Absorption Time of the Moran Process[*]

Josep Díaz[†]   Leslie Ann Goldberg[‡]   David Richerby[‡]   Maria Serna[†]

June 29, 2015

Abstract

The Moran process models the spread of mutations in populations on graphs. We investigate the absorption time of the process, which is the time taken for a mutation introduced at a randomly chosen vertex to either spread to the whole population, or to become extinct. It is known that the expected absorption time for an advantageous mutation is $O(n^4)$ on an n-vertex undirected graph, which allows the behaviour of the process on undirected graphs to be analysed using the Markov chain Monte Carlo method. We show that this does not extend to directed graphs by exhibiting an infinite family of directed graphs for which the expected absorption time is exponential in the number of vertices. However, for regular directed graphs, we show that the expected absorption time is $\Omega(n \log n)$ and $O(n^2)$. We exhibit families of graphs matching these bounds and give improved bounds for other families of graphs, based on isoperimetric number. Our results are obtained via stochastic dominations which we demonstrate by establishing a coupling in a related continuous-time model. The coupling also implies several natural domination results regarding the fixation probability of the original (discrete-time) process, resolving a conjecture of Shakarian, Roos and Johnson.

1   Introduction

The Moran process [?], as adapted by Lieberman, Hauert and Nowak [?], is a stochastic model for the spread of genetic mutations through populations of organisms. Similar processes have been used to model the spread of epidemic diseases, the behaviour of voters, the spread of ideas in social networks, strategic interaction in evolutionary game theory, the emergence of monopolies, and cascading failures in power grids and transport networks [?, ?, ?, ?, ?].

In the Moran process, individuals are modelled as the vertices of a graph and, at each step of the discrete-time process, an individual is selected at random to reproduce. This vertex chooses one of its neighbours uniformly at random and replaces that neighbour with its offspring, a copy of itself. The probability that any given individual is chosen to reproduce is proportional to its *fitness*: individuals with the mutation have fitness $r > 0$ and non-mutants have fitness 1. The initial state has a single mutant placed uniformly at random in the graph, with every other vertex a non-mutant. On any finite, strongly connected graph, the process

---

[†]Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Spain.

[‡]Department of Computer Science, University of Oxford, UK.

will terminate with probability 1, either in the state where every vertex is a mutant (known as *fixation*) or in the state where no vertex is a mutant (known as *extinction*).

The principal quantities of interest are the *fixation probability* (the probability of reaching fixation) and the expected *absorption time* (the expected number of steps before fixation or extinction is reached). In general, these depend on both the graph topology and the mutant fitness. In principle, they can be computed by standard Markov chain techniques but doing so for an n-vertex graph involves solving a set of 2^n linear equations, which is computationally infeasible. Fixation probabilities have also been calculated by producing and approximately solving a set of differential equations that model the process [**?**]. These methods seem to work well in practice but there is no known bound on the error introduced by the conversion to differential equations and the approximations in their solution.

When the underlying graph is undirected and the mutant fitness r is at least 1, it is known how to approximate the fixation probability: The paper [**?**] gave a fully polynomial randomised approximation scheme (FPRAS) for computing the fixation probability of the Moran process on undirected graphs. The approximation scheme uses the Markov chain Monte Carlo method. The fact that it provides a suitable approximation in polynomial time follows from the fact that the expected absorption time on an n-vertex graph is at most $\frac{r}{r-1}n^4$ for $r > 1$.

1.1 Our contributions

The main contribution of this paper is to determine the extent to which the polynomial bound on expected absorption time carries through to directed graphs. Throughout the paper, we assume that the mutant fitness r exceeds 1.

1.1.1 Regular digraphs

We start by considering the absorption time on a strongly connected Δ-regular digraph (where every vertex has in-degree Δ and out-degree Δ). Regularity makes some calculations straightforward because the detailed topology of the graph is not relevant. We describe these first, and then discuss the more difficult questions (where topology does play a role) and state our results.

The following facts hold for Δ-regular graphs, independent of the topology.

- It is well known [**?**] that, on any regular graph on n vertices, a single randomly placed mutant with fitness r reaches fixation with probability

$$\frac{1 - r^{-1}}{1 - r^{-n}} \tag{1}$$

  To see this, note that if there are k mutants, the total fitness of the population is $W_k = n + k(r - 1)$. The probability that the next reproduction happens along the directed edge (u, v) is $\frac{r}{W_k}\frac{1}{\Delta}$ if u is a mutant and $\frac{1}{W_k}\frac{1}{\Delta}$, if it is not. Since the graph is Δ-regular, there are exactly as many directed edges from mutants to non-mutants as there are from non-mutants to mutants. It follows that the probability that the number of mutants increases at the next step is exactly r times the probability that it decreases, regardless of which vertices are currently mutants. Thus, the number of mutants in the population, observed every time it changes, forms a random walk on $\{0, \ldots, n\}$ with initial state 1, upward drift r and absorbing barriers at 0 and n. It is a standard result

(e.g., [**?**, Example 3.9.6]) that such a random walk reaches the barrier at n with the probability given in (1).

- It is also well known (e.g., [**?**, Example 3.9.6]) that the expected number of steps of this walk before absorption (which may be at either 0 or n) is a function of r and n that tends to $n(1 + \frac{1}{r})$ in the limit as $n \to \infty$, independent of the graph structure beyond regularity. However, the number of steps taken by the random walk (often referred to as the "active steps" of the Moran process) is not the same as the original process's absorption time, because the absorption time includes many steps at which the number of mutants does not change, either because a mutant reproduces to a mutant or because a non-mutant reproduces to a non-mutant.

In Section 4, we show that the expected absorption time of the Moran process is polynomial for regular digraphs. In contrast to the number of active steps, the absorption time does depend on the detailed structure of the graph. We prove the following upper and lower bounds, where H_n denotes the n'th harmonic number, which is $\Theta(\log n)$.

Theorem 1. *The expected absorption time of the Moran process on a strongly connected Δ-regular n-vertex digraph G is at least $\left(\frac{r-1}{r^2}\right) nH_{n-1}$ and at most $n^2\Delta$.*

In Section 4.7, we prove the following theorem, which shows that the upper bound in Theorem 1 is tight up to a constant factor (which depends on Δ and r but not on n).

Theorem 2. *Suppose that $r > 1$ and $\Delta > 2$. There is an infinite family $\mathcal{G}$ of Δ-regular graphs such that, when the Moran process is run on an n-vertex graph $G \in \mathcal{G}$, the expected absorption time exceeds $\frac{1}{8r}(1 - \frac{1}{r})\frac{n^2}{(\Delta-1)^2}$.*

The digraphs in the family $\mathcal{G}$ are symmetric, so can be viewed as undirected graphs. The upper bound on the expected absorption time in Theorem 1 can be improved for certain classes of regular undirected graphs using the notion of the isoperimetric number $i(G)$ of a graph G, which is defined in Section 4.6. There, we prove the following theorem.

Theorem 3. *The expected absorption time of the Moran process on a connected Δ-regular n-vertex undirected graph G is at most $2\Delta nH_n/i(G)$.*

Theorem 3 pinpoints the expected absorption time for $G = K_n$, up to a constant factor, since $i(K_n) = \lceil n/2 \rceil$ [**?**] and Theorem 1 gives an $\Omega(n \log n)$ lower bound. Theorem 3 is worse than the upper bound of Theorem 1 by a factor of $O(\log n)$ for the cycle C_n since $i(C_n) = 2/\lfloor n/2 \rfloor$ [**?**]. However, we often get an improvement by using the isoperimetric number. For example, the $\sqrt{n}$-by-$\sqrt{n}$ grid has $i(G) = \Theta(1/\sqrt{n})$ (see [**?**]), giving an $O(n^{3/2} \log n)$ absorption time; the hypercube has $i(G) = 1$ (see, e.g., [**?**]), giving an $O(n \log^2 n)$ absorption time. Bollobás [**?**] showed that, for every $\Delta \geq 3$ there is a positive number $\eta < 1$ such that, for almost all Δ-regular n-vertex undirected graphs G (as n tends to infinity), $i(G) \geq (1-\eta)\Delta/2$, which gives an $O(n \log n)$ absorption time for almost all graphs since these graphs are almost all connected. Of course this gives a similar bound on the expected absorption time for a random Δ-regular n-vertex graph, under the uniform distribution. Note that the upper bounds in Theorems 1 and 3 are independent of r. However, since our bounds are not tight, it is still possible that the precise expected absorption time depends on r.

1.1.2 Slow absorption

Theorem 1 shows that regular digraphs, like undirected graphs, reach absorption in expected polynomial time. In Section 5 we show that the same does not hold for general digraphs. In particular, we construct an infinite family $\{G_{r,N}\}$ of strongly connected digraphs indexed by a positive integer N. We then prove the following theorem.

Theorem 4. *Fix $r > 1$ and let $\varepsilon_r = \min(r-1, 1)$. For any positive integer N that is sufficiently large with respect to r, the expected absorption time of the Moran process on $G_{r,N}$ is at least*

$$\frac{1}{16} \left\lfloor \left(\frac{\varepsilon_r}{32}\right) (2^N - 1) \right\rfloor \frac{\varepsilon_r}{4\lceil r \rceil + 3}.$$

Theorem 4 shows that there is an infinite family of strongly connected digraphs such that the absorption time on n-vertex graphs in this family is exponentially large, as a function of n. It follows that the techniques from [?] do not give a polynomial-time algorithm for approximating the fixation probability on digraphs.

The underlying structure of the graph $G_{r,N}$ is a large undirected clique on N vertices and a long directed path. Each vertex of the clique sends an edge to the first vertex of the path, and each vertex of the clique receives an edge from the path's last vertex. We refer to the first N vertices of the path as P and the remainder as Q. Each vertex of P has out-degree 1 but receives $4\lceil r \rceil$ edges from Q. (See Figure 1 in Section 5.1.)

Suppose that N is sufficiently large with respect to r and consider the Moran process on $G_{r,N}$. Given the relative sizes of the clique and the path, there is a reasonable probability (about $\frac{1}{4r+2}$) that the initial mutant is in the clique. The edges to and from the path have a negligible effect so it is reasonably likely (probability at least $1 - \frac{1}{r}$) that we will then reach the state where half the clique vertices are mutants. To reach absorption from this state, one of two things must happen.

For the process to reach extinction, the mutants already in the clique must die out. Because the interaction between the clique and path is small, the number of mutants in the clique is very close to a random walk on $\{0, \ldots, N\}$ with upward drift r, and the expected time before such a walk reaches zero from $N/2$ is exponential in N.

On the other hand, suppose the process reaches fixation. In particular, the vertices of P, the first part of the path, must become mutants. Note that no vertex of Q can become a mutant before the last vertex of P has done so. While all the vertices in Q are non-mutants, the edges from that part of the path to P ensure that each mutant in P is more likely to be replaced by a non-mutant from Q than it is to create a new mutant in P. As a result, the number of mutants in P is stochastically bounded above by a random walk on $\{0, \ldots, N\}$ with a strictly greater probability of decreasing than increasing. Again, this walk is expected to take exponentially many steps before reaching N.

1.1.3 Stochastic domination

Our main technical tool is stochastic domination. Intuitively, one expects that the Moran process has a higher probability of reaching fixation when the set of mutants is S than when it is some subset of S, and that it is likely to do so in fewer steps. It also seems obvious that modifying the process by continuing to allow all transitions that create new mutants but forbidding some transitions that remove mutants should make fixation faster and more

probable. Such intuitions have been used in proofs in the literature; it turns out that they are essentially correct, but for rather subtle reasons.

The Moran process can be described as a Markov chain $(Y_t)_{t\geq 1}$ where Y_t is the set $S \subseteq V(G)$ of mutants at the t'th step. The normal method to make the above intuitions formal would be to demonstrate a stochastic domination by coupling the Moran process $(Y_t)_{t\geq 1}$ with another copy $(Y'_t)_{t\geq 1}$ of the process where $Y_1 \subseteq Y'_1$. The coupling would be designed so that $Y_1 \subseteq Y'_1$ would ensure that $Y_t \subseteq Y'_t$ for all $t > 1$. However, a simple example shows that such a coupling does not always exist for the Moran process. Let G be the undirected path with two edges: $V(G) = \{1, 2, 3\}$ and $E(G) = \{(1, 2), (2, 1), (2, 3), (3, 2)\}$. Let $(Y_t)_{t\geq 1}$ and $(Y'_t)_{t\geq 1}$ be Moran processes on G with $Y_1 = \{2\}$ and $Y'_1 = \{2, 3\}$. With probability $\frac{r}{2(r+2)}$, we have $Y_2 = \{1, 2\}$. The only possible value for Y'_2 that contains Y_2 is $\{1, 2, 3\}$ but this occurs with probability only $\frac{r}{2(2r+1)}$. Therefore, any coupling between the two processes fails because

$$\Pr(Y_2 \not\subseteq Y'_2) \geq \frac{r(r-1)}{2(r+2)(2r+1)},$$

which is strictly positive for any $r > 1$. The problem is that, when vertex 3 becomes a mutant, it becomes more likely to be the next vertex to reproduce and, correspondingly, every other vertex becomes less likely. This can be seen as the new mutant "slowing down" all the other vertices in the graph.

To get around this problem, we consider a continuous-time version of the process, $\widetilde{Y}[t]$ $(t \geq 0)$. Given the set of mutants $\widetilde{Y}[t]$ at time t, each vertex waits an amount of time before reproducing. For each vertex, this period of time is chosen according to the exponential distribution with parameter equal to the vertex's fitness, independently of the other vertices. (Thus, the parameter is r if the vertex is a mutant and 1, otherwise.) If the first vertex to reproduce is v at time $t + \tau$ then, as in the standard, discrete-time version of the process, one of its out-neighbours w is chosen uniformly at random, the individual at w is replaced by a copy of the one at v and the time at which w will next reproduce is exponentially distributed with parameter given by its new fitness. The discrete-time process is recovered by taking the sequence of configurations each time a vertex reproduces.

In continuous time, each member of the population reproduces at a rate given by its fitness, independently of the rest of the population whereas, in discrete time, the population has to co-ordinate to decide who will reproduce next. It is still true in continuous time that vertex w becoming a mutant makes it less likely that each vertex $v \neq w$ will be the next to reproduce. However, the vertices are not slowed down as they are in discrete time: they continue to reproduce at rates determined by their fitnesses. This distinction allows us to establish the following coupling lemma, which formalises the intuitions discussed above.

Lemma 5 (Coupling lemma). *Let $G = (V, E)$ be any digraph, let $Y \subseteq Y' \subseteq V(G)$ and $1 \leq r \leq r'$. Let $\widetilde{Y}[t]$ and $\widetilde{Y}'[t]$ $(t \geq 0)$ be continuous-time Moran processes on G with mutant fitness r and r', respectively, and with $\widetilde{Y}[0] = Y$ and $\widetilde{Y}'[0] = Y'$. There is a coupling between the two processes such that $\widetilde{Y}[t] \subseteq \widetilde{Y}'[t]$ for all $t \geq 0$.*

In the paper, we use the coupling lemma to establish stochastic dominations between *discrete* Moran processes. It also has consequences concerning fixation probabilities. Recall that "fixation" is the state of a (discrete) Moran process in which every vertex is a mutant. The fixation probability $f_{G,r}$ is the probability that this state is reached when the Moran process is run on a digraph $G = (V, E)$, starting from a state in which a single initial mutant is placed

uniformly at random. For any set $S \subseteq V$, let $f_{G,r}(S)$ be the probability of reaching fixation when the set of vertices initially occupied by mutants is S. Thus, $f_{G,r} = \frac{1}{|V|} \sum_{v \in V} f_{G,r}(\{v\})$. Using the coupling lemma, we can prove the following theorem.

Theorem 6. *For any digraph G, if $0 < r \leq r'$ and $S \subseteq S' \subseteq V(G)$, then $f_{G,r}(S) \leq f_{G,r'}(S')$.*

This theorem has two immediate corollaries. The first was conjectured by Shakarian, Roos and Johnson [?, Conjecture 2.1]. It was known from [?] that $f_{G,r} \geq f_{G,1}$ for any $r > 1$.

Corollary 7 (Monotonicity)**.** *If $0 < r \leq r'$ then, for any digraph G, $f_{G,r} \leq f_{G,r'}$.*

Corollary 7 follows immediately from Theorem 6 since $f_{G,r}(\{v\}) \leq f_{G,r'}(\{v\})$ for all $v \in V(G)$.

The second corollary can be stated informally as, "Adding more mutants can't decrease the fixation probability," and has been assumed in the literature, without proof. However, although it appears obvious at first, it is somewhat subtle: the example at the beginning of this section shows that adding more mutants can actually decrease the probability of a particular vertex becoming a mutant at the next step of the process.

Corollary 8 (Subset domination)**.** *For any digraph G and any $r > 0$, if $S \subseteq S' \subseteq V(G)$, then $f_{G,r}(S) \leq f_{G,r}(S')$.*

Note that, although we have introduced the continuous-time version of the process for technical reasons, to draw conclusions about the original discrete-time Moran process, the continuous-time version may actually be a more realistic model than the discrete-time version.

1.2   Previous work

The Moran process is similar to a discrete version of directed percolation known as the contact process. There is a vast literature (e.g., [?, ?, ?]) on the contact process and other related infection processes. The particular details of the Moran process will be important for us, so we concentrate here on previous work about calculating the fixation probability of the Moran process itself. Fixation probabilities are known for regular graphs [?] and for stars (complete bipartite graphs $K_{1,k}$) [?]. Lieberman et al. [?] have defined classes of directed graphs with a parameter k, for which they claim that the fixation probability tends to $1 - r^{-k}$ for large graphs. While these graphs do seem to have very large fixation probability, we have shown this specific claim to be incorrect for $k = 5$ [?]. Very recently, it has been claimed [?] that for large k, the fixation probability is close to $1 - \frac{1}{(k-2)r^4}$. Other work has investigated the possibility of so-called "suppressors", graphs having fixation probability less than that given by (1) for at least some range of values for r [?, ?].

There is a more complicated version of the Moran process in which the fitness of a vertex is determined by its expected payoff when playing some two-player game against a randomly chosen neighbour [?, ?]. In this version of the process, mutants play the game with one strategy and non-mutants play the game with another. The ordinary Moran process corresponds to the special case of this game in which the mutant and non-mutant strategies give payoffs r and 1, respectively, regardless of the strategy used by the opponent.

Most previous work on absorption times has been in the game-based version of the process, where the added complexity of the model limits analysis to very simple graphs, such as complete graphs, stars and cycles [?, ?].

2 Preliminaries

When k is a positive integer, $[k]$ denotes $\{1,\ldots,k\}$. We consider the evolution of the Moran process [?] on a strongly connected directed graph (digraph). Consider such a digraph $G = (V,E)$ with $n = |V|$. When the process is run on G, each state is a set of vertices $S \subseteq V$. The vertices in S are said to be "mutants". If $|S| = k$ then the total fitness of the state is given by $W_k = n + (r-1)k$ — each of the k mutants contributes fitness r to the total fitness and each other vertex contributes fitness 1. Except where stated otherwise, we assume that $r > 1$. The starting state is chosen uniformly at random from the one-element subsets of $V(G)$. From a state S with $|S| = k$, the process evolves as follows. First, a vertex u is chosen to reproduce. The probability that vertex u is chosen is r/W_k if u is a mutant and $1/W_k$ otherwise. Given that u will reproduce, a directed edge (u,v) is chosen uniformly at random from $\{(u,v') \mid (u,v') \in E\}$. The state of vertex u in S is copied to vertex v to give the new state

$$S|_{u \to v} = \begin{cases} S \cup \{v\} & \text{if } u \text{ is a mutant,} \\ S \setminus \{v\} & \text{if } u \text{ is a non-mutant.} \end{cases}$$

Let $d^+(u)$ denote the out-degree of vertex u and let $d^-(u)$ denote its in-degree. A digraph $G = (V,E)$ is Δ-*regular* if, for every vertex $u \in V$, $d^+(u) = d^-(u) = \Delta$. G is *regular* if it is Δ-regular for some natural number Δ. If the Moran process is run on a strongly connected digraph G, there are exactly two absorbing states — $\emptyset$ and $V(G)$. Once one of these states is reached, the process will stay in it forever. The *absorption time* is the number of reproduction steps until such a state is reached.

A digraph $G = (V,E)$ is weakly connected if the underlying undirected graph is connected. Given a subset $S \subseteq V$, let m_S^+ be the number of edges from vertices in S to vertices in $V \setminus S$. Let m_S^- be the number of edges from vertices in $V \setminus S$ to S. If G is regular then, for every $S \subseteq V$,

$$\begin{aligned} m_S^+ &= |\{(u,v) \in E \mid u \in S\}| - |\{(u,v) \in E \mid u,v \in S\}| \\ &= |S|\Delta - |\{(u,v) \in E \mid u,v \in S\}| \\ &= |\{(u,v) \in E \mid v \in S\}| - |\{(u,v) \in E \mid u,v \in S\}| \\ &= m_S^- \,. \end{aligned}$$

Thus, every regular digraph that is weakly connected is strongly connected.

We sometimes consider the Moran process on an undirected graph $G = (V,E)$. We view the undirected graph as a digraph in which the set E of edges is symmetric (so $(u,v) \in E$ if and only if $(v,u) \in E$). If G is undirected then, for every vertex u, $d^+(u) = d^-(u)$ and in this case we just write $d(u)$ to denote this quantity.

3 Domination

A useful proof technique is to show that the behaviour of the Moran process is stochastically dominated by that of a related process that is easier to analyse. Similarly, it is useful to compare the behaviour of the Moran process, evolving on a digraph G, with that of another Moran process on the same digraph, where the second process starts with more mutants. Recall that the Moran process can be described as a Markov chain $(Y_t)_{t \geq 1}$ where Y_t is the set $S \subseteq V(G)$ of mutants at the t'th step. It would be natural to attempt to establish a coupling

between Moran processes $(Y_t)_{t\geq 1}$ and $(Y'_t)_{t\geq 1}$ such that, if $Y_1 \subseteq Y'_1$, then $Y_t \subseteq Y'_t$ for all $t \geq 1$ but, as we showed in Section 1.1.3, this cannot be done.

To obtain useful dominations, we will consider a continuous-time version of the Moran process. The domination that we construct for the continuous-time process will allow us to draw conclusions about the original (discrete-time) Moran process. In a digraph $G = (V, E)$ where the set of mutants Y have fitness r, let $r_{v,Y} = r$ if $v \in Y$ and $r_{v,Y} = 1$, otherwise. We define the continuous-time version of the Moran process on a digraph $G = (V, E)$ as follows. Starting in configuration $\widetilde{Y}[t]$ at time t, each vertex v waits for a period of time before reproducing. This period of time is chosen, independently of other vertices, according to an exponential distribution with parameter $r_{v,\widetilde{Y}[t]}$. Therefore, the probability that two vertices reproduce at once is zero. Suppose that the first vertex to reproduce after time t is vertex v, at time $t + \tau$. As in the discrete-time version of the process, an out-neighbour w of v is chosen u.a.r. and the new configuration is given by $\widetilde{Y}[t + \tau] = \widetilde{Y}[t]|_{v \to w}$.

From the definition of the exponential distribution, it is clear that the probability that a particular vertex v is the next to reproduce, from configuration $\widetilde{Y}[t]$, is $r_{v,\widetilde{Y}[t]}/W_{|\widetilde{Y}[t]|}$. Thus, the Moran process (as generalised by Lieberman et al.) is recovered by taking the sequence of configurations each time a vertex reproduces.[1] We can now give the proof of Lemma 5 and Theorem 6.

Lemma 5. Let $G = (V, E)$ be any digraph, let $Y \subseteq Y' \subseteq V(G)$ and $1 \leq r \leq r'$. Let $\widetilde{Y}[t]$ and $\widetilde{Y}'[t]$ $(t \geq 0)$ be continuous-time Moran processes on G with mutant fitness r and r', respectively, and with $\widetilde{Y}[0] = Y$ and $\widetilde{Y}'[0] = Y'$. There is a coupling between the two processes such that $\widetilde{Y}[t] \subseteq \widetilde{Y}'[t]$ for all $t \geq 0$.

Proof. Suppose that $\widetilde{Y}[t] \subseteq \widetilde{Y}'[t]$ for some t. We couple the evolution of $\widetilde{Y}[t']$ and $\widetilde{Y}'[t']$ for $t' \geq t$ as follows. For ease of notation, we write $r_{v,t}$ and $r'_{v,t}$ for $r_{v,\widetilde{Y}[t]}$ and $r'_{v,\widetilde{Y}'[t]}$, respectively. Let

$$S = \{v \in V(G) \mid r_{v,t} < r'_{v,t}\} \subseteq \widetilde{Y}'[t]$$

and note that, for $v \in V \setminus S$, $r'_{v,t} = r_{v,t}$. For $v \in V$, let t_v be a random variable drawn from $\text{Exp}(r_{v,t})$ and, for $v \in S$, let $t'_v \sim \text{Exp}(r'_{v,t} - r_{v,t})$. From the definition of the exponential distribution, it is easy to see that, for each $v \in S$, $\min(t_v, t'_v) \sim \text{Exp}(r'_{v,t})$.

If some t_v is minimal among $\{t_v \mid v \in V\} \cup \{t'_v \mid v \in S\}$, then choose an out-neighbour w of v u.a.r. and set $\widetilde{Y}[t + t_v] = \widetilde{Y}[t]|_{v \to w}$ and $\widetilde{Y}'[t + t_v] = \widetilde{Y}'[t]|_{v \to w}$. It is clear that $\widetilde{Y}[t + t_v] \subseteq \widetilde{Y}'[t + t_v]$.

Otherwise, some t'_v is minimal. In this case, set $\widetilde{Y}[t + t'_v] = \widetilde{Y}[t]$; choose an out-neighbour w of v u.a.r. and set $\widetilde{Y}'[t + t'_v] = \widetilde{Y}'[t]|_{v \to w}$. Since $v \in S \subseteq \widetilde{Y}'[t]$, we have

$$\widetilde{Y}[t + t'_v] = \widetilde{Y}[t] \subseteq \widetilde{Y}'[t] \subseteq \widetilde{Y}'[t + t'_v].$$

In both cases, the continuous-time Moran process has been faithfully simulated up to time $t + \tau$, where $\tau = t_v$ in the first case and $\tau = t'_v$ in the second case, and the memorylessness of the exponential distribution allows the coupling to continue from $\widetilde{Y}[t + \tau]$ and $\widetilde{Y}'[t + \tau]$. □

The coupling provided in Lemma 5 could be translated to a coupling for the discrete-time Moran process, though the time steps in the two copies would not be the same since such a coupling was ruled out in Section 1.1.3. In fact, is easier for us to use Lemma 5 directly.

---

[1]This is closely related to the jump chain, which is defined to be the discrete-time chain whose successive states are the states $\widetilde{Y}[t]$ for the successive times t immediately after the state changes. Thus, the jump chain is the chain of "active" steps of the discrete-time Moran process (see Section 4.3).

Theorem 6. For any digraph G, if $0 < r \leq r'$ and $S \subseteq S' \subseteq V(G)$, then $f_{G,r}(S) \leq f_{G,r'}(S')$.

Proof. We split the proof into two parts: $1 \leq r \leq r'$ and $r \leq r' \leq 1$. The remaining case $r \leq 1 \leq r'$ follows because $f_{G,r}(S) \leq f_{G,1}(S) \leq f_{G,r'}(S')$.

First, suppose that $1 \leq r \leq r'$. Let $\widetilde{Y}[t]$ and $\widetilde{Y}'[t]$ be Moran processes on $G = (V, E)$ with mutant fitnesses r and r', respectively, with $\widetilde{Y}[0] = S$ and $\widetilde{Y}'[0] = S'$. By the coupling lemma, we can couple the processes such that $\widetilde{Y}[t] \subseteq \widetilde{Y}'[t]$ for all $t \geq 0$. In particular, if there is a t such that $\widetilde{Y}[t] = V$, we must have $\widetilde{Y}'[t] = V$ also. Therefore, $f_{G,r'}(S') \geq f_{G,r}(S)$.

Now, suppose that $r \leq r' \leq 1$. Observe that the behaviour of the Moran process is independent of any consistent scaling of the mutant and non-mutant fitnesses, in the following sense. For any $\alpha > 0$, the process where mutants have fitness r and non-mutants have fitness 1 is identical to the one where they have fitness αr and α, respectively. Let $\widetilde{Y}[t]$ be the process where mutants have fitness $\frac{1}{r} \cdot r = 1$ and non-mutants have fitness $\frac{1}{r} \geq 1$, and let $\widetilde{Y}'[t]$ have mutant fitness 1, non-mutant fitness $\frac{1}{r'} \geq 1$. Let $\widetilde{Y}[0] = S$ and $\widetilde{Y}'[0] = S'$. Now, $f_{G,r}(S)$ is the probability that the individuals with fitness 1 take over the graph in $\widetilde{Y}[t]$, which is $1 - f_{G,1/r}(V \setminus S)$; similarly, $f_{G,r'}(S') = 1 - f_{G,1/r'}(V \setminus S')$. By the first part, $f_{G,1/r'}(V \setminus S') \leq f_{G,r}(V \setminus S)$ and the result follows. $\qquad \square$

Corollaries 7 (monotonicity) and 8 (subset domination) follow immediately from Theorem 6, as shown in Section 1.1.3.

4   Regular digraphs

This section provides upper and lower bounds on the absorption time of the Moran process on regular digraphs.

4.1   An upper bound for undirected graphs

Theorem 4, which was stated in the introduction and will be proved in Section 5.3, makes it clear that the absorption time bounds from [?] do not apply to digraphs. For example, Theorem 7 of [?] gives a polynomial upper bound on the expected absorption time for all connected undirected graphs, but Theorem 4 shows that process takes exponential time on some strongly connected digraphs.

Since we will be discussing both undirected graphs and digraphs in this section, we start by observing that Theorem 7 of [?] can be improved to give an $O(n^3)$ bound in the special case in which the undirected graphs to which it applies are regular. This is certainly not tight (as we shall see below) but it is a natural place to begin.

Proposition 9. The expected absorption time of the Moran process on a connected Δ-regular n-vertex undirected graph is at most $(r/(r-1))n^2\Delta$.

Proof. Given an undirected graph G and a set $S \subseteq V(G)$, let $\phi(S) = \sum_{v \in S} \frac{1}{d(v)}$. Let ∂S be the set of (undirected) edges between vertices in S and vertices in $V(G) \setminus S$. If G is Δ-regular and has n vertices, then $\phi(V(G)) = n/\Delta$. The proof of Theorem 7 and Equation (1) of [?] show that the absorption time is at most

$$\phi(V(G)) \max \left\{ \left(\frac{(n + (r-1)|S|)\,\Delta^2}{(r-1)\,|\partial S|} \right) \mid \emptyset \subset S \subset V(G) \right\}.$$

The bound follows using $|S| \leq n$ and $|\partial S| \geq 1$. $\qquad \square$

4.2 Definitions

We use the following standard Markov chain definitions. For more detail, see, for example, [?]. We use $(X_t)_{t\geq 0}$ to denote a discrete-time Markov chain $\mathcal{M}$ with finite state space Ω and transition matrix P. $T_k = \inf\{t \geq 1 \mid X_t = k\}$ is the first passage time for visiting state k (not counting the initial state X_0). The time spent in state i between visits to state k is given by

$$\gamma_i^k = \sum_{t=0}^{T_k-1} 1_{X_t=i}, \text{ where } X_0 = k.$$

The chain is said to be irreducible if, for every pair of states (i, j) there is some $t \geq 0$ such that $\Pr(X_t = j \mid X_0 = i) > 0$. Since Ω is finite, this implies that the chain is recurrent, which means that, for every state $i \in \Omega$, $\Pr(X_t = i$ for infinitely many $t) = 1$. We use the following proposition, which (up to minor notational differences) is the special case of [?, Theorem 1.7.6] corresponding to finite state spaces.

Proposition 10. Let $\mathcal{M}$ be an irreducible discrete-time Markov chain with finite state space $\Omega = \{0, \ldots, \omega - 1\}$ and transition matrix P. For $k \in \Omega$, let $\lambda = (\lambda_0, \ldots, \lambda_{\omega-1})$ be a vector of non-negative real numbers with $\lambda_k = 1$ satisfying $\lambda P = \lambda$. Then, for every $j \in \Omega$, $E[\gamma_j^k] = \lambda_j$.

4.3 Active steps of the Moran process

We fix $r > 1$ and study the Moran process on a strongly connected Δ-regular n-vertex digraph $G = (V, E)$ with $n > 1$. We refer to the steps of the process during which the number of mutants changes as "active steps". As explained in the introduction, the evolution of the number of mutants, sampled after each active step, corresponds to a one-dimensional random walk on $\{0, \ldots, n\}$ which starts at state 1, absorbs at states 0 and n, and has upwards drift $p = r/(r + 1)$. (To see this, note that the probability that the number of mutants increases from a size-k state S is $\sum_{e \in E \cap (S \times V(G)\setminus S)} \frac{r}{W_k\Delta}$ and the probability that it decreases is $\sum_{e \in E \cap (V(G)\setminus S \times S)} \frac{1}{W_k\Delta}$ but we showed earlier that the number of edges in each summation is equal when G is Δ-regular, so the ratio between these two probabilities is r to 1.)

To derive the properties that we need, we consider a Markov chain $\mathcal{M}$ with state space $\Omega = \{0, \ldots, n + 1\}$. The non-zero entries of the transition matrix P of $\mathcal{M}$ are as follows. $P_{0,n+1} = P_{n,n+1} = 1$. Also, $P_{n+1,1} = 1$. Finally, for $1 \leq i \leq n - 1$, $P_{i,i+1} = p$ and $P_{i,i-1} = 1 - p$. Starting from state 1, the chain simulates the one-dimensional walk discussed above. State $n + 1$ is a special state of the Markov chain that is visited after an absorbing state of the random walk is reached. From state $n + 1$, the chain goes back to state 1 and repeats the random walk. We use the following property of $\mathcal{M}$.

Lemma 11. Let $f = (r^n - r^{n-1})/(r^n - 1)$. Define the vector $\lambda = (\lambda_0, \ldots, \lambda_{n+1})$ as follows.

$$\lambda_0 = 1 - f,$$
$$\lambda_j = (1 + r)(1 - f)(r^n - r^j)/(r^n - r), \text{ for } 1 \leq j \leq n - 1,$$
$$\lambda_n = f,$$
$$\lambda_{n+1} = 1.$$

Then, for every $j \in \Omega$, $E[\gamma_j^{n+1}] = \lambda_j$.

Proof. Note that P is irreducible. By Proposition 10, it suffices to show that $\lambda P = \lambda$. First, consider the column vector $P_{*,0}$. This is all zero except the entry $P_{1,0} = 1 - p$ so $\lambda P_{*,0} = (1-p)\lambda_1 = \lambda_0$, as required. Then note that $1/(1-f) = r(r^n - 1)/(r^n - r)$. So

$$\lambda P_{*,1} = (1-p)\lambda_2 + \lambda_{n+1}$$
$$= (1-f)\left(\frac{r^n - r^2}{r^n - r} + \frac{1}{1-f}\right)$$
$$= (1-f)\left(\frac{r^n - r^2 + r(r^n - 1)}{r^n - r}\right)$$
$$= (1-f)(r+1) = \lambda_1,$$

as required. Next, consider the column vector $P_{*,j}$ for $1 < j < n - 1$. In this case,

$$\lambda P_{*,j} = p\lambda_{j-1} + (1-p)\lambda_{j+1}$$
$$= (1-f)\left(\frac{r(r^n - r^{j-1}) + (r^n - r^{j+1})}{r^n - r}\right)$$
$$= (1-f)\left(\frac{(1+r)(r^n - r^j)}{r^n - r}\right) = \lambda_j,$$

as required. Then

$$\lambda P_{*,n-1} = p\lambda_{n-2}$$
$$= \left(\frac{(1+r)(1-f)}{r^n - r}\right)\left(\frac{r(r^n - r^{n-2})}{r+1}\right)$$
$$= \left(\frac{(1+r)(1-f)}{r^n - r}\right)(r^n - r^{n-1}) = \lambda_{n-1},$$

as required. Furthermore,

$$\lambda P_{*,n} = p\lambda_{n-1} = (1-f)\left(\frac{r^{n+1} - r^n}{r^n - r}\right).$$

Also,
$$\frac{1-f}{f} = \frac{r^n - r}{r^{n+1} - r^n},$$

so
$$\frac{\lambda P_{*,n}}{\lambda_n} = \frac{p\lambda_{n-1}}{f} = \frac{(1-f)}{f}\left(\frac{r^{n+1} - r^n}{r^n - r}\right) = \left(\frac{r^n - r}{r^{n+1} - r^n}\right)\left(\frac{r^{n+1} - r^n}{r^n - r}\right) = 1,$$

as required. Finally, $\lambda P_{*,n+1} = \lambda_0 + \lambda_n = 1 = \lambda_{n+1}$, as required. This completes the proof. $\square$

It is well known [?] that f is the fixation probability of the Moran process on a regular graph. This is an easy consequence of Lemma 11, but we don't need it here. We instead use the following corollary.

Corollary 12. *For all $j \in \{1, \ldots, n-1\}$, $1 - \frac{1}{r^2} \leq E[\gamma_j^{n+1}] \leq 1 + \frac{1}{r}$.*

Proof. From Lemma 11,

$$E[\gamma_j^{n+1}] = \left(\frac{r+1}{r}\right)\left(\frac{r^n - r^j}{r^n - 1}\right).$$

The upper bound follows from the fact that $r^n - r^j \leq r^n - 1$ (a consequence of $r > 1$ and $j \geq 0$). For the lower bound, note that $E[\gamma_j^{n+1}]$ is minimised at $j = n - 1$ and

$$E[\gamma_{n-1}^{n+1}] = \left(\frac{r^n}{r^n - 1}\right)\left(1 - \frac{1}{r^2}\right).$$

The lower bound then follows from $r^n/(r^n - 1) \geq 1$. □

4.4 Absorption time

Now let the Moran process be $(Y_t)_{t\geq 1}$ where each state Y_t is the set $S \subseteq V(G)$ of vertices of G that are mutants at the t'th step. The state Y_1 is selected uniformly at random from the size-1 subsets of $V(G)$. For each state S, let $p(S) = \Pr(Y_{t+1} \neq S \mid Y_t = S)$ and let $\mu(S) = \inf\{t \geq 1 \mid Y_{t+1} \neq S, Y_1 = S\}$. $\mu(S)$ is a random variable representing the number of times that the state S appears when the process is run, starting from S, before another state is reached. It is geometrically distributed with parameter $p(S)$, so $E[\mu(S)] = 1/p(S)$. The absorption time T_A is the number of steps needed to get to state 0 or state n, which is $T_A = \inf\{t \geq 1 \mid |Y_t| \in \{0, n\}\} - 1$.

Let $\tau_1 = 1$. For $j > 1$, let $\tau_j = \inf\{t > \tau_{j-1} \mid Y_t \neq Y_{t-1}\}$. The values $\tau_2, \tau_3, \ldots$ are the times at which the state changes. These are the active steps of the process. The sequence $Y_{\tau_1}, Y_{\tau_2}, \ldots$ is the same as the Moran process except that repeated states are omitted. Now recall the Markov chain $(X_t)_{t\geq 0}$ with start state $X_0 = n + 1$ and recall the definition of the first passage time T_{n+1} which is the first time that the chain returns to state $n + 1$. Note that the sequence $n+1, |Y_{\tau_1}|, |Y_{\tau_2}|, \ldots, |Y_{\tau_{(T_{n+1}-1)}}|$ is a faithful simulation of the Markov chain $X_0, X_1, \ldots, X_{T_{n+1}-1}$ starting from state $X_0 = n + 1$, up until it reaches state 0 or state n. Also, the absorption time satisfies

$$T_A = \tau_{(T_{n+1}-1)} - 1 = \sum_{j=2}^{T_{n+1}-1} (\tau_j - \tau_{(j-1)})$$

and for $j \geq 2$, $\tau_j - \tau_{j-1}$ is distributed as $\mu(Y_{\tau_{j-1}})$, which is geometric with parameter $p(Y_{\tau_{j-1}})$.

To derive upper and lower bounds for $E[T_A]$, we break the sum into pieces. For $k \in \{1, \ldots, n-1\}$, let

$$T_{A,k} = \sum_{j=2}^{T_{n+1}-1} \Psi_{k,j},$$

where $\Psi_{k,j}$ is geometrically distributed with parameter $p(Y_{\tau_{j-1}})$ if $|Y_{\tau_{j-1}}| = k$ and $\Psi_{k,j} = 0$, otherwise. Then T_A is distributed as $\sum_{k=1}^{n-1} T_{A,k}$. In order to derive upper and lower bounds, let $p_k^+ = \max\{p(S) \mid |S| = k\}$ and $p_k^- = \min\{p(S) \mid |S| = k\}$. Let $T_{A,k}^+ = \sum_{j=2}^{T_{n+1}-1} \Psi_{k,j}^+$ where $\Psi_{k,j}^+$ is geometrically distributed with parameter p_k^- if $|Y_{\tau_{j-1}}| = k$ and $\Psi_{k,j}^+ = 0$, otherwise. Let $T_{A,k}^- = \sum_{j=2}^{T_{n+1}-1} \Psi_{k,j}^-$ where $\Psi_{k,j}^-$ is geometrically distributed with parameter p_k^+ if $|Y_{\tau_{j-1}}| = k$ and $\Psi_{k,j}^- = 0$, otherwise. Then by stochastic domination for the geometric distribution,

$$\sum_{k=1}^{n-1} E[T_{A,k}^-] \leq E[T_A] \leq \sum_{k=1}^{n-1} E[T_{A,k}^+]. \tag{2}$$

Theorem 13.
$$\left(1 - \frac{1}{r^2}\right) \sum_{k=1}^{n-1} \frac{1}{p_k^+} \le E[T_A] \le \left(1 + \frac{1}{r}\right) \sum_{k=1}^{n-1} \frac{1}{p_k^-}.$$

Proof. By (2), $E[T_A]$ is at most $\sum_{k=1}^{n-1} E[T_{A,k}^+]$. Now $T_{A,k}^+$ is a sum of geometric random variables with parameter p_k^-. The number of random variables in the sum is γ_k^{n+1} which is the number of times that state k is visited between visits to state $n+1$ in the Markov chain (X_i). Since $1/p_k^-$ and $E[\gamma_k^{n+1}]$ are both finite (see Corollary 12), Wald's identity (see [**?**, Theorem 12.3]) guarantees that $E[T_{A,k}^+] = E[\gamma_k^{n+1}]/p_k^-$. The upper bound follows from Corollary 12. The lower bound is similar. $\square$

4.5   Upper and lower bounds

We start with the following observation.

Observation 14. If $|S| = k$ then $p(S) = \frac{rm_S^+}{W_k\Delta} + \frac{m_S^-}{W_k\Delta}$ so $\frac{1}{p(S)} = \frac{W_k\Delta}{rm_S^+ + m_S^-}$.

Putting Theorem 13 together with Observation 14 we get the following.

Corollary 15. *The expected absorption time of the Moran process on a strongly connected Δ-regular n-vertex digraph G is at least*
$$\left(1 - \frac{1}{r^2}\right) W_1\Delta \sum_{k=1}^{n-1} \frac{1}{\max\left\{rm_S^+ + m_S^- \mid |S| = k\right\}}$$

and is at most
$$\left(1 + \frac{1}{r}\right) W_n\Delta \sum_{k=1}^{n-1} \frac{1}{\min\left\{rm_S^+ + m_S^- \mid |S| = k\right\}}.$$

We can now prove Theorem 1.

Theorem 1. The expected absorption time of the Moran process on a strongly connected Δ-regular n-vertex digraph G is at least $\left(\frac{r-1}{r^2}\right) nH_{n-1}$ and at most $n^2\Delta$.

Proof. If $|S| = k$ then we have the trivial bound $rm_S^+ + m_S^- \le (r+1)k\Delta$, which, together with Corollary 15, establishes the lower bound. If a digraph is strongly connected, then m_S^+ and m_S^- are at least 1 when $1 \le |S| \le n - 1$ so $rm_S^+ + m_S^- \ge r + 1$. This, together with Corollary 15, establishes the upper bound. $\square$

Note that the upper bound in Theorem 1 generalises the one given in Proposition 9 to the directed case. The following observations follow from special cases of Corollary 15.

Observation 16. Suppose that the graph G is the undirected clique K_n (which is Δ-regular with $\Delta = n - 1$). In this case, for S of size k, $m_S^+ = m_S^- = k(n - k)$, so Corollary 15 shows that the absorption time is at most
$$n \sum_{k=1}^{n-1} \frac{n-1}{k(n-k)} \le n \sum_{k=1}^{n-1} \frac{n-k}{k(n-k)} + n \sum_{k=1}^{n-1} \frac{k}{k(n-k)} \le 2nH_{n-1},$$

matching the lower bound from Theorem 1 up to a constant factor (that depends only on r but not on n).

Observation 17. Suppose that the graph G is the undirected cycle C_n (which is Δ-regular with $\Delta = 2$). Since the process starts with a single mutant, it is easy to see that the set of mutant vertices must be connected, if it is non-empty. Therefore, $m_S^+ = m_S^- = 2$ for any non-trivial S that is reachable from the initial configuration, so the absorption time is at least $\left(1 - \frac{1}{r^2}\right) \frac{2n}{r+1} \sum_{k=1}^{n-1} \frac{1}{2} = \Omega(n^2)$, matching the upper bound from Theorem 1 up to a constant factor.

Observation 18. Suppose that the graph G is the directed n-vertex cycle (which is Δ-regular with $\Delta = 1$). Again, the mutants remain connected; in this case $m_S^+ = m_S^- = 1$ for any non-trivial S so the absorption time is at least $\left(1 - \frac{1}{r^2}\right) \frac{n}{r+1} \sum_{k=1}^{n-1} \frac{1}{2} = \Omega(n^2)$, matching the upper bound from Theorem 1 up to a constant factor.

4.6 Better upper bounds for undirected graphs via isoperimetric numbers

Suppose that a graph G is undirected. As in the proof of Proposition 9, let ∂S be the set of (undirected) edges between vertices in S and vertices in $V(G) \setminus S$. Then $m_S^+ = m_S^- = |\partial S|$. The isoperimetric number of the graph G was defined by Buser [?] as follows

$$i(G) = \min\left\{ \frac{|\partial S|}{|S|} \mid S \subseteq V(G), 0 < |S| \leq |V(G)|/2 \right\}.$$

The quantity $i(G)$ is a discrete analogue of the Cheeger isoperimetric constant. For graphs with good expansion, Theorem 3 improves the upper bound in Theorem 1.

Theorem 3. The expected absorption time of the Moran process on a connected Δ-regular n-vertex undirected graph G is at most $2\Delta n H_n / i(G)$.

Proof. From Corollary 15, the expected absorption time is at most

$$\frac{\Delta W_n}{r} \sum_{k=1}^{n-1} \frac{1}{\min\{|\partial_S| \mid S \subseteq V(G), |S| = k\}}.$$

This is at most

$$\frac{\Delta W_n}{r} \left(2 \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{1}{\min\{|\partial_S| \mid S \subseteq V(G), |S| = k\}} \right)$$

$$= \frac{\Delta W_n}{r} \left(2 \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{1}{k \min\left\{ \frac{|\partial_S|}{k} \mid S \subseteq V(G), |S| = k \right\}} \right)$$

$$\leq \frac{\Delta W_n}{i(G)r} \left(2 \sum_{k=1}^{\lfloor n/2 \rfloor} \frac{1}{k} \right) = \frac{2\Delta W_n H_{\lfloor n/2 \rfloor}}{i(G)r}. \qquad \square$$

4.7 Families for which the upper bound is optimal

For every fixed $\Delta > 2$, we construct an infinite family of connected, Δ-regular undirected graphs for which the upper bound in Theorem 1 is optimal, up to a constant factor (which may depend upon r and Δ but not on n).

To do this, we define the graph H_Δ to be $K_{\Delta-2,\Delta-1}$ with the addition of edges forming a cycle on the side with $\Delta - 1$ vertices. Note that $\Delta - 2$ vertices have degree $\Delta - 1$ and the others have degree Δ.

Now, let $G_{\ell,\Delta}$ be the Δ-regular graph formed from a cycle $x_1 \ldots x_\ell x_1$ and ℓ disjoint copies of H_Δ by adding an undirected edge between x_i and each of the vertices of degree $\Delta - 1$ in the i'th copy of H_Δ, for each $i \in [\ell]$. Note that $|V(G_{\ell,\Delta})| = 2\ell(\Delta - 1)$.

Theorem 19. *For $r > 1$ and sufficiently large ℓ (with respect to r), the expected absorption time of the Moran process on $G_{\ell,\Delta}$ exceeds $\frac{1}{2r}(1 - \frac{1}{r})\ell^2$.*

Proof. Let $n = n(\ell, \Delta) = |V(G_{\ell,\Delta})|$ and let $T_{\mathrm{d}} = \ell^2/r$.

Let $(Y_t)_{t\geq 1}$ be the discrete-time Moran process on $G_{\ell,\Delta}$ and consider the following events. Let $\mathcal{F}^*$ be the event that $Y_{T_{\mathrm{d}}+1} = V(G_{\ell,\Delta})$, i.e., that the process reaches fixation in at most T_{d} steps. Let $\mathcal{E}^*$ be the event that $Y_{T_{\mathrm{d}}+1} = \emptyset$, i.e., that the process reaches extinction in at most T_{d} steps. $\Pr(\mathcal{E}^*)$ is at most the extinction probability, which is less than $\frac{1}{r}$ since G is regular, so the fixation probability is the quantity f from Lemma 11 [?] which exceeds $1 - \frac{1}{r}$ for $r > 1$.

To bound $\Pr(\mathcal{F}^*)$, consider the continuous-time version of the process, $\widetilde{Y}[t]$. We will show that, by time $T_{\mathrm{c}} = 2T_{\mathrm{d}}/n = 2\ell^2/rn$, it is very likely that the continuous process will have had at least T_{d} reproductions. Let $\mathcal{S}$ (for "slow") be the event that the continuous process has not had T_{d} reproductions by time T_{c}. Let $\widetilde{\mathcal{F}}^*$ be the event that it has reached fixation by time T_{c}. We have

$$\Pr(\mathcal{F}^*) = \Pr(\mathcal{F}^* \wedge \neg\mathcal{S}) + \Pr(\mathcal{F}^* \wedge \mathcal{S}) \leq \Pr(\widetilde{\mathcal{F}}^*) + \Pr(\mathcal{S}).$$

In the continuous process, each of the n vertices reproduces at rate at least 1 so the number N of reproductions up to time T_{c} is stochastically bounded below by a Poisson random variable with parameter $nT_{\mathrm{c}} = 2\ell^2/r$. By a Chernoff-type argument [?, Theorem 5.4], we have

$$\Pr(\mathcal{S}) = \Pr(N \leq T_{\mathrm{d}}) \leq e^{-nT_{\mathrm{c}}}\left(\frac{enT_{\mathrm{c}}}{T_{\mathrm{d}}}\right)^{T_{\mathrm{d}}} = e^{-2\ell^2/r}(2e)^{\ell^2/r} < \tfrac{1}{4}\left(1 - \tfrac{1}{r}\right),$$

for large enough ℓ.

To bound $\Pr(\widetilde{\mathcal{F}}^*)$, consider the process $\widetilde{Z}[t]$ on $G_{\ell,\Delta}$ that behaves like $\widetilde{Y}[t]$ except for the two following points.

- For some i, we have $\widetilde{Y}[0] = \{x_i\}$ or $\widetilde{Y}[0]$ is in the copy of H_Δ attached to x_i. Let $\widetilde{Z}[0] = \widetilde{Y}[0] \cup \{x_i\}$.

- No mutant in the cycle $x_1 \ldots x_\ell x_1$ can ever be replaced by a non-mutant. That is, if, at time t, a non-mutant neighbour of some x_i ($i \in [\ell]$) is selected to reproduce to x_i, then the state does not change.

We couple the processes $\widetilde{Y}[t]$ and $\widetilde{Z}[t]$ as follows. Let t be such that $\widetilde{Y}[t] \subseteq \widetilde{Z}[t]$, noting that $t = 0$ has this property. The coupling lemma (Lemma 5) allows us to maintain $\widetilde{Y}[t+\tau] \subseteq \widetilde{Z}[t+\tau]$ until the next time, t', at which a mutant at one of the x_i is replaced by a non-mutant in $\widetilde{Y}[t]$. This maintains the property that $\widetilde{Y}[t'] \subseteq \widetilde{Z}[t']$, so the coupling can be restarted from this point.

Now $\Pr(\widetilde{\mathcal{F}}^*) = \Pr(\widetilde{Y}[T_{\mathrm{c}}] = V(G_{\ell,\Delta})) \leq \Pr(\widetilde{Z}[T_{\mathrm{c}}] = V(G_{\ell,\Delta}))$, so we will find an upper bound for $\Pr(\widetilde{Z}[T_{\mathrm{c}}] = V(G_{\ell,\Delta}))$. Let $C = \{x_1, \ldots, x_\ell\}$. The set $\widetilde{Z}[t] \cap C$ is non-empty,

non-decreasing and connected in $G_{\ell,\Delta}$. If $\widetilde{Z}[t] \cap C$ is a proper subset of C then it increases exactly when one of the two mutants in C reproduces to its non-mutant neighbour in C or, if there is only one mutant in C, when that mutant reproduces to either of its neighbours in the cycle. In both cases, this happens with rate $\frac{2r}{\Delta}$, so $|\widetilde{Z}[t] \cap C|$ is bounded from above by a Poisson random variable with parameter $\frac{2r}{\Delta}t$. Let $\lambda^* = 4\ell/9$. For $t = T_{\mathrm{c}}$ the parameter is $\frac{2r}{\Delta}T_{\mathrm{c}} = \frac{4\ell^2}{n\Delta} < \frac{4\ell}{\Delta^2} \leq \lambda^*$ so $|\widetilde{Z}[t] \cap C|$ is bounded above by a Poisson random variable Ψ^* with parameter λ^*. Therefore, $E[|\widetilde{Z}[T_{\mathrm{c}}] \cap C|] \leq \lambda^*$, and (see, for example, Theorem 5.4 of [?]) we have

$$\Pr\left(|\widetilde{Z}[T_{\mathrm{c}}] \cap C| \geq \tfrac{8}{9}\ell\right) \leq \Pr(\Psi^* \geq 2\lambda^*) \leq e^{-\lambda^*}\left(\tfrac{e}{2}\right)^{2\lambda^*} = \left(\tfrac{e}{4}\right)^{\lambda^*} < \tfrac{1}{4}\left(1 - \tfrac{1}{r}\right),$$

for large enough ℓ (with respect to r). Now,

$$E[T_{\mathrm{A}}] \geq E\left[T_{\mathrm{A}} \mid \overline{\mathcal{F}^* \cup \mathcal{E}^*}\right] \times \Pr\left(\overline{\mathcal{F}^* \cup \mathcal{E}^*}\right).$$

Clearly, the expected absorption time of the discrete process conditioned on absorption not occurring within ℓ^2/r steps is at least ℓ^2/r. Meanwhile,

$$\Pr\left(\overline{\mathcal{F}^* \cup \mathcal{E}^*}\right) > 1 - \tfrac{1}{r} - 2\tfrac{1}{4}\left(1 - \tfrac{1}{r}\right) = \tfrac{1}{2}\left(1 - \tfrac{1}{r}\right)$$

for large enough ℓ (with respect to r). $\qquad\square$

Thus, we have shown that the $O(n^2)$ upper bound of Theorem 1 is tight up to a constant factor (which may depend on r and Δ, but not on n).

Theorem 2. Suppose that $r > 1$ and $\Delta > 2$. There is an infinite family $\mathcal{G}$ of Δ-regular graphs such that, when the Moran process is run on an n-vertex graph $G \in \mathcal{G}$, the expected absorption time exceeds $\frac{1}{8r}(1 - \frac{1}{r})\frac{n^2}{(\Delta-1)^2}$.

Proof. For a given value of r, Let ℓ_r be the smallest value of ℓ for which that Theorem 19 applies. Take $\mathcal{G} = \{G_{\ell,\Delta} \mid \ell \geq \ell_r\}$ and the result is immediate from Theorem 19 and the fact that $G_{\ell,\Delta}$ has $2\ell(\Delta - 1)$ vertices. $\qquad\square$

5 General digraphs

Fix $r > 1$ and let $\varepsilon_r = \min(r - 1, 1)$. Theorem 7 of [?] shows that the expected absorption time of the Moran process on a connected n-vertex undirected graph is at most $(r/(r-1))n^4$. Theorem 1 shows that the expected absorption time on a strongly connected Δ-regular digraph is at most $n^2\Delta$. In contrast, we show that there is an infinite family of strongly connected digraphs such that the expected absorption time of the Moran process on an n-vertex graph from the family is $2^{\Omega(n)}$.

5.1 The family of graphs

Let $G_{r,N}$ be the disjoint union of the complete graph K_N (with bidirectional edges), a directed path $P = u_1 \ldots u_N$ and a directed path $Q = v_{4\lceil r \rceil N} \ldots v_0$, along with the directed edge $(u_N, v_{4\lceil r \rceil N})$ and the following directed edges (see Figure 1):
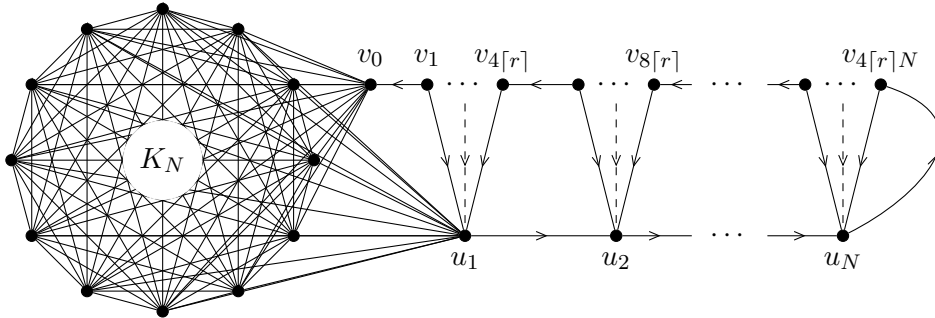
- (x, u_1) and (v_0, x) for every $x \in K_N$;

Figure 1: The graph $G_{r,N}$. The edges within the clique are bidirectional; v_0 sends a directed edge to every vertex in the clique and u_1 receives one from each. Other edges are directed as indicated.

- $(v_{4\lceil r \rceil (i-1)+j}, u_i)$ for each $i \in [N]$, $j \in [4 \lceil r \rceil]$.

The intuition is as follows. Consider the Moran process on $G_{r,N}$. With probability close to $\frac{1}{4\lceil r \rceil +2}$, the initial mutant is in the clique (Observation 21). Conditioned on this, Lemmas 22, 24 and 25 allow us to show that it is fairly likely that there is a time during the first N^3 steps when the clique is half full, but that absorption does not happen in the first $T^*(N)$ steps for a function T^* which is exponential in N. Of course, the expected absorption time conditioned on this is at least $T^*(N)$, so we conclude (Theorem 4) that the overall expected absorption time is at least $T^*(N)$.

The main challenge of the proof is the second step — showing that is it is fairly likely that there is a time during the first N^3 steps when the clique is half full, but that absorption does not happen in the first $T^*(N)$ steps. The fact that the clique becomes half full (Lemma 22) follows by dominating the number of mutants in the clique during the initial stages of the process by an appropriate one-dimensional random walk, and then showing that sufficiently many random-walk steps are actually taken during the first N^3 steps of the process. The fact that extinction is then unlikely in the first $T^*(N)$ steps (Lemma 24) follows from the fact that the many mutants in the clique are unlikely to become extinct very quickly. On the other hand, the fact that fixation is unlikely in the first $T^*(N)$ steps (Lemma 25) follows from the fact that mutants make slow progress along the path P because because vertices in Q tend to push the "mutant frontier" backwards towards u_1. However, the chain of mutants has to push all the way around this chain in order for fixation to occur.

5.2 The one-dimensional random walk

The following Lemma is Example 3.9.6 from [**?**].

Lemma 20. *Let $(Z_t)_{t \geq 0}$ be the random walk on $\{0, \ldots, n\}$ with absorbing barriers at 0 and n and, for $0 < Z_t < n$, let $Z_{t+1} = Z_t + 1$ with probability $p \neq \frac{1}{2}$ and $Z_{t+1} = Z_t - 1$ with probability $q = 1 - p$. Let p_i be the probability of absorption at 0, given that $Z_0 = i$. Writing $\rho = q/p$,*

$$p_i = \frac{\rho^i - \rho^n}{1 - \rho^n} .$$

5.3 Bounding the absorption time

Consider the Moran process $(Y_t)_{t \geq 1}$ on $G_{r,N}$. We will assume that N is sufficiently large with respect to r — the exact inequalities that we need will be presented as they arise in the proof. Let $n = 1 + (4\lceil r \rceil + 2)N$ be the number of vertices of $G_{r,N}$. Let $W_t = n + (r-1)|Y_t|$ be the total fitness of Y_t. Let T_A be the absorption time of the process. Our goal is to show that $E[T_A]$ is exponentially large, as a function of N. Let

$$T^*(N) = \left\lfloor \left(\frac{\varepsilon_r}{32}\right)(2^N - 1) \right\rfloor$$

and let N' denote $\lfloor N/2 \rfloor$. We now identify various events which we will study in the lemmas that follow.

- Let $\mathcal{S}_C$ be the event that $Y_1 \subseteq K_N$ (mnemonic: $\mathcal{S}_C$ is the event that the initial mutant starts in the clique; $\mathcal{S}$ is for "Starts" and $\mathcal{C}$ is for "Clique").

- Let $\mathcal{H}_{C,t}$ be the event that $|Y_t \cap K_N| \geq N'$ (mnemonic: $\mathcal{H}_{C,t}$ is the event that the clique is half full at time t. $\mathcal{H}$ is for "Half").

- Let $\mathcal{H}_C = \bigcup_{t \in [N^3 + 1]} \mathcal{H}_{C,t}$.

- Let $\mathcal{F}^*$ be the event that $Y_{T^*(N)+1} = V(G_{r,N})$ (mnemonic: $\mathcal{F}^*$ is the event that fixation occurs after at most $T^*(N)$ steps; $\mathcal{F}$ is for "Fixation").

- Let $\mathcal{E}^*$ be the event that $Y_{T^*(N)+1} = \emptyset$ (mnemonic: $\mathcal{E}^*$ is the event that extinction occurs after at most $T^*(N)$ steps; $\mathcal{E}$ is for "Extinction").

Observation 21.
$$\Pr(\mathcal{S}_C) = \frac{N}{n} = \frac{N}{N(4\lceil r \rceil + 2) + 1}.$$

Lemma 22. $\Pr(\mathcal{H}_C \mid \mathcal{S}_C) \geq \varepsilon_r/8$.

Proof. Let $Z'_t = |Y_t \cap K_N|$. We will condition on the fact that $\mathcal{S}_C$ occurs, so $Z'_1 = 1$. If $Z'_t \in \{1, \ldots, N' - 1\}$ then for all z_t and w_t,

$$\Pr(Z'_{t+1} = z'_t + 1 \mid Z'_t = z_t, W_t = w_t) \geq \left(\frac{rz'_t}{w_t}\right)\left(\frac{N - z'_t}{N}\right).$$

(The probability is greater than this if v_0 is in Y_t.) Also,

$$\Pr(Z'_{t+1} = z'_t - 1 \mid Z'_t = z'_t, W_t = w_t) \leq \left(\frac{1}{w_t}\right)\left(\frac{z'_t}{N}\right) + \left(\frac{N - z'_t}{w_t}\right)\left(\frac{z'_t}{N}\right),$$

where the first term comes from reproduction from a non-mutant at v_0 and the second from reproduction within the clique. Also, $Z'_{t+1} \in \{Z'_t - 1, Z'_t, Z'_t + 1'\}$. Now let

$$p' = \frac{r}{r + 1 + \frac{2}{N}}$$

and note that

$$\frac{\Pr(Z'_{t+1} = z'_t + 1 \mid Z'_t = z'_t)}{\Pr(Z'_{t+1} = z'_t + 1 \mid Z'_t = z'_t) + \Pr(Z'_{t+1} = z'_t - 1 \mid Z'_t = z'_t)} \geq \frac{r(N - z'_t)}{r(N - z'_t) + 1 + (N - z'_t)} \geq p'.$$

The restriction of (Z'_t) to steps where the state changes, stopping when Z'_t reaches 0 or N' is a process which is dominated below by (Z_t), a random walk on $\{0, \ldots, N'\}$ that starts at 1 and absorbs at 0 and N' and has parameter p'.

Now let $\mathcal{E}_\mathcal{C}$ be the event that there is a t with $Y_t \cap K_N = \emptyset$ such that, for all $t' < t$, we have $|Y_{t'} \cap K_N| < N'$. $\mathcal{E}_\mathcal{C}$ is the event that the clique becomes empty before it becomes half full. Then applying Lemma 20 to the dominating random walk, $\Pr(\mathcal{E}_\mathcal{C} \mid \mathcal{S}_\mathcal{C}) \leq \frac{\rho - \rho^{N'}}{1 - \rho^{N'}} \leq \rho$, where $\rho = (1 - p')/p' \leq (N + 2)/(N + \varepsilon_r N)$. Since we are taking N to be sufficiently large with respect to ε_r (in particular, we will take $N \geq 4/\varepsilon_r$) we have $\rho \leq (1 + \varepsilon_r/2)/(1 + \varepsilon_r)$ which is at most $1 - \varepsilon_r/4$ since $\varepsilon_r \leq 1$.

Let $\mathcal{Q}$ be the event that there is a $t \in [N^3 + 1]$ with $|Y_t \cap K_n| \notin \{1, \ldots, N' - 1\}$. ($\mathcal{Q}$ is the event that the number of mutants in the clique changes quickly — the clique takes at most N^3 steps to either become empty or to become at least half full). We will show below that $\Pr(\mathcal{Q} \mid \mathcal{S}_\mathcal{C}) \geq 1 - \varepsilon_r/8$. Note that if $\mathcal{Q}$ occurs but $\mathcal{E}_\mathcal{C}$ does not occur then $\mathcal{H}_\mathcal{C}$ occurs. So

$$\Pr(\mathcal{H}_\mathcal{C} \mid \mathcal{S}_\mathcal{C}) \geq \Pr(\mathcal{Q} \setminus \mathcal{E}_\mathcal{C} \mid \mathcal{S}_\mathcal{C}) \geq \Pr(\mathcal{Q} \mid \mathcal{S}_\mathcal{C}) - \Pr(\mathcal{E}_\mathcal{C} \mid \mathcal{S}_\mathcal{C}) \geq (1 - \varepsilon_r/8) - (1 - \varepsilon_r/4) \geq \varepsilon_r/8,$$

which would complete the proof.

We conclude the proof, then, by showing $\Pr(\neg \mathcal{Q} \mid \mathcal{S}_\mathcal{C}) \leq \varepsilon_r/8$. So we will show that $\Pr(Z'_1, \ldots, Z'_{N^3+1} \in \{1, \ldots, N' - 1\} \mid Z'_1 = 1) \leq \varepsilon_r/8$. For this, let $\delta = \varepsilon_r/20$ and let $\Psi = N'/\delta$. We require that N is sufficiently large with respect to r, so $N^3 \geq 4n\Psi$.

Let $\Upsilon_1, \ldots, \Upsilon_{1+\Psi}$ be Ψ steps of a random walk on the integers that starts with $\Upsilon_1 = 1$ and has $\Upsilon_{t+1} = \Upsilon_t + 1$ with probability p' and $\Upsilon_{t+1} = \Upsilon_t - 1$ with probability $1 - p'$. It is likely that the state Υ_t increases at least $(1 - \delta)p'\Psi$ times. By a Chernoff bound (e.g., [?, Theorem 4.5]), the probability that this does not happen is at most $\exp(-p'\Psi\delta^2/2)$. Since N is sufficiently large with respect to r (and therefore N' is sufficiently large with respect to r) and $p' \geq r/(r + 2)$, this probability is at most $\varepsilon_r/16$. (This calculation is not tight in any way — $\varepsilon_r/16$ happens to be sufficient.) If there are at least $(1 - \delta)p'\Psi$ increases then

$$\Upsilon_{\Psi+1} - \Upsilon_1 \geq (2(1 - \delta)p' - 1)\Psi.$$

But since

$$\delta = \frac{\left(\frac{\varepsilon_r}{4}\right)}{5} < \frac{\frac{\varepsilon_r}{2} - \frac{1}{N}}{2 + \frac{3\varepsilon_r}{2} + \frac{1}{N}},$$

we have

$$p' = \frac{1 + \varepsilon_r}{2(1 + \frac{\varepsilon_r}{2} + \frac{1}{N})} > \frac{1 + \delta}{2(1 - \delta)}.$$

so

$$\Upsilon_{\Psi+1} - \Upsilon_1 \geq (2(1 - \delta)p' - 1)\Psi \geq \delta\Psi = N'. \tag{3}$$

Now if $Z'_t \in \{1, \ldots, N' - 1\}$ then

$$\Pr(Z'_{t+1} \neq z'_t \mid Z'_t = z'_t, W_t = w_t) \geq \frac{(r + 1)z'_t(N - z'_t)}{Nw_t} \geq \frac{1}{2n}.$$

(Again, this calculation is not tight, but $1/(2n)$ suffices.) If we select N^3 Bernoulli random variables, each with success probability $\frac{1}{2n}$, then, by another Chernoff bound, the probability that we fail to get at least $N^3/(4n) \geq \Psi$ successes is at most $\exp(-N^3/(16n)) \leq \varepsilon_r/16$.

Now $\Pr(Z'_1, \ldots, Z'_{N^3+1} \in \{1, \ldots, N' - 1\} \mid Z'_1 = 1)$ is at most the sum of two probabilities.

- The probability that $Z'_1, \ldots, Z'_{N^3+1}$ are all in $\{1, \ldots, N'-1\}$ and there fewer than Ψ values t with $Z'_{t+1} \neq Z'_t$. This is dominated by the the selection of N^3 Bernoulli random variables as above, and the probability is at most $\varepsilon_r/16$.

- The probability that $Z'_1, \ldots, Z'_{N^3+1}$ are all in $\{1, \ldots, N'-1\}$ and there are at least Ψ values t with $Z'_{t+1} \neq Z'_t$. In order to bound this probability, imagine the evolution of the process proceeding in two sub-steps at each step. First, decide whether $Z'_{t+1} \neq Z'_t$ with the appropriate probability. If so, select $Z'_{t+1} \in \{Z'_t + 1, Z'_t - 1\}$ with the appropriate probability. The probability of the whole event is then dominated above by the probability that $\Upsilon_{\Psi+1} - \Upsilon_1 < N'$, which is at most $\varepsilon_r/16$, as we showed above. (If this difference is at least N' then either the Z'_t process hits 0 before it changes for the Ψ'th time, or Z'_t reaches N'.) $\qquad\square$

Lemma 23. $\Pr(Y_{T^*(N)+t} = \emptyset \mid \mathcal{H}_{C,t} \wedge \mathcal{S}_\mathcal{C}) \leq \varepsilon_r/(32(N^3+1))$.

Proof. As in the proof of Lemma 22, let $Z'_t = |Y_t \cap K_N|$. If $\mathcal{H}_{C,t}$ holds then $Z'_t \geq N'$. Let Γ be the number of distinct values t' with $t < t' < \inf\{t'' > t \mid Z'_{t''} = 0\}$ satisfying $Z'_{t'} = N'$. We will show that $\Pr(\Gamma < T^*(N)) \leq \varepsilon_r/(32(N^3+1))$.

For any $T > t$, suppose that $Z'_T = N' - 1$ and let $T' = \inf\{t'' > t \mid Z'_{t''} \in \{0, N'\}\}$. Let $\pi = \Pr(Z'_{T'} = 0)$. By the argument in the proof of Lemma 22, π is at most the probability that a random walk Z_t on $\{0, \ldots, N'\}$ which starts at $N' - 1$ and absorbs at 0 and N' and has parameter p' absorbs at 0. By Lemma 20,

$$\pi \leq \frac{\rho^{N'-1} - \rho^N}{1 - \rho^N} \leq \rho^{N'-1},$$

where $\rho = (N+2)/(N + \varepsilon_r N) < 1$. Then $\Pr(\Gamma < T^*(N)) \leq T^*(N)\rho^{N'-1}$. We choose N to be sufficiently large so that

$$T^*(N) = \left\lfloor \left(\frac{\varepsilon_r}{32}\right)(2^N - 1) \right\rfloor \leq \left(\frac{\varepsilon_r}{32(N^3+1)}\right)\left(\frac{N + \varepsilon_r N}{N + 2}\right)^{N'-1}.$$

Then $\Pr(\Gamma < T^*(N)) \leq T^*(N)\rho^{N'-1} \leq \varepsilon_r/(32(N^3+1))$, which completes the proof. $\qquad\square$

Lemma 24. $\Pr(\mathcal{E}^* \mid \mathcal{H}_\mathcal{C} \wedge \mathcal{S}_\mathcal{C}) \leq \varepsilon_r/32$.

Proof. This follows easily from Lemma 23 using the following summation.

$$\begin{aligned}
\Pr(\mathcal{E}^* \mid \mathcal{H}_\mathcal{C} \wedge \mathcal{S}_\mathcal{C}) = \frac{\Pr(\mathcal{E}^* \wedge \mathcal{H}_\mathcal{C} \wedge \mathcal{S}_\mathcal{C})}{\Pr(\mathcal{H}_\mathcal{C} \wedge \mathcal{S}_\mathcal{C})} &\leq \frac{\sum_{t \in [N^3+1]} \Pr(\mathcal{E}^* \wedge \mathcal{H}_{C,t} \wedge \mathcal{S}_\mathcal{C})}{\Pr(\mathcal{H}_\mathcal{C} \wedge \mathcal{S}_\mathcal{C})} \\
&= \sum_{t \in [N^3+1]} \frac{\Pr(\mathcal{E}^* \mid \mathcal{H}_{C,t} \wedge \mathcal{S}_\mathcal{C}) \Pr(\mathcal{H}_{C,t} \mid \mathcal{S}_\mathcal{C}) \Pr(\mathcal{S}_\mathcal{C})}{\Pr(\mathcal{H}_\mathcal{C} \mid \mathcal{S}_\mathcal{C}) \Pr(\mathcal{S}_\mathcal{C})} \\
&\leq \sum_{t \in [N^3+1]} \Pr(\mathcal{E}^* \mid \mathcal{H}_{C,t} \wedge \mathcal{S}_\mathcal{C}) \\
&\leq \sum_{t \in [N^3+1]} \Pr(Y_{T^*(N)+t} = \emptyset \mid \mathcal{H}_{C,t} \wedge \mathcal{S}_\mathcal{C}). \qquad\square
\end{aligned}$$

Lemma 25. $\Pr(\mathcal{F}^* \mid \mathcal{S}_\mathcal{C}) \leq \varepsilon_r/32$.

Proof. Recall the paths $P = u_1 \ldots u_N$ and $Q = v_{4\lceil r \rceil N} \cdots v_0$ in $G_{r,N}$. Define U_t as follows.

- If $Y_t \cap Q$ is non-empty then $U_t = N$.

- If $Y_t \cap Q$ and $Y_t \cap P$ are both empty then $U_t = 0$.

- If $Y_t \cap Q$ is empty and $Y_t \cap P$ is non-empty then $U_t = \max\{i \mid u_i \in Y_t\}$.

Let $\tau = \inf\{t \mid U_t = N\}$. We will show that $\Pr(\tau < T^*(N)) \leq \varepsilon_r/32$. If $U_t \in \{1, \ldots, N-1\}$ then $\Pr(U_{t+1} = U_t + 1) = \frac{r}{W_t}$ and $\Pr(U_{t+1} = U_t - 1) \geq \left(\frac{4\lceil r \rceil}{W_t}\right)\left(\frac{1}{2}\right) \geq \frac{2r}{W_t}$. Also, $U_{t+1} \in \{U_{t-1}, U_t, U_t + 1\}$.

Let $F_t = \inf\{t' > t \mid U_{t'} \in \{0, N\}\}$ and $\gamma = \Pr(U_{F_t} = N \mid U_t = 1)$. γ is at most the probability of absorbing at N in a random walk on $\{0, \ldots, N\}$ that starts at 1, absorbs at 0 and N and has parameter $1/3$ (twice as likely to go down as to go up). By Lemma 20

$$\gamma \leq 1 - \left(\frac{\rho - \rho^N}{1 - \rho^N}\right),$$

where $\rho = 2$, so $\gamma \leq 1/(2^N - 1)$. Now let Ψ be the number of times $t < \tau$ with $U_t = 0$. Then $\Psi \leq \tau$ so

$$\Pr(\tau < T^*(N)) \leq \Pr(\Psi < T^*(N)) \leq T^*(N)\gamma \leq \varepsilon_r/32. \qquad \square$$

Putting together the lemmas in this section, we prove **Theorem 4**. Recall that $T^*(N) = \left\lfloor \left(\frac{\varepsilon_r}{32}\right)(2^N - 1)\right\rfloor$. For convenience, we restate the theorem using this notation.

Theorem 26. *Fix $r > 1$ and let $\varepsilon_r = \min(r - 1, 1)$. Suppose that N is sufficiently large with respect to r and consider the Moran process $(Y_t)_{t \geq 1}$ on $G_{r,N}$. Let T_A be the absorption time of the process. Then $E[T_A] > \frac{1}{16}T^*(N)\,\varepsilon_r/(4\lceil r \rceil + 3)$.*

Proof.

$$E[T_A] \geq E[T_A \mid \mathcal{S}_\mathcal{C} \wedge \mathcal{H}_\mathcal{C} \setminus (\mathcal{F}^* \cup \mathcal{E}^*)] \times \Pr(\mathcal{H}_\mathcal{C} \setminus (\mathcal{F}^* \cup \mathcal{E}^*) \mid \mathcal{S}_\mathcal{C}) \times \Pr(\mathcal{S}_\mathcal{C}).$$

The first term on the right-hand side is greater than $T^*(N)$ by the definition of the excluded events $\mathcal{F}^*$ and $\mathcal{E}^*$. Lemmas 22, 24 and 25 show that the second term on the right-hand side is at least $\varepsilon_r/8 - 2\varepsilon_r/32 \leq \varepsilon_r/16$. Finally, Observation 21 shows that the third term on the right-hand side is at least $1/(4\lceil r \rceil + 3)$. $\qquad \square$