# Lower Bounds and Algorithms for Dominating Sets in Web Graphs

Colin Cooper, Ralf Klasing, and Michele Zito

**Abstract.** In this paper we study the size of generalised dominating sets in two graph processes that are widely used to model aspects of the World Wide Web. On the one hand, we show that graphs generated this way have fairly large dominating sets (i.e., linear in the size of the graph). On the other hand, we present efficient strategies to construct small dominating sets.

The algorithmic results represent an application of a particular analysis technique which can be used to characterise the asymptotic behaviour of a number of dynamic processes related to the web.

## 1. Introduction

In recent years the World Wide Web has grown dramatically. Its current size is measured in billions of pages [Levene and Wheeldon 01], and pages are added to it every day. As this graph (nodes correspond to web pages and edges to links between pages) continues to grow it becomes increasingly important to study mathematical models that capture its structural properties [Broder et al. 00, Kumar et al. 00]. Such models can be used to design efficient algorithms for web applications and may even uncover unforeseen properties of this huge evolving structure. Several mathematical models for analysing the web have been proposed (for instance, [Bollobás et al. 01, Cooper and Frieze 03, Kumar et al. 00]). The (evolution of the) web graph is usually modelled by a (random)

process in which new vertices appear from time to time. Such vertices may be linked randomly to the existing structure through some form of *preferential attachment*: existing vertices with many neighbours are somewhat more likely to be linked to the newcomers.

The main focus of research so far has been on capturing empirically observed features of the web [Barabási and Albert 99, Broder et al. 00]. No attempt has been made to characterise graph-theoretic substructures of such graphs. We initiate such an investigation by looking at sets of vertices that, in a sense, cover all other vertices. More formally, a vertex in a graph *dominates* all vertices that are adjacent to it (parallel edges give multiple domination). In the spirit of Harary and Haynes [Harary and Haynes 00], an *h-dominating set* for a graph $G = (V, E)$ is a set $\mathcal{S} \subseteq V$ such that each vertex in $V \setminus \mathcal{S}$ is dominated at least $h$ times by vertices in $\mathcal{S}$. Let $\gamma_h = \gamma_h(G)$ denote the size of the smallest $h$-dominating sets in $G$. The *minimum h-dominating set* problem (M$h$DS) asks for an $h$-dominating set of size $\gamma_h$.

Dominating sets play an important role in many practical applications, e.g., in the context of distributed computing or mobile ad-hoc networks [Alzoubi et al. 02, Duckworth and Zito 00, Stojmenovic et al. 02]. The reader is referred to [Haynes et al. 98a, Haynes et al. 98b] for an in-depth view of the subject. The typical fundamental task in such applications is to select a subset of nodes in the network that will "provide" a certain service to all other vertices. For this to be time-efficient, all other vertices must be directly connected to the selected nodes, and in order for it to be cost-effective, the number of selected nodes must be minimal. In relation to web graphs, a dominating set may be used to devise efficient web searches. For $h > 1$, an $h$-dominating set can be considered as a more fault-tolerant structure. If up to $h-1$ vertices or edges fail, the domination property is still maintained (i.e., it is still possible to provide the service).

The M$h$DS problem is NP-hard [Garey and Johnson 78, Klasing and Laforest 04]; moreover, it is not likely that it may be approximated effectively [Feige 98]. Polynomial time algorithms exist on special classes of graphs (e.g., [Liao and Chang 03]). The M1DS problem has been studied also in random graphs. In the *binomial model* $G(n, p)$ [Wieland and Godbole 01] the value of $\gamma_1$ can be pinpointed quite precisely, provided that $p$ is not too small compared to $n$. In random regular graphs of degree $r$ (see, for example, results in the *configuration model* [Zito 01] and references therein), upper and lower bounds are known on $\gamma_1$.

In this paper we look at simple and efficient algorithms for building small $h$-dominating sets in graphs. The performance guarantees of these algorithms are analysed under the assumption that the input is a random web graph. We also analyse the tightness of the performances of such algorithms, by proving

combinatorial lower bounds on $\gamma_h$, for any fixed $h \geq 1$. Such bounds, often disappointingly weak, offer nevertheless a proof that, most of the time, the sets returned by the various algorithms are only a constant factor away from an optimal answer. Finally, we compare the quality of the solutions returned by (some of) our algorithms with empirical average values of $\gamma_1$.

The main outcome of this paper can be stated informally by saying that web graphs have fairly large dominating sets. Hence, a crawler who wants to use a dominating set to explore the web will need to store a large proportion of the whole graph. Interestingly, the results in this paper also uncover a(nother) difference between models of the web based on preferential attachment and more traditional random graph models. The tendency to choose neighbours of high degree affects the size of the smallest dominating sets.

Most of our algorithms are *on-line* in the sense that the decision to add a particular vertex to the dominating set is taken without total information about the web graph under consideration and *greedy* in the sense that decisions, once taken, are never changed. The algorithms are also quite efficient: only a constant amount of time is used per update of the dominating set. Such algorithms are of particular interest in the context of web graphs. As the web graph is evolving, one wants to decide whether a new vertex is to be added to the already-existing dominating set without recomputing the existing dominating set and with minimal computational effort. On-line strategies for the dominating-set problem have been considered in the past for general graphs [Eidenbenz 02, King and Tzeng 97]. However, the authors are not aware of any result on on-line algorithms for this problem in random graphs.

Our results hold *asymptotically almost surely* (a.a.s.), i.e., with probability approaching one as the size of the web graph grows to infinity. The algorithmic results are based on the analysis of a number of (Markovian) random processes. In each case the properties of the process under consideration lead to the definition of a (deterministic) continuous function that is very close (in probability) to the values of the process, as the size of the graph grows. It should be pointed out at this stage that the proposed analysis methodology is quite general. We apply it to analyse heuristics for the M$h$DS problem only, but it could be used to prove results about other graph parameters such as the independence or the chromatic number. The method is closely related to the so-called *differential equation method* [Wormald 99]. In fact, a version of the main analytical tool proposed by Wormald can be adapted to work for the processes considered in this paper. However, the machinery employed in [Wormald 99] is not needed to analyse the processes considered in this paper. Our results are obtained by proving the concentration of the various processes of interest around their mean and by devising a method for getting close estimates on the relevant expectations.

In Section 2, we review the definitions of the models of web graphs that we will use. We also state our main result in the context of these models and present more detailed comments on our general analysis method. In the following section we consider a very simple algorithm and apply the proposed method to obtain nontrivial upper bounds on $\gamma_1$. Refined algorithms are introduced and analysed in Sections 4 and 5. In Section 6 we discuss generalisations to $h > 1$. Then, we turn to lower bounds. In Sections 7 and 8 we present our argument for the lower bounds stated in Section 2. Finally, we briefly comment on some empirical work carried out on a sub-class of the graphs considered in this paper.

## 2.   Models and Results

The models used in this paper are based on the work of Albert and Barabasi [Barabási and Albert 99]. A *web graph* (see also [Cooper and Frieze 03]) can be defined as an ever-growing structure in which, at each step, new vertices, new edges, or a combination of these can be added. Decisions on what to add to the existing graph are made at random based on the values of a number of defining parameters. The existence of these parameters makes the model very general. For the purposes of this paper, to avoid cluttering the description of our results, we prefer to make a drastic simplification. We will consider graphs generated according to two rather extreme procedures derived from the general model. In each case the generation process is governed by a single integer parameter $m$. The second of these mimics the preferential attachment phenomenon. The first one, related to more traditional random graph models, is considered mainly for comparison.

**Uar graph process.** The initial graph $G_0^{R,m}$ is a single vertex $v_0$ with $m$ loops attached to it. For $t \geq 1$, let $G_{t-1}^{R,m}$ be the graph generated in the first $t - 1$ steps of this process. To define $G_t^{R,m}$, a new vertex $v_t$ is generated, and it is connected to $G_{t-1}^{R,m}$ through $m$ (undirected) edges. The neighbours of $v_t$ are chosen uniformly at random (uar) with replacement from $\{v_0, \ldots, v_{t-1}\}$.

**Preferential attachment graph process.** The initial graph $G_0^{C,m}$ is a single vertex $v_0$ with $m$ loops attached to it. For $t \geq 1$, graph $G_t^{C,m}$ is defined from $G_{t-1}^{C,m}$ by generating a new vertex $v_t$ and connecting it to $G_{t-1}^{C,m}$ through $m$ (undirected) edges. A vertex $u \in \{v_0, \ldots, v_{t-1}\}$ is connected to $v_t$ with probability $\frac{|\Gamma(u)|}{2mt}$ (where $\Gamma(u)$ is the (multi)set of vertices adjacent to $u$ in $G_{t-1}^{C,m}$).

| $m$ | $\alpha_{\text{lo}}^{R}$ | $\alpha_{\text{up}}^{R}$ | $\alpha_{\text{lo}}^{C}$ | $\alpha_{\text{up}}^{C}$ |
|---|---|---|---|---|
| 1 | 0.3678 | 0.5 | 0.2939 | 0.3333 |
| 2 | 0.1868 | 0.3714 | 0.0286 | 0.2342 |
| 3 | 0.1488 | 0.3054 | 0.0148 | 0.1777 |
| 4 | 0.1259 | 0.2634 | 0.0097 | 0.1422 |
| 5 | 0.1102 | 0.2335 | 0.0066 | 0.1178 |
| 6 | 0.0987 | 0.2110 | 0.0049 | 0.1000 |
| 7 | 0.0898 | 0.1932 | 0.0038 | 0.0865 |

**Table 1**. Numerical bounds for the minimum dominating set problem.

We will often refer to the uar graph process as a random graph process. The edges have an intrinsic direction that we ignore.

Both models are dynamic, with new vertices and edges continuously increasing the size of the graph. However, they represent two extreme cases. In the uar graph process the terminal vertex of an edge is chosen randomly from the set of available vertices, whereas in the preferential attachment model the terminal vertex of an edge is chosen with probability proportional to the current degree of the various vertices.

Let $M \in \{R, C\}$ and $m$ be a positive integer. The results in this paper can be summarised by saying that for each $h \geq 1$ there exist positive real constants $\alpha_{\text{lo}}^{M}$ and $\alpha_{\text{up}}^{M}$ (dependent on $M$, $m$, and $h$ but independent of $t$) with $\alpha_{\text{lo}}^{M} < \alpha_{\text{up}}^{M} < 1$ such that $\alpha_{\text{lo}}^{M} \cdot t \leq \gamma_h(G_t^{M,m}) \leq \alpha_{\text{up}}^{M} \cdot t$ a.a.s. For $h = 1$, the values of the aforementioned constants are reported in Table 1. Bounds for $h > 1$ are reported (and briefly commented on) in Section 6. In particular, the upper bounds are proved by analysing the size of the dominating set returned by a number of simple polynomial-time algorithms, whereas the lower bounds come from combinatorial arguments.

The proof of the algorithmic results are based on the fact that natural edge-exposure martingales can be defined on the graph processes under consideration [Cooper and Frieze 03]. More precisely, if $f(G)$ is any graph theoretic function (e.g., the size of the dominating set returned by a particular algorithm), the random process defined by setting $Z_0 = \mathbf{E}\left(f(G_t^{M,m})\right)$ and $Z_i$ (for $i \in \{1, \ldots, mt\}$) to be the expectation of $f(G_t^{M,m})$ conditioned on the "exposure" of the first $i$ edges in the graph process is a martingale. Notice that the space of all graphs which can be generated according to the given model $G_t^{M,m}$ is partitioned into classes (or $i$-blocks) containing all those graphs that coincide with respect to the first $i$ edge exposures.

In the forthcoming sections we will repeatedly use the following concentration result (for a proof see, for instance, [Alon et al. 92]).

**Theorem 2.1.** *Let* $c = Z_0, \ldots, Z_n$ *be a martingale with* $|Z_{i+1} - Z_i| \leq 1$ *for all* $i \in \{0, \ldots, n-1\}$. *Then* $\mathbf{Pr}(|Z_n - c| > \lambda\sqrt{n}) \leq 2\mathrm{e}^{-\lambda^2/2}$.

In all our applications $c = \mathbf{E}\ (f(G_t^{M,m}))$, $n = mt$, and $\lambda = O(\log t)$. In order to apply Theorem 2.1, one needs to prove that $|Z_{i+1} - Z_i| \leq 1$. Such an inequality follows from the smoothness of $f$ (i.e., $|f(G) - f(H)| \leq 1$ if $G$ and $H$ differ with respect to the presence of a single edge) and the ability to demonstrate the existence of a measure preserving bijection between $(i+1)$-blocks in a same $i$-block. This is obvious in the random graph process as edges are inserted independently. In the case of the preferential attachment process, it is convenient to identify each graph with the projection of a particular type of *configuration*, i.e., an ordered collection of $mt$ labelled pairs of points (see [Bender and Canfield 78]). Let $C_1$ and $C_2$ be two such configurations that are identical up to the $i$th pair. Suppose that the $(i+1)$-st pair is $\{a, b\}$ in $C_1$ and $\{a, c\}$ in $C_2$. If $C_1$ never uses point $b$ again, then the image of $C_1$ under the measure-preserving bijection will be a configuration $C'$ identical to $C_1$ except that pair $\{a, b\}$ is replaced by pair $\{a, c\}$. If $b$ is used in a following pair (say $\{d, b\}$) of $C_1$, then $C'$ will have $\{a, c\}$ instead of $\{a, b\}$ and $\{d, c\}$ instead of $\{d, b\}$, and so on. A similar construction is presented in [Cooper and Frieze 03].

Finally, we will need the following.

**Lemma 2.2.** *If* $Z_n$ *is a martingale and there exist constants* $c_1, c_2 > 0$ *such that* $\mu_n = \mathbf{E}\ (Z_n) \in [c_1 n, c_2 n]$, *then for each fixed integer* $j > 0$, *there exist positive constants* $K$ *and* $\epsilon$ *such that* $|(Z_n)^j - (\mu_n)^j| \leq K n^{j-\epsilon}$ *a.a.s.*

**Proof.** If $Z_n$ is a martingale, then by Theorem 2.1, $\mu_n - \lambda\sqrt{n} \leq Z_n \leq \mu_n + \lambda\sqrt{n}$ where without loss of generality we assume $\lambda = o(\sqrt{n})$. From this we also have, for each fixed integer $j > 0$,

$$(\mu_n)^j \left(1 - \frac{\lambda\sqrt{n}}{\mu_n}\right)^j \leq (Z_n)^j \leq (\mu_n)^j \left(1 + \frac{\lambda\sqrt{n}}{\mu_n}\right)^j.$$

The result now follows (provided that $K$ is chosen big enough), since the assumptions on $\lambda$ and $\mu_n$ entail that $(1 + \frac{\lambda\sqrt{n}}{\mu_n})^j$ is at most $1 + \frac{j^2\lambda\sqrt{n}}{\mu_n}$, whereas $(1 - \frac{\lambda\sqrt{n}}{\mu_n})^j$ is at least $1 - \frac{2j\lambda\sqrt{n}}{\mu_n}$. $\qquad\square$

## 3. Simplest Algorithm

The algorithm presented in this section is a very simple "first attempt" solution for the problem at hand. Although in many cases it does not lead to a very

| $m$ | $x$ |
|---|---|
| 1 | 0.5 |
| 2 | 0.382 |
| 3 | 0.3177 |
| 4 | 0.2755 |
| 5 | 0.2451 |
| 6 | 0.2219 |
| 7 | 0.2035 |

**Table 2**. Numerical values defined in Lemma 3.1.

small dominating set, it represents a natural benchmark for any more refined heuristic.

**Algorithm 1.** Before the first step of the algorithm, the graph consists of a single vertex $v_0$ and $\mathcal{S} = \{v_0\}$. At step $t$, if the newly generated vertex $v_t$ does not have any neighbours in $\mathcal{S}$ (i.e., $\Gamma(v_t) \cap \mathcal{S} = \emptyset$), then $v_t$ is added to $\mathcal{S}$.

In the forthcoming discussion $m$ is a fixed positive integer. Let $X_t$ denote the size of the dominating set $\mathcal{S}$ computed by Algorithm 1 before $v_t$ is added to the current graph, and let $\mu_t = \mathbf{E}(X_t)$. For graphs generated according to the $G_t^{R,m}$ model, the probability that $v_t$ misses the dominating set is $(1 - \frac{X_t}{t})^m$. Hence, we can write

$$\mu_{t+1} = \mu_t + \mathbf{E}\left[\left(1 - \tfrac{X_t}{t}\right)^m\right].$$

Let $x = x(m)$ be the unique solution of the equation $x = (1-x)^m$ in $(0,1)$. Table 2 gives the values of $x$ for the first few values of $m$.

**Lemma 3.1.** *For any $\frac{1}{2} < \rho < 1$ constant, there exists an absolute positive constant $C$ such that for all $t > 0$, $|\mu_t - xt| \le Ct^\rho$ a.a.s.*

**Proof.** We claim that the difference $|\mu_t - xt|$ satisfies a recurrence of the form

$$|\mu_{t+1} - x(t+1)| \le |\mu_t - xt| + O\left(\sqrt{\frac{\log t}{t}}\right).$$

This can be proved by induction on $t$. By definition $X_1 = 1$; hence, $|\mu_1 - x| = 1 - x$. We also have that $|\mu_{t+1} - x(t+1)| = |\mu_t - xt + \mathbf{E}[(1 - \frac{X_t}{t})^m] - (1-x)^m|$. The difference $\mathbf{E}[(1 - \frac{X_t}{t})^m] - (1-x)^m$ can be rewritten as $-\frac{m}{t}(\mu_t - xt) + \{\mathbf{E}[(1 - \frac{X_t}{t})^m] - 1 + m\frac{\mu_t}{t} - (1-x)^m + 1 - mx\}$. Hence,

$$|\mu_{t+1} - x(t+1)| = \left| \left( 1 - \frac{m}{t} \right) (\mu_t - xt) \right.$$

$$\left. + \left\{ \mathbf{E} \left[ \left( 1 - \frac{X_t}{t} \right)^m \right] - 1 + m\frac{\mu_t}{t} - (1-x)^m + 1 - mx \right\} \right|.$$

To complete the proof, notice that, by Lemma 2.2,

$$\mathbf{E} \left[ \left( 1 - \frac{X_t}{t} \right)^m \right] - 1 + m\frac{\mu_t}{t} = \left( 1 - \frac{\mu_t}{t} \right)^m - 1 + m\frac{\mu_t}{t} + O\left( \sqrt{\frac{\log t}{t}} \right)$$

and the function $f(z) = (1-z)^m - 1 + mz$ satisfies $|f(z_1) - f(z_2)| \le m|z_1 - z_2|$, for $z_1, z_2 \in [0,1]$. $\qquad \square$

The following theorem is a direct consequence of Lemma 3.1 and the concentration result mentioned in Section 2.

**Theorem 3.2.** $X_t \sim xt$ *a.a.s.*

## 4.   Improved Approximations in the Random Graph Process

Although Algorithm 1 is quite simple, it seems difficult to beat, as a glance at $\alpha_{\text{up}}^R$ in Tables 1 and 2 shows. This is especially true for $m = 1$ where no improvement could be obtained. For larger values of $m$, a better way of finding small dominating sets is obtained by occasionally allowing vertices to be dropped from $\mathcal{S}$. It is convenient to classify the vertices in the dominating set as *permanent* (set $\mathcal{P}$) and *replaceable* (set $\mathcal{R}$). Thus, $\mathcal{S} = \mathcal{P} \cup \mathcal{R}$. Let $P_t$ and $R_t$ denote the sizes of such sets at time $t$ (set $P_1 = 0$ and $R_1 = 1$).

**Algorithm 2.** Before the first step of the algorithm, the graph consists of a single vertex $v_0$ and $\mathcal{R} = \{v_0\}$. After $v_t$ is created and connected to $m$ neighbours, if $\Gamma(v_t) \cap \mathcal{P} \neq \emptyset$ then $v_t$ is moved to $V \setminus \mathcal{S}$. Otherwise, $v_t$ is added to $\mathcal{R}$ if $\Gamma(v_t) \cap \mathcal{R} = \emptyset$, and otherwise $v_t$ is added to $\mathcal{P}$ and all vertices in $\Gamma(v_t) \cap \mathcal{R}$ are moved to $V \setminus \mathcal{S}$.

The expectations $\pi_t = \mathbf{E}(P_t)$ and $\rho_t = \mathbf{E}(R_t)$ satisfy

$$\pi_{t+1} = \pi_t + \mathbf{E} \left[ \left( 1 - \frac{P_t}{t} \right)^m \right] - \mathbf{E} \left[ \left( 1 - \frac{P_t}{t} - \frac{R_t}{t} \right)^m \right],$$

$$\rho_{t+1} = \rho_t + \mathbf{E} \left[ \left( 1 - \frac{R_t}{t} - \frac{P_t}{t} \right)^m \right] - m\mathbf{E} \left[ \frac{R_t}{t} \left( 1 - \frac{P_t}{t} \right)^{m-1} \right].$$

Define $\alpha_{\mathrm{up}}^R$ as $p + r$, where $p = p(m)$ and $r = r(m)$ satisfy

$$r = \frac{(1-p)^m - p}{1 + m(1-p)^{m-1}},$$

$$p = (1-p)^m - (1-p-r)^m.$$

**Lemma 4.1.** *For any $\frac{1}{2} < \rho < 1$ constant, there exist absolute positive constants $C_1$ and $C_2$ such that for all $t > 0$, $|\pi_t - pt| \le C_1 t^\rho$ and $|\rho_t - rt| \le C_2 t^\rho$ a.a.s.*

**Proof.** The proof is, essentially, a generalisation of that of Lemma 3.1. We present the argument in some details for $\pi_t$. At the inductive step,

$$|\pi_{t+1} - p(t+1)| = \left| \pi_t - pt + \mathbf{E}\left[ \left(1 - \frac{P_t}{t}\right)^m \right] - (1-p)^m \right.$$
$$\left. - \left\{ \mathbf{E}\left[ \left(1 - \frac{P_t}{t} + \frac{R_t}{t}\right)^m \right] - (1-p-r)^m \right\} \right|.$$

The proof is completed by decomposing the differences $\mathbf{E}\left[(1-\frac{P_t}{t})^m\right] - (1-p)^m$ and $\mathbf{E}\left[(1 - \frac{P_t}{t} - \frac{R_t}{t})^m\right] - (1-p-r)^m$ into parts that are proportional to either $P_t - \pi_t$ or $\pi_t - pt$.

The result about $\rho_t$ is proved similarly after noticing that $r$ satisfies

$$r = (1-p-r)^m - mr(1-p)^{m-1}.$$

$\square$

Let $X_t^2$ denote the size of the dominating set returned by Algorithm 2 when run on a random graph process until time $t$.

**Theorem 4.2.** $X_t^2 \sim (p+r)t$ a.a.s.

**Proof.** Theorem 2.1 implies that the sum $p+r$ is a.a.s. very close to $\frac{|\mathcal{S}|}{t} = \frac{P_t}{t} + \frac{R_t}{t}$. The result follows. $\square$

The values of $p + r$ for $m \le 7$ are reported in the column labelled $\alpha_{\mathrm{up}}^R$ of Table 1.

## 5.  Improved Approximations in the Preferential Attachment Process

Algorithm 1, described in Section 3, can be analysed in the preferential attachment model as well. The expected change in the variable $X_t$ can be computed

by keeping track of the total degree of the dominating set, $D_t$. In particular, the following relationships hold:

$$\mathbf{E}\left(X_{t+1}\right) \;=\; \mathbf{E}\left(X_t\right) + \mathbf{E}\left[\left(1 - \tfrac{D_t}{2mt}\right)^m\right],$$

$$\mathbf{E}\left(D_{t+1}\right) \;=\; \left(1 + \tfrac{1}{2t}\right)\mathbf{E}\left(D_t\right) + m\mathbf{E}\left[\left(1 - \tfrac{D_t}{2mt}\right)^m\right].$$

Not surprisingly, an analysis similar to the one described in the previous sections implies that such an algorithm returns dominating sets of size $xt$ in $G_t^{C,m}$. However, in the model based on preferential attachment we can improve on this by pushing high degree vertices in the dominating set. The natural way to accomplish this would be, for any newly generated uncovered vertex, to select a neighbour of maximum degree and add it to $\mathcal{S}$. Unfortunately such algorithm is not easy to analyse because in the graph processes that we consider there may be few vertices of rather large degree (this is a consequence of the power law distribution of vertex degrees [Barabási and Albert 99, Cooper 02]). However a draconian version of this heuristic can be analysed. The following algorithm takes as input an additional integer parameter $k > 0$.

**Algorithm 3.** Before the first step of the algorithm, the graph consists of a single isolated vertex $v_0$ and $\mathcal{S} = \{v_0\}$. After $v_t$ is created and connected to $m$ neighbours, let $Z$ be the set of all neighbours of $v_t$ in $V \setminus \mathcal{S}$ of degree $km + 1$. If $Z \neq \emptyset$, then all vertices in $Z$ are added to $\mathcal{S}$. Otherwise, if $v_t$ is not dominated by some element of $\mathcal{S}$, then a vertex of maximum degree in $\Gamma(v_t)$ is added to the dominating set.

Notice that after each $v_t$ is generated and connected to $G_{t-1}^{C,m}$, all vertices with degree that has become larger than $km$ are moved inside $\mathcal{S}$. The analysis of the evolution of $|\mathcal{S}|$ is based again on the definition of a random process that describes the algorithm dynamics and on the proof that such a process behaves in a predictable way for large $t$.

Let $n = (k-1)m + 1$. For each $i \in \{0, \ldots, n-1\}$ and $t > 0$, define $Y_t^i = |V_{m+i} \setminus \mathcal{S}|$ in $G_{t-1}^{C,m}$ ($V_i$ is the set of vertices of degree $i$) before $v_t$ is added to the graph. Let $Y_t^n$ denote the total degree inside $\mathcal{S}$ (i.e., $Y_t^n = \sum_{v \in \mathcal{S}} |\Gamma(v)|$) and $X_t$ the size of the dominating set before $v_t$ is added to the graph. The state of the system, at each step $t$, is modelled by the (random) vector $(Y_t^0, \ldots, Y_t^n, X_t)$. Notice that, for each $t > 0$, the variation in each of the variables is at most $m$.

Also, $Y_t^n + \sum_{i=0}^{(k-1)m}(m+i)Y_t^i = 2mt$, and at each step $t \geq 1$, when $v_t$ is created, the probability that it hits a vertex of degree $m + i$, for $i \in \{0, \ldots, (k-1)m\}$ (respectively, the dominating set for $i = n$) in any one of the $m$

trials available to it is approximately (omitting $o(1)$ factors, for $t$ large) equal to $P_i = \frac{((m+i-1)(1-\delta_{i,n})+1)Y_t^i}{2mt}$ (where $\delta_{i,n} = 1$ if $i = n$ and zero otherwise).

For $d \in \{0, \ldots, n-1\}$ (respectively, $d = n$), let $E_d$ denote the event "$v_t$ missed $\mathcal{S}$ and the maximum degree in $\Gamma(v_t)$ is $m + d$" (respectively, "$v_t$ did not miss $\mathcal{S}$"). The expected change to $Y_t^i$, conditioned to the process history up to time $t$, can be computed by further conditioning on the family of events $(E_d)_{d \in \{0, \ldots, n\}}$. We can write such a quantity as

$$\sum_{d=0}^{n} \mathbf{E}\left(Y_{t+1}^i - Y_t^i \mid E_d\right) \mathbf{Pr}(E_d).$$

For $t$ large, the probability in the expression above is approximately $\chi_d = (S_0^d)^m - (S_0^{d-1})^m$ (the notation $S_a^b$ stands for $P_a + \ldots + P_b$, with $S_a^b \equiv 0$ if $a > b$). Furthermore, we can approximate $\mathbf{E}\left(Y_{t+1}^i - Y_t^i \mid E_d\right)$ by the expression

$$\sum C(h_0, \ldots, h_d, 0, \ldots, 0) \frac{m!}{h_0! \, h_1! \, \ldots \, h_d!} \prod_{i=0}^{d} P_i^{h_i} \frac{1}{\chi_d},$$

where $C$ has $n + 1$ arguments, the righmost $n - d$ of which are zero, the sum is over all possible ordered tuples of values $h_0, \ldots, h_d$ such that $\sum_{i=0}^{d} h_i = m$ and $h_d > 0$, and $C(h_0, \ldots, h_d, 0, \ldots, 0)$ contains

- a term for the addition of $v_t$ to $G_{t-1}^{C,m}$,

- a term $\phi_{i,d}$ for the change to $Y_t^i$ due to the handling of the chosen vertex of maximum degree $m+d$ in $\Gamma(v_t)$ (for $d = n$ this is just one of the vertices hitting $\mathcal{S}$), and

- a term $\psi_{i,s}$ for the change to $Y_t^i$ due to the handling of a vertex accounted for by $Y_t^s$ in $\Gamma(v_t)$, for $s \leq d$.

The first of these is simply $\delta_{i,0}$. We also have

$$\phi_{i,d} = \delta_{d,n} \times \delta_{i,d} + (1 - \delta_{d,n}) \times \{(m + d + 1)\delta_{i,n} - \delta_{i,d}\}$$

(if $d = n$ (i.e., if $v_t$ hits the dominating set), then one is added to $Y_t^n$, and otherwise $Y_t^d$ is decreased and $m + d + 1$ units are added to $Y_t^n$), and

$$\psi_{i,s} = \delta_{s,n} \times \delta_{i,s} + (1 - \delta_{s,n}) \times \{((m + s)\delta_{s,n-1} + 1)\delta_{i,s+1} - \delta_{i,s}\}$$

(if $s = n$, then $Y_t^n$ is increased by one, if $s = n - 1$ the newly created vertex of degree $n$ must be added to $\mathcal{S}$, and finally, in any other case the vertex that has been hit is moved from $V_{m+s}$ to $V_{m+s+1}$). Therefore,

$$C(h_0, \ldots, h_d, 0, \ldots, 0) = \delta_{i,0} + \phi_{i,d} + (h_d - 1)\psi_{i,d} + \sum_{s=0}^{d-1} h_s \psi_{i,s}.$$

Doing all the sums, for each $i \in \{0, \ldots, n\}$, the expected change to $Y_t^i$ is approximately equal to

$$\delta_{i,0} + \sum_{d=0}^{n} \left\{ \chi_d \phi_{i,d} + \psi_{i,d} \left[ m P_d (S_0^d)^{m-1} - \chi_d \right] + \frac{m[\chi_d - P_d(S_0^d)^{m-1}]}{S_0^{d-1}} \sum_{s=0}^{d-1} \psi_{i,s} P_s \right\}.$$

Also, $\mathbf{E}\,(X_{t+1} - X_t)$ is approximately equal to

$$\frac{km\mathbf{E}\,(Y_t^{n-1})}{2t} + \mathbf{E}\left[ \left( 1 - \frac{Y_t^n}{2mt} - \frac{kY_t^{n-1}}{2t} \right)^m \right].$$

For each positive integer $m$, let $x$ be the value $kmy^{n-1} + (1 - \frac{y^n}{2m} - \frac{ky^{n-1}}{2})^m$, where $y^i$, for $i \in \{0, \ldots, n\}$, satisfy

$$y^i = \left[ \mathbf{E}\,(Y_{t+1}^i) - \mathbf{E}\,(Y_t^i) \right] \big|_{Y^i = y^i t, i \in \{0, \ldots, n\}}$$

and the parameter $k$ can be chosen arbitrarily (larger values of $k$ give slightly smaller values of $x$). The main result about Algorithm 3 is the following theorem, the proof of which is a consequence of the argument above and Theorem 2.1.

**Theorem 5.1.** $X_t \sim xt$, a.a.s.

For each $m \leq 7$, values of $k \leq 20$ give the values of $x$ reported in Table 1 in the column labelled $\alpha_{\mathrm{up}}^C$.

## 6.  Generalisations

The algorithms presented in the previous sections generalise naturally to larger values of $h$. We present here the generalisation of Algorithm 2 for finding an $h$-dominating set in the random graph process and of Algorithm 3 for the preferential attachment process. We then briefly sketch the analysis of the first heuristic and comment on the numerical results presented in Table 3.

**Algorithm 4.** A vertex in the $h$-dominating set can be either *permanent* (set $\mathcal{P}$), meaning that it will never be removed from the set, or *i-removable* (set $\mathcal{R}_i$), meaning that it is dominated $i \in \{0, \ldots, h-1\}$ times by other permanent vertices.

Before the first step of the algorithm, the graph consists of a single vertex $v_0$ and $\mathcal{R}_0 = \{v_0\}$. After $v_t$ is created and connected to $m$ neighbours, $v_t$ is added to $\mathcal{R}_0$ if $\Gamma(v_t) \cap (\mathcal{P} \cup \bigcup_i \mathcal{R}_i) = \emptyset$; otherwise, if $\Gamma(v_t) \cap \mathcal{P} = \emptyset$, $v_t$ is added to

$\mathcal{P}$, any vertex in $\Gamma(v_t) \cap \mathcal{R}_i$ for $i < h-1$ is moved to $\mathcal{R}_{i+1}$, and any vertex in $\Gamma(v_t) \cap \mathcal{R}_{h-1}$ is moved to $V \setminus \mathcal{S}$.

The process in Algorithm 4 can be modelled by $h$ sequences of random variables: $R_t^i$ for $i \in \{0, \ldots, h-1\}$ with $R_t^i = |\mathcal{R}_i|$ in the graph $G_{t-1}^{R,m}$ just before $v_t$ is added to it, and $P_t = |\mathcal{P}|$. The expected changes in these variables satisfy

$$\mathbf{E}\left(P_{t+1}\right) = \mathbf{E}\left(P_t\right) + \mathbf{E}\left[\left(1 - \tfrac{P_t}{t}\right)^m\right] - \mathbf{E}\left[\left(1 - \tfrac{P_t}{t} - \sum_i \tfrac{R_t^i}{t}\right)^m\right],$$

$$\mathbf{E}\left(R_{t+1}^i\right) = \mathbf{E}\left(R_t^i\right) + \mathbf{E}\left[\left(1 - \tfrac{P_t}{t} - \sum_i \tfrac{R_t^i}{t}\right)^m\right]\delta_{i,0}$$
$$+ m\mathbf{E}\left[\left(\tfrac{R_t^{i-1}}{t}(1-\delta_{i,0}) - \tfrac{R_t^i}{t}\right)\left(1 - \tfrac{P_t}{t}\right)^{m-1}\right].$$

Therefore, we have $P_t \sim pt$ where $p$ satisfies

$$p = (1-p)^m - \left\{1 - \left[(1-p)^m - p\right]\left[1 - \left(\tfrac{m(1-p)^{m-1}}{1+m(1-p)^{m-1}}\right)^h\right] - p\right\}^m$$

and $R_t^i \sim r^i t$ where $r^0 = \frac{(1-p)^m - p}{1+m(1-p)^{m-1}}$ and $r^i = \frac{m(1-p)^{m-1}}{1+m(1-p)^{m-1}} r^{i-1}$. The values reported in Table 3 below are given by $p + \sum_{i=0}^{h-1} r^i$.

**Algorithm 5.** Before the first step of the algorithm, the graph consists of a single vertex $v_0$ and $\mathcal{S} = \{v_0\}$. After $v_t$ is created and connected to $m$ neighbours,

| $m$ | $h=2$ | $h=3$ | $h=4$ | $h=5$ |
|---|---|---|---|---|
| 2 | 0.4484 | | | |
| 3 | 0.368 | 0.3836 | | |
| 4 | 0.3162 | 0.3307 | 0.332 | |
| 5 | 0.2795 | 0.2929 | 0.2944 | 0.2931 |
| 6 | 0.2517 | 0.2641 | 0.2658 | 0.2647 |
| 7 | 0.2298 | 0.2413 | 0.2431 | 0.2422 |

Algorithm 4

| $m$ | $h=2$ | $h=3$ | $h=4$ | $h=5$ |
|---|---|---|---|---|
| 2 | 0.5 | | | |
| 3 | 0.4075 | 0.6 | | |
| 4 | 0.3359 | 0.4852 | 0.6663 | |
| 5 | 0.282 | 0.4073 | 0.5423 | 0.7036 |
| 6 | 0.2428 | 0.3523 | 0.4613 | 0.5862 |
| 7 | 0.2132 | 0.3066 | 0.4035 | 0.5056 |

Algorithm 5

**Table 3**. Upper bounds on $\gamma_h/t$, for $h > 1$.

the set $Z$ of all newly generated vertices of degree more than $km$ are added to $\mathcal{S}$. If $v_t$ is dominated $h - x$ times by elements of $\mathcal{S}$, then the $x$ vertices of highest degree in $\Gamma(v_t) \setminus Z$ are added to the dominating set.

## 7. Tightness of the Algorithmic Results

An interesting and simple argument can be used to complement the algorithmic results presented in the previous sections. The argument is based on the following result.

**Lemma 7.1.** *Let $\mathcal{S}$ be an $h$-dominating set in a graph $G = (V, E)$. If the total degree of the vertices in $\mathcal{S}$ is at least $d$, then $|\mathcal{S}| \geq \sum_{i > i_0} |V_i|$, where $i_0$ is the largest index $i$ for which $\sum_{j > i} j|V_j| \geq d$.*

**Proof.** Let $i_0$ be defined as in the statement of the result. If $\Delta = \max_{v \in V} |\Gamma(v)|$, then the set $V_{i_0} \cup \ldots \cup V_\Delta$ is the smallest set of vertices in $G$ with total degree at least $d$ (any other vertex in $G$ would have smaller degree, and therefore it would contribute less to the total degree).  $\square$

| $m$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ |
|-----|---------|---------|---------|---------|
| 2 | 0.1317 | | | |
| 3 | 0.1001 | 0.178 | | |
| 4 | 0.0687 | 0.1342 | 0.1678 | |
| 5 | 0.0649 | 0.0935 | 0.1346 | 0.1938 |
| 6 | 0.0535 | 0.0849 | 0.1155 | 0.1573 |
| 7 | 0.0406 | 0.0692 | 0.1033 | 0.1349 |

uar process

| $m$ | $h = 2$ | $h = 3$ | $h = 4$ | $h = 5$ |
|-----|---------|---------|---------|---------|
| 2 | 0.0545 | | | |
| 3 | 0.0316 | 0.0351 | | |
| 4 | 0.023 | 0.0333 | 0.0246 | |
| 5 | 0.0183 | 0.0284 | 0.0302 | 0.0183 |
| 6 | 0.0141 | 0.0233 | 0.0299 | 0.0269 |
| 7 | 0.0113 | 0.0203 | 0.0271 | 0.0283 |

preferential attachment process

**Table 4**. Lower bounds on $\gamma_h/t$, for $h > 1$.

The total degree of an $h$-dominating set in $G_t^{M,m}$ must be at least $h(t-|\mathcal{S}|) \geq ht(1 - \alpha_{up}^M)$. Hence, a lower bound on the size of any dominating set in a web graph is obtained by using information on the proportional degree sequence.

Bollobas et al. [Bollobás et al. 01] (see also Cooper [Cooper 02]) proved that for a random graph built using a model equivalent to our preferential attachment process, $|V_i| = tn_i + O(\sqrt{t}\log t)$ a.a.s. for any $i \geq m$, where

$$n_i = \frac{2m(m+1)}{i(i+1)(i+2)}.$$

In the same paper (p. 288) it is possible to find a result about $G_t^{R,m}$. In the uar graph process, for any $i \geq m$,

$$n_i = \frac{1}{m+1}\left(\frac{m}{m+1}\right)^{i-m}.$$

Again it is possible to prove that $|V_i|$ is concentrated around $n_i t$.

For $m \geq 2$, the lower bounds given for $G_t^{C,m}$ in Table 1 are obtained using Lemma 7.1 and the approximations above for $|V_i|$ in each case. Bounds for $h > 1$ are in Table 4.

For the uar graph process $G_t^{R,m}$, we can produce better a.a.s. lower bounds using the Markov inequality. The method is explained in Section 7.1. For $m \geq 2$, these are the lower bounds given for $G_t^{R,m}$ in Table 1.

## 7.1. Existential Lower Bounds for the Uar Process

For simplicity we present the argument for $h = 1$ (i.e., dominating sets) only.

**Lemma 7.2.** *There exists a function $f(x,y) : \mathbb{R} \to \mathbb{R}$ such that, for each positive integer $m$, there exists a positive real number $d = d(m)$, defined by $f(m,d) = 1$, such that a.a.s. there is no dominating set in $G_t^{R,m}$ of size at most $dt$.*

*The values of $d$ for the first few values of $m$ are given in the following table:*

| $m$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $d(m)$ | 0.26894 | 0.18683 | 0.14881 | 0.12591 | 0.11027 | 0.09876 | 0.08986 |

**Proof.** In the proof we assume that the $t$ vertices in $G_t^{R,m}$ are labelled $1, 2, \ldots, t$. The argument is as follows:

(i) Let $\mathcal{S}^* = [s] = \{1, \ldots, s\}$, and let $|\mathcal{S}| = s$. Then,

$$\mathbf{Pr}(\mathcal{S} \text{ dom. } \overline{\mathcal{S}}) \leq \mathbf{Pr}(\mathcal{S}^* \text{ dom. } \overline{\mathcal{S}^*}),$$

where dom. is an abbreviation of "dominates."

(ii) For $s \leq dt$, the expected number of dominating sets of size $s$ is at most

$$\binom{t}{s} \mathbf{Pr}(\mathcal{S}^* \text{ dom. } \overline{\mathcal{S}^*}) \to 0.$$

Thus, Marvov inequality implies that a.a.s. there is no dominating set $\mathcal{S}$ of size at most $dt$.

**Proof of (i).** We regard the $m$ out-edges $(e_1, \ldots, e_m)$ of any vertex as labelled and ordered so that all $t$ vertex graphs $G$ have the same probability in the uar graph process; i.e., $\mathbf{Pr}(G) = 1/((t-1)!)^m$. (Vertex $v_0$ has $m$ loops.)

Given a partition $(\mathcal{S}, \overline{\mathcal{S}})$ of $[t]$ with $|\mathcal{S}| = s$, let $u$ be the smallest element $u = v_i \in \overline{\mathcal{S}}$ such that the next vertex is $v = v_{i+1} \in \mathcal{S}$ in the natural ordering $1, 2, \ldots, t$. If $u$ does not exist, then $\mathcal{S} = \mathcal{S}^*$.

We swap $u$ and $v$ to get the partition $(\mathcal{S}', \overline{\mathcal{S}'})$ where $\mathcal{S}' = \mathcal{S} - v + u$ and $\overline{\mathcal{S}'} = \overline{\mathcal{S}} - u + v$. We prove that

$$\mathbf{Pr}(\mathcal{S} \text{ dom. } \overline{\mathcal{S}}) \leq \mathbf{Pr}(\mathcal{S}' \text{ dom. } \overline{\mathcal{S}'}).$$

We do this by showing that for every graph $G \in \{\mathcal{S} \text{ dom. } \overline{\mathcal{S}}\}$ for which the swap destroys the dominance property, i.e., $G \notin \{\mathcal{S}' \text{ dom. } \overline{\mathcal{S}'}\}$ ($G$ is good but goes bad on swapping), there is a unique bad graph $H = H(G)$ ($H \notin \{\mathcal{S} \text{ dom. } \overline{\mathcal{S}}\}$) that becomes good on swapping ($H \in \{\mathcal{S}' \text{ dom. } \overline{\mathcal{S}'}\}$). Thus, the set of good graphs is of nondecreasing size (and hence probability) on swapping.

More formally, let $\boldsymbol{G}(t) = G_t^{R,m}$ denote the space of uar graph processes after the $t$th vertex has been added. Let $\mathcal{S}$ be the distinguished set. Let $\mathsf{GG}$ denote the subset of $\boldsymbol{G}$ that is good for both $\mathcal{S}$ and $\mathcal{S}'$ (i.e., $\mathsf{GG} = \{G : \mathcal{S} \text{ dom. } \overline{\mathcal{S}}\} \cap \{G : \mathcal{S}' \text{ dom. } \overline{\mathcal{S}'}\}$) and similarly $\mathsf{GB}, \mathsf{BG}$, and $\mathsf{GB}$, so that $\mathsf{GB}$ is the set of graphs that are good for $S$ but not $\mathcal{S}'$, etc. Thus, $\mathsf{GG}, \mathsf{GB}, \mathsf{BG}$, and $\mathsf{GB}$ are a partition of $\boldsymbol{G}$. We define a one-to-one map $\phi : \boldsymbol{G} \to \boldsymbol{G}$ such that $\phi(\mathsf{GB}) \subseteq \mathsf{BG}$. It follows that

$$
\begin{aligned}
|\{G : \mathcal{S} \text{ dom. } \overline{\mathcal{S}}\}| &= |\mathsf{GG}| + |\mathsf{GB}| \\
&= |\mathsf{GG}| + |\phi(\mathsf{GB})| \\
&\leq |\mathsf{GG}| + |\mathsf{BG}| = |\{G : \mathcal{S}' \text{ dom. } \overline{\mathcal{S}'}\}|.
\end{aligned}
$$

With some abuse of notation, we associate $\text{out}(x)$, the sequence of out-edges of vertex $x$, with the multi-set of terminal vertices, and similarly for $\text{in}(x)$ (the sequence of in-edges of $x$ ordered naturally) with their initial vertices.

*Construction of $\phi$.* Let $[t] = (1, 2, \ldots, u - 1, u, v, v + 1, \ldots, t)$ where $u, v$ are the swap pair. Consider the graph $G$. Let $i(x) = \text{in}(x) \cap [v + 1, \ldots, n]$ and $o(x) = \text{out}(x) \cap [1, \ldots, v]$. Let $i'(u) = i(v)$ and $i'(v) = i(u)$.

*Case that there is no edge* $(v, u)$. Let $o'(u) = o(v)$ and $o'(v) = o(u)$. Construct $\phi(G) = H(G)$ from $G$ by replacing $i(x), o(x)$ by $i'(x), o'(x)$, $x = u, v$. The graph $H$ is uniquely associated with $G$, as we can reconstruct $G$ from $H$.

*Case that there are edges* $(v, u)$. Let the $(v, u)$ edges be $\{e_k, k \in K\}$ where $K$ indexes their occurrences in $(e_1, \ldots, e_m)$. Let $\{f_k\}$ be the out-edges of $u$ with the same index set. Let $o'(u) = o(v) - \{e_k\} + \{f_k\}$, and let $o'(v) = o(u) + \{e_k\} - \{f_k\}$. Thus, $v$ keeps its out-edges to $u$, and $u$ keeps its out-edges with the same labels.

If there is no $(v, u)$ edge and $G \in \mathsf{GB}$, then $H$ was bad but becomes good (i.e., $H \in \mathsf{BG}$). For, in $H$, $u$ (now in the dominating set) has the $v$ edges from $G$, and $v$ (no longer in the dominating set) got the $u$ edges from $G$.

If there is a $(v, u)$ edge and $G \in \mathsf{GB}$, then $u$ dominates $v$ after the swap, just as $v$ dominated $u$ before. The in-edges have been swapped over to preserve any dominance that $v$ made with them (now made by $u$). The edge set $\{f_k\}$ was not necessary for $G$ to be good, as $v$ dominated $u$ and $u$ did not dominate anybody. If $v$ dominated anybody else with $o(v) - \{e_k\}$, then $u$ still does it with $o'(u) = o(v) - \{e_k\} + \{f_k\}$. Thus, $H$ is good. $\qquad \square$

**Proof of (ii).** Let $S^* = [s]$. We have that

$$\mathbf{Pr}(\exists \text{ dominating set } S, |S| = s) \le \binom{t}{s}\mathbf{Pr}(S^* \text{ dom. } \overline{S^*}).$$

Let $s = dt$. Then $\binom{t}{s} \le \left(\frac{1}{d^d(1-d)^{(1-d)}}\right)^t$, and

$$
\begin{aligned}
\mathbf{Pr}(S^* \text{ dom. } \overline{S^*}) &= \prod_{v=s}^{t-1} \left(1 - \left(1 - \tfrac{s}{v}\right)^m\right) \\
&= \exp\left(\sum_s^{t-1} \log\left(1 - \left(1 - \tfrac{s}{v}\right)^m\right)\right) \\
&= \exp\left((1 + o(1)) \int_{s+1}^t \log\left(1 - \left(1 - \tfrac{s}{v}\right)^m\right) dv\right) \\
&= \exp\left((1 + o(1)) \, dt \int_d^1 x^{-2} \log\left(1 - (1 - x)^m\right) dx\right).
\end{aligned}
$$

To obtain $(m, d)$ as given, we compute the largest $d$ such that

$$\left(\frac{1}{d^d(1-d)^{(1-d)}}\right) \exp\left(d \int_d^1 x^{-2} \log\left(1 - (1 - x)^m\right) dx\right) < 1.$$

$\qquad \square$

### 7.2. Lower bounds for the preferential attachment process.

**Lemma 7.3.** *There exists a function* $f(x, y) : \mathbb{R} \to \mathbb{R}$ *such that, for each positive integer* $m$, *there exists a positive real number* $d = d(m)$, *defined by* $f(m, d) = 1$, *such that a.a.s. there is no dominating set in* $G_t^{C,m}$ *of size at most* $dt$.

*The values of d for the first few values of m are given in the following table:*

| $m$    | 1       | 2       | 3       | 4       | 5       | 6       | 7       |
|--------|---------|---------|---------|---------|---------|---------|---------|
| $d(m)$ | 0.17379 | 0.10098 | 0.07099 | 0.05443 | 0.04391 | 0.03663 | 0.03130 |

**Proof.** The argument is the same as the one used for the uar graph process; i.e., we prove that for any set $S$ of size $s$, $\mathbf{Pr}(S \text{ dom. } \overline{S}) \leq \mathbf{Pr}(S^* \text{ dom. } \overline{S^*})$ where $S^* = [s] = \{1, \ldots, s\}$. As before, we do this by moving the vertices of the dominating set $S$ downward. Now, however, in the case where there is a $(v, u)$ edge, swapping edges between the vertices $u$ and $v$ can alter the probability of the graphs, i.e., $\mathbf{Pr}(G) \neq \mathbf{Pr}(\phi(G))$, and some care is needed to overcome this problem. In order to get a precise value for $\mathbf{Pr}(G)$, we work in the scale-free model of Bollobas and Riordan. A full description of the model can be found in [Bollobás et al. 01]. The scale-free model differs slightly from our preferential attachment model in that loops are allowed. We consider the case $m = 1$, as the general case is easily derived from this. We have

$$\mathbf{Pr}(v_t \text{ chooses } u) = \begin{cases} \frac{\deg(u, t-1)}{2t-1} & u \neq v_t \\[2mm] \frac{1}{2t-1} & \text{otherwise} \end{cases}$$

where $\deg(v, t)$ is the degree of vertex $v$ at the end of step $t$. For fixed $t = n$, the scale-free model is formulated as a configuration model on the edge endpoints, labelled $1, \ldots, 2n$. A pairing $\mathcal{P}$ of the endpoints describes the set of edges, and an edge is completed at step $k$ in the linear order $1, \ldots, 2n$ if the larger point in the pairing is $k$. Given a point $k$, the left partial pairing $\mathcal{L}$ is the partial pairing on $1, \ldots, k$, and the right partial pairing $\mathcal{R}$ is the partial pairing on $k+1, \ldots, 2n$, both of which may have some completed edges. A further matching of unpaired points between the partial pairings completes $\mathcal{P}$. See [Bollobás et al. 01] for more details.

As usual, let $i'(x), x = u, v$, be the (modified) in-degree (or associated set of in-edges) of vertex $x$. Thus, $i'(x)$ is the in-degree except that $i'(u)$ does not include the edge $(v, u)$ if it exists. Let $i'(u) = a_1$ and $i'(v) = a_2$. Thus, the in-degree of $u$ is $a_1 + 1$ if $(v, u)$ exists and $a_1$ if not. Let $G(a_1, a_2)$ be a graph with these in-degrees at $u$ and $v$, respectively. To save on notation, we let $a_i$ be the multi-set of distinguished in-edges as well as the set size. Writing $G(a_2, a_1)$ means that the in-edges of $u$ and $v$ other than $(v, u)$ have been swapped, and $\phi(G(a_1, a_2))$ means that the out-edges other than $(v, u)$ and its corresponding labelled edge(s) at $u$ have also been swapped. We use $a_1 + a_2$ for set union, preserving multiplicity. The meaning will always be clear from the context.

Let $u = x_k$ be the $k$th vertex, and so $v = x_{k+1}$. Suppose that $x_k$, the $k$th right endpoint, has label $2k + s + a_1 + 1$ and the $(k + 1)$th right endpoint has label $2(k + 1) + s + a_1 + a_2$ in the linear order. If there is no edge $(v, u)$, let the $(k - 1)$th right endpoint have label $2(k - 1) + s + 2$, and if there is a $(v, u)$ edge, let it have label $2(k - 1) + s + 1$. Thus, vertex $u$ has in-degree $a_1$ in the first case and $a_1 + 1$ in the second case. In either case vertex $v$ has in-degree $a_2$, and there are $s$ unpaired labels in the left partial pairing of vertex $x_{k-1}$.

For $t$ even, let $\Phi(t)$ denote the number of pairings of $t$ points, and thus $\Phi(t) = t!/((t/2)!2^{t/2})$. We can write down the probability of the graph $G(a_1, a_2)$ in the two cases as follows: $\mathbf{Pr}(G(a_1, a_2)) = F_L F_{R,L}/\Phi(2n)$ where $F_L$ is the number of left partial pairings of $2(k + 1) + s + a_1 + a_2$ labels and $F_{R,L}$ is the number of completions with the right partial pairing on $2n - (2(k + 1) + s + a_1 + a_2)$ labels. In either case

$$F_{R,L} = \binom{2n - (2(k + 1) + s + a_1 + a_2)}{s + a_1 + a_2}(s+a_1+a_2)!\Phi(2(n-((k+1)+s+a_1+a_2))).$$
$$(7.1)$$

If there is no $(v, u)$ edge, then

$$F_L = (2k + s - 1)(2k + s - 2)(2k + s - 3)\binom{2k + s - 4}{s}\Phi(2(k - 2)), \quad (7.2)$$

and if there is a $(v, u)$ edge, then

$$F_L = (2k + s - 2)(2k + s - 3)(a_1 + 1)\binom{2k + s - 4}{s}\Phi(2(k - 2)). \quad (7.3)$$

The first two terms on the right are for $x_{k-1}, x_k = u$, completing a pairing, and the third is for $x_{k+1} = v$. We then choose $s$ points for the left-right pairing and pair the remaining points on the left.

From the above expressions we see that if there is no edge $(v, u)$ in $G$, then expressions (7.1) and (7.2) are unchanged on swapping and $\mathbf{Pr}(G) = \mathbf{Pr}(\phi(G))$. If there is an edge $(v, u)$, then

$$\begin{aligned} \mathbf{Pr}(G(a_1, a_2)) &= c(a_1 + 1), & (7.4) \\ \mathbf{Pr}(\phi(G(a_1, a_2))) &= \mathbf{Pr}(G(a_2, a_1)) = c(a_2 + 1), & (7.5) \end{aligned}$$

where $c$ is given above. We see that if $a_2 < a_1$ we cannot use the uar proof without modification.

Recall that GB is the set of graphs that are good for $S$ ($S$ dominates $\overline{S}$) but bad for $S' = S + u - v$ ($S'$ does not dominate $\overline{S'}$). Our approach is to partition GB into sets $C$ that have the property $\mathbf{Pr}(C) \leq \mathbf{Pr}(\phi(C))$. In preparation for

this we let $\sigma = \lceil n/\log n \rceil$ and let

$$
\begin{aligned}
B_I &= \{G : \exists i \geq \sigma \text{ such that } G \text{ has at least 2 edges } (x_{i+1}, x_i)\}, \\
B_{II} &= \left\{ \begin{array}{l} G : \exists j > i \geq \sigma \text{ such that } G \text{ has an edge } (x_{i+1}, x_i) \\ \text{and at least 2 edges } (x_j, x_{i+1}) \end{array} \right\}.
\end{aligned}
$$

Based on the almost sure maximum degree of vertex $x$ at step $n$ of $(n/x)^{1/2} c \log^3 n$ (see, e.g., [Bollobás et al. 01] and [Cooper 02]) for all $x \geq \sigma$, we see that $\mathbf{Pr}(B_I) = O(\log n/n)$ and that $\mathbf{Pr}(B_{II}) = O(\log^6 n/n)$. To take advantage of this, we note that

$$
\mathbf{Pr}(S \text{ dom. } \overline{S}) \leq \mathbf{Pr}(S \cup [\sigma] \text{ dom. } \overline{S} \setminus [\sigma]).
$$

Thus, we need only consider swaps on $[\sigma + 1, \ldots, n]$ as $[\sigma]$ is always in the dominating set.

Let

$$
A_u = \{G \in \mathsf{GB} : \exists (v, u) \text{ edge in } G\} \cap \overline{(B_I \cup B_{II})},
$$

let

$$
B_u = \mathsf{GB} \cap (B_I \cup B_{II}),
$$

and let

$$
C_u = \{G \in \mathsf{GB} : \text{ no edge } (v, u)\}.
$$

The proof for $C$ follows directly from the uar case as $\mathbf{Pr}(G) = \mathbf{Pr}(\phi(G))$. Also, $\mathbf{Pr}(\cup_u B_u) = O(\log^6 n/n)$. The proof for $A$ is simplified by the event $\overline{B_I}$ as there is exactly one $(v, u)$ edge to consider.

For $G(a_1, a_2) \in A$, we define the minimal graph $G(\alpha_1, \alpha_2)$ of $G$ as follows: Let $a_2 = \alpha_2 + \beta_2$ where every edge of $\alpha_2$ is necessary for $S$ to dominate $\overline{S}$ in $G$. Thus, if $e \in \alpha_2$, then $G(a_1, \alpha_2 - e)$ does not satisfy $S$ dominates $\overline{S}$ but $G(a_1, \alpha_2)$ does. It is quite possible that $\alpha_2 = \emptyset$, e.g., only $out(v)$ dominates something or $v$ dominates itself. Let $\alpha_1 = a_1 + \beta_2$. Every $G(a_1, a_2)$ has at least one minimal graph. We form the set

$$
\boldsymbol{G}(a_1, a_2) = \{G(\alpha_1 - I, \alpha_2 + I) : I \subseteq \alpha_1\}.
$$

Thus, $\boldsymbol{G}(a_1, a_2) = \{G(\alpha_1, \alpha_2), \ldots, G(a_1, a_2), \ldots, G(\emptyset, \alpha_1 + \alpha_2)\}$. We claim the following:

(i) Given $G(a_1, a_2) \in A$, then $G(\alpha_1, \alpha_2)$ is unique.

(ii) $\boldsymbol{G}(a_1, a_2) \subseteq A$.

(iii) If $G(a, b), G(a', b')$ are distinct minimal graphs, then

$$\boldsymbol{G}(a, b) \cap \boldsymbol{G}(a', b') = \emptyset.$$

(iv) The sets $\boldsymbol{G}(\alpha_1, \alpha_2)$ partition $A$.

Let $G(a_1, a_2)$ have two distinct minimal graphs $G(\alpha_1, \alpha_2)$ and $G(\alpha'_1, \alpha'_2)$; thus, $\alpha_2 \neq \emptyset$. Since $\alpha_2$ is minimal, there is some edge $e = (x, v) \in \alpha_2$ such that $v$ is the only vertex of $S$ dominating $x$. Thus, $\alpha'_2$ must also have an edge $e' = (x, v)$ for $v$ to dominate $x$. The event $B_{II}$ that $(v, u)$ exists and that distinct parallel edges $e, e'$ to $v$ also exist has been excluded from $A$, so $\alpha'_2 = \alpha_2$.

We next prove that if $G(a_1, a_2) \in A$, then $G(\alpha_1, \alpha_2) \in A$. Since $G(a_1, a_2) \in$ GB, there exists a vertex $x$ uniquely dominated by $v$. Suppose that $x > v$ so that $\alpha_2 \neq \emptyset$ and that $a_1$ does not include an edge from $x$. Also, $\beta_2 = a_2 - \alpha_2$ does not include an edge from $x$, as there are no parallel edges. Thus, there is no edge from $x$ in $\alpha_1 = a_1 + \beta_2$, so $S' = S - v + u$ cannot dominate $x$ and $G(\alpha_1, \alpha_2)$ is bad for $S'$.

The rest of the claims follow in a similar manner. Finally, we claim that

$$\mathbf{Pr}(\boldsymbol{G}(\alpha_1, \alpha_2)) \leq \mathbf{Pr}(\phi(\boldsymbol{G}(\alpha_1, \alpha_2))). \tag{7.6}$$

Note that $\mathbf{Pr}(\phi(G(b_1, b_2))) = \mathbf{Pr}(G(b_2, b_1))$ as the out-edge swaps of $u$ and $v$ are measure preserving (see (7.1) and (7.2)). It also follows from (7.1) and (7.2) that if there is no $(v, u)$ edge, then $\mathbf{Pr}(G(b_1, b_2)) = \mathbf{Pr}(G(b_2, b_1))$. If there is a $(v, u)$ edge, then from (7.3), (7.4), and (7.5),

$$\mathbf{Pr}(G(b_1, b_2)) = c(b_1 + 1), \qquad \mathbf{Pr}(G(b_2, b_1)) = c(b_2 + 1).$$

As $\boldsymbol{G}(\alpha_1, \alpha_2) = \{G(\alpha_1, \alpha_2), \dots, G(\alpha_1 - I, \alpha_2 + I), \dots, G(\emptyset, \alpha_1 + \alpha_2)\}$, we have

$$
\begin{aligned}
\mathbf{Pr}(\boldsymbol{G}(\alpha_1, \alpha_2)) &= c \sum_{j=0}^{\alpha_1} \binom{\alpha_1}{j} (j + 1) \\
&= c \left( 2^{\alpha_1} + \alpha_1 2^{\alpha_1 - 1} \right),
\end{aligned}
$$

whereas $\phi(\boldsymbol{G}(\alpha_1, \alpha_2))$ has the same probability as $\{G(\alpha_2, \alpha_1), \dots, G(\alpha_2 + I, \alpha_1 - I), \dots, G(\alpha_1 + \alpha_2, \emptyset)\}$, and thus

$$
\begin{aligned}
\mathbf{Pr}(\phi(\boldsymbol{G}(\alpha_1, \alpha_2))) &= c \sum_{j=0}^{\alpha_1} \binom{\alpha_1}{j} (\alpha_2 + \alpha_1 - j + 1) \\
&= c \left( (\alpha_1 + \alpha_2 + 1) 2^{\alpha_1} - \alpha_1 2^{\alpha_1 - 1} \right).
\end{aligned}
$$

Thus,

$$\mathbf{Pr}(\phi(\boldsymbol{G}(\alpha_1,\alpha_2))) - \mathbf{Pr}(\boldsymbol{G}(\alpha_1,\alpha_2)) = c\alpha_2 2^{\alpha_1} \geq 0,$$

as required by (7.6).

Finally, we explain the generalization of this proof to scale-free graphs of out-degree $m \geq 2$. The scale-free model treats vertex $x$ as a sequence $(x(1),\ldots,x(m))$ of sub-vertices of degree 1 arising from $m$ consecutive right endpoints of edges in the linear order $1,\ldots,2mn$. Let the $(v,u)$ edge in question be $e_i = (v(i),u(j))$. We swap the edges of $v(i)$ and $u(j)$ except $e_i$ and the corresponding out-edge $f_j$ of $u(j)$. The other sub-vertices are paired cyclically $v(i+k)$ with $u(j+k)$ for swapping.

Let $\boldsymbol{G}_S = \{G_t^m\}$ be the space of scale-free graphs of out-degree $m$. The space $\boldsymbol{G}_S$ is very close to the space of preferential attachment processes $\boldsymbol{G}_C = \{G_t^{C,m}\}$ described in this paper. In either model, on addition of the out-edges $e_j$, $j = 1,\ldots,m$, of vertex $v_t$, the terminal vertex for the edge $e_j$ is chosen by selecting a random end-point of a random edge in a designated set. In the scale-free model, all existing edges and half edges are included in the choice. In the preferential attachment graph process, only out-edges of previously added vertices $v_1,\ldots,v_{t-1}$ are included. Thus, in the scale-free model,

$$\mathbf{Pr}(e_j \text{ chooses from out}(v_1,\ldots,v_{t-1})) = \tfrac{2m(t-1)}{2m(t-1)+2(j-1)+1}, \qquad j = 1,\ldots,m.$$

Thus, in the scale-free model, the subset $C$ of preferential attachment graphs $\boldsymbol{G}_C$ has probability

$$\mathbf{Pr}(C) \geq \prod_{s=2}^{t} \left(1 - \frac{1}{s-1}\right)^m = \Omega(t^{-m}),$$

and conditional probability of graphs in $C$ is the correct measure for $\boldsymbol{G}_C$. The existential lower bounds (derived below) for the smallest dominating set in the scale-free model have probability of the form

$$\mathbf{Pr}(\text{there exists a dominating set } S \text{ of size } |S| \leq dt) = O(c^t),$$

for $c < 1$ constant. We conclude that the results are valid for the preferential attachment model.

Let $S^* = [dt]$; then,

$$\mathbf{Pr}(S^* \text{ dom. } \overline{S^*}) \leq \prod_{v=dn+1}^{n} \left(1 - \left(1 - \frac{\deg(S^*,v-1)}{2m(v-1)}\right)^m\right),$$

and $\deg(S^*, v) = 2mdt(v/dt)^{1/2}(1 + o(1))$ almost surely for $d > 0$ constant (see [Bollobás et al. 01]). Thus

$$\mathbf{Pr}(S^* \text{ dom. } \overline{S^*}) = \exp\left(\sum_{dt+1}^{t} \log(1 - (1 - (dt/v)^{1/2}(1 + o(1)))^m)\right).$$

Let $x = (dt/v)^{1/2}$. Choosing $d$ such that

$$\frac{1}{d^d(1 - d)^{1-d}} \exp\left(2d \int_{\sqrt{d}}^{1} x^{-3} \log(1 - (1 - x)^m)\, dx < 1\right)$$

gives the required lower bound. Values of $(m, d)$ are given in the table in Lemma 7.3.                                                                              □

## 8.   Trees

For $m = 1$ the graph processes under consideration generate a connected graph without cycles. Such a structural property can be exploited to obtain improved lower bounds on $\gamma_1$. Without loss of generality, any vertex in such graphs that has at least one neighbour $u \in V_1$ must be part of a minimum-size dominating set. The number of such vertices is precisely $t - |V_1| - |I|$, where $|I|$ is the number of vertices that have no neighbour in $V_1$. The cardinality of $I$ can be estimated in both models via either a martingale argument similar to those used in previous sections or through the technique exploited in [Cooper 02] to estimate $|V_i|$. The lower bounds in Table 1 for $m = 1$ come from this argument.

We end this section reporting on some simple empirical results that help put the mathematical analysis performed so far into context. It is well known [Cockayne et al. 75] that minimum-size dominating sets can be found efficiently in trees. We implemented Cockayne et al.'s algorithm and tested its performance. For different values of $t$, we repeatedly ran the two graph processes up to time $t$ and then applied Cockayne's algorithm. Table 5 reports the average values that we obtained. The least square approximation lines over the full set of data that we collected are (coefficients rounded to the sixth decimal place) $y = 0.374509x - 0.214185$ for the random graph case and $y = 0.294294x + 0.32284$ for the preferential attachment case. These results indicate that our algorithms are able to get better results for graphs generated according to the preferential attachment graph process ($\alpha_{\mathrm{up}}^C = 0.3333$) than for graphs generated by the uar process ($\alpha_{\mathrm{up}}^R = 0.5$). We leave the finding of improved algorithms, especially in the random graph process, or indeed better lower bounds in either models as an interesting open problem of this work.

| $t$ | $\gamma_1(G_{1,t}^R)/t$ | $\gamma_1(G_{1,t}^C)/t$ |
|---|---|---|
| 10000 | 3745.053 | 2943.157 |
| 20000 | 7489.3 | 5887.301 |
| 30000 | 11233.68 | 8829.288 |
| 40000 | 14980.384 | 11772.175 |
| 50000 | 18725.448 | 14714.073 |
| ... | | |
| 100000 | 37451.20312 | 29424.216 |

**Table 5**. Average values obtained over 1000 experiments for each value of $t$.

# References

[Alon et al. 92] N. Alon, J. H. Spencer, and P. Erdős. *The Probabilistic Method.* New York: John Wiley & Sons, 1992.

[Alzoubi et al. 02] K. Alzoubi, P. J. Wan, and O. Frieder. "Message-Optimal Connected Dominating Sets in Mobile ad hoc Networks." In *Proceedings of the 3rd International Symposium on Mobile ad hoc Networking & Computing*, pp. 157–164. New York: ACM Press, 2002.

[Barabási and Albert 99] A. Barabási and R. Albert. "Emergence of Scaling in Random Networks." *Science* 286 (1999), 509–512.

[Bender and Canfield 78] E. A. Bender and E. R. Canfield. "The Asymptotic Number of Labeled Graphs with Given Degree Sequences." *Journal of Combinatorial Theory* A 24 (1978), 296–307.

[Bollobás et al. 01] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. "The Degree Sequence of a Scale-Free Random Graph Process." *Random Structures and Algorithms* 18 (2001), 279–290.

[Broder et al. 00] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. "Graph Structure in the Web." *Computer Networks* 33:1–6 (2000), 309–320.

[Cockayne et al. 75] E. Cockayne, S. Goodman, and S. Hedetniemi. "A Linear Algorithm for the Domination Number of a Tree." *Information Processing Letters* 4 (1975), 41–44.

[Cooper 02] C. Cooper. "The Age Specific Degree Distribution of Web-Graphs." Submitted to *Combinatorics Probability and Computing*, 2002.

[Cooper and Frieze 03] C. Cooper and A. Frieze. "A General Model of Web Graphs." *Random Structures and Algorithms* 22 (2003), 311–335.

[Duckworth and Zito 00] W. Duckworth and M. Zito. "Sparse Hypercube 3-Spanners." *Discrete Applied Mathematics* 103 (2000), 289–295.

[Eidenbenz 02] S. Eidenbenz. "Online Dominating Set and Variations on Restricted Graph Classes." Technical Report 380, Department of Computer Science, ETH Zürich, 2002.

[Feige 98] U. Feige. "A Threshold of $\ln n$ for Approximating Set Cover." *Journal of the ACM* 45 (1998), 634–652.

[Garey and Johnson 78] M. R. Garey and D. S. Johnson. "Strong NP-Completeness Results: Motivation, Examples, and Implications." *Journal of the ACM* 25:3 (1978), 499–508.

[Harary and Haynes 00] F. Harary and T. W. Haynes. "Double Domination in Graphs." *Ars Comb.* 55 (2000), 201–213.

[Haynes et al. 98a] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater, editors. *Domination in Graphs: Advanced Topics.* New York: Marcel Dekker, 1998.

[Haynes et al. 98b] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Fundamentals of Domination in Graphs.* New York: Marcel Dekker, 1998.

[King and Tzeng 97] G.-H. King and W.-G. Tzeng. "On-Line Algorithms for the Dominating Set Problem." *Information Processing Letters* 61 (1997), 11–14.

[Klasing and Laforest 04] R. Klasing and C. Laforest. "Hardness Results and Approximation Algorithms of $k$-tuple Domination in Graphs." *IPL* 89 (2004), 75–83.

[Kumar et al. 00] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. "The Web as a Graph." In *Proceedings of the Nineteenth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pp. 1–10. New York: ACM Press, 2000.

[Levene and Wheeldon 01] M. Levene and R. Wheeldon. "Web Dynamics." *Software Focus* 2 (2001), 31–38.

[Liao and Chang 03] C.-S. Liao and G. J. Chang. "$k$-Tuple Domination in Graphs." *Information Processing Letters* 87 (2003), 45–50.

[Stojmenovic et al. 02] I. Stojmenovic, M. Seddigh, and J. Zunic. "Dominating Sets and Neighbor Elimination-Based Broadcasting Algorithms in Wireless Networks." *IEEE Trans. Parallel and Dist. Systems* 13 (2002), 14–25.

[Watts and Strogatz 98] D. J. Watts and S. H. Strogatz. "Collective Dynamics of 'Small-World' Networks." *Nature* 393:4 (June 1998), 440–442.

[Wieland and Godbole 01] B. Wieland and A. P. Godbole. "On the Domination Number of a Random Graph." *Elec. J. Combinat.* 8 (2001), R37.

[Wormald 99] N. C. Wormald. "The Differential Equation Method for Random Graph Processes and Greedy Algorithms." In *Lectures on Approximation and Randomized Algorithms*, edited by M. Karoński and H. J. Prömel, pp. 73–155. Warsaw: PWN, 1999.

[Zito 01] M. Zito. "Greedy Algorithms for Minimisation Problems in Random Regular Graphs." In *Algorithms—ESA 2001: 9th Annual European Symposium, Aurhus, Denmark, August 28–31, 2001, Procdeedings*, Lecture Notes in Computer Science 2161, edited by F. Meyer auf der Heide, pp. 524–536. New York: Springer-Verlag, 2001.

Colin Cooper, Department of Computer Science, King's College, London WC2R 2LS, United Kingdom (ccooper@dcs.kcl.ac.uk)

Ralf Klasing, LaBRI–Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France (klasing@labri.fr)

Michele Zito, Department of Computer Science, University of Liverpool, Peach Street, Liverpool L69 7ZF, United Kingdom (M.Zito@csc.liv.ac.uk)