_r(\emptyset) = 0$ and $p_r(V) = 1$.

With a slight abuse of notation, whenever $S = \{u_1, u_2, \ldots, u_k\}$, we denote $p_r(u_1, u_2, \ldots, u_k) = p_r(S)$.

Observation 1 *The linear system L_0 in Definition 3 corresponds naturally to the Markov chain $\mathcal{M}_0$ with one state for every subset $S \subseteq V$, where the states $\emptyset$ and V are absorbing, and every non-absorbing state S has exactly two transitions to the states $S + y(S)$ and $S - x(S)$ with transition probabilities $q_S = \frac{r d_{x(S)}}{r d_{x(S)} + d_{y(S)}}$ and $1 - q_S$, respectively.*

Observation 2 *Let $G = (V, E)$ be a graph and $r > 1$. For every vertex $x \in V$ let $d_x = \frac{1}{\deg x}$ be the temperature of x. Then $f_r(S) \geq p_r(S)$ for every $S \subseteq V$, where the values $p_r(S)$ are the solution of the linear system L_0.*

Before we provide the Thermal Theorem (Theorem 4), we first prove an auxiliary result in the next lemma which generalizes the Isothermal Theorem of [20] for regular graphs, i.e. for graphs with the same number of neighbors for every vertex.

Lemma 7. *Let $G = (V, E)$ be a graph with n vertices, $r > 1$, and d_u be the same for all vertices $u \in V$. Then for every vertex $u \in V$,*

$$p_r(u) = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^n}} \geq 1 - \frac{1}{r}$$

Proof. Since d_u is the same for all vertices $u \in V$, it follows by (69) that for every set $S \notin \{\emptyset, V\}$, the forward probability is $q_S = \frac{r}{r+1}$ and the backward probability is $1 - q_S = \frac{1}{r+1}$. Therefore, by symmetry, $p_r(S) = p_r(S')$ whenever $|S| = |S'|$. For every $0 \leq k \leq n$ denote by $p_k = p_r(S)$, where $|S| = k$. Note that $p_0 = 0$ and $p_n = 1$. Then it follows by (69) that, whenever $1 \leq k \leq n - 1$,

$$p_{k+1} - p_k = \frac{1}{r}(p_k - p_{k-1}) = \ldots = \frac{1}{r^k}(p_1 - p_0)$$

Therefore, summing up these equations for every $1 \leq k \leq n - 1$ it follows that

$$p_n - p_1 = (p_1 - p_0)(\frac{1}{r} + \frac{1}{r^2} + \ldots + \frac{1}{r^{n-1}})$$

and thus, since $p_0 = 0$ and $p_n = 1$,

$$p_1 = \frac{1}{1 + \frac{1}{r} + \frac{1}{r^2} + \ldots + \frac{1}{r^{n-1}}} \geq 1 - \frac{1}{r}$$

$\square$

We are now ready to provide our main result in this section which provides a lower bound for the fixation probability on arbitrary graphs, parameterized by the maximum ratio between two different temperatures in the graph.

Theorem 4 (Thermal Theorem). *Let $G = (V, E)$ be a connected undirected graph and $r > 1$. Then $f_r(v) \geq \frac{r-1}{r + \frac{\deg v}{\deg_{\min}}}$ for every $v \in V$.*

Proof. Let G have n vertices, i.e. $|V| = n$. Our proof is based on the linear system L_0 of Definition 3. Namely, we consider the linear system L_0 with weight $d_v = \frac{1}{\deg v}$ for every vertex $v \in V$. Note that $d_{\min} = \frac{1}{\deg_{\max}}$ and $d_{\max} = \frac{1}{\deg_{\min}}$. Recall that $\mathcal{M}_0$ is the Markov chain that can be defined from the linear system L_0 (cf. Observation 1), and that every state $S \notin \{\emptyset, V\}$ of $\mathcal{M}_0$ has exactly two transitions, namely to states $S + y(S)$ and $S - x(S)$.

We now define the Markov chain $\mathcal{M}_0^*$ from $\mathcal{M}_0$ as follows. Consider an arbitrary state $S \subseteq V$ such that $1 \leq |S| \leq n - 2$. Denote $x(S) = u$ and $y(S) = v$ (note that $v \notin S$). Furthermore denote $x(S + v) = x_0$ and $y(S + v) = y_0$. Then perform the following changes to the Markov chain $\mathcal{M}_0$:

Step A. add a new dummy state X_S to $\mathcal{M}_0$,

Step B. replace the transition from S to $S + v$ by a transition from S to X_S (with the same transition probability q_S),

Step C. add to state X_S the transitions to states $S + v + y_0$ and $S + v - x_0$, with transition probabilities q_{S+v} and $1 - q_{S+v}$, respectively.

An example of the application of the above Steps A, B, C is illustrated in Figure 7. Denote by $\mathcal{M}_0^*$ the Markov chain obtained after applying these steps to $\mathcal{M}_0$ for every state $S \subseteq V$ with $1 \leq |S| \leq n - 2$. Furthermore denote by L_0^* the linear set that corresponds to $\mathcal{M}_0^*$. Note that L_0^* has the additional variables $\{p_r(X_S) : 1 \leq |S| \leq n - 2\}$ that do not exist in L_0. Moreover, for every state X_S of $\mathcal{M}_0^*$, note by the construction of $\mathcal{M}_0^*$ that $p_r(X_S) = p_r(S + y(S))$ in the solution of L_0^*.
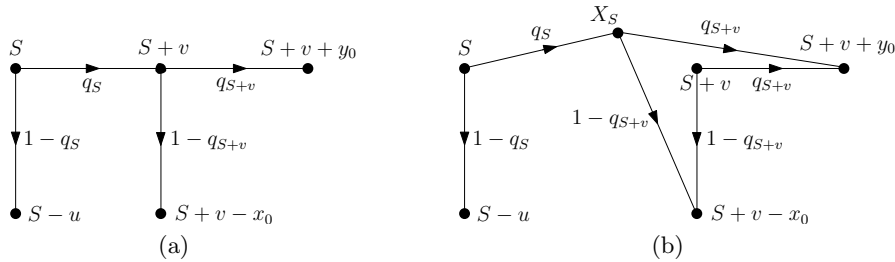

Fig. 7. Parts of (a) the Markov chain $\mathcal{M}_0$ and of (b) the Markov chain $\mathcal{M}_0^*$ (after the execution of Steps A, B, and C).

In the remainder of the proof we fix an arbitrary vertex v_0 and we prove that $f_r(v_0) \geq \frac{r-1}{r + \frac{\deg v}{\deg_{\min}}}$. To do so, we first consider an arbitrary numbering $v_0, v_1, \ldots, v_{n-1}$ of the n vertices of G, starting at our fixed vertex v_0. Then, starting from the Markov chain $\mathcal{M}_0^*$ that we described above (cf. Figure 7), we iteratively construct a sequence of Markov chains $\mathcal{M}_1^*, \mathcal{M}_2^*, \ldots, \mathcal{M}_{n-1}^*$, which correspond to a sequence of linear systems $L_1^*, L_2^*, \ldots, L_{n-1}^*$, respectively (cf. Observation 1). Each Markov chain $\mathcal{M}_i^*$

is obtained from the previous chain $\mathcal{M}^*_{i-1}$ by applying two local replacement rules (see Step 1 and Step 2 below) to some transitions of $\mathcal{M}^*_{i-1}$ which involve vertex v_i. Denote by $p^i_r(v_0)$ the value of $p_r(v_0)$ in the solution of the system L^*_i. We can prove that $p^0_r(v_0) \geq \ldots \geq p^{n-2}_r(v_0) \geq p^{n-1}_r(v_0)$. Therefore, since $f_r(v_0) \geq p^0_r(v_0)$ by Observation 2, each of the values $p^i_r(v_0)$ is smaller than or equal to the fixation probability $f_r(v_0)$ of v. Moreover, it turns out that the value $p^i_r(v_0)$ is a *monotone decreasing* function of d_{v_i}. Thus we increase the value of d_{v_i} to $d_{\max} = \frac{1}{\deg_{\min}}$ in L^*_i and the value $p^i_r(v_0)$ decreases even more after this change. Using this fact we can prove at the end our desired lower bound $f_r(v_0) \geq \frac{r-1}{r + \frac{\deg v}{\deg_{\min}}}$.

For every $i \in \{1, 2, \ldots, n-1\}$, we iteratively construct the Markov chain $\mathcal{M}^*_i$ from the corresponding Markov chain $\mathcal{M}^*_{i-1}$, as follows. Consider a state $S \notin \{\emptyset, V\}$, where $y(S) = v_i$ (note that in this case $v_i \notin S$). Denote $x(S) = u$, $x(S + v_i) = x_0$, and $y(S + v_i) = y_0$. Then perform the following changes to the Markov chain $\mathcal{M}^*_{i-1}$:

Step 1. replace the transition from X_S to $S + v_i + y_0$ by a transition from X_S to $S + y_0$ (with the same transition probability q_{S+v_i}),

Step 2. replace the transition from X_S to $S + v_i - x_0$ by a transition from X_S to $S - x_0$ (with the same transition probability $1 - q_{S+v_i}$).

Denote by $\mathcal{M}^*_i$ the Markov chain obtained by iteratively applying Steps 1 and 2 to $\mathcal{M}^*_{i-1}$ for every state $S \notin \{\emptyset, V\}$ with $y(S) = v_i$. Note that for every state $S_1 \notin \{\emptyset, V\}$ and for every state $S_2 \neq V$, where $v_i \in S_2$ and $v_i \notin S_1$, there exists no transition path in $\mathcal{M}^*_i$ from S_1 to S_2. Furthermore denote by L^*_i the linear system that corresponds to $\mathcal{M}^*_i$ (cf. Observation 1). Note that, in the above Step 2, if $x_0 = v_i \notin S$ then $S - x_0 = S$, and thus we have in this case a transition from X_S to S in $\mathcal{M}^*_i$. The changes made to $\mathcal{M}^*_{i-1}$ by Steps 1 and 2 are illustrated in Figures 8 and 9 for the cases where $x_0 = v_i$ and $x_0 \neq v_i$, respectively.
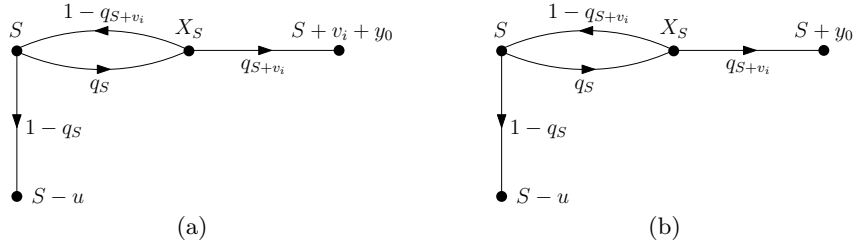

Fig. 8. Parts of (a) the Markov chain $\mathcal{M}^*_{i-1}$ and of (b) the Markov chain $\mathcal{M}^*_i$, where $x_0 = v_i$.
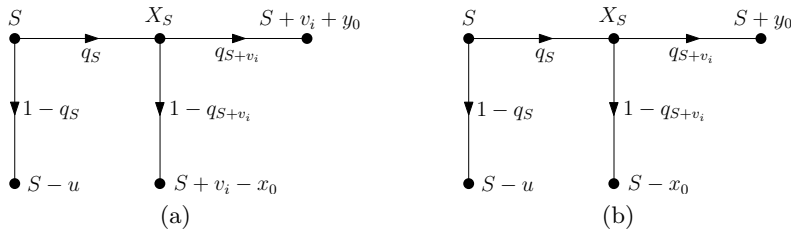

Fig. 9. Parts of (a) the Markov chain $\mathcal{M}^*_{i-1}$ and of (b) the Markov chain $\mathcal{M}^*_i$, where $x_0 \neq v_i$.

For every state S (resp. X_S) of the Markov chain $\mathcal{M}^*_i$, and for any $i \in \{0, 1, \ldots, n-1\}$, we denote in the following by $p^i_r(S)$ (resp. $p^i_r(X_S)$) the value of $p_r(S)$ (resp. $p_r(X_S)$) in the solution of the linear system L^*_i. Note that for every state $S \notin \{\emptyset, V\}$ with $y(S) = v_i$, the transitions of

state X_S in $\mathcal{M}_{i-1}^*$ are to the states $S + v_i + y_0$ and $S + v_i - x_0$, while the transitions of state X_S in $\mathcal{M}_i^*$ are to the states $S + y_0 \subseteq S + v_i + y_0$ and $S - x_0 \subseteq S + v_i - x_0$, respectively. Therefore $p_r^i(X_S) \le p_r^{i-1}(X_S)$, and thus $p_r^i(S) \le p_r^{i-1}(S)$. Similarly $p_r^i(T) \le p_r^{i-1}(T)$, for every state $T \subseteq V$ with $v_i \notin T$. Thus in particular $p_r^i(v_j) \le p_r^{i-1}(v_j)$ for every $j \neq i$.

In order to continue our analysis, we distinguish now the cases where $x_0 = v_i$ and $x_0 \neq v_i$.

Case 1: $x_0 = v_i$, cf. Figure 8. In this case we have in the linear system L_i^* that

$$p_r^i(S) = q_S p_r^i(X_S) + (1 - q_S) p_r^i(S - u)$$
$$p_r^i(X_S) = q_{S+v_i} p_r^i(S + y_0) + (1 - q_{S+v_i}) p_r^i(S)$$

and thus

$$p_r^i(S) = \frac{q_S q_{S+v_i} p_r^i(S + y_0) + (1 - q_S) p_r^i(S - u)}{q_S q_{S+v_i} + (1 - q_S)}$$

where

$$q_S = \frac{r d_u}{r d_u + d_{v_i}}$$

$$q_{S+v_i} = \frac{r d_{v_i}}{r d_{v_i} + d_{y_0}}$$

Therefore the forward probability of state S in $\mathcal{M}_i^*$ is (after eliminating the state X_S) equal to

$$\frac{q_S q_{S+v_i}}{q_S q_{S+v_i} + (1 - q_S)} = \frac{r^2 d_u}{r^2 d_u + r d_{v_i} + d_{y_0}} \tag{70}$$

Case 2: $x_0 \neq v_i$, cf. Figure 9. In this case we have in the linear system L_i^* that

$$p_r^i(S) = q_S p_r^i(X_S) + (1 - q_S) p_r^i(S - u)$$
$$p_r^i(X_S) = q_{S+v_i} p_r^i(S + y_0) + (1 - q_{S+v_i}) p_r^i(S - x_0)$$

and thus

$$p_r^i(S) = q_S q_{S+v_i} p_r^i(S + y_0) + q_S(1 - q_{S+v_i}) p_r^i(S - x_0) + (1 - q_S) p_r^i(S - u)$$

where

$$q_S = \frac{r d_u}{r d_u + d_{v_i}}$$

$$q_{S+v_i} = \frac{r d_{x_0}}{r d_{x_0} + d_{y_0}}$$

Therefore the forward probability of state S in $\mathcal{M}_i^*$ is (after eliminating the state X_S) equal to

$$q_S q_{S+v_i} = \frac{r d_u}{r d_u + d_{v_i}} \cdot \frac{r d_{x_0}}{r d_{x_0} + d_{y_0}} \tag{71}$$

It follows now by Cases 1 and 2 (cf. (70) and (71)) that the forward probability of state S in $\mathcal{M}_i^*$ (after eliminating the state X_S) is a monotone decreasing function of d_{v_i}. Therefore, for every state $S' \subseteq V$ with $v_i \notin S'$, the value $p_r^i(S')$ is also a monotone decreasing function of d_{v_i}. Thus, in particular, also the value $p_r^i(v_j)$, where $j \neq i$, is a monotone decreasing function of d_{v_i}. We now increase the value of d_{v_i} to $d_{\max}$ in L_i^*. Thus for every $j \neq i$, the value $p_r^i(v_j)$ decreases after this change.

Recall that $f_r(v_0) \ge p_r^0(v_0)$ by Observation 2. Therefore, since also $p_r^0(v_0) \ge \ldots \ge p_r^{n-2}(v_0) \ge p_r^{n-1}(v_0)$, it follows that $f_r(v_0) \ge p_r^{n-1}(v_0)$, i.e. $p_r^{n-1}(v_0)$ is a lower bound for the fixation probability

$f_r(v_0)$ in the Markov chain $\mathcal{M}_r(G)$. Furthermore $d_{v_1} = d_{v_2} = \ldots = d_{v_{n-1}} = d_{\max}$ in the linear system L_{n-1}^*. Consider now the state $S = \{v_0\}$ in the Markov chain $\mathcal{M}_{n-1}^*$, and let $y(S) = v_{i_0}$, where $1 \leq i_0 \leq n - 1$. Note that $x(S) = v_0$. Then the value $p_r^{n-1}(v_0)$ equals

$$p_r^{n-1}(v_0) = \frac{rd_{v_0} p_r^{n-1}(v_0, v_{i_0}) + d_{v_{i_0}} p_r^{n-1}(\emptyset)}{rd_{v_0} + d_{v_{i_0}}}$$

$$\geq \frac{rd_{v_0}}{rd_{v_0} + d_{\max}} p_r^{n-1}(v_{i_0}) \tag{72}$$

cf. Definition 3. Recall that $d_{v_0} = \frac{1}{\deg v_0}$ and $d_{\max} = \frac{1}{\deg_{\min}}$ by definition. Thus, since $f_r(v_0) \geq p_r^{n-1}(v_0)$ as we proved above, (72) implies that

$$f_r(v_0) \geq p_r^{n-1}(v_0) \geq \frac{r}{r + \frac{\deg v_0}{\deg_{\min}}} p_r^{n-1}(v_{i_0}) \tag{73}$$

Now, similarly to the above transformations of the linear system L_{i-1}^* to L_i^*, where $1 \leq i \leq n - 1$, we construct the linear system L_n^* (and the corresponding Markov chain $\mathcal{M}_n^*$) from L_{n-1}^* (and from the corresponding Markov chain $\mathcal{M}_{n-1}^*$), by applying iteratively the above Steps 1 and 2 to the states $S \subseteq V$, where $y(S) = v_0$ (instead of $y(S) = v_i$ above). Furthermore we increase the value of d_{v_0} to $d_{\max}$ in the resulting linear system L_n^*. Then, similarly to the construction of L_i^*, where $1 \leq i \leq n - 1$, it follows that $p_r^n(v_j) \leq p_r^{n-1}(v_j)$ for every $j \neq 0$. Thus, in particular, $p_r^n(v_{i_0}) \leq p_r^{n-1}(v_{i_0})$. Furthermore $d_{v_0} = d_{v_1} = \ldots = d_{v_{n-1}} = d_{\max}$ in L_n^*, and thus $p_r^n(v_{i_0}) \geq 1 - \frac{1}{r}$ by Lemma 7. Therefore, since $p_r^n(v_{i_0}) \leq p_r^{n-1}(v_{i_0})$, it follows by (73) that

$$f_r(v_0) \geq \frac{r}{r + \frac{\deg v_0}{\deg_{\min}}} p_r^n(v_{i_0}) \geq \frac{(r-1)}{r + \frac{\deg v_0}{\deg_{\min}}} \tag{74}$$

Since v_0 has been chosen arbitrarily, this completes the proof of the theorem. $\qquad \square$

The lower bound for the fixation probability in Theorem 4 is almost tight. Indeed, if a graph $G = (V, E)$ with n vertices is regular, i.e. if $\deg u = \deg v$ for every $u, v \in V$, then $f_r(G) = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^n}}$ by Lemma 7 (cf. also the Isothermal Theorem in [20]), and thus $f_r(G) \cong \frac{r-1}{r}$ for large enough n. On the other hand, Theorem 4 implies for a regular graph G that $f_r(G) \geq \frac{r-1}{r+1}$.

4.2   A class of selective suppressors

In this section we present for every function $\phi(n)$, where $\phi(n) = \omega(1)$ and $\phi(n) \leq \sqrt{n}$, the class $\mathcal{G}_{\phi(n)} = \{G_{\phi(n),n} : n \geq 1\}$ of $(\frac{n}{\phi(n)+1}, \frac{n}{\phi(n)})$-selective suppressors. We call these graphs $\phi(n)$-*urchin graphs*, since for $\phi(n) = 1$ they coincide with the class of urchin graphs in Section 3.2. For every n, the graph $G_{\phi(n),n} = (V_{\phi(n),n}, E_{\phi(n),n})$ has n vertices. Its vertex set $V_{\phi(n),n}$ can be partitioned into two sets $V_{\phi(n),n}^1$ and $V_{\phi(n),n}^2$, where $|V_{\phi(n),n}^1| = \frac{n}{\phi(n)+1}$ and $|V_{\phi(n),n}^2| = \frac{\phi(n)}{\phi(n)+1} n$, such that $V_{\phi(n),n}^1$ induces a clique and $V_{\phi(n),n}^2$ induces an independent set in $G_{\phi(n),n}$. Furthermore, every vertex $u \in V_{\phi(n),n}^2$ has $\phi(n)$ neighbors in $V_{\phi(n),n}^1$, and every vertex $v \in V_{\phi(n),n}^1$ has $\phi^2(n)$ neighbors in $V_{\phi(n),n}^2$. Therefore $\deg v = n + \phi^2(n) - 1$ for every $v \in V_{\phi(n),n}^1$ and $\deg u = \phi(n)$ for every $u \in V_{\phi(n),n}^2$. An example of a graph $G_{\phi(n),n}$ is illustrated in Figure 10(a).

Lemma 8. *For every $v \in V_{\phi(n),n}^1$ and sufficiently large n,*
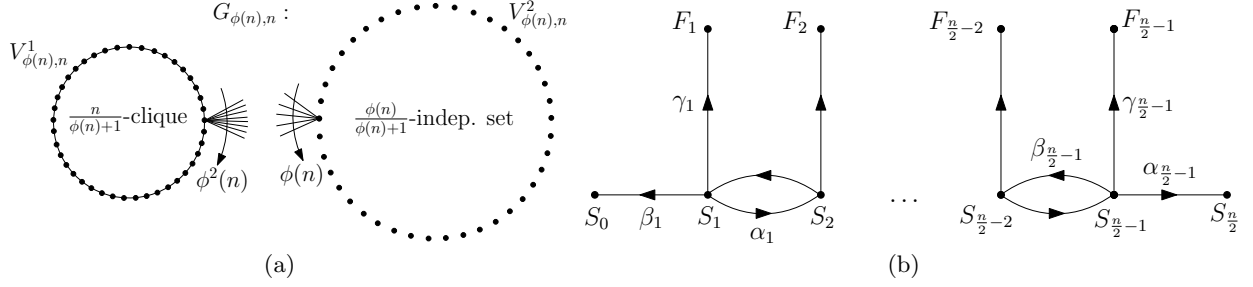
$$f_r(v) < 5r \cdot \frac{\phi(n)}{n}$$

Fig. 10. (a) The graph $G_{\phi(n),n}$ with n vertices and (b) the relaxed Markov chain.

Proof. Denote by S_k the state, in which exactly $k \geq 0$ vertices of $V_{\phi(n),n}^1$ are infected and all vertices of $V_{\phi(n),n}^2$ are not infected. Note that S_0 is the empty state. Furthermore denote by F_k the state where exactly $k \geq 0$ vertices of $V_{\phi(n),n}^1$ and at least one vertex of $V_{\phi(n),n}^2$ are infected. In order to compute an upper bound for the fixation probability $f_r(S_1)$ (i.e. of the fixation probability $f_r(v)$ where $v \in V_{\phi(n),n}^1$), we can set the value $f_r(S_{\frac{n}{2}})$ and the values $f_r(F_k)$ for every $k \geq 1$ to their trivial upper bound 1. That is, we assume that the state $S_{\frac{n}{2}}$, as well as all states F_k, where $k \geq 1$, are absorbing. After performing these relaxations (and eliminating self loops), we obtain a Markov chain, whose state graph is illustrated in Figure 10(b). For any $1 \leq k \leq \frac{n}{2} - 1$ in this relaxed Markov chain,

$$f_r(S_k) = \alpha_k f_r(S_{k+1}) + \beta_k f_r(S_{k-1}) + \gamma_k \tag{75}$$

where

$$\alpha_k = \frac{rk\frac{n-k}{n+\phi^2(n)-1}}{\sum_k}$$

$$\beta_k = \frac{k\left(\frac{\phi^2(n)}{\phi(n)} + \frac{n-k}{n+\phi^2(n)-1}\right)}{\sum_k} \tag{76}$$

$$\gamma_k = \frac{rk\frac{\phi^2(n)}{n+\phi^2(n)-1}}{\sum_k}$$

where $\sum_k = rk\frac{n-k}{n+\phi^2(n)-1} + k\left(\frac{\phi^2(n)}{\phi(n)} + \frac{n-k}{n+\phi^2(n)-1}\right) + rk\frac{\phi^2(n)}{n+\phi^2(n)-1}$. Note now by (76) that

$$\frac{\beta_k}{\alpha_k} = \frac{1}{r}\left(1 + \frac{\phi^2(n)\left(n + \phi^2(n) - 1\right)}{\phi(n)\left(n - k\right)}\right) > \frac{\phi(n)}{r} > 1 \tag{77}$$

Furthermore, since $k < \frac{n}{2}$, it follows by (76) that

$$\frac{\gamma_k}{\alpha_k} = \frac{\phi^2(n)}{n-k} < 2\frac{\phi^2(n)}{n} \tag{78}$$

Now, since $\alpha_k + \beta_k + \gamma_k = 1$ and $f_r(S_k) \geq f_r(S_{k-1})$ for every k, (75) implies by (77) and (78) that

$$\begin{aligned} f_r(S_{k+1}) - f_r(S_k) &= \frac{\beta_k}{\alpha_k}(f_r(S_k) - f_r(S_{k-1})) - \frac{\gamma_k}{\alpha_k}(1 - f_r(S_k)) \\ &> \frac{\phi(n)}{r}(f_r(S_k) - f_r(S_{k-1})) - 2\frac{\phi^2(n)}{n} \\ &> \dots \\ &> \left(\frac{\phi(n)}{r}\right)^k \cdot f_r(S_1) - 2\frac{\phi^2(n)}{n} \cdot \frac{\left(\frac{\phi(n)}{r}\right)^k - 1}{\frac{\phi(n)}{r} - 1} \end{aligned}$$

26

Thus, since $f_r(S_{\frac{n}{2}}) = 1$ in the relaxed Markov chain, we have that

$$1 - f_r(S_1) = \sum_{k=1}^{\frac{n}{2}-1} (f_r(S_{k+1}) - f_r(S_k))$$

$$> \sum_{k=1}^{\frac{n}{2}-1} \left[\left(\frac{\phi(n)}{r}\right)^k \cdot f_r(S_1) - 2\frac{\phi^2(n)}{n} \cdot \frac{\left(\frac{\phi(n)}{r}\right)^k - 1}{\frac{\phi(n)}{r} - 1} \right]$$

Therefore

$$f_r(S_1) \sum_{k=0}^{\frac{n}{2}-1} \left(\frac{\phi(n)}{r}\right)^k < 1 + 2\frac{\phi^2(n)}{n\left(\frac{\phi(n)}{r} - 1\right)} \sum_{k=0}^{\frac{n}{2}-1} \left[\left(\frac{\phi(n)}{r}\right)^k - 1 \right]$$

and thus

$$f_r(S_1) < 2\frac{\phi^2(n)}{n\left(\frac{\phi(n)}{r} - 1\right)} + \frac{1}{\sum_{k=0}^{n/2-1} \left(\frac{\phi(n)}{r}\right)^k}$$

$$= 2r\frac{\phi^2(n)}{n\left(\phi(n) - r\right)} + \frac{1}{\sum_{k=0}^{n/2-1} \left(\frac{\phi(n)}{2}\right)^k}$$

Therefore, since $\phi(n) = \omega(1)$ and r is constant by assumption, it follows that $r \leq \frac{\phi(n)}{2}$ for sufficiently large n, and thus

$$f_r(S_1) < 4r\frac{\phi(n)}{n} + \frac{1}{n} < 5r\frac{\phi(n)}{n}$$

$\square$

Using Lemma 8 we can now prove the next theorem.

Theorem 5. *For every function $\phi(n)$, where $\phi(n) = \omega(1)$ and $\phi(n) \leq \sqrt{n}$, the class $\mathcal{G}_{\phi(n)} = \{G_{\phi(n),n} : n \geq 1\}$ of $\phi(n)$-urchin graphs is a class of $(\frac{n}{\phi(n)+1}, \frac{n}{\phi(n)})$-selective suppressors.*

Proof. It follows by Lemma 8 that, if $v \in V_{\phi(n),n}^1$, then $f_r(v) < 5r\frac{\phi(n)}{n} = \frac{5r}{n/\phi(n)}$ for any $r > 1$ and sufficiently large n. Therefore, since $|V_{\phi(n),n}^1| = \frac{n}{\phi(n)+1}$ for every graph $G_{\phi(n),n}$, it follows by Definition 2 that the class $\mathcal{G}_{\phi(n)}$ of graphs is a class of $(\frac{n}{\phi(n)+1}, \frac{n}{\phi(n)})$-selective suppressors. $\square$

References

1. D. Aldous and J. Fill. *Reversible Markov chains and random walks on graphs*. 2002. Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/aldous/RWG/book.html.

2. T. Antal and I. Scheuring. Fixation of strategies for an evolutionary game in finite populations. *Bulletin of Mathematical Biology*, 68:1923–1944, 2006.

3. M. Broom, C. Hadjichrysanthou, and J. Rychtar. Evolutionary games on graphs and the speed of the evolutionary process. *Proceedings of the Royal Society A*, 466(2117):1327–1346, 2010.

4. M. Broom, C. Hadjichrysanthou, and J. Rychtar. Two results on evolutionary processes on general non-directed graphs. *Proceedings of the Royal Society A*, 466(2121):2795–2798, 2010.

5. M. Broom and J. Rychtar. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A*, 464(2098):2609–2627, 2008.

6. M. Broom, J. Rychtar, and B. Stadler. Evolutionary dynamics on small order graphs. *Journal of Interdisciplinary Mathematics*, 12:129–140, 2009.

7. A. S. Christine Taylor, Drew Fudenberg and M. A. Nowak. Evolutionary game dynamics in finite populations. *Bulletin of Mathematical Biology*, 66(6):1621–1644, 2004.

8. J. Diáz, L. Goldberg, G. Mertzios, D. Richerby, M. Serna, and P. Spirakis. On the fixation probability of superstars. *Proceedings of the Royal Society A*, 469(2156):193–203, 2013.

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

10. R. Durrett. *Lecture notes on particle systems and percolation*. Wadsworth Publishing Company, 1988.

11. D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

12. A. Galanis, A. Göbel, L. A. Goldberg, J. Lapinskas, and D. Richerby. Amplifiers for the Moran process. *Journal of the ACM*, 64(1):5:1–5:90, 2017.

13. G. Giakkoupis. Amplifiers and suppressors of selection for the Moran process on undirected graphs. *CoRR*, 2016. Available online at http://arxiv.org/abs/1611.01585.

14. H. Gintis. *Game theory evolving: A problem-centered introduction to modeling strategic interaction*. Princeton University Press, 2000.

15. L. A. Goldberg, J. Lapinskas, J. Lengler, F. Meier, K. Panagiotou, and P. Pfister. Asymptotically optimal amplifiers for the Moran process. *CoRR*, 2016. Available online at http://arxiv.org/abs/1611.04209.

16. J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

17. L. A. Imhof. The long-run behavior of the stochastic replicator dynamics. *Annals of applied probability*, 15(1B):1019–1045, 2005.

18. M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):29–56, 1993.

19. S. Karlin and H. Taylor. *A First Course in Stochastic Processes*. NY: Academic Press, 2nd edition, 1975.

20. E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–316, 2005.

21. T. M. Liggett. *Interacting Particle Systems*. Springer-Verlag, 1985.

22. G. B. Mertzios, S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. Natural models for evolution on networks. *Theoretical Computer Science*, 477:76–95, 2013.

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

24. M. A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, 2006.

25. H. Ohtsuki and M. A. Nowak. Evolutionary games on cycles. In *Proceedings of the Royal Society B: Biological Sciences*, volume 273, pages 2249–2256, 2006.

26. J. Rychtář and B. Stadler. Evolutionary dynamics on small-world networks. *International Journal of Computational and Mathematical Sciences*, 2(1):1–4, 2008.

27. W. H. Sandholm. *Population games and evolutionary dynamics*. MIT Press, 2011.

28. C. Taylor, Y. Iwasa, and M. A. Nowak. A symmetry of fixation times in evoultionary dynamics. *Journal of Theoretical Biology*, 243(2):245–251, 2006.

29. A. Traulsen and C. Hauert. Stochastic evolutionary game dynamics. In *Reviews of Nonlinear Dynamics and Complexity*, volume 2. NY: Wiley, 2008.