Greedy Maximal Independent Sets via Local Limits

Michael Krivelevich

School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel
krivelev@tauex.tau.ac.il

Tamás Mészáros

Fachbereich Mathematik und Informatik, Institut für Mathematik, Freie Universität Berlin, 14195 Berlin, Germany
tamas.meszaros@fu-berlin.de

Peleg Michaeli

School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel
peleg.michaeli@math.tau.ac.il

Clara Shikhelman

Simons Institute for the Theory of Computing, Berkeley, CA, USA
School of Mathematical Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel
clara.shikhelman@gmail.com

Abstract

The random greedy algorithm for finding a maximal independent set in a graph has been studied extensively in various settings in combinatorics, probability, computer science — and even in chemistry. The algorithm builds a maximal independent set by inspecting the vertices of the graph one at a time according to a random order, adding the current vertex to the independent set if it is not connected to any previously added vertex by an edge.

In this paper we present a natural and general framework for calculating the asymptotics of the proportion of the yielded independent set for sequences of (possibly random) graphs, involving a useful notion of local convergence. We use this framework both to give short and simple proofs for results on previously studied families of graphs, such as paths and binomial random graphs, and to study new ones, such as random trees.

We conclude our work by analysing the random greedy algorithm more closely when the base graph is a tree. We show that in expectation, the cardinality of a random greedy independent set in the path is no larger than that in any other tree of the same order.

2012 ACM Subject Classification Mathematics of computing → Graph algorithms; Mathematics of computing → Random graphs; Mathematics of computing → Probabilistic algorithms

Keywords and phrases Greedy maximal independent set, random graph, local limit

Digital Object Identifier 10.4230/LIPIcs...

Funding *Michael Krivelevich*: Partially supported by USA–Israel BSF grants 2014361 and 2018267, and by ISF grant 1261/17.
Tamás Mészáros: Research supported by the Dahlem Research School and by the Berlin Mathematics Research Center Math+.
Peleg Michaeli: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 676970, RANDGEOM).

Acknowledgements The authors wish to thank the organisers of the Joint FUB–TAU Workshop on Graph and Hypergraph Colouring, hosted by the Freie Universität Berlin in 2018, and Michal Amir, Shagnik Das, Lior Gishboliner, Matan Harel, Frank Mousset, Matan Shalev and Yinon Spinka for useful discussions and ideas.

1 Introduction

Algorithmic problems related to finding or approximating the independence number of a graph, or to producing large independent sets, have long been in the focus of the computer science community. Computing the size of a *maximum* independent set is known to be NP-complete [30] and the groundbreaking work [16] on the difficulty of approximating it even made its way to The New York Times. A natural way to try to efficiently produce a large independent set in an input graph G is to output a *maximal* independent set (MIS) in G, where a vertex subset $I \subseteq V(G)$ is a MIS in G if I is maximal by inclusion. While in principle a badly chosen MIS can be very small (like, say, the star center in a star), one might hope that quite a few of the maximal independent sets will have size comparable in some sense to the independence number of G.

In this paper, we study the *random* greedy algorithm for producing a maximal independent set, which is defined as follows. Consider an input graph G on n vertices. The algorithm first orders the vertices of G uniformly at random, and then builds an independent set $\mathbf{I}(G)$ by considering each of the vertices v one by one in order, adding v to $\mathbf{I}(G)$ if the resulting set does not span any edge. Observe that the set $\mathbf{I}(G)$ is in fact the set of vertices coloured in the first colour in a random greedy proper colouring of G. A basic quantity to study, which turns out to have numerous applications, is the proportion of the yielded independent set (which we call the **greedy independence ratio**). In particular, it is of interest to study the asymptotic behaviour of this quantity for natural graph sequences.

Due to its simplicity, this greedy algorithm has been studied extensively by various authors in different fields, ranging from combinatorics [48], probability [42] and computer science [18] to chemistry [20]. As early as 1931 this model was studied by chemists under the name *random sequential adsorption* (RSA), focusing mostly on d-dimensional grids. The 1-dimensional case was solved by Flory [20] (see also [38]), who showed that the expected greedy independence ratio tends to $\zeta_2 = (1 - e^{-2})/2$ as the path length tends to infinity.

A continuous analogue, where "cars" of unit length "park" at random free locations on the interval $[0, X]$, was introduced (and solved) by Rényi [43], under the name *car-parking process*. The limiting density, as X tends to infinity, is therefore called **Rényi's parking constant**, and ζ_2 may be considered as its discrete counterpart (see, e.g., [17]). Following this terminology, the final state of the car-parking process is often called the *jamming limit* of the graph, and the density of this state is called the *jamming constant*. For dimension 2, Palásti [39] conjectured, in the continuous case (where unit square "cars" park in a larger square), that the limiting density is Rényi's parking constant squared. This conjecture may be carried over to the discrete case, but to the best of our knowledge, in both cases it remains open. For further details see [17] (see also [15] for an extensive survey on RSA models).

In combinatorics, the greedy algorithm for finding a maximal independent set was analysed in order to give a lower bound on the (usually asymptotic) typical independence number of (random) graphs[1]. The asymptotic greedy independence ratio of binomial random graphs with linear edge density was studied by McDiarmid [35] (but see also [9, 25]). The asymptotic greedy independence ratio of random regular graphs was studied by Wormald [48], who used the so-called *differential equation method* (see [49] for a comprehensive survey; see also [47] for a short proof of Wormald's result). His result was further extended in [33] for any regular graph sequence with growing girth (see also [28, 29] for similar extensions for more

[1] In this regard, the greedy algorithm has long been superseded by more sophisticated algorithms; these algorithms often lack, however, the local properties of the greedy algorithm.

sophisticated algorithms). Recently, the case of uniform random graphs with given degree sequences was studied (independently) in [5] and [11].

In a more general setting, where we run the random greedy algorithm on a *hypergraph*, the model recovers in particular the *triangle-free process* (or, more generally, the *H-free process*). In this process, which was first introduced in [14], we begin with the empty graph, and at each step add a random edge as long as it does not create a copy of a triangle (or a copy of H). To recover this process we take the hypergraph whose vertices are the edges of the complete graph, and whose hyperedges are the triples of edges that span a triangle (or k-sets of edges that form a copy of H, if H has k edges). Bohman's key result [7] is that for this hypergraph, $|\mathbf{I}|$ is asymptotically almost surely $\Theta(n^{3/2}\sqrt{\ln n})$, where n is the number of vertices. The exact asymptotics was later found by Bohman and Keevash [8] and by Fiz Pontiveros, Griffiths and Morris [19]. Similar results were obtained for the complete graph on 4 vertices by Warnke [46] and for cycles independently by Picollelli [41] and by Warnke [45]. For a discussion about the general setting, see [4].

Consider the following alternative but equivalent definition of the model. Assign an independent uniform *label* from $[0, 1]$ to each vertex of the graph, and consider it as the *arrival time* of a particle at that vertex. All vertices are initially vacant, and a vertex becomes occupied at the time denoted by its label if and only if all of its neighbours are still vacant at that time. Clearly, we do not need to worry that two particles will arrive at the same time. The set of occupied vertices at time 1 is exactly the greedy MIS. The advantage of this formulation of the model is that under mild assumptions, it can be defined on an infinite graph. We may think of the resulting MIS as a *factor of iid* (**fiid**)[2], meaning, informally, that there exists a local rule which is unaware of the "identity" of a given vertex, that determines whether that vertex is occupied. It was conjectured (formally by Hatami, Lovász and Szegedy [26]) that, using a proper rule, **fiid** can produce an asymptotically maximum independent set in random regular graphs. However, this was disproved recently by Gamarnik and Sudan [23]. In fact, they showed that this kind of local algorithms has a uniformly limited power for sufficiently large degree, and later Rahman and Virág [42] showed that the density of **fiid** independent sets in regular trees and in Poisson Galton–Watson trees, with large average degree, is asymptotically at most half-optimal, concluding (after projecting to random regular graphs or to binomial random graphs) that local algorithms cannot achieve better.

However, on other families of graphs, local algorithms may clearly do better than that. A trivial example is the set of stars, where the greedy algorithm typically performs perfectly. A less trivial example is that of uniform random trees. The expected independence ratio of a uniform random tree is the unique solution of the equation $x = e^{-x}$ (see [36]), which is approximately 0.5671..., while the greedy algorithm yields an independent set of expected density 1/2 as we will see in Section 2.3.

Finally, we note that the following parallel/distributed algorithm gives a further way to look at the maximal independent set generated by the greedy algorithm. After (randomly) ordering the vertices, we colour "red" all the *sinks*, that is, all the vertices which appear before their neighbours in the order, and then remove them and their neighbours from the graph and continue. Formulated this way, the algorithm is very easy to implement, and requires only local communication between the nodes. Also, conditioning on the initial random ordering, it is deterministic, a property which appears to be of importance (see, e.g., [6]). A main question of interest is the number of rounds it takes the algorithm to

[2] The letters **iid** abbreviate "independent and identically distributed".

terminate. In [18] it was shown that with high probability (**whp**)[3] it terminates in $O(\ln n)$ steps on any n-vertex graph, and that this is tight. Thus, even though these algorithms may be suboptimal, they are strikingly simple and can be surprisingly efficient.

1.1 Our Contribution

The goal of this work is to introduce a simple and fairly general framework for calculating the asymptotics of the greedy independence ratio for a wide variety of (random) graph sequences. The general approach is to study a suitable limiting object, typically a random rooted infinite graph, which captures the local view of a typical vertex, and to calculate the probability that its root appears in a random independent set in this graph, which is created according to some natural "local" rule, to be described later. We show that this probability approximates the expected greedy independence ratio.

Let us formulate this more precisely. For a (random) finite graph G let $\mathbf{I}(G)$ be the random greedy maximal independent set of G, let $\iota(G) := |\mathbf{I}(G)|/|V(G)|$ be its density, and let $\bar{\iota}(G)$ be its expected density (taken over the distribution of G and over the random greedy maximal independent set). Suppose (U, ρ) is a random rooted infinite graph (that is, (U, ρ) is a distribution on rooted infinite graphs). A random labelling $\sigma = (\sigma_v)_{v \in V(U)}$ of U is a process consisting of **iid** random variables σ_v, each distributed uniformly in $[0, 1]$. The past of a vertex v, denoted $\mathcal{P}_v$, is the set of vertices in U reachable from v by a monotone decreasing path (with respect to σ). We say that (U, ρ) has **nonexplosive growth** if the past of ρ is almost surely finite. For such (U, ρ) we may define

$$\iota(U, \rho) = \mathbb{P}[\rho \in \mathbf{I}(U[\mathcal{P}_\rho])].$$

We say that a graph sequence G_n converges locally to (U, ρ), and denote it by $G_n \xrightarrow{\text{loc}} (U, \rho)$, if for every $r \geq 0$, the ball of radius r around a uniformly chosen point from G_n converges in distribution to the ball of radius r around ρ in U. To make this notion precise, we need to endow the space of rooted locally finite connected graphs with a topology. This will be done rigorously in Section 3. The following key theorem gives motivation for the definitions above.

▶ **Theorem 1.1.** *If $G_n \xrightarrow{\text{loc}} (U, \rho)$ and (U, ρ) has nonexplosive growth then $\bar{\iota}(G_n) \to \iota(U, \rho)$.*

We remark that $\iota(U, \rho)$ is almost surely positive, implying that for locally convergent graph sequences the expected size of the random greedy maximal independent set is linear.

With some mild growth assumptions on the graph sequence, we can also obtain asymptotic concentration of the greedy independence ratio around its mean. For a graph G let $\mathcal{N}_G(r)$ be the random variable counting the number of paths of length at most r from a uniformly chosen random vertex of G. For two real numbers x, y denote by $x \wedge y$ their minimum. Let

$$\mu^*(r) = \lim_{M \to \infty} \limsup_{n \to \infty} \mathbb{E}[\mathcal{N}_{G_n}(r) \wedge M].$$

We say that G_n has subfactorial path growth (**sfpg**) if $\mu^*(r) \ll_r r!$.[4] Note that every graph sequence with uniformly bounded degrees has **sfpg**, but there are graph sequences with unbounded degrees, and even with unbounded average degree, which still have **sfpg**. For most cases, and for all of the applications presented in this paper, requiring that the somewhat simpler expression $\limsup_{n \to \infty} \mathbb{E}[\mathcal{N}_{G_n}(r)]$ is subfactorial would have sufficed;

[3] That is, with probability tending to 1 as n tends to infinity.
[4] By $g_1(r) \ll_r g_2(r)$ we mean that $\lim_{r \to \infty} g_1(r)/g_2(r) = 0$.

however, requiring that $\mu^*(r)$ is subfactorial is less strict, and is more natural for the following reason: if the graph sequence converges locally, then $\mu^*(r)$ is the expected number of paths of length at most r in the limit. For two functions $f_1(n), f_2(n)$ write $f_1(n) \sim f_2(n)$ if $f_1(n) = (1 + o(1))f_2(n)$. We are now ready to state our concentration result.

▶ **Theorem 1.2.** *If G_n has **sfpg** and $G_n \xrightarrow{\text{loc}} (U, \rho)$ then $\iota(G_n) \sim \iota(U, \rho)$ with high probability.*

▶ Remark. Gamarnik and Goldberg [22] have established concentration of $\iota(G_n)$ around its mean, under the assumption that the degrees of G_n are uniformly bounded. Here we relax that assumption.

▶ Remark. A sequence of graphs which has **sfpg** does not necessarily have a local limit, but it does have a locally convergent subsequence. Any limit of such a sequence will have nonexplosive growth.

When the limiting object is supported on rooted trees, we call the (random) graph sequence **locally tree-like**. Our next result is a general differential-equations based tool for analysing the asymptotics of the greedy independence ratio of locally tree-like (random) **sfpg** graph sequences, with the restriction that their limit may be emulated by a *simple* branching process with at most countably many types. Roughly speaking, a **multitype branching process** is a rooted tree, in which each node is assigned a *type*, and the number and types of each node's "children" follow a law which depends solely on the node's type, and is independent for distinct nodes. Such a branching process is called **simple** if each such law is a product measure. Formal definitions will be given in Section 5. The following theorem reduces the problem of calculating $\iota(U, \rho)$ in these cases to the problem of solving a (possibly infinite) system of ODEs.

▶ **Theorem 1.3.** *Let (U, ρ) be a simple multitype branching process with finite or countable type set T, root distribution $\dot{\mu}$ and offspring distributions $\mu^{k\to j}$. For every $x \in [0, 1]$ and $k, j \in T$ let $\mu_x^{k\to j} = \mathsf{Bin}(\mu^{k\to j}, x)$ denote the distribution of the number of children of type j of a node of type k with random label at most x. Then,*

$$\iota(U, \rho) = \sum_{k\in T} y_k(1)\dot{\mu}(k), \tag{1}$$

where $\{y_k\}_{k\in T}$ is a solution to the following system of ODEs:

$$y_k'(x) = \sum_{\ell\in\mathbb{N}^T} \prod_{j\in T} \mu_x^{k\to j}(\ell_j) \left(1 - \frac{y_j(x)}{x}\right)^{\ell_j}, \qquad y_k(0) = 0. \tag{$*$}$$

We call $(*)$ the **fundamental system of ODEs** of the branching process (U, ρ). While this system of ODEs may seem complicated, in many important cases it reduces to a fairly simple system, as we will demonstrate in Section 2. In particular, the proof of Theorem 1.3 implies that a solution to $(*)$ exists, and in the presented applications it will be unique. In the cases where (U, ρ) is either a single type branching process or a random tree with **iid** degrees, we provide an easy probability generating function tool that may be used to "skip" solving $(*)$. This is described in Appendix B. We mention that a somewhat related, but apparently less applicable statement, providing differential equations for the occupancy probability of a given vertex in bounded degree graphs, appears in [40].

▶ Remark. The proof of Theorem 1.3 actually yields a stronger statement. Replacing $y_k(1)$ with $y_k(x)$ in the RHS of (1), the obtained quantity is the probability that the root is occupied "at time x", namely, when vertices whose label is above x are ignored.

We conclude our work with a theorem, according to which on the set of all trees of a given order the expected size of the greedy MIS achieves its minimum on the path.

▶ **Theorem 1.4.** *Let $n \geq 1$, let T be a tree on n vertices and let P_n be the path on n vertices. Then $\bar{\iota}(P_n) \leq \bar{\iota}(T)$.*

This theorem gives us an exact (non-asymptotic) explicit lower bound for the expected greedy independence ratio of trees (an asymptotic upper bound is trivial). The methods used to prove it are different from the ones used in the rest of this paper, and are more combinatorial in nature. In particular, we make use of a transformation on trees, originally introduced by Csikvári in [12], which gives rise to a graded poset of all trees of a given order, in which the path is the unique minimum (say). While we are not able to show that this transformation can only increase the expected greedy independence ratio, we show it can only increase some other quantitative property of trees, which allows us to argue that paths indeed achieve the minimum expected greedy independence ratio.

1.2 Organisation of the Paper

We start by a short list of important applications in Section 2, where we prove some new results and reprove some known ones, using the machinery of Theorems 1.2 and 1.3. In a few cases, we are assisted by the claims from Appendix B. In particular, we calculate the asymptotics of the greedy independence ratio for paths and cycles (reproving results from [20, 38]), binomial random graphs (reproving a result from [35]), uniform random trees and random functional digraphs (new results) and random regular graphs or regular graphs with high girth (reproving results from [33, 48]).

We then shift our focus to the formal definitions and proofs. We begin by introducing the metric that is used to define the notion of *local convergence* in Section 3, where we also prove Theorem 1.1. In Section 4 we prove Theorem 1.2, by essentially proving a decay of correlation between vertices in terms of their distance, and showing that typical pairs of vertices are distant. In fact, the results of Section 4 imply that even without local convergence, under mild growth assumptions, the variance of the greedy independence ratio is decaying.

In Section 5 we turn our attention to locally tree-like graph sequences, define (simple, multitype) branching processes, and prove Theorem 1.3. We enhance this in Appendix B by introducing a probability generating functions based "trick", which allows, in some cases, a significant simplification. In Section 6 we focus further on tree sequences, where we prove Theorem 1.4. To this end we pinpoint several interesting properties of the expected greedy independence ratio of the path.

2 Applications

The goal of this section is to demonstrate the power of the introduced framework by finding ι for several natural (random) graph sequences, via finding their local limit and solving its fundamental system of ODEs, as described in Theorem 1.3. In some cases, we may use probability generating functions, as described in Appendix B, to ease calculations.

2.1 Infinite-Ray Stars

For $d \geq 1$, let $\mathcal{S}_d$ be the **infinite-ray star** with d branches. Formally, the vertex set of $\mathcal{S}_d$ is $\{(0,0)\} \cup \{(i,j) : i \in [d], j = 1, 2, \dots\}$, and $(i,j) \sim (i',j')$ if $|j - j'| = 1$ and either $i = i'$ or $ii' = 0$. Note that $\mathcal{S}_1 = \mathbb{N}$ and $\mathcal{S}_2 = \mathbb{Z}$. This is a two-type branching process, with types

d for the root and 1 for a branch vertex. The fundamental system of ODEs in this case is $y_d'(x) = (1 - y_1(x))^d$, and for $d = 1$ we obtain the equation $y_1' = 1 - y_1$ of which the solution is $y_1(x) = 1 - e^{-x}$. For $d > 1$ we obtain the equation $y_d' = e^{-dx}$ of which the solution is $y_d(x) = \frac{1}{d}(1 - e^{-dx})$. Since $\tau = d$ a.s., it follows that $\iota(\mathcal{S}_d) = y_d(1) = \zeta_d := \frac{1}{d}(1 - e^{-d})$. In particular, $\iota(\mathbb{N}) = 1 - e^{-1} \approx 0.6321...$ and $\iota(\mathbb{Z}) = \frac{1}{2}(1 - e^{-2}) \approx 0.43233....$

As $\mathbb{N}$ is a single type branching process and $\mathbb{Z}$ is a random tree with **iid** degrees, we may use the alternative approach for calculating $\iota(\mathbb{N})$ and $\iota(\mathbb{Z})$, as described in Appendix B. Solving $\int_h^1 \frac{dz}{z} = 1$ gives $h = e^{-1}$, hence by Claim B.1, $\iota(\mathbb{N}) = 1 - e^{-1}$, and by Claim B.2, $\iota(\mathbb{Z}) = \frac{1}{2}\left(1 - e^{-2}\right)$.

The local limit of the sequences P_n of paths and C_n of cycles is clearly $\mathbb{Z}$. It follows from the discussion above that $\iota(P_n), \iota(C_n) \sim \frac{1}{2}(1 - e^{-2})$ **whp**. This was already calculated by Flory [20] (who only considered the expected ratio) and independently by Page [38], and can be thought of as the discrete variant of Rényi's parking constant (see [17]).

2.2 Poisson Galton–Watson Trees

A Poisson Galton–Watson tree $\mathcal{T}_\lambda$ is a single type branching process with offspring distribution $\mathsf{Pois}(\lambda)$ for some parameter $\lambda \in (0, \infty)$. The fundamental ODE in this case is $y'(x) = e^{-\lambda y(x)}$. (This can be calculated directly using (4)). The solution for this differential equation is $y(x) = \ln(1+\lambda x)/\lambda$, hence $\iota(\mathcal{T}_\lambda) = y(1) = \ln(1+\lambda)/\lambda$. The same result can be obtained using the probability generating function of the Poisson distribution, as described in Appendix B.

Consider the **binomial random graph** $G(n, \lambda/n)$, which is the graph on n vertices in which every pair of nodes is connected by an edge independently with probability λ/n. It is easy to check that it converges locally to $\mathcal{T}_\lambda$, hence $\iota(G(n, \lambda/n)) \sim \ln(1 + \lambda)/\lambda$ **whp**, recovering a known result (see [35]).

2.3 Size-Biased Poisson Galton–Watson Trees

For $0 < \lambda \leq 1$, a size-biased Poisson Galton–Watson tree $\hat{\mathcal{T}}_\lambda$ can be defined (see [34]) as a two-type simple branching process, with types s (*spine* vertices) and t (*tree* vertices), where a spine vertex has 1 spine child plus $\mathsf{Pois}(\lambda)$ tree children, a tree vertex has $\mathsf{Pois}(\lambda)$ tree children, and the root is a spine vertex. The fundamental system of ODEs in this case is

$$y_{\mathsf{s}}'(x) = x \sum_{d=0}^{\infty} \frac{(\lambda x)^d}{e^{\lambda x} d!} \left(1 - \frac{y_{\mathsf{s}}(x)}{x}\right) \left(1 - \frac{y_{\mathsf{t}}(x)}{x}\right)^d + (1 - x) \sum_{d=0}^{\infty} \frac{(\lambda x)^d}{e^{\lambda x} d!} \left(1 - \frac{y_{\mathsf{t}}(x)}{x}\right)^d$$

$$= (1 - y_{\mathsf{s}}(x)) \sum_{d=0}^{\infty} \frac{(\lambda x)^d}{e^{\lambda x} d!} \left(1 - \frac{y_{\mathsf{t}}(x)}{x}\right)^d = (1 - y_{\mathsf{s}}(x)) e^{-\lambda y_{\mathsf{t}}(x)},$$

and from Section 2.2 we obtain $y_{\mathsf{t}}(x) = \ln(1 + \lambda x)/\lambda$. Hence $y_{\mathsf{s}}'(x) = (1 - y_{\mathsf{s}}(x))/(1 + \lambda x)$, and the solution for that equation is $y_{\mathsf{s}}(x) = 1 - \exp(-\ln(1 + \lambda x)/\lambda)$. Thus $\iota(\hat{\mathcal{T}}_\lambda) = y_{\mathsf{s}}(1) = 1 - (1 + \lambda)^{-1/\lambda} = 1 - e^{-\iota(\mathcal{T}_\lambda)}$. In particular, $\iota(\hat{\mathcal{T}}_1) = 1/2$.

It is a classical (and beautiful) fact (see, e.g., [24, 32]) that if T_n is a uniformly chosen random tree drawn from the set of n^{n-2} trees on (labelled) n vertices, then T_n converges locally to $\hat{\mathcal{T}}_1$, hence $\iota(T_n) \sim 1/2$ **whp**. To the best of our knowledge, this intriguing fact was not previously known. In fact, it was shown recently in [27] that if G_n is a sequence of connected regular graphs that converges to a nondegenerate graphon, and T_n is the uniform spanning tree of G_n, then T_n also converges locally to $\hat{\mathcal{T}}_1$, hence it follows that $\iota(T_n) \sim 1/2$ **whp** in this case as well.

It can be easily verified that the local limit of a random functional digraph $\vec{G}_1(n)$ (the digraph on n vertices whose edges are $(i, \pi(i))$ for a uniform random permutation π), with orientations ignored, is also $\hat{\mathcal{T}}_1$, hence $\iota(\vec{G}_1) \to 1/2$ **whp**.

2.4 d-ary Trees

For $d > 1$, let T_d be the d-ary tree. It may be viewed as a (single type) branching process. It thus immediately follows from (4) that $y'(x) = (1 - y(x))^d$. The solution for this differential equation is $y(x) = 1 - ((d-1)x + 1)^{-1/(d-1)}$. It follows that $\iota(\mathsf{T}_d) = y(1) = 1 - d^{-1/(d-1)}$. This fact also follows easily using the generating functions approach described in Appendix B. A remarkable example is $\iota(\mathsf{T}_2) = 1/2$.

2.5 Regular Trees

For $d \geq 3$, let $\mathbb{T}_d$ be the d-regular tree. It may viewed as a two-type branching process with types d for the root and 1 for the rest of the vertices. The fundamental system of ODEs in this case is $y_d'(x) = (1 - y_1(x))^d$, and from Section 2.4 we obtain $y_1(x) = 1 - ((d-2)x + 1)^{-1/(d-2)}$. It follows that $y_d'(x) = ((d-2)x+1)^{-d/(d-2)}$, of which the solution is $y_d(x) = (1 - ((d-2)x + 1)^{-2/(d-2)})/2$. Therefore,

$$\iota(\mathbb{T}_d) = y_d(1) = \frac{1}{2}\Big(1 - (d-1)^{-2/(d-2)}\Big).$$

As with d-ary trees, here again the generating functions approach works easily: the solution to $\int_{h(x)}^1 z^{d-1} dz = x$ is $h(x) = (1 - (2-d)x)^{1/(2-d)}$, and the result follows from Claim B.2. Remarkable examples include $\iota(\mathbb{T}_3) = 3/8$ and $\iota(\mathbb{T}_4) = 1/3$.

Since the **random regular graph** $G(n, d)$ (a uniformly sampled graph from the set of all d-regular graphs on n vertices, assuming dn is even) converges locally to $\mathbb{T}_d$ (see, e.g., [50]), the above result for this case is exactly [48, Theorem 4]. In fact, since any sequence of d-regular graphs with girth tending to infinity converges locally to $\mathbb{T}_d$, we also recover [33, Theorem 2].

3 Local Limits

In order to study asymptotics, it is often useful to construct a suitable limiting object first. Local limits were introduced by Benjamini and Schramm [3] and studied further by Aldous and Steele [2] (A very similar approach has already been introduced by Aldous in [1]). Local limits, when they exist, encapsulate the asymptotic data of local behaviour of the convergent graph sequence, and in particular, that of the performance of the greedy algorithm.

We start with basic definitions. Consider the space $\mathcal{G}_\bullet$ of rooted locally finite connected graphs viewed up to root preserving graph isomorphisms. We provide $\mathcal{G}_\bullet$ with the metric $d_{\text{loc}}((G_1, \rho_1), (G_2, \rho_2)) = 2^{-R}$, where R is the largest integer for which $B_{G_1}(\rho_1, R) \simeq B_{G_2}(\rho_2, R)$. Here we understand $B_G(\rho, R)$ as the *rooted* subgraph of (G, ρ) spanned by the vertices of distance at most R from ρ, and $\simeq$ as *rooted-isomorphic*. It is an easy fact that $(\mathcal{G}_\bullet, d_{\text{loc}})$ is a separable complete metric space, hence it is a Polish space. $(\mathcal{G}_\bullet, d_{\text{loc}})$, while being bounded, is not compact (the sequence of rooted stars S_n does not have a convergent subsequence).

Recall that a sequence of random elements $\{X_n\}_{n=1}^\infty$ **converges in distribution** to a random element X, if for every bounded continuous function f we have that $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$. Let G_n be a sequence of (random) finite graphs. We say that G_n **converges**

locally to a (random) element (U, ρ) of $\mathcal{G}_\bullet$ if for every $r \geq 0$, the sequence $B_{G_n}(\rho_n, r)$ converges in distribution to $B_U(\rho, r)$, where ρ_n is a uniformly chosen vertex of G_n. Since the inherited topology on all rooted balls in $\mathcal{G}_\bullet$ with radius r is discrete, this implies convergence in total variation distance.

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Fix $\varepsilon > 0$. For a given labelling σ of U, let ℓ_σ be the length of the longest decreasing sequence (w.r.t. σ) starting from ρ. Since (U, ρ) has nonexplosive growth, there exists r_ε for which for every $r \geq r_\varepsilon$, $\mathbb{P}[\ell_\sigma \geq r] < \varepsilon$. For $r \geq 0$, let $G_n^r = B_{G_n}(\rho_n, r)$ and $U^r = B_U(\rho, r)$. We couple G_n^r and a random permutation π on its vertices with U^r and a random labelling σ as follows. First, since G_n^r converges in distribution (and hence in total variation distance) to U^r, there exists n_r such that for all $n \geq n_r$ we have that $\mathbb{P}[G_n^r \not\cong U^r] \leq \varepsilon$. If this event occurs, we say that the coupling has failed. Otherwise, for some isomorphism $\varphi : G_n^r \to U^r$, we let π be the permutation on the vertices of G_n^r which agrees with the ordering of the labels on the vertices of the isomorphic image (that is, $\pi_u < \pi_v \iff \sigma_{\varphi(u)} < \sigma_{\varphi(v)}$). Note that under this coupling, if it succeeds, $\rho_n \in \mathbf{I}(G_n^r) \iff \rho \in \mathbf{I}(U^r)$. However, on the event "$\ell_\sigma \leq r$", $\rho_n \in \mathbf{I}(G_n^r) \iff \rho_n \in \mathbf{I}(G_n)$ and $\rho \in \mathbf{I}(U^r) \iff \rho \in \mathbf{I}(U[\mathcal{P}_\rho])$. Observing that $\bar{\iota}(G_n) = \mathbb{P}[\rho_n \in \mathbf{I}(G_n)]$ we obtain that for $r \geq r_\varepsilon$ and $n \geq n_r$, $|\bar{\iota}(G_n) - \iota(U, \rho)| < 2\varepsilon$. ◀

4 Concentration

With some mild growth assumptions on the graph sequence, without assuming local convergence, we obtain asymptotic concentration of the greedy independence ratio around its mean. Under these assumptions we show that the dependence between the inclusion of distinct nodes in the maximal independent set decays as a functions of their distance, a phenomenon which is sometimes called *correlation decay* or *long-range independence*. To prove that the model exhibits this phenomenon, we show that with high probability there are no "long" monotone paths emerging from a typical vertex, which is the contents of the next claim. We then observe that two independent random vertices are typically distant, and use a general lemma about exploration algorithms to prove decay of correlation. We remark that similar locality arguments appear in [37]. Some of the proofs are given in Appendix A.

▷ Claim 4.1. Suppose that G_n has **sfpg**. Let π be a uniform random permutation of the vertices of G_n, and let u be a uniformly chosen vertex from G_n. Then, for every $\varepsilon > 0$, there exists $r > 0$ such that for every large enough n, the probability that there exists a monotone decreasing path of length r (w.r.t. π), emerging from u, is at most ε.

▷ Claim 4.2. Suppose that G_n has **sfpg**. Let u, v be two independently and uniformly chosen vertices from G_n. Then, for every $\varepsilon, r \geq 0$ we have that for every large enough n, $\mathbb{P}[\mathrm{dist}_{G_n}(u, v) \leq r] \leq \varepsilon$.

Let $G = (V, E)$ be a graph. An **exploration-decision rule** for G is a (deterministic) function $\mathcal{Q}$, whose input is a pair (S, g), where S is a non-empty sequence of distinct vertices of V, and $g : S \to [0, 1]$, and whose output is either a vertex $v \in V \setminus S$ or a "decision" T or F. An **exploration-decision algorithm** for G, with rule $\mathcal{Q}$, is a (deterministic) algorithm A, whose input is an initial vertex $v \in V$ and a function $f : V \to [0, 1]$, which outputs T or F, and operates as follows. Set $u_1 = v$. Suppose A has already set $u_1, \ldots, u_i$. Let $x = \mathcal{Q}((u_1, \ldots, u_i), f \restriction_{\{u_1, \ldots, u_i\}})$. If $x \in V$, set $u_{i+1} = x$ and continue. Otherwise stop and return x. We call the set $u_1, \ldots, u_i$ at this stage the **range** of the algorithm's run. We denote the output of the algorithm by $\mathsf{A}(v, f)$ and its range by $\text{rng}_{\mathsf{A}}(v, f)$. The radius of the algorithm's run, denoted $\text{rad}_{\mathsf{A}}(v, f)$, is the maximum distance between v and an element of its range.

▶ **Lemma 4.3.** *Let $\varepsilon > 0$. Let $G = (V, E)$ be a graph, let σ be a random labelling of its vertices, let A be an exploration-decision algorithm for G and let $r \geq 1$. Let u, v be sampled independently from some distribution over V. Suppose that w.p. at least $1 - \varepsilon$ both $\text{dist}_G(u, v) \geq 3r$, and $\text{rad}_{\mathsf{A}}(u, \sigma), \text{rad}_{\mathsf{A}}(v, \sigma) \leq r$. Then $|\operatorname{cov}[\mathsf{A}(u, \sigma), \mathsf{A}(v, \sigma)]| = o_\varepsilon(1)$.*

We now apply the lemma in our setting.

▷ Claim 4.4. Suppose that G_n has **sfpg**. Let u, v be two independently and uniformly chosen vertices from G_n. Denote by R_u, R_v the events that $u \in \mathbf{I}(G_n)$, $v \in \mathbf{I}(G_n)$, respectively. Then $|\operatorname{cov}[R_u, R_v]| = o(1)$.

Proof. Let $\varepsilon > 0$. We describe an exploration-decision algorithm A by defining its rule. Given a vertex sequence $S = (u_1, \ldots, u_i)$ and labels $g : S \to [0, 1]$, the rule checks for monotone decreasing sequences emerging from u_1, in S, with respect to g. Denote by $\mathcal{E}$ the set of ends of these sequences. If there are vertices in $V \setminus S$ with neighbours in $\mathcal{E}$, return an arbitrary vertex among these. Otherwise, perform the Greedy MIS algorithm on the past of u_1 inside S, and return T if u_1 ends up in the MIS, or F otherwise. We observe that if σ is a random labelling of G_n then for $w \in \{u, v\}$ the event $\mathsf{A}(w, \sigma) = \mathsf{T}$ is in fact the event R_w. We also note that if the longest monotone decreasing sequence, w.r.t. σ, emerging from w is of length $r - 1$, then $\text{rad}_{\mathsf{A}}(w, \sigma) \leq r$.

By Claim 4.1 there exists $r > 0$ such that for every large enough n the probability that there exists a monotone decreasing path of length $r - 1$ from either u or v is at most ε. By Claim 4.2, for large enough n, the probability that the distance between u and v is at most $3r$ is at most ε. Therefore, by Lemma 4.3, $|\operatorname{cov}[\mathsf{A}(u, \sigma), \mathsf{A}(v, \sigma)]| = o_\varepsilon(1)$. ◀

▷ Claim 4.5. Suppose that G_n has **sfpg**. Then $\operatorname{Var}[\iota(G_n)] = o(1)$.

Proof. For a vertex w, denote by R_w the event that $w \in \mathbf{I}(G_n)$. Let u, v be two independently and uniformly chosen vertices from G_n. Since the random variables $\mathbb{E}[R_u \mid u]$ and $\mathbb{E}[R_v \mid v]$ are independent, by Claim 4.4,

$$\begin{aligned}\operatorname{Var}[\iota(G_n)] &= \mathbb{E}[\operatorname{cov}[R_u, R_v \mid u, v]] \\ &= \operatorname{cov}[R_u, R_v] - \operatorname{cov}[\mathbb{E}[R_u \mid u]\mathbb{E}[R_v \mid v]] = \operatorname{cov}[R_u, R_v] = o(1).\end{aligned}$$ ◀

Proof of Theorem 1.2. Let $\varepsilon > 0$. Note that since G_n has **sfpg**, (U, ρ) has nonexplosive growth, hence by Theorem 1.1 there exists n_0 such that for every $n \geq n_0$, $|\bar{\iota}(G_n) - \iota(U, \rho)| \leq \varepsilon$. Thus, by Chebyshev's inequality and Claim 4.5,

$$\mathbb{P}[|\iota(G_n) - \iota(U, \rho)| > 2\varepsilon] \leq \mathbb{P}[|\iota(G_n) - \bar{\iota}(G_n)| > \varepsilon] \leq \varepsilon^{-2} \operatorname{Var}[\iota(G_n)] = o(1).$$ ◀

5 Branching Processes and Differential Equations

As promised, we begin with a formal definition of multitype branching processes. Let T be a finite or countable set, which we call the **type set**. Let $\dot{\mu}$ be a distribution on T, which we call the **root distribution**, and for each $k \in T$ let $(\mu^{k \to j})_{j \in T}$ be an **offspring distribution**, which is a distribution on vectors with nonnegative integer coordinates. Let $\tau \sim \dot{\mu}$ and for every finite sequence of natural numbers $\mathbf{v}$ let $(\xi^{k \to j}_{\mathbf{v}})_{j \in T} \sim (\mu^{k \to j})_{j \in T}$ be a random vector, where these random vectors are independent for different indices $\mathbf{v}$ and are independent of τ. A **multitype branching process** $(\mathbf{Z}_t)_{t\in\mathbb{N}}$ with type set T, root distribution $\dot{\mu}$ and offspring distributions $(\mu^{k\to j})_{j\in T}$ is a Markov process on labelled trees, in which each vertex is assigned a type in T, which may be described as follows. At time $t=0$ the tree $\mathbf{Z}_0$ consists of a single vertex of type τ, labelled by the empty sequence. At time $t+1$ the tree $\mathbf{Z}_{t+1}$ is obtained from $\mathbf{Z}_t$ as follows. For each $k\in T$ and $\mathbf{v}$ of length t and type k in $\mathbf{Z}_t$, we add the vertices $\mathbf{v}^\frown i$ for all $0\le i<\sum_{j\in T}\xi_{\mathbf{v}}^{k\to j}$, having exactly $\xi_{\mathbf{v}}^{k\to j}$ of them being assigned type j, uniformly at random, and connecting them with edges to $\mathbf{v}$.[5] If in addition $(\mu^{k\to j})_{j\in T}$ is a product measure, namely, if $\xi_{\mathbf{v}}^{k\to j}\sim\mu^{k\to j}$ are sampled independently for distinct $j\in T$, the process is called **simple**. We often think of a multitype branching process as the possibly infinite (random) rooted graph $\mathbf{Z}_\infty=\bigcup_{t\ge0}\mathbf{Z}_t$, rooted at the single vertex of $\mathbf{Z}_0$.

Proof of Theorem 1.3. Let σ be a random labelling of U. To ease notation, set $\iota=\iota(U,\rho)$ and $\mathbf{I}=\mathbf{I}(U[\mathcal{P}_\rho])$, and recall that $\iota=\mathbb{P}[\rho\in\mathbf{I}]$. Let $\tau\sim\dot{\mu}$ be the type of the root. For $k\in T$ and $x\in[0,1]$, define $\iota^{(k)}=\mathbb{P}[\rho\in\mathbf{I}\mid\tau=k]$ and $\iota_x^{(k)}=\mathbb{P}[\rho\in\mathbf{I}\mid\sigma_\rho=x,\ \tau=k]$. Note that this is well defined, even if the event that $\sigma_\rho=x$ has probability 0. Let further

$$\iota_{<x}^{(k)}=\int_0^x\iota_z^{(k)}dz,$$

so $\iota^{(k)}=\iota_{<1}^{(k)}$, hence

$$\iota=\sum_{k\in T}\iota_{<1}^{(k)}\cdot\mathbb{P}[\tau=k].$$

It therefore suffices to show that the family $y_k(x):=\iota_{<x}^{(k)}$ satisfies $(*)$ (it clearly satisfies the boundary conditions). The key observation is that distinct children in the past of the root are roots to independent subtrees. Formally, conditioning on the event that $v_1,\dots,v_a$ are the children of ρ in its past, the events "$v_i\in\mathbf{I}$" for $i=1,\dots,a$ are mutually independent. Since $\rho\in\mathbf{I}$ if and only if $v_i\notin\mathbf{I}$ for every $i=1,\dots,a$,

$$\begin{aligned}y_k'(x)=(\iota_{<x}^{(k)})'=\iota_x^{(k)}&=\sum_{\ell\in\mathbb{N}^T}\prod_{j\in T}\mu_x^{k\to j}(\ell_j)(1-\mathbb{P}[\rho\in\mathbf{I}\mid\sigma_\rho<x,\ \tau=j])^{\ell_j}\\&=\sum_{\ell\in\mathbb{N}^T}\prod_{j\in T}\mu_x^{k\to j}(\ell_j)\left(1-\frac{y_j(x)}{x}\right)^{\ell_j}.\end{aligned}$$ ◀

6 Lower Bound on Tree Sequences

Let us focus on tree sequences. How large can the expected greedy independent ratio be? How small can it be? The sequence of stars is a clear witness that the only possible asymptotic upper bound is the trivial one, namely 1. Apparently, the lower bound is not trivial. An immediate corollary of Theorems 1.1 and 1.4 is that a tight asymptotic lower bound is $\iota(\mathbb{Z})=(1-e^{-2})/2$ (compare with [44]). The statement of Theorem 1.4 is, however, much stronger: paths achieve the *exact* (non-asymptotic) lower bound for the expected greedy independence ratio among the set of all trees of a given order.

To prove Theorem 1.4 we will need to first understand the behaviour of the greedy algorithm on the path.

[5] By $\mathbf{v}^\frown i$ we mean the sequence obtained from $\mathbf{v}$ by appending the element i.

For a graph G denote by $\mathfrak{i}(G)$ the cardinality of its greedy independent set, and let $\bar{\mathfrak{i}}(G) = \mathbb{E}[\mathfrak{i}(G)]$. Let $\alpha_n = \bar{\mathfrak{i}}(P_n)$. Suppose the vertices of P_n are $1, \dots, n$, and let S be the vertex which is first in the permutation of the vertices. Setting $\alpha_{-1} = \alpha_0 = 0$, we obtain the recursion

$$\alpha_n = \mathbb{E}[\mathbb{E}[\mathfrak{i}(P_n) \mid S]] = \frac{1}{n}\sum_{i=1}^{n}(1 + \alpha_{i-2} + \alpha_{n-i-1}) = 1 + \frac{2}{n}\sum_{i=1}^{n}\alpha_{i-2}, \tag{2}$$

from which an explicit formula for α_n can be derived (see [21]). We will need the following two properties of α_n, whose proofs (which are rather long and technical) we omit in this extended abstract.

▷ Claim 6.1. α_n is monotone increasing and subadditive.

Let $\xi_{n,\ell} = \sum_{j=1}^{\ell} \alpha_{n+j}$.

▷ Claim 6.2. For every $\ell, a, b \geq 1$ it holds that $\xi_{a,\ell} + \xi_{b,\ell} \leq \xi_{a+b,\ell} + \xi_{0,\ell}$.

6.1 KC-Transformations

In this section we introduce the main tool that will be used to prove Theorem 1.4. Let T be a tree and let x, y be two vertices of T. We say that the path between x and y is **bare** if for every vertex $v \neq x, y$ on that path, $d_T(v) = 2$. Suppose x, y are such that the unique path P in T between them is bare, and let z be the neighbour of y in that path. For a vertex v, denote by $N(v)$ the neighbours of v in T. The **KC-transformation** $\mathrm{KC}(T, x, y)$ of T with respect to x, y is the tree obtained from T by deleting every edge between y and $N(y) \setminus z$ and adding the edges between x and $N(y) \setminus z$ instead. Note that $\mathrm{KC}(T, x, y) \simeq \mathrm{KC}(T, y, x)$, so if we care about unlabelled trees, we may simply write $\mathrm{KC}(T, P)$, for a bare path P in T. The term "KC-transformation" was coined by Bollobás and Tyomkyn [10] after Kelmans, who defined a similar operation on graphs [31], and Csikvári, who defined it in this form [12] under the name "generalized tree shift" (GTS).

A nice property of KC-transformations, first observed by Csikvári [12], is that they induce a graded poset on the set of unlabelled trees of a given order, which is graded by the number of leaves. In particular, this means that in that poset, the path is the unique minimum (say) and the star is the unique maximum. Note that if P contains a leaf then $\mathrm{KC}(T, P) \simeq T$, and otherwise $\mathrm{KC}(T, P)$ has one more leaf than T. In the latter case, we say that the transformation is **proper**.

Here is the plan for how to prove Theorem 1.4. For a tree T and a vertex v, denote by $T \star v$ the forest obtained from T by **shattering** T at v, that is, by removing from T the set $\{v\} \cup N(v)$. Denote by $\kappa_v(T)$ the multiset of orders of trees in the forest $T \star v$, and by $\kappa(T)$ the sum of $\kappa_v(T)$ for all vertices v in T. Note that for trees with up to 3 vertices, Theorem 1.4 is trivial; we proceed by induction. By the induction hypothesis,

$$\bar{\mathfrak{i}}(T) = \frac{1}{n}\sum_{v \in V(T)}\sum_{S \in T \star v}(1 + \bar{\mathfrak{i}}(S)) \geq 1 + \frac{1}{n}\sum_{v \in V(T)}\sum_{k \in \kappa_v(T)}\alpha_k = 1 + \frac{1}{n}\sum_{k \in \kappa(T)}\alpha_k. \tag{3}$$

Therefore, it makes sense to study the quantities $\nu_v(T) = \sum_{k \in \kappa_v(T)} \alpha_k$ and $\nu(T) = \sum_{k \in \kappa(T)} \alpha_k$. In fact, it would suffice to show that for any tree T on n vertices $\nu(T) \geq \nu(P_n)$, since by (2) and (3) we would obtain

$$\bar{\mathfrak{i}}(T) \geq 1 + \frac{1}{n}\nu(T) \geq 1 + \frac{1}{n}\nu(P_n) = \bar{\mathfrak{i}}(P_n).$$

We therefore reduced our problem to proving the following theorem about KC-transformations.

▶ **Theorem 6.3.** *If T is a tree and P is a bare path in T then* $\nu(\mathrm{KC}(T,P)) \geq \nu(T)$.

It would have been nice if for every $v \in V(T)$ we would have had $\nu_v(\mathrm{KC}(T,P)) \geq \nu_v(T)$; unfortunately, this is not true in general. However, the following statement, whose proof can be found in Appendix C, would suffice.

▶ **Theorem 6.4.** *Let T be a tree and let $x \neq y$ be two vertices with the path between them being bare. Denote $T' = \mathrm{KC}(T,x,y)$. Let A be the set of vertices $v \neq x$ in T for which every path between v and y passes via x, and similarly, let B be the set of vertices $v \neq y$ in T for which every path between v and x passes via y. Let P be the set of vertices on the bare path between x and y, so $A \cup B \cup P$ is a partition of $V(T)$. Then*

1. *For $v \in A \cup B$ we have that $\nu_v(T') \geq \nu_v(T)$.*
2. $\sum_{v\in P} \nu_v(T') \geq \sum_{v\in P} \nu_v(T)$.

7 Concluding Remarks and Open Questions

Non Locally Tree-Like Graph Sequences

Our local limit approach does not assume that the converging sequence is locally tree-like. However, the differential equation tool fails completely if short cycles appear in a typical local view. As it seems, to date, there is no general tool to handle these cases, and indeed, even the asymptotic behaviour of the random greedy MIS algorithm on d-dimensional tori (for $d \geq 2$) remains unknown.

Better Local Rules

The random greedy algorithm presented here follows a very simple local rule. More complicated local rules may yield, in some cases, larger maximal independent sets; for example, the initial random ordering may "favour" low degree vertices. It would be nice to adapt our framework, or at least some of its components, to other settings. For *adaptive* "better" local algorithms we refer the reader to [48, 51].

The Second Colour

In this work we have analysed the output of the random greedy algorithm for producing a maximal independent set. As already remarked, this is in fact the set of vertices in the first colour class in the random greedy colouring algorithm. It is rather easy to see, that, after slight modifications (in particular, in Theorem 1.3) this approach allow us to calculate the asymptotic proportion of the size of the set of vertices in the second colour class (or in the k'th colour class in general, for any fixed k) as well. Non-asymptotic questions about the expected cardinality of the set of vertices in the second colour class might be also of interest. For example, is it true that the path has the smallest expected number of vertices in the first two colour classes among all trees of the same order? It is not hard to see that this statement is not true for the first three colour classes (as three colours suffice to greedily colour the path).

Monotonicity With Respect to KC-Transformations

It is likely that the expected greedy independence ratio in trees is monotone with respect to KC-transformations, and strictly monotone with respect to *proper* KC-transformations. If true, this would imply that the greedy independence ratio in trees achieves its unique minimum on the path and its unique maximum on the star.

References

[1] David Aldous, *Asymptotic fringe distributions for general families of random trees*, The Annals of Applied Probability **1** (1991), no. 2, 228–266. MR1102319

[2] David Aldous and J. Michael Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, Probability on discrete structures, 2004, pp. 1–72. MR2023650

[3] Itai Benjamini and Oded Schramm, *Recurrence of distributional limits of finite planar graphs*, Electronic Journal of Probability **6** (2001), no. 23, 13. MR1873300

[4] Patrick Bennett and Tom Bohman, *A note on the random greedy independent set algorithm*, Random Structures & Algorithms **49** (2016), no. 3, 479–502. MR3545824

[5] Paola Bermolen, Matthieu Jonckheere and Pascal Moyal, *The jamming constant of uniform random graphs*, Stochastic Processes and their Applications **127** (2017), no. 7, 2138–2178. MR3652409

[6] Guy E. Blelloch, Jeremy T. Fineman, Phillip B. Gibbons and Julian Shun, *Internally deterministic parallel algorithms can be fast*, Proceedings of the 17th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, 2012, pp. 181–192.

[7] Tom Bohman, *The triangle-free process*, Advances in Mathematics **221** (2009), no. 5, 1653–1677. MR2522430

[8] Tom Bohman and Peter Keevash, *Dynamic concentration of the triangle-free process*, The Seventh European Conference on Combinatorics, Graph Theory and Applications, 2013, pp. 489–495. MR3185850

[9] Béla Bollobás and Paul Erdős, *Cliques in random graphs*, Mathematical Proceedings of the Cambridge Philosophical Society **80** (1976), no. 3, 419–427. MR0498256

[10] Béla Bollobás and Mykhaylo Tyomkyn, *Walks and paths in trees*, Journal of Graph Theory **70** (2012), no. 1, 54–66. MR2916067

[11] Graham Brightwell, Svante Janson and Malwina Luczak, *The greedy independent set in a random graph with given degrees*, Random Structures & Algorithms **51** (2017), no. 4, 565–586. MR3718590

[12] Péter Csikvári, *On a poset of trees*, Combinatorica **30** (2010), no. 2, 125–137. MR2676831

[13] Herold G. Dehling, Sjoert R. Fleurke and Christof Külske, *Parking on a random tree*, Journal of Statistical Physics **133** (2008), no. 1, 151–157. MR2438901

[14] Paul Erdős, Stephen Suen and Peter Winkler, *On the size of a random maximal graph*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993), 1995, pp. 309–318. MR1370965

[15] James W. Evans, *Random and cooperative sequential adsorption*, Reviews of modern physics **65** (1993), no. 4, 1281.

[16] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra and Mario Szegedy, *Interactive proofs and the hardness of approximating cliques*, Journal of the ACM **43** (1996), no. 2, 268–292. MR1408323

[17] Steven R. Finch, **Mathematical constants**, Encyclopedia of Mathematics and its Applications, vol. 94, Cambridge University Press, Cambridge, 2003. MR2003519

[18] Manuela Fischer and Andreas Noever, *Tight analysis of parallel randomized greedy MIS*, Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 2018, pp. 2152–2160. MR3775928

[19] Gonzalo Fiz Pontiveros, Simon Griffiths and Robert Morris, *The triangle-free process and the Ramsey number $R(3,k)$*, arXiv e-prints (2013Feb), arXiv:1302.6279, available at `https://arxiv.org/abs/1302.6279`.

[20] Paul J. Flory, *Intramolecular reaction between neighboring substituents of vinyl polymers*, Journal of the American Chemical Society **61** (1939), no. 6, 1518–1521.

[21] Dave Freedman and Larry Shepp, *An Unfriendly Seating Arrangement*, SIAM Review **4** (1962Apr), no. 2, 150–150.

[22] David Gamarnik and David A. Goldberg, *Randomized greedy algorithms for independent sets and matchings in regular graphs: exact results and finite girth corrections*, Combinatorics, Probability and Computing **19** (2010), no. 1, 61–85. MR2575098

[23] David Gamarnik and Madhu Sudan, *Limits of local algorithms over sparse random graphs*, The Annals of Probability **45** (2017), no. 4, 2353–2376. MR3693964

[24] Geoffrey R. Grimmett, *Random labelled trees and their branching networks*, Australian Mathematical Society. Journal. Series A **30** (1980/81), no. 2, 229–237. MR607933

[25] Geoffrey R. Grimmett and Colin J. H. McDiarmid, *On colouring random graphs*, Mathematical Proceedings of the Cambridge Philosophical Society **77** (1975), 313–324. MR0369129

[26] Hamed Hatami, László Lovász and Balázs Szegedy, *Limits of locally-globally convergent graph sequences*, Geometric and Functional Analysis **24** (2014), no. 1, 269–296. MR3177383

[27] Jan Hladký, Asaf Nachmias and Tuan Tran, *The local limit of the uniform spanning tree on dense graphs*, Journal of Statistical Physics **173** (2018), no. 3-4, 502–545. MR3876899

[28] Carlos Hoppen and Nicholas Wormald, *Properties of regular graphs with large girth via local algorithms*, Journal of Combinatorial Theory. Series B **121** (2016), 367–397. MR3548298

[29] Carlos Hoppen and Nicholas Wormald, *Local algorithms, regular graphs of large girth, and random regular graphs*, Combinatorica **38** (2018), no. 3, 619–664. MR3876879

[30] Richard M. Karp, *Reducibility among combinatorial problems*, Proceedings of a symposium on the Complexity of Computer Computations, held March 20-22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, USA, 1972, pp. 85–103. MR0378476

[31] Alexander K. Kelmans, *On graphs with randomly deleted edges*, Acta Mathematica Academiae Scientiarum Hungaricae **37** (1981), no. 1-3, 77–88. MR616881

[32] Valentin F. Kolchin, *Branching processes, random trees, and a generalized scheme of arrangements of particles*, Mathematical notes of the Academy of Sciences of the USSR **21** (1977), no. 5, 386–394.

[33] Joseph Lauer and Nicholas C. Wormald, *Large independent sets in regular graphs of large girth*, Journal of Combinatorial Theory. Series B **97** (2007), no. 6, 999–1009. MR2354714

[34] Russell Lyons, Robin Pemantle and Yuval Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, The Annals of Probability **23** (1995), no. 3, 1125–1138. MR1349164

[35] Colin J. H. McDiarmid, *Colouring random graphs*, Annals of Operations Research **1** (1984), no. 3, 183–200.

[36] Amram Meir and John W. Moon, *The expected node-independence number of random trees*, Nederlandse Akademie van Wetenschappen. Proceedings. Series A. Indagationes Mathematicae **76** (1973), 335–341. MR0345877

[37] Huy N. Nguyen and Krzysztof Onak, *Constant-time approximation algorithms via local improvements*, 2008 49th Annual IEEE Symposium on Foundations of Computer Science, 2008Oct, pp. 327–336.

[38] Ewan S. Page, *The distribution of vacancies on a line*, Journal of the Royal Statistical Society. Series B. **21** (1959), 364–374. MR0119218

[39] Ilona Palásti, *On some random space filling problems*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences **5** (1960), 353–359.

[40] Mathew D. Penrose and Aidan Sudbury, *Exact and approximate results for deposition and annihilation processes on graphs*, The Annals of Applied Probability **15** (2005), no. 1B, 853–889. MR2114992

[41] Michael E. Picollelli, *The final size of the C_ℓ-free process*, SIAM Journal on Discrete Mathematics **28** (2014), no. 3, 1276–1305. MR3252802

[42] Mustazee Rahman and Bálint Virág, *Local algorithms for independent sets are half-optimal*, The Annals of Probability **45** (2017), no. 3, 1543–1577. MR3650409

[43] Alfréd Rényi, *On a one-dimensional problem concerning random space filling.*, Publications of the Mathematical Institute of the Hungarian Academy of Sciences **3** (1958), 109–127 (Hungarian).

[44] Aidan Sudbury, *Random sequential adsorption on random trees*, Journal of Statistical Physics **136** (2009), no. 1, 51–58. MR2525226

[45] Lutz Warnke, *The C_ℓ-free process*, Random Structures & Algorithms **44** (2014), no. 4, 490–526. MR3214202

[46] Lutz Warnke, *When does the K_4-free process stop?*, Random Structures & Algorithms **44** (2014), no. 3, 355–397. MR3188600

[47] Lutz Warnke, *On Wormald's differential equation method*, arXiv e-prints (May 2019), arXiv:1905.08928, available at `https://arxiv.org/abs/1905.08928`.

[48] Nicholas C. Wormald, *Differential equations for random processes and random graphs*, The Annals of Applied Probability **5** (1995), no. 4, 1217–1235. MR1384372

[49] Nicholas C. Wormald, *The differential equation method for random graph processes and greedy algorithms*, Lectures on Approximation and Randomized Algorithms, 1999, pp. 73–155.

[50] Nicholas C. Wormald, *Models of random regular graphs*, Surveys in combinatorics, 1999 (Canterbury), 1999, pp. 239–298. MR1725006

[51] Nicholas C. Wormald, *Analysis of greedy algorithms on graphs with bounded degrees*, Discrete Mathematics **273** (2003), no. 1-3, 235–260. EuroComb'01 (Barcelona). MR2025954

A Proofs for Section 4

Proof of Claim 4.1. Let $\varepsilon \geq 0$. Since $\mu^*(r) \ll_r r!$ for every large enough r we have $\mu^*(r) \leq \varepsilon r!$. We couple $\mathcal{N}_{G_n}(r)$ and u such that the former counts the number of paths of length at most r emerging from the latter. Denote by A_n^r the event that there exists a monotone decreasing path in G_n (w.r.t. π) emerging from u of length r. Note that the probability that a given path of length r is monotone decreasing w.r.t. π is $1/r!$. Since $\mu^*(r)$ is finite, there exists $M \geq 0$ such that $\mathbb{P}[\mathcal{N}_{G_n}(r) > M] < \varepsilon$ for every large enough n. In addition, for large enough n we have $\mathbb{E}[\mathcal{N}_{G_n}(r) \wedge M] \leq 2\mu^*(r)$. Hence, for large enough n,

$$\begin{aligned}\mathbb{P}[A_n^r] &\leq \sum_{m=0}^{M} \mathbb{P}[A_n^r \mid \mathcal{N}_{G_n}(r) = m] \cdot \mathbb{P}[\mathcal{N}_{G_n}(r) = m] + \mathbb{P}[\mathcal{N}_{G_n}(r) > M] \\ &\leq \frac{1}{r!} \cdot \mathbb{E}[\mathcal{N}_{G_n}(r) \wedge M] + \varepsilon \leq 3\varepsilon. \qquad \blacktriangleleft\end{aligned}$$

Proof of Claim 4.2. Let $\varepsilon, r \geq 0$. We couple $\mathcal{N}_{G_n}(r)$ and u such that the former counts the number of paths of length at most r emerging from the latter. Note that under this coupling, $|B_{G_n}(u,r)| \leq \mathcal{N}_{G_n}(r)$. Since $\mu^*(r)$ is finite, there exists $M \geq 0$ such that $\mathbb{P}[\mathcal{N}_{G_n}(r) > M] < \varepsilon$ for every large enough n. Hence, for large enough n,

$$\begin{aligned}&\mathbb{P}[\mathrm{dist}_{G_n}(u,v) \leq r] = \mathbb{P}[v \in B_{G_n}(u,r)] \\ &\leq \sum_{m=0}^{M} \mathbb{P}[v \in B_{G_n}(u,r) \mid \mathcal{N}_{G_n}(r) = m] \cdot \mathbb{P}[\mathcal{N}_{G_n}(r) = m] + \mathbb{P}[\mathcal{N}_{G_n}(r) > M] \\ &\leq \frac{1}{n} \cdot \mathbb{E}[\mathcal{N}_{G_n}(r) \wedge M] + \varepsilon \leq \frac{M}{n} + \varepsilon \leq 2\varepsilon. \qquad \blacktriangleleft\end{aligned}$$

▶ Remark. We only used the fact that $\mathcal{N}_{G_n}(r)$ are uniformly integrable for every $r \geq 0$.

Proof of Lemma 4.3. Let $\mathcal{Q}$ be the rule of the algorithm A. The **r-truncated** version of $\mathcal{Q}$, denoted $\mathcal{Q}^r$, is defined as follows. To determine $\mathcal{Q}^r((u_1,\ldots,u_i),g)$, $\mathcal{Q}^r$ checks the value $x = \mathcal{Q}((u_1,\ldots,u_i),g)$. If $x \in \{\mathsf{T},\mathsf{F}\}$ or $\mathrm{dist}_G(u_1,x) \leq r$, $\mathcal{Q}$ returns x. Otherwise it returns F. The **r-truncated** version of the algorithm A, denoted A^r, is the exploration-decision algorithm with rule $\mathcal{Q}^r$. Note that for every v and f, $\mathrm{rad}_{\mathsf{A}^r}(v,f) \leq r$.

For a vertex $w \in \{u,v\}$, let X_w be the event "$\mathsf{A}(w,\sigma) = \mathsf{T}$", let Y_w be the event "$\mathsf{A}^r(w,\sigma) = \mathsf{T}$", and let $r_w = \mathrm{rad}_{\mathsf{A}}(w,\sigma)$. Note that $\mathbb{P}[X_w \wedge r_w \leq r] = \mathbb{P}[Y_w \wedge r_w \leq r] = \mathbb{P}[Y_w]$, thus $\mathbb{P}[X_w] = \mathbb{P}[Y_w] + o_\varepsilon(1)$. Since for x,y satisfying $\mathrm{dist}_G(x,y) \geq 3r$ we have that Y_x, Y_y are independent, it follows that $\mathbb{P}[Y_u \wedge Y_v] = \mathbb{P}[Y_u]\mathbb{P}[Y_v] + o_\varepsilon(1)$.

$$\begin{aligned}\mathbb{P}[X_u \wedge X_v] &= \mathbb{P}[X_u \wedge X_v \wedge (\max\{r_u,r_v\} \leq r)] + \mathbb{P}[X_u \wedge X_v \wedge (\max\{r_u,r_v\} > r)] \\ &= \mathbb{P}[Y_u \wedge Y_v \wedge (\max\{r_u,r_v\} \leq r)] + o_\varepsilon(1) \\ &= \mathbb{P}[Y_u \wedge Y_v] + o_\varepsilon(1) = \mathbb{P}[Y_u]\mathbb{P}[Y_v] + o_\varepsilon(1) = \mathbb{P}[X_u]\mathbb{P}[X_v] + o_\varepsilon(1). \qquad \blacktriangleleft\end{aligned}$$

B Probability Generating Functions

The goal of this section is to demonstrate how generating functions may aid solving the fundamental system of ODEs $(*)$ (and thus finding ι) for certain simple branching processes. In the following sections, we will use the notation $y_k(x)$ as in $(*)$, and omit the subscript k when the branching process has a single type.

Single Type Branching Processes

For a probability distribution $\mathbf{p} = (p_d)_{d=0}^{\infty}$, let $\mathsf{T}_{\mathbf{p}}$ be the $\mathbf{p}$-ary tree, namely, it is a (single type) branching process, for which the offspring distribution is $\mathbf{p}$. The fundamental ODE in this case is

$$y'(x) = \sum_{d=0}^{\infty} p_d \sum_{\ell=0}^{d} \binom{d}{\ell} (1-x)^{d-\ell} x^{\ell} \left(1 - \frac{y(x)}{x}\right)^{\ell} = \sum_{d=0}^{\infty} p_d (1 - y(x))^d. \qquad (4)$$

This differential equation may not be solvable, but in many important cases it is, and we will use it. Denote by $g_{\mathbf{p}}(z)$ the probability generating function (**pgf**) of $\mathbf{p}$, that is,

$$g_{\mathbf{p}}(z) = \sum_{d=0}^{\infty} p_d z^d. \qquad (5)$$

Let $\mathsf{h}_{\mathbf{p}}(x)$ be the solution to the equation

$$\int_{\mathsf{h}_{\mathbf{p}}(x)}^{1} \frac{dz}{g_{\mathbf{p}}(z)} = x. \qquad (6)$$

▷ Claim B.1. $y(x) = 1 - \mathsf{h}_{\mathbf{p}}(x)$.

Proof. Fix $x \in [0,1]$, let $\mathsf{h} = \mathsf{h}_{\mathbf{p}}(x)$ and $g(z) = g_{\mathbf{p}}(z)$. Define $\varphi : [0, \beta] \to [\mathsf{h}, 1]$, where $\beta = y^{-1}(1 - \mathsf{h})$, as follows: $\varphi(u) = 1 - y(u)$. Note that by (4),

$$\varphi'(u) = -y'(u) = -g(\varphi(u)).$$

Thus

$$x = \int_{\mathsf{h}}^{1} \frac{dz}{g(z)} = -\int_{\varphi(0)}^{\varphi(\beta)} \frac{dz}{g(z)} = -\int_{0}^{\beta} \frac{\varphi'(z)dz}{g(\varphi(z))} = \beta,$$

hence $y(x) = 1 - \mathsf{h}$. ◀

In particular, it follows from Claim B.1 that $\iota(\mathsf{T}_{\mathbf{p}}) = 1 - \mathsf{h}_{\mathbf{p}}(1)$.

Random Trees With IID Degrees

For a probability distribution $\mathbf{p} = (p_d)_{d=1}^{\infty}$, let $\mathbb{T}_{\mathbf{p}}$ be the $\mathbf{p}$-tree, namely, it is a random tree in which the degrees of the vertices are independent random variables with distribution p. We may view it as a two-type branching process, with type 0 for the root and 1 for the rest of the vertices. Let $g_{\mathbf{p}}(z)$ be the **pgf** of $\mathbf{p}$ (see (5), and note that $p_0 = 0$). The fundamental system of ODEs in this case is

$$y_0'(x) = \sum_{d=1}^{\infty} p_d \sum_{\ell=0}^{d} \binom{d}{\ell} (1-x)^{d-\ell} x^{\ell} \left(1 - \frac{y_1(x)}{x}\right)^{\ell} = \sum_{d=1}^{\infty} p_d (1 - y_1(x))^d = g_{\mathbf{p}}(1 - y_1(x)), \qquad (7)$$

and by (4),

$$y_1'(x) = \sum_{d=0}^{\infty} p_{d+1} (1 - y_1(x))^d = \frac{1}{1 - y_1(x)} \sum_{d=1}^{\infty} p_d (1 - y_1(x))^d = \frac{g_{\mathbf{p}}(1 - y_1(x))}{1 - y_1(x)}. \qquad (8)$$

Let $\mathfrak{h}_{\mathbf{p}}(x)$ be the solution to the equation

$$\int_{\mathfrak{h}_{\mathbf{p}}(x)}^{1} \frac{z dz}{g_{\mathbf{p}}(z)} = x.$$

The next claim is [13, Theorem 1].[6]

▷ Claim B.2. $y_0(x) = \frac{1}{2}\big(1 - \mathfrak{h}_{\mathbf{p}}^2(x)\big)$.

Proof. Fix $x \in [0,1]$, let $\mathfrak{h} = \mathfrak{h}_{\mathbf{p}}(x)$ and $g(z) = g_{\mathbf{p}}(z)$. Define $\varphi : [0,\beta] \to [\mathfrak{h}, 1]$, where $\beta = y_1^{-1}(1-\mathfrak{h})$, as follows: $\varphi(u) = 1 - y_1(u)$. Note that by (8),

$$\varphi'(u) = -y_1'(u) = -\frac{g(\varphi(u))}{\varphi(u)}.$$

Thus

$$x = \int_{\mathfrak{h}}^1 \frac{zdz}{g(z)} = -\int_{\varphi(0)}^{\varphi(\beta)} \frac{zdz}{g(z)} = -\int_0^\beta \frac{\varphi'(z)\varphi(z)dz}{g(\varphi(z))} = \beta,$$

hence $y_1(x) = 1 - \mathfrak{h}$. From (7) and (8) it follows that $y_0'(x) = g(\mathfrak{h}) = y_1'(x) \cdot \mathfrak{h} = -\mathfrak{h}\mathfrak{h}'$, and since $y_0(0) = 0$ it follows that $y_0(x) = \frac{1}{2}(1-\mathfrak{h}^2)$. ◀

In particular, it follows from Claim B.2 that $\iota(\mathbb{T}_{\mathbf{p}}) = \frac{1}{2}\big(1 - \mathfrak{h}_{\mathbf{p}}^2(1)\big)$.

C Proof of Theorem 6.4

1. It suffices to prove the claim for $v \in A$. First note that there exists a unique tree S_v in $T \star v$ which is not fully contained in A, and the rest of the trees are retained in the KC-transformation. The set of trees in $T' \star v$ which are not fully contained in A may be different from S_v, but they are on the same vertex set, so the result follows from subadditivity of α_n (Claim 6.1).
2. Write $|A| = a$, $|B| = b$ and $|P| = \ell + 1$. Let $A_1, \ldots, A_s$ be the trees of $T \star x$ which are fully contained in A, and denote $a_i = |A_i|$. Let $B_1, \ldots, B_t$ be the trees of $T \star y$ which are fully contained in B, and denote $b_i = |B_i|$. Let $\alpha_A = \sum_{i=1}^s \alpha_{a_i}$, $\alpha_A^+ = \sum_{i=1}^s \alpha_{1+a_i}$, $\alpha_B = \sum_{i=1}^t \alpha_{b_i}$ and $\alpha_B^+ = \sum_{i=1}^t \alpha_{1+b_i}$. Denote the vertices of P by $x = u_0, u_1, \ldots, u_\ell$. The following table summarises the values of ν in T, T' along vertices of P, in the case where $\ell \geq 3$ (similar tables can be made for the cases $\ell = 1, 2$).

	$\nu_{u_j}(T)$	$\nu_{u_j}(T')$
$j = 0$	$\alpha_A + \alpha_{b+\ell-1}$	$\alpha_A + \alpha_B + \alpha_{\ell-1}$
$j = 1$	$\alpha_A^+ + \alpha_{b+\ell-2}$	$\alpha_A^+ + \alpha_B^+ + \alpha_{\ell-2}$
$2 \leq j \leq \ell - 2$	$\alpha_{a+j-1} + \alpha_{b+\ell-j-1}$	$\alpha_{a+b+j-1} + \alpha_{\ell-j-1}$
$j = \ell - 1$	$\alpha_{a+\ell-2} + \alpha_B^+$	$\alpha_{a+b+\ell-2}$
$j = \ell$	$\alpha_{a+\ell-1} + \alpha_B$	$\alpha_{a+b+\ell-1}$

It follows (for every $\ell \geq 1$) that

$$\sum_{v \in P} (\nu_v(T') - \nu_v(T)) = \sum_{j=1}^{\ell-1} (\alpha_{a+b+j} + \alpha_j - \alpha_{a+j} - \alpha_{b+j})$$
$$= \xi_{a+b,\ell-1} + \xi_{0,\ell-1} - \xi_{a,\ell-1} - \xi_{b,\ell-1},$$

which is, by Claim 6.2, nonnegative. ◀

[6] In [13] the authors required that the the degrees of the tree are all at least 2; we do not require this here.