

# Perfect and nearly perfect separation dimension of complete and random graphs \*

Raphael Yuster †

## Abstract

The separation dimension of a graph  $G$  is the smallest natural number  $d$  for which there is an embedding of  $G$  into  $\mathbb{R}^d$ , such that any pair of disjoint edges is separated by some hyperplane normal to one of the axes. The perfect separation dimension further requires that any pair of disjoint edges is separated by the same amount of such (pairwise nonparallel) hyperplanes. While it is known that the separation dimension of  $K_n$  is  $\Theta(\log n)$ , the perfect separation dimension is much larger.

In our first result, we prove that the perfect separation dimension of  $K_n$  is linear in  $n$ , up to a small polylogarithmic factor. In fact, we prove it is at least  $n/2 - 1$  and at most  $n(\log n)^{1+o(1)}$ .

Our second result shows that, in fact, for almost all graphs, the separation dimension is also linear in  $n$ , up to a logarithmic factor. This follows as a special case of a more general result showing that the perfect separation dimension of the random graph  $G(n, p)$  is w.h.p.  $\Omega(np/\log n)$  for a wide range of values of  $p$ , including all constant  $p$ .

Finally, in our third result, we show that significantly relaxing perfection to just requiring that any pair of disjoint edges is separated the same number of times up to a difference of  $c \log n$  for some absolute constant  $c$ , still requires the dimension to be  $\Omega(n)$ . This is perhaps surprising as it is known that if we allow a difference of  $7 \log n$ , then the dimension reduces to only  $\Theta(\log n)$ .

**AMS subject classifications:** 05B30, 05A05, 05C50, 05C62, 05C80

**Keywords:** separation dimension; complete graph; random graph; perfect separation

## 1 Introduction

The separation dimension of a graph  $G$  is the smallest natural number  $d$  for which there is an embedding of  $G$  into  $\mathbb{R}^d$ , such that any pair of disjoint edges is separated by some hyperplane normal to one of the axes. Equivalently, it is the smallest possible size of a *separating family* (also called *pairwise-suitable family* in [6]) of total orders of  $V(G)$ , where a family  $\mathcal{F}$  of total orders of

---

\*This research was supported by the Israel Science Foundation (grant No. 1082/16).

†Department of Mathematics, University of Haifa, Haifa 3498838, Israel. Email: raphael.yuster@gmail.com

$V(G)$  is separating if for any two disjoint edges of  $G$ , there exists at least one element of  $\mathcal{F}$  in which all the vertices of one edge precede those of the other. Let  $\pi(G)$  denote the separation dimension of  $G$ .

Separation dimension has been studied by several researchers. Basavaraju et al. [6] were the first to define and study it, motivated by its interesting connection with *boxicity*, a well-studied geometric representation of graphs. In particular, they have shown that  $\pi(K_n) = \Theta(\log n)$ . Alon et al. [3] proved that if  $G$  has maximum degree at most  $d$ , then  $\pi(G)$  is upper bounded by an almost linear function of  $d$  and provided constructions where  $\pi(G) \geq d/2$ . Scott and Wood [15] improved the aforementioned result on bounded degree graphs showing that  $\pi(G) \leq 20d$ . Recently, Alon et al. [4] proved that the separation dimension of a  $d$ -degenerate graph on  $n$  vertices is  $O(d \log \log n)$  and that the  $\log \log n$  term is required already for 2-degenerate graphs. They have also shown that graphs with bounded separation dimension cannot be very dense. Additional papers on separation dimension and some of its variants can be found in [7, 8, 12, 19].

In this paper we consider separating families that are also required to be *balanced*. For a separating family  $\mathcal{F}$  (hereafter we allow  $\mathcal{F}$  to be a multiset) and for a pair  $\{e, f\}$  of disjoint edges of  $G$ , let  $c_{\mathcal{F}}(\{e, f\})$  be the number of total orders in  $\mathcal{F}$  that separate  $e$  and  $f$ .

**Definition 1.1** ( $\Delta$ -balanced separation dimension). *For a non-negative integer  $\Delta$ , a separating family  $\mathcal{F}$  is  $\Delta$ -balanced if for any two pairs of edges  $\{e_1, f_1\}, \{e_2, f_2\}$  where  $e_1 \cap f_1 = \emptyset$  and  $e_2 \cap f_2 = \emptyset$  it holds that  $|c_{\mathcal{F}}(\{e_1, f_1\}) - c_{\mathcal{F}}(\{e_2, f_2\})| \leq \Delta$ . Let  $\pi_{\Delta}(G)$  be the  $\Delta$ -balanced separation dimension, namely the smallest possible size of a  $\Delta$ -balanced separating family.*

Observe that, trivially,  $\pi(G) = \pi_{\pi(G)}(G)$ . On the other extreme, perhaps most intriguing is the case  $\Delta = 0$ , which merits its separate definition.

**Definition 1.2** (perfect separation dimension). *A separating family  $\mathcal{F}$  is perfect if there exists a positive integer  $\lambda$  such that for any pair  $\{e, f\}$  of disjoint edges of  $G$  it holds that  $c_{\mathcal{F}}(\{e, f\}) = \lambda$ . We call  $\lambda$  the multiplicity of  $\mathcal{F}$ . Let  $\text{PSD}(G)$  be the perfect separation dimension of  $G$ , namely the smallest possible size of a perfect separating family.*

Notice that  $\text{PSD}(G)$  (and therefore  $\pi_{\Delta}(G)$  for any  $\Delta \geq 0$ ) is well-defined, as the family of all total orders of  $V(G)$  is a perfect separating family. Hence,  $\text{PSD}(G) \leq n!$  is a trivial upper bound for any graph with  $n$  vertices. Another easily observed fact is that if  $G'$  is a subgraph of  $G$ , then  $\text{PSD}(G') \leq \text{PSD}(G)$  and  $\pi_{\Delta}(G') \leq \pi_{\Delta}(G)$ . In particular, if  $G$  is an  $n$ -vertex graph, then  $\text{PSD}(G) \leq \text{PSD}(K_n)$ . As mentioned earlier,  $\pi(K_n) = \Theta(\log n)$  [6]. On the other hand, determining the order of magnitude of  $\text{PSD}(K_n)$  seems more involved. Our first main result determines  $\text{PSD}(K_n)$  up to a small polylogarithmic factor.

**Theorem 1.3.** *For all  $n \geq 4$  it holds that*

$$n/2 - 1 \leq \text{PSD}(K_n) \leq n(\log n)^{1+o(1)}.$$

Hence, the following open problem is naturally raised.

**Problem 1.4.** *Is  $\text{PSD}(K_n)$  linear in  $n$ ?*

One may wonder whether the “completeness” of  $K_n$  is the reason that its perfect separation dimension is essentially linear. Our next main result shows that this is not so. In fact, it will follow as a special case from the next theorem that almost all  $n$ -vertex graphs have perfect separation dimension almost linear in  $n$ . Recall the Erdős-Rényi random graph  $G(n, p)$ , the probability space of all labeled  $n$ -vertex graphs with edge probability  $p = p(n)$ . In what follows “with high probability” (w.h.p.) means with probability approaching 1 as  $n$  goes to infinity. The next theorem shows that for a wide range of  $p = p(n)$ , a sampled element of  $G(n, p)$  has perfect separation dimension close to its average degree up to a logarithmic factor. In particular, the statement above about almost all  $n$ -vertex graphs follows from the case  $p = \frac{1}{2}$ .

**Theorem 1.5.** *Let  $p = p(n)$  satisfy  $n^{-0.4} \leq p < 1$ . Then w.h.p.*

$$\text{PSD}[G(n, p)] \geq \frac{np}{400 \ln n} .$$

Another intriguing question is whether the perfection requirement is the only cause for  $\text{PSD}(K_n)$  to be  $\Omega(n)$ . Perhaps if we relax the condition and just ask for a  $\Delta$ -balanced separating family where  $\Delta$  is relatively small, we can reduce the dimension considerably? Recall again that  $\pi(K_n) = \Theta(\log n)$  and in fact it is proved in [6] that  $\pi(K_n) \leq 7 \log n$ <sup>1</sup>. In particular, one trivially get that  $\pi_\Delta(K_n) \leq 7 \log n$  for  $\Delta = 7 \log n$ . Now, suppose we decrease  $\Delta$  by just a constant factor to  $c \log n$  for some small absolute constant  $c$ . It is reasonable to suspect that  $\pi_\Delta(K_n)$  should not increase by much. Perhaps surprisingly, however, it does. Moreover, it increases to  $\Omega(n)$ , just as for  $\Delta = 0$  (namely,  $\text{PSD}(K_n)$ ).

**Theorem 1.6.** *There exist absolute positive constants  $c$  and  $K$  such that for all  $0 \leq \Delta \leq c \ln n$*

$$\pi_\Delta(K_n) \geq \frac{n}{K} .$$

The rest of this paper proceeds as follows. In Section 2 we construct the  $n(\log n)^{1+o(1)}$  upper bound for  $\text{PSD}(K_n)$ . In Section 3 we prove the  $n/2 - 1$  lower bound for  $\text{PSD}(K_n)$ , thus completing the proof of Theorem 1.3. The perfect separation dimension of random graphs is considered in Section 4 where we prove Theorem 1.5. Finally, in Section 5 we consider  $\Delta$ -balanced separation dimension and prove Theorem 1.6.

---

<sup>1</sup>Hereafter, when necessary, assume that  $\log n = \log_2 n$ .

## 2 Constructing a perfect separating family

In this section we prove the upper bound part of Theorem 1.3. We first set some terminology that we use throughout the rest of this section (in fact, throughout the rest of the paper). Let  $S_n$  denote the symmetric group of order  $n$ . Hence, a separating family  $\mathcal{F}$  of a graph with vertex set  $[n] = \{1, \dots, n\}$  is a (multi)subset of  $S_n$ . When explicitly specifying an element  $\sigma \in S_n$  we use the notation  $\sigma = (a_1, \dots, a_n)$  to mean that  $\sigma(i) = a_i$  and  $\sigma^{-1}(a_i) = i$  (this notation should not be confused with the cycle decomposition notation of a permutation). We use  $\sigma \circ \pi$  to denote the usual group element product of two elements  $\sigma, \pi \in S_n$ . We also view  $\sigma = (a_1, \dots, a_n)$  as a sequence whose  $i$ 'th element is  $a_i$ , hence the notion of a subsequence of  $\sigma$  is well-defined, as well as the notion of *reverse* where  $rev(\sigma) = (a_n, \dots, a_1)$ .

Our construction of a perfect separating family of  $K_n$  requires several ingredients which we next specify. Our first ingredient is the case  $k = 3$  of the following definition.

**Definition 2.1.** *[perfect  $k$ -sequence covering array] A multiset  $\mathcal{T} \subseteq S_n$  is a perfect  $k$ -sequence covering array if there exists a positive integer  $t$  such that every sequence of  $k$  distinct elements of  $[n]$  is a subsequence of precisely  $t$  elements of  $\mathcal{T}$ . We call  $t$  the multiplicity of  $\mathcal{T}$ .*

Notice that if  $\mathcal{T}$  is a perfect  $k$ -sequence covering array of multiplicity  $t$ , then  $|\mathcal{T}| = k!t$ . It is perhaps worth mentioning that a (not necessarily perfect)  *$k$ -sequence covering array* just asks for a subset of  $S_n$  that has the property that each  $k$ -sequence of distinct elements of  $[n]$  is a subsequence of *some* element of that subset. Fairly close lower and upper bounds are known for the minimum size of a  $k$ -sequence covering array of  $[n]$ , the state of the art for the case  $k = 3$  given in [9] (lower bound), [17] (upper bound) and for general  $k$  given in [14] (lower bound), [16] (upper bound). Fairly close upper and lower bounds are known for the minimum size of a perfect 3-sequence covering array of  $[n]$  [18]. Trivially, every perfect 4-sequence covering array is a perfect separating family. Unfortunately, however, a lower bound of  $\Omega(n^2)$  is known for the minimum size of a perfect 4-sequence covering array [18]. Since Theorem 1.3 asserts a quasi-linear upper bound for a perfect separating family, we cannot use perfect 4-sequence covering arrays for our construction. Nevertheless, we will prove that a modification of the construction of perfect 3-sequence covering arrays given in [18] (requiring a considerable modification of the proof), yields a construction of a perfect separating family. Notice that a priori, the requirement of being a perfect separating family seems more demanding than being a perfect 3-sequence covering array, as the former requires satisfying  $\Theta(n^4)$  demands while the latter requires satisfying  $\Theta(n^3)$  demands.

We will need the following simple lemma that merely states that a perfect separating family as well as a perfect  $k$ -sequence covering array is closed under relabeling.

**Lemma 2.2.** *Let  $\pi$  be a total order of an  $n$ -element set  $V$ . Let  $\mathcal{F}$  be a perfect separating family of  $K_n$ , and let  $\mathcal{T}$  be a perfect  $k$ -sequence covering array. Then  $\pi\mathcal{F} = \{\pi \circ \sigma \mid \sigma \in \mathcal{F}\}$  is a perfect*

separating family of the complete graph on vertex set  $V$ . Similarly,  $\pi\mathcal{T} = \{\pi \circ \sigma \mid \sigma \in \mathcal{T}\}$  is a perfect  $k$ -sequence covering array of the set  $V$ <sup>2</sup>.

*Proof.* We prove that  $\pi\mathcal{F}$  is a perfect separating family of  $K_n$  on vertex set  $V$ . The second part of the lemma is proved analogously. Let  $\lambda$  be the multiplicity of  $\mathcal{F}$ . Observe that  $\pi \circ \sigma$  separates  $(u_1, v_1)$  from  $(u_2, v_2)$  if and only if  $\sigma$  separates  $(\pi^{-1}(u_1), \pi^{-1}(v_1))$  from  $(\pi^{-1}(u_2), \pi^{-1}(v_2))$ . As  $\pi$  and each  $\sigma \in \mathcal{F}$  are bijective, this implies that  $\pi\mathcal{F}$  separates  $(u_1, v_1)$  from  $(u_2, v_2)$  precisely  $\lambda$  times.  $\square$

A *finite affine plane of order  $n$*  is a set  $\{P_1, \dots, P_{n+1}\}$  of partitions of  $[n^2]$  with the following properties: (i) Each  $P_i$  contains  $n$  parts of order  $n$  each, denoted by  $P_{i,j}$  for  $j = 1, \dots, n$ . (ii) For any pair of distinct elements of  $[n^2]$ , there is exactly one partition  $P_i$  that contains both of them in the same part of  $P_i$ . It is well-known that if  $n$  is a prime power, then there is a finite affine plane of order  $n$  [11]. In geometric terminology, the  $P_{i,j}$  are called *lines* and the  $P_i$  are called *parallel classes*. In our proof we will view each  $P_{i,j}$  not just as a set, but as an ordered set of  $n$  elements of  $[n^2]$ . In Figure 1 we list an affine plane of order 4 viewed in this way. In fact, this will be the plane we use in the first step of our construction and we refer to it in various examples.

	$P_1$	$P_2$	$P_3$	$P_4$	$P_5$
1	(1, 2, 3, 4)	(1, 5, 9, 14)	(4, 5, 10, 13)	(2, 6, 10, 14)	(4, 8, 12, 14)
2	(5, 6, 7, 8)	(2, 7, 12, 13)	(1, 6, 12, 16)	(3, 5, 12, 15)	(1, 7, 10, 15)
3	(9, 10, 11, 12)	(3, 8, 10, 16)	(2, 8, 9, 15)	(4, 7, 9, 16)	(3, 6, 9, 13)
4	(13, 14, 15, 16)	(4, 6, 11, 15)	(3, 7, 11, 14)	(1, 8, 11, 13)	(2, 5, 11, 16)

Figure 1: A finite affine plane of order 4. Row  $j$  column  $i$  is the line  $P_{i,j}$  listed as a sequence of distinct elements.

Our construction is established in the following lemma. However, before stating the lemma we require the definition of some integer function.

$$g(n) = \begin{cases} 6 & \text{if } n = 3, 4 \\ 2 \cdot (1+r)g(r) & \text{where } r \text{ is the smallest prime power satisfying } r^2 \geq n. \end{cases} \quad (1)$$

We note that  $g(n)$  is recursively well-defined and monotone non-decreasing for every  $n \geq 3$ . Indeed, it is defined for  $n = 3, 4$  and for larger  $n$ , we have that  $r^2$  is less than  $4n$  as some integer between  $n$  and  $4n - 1$  is an even power of 2. Now,  $r < 2\sqrt{n} \leq n$ . For example, we have  $g(5) = g(9) = 2(1+3)g(3) = 48$ ,  $g(10) = g(16) = 2(1+4)g(4) = 60$  and  $g(17) = g(25) = 2(1+5)g(5) = 576$ .

---

<sup>2</sup>Although in Definition 2.1 the ground set is  $[n]$ , we can equivalently consider any  $n$ -set  $V$  as a ground set and ask for a multiset of total orders of  $V$ .

**Lemma 2.3.** *For all  $n \geq 3$ ,  $\text{PSD}(K_n) \leq g(n)$ . Furthermore, there is a perfect separating family of  $K_n$  of size  $g(n)$  that is also a perfect 3-sequence covering array of  $[n]$ .*

*Proof.* We prove the lemma by induction on  $n$ . For the base cases  $n = 3, 4$ , we have  $g(3) = g(4) = 6$ . Trivially,  $S_3$  is a perfect 3-sequence covering array of  $\{1, 2, 3\}$  of size 6. It is also vacuously a perfect separating family of  $K_3$  as there are no two disjoint edges in  $K_3$ . For  $n = 4$ , consider the subset of permutations of  $S_4$  given by  $\{1234, 1432, 4231, 2413, 3412, 3214\}$ . It consists of 6 elements and it is easy to verify that it is a perfect 3-sequence covering array of  $\{1, 2, 3, 4\}$  (with multiplicity 1) as well as a perfect separating family of  $K_4$  of multiplicity 2. Notice that this set is not the smallest possible perfect separating family of  $K_4$ . Indeed, its subset  $\{1234, 1432, 4213\}$  is a perfect separating family of  $K_4$  of multiplicity 1.

For the general case, let us first observe that it suffices to prove the lemma for values of  $n$  such that  $n = r^2$  and  $r$  is a prime power. This immediately follows from the definition  $g(r)$ , as if  $n$  is not such, and  $r$  is the smallest prime power with  $r^2 > n$ , then  $g(n) = g(r^2)$  and from the fact mentioned in the introduction that if  $m \geq n$ , then  $\text{PSD}(K_m) \geq \text{PSD}(K_n)$  and similarly a perfect 3-sequence covering array of  $[m]$  induces a perfect 3-sequence covering array of  $[n]$  by keeping only the integers in  $[n]$  in each permutation. So, we assume until the end of the lemma that  $n = r^2 \geq 9$  and  $r$  is a prime power.

By induction we may assume the existence of a multiset  $\mathcal{F}$  consisting of elements of  $S_r$ , of size  $|\mathcal{F}| = g(r)$  that is both a perfect separating family of  $K_r$  and a perfect 3-sequence covering array of  $[r]$ . We must show that there exists a multiset  $\mathcal{F}^*$  of elements of  $S_{r^2}$  such that  $|\mathcal{F}^*| = g(r^2) = 2 \cdot (1+r)g(r)$  and such that  $\mathcal{F}^*$  is both a perfect separating family of  $K_{r^2}$  and a perfect 3-sequence covering array of  $[r^2]$ .

Fix some affine plane  $\{P_1, \dots, P_{r+1}\}$  of order  $r$  (which exists since  $r$  is a prime power) where each parallel class  $P_i$  is a partition of  $[r^2]$  into parts (lines)  $P_{i,1}, \dots, P_{i,r}$  and recall that each  $P_{i,j}$  is viewed as an ordered set. The choice of ordering is arbitrary; any total order of the elements of  $P_{i,j}$  is fine.

We construct  $\mathcal{F}^*$  as follows. Each  $P_i$  gives rise to a multiset  $W_i$  of elements of  $S_{r^2}$ . Furthermore,  $W_i$  is the union of two multisets  $Y_i$  and  $Z_i$ , each of size  $g(r)$ , hence  $|W_i| = 2g(r)$ . We then define  $\mathcal{F}^*$  to be the multiset obtained by taking all the  $W_i$ 's. Hence,  $\mathcal{F}^*$  indeed contains  $g(r^2) = 2(r+1)g(r)$  elements.

We next show how to construct  $Y_i$ . For  $\sigma \in S_r$ , define  $q(P_{i,j}, \sigma) = P_{i,j} \circ \sigma$  to be the permutation of  $P_{i,j}$  corresponding to  $\sigma$  (recall, for  $1 \leq k \leq r$ ,  $P_{i,j}(k)$  is the element at location  $k$  of the ordered set  $P_{i,j}$ , so  $P_{i,j} \circ \sigma$  is well-defined). For  $\sigma \in S_r$ , let  $q(P_i, \sigma)$  be the concatenation of  $q(P_{i,\sigma(1)}, \sigma), \dots, q(P_{i,\sigma(r)}, \sigma)$ . We call each part of this concatenation a *block*, so  $q(P_i, \sigma)$  consists of  $r$  blocks of size  $r$  each. We observe that  $q(P_i, \sigma) \in S_{r^2}$  and set  $Y_i = \{q(P_i, \sigma) \mid \sigma \in \mathcal{F}\}$ . Indeed,  $|Y_i| = |\mathcal{F}| = g(r)$ , as required.

$Z_i$  is defined analogously to  $Y_i$ , but the concatenation order of the blocks is reversed. More formally, for  $\sigma \in S_r$  let  $r(P_i, \sigma)$  be the concatenation of  $q(P_{i, \sigma(r)}, \sigma), \dots, q(P_{i, \sigma(1)}, \sigma)$  (note: we reverse the concatenation order of the blocks, but the order within each block is not reversed). We observe that  $r(P_i, \sigma) \in S_{r,2}$  and set  $Z_i = \{r(P_i, \sigma) \mid \sigma \in \mathcal{F}\}$ . Indeed,  $|Z_i| = |\mathcal{F}| = g(r)$ , as required.

As an example, consider the case  $r = 4$ . We use here  $\mathcal{F} = \{1234, 1432, 4231, 2413, 3412, 3214\}$  which we constructed initially in the case  $r = 4$  and use the affine plane of Figure 1. Suppose, say, that  $i = 3$  and  $\sigma = (1432) \in \mathcal{F}$ . Then  $P_{3,1} = (4, 5, 10, 13)$  so  $q(P_{3,1}, \sigma) = (4, 13, 10, 5)$ ,  $P_{3,2} = (1, 6, 12, 16)$  so  $q(P_{3,2}, \sigma) = (1, 16, 12, 6)$ ,  $P_{3,3} = (2, 8, 9, 15)$  so  $q(P_{3,3}, \sigma) = (2, 15, 9, 8)$ ,  $P_{3,4} = (3, 7, 11, 14)$  so  $q(P_{3,4}, \sigma) = (3, 14, 11, 7)$ . Concatenating them as ordered by  $\sigma$ , the first block is  $q(P_{3,1}, \sigma)$ , the second block is  $q(P_{3,4}, \sigma)$ , the third block is  $q(P_{3,3}, \sigma)$ , and the fourth block is  $q(P_{3,2}, \sigma)$ . Hence,  $q(P_3, \sigma) = (4, 13, 10, 5, 3, 14, 11, 7, 2, 15, 9, 8, 1, 16, 12, 6)$ . The corresponding  $r(P_3, \sigma) = (1, 16, 12, 6, 2, 15, 9, 8, 3, 14, 11, 7, 4, 13, 10, 5)$  as it reverses the order of the blocks of  $q(P_3, \sigma)$ .

It remains to show that  $\mathcal{F}^*$  is indeed a perfect separating family of  $K_{r,2}$  and is also a perfect 3-sequence covering array of  $[r^2]$ . It is important to stress that, although our goal in Theorem 1.3 is to construct a perfect separating family, our construction needs the fact that it is also a perfect 3-sequence covering array in order for the inductive proof to work (just assuming inductively that  $\mathcal{F}$  is a perfect separating family of  $K_r$  will not suffice to prove that  $\mathcal{F}^*$  is such for  $K_{r,2}$ ).

We begin with proving that  $\mathcal{F}^*$  is a perfect 3-sequence covering array of  $[r^2]$ . The proof of this fact is very similar to the proof of the construction given in [18], but we repeat it for completeness (note, however, that the proof that  $\mathcal{F}^*$  is a perfect separating family given later is more involved and is not needed in the proof given in [18]). So, let  $a, b, c$  be three distinct integers in  $[r^2]$ . We must prove that  $a, b, c$  is a subsequence of precisely  $|\mathcal{F}^*|/6 = g(r^2)/6$  elements of  $\mathcal{F}^*$ . There are two cases to consider: (i)  $\{a, b, c\}$  is contained in some  $P_{i,j}$ . (ii)  $\{a, b, c\}$  is not a subset of any  $P_{i,j}$ .

Case (i): Suppose  $\{a, b, c\}$  appears in some  $P_{i,j}$ . Recall that  $\mathcal{F}$  is a perfect 3-sequence covering array of multiplicity  $|\mathcal{F}|/6 = g(r)/6$ . In each  $q(P_i, \sigma) \in Y_i$ , all  $\{a, b, c\}$  appear in the same block, namely block number  $\sigma^{-1}(j)$ . Inside the block, the internal order of  $\{a, b, c\}$  is their order in  $q(P_{i,j}, \sigma)$ . But by Lemma 2.2,  $P_{i,j}\mathcal{F} = \{q(P_{i,j}, \sigma) \mid \sigma \in \mathcal{F}\}$  is a perfect 3-sequence covering array of  $P_{i,j}$ . Hence,  $a, b, c$  is a subsequence of precisely  $g(r)/6$  elements of  $Y_i$ . Similarly,  $a, b, c$  is a subsequence of precisely  $g(r)/6$  elements of  $Z_i$  (recall that the internal block order of each block of  $r(P_i, \sigma)$  is the same as the internal block order of each block of  $q(P_i, \sigma)$ ). Altogether  $a, b, c$  is a subsequence of  $g(r)/3$  elements of  $W_i$ . Now, if  $i' \neq i$ , then each of  $a, b, c$  appears in a distinct part of  $P_{i'}$  (recall, in an affine plane no pair appears twice in any of the parts). But in  $q(P_{i'}, \sigma)$  the order of these parts is determined by  $\sigma$ . Since  $\mathcal{F}$  is a perfect 3-sequence covering array, precisely in  $g(r)/6$  of the elements of  $Y_{i'}$ , the part containing  $a$  appears before the part containing  $b$  and the part containing  $b$  appears before the part containing  $c$ . Hence  $a, b, c$  is a subsequence of precisely  $g(r)/6$  elements of  $Y_{i'}$ . The same argument holds for  $Z_{i'}$  as the parts just reverse their order (so

$c$  appears before  $b$  and  $b$  appears before  $a$  in  $r(P_{i'}, \sigma)$  if and only if  $a$  is before  $b$  and  $b$  is before  $c$  in  $q(P_{i'}, \sigma)$ . Altogether we have that  $a, b, c$  is a subsequence of precisely  $(r+1)g(r)/3 = g(r^2)/6$  elements of  $\mathcal{F}^*$ .

Case(ii): Suppose  $\{a, b, c\}$  is not a subset of any  $P_{i,j}$ . Let  $\gamma$  be the unique index such that  $\{a, b\}$  is a subset of some part of  $P_\gamma$ , let  $\beta$  be the unique index such that  $\{a, c\}$  is a subset of some part of  $P_\beta$  and let  $\alpha$  be the unique index such that  $\{b, c\}$  is a subset of some part of  $P_\alpha$ . Note that  $\alpha, \beta, \gamma$  are indeed unique and distinct as follows from the properties of an affine plane. As in the previous case, we have that if  $i \notin \{\alpha, \beta, \gamma\}$  then  $a, b, c$  is a subsequence of  $g(r)/3$  elements of  $W_i$ . How many times is  $a, b, c$  a subsequence in  $W_\beta$ ? The answer is 0, since in each element of  $W_\beta$ ,  $a$  and  $c$  appear in the same block while  $b$  appears in another block. How many times is  $a, b, c$  a subsequence of  $W_\alpha$ ? Since  $b, c$  are in the same block of each element of  $W_\alpha$  and since in precisely half of the elements of each of  $Y_\alpha$  and  $Z_\alpha$ ,  $b$  appears before  $c$  (we use here the fact that a perfect 3-sequence covering array is trivially also a perfect 2-sequence covering array) we have that if  $\sigma$  is such that in  $q(P_\alpha, \sigma)$ ,  $b$  appears before  $c$ , then in precisely one of  $q(P_\alpha, \sigma)$  or  $r(P_\alpha, \sigma)$ ,  $a, b, c$  is a subsequence. Hence,  $a, b, c$  is a subsequence of  $g(r)/2$  elements of  $W_\alpha$ . The same argument holds for  $W_\gamma$ . Altogether we have that  $a, b, c$  is a subsequence of precisely  $(r-2)g(r)/3 + 0 + 2g(r)/2 = (r+1)g(r)/3 = g(r^2)/6$  elements of  $\mathcal{F}^*$ .

We now turn to proving that  $\mathcal{F}^*$  is a perfect separating family of  $K_{r,2}$ . So, consider two disjoint edges  $ab$  and  $cd$  of  $K_n$ . We must show that precisely  $|\mathcal{F}^*|/3 = g(r^2)/3$  elements of  $\mathcal{F}^*$  separate them. Here there are more cases to consider, according to the distribution of the 6 pairs induced by  $\{a, b, c, d\}$  among the parts of the affine plane. To effectively examine these cases, it is beneficial to classify how a parallel class  $P_i$  of the affine plane distributes  $a, b, c, d$  among its lines. This classification is made according to *types*, as given in Table 1. As can be seen from that table, there are seven possible types, up to isomorphism. A parallel class  $P_i$  is of type 1 if some line of it contains all of  $a, b, c, d$ . It is of type 2 if some line contains three of  $\{a, b, c, d\}$  and some other line contains the fourth vertex (in Table 1 we write for type 2 that  $\{a, b, c\}$  is the triple contained in one line, but this is isomorphic to any other triple). It is of type 3 if the edge  $ab$  is in one line and the edge  $cd$  is in another line. It is of type 4 if precisely one of the edges  $ab$  or  $cd$  is in some line, and the other two vertices are each in a separate line. It is of type 5 if two pairs are in two lines and neither of these pairs are the edges  $ab$  or  $cd$ . It is of type 6 if one pair is in some line and this pair is not one of the edges  $ab$  or  $cd$ , and the other two vertices are each in a separate line. Finally, it is of type 7 if each vertex is in a separate line.

To illustrate these types, consider the affine plane given in Figure 1 and the parallel class  $P_1$ . Then  $P_1$  is of type 1 with respect, say, to the disjoint edges 12, 34. It is of type 2 w.r.t. 12, 35. It is of type 3 w.r.t. 12, 56. It is of type 4 w.r.t. 12, 59. It is of type 5 w.r.t. 15, 26. It is of type 6 w.r.t. 15, 29. It is of type 7 w.r.t. the disjoint edges whose union is, say,  $\{1, 5, 9, 13\}$ .

Having described the types, we can now consider all possible cases for a pair of disjoint edges



type	distribution of $\{a, b, c, d\}$ among parts	# of times $ab, cd$ separated in $W_i$
1	$\{a, b, c, d\}$	$2g(r)/3$
2	$\{a, b, c\} \{d\}$	$2g(r)/3$
3	$\{a, b\} \{c, d\}$	$2g(r)$
4	$\{a, b\} \{c\} \{d\}$	$4g(r)/3$
5	$\{a, c\} \{b, d\}$	0
6	$\{a, c\} \{b\} \{d\}$	$g(r)/3$
7	$\{a\} \{b\} \{c\} \{d\}$	$2g(r)/3$

Table 1: The possible types (up to isomorphism) of a parallel class of an affine plane with respect to the two disjoint edges  $ab, cd$ . For each type, the number of times  $ab, cd$  are separated in  $W_i$  is given.

$ab, cd$ . For each of the  $r + 1$  parallel classes of the affine plane, we consider its type with respect to  $ab, cd$ , and count how many parallel classes are from each type. This gives a 7-tuple of non-negative integers summing up to  $r + 1$ , where the  $j$ 'th tuple coordinate is the number of parallel classes of type  $j$ . Now, each possible 7-tuple corresponds to a distinct *case*. The possible cases, together with their corresponding 7-tuple, are given in Table 2. As can be seen, there are eight possible cases (not all cases are necessarily realizable in every affine plane). For example, a pair of disjoint edges  $ab, cd$  belongs to Case 1, if some parallel class has type 1 with respect to it (namely, all of  $a, b, c, d$  are in the same line), and all other  $r$  parallel classes are necessarily of type 7, namely each of  $a, b, c, d$  is in a distinct line in each of these classes.

To illustrate some cases, consider again the affine plane given in Figure 1. The pair of edges 12, 34 belongs to Case 1, as  $P_1$  is of type 1 with respect to it and  $P_2, P_3, P_4, P_5$  are each of type 7 with respect to it. The pair 12, 35 belongs to Case 2, as  $P_1$  is of type 2 with respect to it,  $P_4$  is of type 4,  $P_2, P_5$  are of type 6, and  $P_3$  is of type 7. Similarly, one can check that the pair of edges 12, 57 belongs to Case 3, the pair of edges 12, 56 belongs to Case 5 and the pair 13, 89 to Case 7.

case	type 1	type 2	type 3	type 4	type 5	type 6	type 7
1	1						$r$
2		1		1		2	$r - 3$
3			1		2		$r - 2$
4			1		1	2	$r - 3$
5			1			4	$r - 4$
6				2	2		$r - 3$
7				2	1	2	$r - 4$
8				2		4	$r - 5$

Table 2: The possible cases of how the pairs induced by  $\{a, b, c, d\}$  given two disjoint edges  $ab, cd$  are distributed among the parallel classes of an affine plane. For each type, the number of parallel classes of this type is listed, (empty cells correspond to zero).

Suppose now that  $ab, cd$  is a pair of disjoint edges and  $P_i$  is some parallel class. We compute the number of elements of  $W_i$  that separate  $ab$  and  $cd$ . We do this according to  $P_i$ 's type w.r.t.  $ab, cd$ .

*$P_i$  is of type 1.* In this case,  $\{a, b, c, d\}$  appears in some  $P_{i,j}$ . Recall that  $\mathcal{F}$  is a perfect separating family of multiplicity  $|\mathcal{F}|/3 = g(r)/3$ . In each  $q(P_i, \sigma) \in Y_i$ , all  $\{a, b, c, d\}$  appear in the same block, namely block number  $\sigma^{-1}(j)$ . Inside the block, the internal order of  $\{a, b, c, d\}$  is their order in  $q(P_{i,j}, \sigma)$ . But by Lemma 2.2,  $P_{i,j}\mathcal{F} = \{q(P_{i,j}, \sigma) \mid \sigma \in \mathcal{F}\}$  is a perfect separating family of  $P_{i,j}$ . Hence,  $ab$  and  $cd$  are separated by precisely  $g(r)/3$  elements of  $Y_i$ . Similarly,  $a, b, c, d$  are separated by precisely  $g(r)/3$  elements of  $Z_i$  (recall that the internal block order of each block of  $r(P_i, \sigma)$  is the same as the internal block order of each block of  $q(P_i, \sigma)$ ). Altogether  $ab$  and  $cd$  are separated by  $2g(r)/3$  elements of  $W_i$ .

*$P_i$  is of type 2.* Without loss of generality,  $\{a, b, c\}$  appears in some  $P_{i,j}$  and  $d$  appears in some  $P_{i,k}$  with  $k \neq j$ . Since  $\mathcal{F}$  is a perfect 3-sequence covering array of multiplicity  $g(r)/6$ , in precisely  $4g(r)/6$  of the elements of  $q(P_i, \sigma) \in Y_i$ , the block containing  $a, b, c$  does not have  $c$  between  $a$  and  $b$  (so the subsequence on  $\{a, b, c\}$  is one of  $(a, b, c), (b, a, c), (c, a, b), (c, b, a)$ ). For each such element, precisely one of  $q(P_i, \sigma)$  or  $r(P_i, \sigma)$ , separates  $ab$  from  $cd$  (for example, if the subsequence on  $\{a, b, c\}$  is  $(b, a, c)$ , we would like the block containing  $d$  to appear after the block containing  $a, b, c$ , and this happens in precisely one of  $q(P_i, \sigma)$  or  $r(P_i, \sigma)$ ). Altogether  $ab$  and  $cd$  are separated by  $4g(r)/6 = 2g(r)/3$  elements of  $W_i$ .

*$P_i$  is of type 3.* Here  $\{a, b\}$  is in some  $P_{i,j}$  and  $\{c, d\}$  are in some  $P_{i,k}$  with  $k \neq j$ . So each element of  $W_i$  separates  $a, b$  from  $c, d$ . Altogether  $ab$  and  $cd$  are separated by  $|W_i| = 2g(r)$  elements of  $W_i$ .

*$P_i$  is of type 4.* Here  $\{a, b\}$  is in some  $P_{i,j}$ ,  $c \in P_{i,k}$  and  $d \in P_{i,\ell}$  where  $j, k, \ell$  are distinct. In order for  $a, b$  to be separated from  $c, d$  in  $q(P_i, \sigma) \in Y_i$ , we would like the block containing  $a, b$  to not be in between the block containing  $c$  and the block containing  $d$ . But since  $\mathcal{F}$  is a 3-sequence covering array, this occurs in precisely  $2g(r)/3$  elements of  $Y_i$ , and the same holds for  $Z_i$ . Altogether  $ab$  and  $cd$  are separated by  $4g(r)/3$  elements of  $W_i$ .

*$P_i$  is of type 5.* Here  $\{a, c\}$  is in some  $P_{i,j}$  and  $\{b, d\}$  is in some  $P_{i,k}$  with  $k \neq j$ . In this case, no element of  $W_i$  separates  $ab$  from  $cd$  as  $a, b$  are always in the same block and  $c, d$  are always in the same block of each element of  $W_i$ .

*$P_i$  is of type 6.* Here  $\{a, c\}$  is in some  $P_{i,j}$ ,  $b \in P_{i,k}$  and  $d \in P_{i,\ell}$  where  $j, k, \ell$  are distinct. The only way that  $ab$  and  $cd$  could be separated by an element of  $W_i$  is if the block containing  $a, c$  is in between the two blocks containing  $b$  and  $d$ . As  $\mathcal{F}$  is a perfect 3-sequence covering array, this occurs in precisely  $g(r)/3$  of the elements of  $Y_i$ . For each such element  $q(P_i, \sigma)$ , if  $a$  is before  $c$ , we would like the block containing  $b$  to be before the block containing  $d$  (recall that the block containing  $a, c$  is in between them). If this does not occur in  $q(P_i, \sigma)$  then it necessarily occurs in  $r(P_i, \sigma)$  and vice versa. Similarly, if  $c$  is before  $a$  we would like the block containing  $d$  to be before the block

containing  $b$ . If this does not occur in  $q(P_i, \sigma)$  then it necessarily occurs in  $r(P_i, \sigma)$  and vice versa. Altogether  $ab$  and  $cd$  are separated by  $g(r)/3$  elements of  $W_i$ .

$P_i$  is of type 7. Here each of  $a, b, c, d$  is in a distinct block. Since  $\mathcal{F}$  is a perfect separating set, we have that in  $Y_i$ , the two blocks containing  $a$  and  $b$  are separated from the blocks containing  $c$  and  $d$  precisely  $g(r)/3$  times. The same holds for  $Z_i$ . Altogether  $ab$  and  $cd$  are separated by  $2g(r)/3$  elements of  $W_i$ .

We have summarized the analysis above in the third column of Table 1. It is now immediate to verify that every pair  $ab, cd$  of disjoint edges is separated precisely  $|F^*|/3 = g(r^2)/3$  times in  $\mathcal{F}^*$ . Indeed, for each possible case from Table 2, if we take the scalar product of its corresponding 7-tuple given in Table 2 and the third column of Table 1, we obtain  $2(r+1)g(r)/3 = g(r^2)/3$ . For instance, in Case 7 we obtain  $g(r)/3$  times the scalar product  $(0, 0, 0, 2, 1, 2, r-4) \cdot (2, 2, 6, 4, 0, 1, 2)$  giving  $(2r+2)g(r)/3$ . The other cases are similarly verified.  $\square$

The next lemma is a consequence of Lemma 2.3 and is the upper bound part of Theorem 1.3.

**Lemma 2.4.** *Let  $\epsilon > 0$ . There is a constant  $C = C(\epsilon)$  such that for all  $n \geq 3$ ,  $\text{PSD}(K_n) \leq Cn(\log n)^{1+\epsilon}$ .*

*Proof.* We require the following result of Baker et al. [5] which states that there is always a prime strictly between  $x$  and  $x + O(x^{21/40})$ . In particular, for all sufficiently large  $n$ , there is a prime between  $\sqrt{n}$  and  $\sqrt{n} + n^{1/3}$ .

Given  $\epsilon > 0$ , let  $\delta = 2^{\epsilon/4} - 1$ . Let  $N$  be the least integer for which for all  $n \geq N$ , the following hold.

- (i)  $\sqrt{n} + n^{1/3} + 1 \leq (1 + \delta)\sqrt{n}$ .
- (ii) If  $r$  is the smallest prime power satisfying  $r^2 \geq n$  then  $\sqrt{n} \leq r \leq \sqrt{n} + n^{1/3}$ .
- (iii)  $(\log(\sqrt{n} + n^{1/3}))^{1+\epsilon} \leq (\log n)^{1+\epsilon}/2^{1+\epsilon/2}$ .

Let  $C = g(N)$  and observe that indeed  $C = C(\epsilon)$ . By Lemma 2.3 it suffices to prove that  $g(n) \leq Cn(\log n)^{1+\epsilon}$ . Indeed, this clearly holds for all  $3 \leq n \leq N$  since  $g(n) \leq g(N) = C$  in this case. Assume now that  $n > N$  and assume inductively that the claim holds for values smaller than  $n$ .

Let  $r$  be the smallest prime power satisfying  $r^2 \geq n$ . So, we have that  $\sqrt{n} \leq r \leq \sqrt{n} + n^{1/3}$ .

Now, by Lemma 2.3 we have that  $g(n) = g(r^2) = 2(r+1)g(r)$ . By the induction hypothesis,

$$\begin{aligned}
g(n) &\leq 2(r+1)Cr(\log r)^{1+\epsilon} \\
&\leq 2(\sqrt{n} + n^{1/3} + 1)C(\sqrt{n} + n^{1/3})(\log(\sqrt{n} + n^{1/3}))^{1+\epsilon} \\
&\leq 2(1+\delta)^2nC(\log n)^{1+\epsilon}/2^{1+\epsilon/2} \\
&= (1+\delta)^2nC(\log n)^{1+\epsilon}/2^{\epsilon/2} \\
&= Cn(\log n)^{1+\epsilon}.
\end{aligned}$$

□

### 3 A lower bound for a perfect separating family of $K_n$

The following lemma establishes the lower bound in Theorem 1.3.

**Lemma 3.1.** *Let  $n \geq 4$  and let  $\mathcal{F}$  be a perfect separating family of  $K_n$ . Then  $|\mathcal{F}| \geq n/2 - 1$ .*

*Proof.* As the lemma holds trivially for  $n = 4$ , we assume that  $n \geq 5$ . For a permutation  $\sigma = (a_1, \dots, a_n)$  of vertices of  $K_n$ , we say that the pairs  $\{a_1, a_2\}$  and  $\{a_{n-1}, a_n\}$  are extremal in  $\sigma$ . Let  $\lambda = |\mathcal{F}|/3$  be the multiplicity of  $\mathcal{F}$ . We say that a pair  $\{u, v\}$  of vertices is *extremal* in  $\mathcal{F}$  if  $\{u, v\}$  is extremal in  $\lambda$  elements of  $\mathcal{F}$ . As each  $\sigma \in \mathcal{F}$  has two pairs that are extremal in  $\sigma$ , the overall count of extremal pairs is  $2|\mathcal{F}| = 6\lambda$ , so there can be at most 6 pairs that are extremal in  $\mathcal{F}$ . But since  $n \geq 5$ , there are more than 6 pairs of vertices so let  $\{u, v\}$  be a pair of vertices that is *not* extremal in  $\mathcal{F}$ . We may assume by relabeling that  $\{u, v\} = \{n-1, n\}$ .

We next define several matrices over  $\mathbb{R}$ . Let  $A$  be the matrix whose rows are indexed by  $[n-2]$  (i.e. all vertices except  $u, v$ ) and whose columns are indexed by  $\mathcal{F}$ . Set  $A[w, \sigma] = 1$  if  $\sigma^{-1}(w) < \min\{\sigma^{-1}(u), \sigma^{-1}(v)\}$ . Otherwise, set  $A[w, \sigma] = 0$ . Let  $B$  be the matrix whose rows are indexed by  $[n-2]$  and whose columns are indexed by  $\mathcal{F}$ . Set  $B[w, \sigma] = 1$  if  $\sigma^{-1}(w) > \max\{\sigma^{-1}(u), \sigma^{-1}(v)\}$ . Otherwise, set  $B[w, \sigma] = 0$ . Clearly,

$$\text{rank}(AA^T) \leq \text{rank}(A) \leq |\mathcal{F}|, \quad \text{rank}(BB^T) \leq \text{rank}(B) \leq |\mathcal{F}|.$$

Let  $C = AA^T + BB^T$ . So,  $C$  is a square matrix of order  $n-2$ . We prove that  $C$  is non-singular. Notice that this suffices since if so, we have that

$$\text{rank}(C) \leq \text{rank}(AA^T) + \text{rank}(BB^T) \leq 2|\mathcal{F}| \tag{2}$$

implying that  $|\mathcal{F}| \geq n/2 - 1$ .

So it remains to prove that  $C$  is non-singular. Consider first some off-diagonal entry of  $AA^T$ , say  $(AA^T)[x, y]$ . Then this entry counts the number of elements of  $\sigma \in \mathcal{F}$  in which both vertices

$x$  and  $y$  appear in  $\sigma$  before  $u, v$ . Similarly,  $(BB^T)[x, y]$  counts the number of elements of  $\sigma \in \mathcal{F}$  in which both vertices  $x$  and  $y$  appear in  $\sigma$  after  $u, v$ . Therefore, since  $\mathcal{F}$  is a perfect separating family, we have that  $(AA^T)[x, y] + (BB^T)[x, y] = \lambda$ . Hence, all the off-diagonal entries of  $C$  equal  $\lambda$ .

A diagonal entry of  $C$ , say  $C[w, w]$  equals the number of elements  $\sigma \in \mathcal{F}$  in which  $w$  is either before both  $u, v$  or after both  $u, v$ . But notice that if  $w' \in [n] - \{u, v, w\}$  then  $w'w$  is separated from  $uv$  precisely  $\lambda$  times, and we therefore have, in particular, that  $C[w, w] \geq \lambda$ . We claim, however, that  $C[w, w] > \lambda$ . Assume otherwise, that  $C[w, w] = \lambda$ . So, there is a subset  $\mathcal{F}_w \subset \mathcal{F}$  of precisely  $\lambda$  elements of  $\mathcal{F}$  in which  $w$  is either before both  $u$  and  $v$  or after both  $u$  and  $v$ . Consider again some  $w' \in [n] - \{u, v, w\}$ . For any  $\sigma \in \mathcal{F} \setminus \mathcal{F}_w$ , we have that  $w$  is in between  $u$  and  $v$  and therefore  $w'w$  is not separated from  $uv$  by  $\sigma$ . As  $w'w$  and  $uv$  need to be separated  $\lambda$  times, this implies that precisely all elements of  $\mathcal{F}_w$  separate  $w'w$  and  $uv$ . Hence,  $w'$  and  $w$  must both either be before both  $u$  and  $v$  or else both be after both  $u$  and  $v$ . As this holds for all  $w' \in [n] - \{u, v, w\}$ , this means that  $\{u, v\}$  is extremal in each  $\sigma \in \mathcal{F}_w$ . Hence  $\{u, v\}$  is extremal in  $\mathcal{F}$  precisely  $\lambda$  times, contradicting the choice of  $\{u, v\}$ . We have proved that  $C[w, w] > \lambda$ . for all  $w \in [n - 2]$ .

To summarize,  $C$  is a matrix whose off-diagonal entries are all equal to  $\lambda$  and whose diagonal entries are all greater than  $\lambda$ . It is an easy exercise that such matrices are non-singular. □

## 4 Perfect separation dimension of random graphs

The following definition plays a central role in the proof of Theorem 1.5.

**Definition 4.1** (minrank). *The minrank of a graph  $G$  on vertex set  $[n]$  over a field  $\mathbb{F}$ , denoted  $\text{minrank}_{\mathbb{F}}(G)$ , is the smallest possible rank of a matrix  $M \in \mathbb{F}^{n \times n}$  with nonzero diagonal entries such that  $M[u, v] = 0$  for any pair  $u, v$  of distinct nonadjacent vertices of  $G$ .*

The notion of minrank (over  $\mathbb{R}$ ) was first studied by Lovász [13] and has found several applications, see [2, 10]. We need the following result of Alon et al. [2] regarding the minrank of a random graph over  $\mathbb{R}$  (in fact, over every field), extending an earlier similar result over finite fields of Golovnev, Regev, and Weinstein [10].

**Lemma 4.2** ([2]). *Assume that  $q = q(n)$  satisfies  $1/n \leq q \leq 1$ . Then w.h.p.*

$$\text{minrank}_{\mathbb{R}}(G(n, q)) \geq \frac{n \log(1/q)}{80 \log n}.$$

Before proving Theorem 1.5, we need to establish some properties of  $G(n, p)$  that hold w.h.p.

**Lemma 4.3.** *Let  $t = \lceil 3 \ln n/p \rceil$ . It holds w.h.p. in  $G(n, p)$  that for any two disjoint vertex sets  $A, B$  of order  $t$  each, there is an edge with one endpoint in  $A$  and one endpoint in  $B$ .*

*Proof.* Consider two disjoint vertex sets  $A, B$  of order  $t$  each. The probability that no edge connects a vertex of  $A$  with a vertex of  $B$  is  $(1 - p)^{t^2}$ . As there are less than  $n^{2t}$  choices for  $A$  and  $B$ , the probability that there is some pair violating the stated property is at most

$$n^{2t}(1 - p)^{t^2} \leq n^{2t} \left(1 - \frac{3 \ln n}{t}\right)^{t^2} \leq n^{-t} = o(1).$$

□

**Lemma 4.4.** *W.h.p. it holds that any set  $X$  of at least  $9 \ln n/p$  vertices of  $G(n, p)$  induces a connected subgraph of order at least  $|X|/3$ .*

*Proof.* Let  $X$  be a set of vertices with  $|X| = x \geq 9 \ln n/p$ . Assume that  $X$  does not induce a connected subgraph of order at least  $x/3$ . Then we can partition  $X$  into two parts  $Y$  and  $X \setminus Y$  with  $2x/3 \geq |Y| \geq |X - Y| \geq x/3$  and there is no edge between  $X$  and  $Y$ . But this implies, in particular, that we have  $A \subset Y$  and  $B \subseteq X \setminus Y$  with  $|A| = |B| = t = \lceil 3 \ln n/p \rceil$  and no edge connecting them, and we have just proved in Lemma 4.3 that w.h.p. this does not occur in  $G(n, p)$ . □

For a permutation  $\sigma$  of vertices of an  $n$ -vertex graph  $G$ , we say that a pair of vertices  $\{u, v\}$  is  $k$ -extremal in  $\sigma$  if  $\max\{\sigma^{-1}(u), \sigma^{-1}(v)\} \leq k$  or  $\min\{\sigma^{-1}(u), \sigma^{-1}(v)\} \geq n - k + 1$ . Let  $\mathcal{F}$  be a perfect separating family of  $G$ , with multiplicity  $\lambda$ . We say that an edge  $uv$  of  $G$  is  $k$ -extremal in  $\mathcal{F}$  if  $\{u, v\}$  is  $k$ -extremal in at least  $\lambda$  elements of  $\mathcal{F}$ .

**Lemma 4.5.** *Assume that  $p = p(n)$  satisfies  $n^{-0.4} \leq p < 1$ . Then w.h.p. the following hold for every perfect separating family  $\mathcal{F}$  of  $G \sim G(n, p)$ : Either  $|\mathcal{F}| \geq np/4$  or else there is an edge  $uv$  of  $G$  such that  $uv$  is not  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\mathcal{F}$ .*

*Proof.* Assume that  $|\mathcal{F}| < np/4$ . For each  $\sigma \in \mathcal{F}$ , there are at most  $2 \binom{\lfloor \sqrt{n} \rfloor}{2} < n$  pairs that are  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\sigma$ . The overall count of  $\lfloor \sqrt{n} \rfloor$ -extremal pairs is therefore less than  $n|\mathcal{F}| < n^2 p/4$ . But the expected number of edges in  $G$  is  $p \binom{n}{2} > pn^2/3$  and distributed binomially  $\mathcal{B}(\binom{n}{2}, p)$ , so w.h.p.  $G$  has more than  $n^2 p/4$  edges. Hence, w.h.p. there is an edge  $uv$  of  $G$  such that  $uv$  is not  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\mathcal{F}$ . □

Notice that the last lemma holds also for values of  $p$  significantly smaller than  $n^{-0.4}$ , but as later considerations require this assumption, we preferred to state the lemma in this way.

*Proof of Theorem 1.5.* According to the theorem's statement, we assume that  $p = p(n)$  satisfies  $n^{-0.4} \leq p < 1$ . We may also assume that  $p \leq 1 - 1/n$  as otherwise, each vertex of  $G(n, p)$  has probability at least  $1/e$  of having degree  $n - 1$ , so (e.g. using Chebyshev's inequality) w.h.p.

$G(n, p)$  has a clique of order at least  $n/4$  and the result follows from Theorem 1.3. Hereafter,  $n^{-0.4} \leq p < 1 - 1/n$  so  $1/n \leq q = 1 - p \leq 1 - n^{-0.4}$ .

Let  $G \sim G(n, p)$  and let  $\mathcal{F}$  be a perfect separating family of  $G$  with multiplicity  $\lambda$ . By Lemmas 4.3, 4.4, 4.5 we may assume that  $G$  satisfies the statements of each of these lemmas. Furthermore,  $\overline{G}$  (the complement of  $G$ ) is chosen uniformly from  $G(n, q)$  so we may also assume by Lemma 4.2 that  $\text{minrank}_{\mathbb{R}}(\overline{G}) \geq \frac{n \log(1/q)}{80 \log n}$ . Finally, if  $|\mathcal{F}| \geq np/4$  then the statement of the theorem holds, so we may assume that  $|\mathcal{F}| < np/4$  and hence by Lemma 4.5, there is an edge  $uv$  of  $G$  such that  $uv$  is not  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\mathcal{F}$ . We may assume by relabeling that  $\{u, v\} = \{n-1, n\}$ . We prove that  $\text{PSD}(G) \geq \frac{np}{400 \ln n}$ .

As in the proof of Lemma 3.1, we consider the matrices  $A, B, C = AA^T + BB^T$  defined there. We will prove that  $\text{rank}(C) \geq \frac{np}{200 \ln n}$  and obtain using (2) that  $|\mathcal{F}| \geq \frac{np}{400 \ln n}$  as required. First observe that for each edge  $xy$  of  $G$  it holds that  $C[x, y] = \lambda$ . However, unlike the proof of Lemma 3.1, we have no information regarding the off-diagonal entries of  $C$  that correspond to non-edges of  $G$  (equivalently, edges of  $\overline{G}$ ). We next consider diagonal entries of  $C$ . We first claim that  $C[w, w] \geq \lambda$  for each  $w \in [n-2]$ . To see this, observe that the expected minimum degree of  $G$  is  $\Theta(np)$  and hence w.h.p. the minimum degree of  $G$  is at least 3, so we may assume this is the case. But this implies that  $w$  has some neighbor  $w' \notin \{u, v\}$ . As the edge  $w'w$  is separated from  $uv$  precisely  $\lambda$  times, and as each  $\sigma \in \mathcal{F}$  which separates them contributes to  $C[w, w]$  (as  $w$  must either be before both  $u, v$  in  $\sigma$  or after both  $u, v$  in  $\sigma$ ), we have that  $C[w, w] \geq \lambda$ . However, it would also be necessary to prove that many diagonal entries are, in fact, larger than  $\lambda$  which is what we show next.

Let  $W \subseteq [n-2]$  be the set of vertices with  $w \in W$  if  $C[w, w] = \lambda$ . We show that  $|W| \leq 9 \ln n/p$ . Assume otherwise, then by Lemma 4.4, there is a subset  $W' \subset W$  with  $|W'| = t = \lceil 3 \ln n/p \rceil$  such that  $W'$  induces a connected subgraph in  $G$ . Consider some  $w \in W'$  and let  $\mathcal{F}_w \subseteq \mathcal{F}$  consist of all  $\sigma \in \mathcal{F}$  such that  $w$  is either before both  $u, v$  in  $\sigma$  or after both  $u, v$  in  $\sigma$ . Then, by the definition of  $W$ , we have that  $|\mathcal{F}_w| = \lambda$ . Now, let  $w' \in W'$  be any neighbor of  $w$ . Then  $w'w$  is separated from  $uv$  precisely by the elements of  $\mathcal{F}_w$ , and since  $w' \in W$  this implies that  $\mathcal{F}_w = \mathcal{F}_{w'}$ . But since  $W'$  induces a connected subgraph, we have that  $\mathcal{F}_w = \mathcal{F}_{w'}$  for any pair of distinct elements of  $W'$  so we can actually denote this common subset by  $\mathcal{F}_{W'}$  as it is independent of the choice of  $w \in W'$ . Now let  $W^* \subseteq [n-2]$  be the set of vertices having a neighbor in  $W'$  (in particular  $W' \subseteq W^*$ ). It follows that for each  $\sigma \in \mathcal{F}_{W'}$ , all vertices in  $W^*$  are either all before  $u, v$  in  $\sigma$  or all after both  $u, v$  in  $\sigma$ . It now follows from Lemma 4.3 that  $[n-2] \setminus W^*$  contains less than  $t$  elements as otherwise  $A = W'$  and  $B \subset [n-2] \setminus W^*$  with  $|B| = t$  violate the statement of the lemma. But this, in turn, implies that  $\{u, v\}$  is  $(t+1)$ -extremal in  $\sigma$ . Indeed, if all elements of  $W^*$  are after  $u, v$  in  $\sigma$  then we have  $\max\{\sigma^{-1}(u), \sigma^{-1}(v)\} \leq t+1$  and if all elements of  $W^*$  are before  $u, v$  in  $\sigma$  then we have  $\min\{\sigma^{-1}(u), \sigma^{-1}(v)\} \geq n-t$ . Now, since  $t+1 = \lceil 3 \ln n/p \rceil + 1 \leq \lfloor \sqrt{n} \rfloor$  (here we use that  $p \geq n^{-0.4}$ ) we have that  $\{u, v\}$  is  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\sigma$ . As this holds for each  $\sigma \in \mathcal{F}_{W'}$ , and since

$|F_{W'}| = \lambda$ , we have that  $\{u, v\}$  is  $\lfloor \sqrt{n} \rfloor$ -extremal in  $\mathcal{F}$ , contradicting our choice of  $u, v$ . We have proved that  $|W| \leq 9 \ln n/p$ .

Consider now the matrix  $D = C - \lambda J$  where  $J$  is the all-one matrix. Then we have that for each edge  $xy$  of  $G$ , equivalently, each non-edge  $xy$  of  $\overline{G}$ , it holds that  $D[x, y] = 0$ . Furthermore, at most  $|W| \leq 9 \ln n/p$  diagonal entries of  $D$  are zero. Arbitrarily set the zero diagonal entries of  $D$  to nonzero values, obtaining a matrix  $D^*$ . As  $D$  and  $D^*$  differ in at most  $9 \ln n/p$  entries, we have that  $\text{rank}(D) \geq \text{rank}(D^*) - 9 \ln n/p$ . Furthermore, by Lemma 4.2 applied to  $\overline{G}$  we have that  $\text{rank}(D^*) \geq \frac{n \log(1/q)}{80 \log n}$ . We therefore have that

$$\begin{aligned} \text{rank}(C) &\geq \text{rank}(D) - 1 \\ &\geq \text{rank}(D^*) - \frac{9 \ln n}{p} - 1 \\ &\geq \frac{n \log(1/q)}{80 \log n} - \frac{9 \ln n}{p} - 1 \\ &\geq \frac{n \log(1+p)}{80 \log n} - \frac{9 \ln n}{p} - 1 \\ &\geq \frac{np}{160 \ln n} - \frac{9 \ln n}{p} - 1 \\ &\geq \frac{np}{200 \ln n} \end{aligned}$$

where in the last inequality we have used that  $p \geq n^{-0.4}$ . □

## 5 $\Delta$ -balanced separation dimension

Here we prove Theorem 1.6. An important ingredient in our proof is the following result of Alon [1].

**Lemma 5.1** ([1]). *There exists an absolute positive constant  $\mu$  so that the following holds. Let  $D$  be an  $n$  by  $n$  real matrix with  $|D[i, i]| \geq \frac{1}{2}$  for all  $i$  and  $|D[i, j]| \leq \epsilon$  for all  $i \neq j$ , where  $\frac{1}{2\sqrt{n}} \leq \epsilon < \frac{1}{4}$ . Then*

$$\text{rank}(D) \geq \frac{\mu \log n}{\epsilon^2 \log(1/\epsilon)}.$$

We also need the following lemma, which is the lower bound for separation dimension of  $K_n$  proved by Basavaraju et al. [6].

**Lemma 5.2** ([6]).

$$\pi(K_n) \geq \log \lfloor n/2 \rfloor.$$

Let  $\mathcal{F}$  be a separating family of  $K_n$ . For a pair of distinct vertices  $u, v$  and for a vertex  $w \notin \{u, v\}$ , let  $s_{\mathcal{F}}(u, v, w)$  be the number of elements  $\sigma \in \mathcal{F}$  such that  $w$  is before both  $u, v$  in  $\sigma$  or after both  $u, v$  in  $\sigma$ . Let  $s_{\mathcal{F}}(u, v) = \sum_{w \in [n] \setminus \{u, v\}} s_{\mathcal{F}}(u, v, w)$ .



**Lemma 5.3.** *There exists a pair of distinct vertices  $u, v$  such that  $s_{\mathcal{F}}(u, v) \geq \frac{2}{3}|\mathcal{F}|(n-2)$ .*

*Proof.* Let  $S$  be the sum of  $s_{\mathcal{F}}(u, v)$  taken over all pairs of distinct vertices  $u, v$ . Each  $\sigma \in \mathcal{F}$  contributes precisely  $\frac{1}{3}n(n-1)(n-2)$  to  $S$  implying that  $S = \frac{1}{3}|\mathcal{F}|n(n-1)(n-2)$ . By averaging, there is pair  $u, v$  with  $s_{\mathcal{F}}(u, v) \geq |S|/\binom{n}{2} = \frac{2}{3}|\mathcal{F}|(n-2)$ .  $\square$

*Proof of Theorem 1.6.* Let  $\mathcal{F}$  be a separating family of  $K_n$  which is  $\Delta$ -balanced for  $\Delta \leq c \log n$  where  $c$  is an absolute constant to be chosen later. Observe first that for any pair of disjoint edges  $e, f$ , it holds that

$$\left| c_{\mathcal{F}}(\{e, f\}) - \frac{|\mathcal{F}|}{3} \right| \leq c \log n. \quad (3)$$

Indeed, suppose  $e = xy$  and  $f = zw$ . Then (3) holds since each element of  $\mathcal{F}$  separates precisely one of the three pairs of edges  $\{xy, zw\}$ ,  $\{xz, yw\}$ ,  $\{xw, yz\}$  and since  $\mathcal{F}$  is  $\Delta$ -balanced.

Let  $\frac{1}{48} > c > 0$  be a real number such that for all  $\alpha \geq \frac{1}{2}$  it holds that

$$\frac{\mu\alpha}{216c^2 \log(\frac{\alpha}{6c})} < 1 \quad (4)$$

where  $\mu$  is the constant from Lemma 5.1.

We will prove that  $|\mathcal{F}| \geq n/K$  where  $K = 3/(2\mu)$ . Assume hereafter, in contradiction, that  $|\mathcal{F}| \leq n/K$ .

Let  $u, v$  be two vertices with  $s_{\mathcal{F}}(u, v) \geq \frac{2}{3}|\mathcal{F}|(n-2)$  which exist by Lemma 5.3. By relabeling, we may assume that  $\{u, v\} = \{n-1, n\}$ . As in the proof of Lemma 3.1, we consider the matrices  $A, B, C = AA^T + BB^T$  defined there. Each diagonal entry  $C[w, w]$  counts the number of  $\sigma \in \mathcal{F}$  such that  $w$  is either before both  $u, v$  or after both  $u, v$ . Hence,  $\text{tr}(C) = s_{\mathcal{F}}(u, v) \geq \frac{2}{3}|\mathcal{F}|(n-2)$ . So, on average, a diagonal element has value  $\frac{2}{3}|\mathcal{F}|$  while trivially, each diagonal entry does not exceed  $|\mathcal{F}|$ . Let  $W \subseteq [n-2]$  be the set of vertices with  $C[w, w] \geq |\mathcal{F}|/2$ . Then we have that  $|W| \geq (n-2)/3$  as otherwise  $\text{tr}(C) < \frac{2}{3}|\mathcal{F}|(n-2)$ . Let  $C_W$  be the submatrix of  $C$  obtained by restricting the rows and columns to  $W$ . Then the dimension of  $C_W$  is at least  $(n-2)/3$  and clearly  $\text{rank}(C_W) \leq \text{rank}(C)$ . We will lower-bound  $\text{rank}(C_W)$  and in fact show that  $\text{rank}(C_W) > 2|\mathcal{F}|$  which contradicts the fact that  $\text{rank}(C) \leq 2|\mathcal{F}|$  by (2).

Let  $\alpha$  be such that  $|\mathcal{F}| = \alpha \log n$ . Observe that by Lemma 5.2 we have that  $\alpha \geq \log\lfloor n/2 \rfloor / \log n$  (so  $\alpha > \frac{1}{2}$ ) and by our assumption,  $\alpha \leq n/(K \log n)$ . We consider two cases, according to  $\alpha$ . The first case is  $6c\sqrt{n} \geq \alpha \geq \log\lfloor n/2 \rfloor / \log n$  and the second case is  $n/(K \log n) \geq \alpha > 6c\sqrt{n}$ .

*Case 1.* Here  $6c\sqrt{n} \geq \alpha \geq \log\lfloor n/2 \rfloor / \log n$ . So,  $C_W$  is a matrix of dimension  $|W|$  whose diagonal entries are all at least  $|\mathcal{F}|/2$  and by (3), the off-diagonal entries differ by at most  $c \log n$  from  $|\mathcal{F}|/3$ . Let  $D = \frac{6}{|\mathcal{F}|}(C_W - \frac{|\mathcal{F}|}{3}J)$  where  $J$  is the all-one matrix. So the diagonal entries of  $D$  are all at least 1 and the off-diagonal entries are all, in absolute value, at most  $\frac{6}{|\mathcal{F}|}c \log n = 6c/\alpha$ . Furthermore,

we have

$$\frac{1}{2\sqrt{|W|}} \leq \frac{1}{2\sqrt{(n-2)/3}} \leq \frac{1}{\sqrt{n}} \leq \frac{6c}{\alpha} < \frac{1}{4}.$$

By Lemma 5.1 and by the choice of  $c$  satisfying (4) we obtain that

$$\begin{aligned} \text{rank}(D) &\geq \frac{\mu\alpha^2 \log |W|}{36c^2 \log(\alpha/(6c))} \\ &\geq \frac{\mu\alpha^2 \log n}{72c^2 \log(\alpha/(6c))} \\ &\geq 3\alpha \log n \\ &= 3|\mathcal{F}|. \end{aligned}$$

But since  $\text{rank}(C) \geq \text{rank}(C_W) \geq \text{rank}(D) - 1 \geq 3|\mathcal{F}| - 1$  we have that  $\text{rank}(C) > 2|\mathcal{F}|$ , contradicting the fact that  $\text{rank}(C) \leq 2|\mathcal{F}|$  by (2).

*Case 2.* Here  $n/(K \log n) \geq \alpha > 6c\sqrt{n}$ . This time we define  $D = \frac{1}{\sqrt{nc \log n}}(C_W - \frac{|\mathcal{F}|}{3}J)$ . So the diagonal entries of  $D$  are all at least 1 and the off-diagonal entries are all, in absolute value, at most  $\frac{1}{\sqrt{n}} \geq \frac{1}{2\sqrt{|W|}}$ . So, we can apply Lemma 5.1 with  $\epsilon = 1/\sqrt{n}$  and obtain that

$$\text{rank}(D) \geq \frac{\mu n \log n}{\log(\sqrt{n})} = 2\mu n = 3n/K \geq 3|\mathcal{F}|.$$

Concluding as in the previous case, we arrive at the exact same contradiction. □

## Acknowledgment

I thank Abhiruk Lahiri for interesting discussions on separation dimension.

## References

- [1] Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability & Computing*, 18(1-2):3, 2009.
- [2] Noga Alon, Igor Balla, Lior Gishboliner, Adva Mond, and Frank Mousset. The minrank of random graphs over arbitrary fields. *Israel Journal of Mathematics*, 235(1):63–77, 2020.
- [3] Noga Alon, Manu Basavaraju, L. Sunil Chandran, Rogers Mathew, and Deepak Rajendraprasad. Separation dimension of bounded degree graphs. *SIAM Journal on Discrete Mathematics*, 29(1):59–64, 2015.

- [4] Noga Alon, Manu Basavaraju, L. Sunil Chandran, Rogers Mathew, and Deepak Rajendraprasad. Separation dimension and sparsity. *Journal of Graph Theory*, 89(1):14–25, 2018.
- [5] Roger Baker, Glin Harman, and János Pintz. The difference between consecutive primes, II. *Proceedings of the London Mathematical Society*, 83(03):532–562, 2001.
- [6] Manu Basavaraju, L. Sunil Chandran, Martin Charles Golumbic, Rogers Mathew, and Deepak Rajendraprasad. Boxicity and separation dimension. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 81–92. Springer, 2014.
- [7] Arpitha P. Bharathi, Minati De, and Abhiruk Lahiri. Circular separation dimension of a subclass of planar graphs. *Discrete Mathematics and Theoretical Computer Science*, 19, 2017.
- [8] L. Sunil Chandran, Rogers Mathew, and Naveen Sivadasan. Boxicity of line graphs. *Discrete Mathematics*, 311(21):2359–2367, 2011.
- [9] Zoltan Füredi. Scrambling permutations and entropy of hypergraphs. *Random Structures & Algorithms*, 8(2):97–104, 1996.
- [10] Alexander Golovnev, Oded Regev, and Omri Weinstein. The minrank of random graphs. *IEEE Transactions on Information Theory*, 64(11):6990–6995, 2018.
- [11] Thomas P. Kirkman. Note on an unanswered prize question. *Cambridge and Dublin Mathematics Journal*, 5:255–262, 1850.
- [12] Sarah J Loeb and Douglas B West. Fractional and circular separation dimension of graphs. *European Journal of Combinatorics*, 69:19–35, 2018.
- [13] László Lovász. On the shannon capacity of a graph. *IEEE Transactions on Information theory*, 25(1):1–7, 1979.
- [14] Jaikumar Radhakrishnan. A note on scrambling permutations. *Random Structures & Algorithms*, 22(4):435–439, 2003.
- [15] Alex Scott and David R. Wood. Separation dimension and degree. In *Mathematical Proceedings of the Cambridge Philosophical Society*, pages 1–10. Cambridge University Press, 2018.
- [16] Joel Spencer. Minimal scrambling sets of simple orders. *Acta Mathematica Hungarica*, 22(3-4):349–353, 1972.
- [17] Jun Tarui. On the minimum number of completely 3-scrambling permutations. *Discrete Mathematics*, 308(8):1350–1354, 2008.

- [18] Raphael Yuster. Perfect sequence covering arrays. *Designs, Codes and Cryptography*, 88(3):585–593, 2020.
- [19] Emile Ziedan, Deepak Rajendraprasad, Rogers Mathew, Martin Charles Golumbic, and Jérémie Dusart. The induced separation dimension of a graph. *Algorithmica*, 80(10):2834–2848, 2018.