

# Rational exponents in extremal graph theory

Boris Bukh\*

David Conlon†

## Abstract

Given a family of graphs  $\mathcal{H}$ , the extremal number  $\text{ex}(n, \mathcal{H})$  is the largest  $m$  for which there exists a graph with  $n$  vertices and  $m$  edges containing no graph from the family  $\mathcal{H}$  as a subgraph. We show that for every rational number  $r$  between 1 and 2, there is a family of graphs  $\mathcal{H}_r$  such that  $\text{ex}(n, \mathcal{H}_r) = \Theta(n^r)$ . This solves a longstanding problem in the area of extremal graph theory.

## 1 Introduction

Given a family of graphs  $\mathcal{H}$ , another graph  $G$  is said to be  $\mathcal{H}$ -free if it contains no graph from the family  $\mathcal{H}$  as a subgraph. The extremal number  $\text{ex}(n, \mathcal{H})$  is then defined to be the largest number of edges in an  $\mathcal{H}$ -free graph on  $n$  vertices. If  $\mathcal{H}$  consists of a single graph  $H$ , the classical Erdős–Stone–Simonovits theorem [8, 9] gives a satisfactory first estimate for this function, showing that

$$\text{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right) \binom{n}{2},$$

where  $\chi(H)$  is the chromatic number of  $H$ .

When  $H$  is bipartite, the estimate above shows that  $\text{ex}(n, H) = o(n^2)$ . This bound is easily improved to show that for every bipartite graph  $H$  there is some positive  $\delta$  such that  $\text{ex}(n, H) = O(n^{2-\delta})$ . However, there are very few bipartite graphs for which we have matching upper and lower bounds.

The most closely studied case is when  $H = K_{s,t}$ , the complete bipartite graph with parts of order  $s$  and  $t$ . In this case, a famous result of Kővári, Sós and Turán [14] shows that  $\text{ex}(n, K_{s,t}) = O_{s,t}(n^{2-1/s})$  whenever  $s \leq t$ . This bound was shown to be tight for  $s = 2$  by Erdős, Rényi and Sós [7] and for  $s = 3$  by Brown [3]. For higher values of  $s$ , it is only known that the bound is tight when  $t$  is sufficiently large in terms of  $s$ . This was first shown by Kollár, Rónyai and Szabó [13], though their construction was improved slightly by Alon, Rónyai and Szabó [1], who showed that there are graphs with  $n$  vertices and  $\Omega_s(n^{2-1/s})$  edges containing no copy of  $K_{s,t}$  with  $t = (s - 1)! + 1$ .

Alternative proofs showing that  $\text{ex}(n, K_{s,t}) = \Omega_s(n^{2-1/s})$  when  $t$  is significantly larger than  $s$  were later found by Blagojević, Bukh and Karasev [2] and by Bukh [4]. In both cases, the basic idea behind the construction is to take a random polynomial  $f : \mathbb{F}_q^s \times \mathbb{F}_q^s \rightarrow \mathbb{F}_q$  and then to consider the graph  $G$  between two copies of  $\mathbb{F}_q^s$  whose edges are all those pairs  $(x, y)$  such that  $f(x, y) = 0$ . A further application of this random algebraic technique was recently given by Conlon [5], who showed that for every natural number  $k \geq 2$  there exists a natural number  $\ell$  such that, for every  $n$ , there is a graph

---

\*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: [bbukh@math.cmu.edu](mailto:bbukh@math.cmu.edu). Research supported in part by NSF grant DMS-1201380.

†Mathematical Institute, Oxford OX2 6GG, United Kingdom. E-mail: [david.conlon@maths.ox.ac.uk](mailto:david.conlon@maths.ox.ac.uk). Research supported by a Royal Society University Research Fellowship.

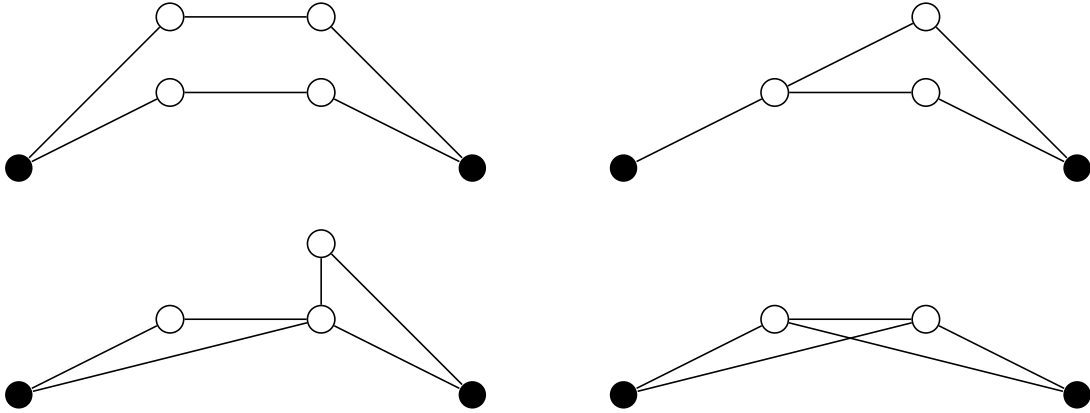


Figure 1: Graphs in  $\mathcal{T}^2$  when  $(T, R)$  is a path of length 3 with rooted endpoints.

on  $n$  vertices with  $\Omega_k(n^{1+1/k})$  vertices for which there are at most  $\ell$  paths of length  $k$  between any two vertices. By a result of Faudree and Simonovits [10], this is sharp up to the implied constant. We refer the interested reader to [5] for further background and details.

In this paper, we give yet another application of the random algebraic method, proving that for every rational number between 1 and 2, there is a family of graphs  $\mathcal{H}_r$  for which  $\text{ex}(n, \mathcal{H}_r) = \Theta(n^r)$ . This solves a longstanding open problem in extremal graph theory that has been reiterated by a number of authors, including Frankl [11] and Füredi and Simonovits [12].

**Theorem 1.1** *For every rational number  $r$  between 1 and 2, there exists a family of graphs  $\mathcal{H}_r$  such that  $\text{ex}(n, \mathcal{H}_r) = \Theta(n^r)$ .*

In order to define the relevant families  $\mathcal{H}_r$ , we need some preliminary definitions.

**Definition 1.1** *A rooted tree  $(T, R)$  consists of a tree  $T$  together with an independent set  $R \subset V(T)$ , which we refer to as the roots. When the set of roots is understood, we will simply write  $T$ .*

Each of our families  $\mathcal{H}_r$  will be of the following form.

**Definition 1.2** *Given a rooted tree  $(T, R)$ , we define the  $p$ th power  $\mathcal{T}_R^p$  of  $(T, R)$  to be the family of graphs consisting of all possible unions of  $p$  distinct labelled copies of  $T$ , each of which agree on the set of roots  $R$ . Again, we will usually omit  $R$ , denoting the family by  $\mathcal{T}^p$  and referring to it as the  $p$ th power of  $T$ .*

We note that  $\mathcal{T}^p$  is a family because we allow the unrooted vertices  $V(T) \setminus R$  to meet in every possible way. For example, if  $T$  is a path of length 3 whose endpoints are rooted, the family  $\mathcal{T}^2$  contains a cycle of length 6 and the various degenerate configurations shown in Figure 1.

The following parameter will be critical in studying the extremal number of the family  $\mathcal{T}^p$ .

**Definition 1.3** *Given a rooted tree  $(T, R)$ , we define the density  $\rho_T$  of  $(T, R)$  to be  $\frac{e(T)}{v(T)-|R|}$ .*

The upper bound in Theorem 1.1 will follow from an application of the next lemma.

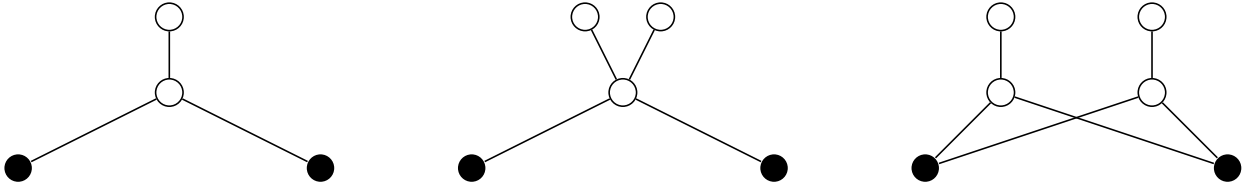


Figure 2: An unbalanced rooted tree  $T$  and two elements of  $\mathcal{T}^2$ .

**Lemma 1.1** *For any rooted tree  $(T, R)$  with at least one root, the family  $\mathcal{T}^p$  satisfies*

$$ex(n, \mathcal{T}^p) = O_p(n^{2-1/\rho_T}).$$

It would be wonderful if there were also a matching lower bound for  $ex(n, \mathcal{T}^p)$ . However, this is in general too much to expect. If, for example,  $(T, R)$  is the star  $K_{1,3}$  with two rooted leaves,  $\mathcal{T}^2$  will contain the graph shown in Figure 2 where the two central vertices agree. However, this graph is a tree, so it is easy to show that  $ex(n, \mathcal{T}^2) = O(n)$ , whereas, since  $\rho_T = 3/2$ , Lemma 1.1 only gives  $ex(n, \mathcal{T}^2) = O(n^{4/3})$ . Luckily, we may avoid these difficulties by restricting attention to so-called balanced trees.

**Definition 1.4** *Given a subset  $S$  of the unrooted vertices  $V(T) \setminus R$  in a rooted tree  $(T, R)$ , we define the density  $\rho_S$  of  $S$  to be  $e(S)/|S|$ , where  $e(S)$  is the number of edges in  $T$  with at least one endpoint in  $S$ . Note that when  $S = V(T) \setminus R$ , this agrees with the definition above. We say that the rooted tree  $(T, R)$  is balanced if, for every subset  $S$  of the unrooted vertices  $V(T) \setminus R$ , the density of  $S$  is at least the density of  $T$ , that is,  $\rho_S \geq \rho_T$ . In particular, if  $|R| \geq 2$ , then this condition guarantees that every leaf in the tree is a root.*

With the caveat that our rooted trees must be balanced, we may now prove a lower bound matching Lemma 1.1 by using the random algebraic method.

**Lemma 1.2** *For any balanced rooted tree  $(T, R)$ , there exists a positive integer  $p$  such that the family  $\mathcal{T}^p$  satisfies*

$$ex(n, \mathcal{T}^p) = \Omega(n^{2-1/\rho_T}).$$

Therefore, given a rational number  $r$  between 1 and 2, it only remains to identify a balanced rooted tree  $(T, R)$  for which  $2 - 1/\rho_T$  is equal to  $r$ .

**Definition 1.5** *Suppose that  $a$  and  $b$  are natural numbers satisfying  $a - 1 \leq b < 2a - 1$  and put  $i = b - a$ . We define a rooted tree  $T_{a,b}$  by taking a path with  $a$  vertices, which are labelled in order as  $1, 2, \dots, a$ , and then adding an additional rooted leaf to each of the  $i + 1$  vertices*

$$1, \left\lfloor 1 + \frac{a}{i} \right\rfloor, \left\lfloor 1 + 2 \cdot \frac{a}{i} \right\rfloor, \dots, \left\lfloor 1 + (i - 1) \cdot \frac{a}{i} \right\rfloor, a.$$

*For  $b \geq 2a - 1$ , we define  $T_{a,b}$  recursively to be the tree obtained by attaching a rooted leaf to each unrooted vertex of  $T_{a,b-a}$ .*



Figure 3: The rooted trees  $T_{4,9}$  and  $T_{4,10}$ .

Note that the tree  $T_{a,b}$  has  $a$  unrooted vertices and  $b$  edges, so that  $\rho_T = b/a$ . Now, given a rational number  $r$  with  $1 < r < 2$ , let  $a/b = 2 - r$  and let  $\mathcal{T}_{a,b}^p$  be the  $p$ th power of  $T_{a,b}$ . To prove Theorem 1.1, it will suffice to prove that  $T_{a,b}$  is balanced, since we may then apply Lemmas 1.1 and 1.2 to  $\mathcal{T}_{a,b}^p$ , for  $p$  sufficiently large, to conclude that

$$\text{ex}(n, \mathcal{T}_{a,b}^p) = \Theta(n^{2-a/b}) = \Theta(n^r).$$

Therefore, the following lemma completes the proof of Theorem 1.1.

**Lemma 1.3** *The tree  $T_{a,b}$  is balanced.*

All of the proofs will be given in the next section: we will prove the easy Lemma 1.1 in Section 2.1; Lemma 1.3 and another useful fact about balanced trees will be proved in Section 2.2; and Lemma 1.2 will be proved in Section 2.3. We conclude, in Section 3, with some brief remarks.

## 2 Proofs

### 2.1 The upper bound

We will use the following folklore lemma.

**Lemma 2.1** *A graph  $G$  with average degree  $d$  has a subgraph  $G'$  of minimum degree at least  $d/2$ .*

With this mild preliminary, we are ready to prove Lemma 1.1, that  $\text{ex}(n, \mathcal{T}^p) = O_p(n^{2-1/\rho_T})$  for any rooted tree  $(T, R)$ .

**Proof of Lemma 1.1:** Suppose that  $G$  is a graph on  $n$  vertices with  $cn^{2-\alpha}$  edges, where  $\alpha = 1/\rho_T$  and  $c \geq 2 \min(|T|, p)$ . We wish to show that  $G$  contains an element of  $\mathcal{T}^p$ . Since the average degree of  $G$  is  $2cn^{1-\alpha}$ , Lemma 2.1 implies that  $G$  has a subgraph  $G'$  with minimum degree at least  $cn^{1-\alpha}$ . Suppose that this subgraph has  $s \leq n$  vertices. By embedding greedily one vertex at a time, the minimum degree condition allows us to conclude that  $G'$  contains at least

$$s \cdot cn^{1-\alpha} \cdot (cn^{1-\alpha} - 1) \cdots (cn^{1-\alpha} - |T| + 2) \geq (c/2)^{|T|-1} sn^{(|T|-1)(1-\alpha)}$$

labelled copies of the (unrooted) tree  $T$ . Since there are at most  $\binom{s}{|R|} \leq s^{|R|}$  choices for the root vertices  $R$ , there must be some choice of root vertices  $R_0$  which corresponds to the root set in at least

$$\frac{(c/2)^{|T|-1} sn^{(|T|-1)(1-\alpha)}}{s^{|R|}} \geq \frac{(c/2)^{|T|-1} n^{(|T|-1)(1-\alpha)}}{n^{|R|-1}} = (c/2)^{|T|-1}$$

distinct labelled copies of  $T$ , where we used that  $s \leq n$  and  $\alpha = 1/\rho_T = (|T| - |R|)/(|T| - 1)$ . Since  $(c/2)^{|T|-1} \geq p$ , this gives the required element of  $\mathcal{T}^p$ .  $\square$

## 2.2 Balanced trees

We will begin by proving Lemma 1.3, that  $T_{a,b}$  is balanced.

**Proof of Lemma 1.3:** Suppose that  $S$  is a proper subset of the unrooted vertices of  $T_{a,b}$ . We wish to show that  $e(S)$ , the number of edges in  $T$  with at least one endpoint in  $S$ , is at least  $\rho_T|S|$ , where  $\rho_T = b/a$ . We may make two simplifying assumptions. First, we may assume that  $a - 1 \leq b < 2a - 1$ . Indeed, if  $b \geq 2a - 1$ , then the bound for  $T_{a,b}$  follows from the bound for  $T_{a,b-a}$ , which we may assume by induction. Second, we may assume that the vertices in  $S$  form a subpath of the base path of length  $a$ . Indeed, given the result in this case, we may write any  $S$  as the disjoint union of subpaths  $S_1, S_2, \dots, S_p$  with no edges between them, so that

$$e(S) = e(S_1 \cup S_2 \cup \dots \cup S_p) = e(S_1) + e(S_2) + \dots + e(S_p) \geq \rho_T(|S_1| + |S_2| + \dots + |S_p|) = \rho_T|S|.$$

Suppose, therefore, that  $S = \{l, l+1, \dots, r\}$  is a proper subpath of the base path  $\{1, 2, \dots, a\}$  and  $b - a = i$ .

As the desired claim is trivially true if  $i = -1$ , we will assume that  $i \geq 0$ . In particular, it follows from this assumption that vertex 1 of the base path is adjacent to a rooted vertex.

Let  $R$  be the number of rooted vertices adjacent to  $S$ . For  $0 \leq j \leq i - 1$ , the  $j$ th rooted vertex is adjacent to  $S$  precisely when  $l \leq 1 + j \left(\frac{a}{i}\right) < r + 1$ , which is equivalent to

$$(l-1)\frac{i}{a} \leq j < r\frac{i}{a}.$$

Therefore, if  $a$  is not contained in  $S$ , it follows that  $R \geq \lfloor |S|\frac{i}{a} \rfloor = \lfloor |S|\frac{b-a}{a} \rfloor$ . Furthermore, if  $l = 1$ , then  $R = \lceil |S|\frac{b-a}{a} \rceil$ . Finally, if  $r = a$  and  $i > 0$ , then, using

$$a - \left\lceil 1 + j \cdot \frac{a}{i} \right\rceil \leq (i-j)\frac{a}{i},$$

it follows that  $S$  is adjacent to the  $j$ th root whenever  $i|S|/a > i - j$ , and so  $R \geq \lceil |S|\frac{b-a}{a} \rceil$ .

*Case 1:*  $i = 0$ . Since  $S$  is a proper subpath, it is adjacent to at least  $|S| = (b/a)|S|$  edges.

*Case 2:*  $R \geq \lceil |S|\frac{b-a}{a} \rceil$ . Then the total number of edges adjacent to  $S$  is at least  $R + |S| \geq (b/a)|S|$ .

*Case 3:*  $i > 0$  and  $R < \lceil |S|\frac{b-a}{a} \rceil$ . Then  $S$  is adjacent to  $|S| + 1$  edges in the base path, for a total of  $\lceil |S|\frac{b-a}{a} \rceil + |S| + 1 \geq (b/a)|S|$  adjacent edges.  $\square$

Before moving on to the proof of Lemma 1.2, it will be useful to note that if  $T$  is balanced then every graph in  $\mathcal{T}^p$  is at least as dense as  $T$ .

**Lemma 2.2** *If  $(T, R)$  is a balanced rooted tree, then every graph  $H$  in  $\mathcal{T}^s$  satisfies*

$$e(H) \geq \rho_T(|H| - |R|).$$

**Proof:** We will prove the result by induction on  $s$ . It is clearly true when  $s = 1$ , so we will assume that it holds for any  $H \in \mathcal{T}^s$  and prove it when  $H \in \mathcal{T}^{s+1}$ .

Suppose, therefore, that  $H$  is the union of  $s + 1$  labelled copies of  $T$ , say  $T_1, T_2, \dots, T_{s+1}$ , each of which agree on the set of roots  $R$ . If we let  $H'$  be the union of the first  $s$  copies of  $T$ , the induction

hypothesis tells us that  $e(H') \geq \rho_T(|H'| - |R|)$ . Let  $S$  be the set of vertices in  $T_{s+1}$  which are not contained in  $H'$ . Then, since  $T$  is balanced, we know that  $e(S)$ , the number of edges in  $T_{s+1}$  (and, therefore, in  $H$ ) with at least one endpoint in  $S$ , is at least  $\rho_T|S|$ . It follows that

$$e(H) \geq e(H') + e(S) \geq \rho_T(|H'| - |R|) + \rho_T|S| = \rho_T(|H| - |R|),$$

as required.  $\square$

### 2.3 The lower bound

The proof of the lower bound will follow [4] and [5] quite closely. We begin by describing the basic setup and stating a number of lemmas which we will require in the proof. We will omit the proofs of these lemmas, referring the reader instead to [4] and [5].

Let  $q$  be a prime power and let  $\mathbb{F}_q$  be the finite field of order  $q$ . We will consider polynomials in  $t$  variables over  $\mathbb{F}_q$ , writing any such polynomial as  $f(X)$ , where  $X = (X_1, \dots, X_t)$ . We let  $\mathcal{P}_d$  be the set of polynomials in  $X$  of degree at most  $d$ , that is, the set of linear combinations over  $\mathbb{F}_q$  of monomials of the form  $X_1^{a_1} \cdots X_t^{a_t}$  with  $\sum_{i=1}^t a_i \leq d$ . By a random polynomial, we just mean a polynomial chosen uniformly from the set  $\mathcal{P}_d$ . One may produce such a random polynomial by choosing the coefficients of the monomials above to be random elements of  $\mathbb{F}_q$ .

The first result we will need says that once  $q$  and  $d$  are sufficiently large, the probability that a randomly chosen polynomial from  $\mathcal{P}_d$  contains each of  $m$  distinct points is exactly  $1/q^m$ .

**Lemma 2.3** *Suppose that  $q > \binom{m}{2}$  and  $d \geq m - 1$ . Then, if  $f$  is a random polynomial from  $\mathcal{P}_d$  and  $x_1, \dots, x_m$  are  $m$  distinct points in  $\mathbb{F}_q^t$ ,*

$$\mathbb{P}[f(x_i) = 0 \text{ for all } i = 1, \dots, m] = 1/q^m.$$

We also need to note some basic facts about affine varieties over finite fields. If we write  $\overline{\mathbb{F}}_q$  for the algebraic closure of  $\mathbb{F}_q$ , a variety over  $\overline{\mathbb{F}}_q$  is a set of the form

$$W = \{x \in \overline{\mathbb{F}}_q^t : f_1(x) = \cdots = f_s(x) = 0\}$$

for some collection of polynomials  $f_1, \dots, f_s: \overline{\mathbb{F}}_q^t \rightarrow \overline{\mathbb{F}}_q$ . We say that  $W$  is defined over  $\mathbb{F}_q$  if the coefficients of these polynomials are in  $\mathbb{F}_q$  and write  $W(\mathbb{F}_q) = W \cap \mathbb{F}_q^t$ . We say that  $W$  has complexity at most  $M$  if  $s, t$  and the degrees of the  $f_i$  are all bounded by  $M$ . Finally, we say that a variety is absolutely irreducible if it is irreducible over  $\overline{\mathbb{F}}_q$ , reserving the term irreducible for irreducibility over  $\mathbb{F}_q$  of varieties defined over  $\mathbb{F}_q$ .

The next result we will need is the Lang–Weil bound [15] relating the dimension of a variety  $W$  to the number of points in  $W(\mathbb{F}_q)$ . It will not be necessary to give a formal definition for the dimension of a variety, though some intuition may be gained by noting that if  $f_1, \dots, f_s: \overline{\mathbb{F}}_q^t \rightarrow \overline{\mathbb{F}}_q$  are generic polynomials then the dimension of the variety they define is  $t - s$ .

**Lemma 2.4** *Suppose that  $W$  is a variety over  $\overline{\mathbb{F}}_q$  of complexity at most  $M$ . Then*

$$|W(\mathbb{F}_q)| = O_M(q^{\dim W}).$$

*Moreover, if  $W$  is defined over  $\mathbb{F}_q$  and absolutely irreducible, then*

$$|W(\mathbb{F}_q)| = q^{\dim W} (1 + O_M(q^{-1/2})).$$

We will also need the following standard result from algebraic geometry, which says that if  $W$  is an absolutely irreducible variety and  $D$  is a variety intersecting  $W$ , then either  $W$  is contained in  $D$  or its intersection with  $D$  has smaller dimension.

**Lemma 2.5** *Suppose that  $W$  is an absolutely irreducible variety over  $\overline{\mathbb{F}}_q$  and  $\dim W \geq 1$ . Then, for any variety  $D$ , either  $W \subseteq D$  or  $W \cap D$  is a variety of dimension less than  $\dim W$ .*

The final ingredient we require says that if  $W$  is a variety which is defined over  $\mathbb{F}_q$ , then there is a bounded collection of absolutely irreducible varieties  $Y_1, \dots, Y_t$ , each of which is defined over  $\mathbb{F}_q$ , such that  $\cup_{i=1}^t Y_i(\mathbb{F}_q) = W(\mathbb{F}_q)$ .

**Lemma 2.6** *Suppose that  $W$  is a variety over  $\overline{\mathbb{F}}_q$  of complexity at most  $M$  which is defined over  $\mathbb{F}_q$ . Then there are  $O_M(1)$  absolutely irreducible varieties  $Y_1, \dots, Y_t$ , each of which is defined over  $\mathbb{F}_q$  and has complexity  $O_M(1)$ , such that  $\cup_{i=1}^t Y_i(\mathbb{F}_q) = W(\mathbb{F}_q)$ .*

We can combine the preceding three lemmas into a single result as follows:

**Lemma 2.7** *Suppose  $W$  and  $D$  are varieties over  $\overline{\mathbb{F}}_q$  of complexity at most  $M$  which are defined over  $\mathbb{F}_q$ . Then one of the following holds, for all  $q$  that are sufficiently large in terms of  $M$ :*

- $|W(\mathbb{F}_q) \setminus D(\mathbb{F}_q)| \geq q/2$ , or
- $|W(\mathbb{F}_q) \setminus D(\mathbb{F}_q)| \leq c$ , where  $c = c_M$  depends only on  $M$ .

**Proof:** By Lemma 2.6, there is a decomposition  $W(\mathbb{F}_q) = \cup_{i=1}^t Y_i(\mathbb{F}_q)$  for some bounded-complexity absolutely irreducible varieties  $Y_i$  defined over  $\mathbb{F}_q$ . If  $\dim Y_i \geq 1$ , Lemma 2.5 tells us that either  $Y_i \subseteq D$  or the dimension of  $Y_i \cap D$  is smaller than the dimension of  $Y_i$ . If  $Y_i \subseteq D$ , then the component does not contribute any point to  $W(\mathbb{F}_q) \setminus D(\mathbb{F}_q)$  and may be discarded. If instead the dimension of  $Y_i$  is smaller than the dimension of  $Y_i$ , the Lang–Weil bound, Lemma 2.4, tells us that for  $q$  sufficiently large

$$|W(\mathbb{F}_q) \setminus D(\mathbb{F}_q)| \geq |Y_i(\mathbb{F}_q)| - |Y_i(\mathbb{F}_q) \cap D| \geq q^{\dim Y_i} - O(q^{\dim Y_i - \frac{1}{2}}) - O(q^{\dim Y_i - 1}) \geq q/2.$$

On the other hand, if  $\dim Y_i = 0$  for every  $Y_i$  which is not contained in  $D$ , Lemma 2.4 tells us that  $|W(\mathbb{F}_q)| \leq \sum |Y_i(\mathbb{F}_q)| = O(1)$ , where the sum is taken over all  $i$  for which  $\dim Y_i = 0$ .  $\square$

We are now ready to prove Lemma 1.2, that for any balanced rooted tree  $(T, R)$  there exists a positive integer  $p$  such that  $\text{ex}(n, \mathcal{T}^p) = \Omega(n^{2-1/\rho_T})$ .

**Proof of Lemma 1.2:** Let  $(T, R)$  be a balanced rooted tree with  $a$  unrooted vertices and  $b$  edges, where  $R = \{u_1, \dots, u_r\}$  and  $V(T) \setminus R = \{v_1, \dots, v_a\}$ . Let  $s = 2br$ ,  $d = sb$ ,  $N = q^b$  and suppose that  $q$  is sufficiently large. Let  $f_1, \dots, f_a: \mathbb{F}_q^b \times \mathbb{F}_q^b \rightarrow \mathbb{F}_q$  be independent random polynomials in  $\mathcal{P}_d$ . We will consider the bipartite graph  $G$  between two copies  $U$  and  $V$  of  $\mathbb{F}_q^b$ , each of order  $N = q^b$ , where  $(u, v)$  is an edge of  $G$  if and only if

$$f_1(u, v) = \dots = f_a(u, v) = 0.$$

Since  $f_1, \dots, f_a$  were chosen independently, Lemma 2.3 with  $m = 1$  tells us that the probability a given edge  $(u, v)$  is in  $G$  is  $q^{-a}$ . Therefore, the expected number of edges in  $G$  is  $q^{-a}N^2 = N^{2-a/b}$ .

Suppose now that  $w_1, w_2, \dots, w_r$  are fixed vertices in  $G$  and let  $C$  be the collection of copies of  $T$  in  $G$  such that  $w_i$  corresponds to  $u_i$  for all  $1 \leq i \leq r$ . We will be interested in estimating the  $s$ -th moment of  $|C|$ . To begin, we note that  $|C|^s$  counts the number of ordered collections of  $s$  (possibly overlapping or identical) copies of  $T$  in  $G$  such that  $w_i$  corresponds to  $u_i$  for all  $1 \leq i \leq r$ . Since the total number of edges  $m$  in a given collection of  $s$  rooted copies of  $T$  is at most  $sb$  and  $q$  is sufficiently large, Lemma 2.3 tells us that the probability this particular collection of copies of  $T$  is in  $G$  is  $q^{-am}$ , where we again use the fact that  $f_1, \dots, f_a$  are chosen independently.

Suppose that  $H$  is an element of  $\mathcal{T}_{\leq}^s \stackrel{\text{def}}{=} \mathcal{T}^1 \cup \mathcal{T}^2 \cup \dots \cup \mathcal{T}^s$ . Within the complete bipartite graph from  $U$  to  $V$ , let  $N_s(H)$  be the number of ordered collections of  $s$  copies of  $T$ , each rooted at  $w_1, \dots, w_r$  in the same way, whose union is a copy of  $H$ . Then

$$\mathbb{E}[|C|^s] = \sum_{H \in \mathcal{T}_{\leq}^s} N_s(H) q^{-ae(H)},$$

while  $N_s(H) = O_s(N^{|H|-|R|})$ . Since  $(T, R)$  is balanced, Lemma 2.2 shows that  $\frac{e(H)}{|H|-|R|} \geq \rho_T = \frac{b}{a}$  for every  $H \in \mathcal{T}_{\leq}^s$ . It follows that

$$\begin{aligned} \mathbb{E}[|C|^s] &= \sum_{H \in \mathcal{T}_{\leq}^s} N_s(H) q^{-ae(H)} = \sum_{H \in \mathcal{T}_{\leq}^s} O_s(N^{|H|-|R|}) q^{-ae(H)} \\ &= O_s \left( \sum_{H \in \mathcal{T}_{\leq}^s} q^{b(|H|-|R|)} q^{-ae(H)} \right) = O_s(1). \end{aligned}$$

By Markov's inequality, we may conclude that

$$\mathbb{P}[|C| \geq c] = \mathbb{P}[|C|^s \geq c^s] \leq \frac{\mathbb{E}[|C|^s]}{c^s} = \frac{O_s(1)}{c^s}.$$

Our aim now is to show that  $|C|$  is either quite small or very large. To begin, note that the set  $C$  is a subset of  $X(\mathbb{F}_q)$ , where  $X$  is the algebraic variety defined as the set of  $(x_1, \dots, x_a) \in \overline{\mathbb{F}}_q^{ba}$  satisfying the equations

- $f_i(w_k, x_\ell) = 0$  for all  $k$  and  $\ell$  such that  $(u_k, v_\ell) \in T$  and
- $f_i(x_k, x_\ell) = 0$  for all  $k$  and  $\ell$  such that  $(v_k, v_\ell) \in T$

for all  $i = 1, 2, \dots, a$ . For each  $k \neq \ell$  such that  $v_k$  and  $v_\ell$  are on the same side of the natural bipartition of  $T$ , we let

$$D_{k\ell} = X \cap \{(x_1, \dots, x_a) : x_k = x_\ell\}.$$

We put

$$D \stackrel{\text{def}}{=} \bigcup_{k, \ell} D_{k\ell}.$$

That is, the  $D_{k\ell}$  capture those elements of  $X$  which are degenerate and so not elements of  $C$ . As a union of varieties is a variety, the set  $D$  is a variety that captures the degenerate elements of  $X$ . Furthermore, the complexity of  $D$  is bounded since the number and complexity of  $D_{k\ell}$  is bounded.



By Lemma 2.7, we see that there exists a constant  $c_T$ , depending only on  $T$ , such that either  $|C| \leq c_T$  or  $|C| \geq q/2$ . Therefore, by the consequence of Markov's inequality noted earlier,

$$\mathbb{P}[|C| > c_T] = \mathbb{P}[|C| \geq q/2] = \frac{O_s(1)}{(q/2)^s}.$$

We call a sequence of vertices  $(w_1, w_2, \dots, w_r)$  bad if there are more than  $c_T$  copies of  $T$  in  $G$  such that  $w_i$  corresponds to  $u_i$  for all  $1 \leq i \leq r$ . If we let  $B$  be the random variable counting the number of bad sequences, we have, since  $s = 2br$  and  $q$  is sufficiently large,

$$\mathbb{E}[B] \leq 2N^r \cdot \frac{O_s(1)}{(q/2)^s} = O_s(q^{br-s}) = o(1).$$

We now remove a vertex from each bad sequence to form a new graph  $G'$ . Since each vertex has degree at most  $N$ , the total number of edges removed is at most  $BN$ . Hence, the expected number of edges in  $G'$  is

$$N^{2-a/b} - \mathbb{E}[B]N = \Omega(N^{2-a/b}).$$

Therefore, there is a graph with at most  $2N$  vertices and  $\Omega(N^{2-a/b})$  edges such that no sequence of  $r$  vertices has more than  $c_T$  labelled copies of  $T$  rooted on these vertices. Finally, we note that this result was only shown to hold when  $q$  is a prime power and  $N = q^b$ . However, an application of Bertrand's postulate shows that the same conclusion holds for all  $N$ .  $\square$

### 3 Concluding remarks

We have shown that for any rational number  $r$  between 1 and 2, there exists a family of graphs  $\mathcal{H}_r$  such that  $\text{ex}(n, \mathcal{H}_r) = \Theta(n^r)$ . However, Erdős and Simonovits (see, for example, [6]) asked whether there exists a single graph  $H_r$  such that  $\text{ex}(n, H_r) = \Theta(n^r)$ . Our methods give some hope of a positive solution to this question, but the difficulties now lie with determining accurate upper bounds for the extremal number of certain graphs.

To be more precise, given a rooted tree  $(T, R)$ , we define  $T^p$  to be the graph consisting of the union of  $p$  distinct labelled copies of  $T$ , each of which agree on the set of roots  $R$  but are otherwise disjoint. Lemma 1.2 clearly shows that  $\text{ex}(n, T^p) = \Omega(n^{2-1/\rho_T})$  when  $T$  is a balanced rooted tree. We believe that a corresponding upper bound should also hold.

**Conjecture 3.1** *For any balanced rooted tree  $(T, R)$ , the graph  $T^p$  satisfies*

$$\text{ex}(n, T^p) = O_p(n^{2-1/\rho_T}).$$

The condition that  $(T, R)$  be balanced is necessary here, as may be seen by considering the graph in Figure 2, namely, a star  $K_{1,3}$  with two rooted leaves. Then  $T^2$  contains a cycle of length 4, so the extremal number is  $\Omega(n^{3/2})$ , whereas the conjecture would suggest that it is  $O(n^{4/3})$ .

In order to solve the Erdős–Simonovits conjecture, it would be sufficient to solve the conjecture for the collection of rooted trees  $T_{a,b}$  with  $a < b$  and  $(a, b) = 1$ . However, even this seems surprisingly difficult and the only known cases are when  $a = 1$ , in which case  $T$  is a star with rooted leaves and  $T^p$  is a complete bipartite graph, or  $b - a = 1$ , when  $T$  is a path with rooted endpoints and  $T^p$  is a theta graph.

**Acknowledgements.** We would like to thank Jacques Verstraete for interesting discussions relating to the topic of this paper.

## References

- [1] N. Alon, L. Rónyai and T. Szabó, Norm-graphs: variations and applications, *J. Combin. Theory Ser. B* **76** (1999), 280–290.
- [2] P. V. M. Blagojević, B. Bukh and R. Karasev, Turán numbers for  $K_{s,t}$ -free graphs: topological obstructions and algebraic constructions, *Israel J. Math.* **197** (2013), 199–214.
- [3] W. G. Brown, On graphs that do not contain a Thomsen graph, *Canad. Math. Bull.* **9** (1966), 281–285.
- [4] B. Bukh, Random algebraic construction of extremal graphs, arXiv:1409.3856 [math.CO].
- [5] D. Conlon, Graphs with few paths of prescribed length between any two vertices, arXiv:1411.0856 [math.CO].
- [6] P. Erdős, On the combinatorial problems which I would most like to see solved, *Combinatorica* **1** (1981), 25–42.
- [7] P. Erdős, A. Rényi and V. T. Sós, On a problem of graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 215–235.
- [8] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.
- [9] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.
- [10] R. J. Faudree and M. Simonovits, On a class of degenerate extremal graph problems, *Combinatorica* **3** (1983), 83–93.
- [11] P. Frankl, All rationals occur as exponents, *J. Combin. Theory Ser. A* **42** (1986), 200–206.
- [12] Z. Füredi and M. Simonovits, The history of degenerate (bipartite) extremal graph problems, in Erdős Centennial, 169–264, Bolyai Soc. Math. Stud., 25, Springer, Berlin, 2013.
- [13] J. Kollár, L. Rónyai and T. Szabó, Norm-graphs and bipartite Turán numbers, *Combinatorica* **16** (1996), 399–406.
- [14] T. Kővári, V. T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.
- [15] S. Lang and A. Weil, Number of points of varieties in finite fields, *Amer. J. Math.* **76** (1954), 819–827.