Local Convergence and Stability of Tight Bridge-Addable Graph Classes

Guillaume Chapuy*[1] and Guillem Perarnau[2]

1 **IRIF, UMR CNRS 8243, Université Paris-Diderot, France; and CRM, UMI CNRS 3457, Université de Montréal, Canada**
guillaume.chapuy@liafa.univ-paris-diderot.fr

2 **School of Mathematics, University of Birmingham, Birmingham, U.K.**
g.perarnau@bham.ac.uk

Abstract

A class of graphs is *bridge-addable* if given a graph G in the class, any graph obtained by adding an edge between two connected components of G is also in the class. The authors recently proved a conjecture of McDiarmid, Steger, and Welsh stating that if $\mathcal{G}$ is bridge-addable and G_n is a uniform n-vertex graph from $\mathcal{G}$, then G_n is connected with probability at least $(1+o(1))e^{-1/2}$. The constant $e^{-1/2}$ is best possible since it is reached for the class of forests.

In this paper we prove a form of uniqueness in this statement: if $\mathcal{G}$ is a bridge-addable class and the random graph G_n is connected with probability close to $e^{-1/2}$, then G_n is asymptotically close to a uniform forest in some "local" sense. For example, if the probability converges to $e^{-1/2}$, then G_n converges for the Benjamini-Schramm topology, to the uniform infinite random forest F_∞. This result is reminiscent of so-called "stability results" in extremal graph theory, with the difference that here the "stable" extremum is not a graph but a graph class.

1998 ACM Subject Classification G.2.1 Combinatorics, G.2.2 Graph Theory

Keywords and phrases bridge-addable classes, random graphs, stability, local convergence, random forests.

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2016.26

1 Introduction and Main Results

In this paper graphs are simple. A graph is labeled if its vertex set is of the form $[1..n]$ for some $n \geq 1$. An unlabeled graph is an equivalence class of labeled graphs by relabeling. Unless mentioned otherwise, graphs are labeled. A class of (labeled) graphs $\mathcal{G}$ is *bridge-addable* if given a graph G in the class, and an edge e of G whose endpoints belong to two distinct connected components, then $G \cup \{e\}$ is also in the class. McDiarmid, Steger and Welsh [11] conjectured that every bridge-addable class contains at least a proportion $(1+o(1))e^{-1/2}$ of connected graphs. This has recently been proved by the authors. In the next statement and later, we denote by $\mathcal{G}_n$ the set of graphs in $\mathcal{G}$ with n vertices, and G_n a uniform random element of $\mathcal{G}_n$.

▶ **Theorem 1** (Chapuy, Perarnau [4]). *For every $\epsilon > 0$, there exists an n_0 such that for every bridge-addable class $\mathcal{G}$ and every $n \geq n_0$, we have*

$$\mathbf{Pr}\left(G_n \textit{ is connected}\right) \geq (1-\epsilon)e^{-1/2}. \tag{1}$$

* G.C. acknowledges support from grant ANR-12-JS02-001-01 "Cartaplus" and from the City of Paris, program "Émergences 2013".

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016).
Editors: Klaus Jansen, Claire Matthieu, José D. P. Rolim, and Chris Umans; Article No. 26; pp. 26:1–26:11
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

If $\mathcal{G}$ is the class of all forests (which is bridge-addable), then Theorem 1 is asymptotically tight, since it is shown in [12] that if F_n is a uniform random forest on n vertices, then as n tends to infinity:

$$\mathbf{Pr}\left(F_n \text{ is connected}\right) \longrightarrow e^{-1/2}. \tag{2}$$

The aim of this paper is to show that, in some appropriate sense, this class is the only one for which Theorem 1 is asymptotically tight. More precisely, we will show that any addable class of graphs that comes close to achieving the constant $e^{-1/2}$ is "close" to a uniform random forest in some local sense.

▶ **Definition 2.** For any $\zeta > 0$, we say that a bridge-addable class of graphs $\mathcal{G}$ is *ζ-tight with respect to connectivity* (or simply *ζ-tight*) if there exists an n_0 such that for every $n \geq n_0$ we have

$$\mathbf{Pr}\left(G_n \text{ is connected}\right) \leq (1+\zeta)e^{-1/2}\ ,$$

where we recall that G_n is chosen uniformly at random from $\mathcal{G}_n$.

In order to state our results, we first need to introduce some notation and terminology. If H is a graph we let $|H|$ be its number of vertices. We denote by $\mathcal{U}$ the set of *unlabeled, unrooted* trees and by $\mathcal{T}$ the set of *unlabeled, rooted* trees, *i.e.* trees with a marked vertex called the root. For every tree $U \in \mathcal{U}$, we denote by $\mathrm{Aut}_u(U)$ the number of automorphisms of U, and for every rooted tree $T \in \mathcal{T}$, we denote by $\mathrm{Aut}_r(T)$ the number of automorphisms of T that fix its root. Moreover given k trees $U_1, \ldots, U_k$ in $\mathcal{U}$, we denote by $\mathrm{Aut}_u(U_1, \ldots, U_k)$ the number of automorphisms of the forest formed by disjoint copies of $U_1, \ldots, U_k$.

Given a graph H, we let $\mathit{Small}(H)$ denote the graph formed by all the components of H that are not the largest one (in case of a tie, we say that the largest component of the graph is the one with the largest vertex label among all candidates). In what follows, we will always see $\mathit{Small}(H)$ as an unlabeled graph. Given a graph G and a rooted tree $T \in \mathcal{T}$, we let $\alpha^G(T)$ be the number of pendant copies of the tree T in G. More precisely, $\alpha^G(T)$ is the number of vertices v of G having the following property: there is at least one cut-edge e incident to v, and if we remove the such cut-edge that separates v from the largest possible component, the vertex v lies in a component of the graph that is a tree, rooted at v, which is isomorphic to T. The following is classical:

▶ **Theorem 3** (see [5]). *Let F_n be a uniform random forest with n vertices. Then for any fixed unlabeled unrooted forest $\mathbf{f}$ we have as n goes to infinity:*

$$\mathbf{Pr}\big(\mathit{Small}(F_n) \equiv \mathbf{f}\big) \longrightarrow p_\infty(\mathbf{f}) := e^{-1/2}\frac{e^{-|\mathbf{f}|}}{\mathit{Aut}_u(\mathbf{f})}, \tag{3}$$

where $\equiv$ denotes unlabeled graph isomorphism. Moreover, p_∞ is a probability distribution on the set of unlabeled unrooted forests.

For any fixed rooted tree $T \in \mathcal{T}$ we have as n goes to infinity:

$$\frac{\alpha^{F_n}(T)}{n} \xrightarrow{(p)} a_\infty(T) := \frac{e^{-|T|}}{\mathit{Aut}_r(T)}, \tag{4}$$

where (p) indicates convergence in probability.

Our main result says that, for bridge-addable classes, if we have an approximate version of (2), then we also have an approximate version of (3) and (4). In the next statement and everywhere in the paper, the constants $\epsilon, \eta, \rho, \nu, \zeta$ are implicitly assumed (in addition to other written quantifications or assumptions) to be positive and smaller than c where c is a small, absolute, constant.

▶ **Theorem 4** (Main result). *For every $\epsilon, \eta > 0$, there exists a $\zeta > 0$ and an n_0 such that for every ζ-tight bridge-addable class $\mathcal{G}$ and $n \geq n_0$, the following holds:*

(i) *The small components of G_n are close to those of a large random forest, in the sense that for every unrooted unlabeled forest $\mathbf{f}$ we have:*

$$\Big|\mathbf{Pr}\big(\mathit{Small}(G_n) \equiv \mathbf{f}\big) - p_\infty(\mathbf{f})\Big| < \epsilon.$$

(ii) *The statistics of pendant trees in G_n are close to those of a large random forest, in the sense that:*

$$\mathbf{Pr}\left(\forall T \in \mathcal{T} : \left|\frac{\alpha^{G_n}(T)}{n} - a_\infty(T)\right| < \eta\right) > 1 - \epsilon.$$

▶ Remark. It is easy to see (up to adapting the dependence of ζ in ϵ, η) that we can replace *i*) by:

(i') The total variation distance between the law of $\mathit{Small}(G_n)$ and the probability law p_∞ is at most ϵ.

Similarly we could replace *ii*) by:

(ii') The total variation distance between the measure $\alpha^{G_n}(\cdot)/n$ and the probability law $a_\infty(\cdot)$ (both are measures on $\mathcal{T}$) is at most η with probability at least $1 - \epsilon$.

▶ Remark. Our main result, Theorem 4, can be viewed both as a *unicity result* (since it states that in the limit, and through the lens of local observables, the class of forests is the only one to reach the optimum value $e^{-1/2}$) and as a *stability result* (since it also states that the only classes than come *close* to the extremal value $e^{-1/2}$ are *close* to forests, again through local observables of random graphs). Here we use the terminology "stability result" on purpose, by analogy with the field of extremal graph theory. Indeed the study of graphs that come *close* to achieving extremal properties is a classical topic in this field. *Stability results*, pioneered in the papers [8, 7, 6, 13], show that in many cases, the graphs that are close to being extremal have a structure close to the actual extremal graphs, in some quantifiable sense. Our main result suggests that the question of stability of extremal graph classes, with respect to appropriate graph limit topologies (here, local convergence), should be further examined.

Before going into the proof of the theorem, let us look at some closely related statements and corollaries. Call a bridge-addable class $\mathcal{G}$ *tight* is it is ζ-tight for any ζ, that is to say:

$$\mathbf{Pr}(G_n \text{ is connected}) \to e^{-1/2}.$$

Then we have the following consequence of Theorem 4. Note that it is weaker (it is a unicity, but not a stability result).

▶ **Theorem 5** (Convergence of local statistics in tight graph classes). *Let $\mathcal{G}$ be a tight bridge-addable class of graphs. Then, when n goes to infinity, $\mathit{Small}(G_n)$ converges in total variation distance to the probability law $p_\infty(.)$ given by (3), that is:*

$$d_{TV}\big(\mathit{Small}(G_n), p_\infty\big) \longrightarrow 0.$$

Moreover, for any rooted tree $T \in \mathcal{T}$, the proportion $\frac{\alpha^{G_n}(T)}{n}$ of local pendant copies of the tree T converges in probability to the deterministic constant $a_\infty(T)$ given by (4):

$$\frac{\alpha^{G_n}(T)}{n} \xrightarrow{(p)} a_\infty(T).$$

Theorem 5 states that, from the point of view of statistics of pendant trees and of non-largest components, tight classes are indistinguishable from random forests in the limit. Let us develop in this direction. Let V_n be a uniform random vertex in G_n. Then for a given $T \in \mathcal{T}$, conditionally to G_n, the quantity $\alpha^{G_n}(T)/n$ is the probability that from V_n hangs a copy of the tree T. Readers familiar with the Benjamini-Schramm (BS) convergence of graphs [3] will note the similarity with this notion. If (G, x) and (H, y) are two rooted graphs, define the BS-distance $d_{BS}((G, x); (H, y))$ as 2^{-K} where K is the largest integer such that the balls of radius K in (G, x) and (H, y) are isomorphic (as rooted graphs). This distance (also called the *ball distance*, see [10]) defines a topology (in fact, a Polish space) on the set of rooted graphs, and enables us to talk about convergence in distribution of random rooted graphs, in the BS-sense. An equivalent definition of this convergence is the following: a sequence of random rooted graphs (H_n, x_n) converges to (H_∞, x_∞) if and only if for any rooted graph (H, x) of radius r, the probability that the ball of radius r in (H_n, x_n) is isomorphic to (H, x) converges to the probability of the same event in (H_∞, x_∞).

It is easy to see (for example using generating functions, see [5]) that if F_n is a uniform random forest on n vertices rooted at a random uniform vertex V_n, then

$$(F_n, V_n) \to (F_\infty, V_\infty)$$

in distribution in the BS-sense, where (F_∞, V_∞) is the "infinite uniform random forest". Namely, (F_∞, V_∞) can be constructed as follows: consider a semi-infinite path, starting at a vertex V_∞, and identify each vertex of this path with the root of an independent Galton-Watson tree of offspring distribution $Poisson(1)$.

In our context, passing from pendant trees to balls is an easy task, and one can deduce the following from Theorem 5.

► **Corollary 6** (Local convergence of tight graph classes). *Let $\mathcal{G}$ be a tight bridge-addable graph class. Let G_n be a uniform random graph from $\mathcal{G}_n$ and let V_n be a uniform random vertex of G_n. Then (G_n, V_n) converges to (F_∞, V_∞) in distribution in the Benjamini-Schramm sense.*

The purpose of stating Corollary 6 is to illustrate the link between our local observables and the BS topology, but we could have stated stronger intermediate results. For example Corollary 6 uses only the second part of Theorem 5, and says nothing about small connected components. In fact, it is true that for tight classes, the pair $((G_n, V_n), Small(G_n))$ converges in distribution to $(F_\infty, V_\infty) \otimes p_\infty$ for the product of the BS and the total variation topologies. This follows easily from our proofs.

Also note the last corollary (and the other statements of the same kind that have just mentioned) is of a much weaker nature than Theorem 5. Indeed, Theorem 5 asserts that with high probability, conditional on the random graph G_n, the graph G_n is similar to a random forest when viewed from a random vertex, whereas Corollary 6 is an unconditioned statement that averages both over the graph G_n and the vertex V_n. It is possible to formulate a version Theorem 5 in terms of the BS convergence as follows. Let μ_{G_n} be the law, conditional on G_n, of the random rooted graph (G_n, V_n) where V_n is a uniform vertex of G_n (then μ_{G_n} is a random probability measure on the set of rooted graphs). Then it follows easily from Theorem 5 that if $\mathcal{G}$ is a tight bridge-addable and n is large enough, μ_{G_n} converges in probability to the *deterministic* probability measure μ_∞, defined as the law of (F_∞, V_∞). The underlying distance for the convergence in probability is the Skorokhod distance induced by the BS distance on the set of probability measures on rooted graphs. We will not give more details on these questions, since the related considerations of convergence of probability measures would lead us too far from our main prospect.

▶ Remark. Our main theorem asserts that tight bridge addable classes are “locally similar” to random forests in some precise sense. However, they can be very different from some other perspective. For example, consider the set $\tilde{\mathcal{F}}_n$ of graphs on $[1..n]$ defined as follows: $\tilde{\mathcal{F}}_n$ contains the graph in which all edges linking vertices in $[1..\lceil n^{2/3}\rceil]$ are present and all other vertices are isolated, and $\tilde{\mathcal{F}}_n$ is the smallest bridge-addable class containing this graph. In other words, $\tilde{\mathcal{F}}_n$ is the set of graphs inducing a clique on $[1..\lceil n^{2/3}\rceil]$, and such that contracting this clique gives a forest. Then $\tilde{\mathcal{F}} = \cup_{n\geq 1}\tilde{\mathcal{F}}_n$ is a bridge-addable class, and it is easy to see that it is tight, so our main results apply. However one can argue that the random graph $\tilde{F}_n$ in $\tilde{\mathcal{F}}_n$ is *very* different from a random forest in several senses: first, it has $\Theta(n^{4/3})$ edges whereas a forest has linearly many. Second, with probability $1 - O(n^{-1/3})$ an edge taken uniformly at random from $\tilde{F}_n$ belongs to a clique of size $\lceil n^{2/3}\rceil$, which is very different from what happens in a forest. This last point does not contradict our results, but only recalls that it is important here to think of locality as a measure of what happens around “typical vertices” and not “typical edges”.

▶ Remark. One can modify the example of the previous remark by replacing the clique of size $\lceil n^{2/3}\rceil$ by a path of length $\lceil n^{2/3}\rceil$. One thus obtains a tight bridge-addable class of graphs, in which the *diameter* of the largest component is of order $\Theta(n^{2/3})$, which is again very different from a random forest in which the giant tree has diameter $\Theta(\sqrt{n})$ with high probability. In both examples, the function $n^{2/3}$ plays no special role and may be replaced by $n^{1-\epsilon}$ for any $\epsilon > 0$.

We conclude this list of results with a simpler statement, that does not require the full strength of our main theorems (it is a relatively easy consequence of the results of [4]).

▶ **Theorem 7.** *Let $\mathcal{G}$ be a tight bridge-addable class and G_n a uniform random graph from $\mathcal{G}_n$. Then for any $k \geq 0$, we have*

$$\mathbf{Pr}\left(G_n \textit{ has } k+1 \textit{ connected components}\right) \longrightarrow e^{-1/2}\,\frac{2^{-k}}{k!}.$$

In other words, the number of connected components of G_n converges in distribution to Poisson$(1/2)$.

Structure of this abstract. In this abstract, we will present the main steps of the proof of Theorem 4, refereeing the reader to the full version [5] for complete proofs, including several easy results on enumeration of forests and on random forests.

2 Theorem 4 for bridge-addable classes of forests

2.1 Number of components in bridge-addable graph classes

We first introduce some notation, following [4]. For a bridge-addable class of graphs $\mathcal{G}$ and for $i \geq 1$, we note $\mathcal{G}_n^{(i)}$ the set of n-vertex graphs in $\mathcal{G}$ having i connected components. An elegant double-counting argument going back to [11] asserts that for all $i \geq 1$, and $n \geq 1$ we have:

$$i \cdot \left|\mathcal{G}_n^{(i+1)}\right| \leq \left|\mathcal{G}_n^{(i)}\right|. \qquad (5)$$

This statement follows by double-counting the edges of an auxiliary bipartite graph on the vertex set $\mathcal{G}_n^{(i)} \uplus \mathcal{G}_n^{(i+1)}$, where two graphs G, H are linked by an edge if and only if one can be obtained from the other by adding a bridge: on the one hand, an element of $\mathcal{G}_n^{(i+1)}$ has degree at least $i(n-i)$ in this auxiliary graph, since $\mathcal{G}$ is bridge-addable; on the other hand, an element of $\mathcal{G}_n^{(i)}$ has degree at most $(n-i)$ (which is the maximum number of cut-edges in a graph with i connected components and n vertices). Thus (5) follows. The main achievement of the paper [4] was to improve this bound by roughly a factor $\frac{1}{2}$, asymptotically.

▶ **Lemma 8** (Proposition 4 in [4]). *For every η and every k there exists an n_0 such that for every bridge-addable class $\mathcal{G}$, every $n \geq n_0$ and every $i \leq k$, one has:*

$$i|\mathcal{G}_n^{(i+1)}| \leq \left(\frac{1}{2}+\eta\right)|\mathcal{G}_n^{(i)}| \ . \tag{6}$$

The following lemma, which follows relatively easily from Theorem 1, provides a converse inequality to (6) for ζ-tight classes. Note that it implies Theorem 7.

▶ **Lemma 9.** *For every η and every k there exists a ζ and an n_0 such that for every ζ-tight bridge-addable class $\mathcal{G}$, every $n \geq n_0$ and every $i \leq k$, one has*

$$\left(\frac{1}{2}-\eta\right)|\mathcal{G}_n^{(i)}| \leq i|\mathcal{G}_n^{(i+1)}| \leq \left(\frac{1}{2}+\eta\right)|\mathcal{G}_n^{(i)}| \ .$$

2.2 Partitioning the graph class into highly structured subclasses

Balister, Bollobás and Gerke [2, Lemma 2.1] proposed an elegant construction that reduces the statement of Theorem 1 to the case where all graphs in $\mathcal{G}$ are forests. As we will see in the next section, their idea can be adapted to the present context. We will therefore start by proving Theorem 4 for classes $\mathcal{G}$ composed by forests:

Throughout the rest of Section 2 we will assume that all graphs in $\mathcal{G}$ are forests.

We will first focus on the graphs in $\mathcal{G}_n$ that have either one or two connected components, and, in view of this, we use the shorter notation $\mathcal{A}_n := \mathcal{G}_n^{(1)}$ and $\mathcal{B}_n := \mathcal{G}_n^{(2)}$. We now introduce a partitioning of $\mathcal{A}_n$ and $\mathcal{B}_n$ in terms of some local statistics, which requires the following set-up, that is modeled on [4, proof of Prop 3]. Here ϵ and k_* are two constants, whose value may vary along the course of the paper, that will *in fine* be chosen very small and very large, respectively:

- $\mathcal{U}_\epsilon$ is the set of unrooted trees of order at most $\lceil \epsilon^{-1} \rceil$: $\mathcal{U}_\epsilon := \{U \in \mathcal{U},\ |U| \leq \lceil \epsilon^{-1} \rceil\}$.
- $\mathcal{T}_*$ is the set of rooted trees of order at most k_*: $\mathcal{T}_* := \{T \in \mathcal{T},\ |T| \leq k_*\}$.

More generally for any given $\ell \geq 1$ we will use the notation $\mathcal{T}_{\leq \ell}, \mathcal{U}_{\leq \ell}$ to denote the set of rooted (resp., unrooted) trees of order at most ℓ, so that $\mathcal{U}_\epsilon = \mathcal{U}_{\leq \lceil \epsilon^{-1} \rceil}$ and $\mathcal{T}_* = \mathcal{T}_{\leq k_*}$.

Roughly speaking, we will use elements of $\mathcal{U}_\epsilon$ and $\mathcal{T}_*$ as "test graphs" to measure the shape of small components of G_n and the number of pending subtrees of G_n, respectively. For $\ell \geq 1$ we write $\mathcal{E}_\ell = \{0, \ldots, n-1\}^{\mathcal{T}_{\leq \ell}}$, and we will be particularly concerned with the set $\mathcal{E}_* := \mathcal{E}_{k_*}$, namely the set of integer vectors with one coordinate for each "test tree" in $\mathcal{T}_*$. For $\alpha \in \mathcal{E}_*$ and $w \geq 1$ (width) we define the *box* $[\alpha]^w \subset \mathcal{E}_*$ and its *q-neighborhood* $[\alpha]_q^w$ as the parallelepipeds:

$$[\alpha]^w := \{\alpha' \in \mathcal{E}_* : \ \forall T \in \mathcal{T}_*,\ \alpha(T) \leq \alpha'(T) < \alpha(T) + w\} \ ,$$
$$[\alpha]_q^w := \{\alpha' \in \mathcal{E}_* : \ \forall T \in \mathcal{T}_*,\ \alpha(T) - q \leq \alpha'(T) < \alpha(T) + w + q\} \ .$$

Finally, if $\mathcal{S}_n$ is a set of graphs (where the letter $\mathcal{S}$ could be $\mathcal{A}$, $\mathcal{B}$, and also carry other decorations), we let $\mathcal{S}_{n,[\alpha]^w}$ be the set of graphs G in $\mathcal{S}$ such that $\alpha^G(T) \in [\alpha]^w$ for all $T \in \mathcal{T}_*$, and we use the same notation with $[\alpha]_q^w$. Also, for every forest $\{U_1, \ldots, U_k\}$, we denote by ${\mathcal{S}_n}^{\{U_1,\ldots,U_k\}}$ the set of graphs G in $\mathcal{S}_n$ such that $Small(G)$ is isomorphic to $\{U_1, \ldots U_k\}$. In the case of graphs with two connected components, we just use the notation $\mathcal{S}_n^U$ for $\mathcal{S}_n^{\{U\}}$, where $U \in \mathcal{U}$.

2.3 Good and bad boxes

The main concern of the paper [4] was to obtain a version of the double-counting argument of Section 2.1 that is *local* in the sense that it relates cardinalities of graphs corresponding to fixed boxes. Given ϵ (hence $\mathcal{U}_\epsilon$) and $\mathcal{T}_*$, [4, Lemma 16] asserts that there exist integers K, w, n_0 (independent of $\mathcal{G}$) and a set of K disjoint boxes of width w in $\mathcal{E}_*$, noted $\{[\beta_i]^w, 1 \leq i \leq K\}$, such that for $n \geq n_0$ and $U \in \mathcal{U}_\epsilon$ we have:

$$\sum_{i=1}^{K} |\mathcal{B}^U_{n,[\beta_i]^w}| \geq (1-\epsilon)|\mathcal{B}^U_n|, \tag{7}$$

and such that for each $1 \leq i \neq j \leq K$, $[\beta_i]^w_q \cap [\beta_j]^w_q = \emptyset$, where $q = q_\epsilon := \lceil \epsilon^{-1} \rceil$. In other words, these boxes are $2q$-apart from each other, and yet capture a proportion at least $(1-\epsilon)$ of the set $\mathcal{B}^U_n$ for each $U \in \mathcal{U}_\epsilon$. We will also use the fact (implicit in [4]) that the $[\beta_i]^w_q$ partition the set $\mathcal{E}_*$.

In the present paper, one of the main tasks consists in showing that the global estimates obtained in [4] can be "lowered" down to boxes for ζ-tight classes. For $\gamma, \epsilon > 0$, we say that a box $[\alpha]^w$ is (γ, ϵ)-*good* (or simply *good*) if the two following conditions hold:

(i) $|\mathcal{B}_{n,[\alpha]^w}| \geq \left(\frac{1}{2} - \gamma\right) \cdot |\mathcal{A}_{n,[\alpha]^w_q}|$

(ii) $\sum_{U \notin \mathcal{U}_\epsilon} |\mathcal{B}^U_{n,[\alpha]^w}| < \gamma |\mathcal{B}_{n,[\alpha]^w}|$.

Note that Property i) is a local version of the first inequality of Lemma 9, while Property ii) ensures that the number of graphs in sets that we do not control, is small (as it happens in a global scale). Hence good boxes are, in some sense, boxes that realize the tightness property *locally*. We will be interested in the boxes among the $[\beta_i]$ that are (γ, ϵ)-good:

$$\mathrm{Good}_{\gamma,\epsilon} := \{i \in [1..K] : [\beta_i] \text{ is } (\gamma, \epsilon)\text{-good}\} .$$

An important step in the proof of Theorem 4 is the following result:

▶ **Lemma 10.** *For every γ and every η, if $\epsilon < \epsilon_0(\gamma, \eta)$ and if $k_* \geq k_0(\epsilon)$, then there exist ζ and an n_0 such that for every ζ-tight bridge-addable class $\mathcal{G}$ and every $n \geq n_0$, one has*

$$\frac{\sum_{i \notin Good_{\gamma,\epsilon}} |\mathcal{A}_{n,[\beta_i]^w_q}|}{|\mathcal{A}_n|} < \eta \ , \quad \frac{\sum_{i \notin Good_{\gamma,\epsilon}} |\mathcal{B}_{n,[\beta_i]^w}|}{|\mathcal{B}_n|} < \eta \ .$$

From this lemma, one deduces, after a lengthy and technical proof, the following result (which is a first version of Theorem 4 for subclasses of forests and for $\mathbf{f}$ being a tree). We define the set of vectors in $\mathcal{E}_\ell$ that are δ-close from the distribution p_∞ (recall that for $T \in \mathcal{T}$, $p_\infty(T) = \frac{e^{-|T|}}{\mathrm{Aut}_r(T)}$),

$$\Xi(\delta, \ell) = \left\{ \beta \in \mathcal{E}_\ell : \left| \frac{\beta(T)}{n} - p_\infty(T) \right| < \delta, \text{ for every } T \in \mathcal{T}_{\leq \ell} \right\} .$$

▶ **Proposition 11.** *For every θ_1 and every $U \in \mathcal{U}$, there exist a $\zeta > 0$ and an n_0 such that for every ζ-tight class $\mathcal{G}$ of forests and every $n \geq n_0$, one has*

$$\left| \frac{|\mathcal{B}^U_n|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{Aut_u(U)} \right| < \theta_1 \ .$$

Moreover, for every θ_1, every δ, every ℓ and every $U \in \mathcal{U}$, there exist a $\zeta > 0$ and an n_0 such that for every ζ-tight class $\mathcal{G}$ of forests and every $n \geq n_0$, one has

$$\left| \frac{\sum_{\beta \in \Xi(\delta,\ell)} \left|\mathcal{B}^U_{n,\beta}\right|}{|\mathcal{B}^U_n|} - 1 \right| < \theta_1 \ .$$

Main ideas of the proof. The proof itself is long (see [5]). The main idea is to go back to the optimization problem of [4] and show that, in order for the class to be ζ-tight, one has to be close to the extremal point of that problem. Roughly speaking, [4] provides some inequalities between the local ratios $|\mathcal{B}_{n,[\alpha]^w}|/|\mathcal{A}_{n,[\alpha]_q^w}|$, where $[\alpha]^w$ is a box, in terms of an optimization problem for the quantities $|\mathcal{B}^U_{n,[\alpha]^w}|/|\mathcal{A}_{n,[\alpha]_q^w}|$ for $U \in \mathcal{U}_\epsilon$. By the preceding lemma, if a class is ζ-tight for ζ small enough, we can almost cover the space $\mathcal{E}_*$ with boxes that capture most of the mass of the sets $\mathcal{A}_n$ and $\mathcal{B}_n$, and such that each box is good. By Property i) of good boxes, a good box is close to reaching the ratio $e^{-1/2}$ which is the optimum in the optimization problem of [4]. The main task of the proof is then to go back to the optimization problem of [4] and quantify the stability of its extrema. After a tedious technical work, one finds that, provided ϵ, k_* are respectively small and large enough, the optimization problem is sufficiently stable to conclude that most of the mass in the sets $\mathcal{B}^U_n$ is concentrated around the unique extreme of the optimization problem. More precisely, one finds that the set $\mathcal{B}^U_n$ has most of its mass in subsets $\mathcal{B}^U_{n,[\alpha]^w}$ such that $\alpha(T)$ is close to $a_\infty(T)$ for each $T \in \mathcal{T}_*$, and that for such subsets the ratios $|\mathcal{B}^U_{n,[\alpha]^w}|/|\mathcal{A}_{n,[\alpha]_q^w}|$ are close to $e^{-1/2}\frac{e^{-|U|}}{\mathrm{Aut}_u(U)}$. The result can then be extended to the ratios $|\mathcal{B}^U_n|/|\mathcal{A}_n|$ by an averaging argument, and to the ratios $|\mathcal{B}^U_n|/|\mathcal{G}_n|$ since for ζ-tight classes $|\mathcal{A}_n|/|\mathcal{G}_n|$ is close to $e^{-1/2}$.

We refer the reader again to the full version of the article [5] for the many subtleties hidden in this seemingly simple proof by tightness arguments. ◀

The proof of the previous proposition, although it involves a lot of work, is the part of the present paper that is conceptually more relying on [4]. The next result, that is equivalent to our main theorem (for classes of forests) relies on arguments of a different nature:

▶ **Theorem 12.** *For every $k \geq 1$, every θ_k and every $U_1, \ldots, U_k \in \mathcal{U}$, there exist a $\zeta > 0$ and an n_0 such that for every ζ-tight class $\mathcal{G}$ of forests and every $n \geq n_0$, one has*

$$\left| \frac{\left|\mathcal{G}_n^{k+1,\{U_1,\ldots,U_k\}}\right|}{|\mathcal{G}_n|} - e^{-1/2}\frac{e^{-\sum_{i=1}^k |U_i|}}{Aut_u(U_1,\ldots,U_k)} \right| < \theta_k \ . \tag{8}$$

Moreover, for every $k, \ell \geq 1$, every θ_k, δ and every $U_1, \ldots, U_k \in \mathcal{U}$, there exist a $\zeta > 0$ and an n_0 such that for every ζ-tight class $\mathcal{G}$ of forests and every $n \geq n_0$, one has

$$\left| \frac{\sum_{\beta \in \Xi(\delta,\ell)} \left|\mathcal{G}_{n,\beta}^{k+1,\{U_1,\ldots,U_k\}}\right|}{\left|\mathcal{G}_n^{k+1,\{U_1,\ldots,U_k\}}\right|} - 1 \right| < \theta_k \ . \tag{9}$$

Main ideas of the proof. The proof uses induction on k, with the base case given by Proposition 11. The main idea is that if $\mathcal{G}$ is bridge-addable and $k \geq 2$, and if $\{U_1, \ldots, U_{k-1}\}$ is a forest composed by k trees on a subset W of $[1..n]$, we can form a bridge-addable class by looking at all graphs G in $\mathcal{G}_n$ such that W induces a union of connected components of G, given by $\{U_1, \ldots, U_{k-1}\}$. Roughly speaking, connected graphs in this new class correspond to graphs in $\mathcal{G}_n^{(k)}$ while graphs with two connected components correspond to graphs in $\mathcal{G}_n^{(k+1)}$. Therefore, by applying Proposition 11 to this class, we may obtain information on the ratios of cardinalities of these sets. Moreover, the induction hypothesis ensures that we have a very precise structural information on the typical graphs in $\mathcal{G}_n^{(k)}$. The full proof is given in [5]. ◀

The last theorem implies Theorem 4 for bridge-addable classes of forests. Indeed, the first part of it implies *i*): by selecting ℓ large enough, we use (8) to control $\left|\mathcal{G}_n^{k+1,\{U_1,\ldots,U_k\}}\right|$

for all $k \le \ell$ and $U_1, \dots, U_k \in \mathcal{T}_{\le \ell}$, and since ℓ is large, the set $\mathcal{G}_n^{k+1,\{U_1,\dots,U_k\}}$ is of negligible size for the rest for forest with more than ℓ vertices or including some tree with order larger than ℓ. In a similar way, we can use (9) to prove the statement *ii*).

3 From classes of forests to classes of graphs

In this section we extend the results of the previous section (where we obtained Theorem 4 for classes of forests) to general bridge-addable classes, concluding the proof of Theorem 4.

The method of proof of Theorem 12 will allow us to derive a statement about *removable edges* which will be crucial to transfer the result from forest to general classes. We say that an edge in a graph $G \in \mathcal{G}$ is *removable* if the graph $G' = G \setminus e$ is in $\mathcal{G}$. For a class $\mathcal{H} \subseteq \mathcal{G}$ and a rooted tree $T \in \mathcal{T}$, we define $p(\mathcal{H}, T)$ to be the probability that given a uniform random graph $H \in \mathcal{H}$, and a uniform random pendant copy of T in H, the graph H' obtained by deleting the edge that connects the pendant copy of T to the rest of the graph belongs to $\mathcal{G}$ (and not only to $\mathcal{H}$). In other words, $p(\mathcal{H}, T)$ is the average over all graphs in $\mathcal{H}$ of the proportion of pendant copies of T that are attached using a removable edge. This notion is inspired by bridge-alterable classes, for which $p(\mathcal{H}, T) = 1$, for every $\mathcal{H} \subseteq \mathcal{G}$ and every $T \in \mathcal{T}$ [1, 9]. We do an slight abuse of notation by writing $p(G, T)$ for $p(\{G\}, T)$, for each $G \in \mathcal{G}$. Also, in the cases where $p(G, T)$ is not well-defined (that is, if G has no pendant copy of T), we interpret the probability as 1.

The next theorem says that ζ-tight bridge-addable classes *of graphs* (not only forests) are essentially also bridge-alterable.

▶ **Theorem 13.** *For every θ, there exist a ζ, an n_0 and an ℓ such that for every ζ-tight bridge-addable class $\mathcal{G}$ and $n \ge n_0$, we have that if G_n is a graph chosen uniformly at random in $\mathcal{G}_n$, and v is a vertex chosen uniformly at random in G_n, the following holds with probability at least $1 - \theta$: v is connected to G_n through a removable edge and the corresponding pendant tree has order at most ℓ. In particular, $p(\mathcal{G}_n, T) \ge 1 - \theta$, for every rooted tree $T \in \mathcal{T}_{\le \ell}$.*

We first sketch how the theorem is proved for classes of forests. Fix $k \ge 1$ and $U_1, \dots, U_k \in \mathcal{U}$. Let $T_1, \dots, T_s$ be the possible rooted versions of U_k. By (8), the ratio between $|\mathcal{G}_n^{k+1,\{U_1,\dots,U_k\}}|$ and $|\mathcal{G}_n^{k,\{U_1,\dots,U_{k-1}\}}|$ is essentially fixed. Moreover, by (9), we know that a typical graph $G \in \mathcal{G}_n^{k,\{U_1,\dots,U_{k-1}\}}$ has $\sum_{i=1}^{s} \alpha^G(T_i) \approx \sum_{i=1}^{s} \frac{e^{-|T_i|}}{\mathrm{Aut}_r(T_i)} = \frac{|U_k| e^{-|U_k|}}{\mathrm{Aut}_u(U_k)}$ pendant trees such that, if the edge from where they hang is removable, then they give rise to a graph in $\mathcal{G}_n^{k+1,\{U_1,\dots,U_k\}}$. It turns out that the only way we can realize the desired ratio, is by having almost every such edge removable. Since the choice of k and $U_1, \dots, U_k$ is arbitrary, we are done.

To transfer the result of Theorem 13 from classes of forests to classes of graphs, we use a nice argument introduced in [2]. Every graph admits a unique decomposition into 2-blocks, joined by edges in a tree-like fashion. Consider the partition of $\mathcal{G}_n$ into subclasses $\mathcal{H}_1, \mathcal{H}_2, \dots$ such that every two graphs H and H' in the same subclass, have the same 2-blocks. Since every subclass $\mathcal{H}_i$ is bridge-addable, one can use an averaging argument to show that if $\mathcal{G}$ is ζ-tight, then there exists ζ' such that if n is large enough, then at least $(1 - \zeta')|\mathcal{G}_n|$ graphs are in subclasses $\mathcal{H}_i$ that are ζ'-tight. Let $\mathcal{H}$ be one of such ζ'-tight subclasses of $\mathcal{G}_n$ and let $\mathcal{F}_{\mathcal{H}}$ be the class of forests obtained by selecting the same spanning tree for each 2-block of the graphs in $\mathcal{H}$. Since $\mathcal{F}_{\mathcal{H}}$ is also a ζ'-tight bridge-addable class of forests, we can apply Theorem 13 to it, and the conclusion of the theorem naturally transfers from $\mathcal{F}_{\mathcal{H}}$ to $\mathcal{H}$. Since most of the graphs are in ζ'-tight bridge-addable classes, the statement of Theorem 13 also holds for general classes of graphs. The full proof is presented in [5].

Our next goal is to show that not only the pendant graphs obtained when deleting a removable edge have bounded size, as Theorem 13 ensures, but in fact, they are pendant trees. For every class $\mathcal{G}_n$ and every $t \geq 1$, given G_n chosen uniformly at random from $\mathcal{G}_n$ and v chosen uniformly at random from the vertices of G_n, let $q(\mathcal{G}_n, t)$ be the probability that v is connected to G_n via a removable edge and the corresponding pendant graph is a tree of order at most t. Observe that if $\mathcal{G}$ is subclass of forests, Theorem 13 implies that for every $\theta > 0$, and under some technical conditions, there exists some ℓ such that $q(\mathcal{G}_n, \ell) \geq 1 - \theta$. Next lemma shows that the same holds for general classes of graphs.

▶ **Lemma 14.** *For every $\vartheta > 0$, there exist a ζ, an n_0 and a t, such that if $\mathcal{G}$ is a ζ-tight class and $n \geq n_0$, then $q(\mathcal{G}_n, t) \geq 1 - \vartheta$.*

As before, we split the class $\mathcal{G}_n$ into subclasses $\mathcal{H}_1, \mathcal{H}_2, \ldots$ according to the 2-blocks. Recall that there exists a ζ' such that at least $(1 - \zeta')|\mathcal{G}_n|$ graphs are in subclasses $\mathcal{H}_i$ that are ζ'-tight. Let $\mathcal{H}$ be one of such ζ'-tight subclasses of $\mathcal{G}_n$ and let $\mathcal{F}_{\mathcal{H}}$ the corresponding class of forests. By Theorem 13, if ζ' is small enough and, n and t are large enough (t plays the role of ℓ), then the probability that a random vertex v in a random graph F_n from $\mathcal{F}_{\mathcal{H}}$ connects to F_n through a removable edge and disconnects a pendant tree T_v of order at most t, is close to 1. If this is the case, by construction of $\mathcal{F}_{\mathcal{H}}$, this edge is also a removable cut-edge in the graph in $\mathcal{H}$ that corresponds to F_n.

It remains to show that, with probability close to 1, the pullback H_v of the tree T_v in the original graph in $\mathcal{H}$ is a also tree. This is done by applying Theorem 13 again but using now ℓ much larger than t, which shows that the proportion of vertices that are linked to the rest of the graph by a removable edge is very close to 1, and by noticing that, if H_v is not a tree, then at least one vertex of H_v does not have this property. Details are given in [5].

The last lemma is the key point in proving Theorem 4 for classes that are not forests. Indeed, Theorem 4 now follows relatively easily from Theorem 12 and Lemma 14 (see [5] for a detailed proof).

References

1 Louigi Addario-Berry, Colin McDiarmid, and Bruce Reed. Connectivity for bridge-addable monotone graph classes. *Combin. Probab. Comput.*, 21(6):803–815, 2012.

2 Paul Balister, Béla Bollobás, and Stefanie Gerke. Connectivity of addable graph classes. *J. Combin. Theory Ser. B*, 98(3):577–584, 2008.

3 Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

4 Guillaume Chapuy and Guillem Perarnau. Connectivity in bridge-addable graph classes: the McDiarmid-Steger-Welsh conjecture. *Extended abstract in the proceedings of SODA 2016. Long version submitted for publication, see arXiv:1504.06344.*, 2015.

5 Guillaume Chapuy and Guillem Perarnau. Local convergence and stability of tight bridge-addable graph classes. *In preparation.*, 2016.

6 Paul Erdős. On some new inequalities concerning extremal properties of graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 77–81, 1966.

7 Paul Erdős. Some recent results on extremal problems in graph theory. *Results, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York*, pages 117–123, 1967.

8 Paul Erdős and M Simonovits. A limit theorem in graph theory. In *Studia Sci. Math. Hung.* Citeseer, 1966.

9 Mihyun Kang and Konstantinos Panagiotou. On the connectivity of random graphs from addable classes. *J. Combin. Theory Ser. B*, 103(2):306–312, 2013.

10 László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

11 Colin McDiarmid, Angelika Steger, and Dominic J. A. Welsh. Random graphs from planar and other addable classes. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 231–246. Springer, Berlin, 2006.

12 Alfréd Rényi. Some remarks on the theory of trees. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 4:73–85, 1959.

13 Miklós Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319, 1968.