

# On Connectivity in Random Graph Models with Limited Dependencies

Johannes Lengler  

Department of Computer Science,  
ETH Zürich, Switzerland

Kalina Petrova  

Department of Computer Science,  
ETH Zürich, Switzerland

Raphael Steiner  

Department of Computer Science,  
ETH Zürich, Switzerland

Emo Welzl 

Department of Computer Science,  
ETH Zürich, Switzerland

Anders Martinsson  

Department of Computer Science,  
ETH Zürich, Switzerland

Patrick Schnider  

Department of Computer Science,  
ETH Zürich, Switzerland

Simon Weber  

Department of Computer Science,  
ETH Zürich, Switzerland

---

## Abstract

For any positive edge density  $p$ , a random graph in the Erdős-Rényi  $G_{n,p}$  model is connected with non-zero probability, since all edges are mutually independent. We consider random graph models in which edges that do not share endpoints are independent while incident edges may be dependent and ask: what is the minimum probability  $\rho(n)$ , such that for any distribution  $\mathcal{G}$  (in this model) on graphs with  $n$  vertices in which each potential edge has a marginal probability of being present at least  $\rho(n)$ , a graph drawn from  $\mathcal{G}$  is connected with non-zero probability?

As it turns out, the condition “edges that do not share endpoints are independent” needs to be clarified and the answer to the question above is sensitive to the specification. In fact, we formalize this intuitive description into a strict hierarchy of five independence conditions, which we show to have at least three different behaviors for the threshold  $\rho(n)$ . For each condition, we provide upper and lower bounds for  $\rho(n)$ . In the strongest condition, the *coloring model* (which includes, e.g., random geometric graphs), we show that  $\rho(n) \rightarrow 2 - \phi \approx 0.38$  for  $n \rightarrow \infty$ , proving a conjecture by Badakhshian, Falgas-Ravry, and Sharifzadeh. This separates the coloring models from the weaker independence conditions we consider, as there we prove that  $\rho(n) > 0.5 - o(n)$ . In stark contrast to the coloring model, for our weakest independence condition – pairwise independence of non-adjacent edges – we show that  $\rho(n)$  lies within  $O(1/n^2)$  of the threshold  $1 - 2/n$  for completely arbitrary distributions.

**2012 ACM Subject Classification** Mathematics of computing → Random graphs; Theory of computation → Randomness, geometry and discrete structures

**Keywords and phrases** Random Graphs, Independence, Dependency, Connectivity, Threshold, Probabilistic Method

**Digital Object Identifier** 10.4230/LIPIcs.APPROX/RANDOM.2023.30

**Category** RANDOM

**Related Version** *Full Version*: <https://arxiv.org/abs/2305.02974>

**Funding** *Kalina Petrova*: Swiss National Science Foundation under grant no. CRSII5 173721.

*Raphael Steiner*: ETH Zürich Postdoctoral Fellowship.

*Simon Weber*: Swiss National Science Foundation under project no. 204320.

**Acknowledgements** This research started at the joint workshop of the *Combinatorial Structures and Algorithms* and *Theory of Combinatorial Algorithms* groups of ETH Zürich held in Stels, Switzerland, January 2023. We thank the organizers for providing a very pleasant and inspiring working atmosphere. We thank Victor Falgas-Ravry for pointing out an abundance of related work, thus helping tremendously to put our results in perspective.



© Johannes Lengler, Anders Martinsson, Kalina Petrova, Patrick Schnider, Raphael Steiner, Simon Weber, and Emo Welzl;  
licensed under Creative Commons License CC-BY 4.0

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023).

Editors: Nicole Megow and Adam D. Smith; Article No. 30; pp. 30:1–30:22



Leibniz International Proceedings in Informatics  
LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

## 1 Introduction

The probabilistic method is an important tool in theoretical computer science, graph theory and combinatorics [3]. With this method, one proves that a random construction has some desirable property with positive probability, and concludes that some objects with this property must exist. Early examples of the probabilistic method include simple constructions of expander graphs [7], or graphs with high girth and large chromatic number [16]. Often, it is possible to express the property as the intersection of very simple events  $A_1, \dots, A_k$ . For example, we can express the property that a set  $S$  is a clique as the intersection of the events “ $u$  is adjacent to  $v$ ” for all pairs  $u, v \in S$ . In many cases it is clear that each of the  $A_i$  has a large probability to occur, but this is generally not enough to conclude that the intersection of all  $A_i$  has positive probability to occur. Of course, if the probability space were a *product space* in which all components are *independent* of each other, then this would be trivial.

However, for many applications it is too limiting to restrict oneself to product spaces. The probabilistic method was significantly extended by the realization that it could still be applied to settings without perfect independence, provided there is some bound on the amount of dependence in the system. The seminal result was the Lovász Local Lemma (LLL), which we give here in the slightly stronger version due to Shearer [28]: for events  $A_1, \dots, A_k$  that all occur with probability at least  $p$ , if each of the  $A_i$  depends on at most  $d$  other  $A_j$ , where  $p \geq 1 - 1/(ed)$ , then there is a positive probability that all of the  $A_i$  occur simultaneously.<sup>1</sup> In this context, the dependencies are captured by a *dependency graph*, a graph with vertex set  $\{A_1, \dots, A_k\}$  such that each vertex is independent from all but its neighbors. There are many situations in which the dependency graph can be restricted, and we give some examples in Section 1.1 below.

The LLL allows to re-introduce a product space through the backdoor, because the LLL condition  $p \geq 1 - 1/(ed)$  allows to couple the process to a product space. This was already implicit in the inductive proof of the LLL [29], and was made explicit by Liggett, Schonmann and Spacey [23]. They also generalized this coupling to a countably infinite number of variables and showed tightness of Shearer’s condition (of the precise version in footnote 1). For a finite number of variables, coupling the probability space to a product space implies trivially that the all-one event has positive probability.

Unfortunately, the LLL scales badly in  $d$ , i.e.,  $p$  needs to be very close to one if  $d$  is large. However, in some cases the degree  $d$  may grow. In particular, in this paper we will study the situation that the variables are associated with the edges of a complete graph on  $n$  vertices, and dependencies only run between adjacent edges (edges which share a common endpoint). This is a common situation, and examples are given in Section 1.1. In this case, the number of dependencies per edge is  $2(n - 2)$ , and thus grows with  $n$ . Thus, if every edge is present with constant probability  $p < 1$ , then it is not possible to couple the probability space with a product space for large  $n$ , and it is not true that the complete graph appears with positive probability [23]. However, in this paper we will show that some weaker global properties *can* be guaranteed. Specifically, we will focus on connectivity because this is arguably the most fundamental global graph property. We will study the question:

Consider a random graph in which every potential edge is inserted with probability at least  $p$ . Assume that non-adjacent edges are independent. For which values of  $p$  can we guarantee that the graph is connected with positive probability?

---

<sup>1</sup> The precise condition is  $p > 1 - \frac{(d-1)^{(d-1)}}{d^d}$  for  $d > 1$  and  $p > 1/2$  for  $d = 1$ . The first expression can be simplified by the estimate  $\frac{(d-1)^{(d-1)}}{d^{d-1}} \geq e^{-1}$ , which becomes tight in the limit  $d \rightarrow \infty$ .

It turns out that the answer depends heavily on what exactly we mean by independence. Surprisingly, there are at least five different ways to interpret the innocent-looking condition that non-adjacent edges are independent, which we define next in order of increasing strength. For an edge  $e = \{u, v\}$ , let  $X_e$  be the event that  $e$  is present in the random graph.

► **Definition 1** (Pairwise independence). *For every pair of non-adjacent edges  $e, f$ , the random variables  $X_e$  and  $X_f$  are independent.*

► **Definition 2** (Matching independence). *For any set  $M$  of pairwise non-adjacent edges (also called a matching),  $\{X_e \mid e \in M\}$  are mutually independent.*

► **Definition 3** (Edge-subgraph independence). *For any edge  $e = \{u, v\}$ , any vertex set  $W$  with  $u, v \notin W$ , and any choice of graph  $H_W$  on  $W$ , we have for the random graph  $G$  sampled from the distribution that*

$$Pr[G[W] = H_W \text{ and } X_e] = Pr[G[W] = H_W] \cdot Pr[X_e].$$

► **Definition 4** (Subgraph independence). *For any two disjoint subsets  $V, W$  of vertices, and any choice of graphs  $H_V$  and  $H_W$  on  $V$  and  $W$ , respectively, we have for the random graph  $G$  sampled from the distribution that*

$$Pr[G[V] = H_V \text{ and } G[W] = H_W] = Pr[G[V] = H_V] \cdot Pr[G[W] = H_W].$$

► **Definition 5** (Coloring model). *The graph distribution is given by a set of probability spaces  $(\Omega_v, Pr_v)$ , one for each vertex  $v \in V$ , and a set of deterministic functions  $f_{u,v} : \Omega_u \times \Omega_v \rightarrow \{0, 1\}$  computing  $X_{\{u,v\}}$ , one for each edge  $\{u, v\} \in \binom{V}{2}$ . If every probability space  $(\Omega_v, Pr_v)$  is finite, we call  $\max_v |\Omega_v|$  the number of colors of the coloring model.*

We remark that the standard proof of LLL requires edge-subgraph independence. In the paper [23] that makes the coupling to product spaces explicit, the authors describe subgraph independence as the required property, but inspecting the proof shows that they only use the weaker condition of edge-subgraph independence.

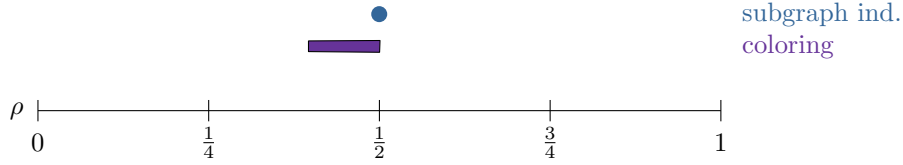
Note that some of our models have been previously studied under different names. We summarize the bounds on the connectivity thresholds obtained before or simultaneously to this paper in Figure 1, and discuss this related work in more detail in Section 1.1. Our main results for large  $n$  are summarized in Table 1 and Figure 2. We give a refined exposition of our results for *all*  $n$  in Theorem 7. We do not have matching upper and lower bounds in all cases, but as can be seen in Figure 2, our bounds imply that there are at least three different thresholds among the five independence conditions discussed above, and four

■ **Table 1** Lower and upper bounds for the threshold  $\rho$  s.t. every constant edge probability  $> \rho$  guarantees connectivity with positive probability for all sufficiently large  $n$ , while a constant edge probability  $< \rho$  does not guarantee connectivity with positive probability for infinitely many  $n$ .

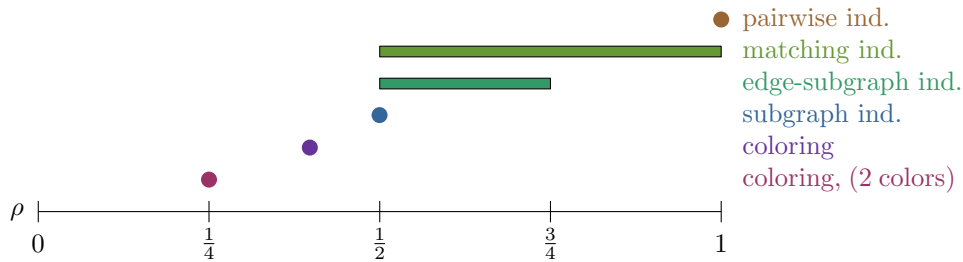
Independence condition	Lower bound	Upper bound
Coloring model (2 colors)	1/4	1/4
Coloring model (general)	0.381966...	0.381966...
Subgraph independence	1/2 [14, Thm. 16]	1/2 [14, Thm. 16]
Edge-subgraph independence	1/2	3/4
Matching independence	1/2	1
Pairwise independence	1	1

## 30:4 On Connectivity in Random Graph Models with Limited Dependencies

different thresholds if we also include the coloring models with only 2 colors. Notably, our proofs give different intervals for the thresholds of all six independence conditions, although this does not imply that the thresholds are all different.



■ **Figure 1** An illustration of the upper and lower bounds on the thresholds known from related work. The bounds for subgraph independence can be found in [14, Thm. 16]. The bounds on the coloring model have been obtained simultaneously and independently to ours [4, Thm. 1.7].



■ **Figure 2** An illustration of the upper and lower bounds mentioned in Table 1 and Theorem 7. As can be seen, there must be at least four thresholds  $\rho$  among the six independence conditions.

Moreover, let us write  $\mathcal{G}_{pw}(n, p)$  for the class of all random graph distributions on  $n$  vertices, with marginal edge probabilities at least  $p$  and with the property “pairwise independence”. We write  $\mathcal{G}_{pw} := \bigcup_{n \in \mathbb{N}} \mathcal{G}_{pw}(n, 0)$  for the same class without restrictions on the number of vertices and the marginal probabilities. Likewise, we write  $\mathcal{G}_{mat}$ ,  $\mathcal{G}_{esub}$ ,  $\mathcal{G}_{sub}$  and  $\mathcal{G}_{col}$  for the random graph distributions satisfying matching independence, edge-subgraph independence, subgraph independence and the coloring models respectively. We show that those five models form a strict hierarchy in Appendix A.

► **Theorem 6.**  $\mathcal{G}_{col} \subsetneq \mathcal{G}_{sub} \subsetneq \mathcal{G}_{esub} \subsetneq \mathcal{G}_{mat} \subsetneq \mathcal{G}_{pw}$ .

### 1.1 Related Work and Examples

There are numerous applications in graph theory, theoretical computer science, and combinatorics in which dependency graphs are bounded or otherwise restricted. For example, if we want to find a vertex coloring in a hypergraph without monochromatic hyperedges, and we color the vertices independently, then any two disjoint hyperedges are independent. Therefore, often times LLL type arguments apply to coloring (hyper)graphs of bounded maximum degree. For instance, the seminal results by Johansson [20] and Molloy [24] stating that every graph of maximum degree  $\Delta$  has (list) chromatic number  $O(\frac{\Delta}{\log \Delta})$  were proved using this method. Similarly, if we want to find a satisfying assignment to a SAT-formula in conjunctive normal form, and assign the variables independently, then any two clauses are independently satisfied if they do not share a common literal, see for instance [18] for more detail and background. These are classical applications of LLL and fall into the class of coloring models.

Another important source of applications comes from percolation theory, which was also the main motivation for [23]. There, an important technique to study locally generated geometric graphs is *rescaling*: the geometric space is covered by boxes that may partially overlap, and boxes are called *good* under some conditions that depend only on the subgraph induced by the box. It is shown recursively that the probability for a box to be good increases with the box size. The goodness of two boxes is independent if they don't overlap, and typically every box only overlaps with constantly many other boxes.

So far, we gave general examples of bounded dependence. We now turn to examples where the random variables are naturally tied to the edge set of a complete graph. An important class of examples come from random graph models that are more complex than Erdős-Rényi graphs, in particular models that capture graph properties of real-world networks like clustering, communities, or heavy-tailed degree distributions. Many such models are generated by drawing some information for each vertex, and connecting two vertices based on this information. There is a large number of such random graph models: in Chung-Lu or Norros-Reittu random graphs, the vertices draw weights which determine their expected degrees [12, 27]; in Random Geometric Graphs (RGG) or Hyperbolic Random Graphs (HRG) the vertices draw some location in a geometric space [22, 26]; in scale-free percolation (SFP) the vertices lie on a grid and also draw random weights [15]; in Geometric Inhomogeneous Random Graphs (GIRG), they draw both a position and a weight [9]. In the Stochastic Block Model (SBM) they draw the community to which they should belong [21]. Many more applications of similar flavour that cannot all be listed here can be found in the literature. All these models fall in the class of coloring models.<sup>2</sup>

The class of coloring models is the most important one.<sup>3</sup> Models that do not fall into this class can arise when the graphs drawn from the distribution must fulfill some global property. For example, consider the following distribution for an *odd* number of vertices  $n \geq 5$ . Every vertex chooses a color from red and blue independently and uniformly at random. The resulting graph consists only of a clique on the color chosen an even number of times. We show later (in the proof of Lemma 38) that this distribution is matching independent, but not edge-subgraph independent (and thus also not a coloring model). As a second example, consider the Erdős-Rényi model  $G_{n,1/2}$  with the additional side constraint that the total number of edges  $|E|$  is divisible by three.<sup>4</sup> One can show quite easily that this distribution is not even pairwise independent. On the other hand, if we instead use the side constraint that  $|E|$  is divisible by *two*, the distribution is actually a coloring model, indicating that independence conditions are surprisingly fickle. A proof of both of these facts can be found in the full version of this paper.

Two of our independence conditions have been previously considered under different names. Firstly, subgraph independent distributions have been studied under the name of *1-independent* random graphs. More generally, a *k-independent* graph distribution is a distribution where any two sets of edges  $E, F$  are independent if the minimum distance

<sup>2</sup> The formulation of many models involves a coin flip for each edge, where the probability of inclusion is a function of the information of the two endpoints. In this formulation, the event  $X_e$  is not a *deterministic* function of the information. But one can simply define the coin flip as part of the random experiments of one of the endpoints.

<sup>3</sup> But note that it is not the model that arises naturally for the LLL. This raises the question whether stronger LLL-type results can be obtained for the coloring model.

<sup>4</sup> More formally, we consider the conditional probability distribution obtained by conditioning the  $G_{n,1/2}$  distribution on the event that  $|E| \equiv_3 0$ .

between any vertex incident to an edge in  $E$  and any vertex incident to an edge in  $F$  is at least  $k$ . These  $k$ -independent graph distributions have been studied extensively in the context of percolation theory, for example on the infinite integer grid [5]. In [14], Day, Falgas-Ravry, and Hancock consider 1-independent random *finite* graphs. Among other results, they pin down the precise value of the connectivity threshold for 1-independent (i.e., subgraph independent) random graphs, as mentioned above in Table 1 and Figure 1.

Coloring models have been studied under two different names. First off, they have been studied as *vertex-based measures* in the context of percolation in [14]. Second off, a deterministic perspective on coloring models lies at the core of the well-studied *density Turán problem for multipartite graphs* [6, 8, 13, 17, 25]: instead of a random graph distribution, the density Turán problem for multipartite graphs considers  $|V(G)|$ -partite blow-ups  $G^*$  of a graph  $G$ , and the minimum density needed between parts of  $G^*$  to guarantee the existence of a graph  $H$  as a *transversal*, i.e., as a subgraph of  $G^*$  that picks precisely one vertex per part of  $G^*$ . This problem has also been considered for the more general case where we are not considering the occurrence of a single graph  $H$  as a transversal, but of any graph from some collection  $\mathcal{H}$ . Furthermore, one can consider the *weighted* case, where the vertices in each part of  $G^*$  each get a positive weight such that all the weights per part add up to 1. The density between two parts is then also computed in a weighted fashion. Now, considering  $G$  as the complete graph  $K_n$ , and  $\mathcal{H}$  as the family of spanning trees on  $n$  vertices, one can see that the weighted density Turán problem for multipartite graphs is equivalent to the problem of determining the minimum edge probability needed to guarantee connectivity with non-zero probability in coloring models (except that the weighted density Turán setting only allows for modeling finitely many colors, which turns out to be without loss of generality, as we discuss in the following subsection). In [4], which appeared simultaneously and independently from this paper, Badakhshian, Falgas-Ravry, and Sharifzadeh consider exactly this question; they prove the same lower bound as us [4, Thm. 1.7], and propose our upper bound as a conjecture [4, Conj. 1.9].

Last but not least, let us mention that in a conceptually similar (but concretely rather different) direction, Alon and Nussboim [2] studied thresholds for the connectivity of random graph models in which only edge-sets of size at most  $k$  for a fixed parameter  $k$  are required to be independent, but dependencies between edge-sets may occur from size  $k + 1$  and up.

## 1.2 Detailed Results

In this section we give the main theorem with more detailed results. We denote by

$$\rho_{\text{pw}}(n) := \inf\{p \in [0, 1] \mid \forall \mathcal{D} \in \mathcal{G}_{\text{pw}}(n, p), \Pr[G_n \sim \mathcal{D} \text{ is connected}] > 0\}$$

the smallest (infimum) marginal edge probability that guarantees a positive probability for  $G_n$  being connected. The quantities  $\rho_{\text{mat}}(n)$ ,  $\rho_{\text{esub}}(n)$ ,  $\rho_{\text{sub}}(n)$ , and  $\rho_{\text{col}}(n)$  are defined similarly. Furthermore, let  $\mathcal{G}_{\text{col}, k}$  denote the coloring models with at most  $k$  colors, and let  $\rho_{\text{col}, k}(n)$  be the corresponding threshold.

► **Theorem 7.** *We have the following bounds on  $\rho$ :*

**Pairwise independence:** For all  $n \in \mathbb{N}$ ,

$$1 - 2/n - \Theta(1/n^2) \leq \rho_{\text{pw}}(n) \leq 1 - 2/n.$$

**Matching independence:** For all  $n \in \mathbb{N}$ ,

$$\frac{1}{2} \left(1 - \tan^2 \frac{\pi}{2n}\right) \leq \rho_{\text{mat}}(n) \leq 1 - 2/n.$$

Moreover, for any  $k \in \mathbb{N}$ ,

$$\frac{1}{2} \leq \rho_{mat}(8k).$$

**Edge-subgraph independence:** For all  $n \in \mathbb{N}$ ,

$$\frac{1}{2} \left(1 - \tan^2 \frac{\pi}{2n}\right) \leq \rho_{esub}(n) \leq 3/4.$$

**Subgraph independence:** [14, Thm 16] For all  $n \geq 2$ ,

$$\rho_{sub}(n) = \frac{1}{2} \left(1 - \tan^2 \frac{\pi}{2n}\right).$$

**Coloring model:** For all  $n \geq 3$ , we have  $1/4 \leq \rho_{col}(n) \leq 1/2$ . Moreover, let  $\phi = \frac{1}{2}(1 + \sqrt{5})$  be the golden ratio. Then

$$\lim_{n \rightarrow \infty} \rho_{col}(n) = 2 - \phi \approx 0.381966.$$

**Coloring model, two colors:** For all  $n \geq 3$ ,  $\rho_{col,2}(n) = 1/4$ .

To the best of our knowledge, this question had not been considered so far for the pairwise independence, matching independence, or edge-subgraph independence models. For the subgraph independence model, the threshold was known precisely [14], but we include it in our theorem for completeness of the hierarchy of our models. Our results for the coloring model confirm a conjecture of Badakhshian, Falgas-Ravry, and Sharifzadeh [4].

We remark that, while  $\mathcal{G}_{col}(n, p)$  includes models with infinitely many outcomes of the random experiment at each vertex, it has been shown<sup>5</sup> (in [8] for  $n = 3$  and in [25, Lemma 2.1] for all  $n$ ) that  $\rho_{col,k}(n) \geq \rho_{col}(n)$  for  $k = n - 1$ , which combined with the trivial  $\rho_{col,k}(n) \leq \rho_{col}(n)$  for any  $k$  gives  $\rho_{col,n-1}(n) = \rho_{col}(n)$ . The proofs of  $\rho_{col,n-1}(n) \geq \rho_{col}(n)$  take a model which has probability 0 of being connected and modify it to use at most  $n - 1$  many colors while preserving the fact that it is never connected and increasing or keeping the same its marginal edge probabilities. However, the probability distribution over all graphs on  $n$  vertices that the modified model gives is usually different from the original one. Here, we prove the even stronger statement that  $\mathcal{G}_{col}(n, p) = \mathcal{G}_{col,k}(n, p)$  for some  $k = k(n)$ , showing that finitely many colors suffice to model any probability distribution over all graphs that can be represented as a coloring model. The proof of Lemma 8 can be found in Appendix C.

► **Lemma 8.** Consider a coloring model on  $n$  vertices along with its corresponding probability distribution  $\mathcal{D}$  over the graphs on  $n$  vertices. Then there is a coloring model on  $n$  vertices with at most  $2^{\binom{n}{2}} + 1$  colors per vertex which results in the same distribution  $\mathcal{D}$ .

### 1.3 Open Questions

There are many interesting questions that remain open. Most obviously, for large  $n$  there are gaps between the upper and lower bounds for matching independence and edge-subgraph independence. For the coloring model we do have matching upper and lower bounds for

<sup>5</sup> In fact, these results are stated only for finitely many colors as they are in the density Turán setting, but the proofs go through in our more general set-up as well.

an unbounded number of colors and for  $k = 2$ , but not for other constant values of  $k$ . Moreover, for fixed  $n$  it is unclear how much richer the coloring model gets by adding more colors. As mentioned above, we know that  $\mathcal{G}_{\text{col}}(n, p) = \mathcal{G}_{\text{col},k}(n, p)$  with  $k = 2^{\binom{n}{2}} + 1$  and  $\rho_{\text{col}}(n) = \rho_{\text{col},n-1}(n)$ . But what is still an open question is the behavior of the functions  $k_{\text{min}}^{\mathcal{G}}(n) = \min\{k \mid \mathcal{G}_{\text{col}}(n, p) = \mathcal{G}_{\text{col},k}(n, p)\}$  and  $k_{\text{min}}^{\rho}(n) = \min\{k \mid \rho_{\text{col}}(n, p) = \rho_{\text{col},k}(n, p)\}$ , that is, how many colors are enough to be able to express all coloring models on  $n$  vertices, respectively to capture the behaviour of the connectivity threshold of the coloring models on  $n$  vertices?

Finally, in this paper we focus on connectivity because that is arguably the most fundamental global property of a graph. We only see this as a starting point and we would find it interesting to explore analogous questions for other global properties. A natural extension might be to study the size of the largest connected component that can be achieved with non-zero probability, but the same questions arise for any other global graph properties such as Hamiltonicity, the chromatic number, and many more.

## 1.4 Proof Techniques

We prove the strict hierarchical structure of our considered independence conditions (Theorem 6) by giving concrete examples of distributions that fulfill the weaker independence condition, but not the stronger one. All of these examples are simple to describe and quite illustrative.

The lower bounds on the thresholds  $\rho$  are shown by concrete series of graph distributions on disconnected graphs. Here, the lower bound on  $\rho_{\text{pw}}$  (pairwise ind.) uses the same distribution as the distribution used to show  $\mathcal{G}_{\text{pw}} \neq \mathcal{G}_{\text{mat}}$ . The lower bounds on  $\rho_{\text{mat}}$ ,  $\rho_{\text{esub}}$ , and  $\rho_{\text{sub}}$  all use the same construction, since we were not able to make use of the additional freedom available in the edge-subgraph or even the matching independence condition (except for the case when  $n$  is divisible by 8 in the matching independence model). This lower bound on  $\rho_{\text{sub}}$  has been previously proven in [14, Thm. 16]. While we do not restate the proof here, note that this bound uses a distribution that can be seen conceptually as a version of a coloring model with *complex probabilities*. The proof that the distribution fulfills subgraph independence follows from the fact that the coloring models are subgraph independent, and from the fundamental theorem of algebra. The lower bounds on  $\rho_{\text{col}}$  and  $\rho_{\text{col},2}$  once again are simple constructions.

For the upper bounds on the thresholds  $\rho$  we use very different proof techniques depending on the independence condition. The upper bound for pairwise and matching independence does not make use of these independence conditions, and just combines a linearity of expectation argument with the maximum number of edges in a disconnected graph. As mentioned above, edge-subgraph independence is the first independence condition which allows us to apply Lovász Local Lemma to achieve a constant bound on  $\rho$ . For the coloring models, we give concrete strategies on how to pick a color for each vertex such that the graph is connected. For the two color case, this strategy is rather simple, while the strategy for the general case is based on adjusting a random coloring, which fulfills useful properties with high probability for large  $n$ . The proof of this bound (Theorem 21) is by far the most technically involved proof in this paper.



## 2 Preliminaries

We make use of the following version of the Lovász Local Lemma.

► **Lemma 9** ([28]). *Let  $A_1, \dots, A_k$  be a sequence of events, such that each event occurs with probability at least  $p$  and is independent of all the other events, except at most  $d$  of them. Then, if*

$$p > \begin{cases} 1 - \frac{(d-1)^{d-1}}{d^d} & \text{for } d \geq 2, \\ 1/2 & \text{for } d = 1, \end{cases}$$

*there is a non-zero probability that all of the events occur.*

The following lemma is a direct consequence of Markov's inequality.

► **Lemma 10.** *Let  $X$  be a random variable such that  $\Pr[X \leq u] = 1$ . Then for  $\ell < u$ ,*

$$\Pr[X \geq \ell] \geq \frac{\mathbb{E}[X] - \ell}{u - \ell}.$$

**Proof.** Note that the random variable  $u - X$  is non-negative and has expectation  $u - \mathbb{E}[X]$ . We can thus apply Markov's inequality.

$$\Pr[X \geq \ell] = \Pr[u - X \leq u - \ell] = 1 - \Pr[u - X > u - \ell] \geq 1 - \frac{u - \mathbb{E}[X]}{u - \ell} = \frac{\mathbb{E}[X] - \ell}{u - \ell}. \blacktriangleleft$$

## 3 Bounds on $\rho$

In this section we prove Theorem 7, showing lower and upper bounds on  $\rho$  for our various models of independence. By the definition of  $\rho$  as in Section 1.2, any lower bound on  $\rho_X$  for some independence condition  $X$  also holds for  $\rho_Y$  for some weaker independence condition  $Y$ , i.e., one such that  $\mathcal{G}_X \subseteq \mathcal{G}_Y$ . Conversely, any upper bound on  $\rho_Y$  also holds for  $\rho_X$ .

### 3.1 Pairwise Independence

► **Lemma 11.** *For even  $n \geq 4$ , we have  $\rho_{pw}(n) \geq 1 - \frac{2}{n} - \Theta\left(\frac{1}{n^2}\right)$ .*

**Proof.** Consider the distribution  $CM(n, q)$  as defined in Definition 40 with  $q = 1 - \Theta\left(\frac{1}{n^2}\right)$  chosen such that  $CM(n, q)$  is pairwise independent, which can be done by Claim 41. The marginal edge probability  $p$  in  $CM(n, q)$  is then  $p = 1 - \frac{2}{n} - \Theta\left(\frac{1}{n^2}\right)$ . Note that the probability of  $G \sim CM(n, q)$  being connected is 0, since in the clique regime there is always an isolated vertex, and in the matching regime  $G$  only consists of a perfect matching. Thus, the threshold  $\rho_{pw}$  for non-zero probability of connectivity in the pairwise independence model is at least  $p = 1 - \frac{2}{n} - \Theta\left(\frac{1}{n^2}\right)$ .  $\blacktriangleleft$

► **Lemma 12.** *For any  $n$ , we have  $\rho_{pw}(n) \leq 1 - \frac{2}{n}$ .*

**Proof.** If the minimum marginal edge probability  $p$  in any graph distribution (not even necessarily pairwise independent) is larger than  $1 - \frac{2}{n}$ , the expected number of edges  $\mathbb{E}[\sum_e X_e]$  is larger than  $(1 - \frac{2}{n})\binom{n}{2} = \binom{n-1}{2}$ . Since any disconnected graph contains at most  $\binom{n-1}{2}$  edges, the graph must be connected with non-zero probability.  $\blacktriangleleft$

### 3.2 Matching Independence

We have no general bounds specifically holding for matching independence. Matching independence is a condition that is surprisingly difficult to exploit. We thus state the following corollary, following directly from Lemma 12 and Lemma 17.

► **Corollary 13.** *For any  $n \geq 2$ , we have  $\frac{1}{2}(1 - \tan^2 \frac{\pi}{2n}) \leq \rho_{mat}(n) \leq 1 - \frac{2}{n}$ .*

The lower bound in Corollary 13 can be improved slightly when  $n$  is divisible by 8, which in particular shows that  $\rho_{mat}(n) \neq \rho_{sub}(n)$  for infinitely many  $n$ , illustrating again the different behavior of the connectivity thresholds for different independence conditions. The proof of this statement can be found in the full version.

► **Lemma 14.** *For any  $n \in \mathbb{N}$  that is divisible by 8, we have  $\rho_{mat}(n) \geq \frac{1}{2}$ .*

### 3.3 Edge-Subgraph Independence

We have no lower bound specifically for edge-subgraph independence. We inherit this bound from the stricter subgraph independence, i.e., Lemma 17.

► **Corollary 15.** *For any  $n \geq 2$ , we have  $\rho_{esub}(n) \geq \frac{1}{2}(1 - \tan^2 \frac{\pi}{2n})$ .*

► **Lemma 16.** *For any  $n$ , we have  $\rho_{esub}(n) \leq \frac{3}{4}$ .*

**Proof.** We show that any edge-subgraph independent distribution with minimum marginal edge probability  $p > \frac{3}{4}$  is connected with non-zero probability. To achieve this, we use the Lovász Local Lemma, as stated in Lemma 9.

We pick the edges of any Hamiltonian path  $e_1, \dots, e_{n-1}$  of  $K_n$  and consider their corresponding events  $X_{e_1}, \dots, X_{e_{n-1}}$ . Since we have edge-subgraph independence, each of these events depends on only at most  $d = 2$  of the other events (the ones corresponding to the neighboring edges). Furthermore, the probability of each of these events is at least  $p > \frac{3}{4}$ . By Lemma 9, since  $p > \frac{3}{4} = 1 - \frac{1}{2^2}$ , with non-zero probability all of the events  $X_{e_1}, \dots, X_{e_{n-1}}$  happen, all edges of the Hamiltonian path are present, and thus the graph is connected. ◀

### 3.4 Subgraph Independence

The exact behavior of  $\rho_{sub}$  has been determined in [14, Thm. 16]. For a proof, we point the reader to [14]. We also include a proof of the lower bound part of this statement in the full version of this paper.

► **Lemma 17** ([14, Thm. 16]). *For any  $n \geq 2$ , we have  $\rho_{sub}(n) = \frac{1}{2}(1 - \tan^2 \frac{\pi}{2n})$ .*

### 3.5 Coloring Models

We first consider coloring models with only 2 colors, since we can find matching lower and upper bounds for that case. Other restrictions of the coloring model could also be interesting to investigate, such as other bounded numbers of colors, or the case where every vertex must pick *uniformly* among its colors. We show that the threshold probability is exactly 1/4 in this case. The proof can be found in Appendix B.

► **Lemma 18.** *For all  $n \geq 3$ , we have  $\rho_{col,2}(n) = 1/4$ .*

We now consider the full generality of coloring models, with an unbounded or even infinite (although by Lemma 8 this can w.l.o.g. be excluded) number of outcomes of the local experiment at each vertex.

► **Lemma 19.** *For all  $n \geq 3$ , we have  $\rho_{\text{col}}(n) \geq (n - 2) \frac{3n-4-\sqrt{5n^2-16n+12}}{2(n-1)^2}$ .*

Note that for  $n = 3$ , this bound is equal to  $1/4$  (which matches the previous subsection, since our construction only uses 2 colors in this case), and for  $n \rightarrow \infty$ , it tends to  $2 - \phi \approx 0.381966$ . We remark that Lemma 19 was obtained simultaneously and independently in [4, Thm 1.7] with the exact same construction.

**Proof.** We show that for each number of vertices  $n \geq 3$ , there exists a coloring model with marginal edge probability  $p(n) = (n - 2) \frac{3n-4-\sqrt{5n^2-16n+12}}{2(n-1)^2}$  which is connected with probability 0.

We define the distribution  $CS(n)$  for any  $n \geq 3$ . A fixed vertex  $v$  picks as its local experiment one other vertex  $v' \in [n] \setminus \{v\}$  uniformly at random. All other vertices pick the color red with probability  $q$  (to be determined later), and blue with probability  $1 - q$ . Then,  $v$  connects to all vertices in  $V \setminus \{v, v'\}$  that are colored blue. Between the vertices in  $V \setminus \{v\}$ , an edge is present if both endpoints are colored red. The resulting graph is clearly not connected: every red vertex is only connected to other red vertices, and in the case that all vertices in  $V \setminus \{v\}$  are blue, only  $n - 2$  edges exist.

For every edge  $e$  not incident to  $v$  we have  $Pr[X_e] = q^2$ , and for every edge  $e'$  incident to  $v$  we have  $Pr[X_{e'}] = (1 - q) \frac{n-2}{n-1}$ . We pick  $q$  maximizing the minimum of these probabilities. Since  $q^2$  is increasing in  $q$  for  $q \geq 0$ , and  $(1 - q) \frac{n-2}{n-1}$  is decreasing in  $q$ , the minimum is maximized when  $q^2 = (1 - q) \frac{n-2}{n-1}$ . We can thus solve

$$q^2 + \frac{n-2}{n-1}q - \frac{n-2}{n-1} = 0$$

to get

$$q = \frac{2 - n \pm \sqrt{5n^2 - 16n + 12}}{2(n - 1)}.$$

Only one of these solutions fulfills  $q > 0$ , namely the one with “+”. Since the marginal edge probability  $p(n)$  is equal to  $(1 - q) \frac{n-2}{n-1}$ , we get the claimed bound. ◀

Let us now consider upper bounds for  $\rho_{\text{col}}(n)$ . We first state the weaker constant bound holding for all  $n$ . The proof can be found in the full version of this paper. Note that the proof of Theorem 1.7 in [4], which came out independently and simultaneously, uses essentially the same techniques as the ones we employ in the proof of the lemma below, except that they optimize the calculations to get a bound of  $\frac{1}{2} - \frac{1}{4n-6}$ .

► **Lemma 20.** *For every  $n$ , we have  $\rho_{\text{col}}(n) \leq \frac{1}{2}$ .*

Next, we present the stronger upper bound holding for large enough  $n$ . This bound tends to  $2 - \phi$  for  $n \rightarrow \infty$  and thus matches the lower bound for large enough  $n$ , as conjectured in [4].

► **Theorem 21.** *For any  $\varepsilon > 0$ , there exists an  $n_0(\varepsilon)$ , such that for any coloring model distribution on graphs on  $n \geq n_0(\varepsilon)$  vertices with minimum marginal edge probability at least  $p' = 2 - \phi + \varepsilon$ , the graph is connected with non-zero probability.*

To prove this theorem, we will need the following notation and setup. We write  $p$  for  $2 - \phi$ .

► **Observation 22.** *For  $p = 2 - \phi$ , we have the nice identities  $\frac{1}{p} - 2 = 1 - p$  and  $(1 - p)^2 = p$ .*

## 30:12 On Connectivity in Random Graph Models with Limited Dependencies

To begin, we color each vertex independently with a random color according to its own distribution.<sup>6</sup> For this fixed coloring, we write  $N(v, c)$  for the set of neighbors of  $v$ , if  $v$  is recolored to the color  $c$ .

► **Lemma 23.** *In a random coloring, with high probability it holds that  $\mathbb{E}_c[|N(v, c)|] > (p + \varepsilon/2)n$  for all  $v$  simultaneously.*

The proof of this lemma uses standard concentration bounds and can be found in the full version.

Thus, from now on, we condition on the high probability event of Lemma 23. This allows us to argue that we can recolor any vertex such that it has a large neighborhood. Furthermore, using Lemma 10, we can also argue that for every vertex, relatively large neighborhoods must exist with a somewhat large probability. This allows us to use union bounds to show that large neighborhoods must exist at the same time as some fixed edges.

Our general strategy for showing the graph is connected with positive probability will be to mostly rely on the random coloring of all vertices that we start with, but recolor some vertices as necessary, using some large neighborhoods to connect vertices to.

Since we will recolor some vertices, the true neighborhoods in the final graph may be different (possibly also smaller) than in the random coloring for which we get these bounds. We use the error term  $C := C'(\varepsilon) \log_2 n$  for an upper bound on the number of vertices that we recolor and aim to connect to (we may recolor more vertices, but we do not argue that we connect to these vertices, thus recoloring them does not matter). If we can guarantee a desired intersection of neighborhoods (or of a neighborhood with a fixed set) to contain at least  $C + 1$  vertices in the random coloring, we know that in the actual coloring at the end the intersection is non-empty.

Let  $(r, c_r)$  be a vertex-color pair which maximizes  $|N(r, c_r)|$ . We recolor  $r$  to  $c_r$ . If every vertex outside of  $N(r, c_r)$  can be recolored to some color such that it has an edge to a vertex in  $N(r, c_r)$ , the graph can be connected. Otherwise, we pick a vertex-color pair  $(b, c_b)$  which maximizes  $|N(b, c_b)|$  among all vertex-color pairs for which  $b \notin N(r, c_r) \cup \{r\}$ , and  $|N(r, c_r) \cap N(b, c_b)| \leq C$ . We recolor  $b$  to  $c_b$ .

In the following, we use the shorthands

$$R := N(r, c_r), \quad B := N(b, c_b), \quad \alpha := |R|/n, \quad \text{and} \quad \beta := |B|/n.$$

The vertices  $r$  and  $b$ , along with their respective neighborhoods  $R$  and  $B$ , will be central to our argument, and we will find various ways to connect all other vertices to those sets.

The next lemma is fairly straightforward and its proof can be found in the full version.

► **Lemma 24.**  *$p < \alpha < 1 - p$ , and  $p < \beta \leq \alpha$ .*

Note that if we have  $R \cap B \neq \emptyset$ , we can easily connect the graph: if we recolor each vertex  $v \notin R \cup B \cup \{r, b\}$  to some color  $c$  such that  $|N(v, c)| > pn$ ,  $N(v, c)$  must intersect  $R \cup B$ . This follows from  $|R \cup B| = |R| + |B| - |R \cap B| \geq \alpha n + \beta n - C \geq 2pn - C$ , and  $(2pn - C) + pn > n$ . We thus assume in the following that  $R \cap B = \emptyset$ .

► **Lemma 25.**  *$\beta \leq \min(1 - \alpha, 1/2)$ .*

**Proof.** This follows directly from  $R \cap B = \emptyset$  and thus  $\alpha + \beta \leq 1$ . ◀

---

<sup>6</sup> We assume that we have a finite number of colors, which we can do by Lemma 8.

To connect the graph in the remaining cases, we need to introduce some more definitions. Intuitively, we say that any vertex  $v$  is *obligate red*, if it cannot robustly be connected to the blue set  $B$ . More formally, we say that a vertex  $v \notin B \cup \{b\}$  is obligate red, if there exists no color  $c$ , such that  $|N(v, c) \cap B| > C$ . Similarly, we say that a vertex  $v \notin R \cup \{r\}$  is *obligate blue*, if there exists no color  $c$  such that  $|N(v, c) \cap R| > C$ . Let  $OR$  be the set of obligate red vertices, and  $OB$  the set of obligate blue vertices.

Note that no vertex can be both obligate blue and obligate red. When a vertex is neither obligate blue nor obligate red, we say it is *non-obligate*.

We start with giving some guarantees on the sizes of neighborhoods of obligate blue and obligate red vertices.

► **Lemma 26.** *Let  $u \in OB, v \in OR$ . Then, over the distribution of the color of  $u$  ( $v$ ), the following hold for the neighborhood of  $u$  ( $v$ ).*

1.  $Pr[|N(u, c)| \geq \frac{p+p\beta-\beta}{p}n + \varepsilon n/2] \geq 1 - p$ .
2.  $Pr[|N(v, c)| \geq \frac{p+p\alpha-\alpha}{p}n + \varepsilon n/2] \geq 1 - p$ .
3.  $Pr[|N(u, c)| \geq (1-p)(1-\beta)n + \varepsilon n/2] \geq 1 - \sqrt{p}$ .
4.  $Pr[|N(v, c)| \geq (1-p)(1-\alpha)n + \varepsilon n/2] \geq 1 - \sqrt{p}$ .

**Proof (sketch).** We can upper bound the size of a neighborhood of any vertex by  $\alpha n$ . Furthermore, we can upper bound the size of a neighborhood of any obligate blue vertex  $u$  by  $\beta n$ . The lemma follows from applying Lemma 10 and simplifying. The full calculations are found in the full version. ◀

The next two corollaries follow from Lemma 26 and some calculations that can be found in the full version.

► **Corollary 27.** *Let  $S$  be a vertex set of size  $|S| \leq C$ . Let  $u \in OB$ . Then,*

$$Pr_c[N(u, c) \cap (B \setminus S) \neq \emptyset] \geq 1 - p.$$

► **Corollary 28.** *Let  $S$  be a vertex set of size  $|S| \leq C$ . Let  $v \in OR$ . Then,*

$$Pr_c[N(v, c) \cap (R \setminus S) \neq \emptyset] \geq 1 - p.$$

We will now prove Theorem 21 by case distinction on the number of obligate red and obligate blue vertices.

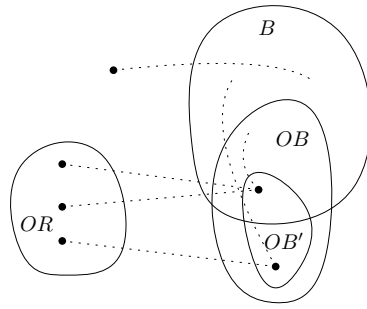
► **Lemma 29.** *If there are either no obligate red or no obligate blue vertices, the graph can be connected.*

**Proof.** If there are no obligate red vertices, we recolor every vertex  $v \notin B \cup \{b\}$  to a color  $c$  such that  $N(v, c)$  intersects  $B$ . Every vertex is thus connected to  $b$  either directly or through a vertex in  $B$ , proving that the graph is connected. Symmetrically, if there are no obligate blue vertices, all vertices can be connected to  $r$  or  $R$ . ◀

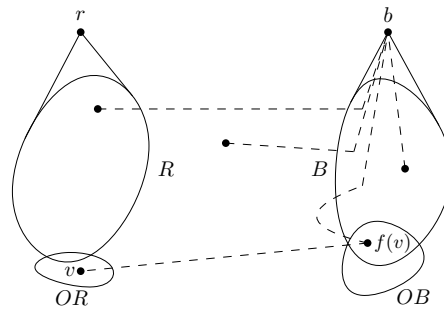
► **Lemma 30.** *If  $|OR| > C$  or  $|OB| > C$ , the graph can be connected.*

The full proof of this lemma can be found in the full version.

**Proof (sketch).** Suppose without loss of generality that  $|OB| > |OR|$  (the other case is symmetric). Let  $OB' \subset OB \setminus \{b\}$  be a subset of  $C$  obligate blue vertices. We recolor all vertices in  $OB'$  as well as all non-obligate vertices such that they connect into  $B \setminus OB'$ . Crucially, each vertex in  $OB'$  has a color that allows it to connect to  $B \setminus OB'$  as well as a



■ **Figure 3** How to connect the graph when  $|OB| > C$ .



■ **Figure 4** How to connect the graph given an injective mapping  $f$  from  $OR$  to  $OB \setminus \{b\}$ .

constant proportion of the remaining vertices in  $OR$ . Since  $C$  is logarithmic in  $n$ , we are able to cover all vertices in  $OR$  in this way, and thus connect the graph. An illustration of this can be seen in Figure 3. ◀

We thus assume in the following that the numbers of obligate red and obligate blue vertices are more than 0 and at most  $C$ .

► **Lemma 31.** *If there is an injective mapping from  $OR$  to  $OB \setminus \{b\}$ , or an injective mapping from  $OB$  to  $OR \setminus \{r\}$ , the graph can be connected.*

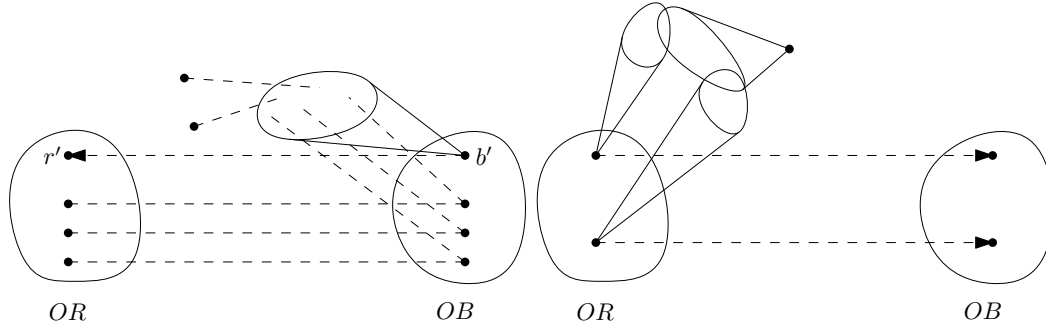
**Proof.** We first assume an injective mapping  $f$  from  $OR$  to  $OB \setminus \{b\}$  exists. The opposite case works symmetrically. In this case we aim to connect every vertex to  $b$ , by connecting each obligate red vertex  $v$  to the vertex  $f(v)$ . These obligate blue vertices, as well as the non-obligate vertices outside  $B \setminus \{b\}$  are then connected to some vertex in  $B \setminus OB$ . Since these vertices are connected to  $b$ , the whole graph is connected. This is shown in Figure 4. Note that we recolor the vertices in  $OB$  and outside  $B$ , but not those in  $B \setminus OB$ .

We let  $v \in OR$  and  $f(v) \in OB$  pick a color independently at random from their respective distributions. We have that  $Pr[vf(v) \in E(G)] \geq p' = p + \varepsilon$ . Furthermore, by Corollary 27,  $Pr[N(f(v), c) \cap (B \setminus OB) \neq \emptyset] \geq (1 - p)$ . Thus, letting  $Y_{v,f(v)}$  be the indicator random variable that is 1 if and only if  $vf(v) \in E(G)$  and  $N(f(v), c) \cap (B \setminus OB) \neq \emptyset$ , we have

$$Pr[Y_{v,f(v)} = 1] \geq p' + (1 - p) - 1 \geq \varepsilon.$$

Thus, we can pick colors for each pair of vertices  $(v, f(v))$ , for  $v \in OR$ , such that  $v$  connects to  $f(v)$  and  $f(v)$  connects to  $B \setminus OB$ . ◀

We can enhance the previous lemma by considering non-obligate vertices which also have a large probability (at least  $1 - p$ ) to robustly connect to either of the two sets  $R$  and  $B$  (as obligate vertices are guaranteed to, by Corollary 27 and Corollary 28, respectively):



(a) Case 1: At least one edge is oriented towards  $OR$ . (b) Case 2: All edges are oriented towards  $OB$ .

■ **Figure 5** Connecting the graph in Lemma 35.

► **Lemma 32.** *If there exists a non-obligate vertex  $v \notin \{r, b\}$  such that for  $S = R$  or  $S = B$ , it holds that  $\Pr_c[|N(v, c) \cap S| > C] \geq 1 - p$ , then the graph can be connected.*

**Proof.** If the conditions of Lemma 30 or Lemma 31 are fulfilled, we can connect the graph. Otherwise,  $|OR| = |OB| \leq C$  and  $r \in OR$  and  $b \in OB$ . Assume there is a vertex  $v$  such that  $\Pr_c[|N(v, c) \cap B| > C] \geq 1 - p$  (the case  $S = R$  works symmetrically). We can make a bijective mapping from  $OR$  to  $(OB \cup \{v\}) \setminus \{b\}$ . Since we have that  $\Pr_c[N(v, c) \cap (B \setminus (OB \cup \{v\} \setminus \{b\})) \neq \emptyset] \geq 1 - p$ , the proof of Lemma 31 also works for this mapping. ◀

Furthermore, if there exists a vertex that can be used to join the red and blue sets, we can also connect the graph:

► **Lemma 33.** *If there exists a vertex  $v \notin \{r, b\}$  and a color  $c$ , such that  $N(v, c) \cap R \neq \emptyset$  and  $N(v, c) \cap B \neq \emptyset$ , we can connect the graph.*

**Proof.** We give the vertex  $v$  the color  $c$ . Every vertex  $w \notin R \cup B \cup \{r, b, v\}$  is recolored to some color  $c_w$  such that  $|N(w, c_w)| > pn$ . Since  $p > 1 - \alpha - \beta$ ,  $w$  is then connected to some vertex in  $R$  or  $B$ . After doing this for all  $w$ , the graph is connected. ◀

Finally, we will consider the cases which were not covered by the previous lemmas. For this, we need the following lower bounds regarding non-obligate vertices, the proofs of which can be found in the full version.

► **Lemma 34.** *For any non-obligate vertex  $v \notin \{r, b\}$ , we define*

$$p_v^r := \max_{c: |N(v, c) \cap B| \leq C} |N(v, c)|, \text{ and } p_v^b := \max_{c: |N(v, c) \cap R| \leq C} |N(v, c)|.$$

*Assuming there exists no vertex fulfilling the conditions of either Lemma 32 or Lemma 33, we have  $p_v^r > \frac{\beta p - \beta + p}{p} n$ , and  $p_v^b > \frac{\alpha p - \alpha + p}{p} n$ .*

Finally, we cover the only remaining case, whose proof can be found in Appendix B. The idea of the proof is sketched in Figure 5. We consider the potential edges  $b'r'$  between  $b' \in OB$  and  $r' \in OR$ . For each such edge we determine the vertex with a higher probability of picking a color such that the other vertex can be connected to it. We then orient the edge from this vertex to the other. If at least one edge is oriented from  $OB$  to  $OR$ , we can use this fact to connect every vertex to  $B \cup OB$ . Otherwise, we connect every vertex to  $R \cup OR$ .

► **Lemma 35.** *If  $|OB| = |OR|$ , and there exists no vertex as in Lemma 32 or Lemma 33, we can connect the graph.*

**Proof of Theorem 21.** Lemmas 29–33 and 35 cover all possible cases. Thus, we can always connect the graph, proving the theorem. ◀

---

## References

- 1 Jon Aaronson, David Gilat, Michael Keane, and Vincent de Valk. An algebraic construction of a class of one-dependent processes. *The annals of probability*, pages 128–143, 1989.
- 2 Noga Alon and Asaf Nussboim.  $k$ -wise independent random graphs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 813–822, 2008. doi:10.1109/FOCS.2008.61.
- 3 Noga Alon and Joel H. Spencer. *The probabilistic method*. John Wiley & Sons, 2016. doi:10.1002/9780470277331.
- 4 Leila Badakhshian, Victor Falgas-Ravry, and Maryam Sharifzadeh. On density conditions for transversal trees in multipartite graphs. *arXiv preprint*, 2023. doi:10.48550/arXiv.2305.05713.
- 5 Paul Balister, Béla Bollobás, and Mark Walters. Random transceiver networks. *Advances in Applied Probability*, 41(2):323–343, 2009. doi:10.1239/aap/1246886613.
- 6 Paul Balister, Tom Johnston, Michael Savery, and Alex Scott. Improved bounds for 1-independent percolation on  $\mathbb{Z}^n$ . *arXiv preprint*, 2022. doi:10.48550/arXiv.2206.12335.
- 7 Béla Bollobás. The isoperimetric number of random regular graphs. *European Journal of Combinatorics*, 9(3):241–244, 1988. doi:10.1016/S0195-6698(88)80014-3.
- 8 Adrian Bondy, Jian Shen, Stéphan Thomassé, and Carsten Thomassen. Density conditions for triangles in multipartite graphs. *Combinatorica*, 26(2):121–131, 2006. doi:10.1007/s00493-006-0009-y.
- 9 Karl Bringmann, Ralph Keusch, and Johannes Lengler. Geometric inhomogeneous random graphs. *Theoretical Computer Science*, 760:35–54, 2019. doi:10.1016/j.tcs.2018.08.014.
- 10 Robert M. Burton, Marc Goulet, and Ronald Meester. On 1-dependent processes and  $k$ -block factors. *The Annals of Probability*, 21(4):2157–2168, 1993. doi:10.1214/aop/1176989014.
- 11 Constantin Carathéodory. Über den variabilitätsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen. *Mathematische Annalen*, 64(1):95–115, 1907. doi:10.1007/BF01449883.
- 12 Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of combinatorics*, 6(2):125–145, 2002. doi:10.1007/PL00012580.
- 13 Péter Csikvári and Zoltán L. Nagy. The density Turán problem. *Combinatorics, Probability and Computing*, 21(4):531–553, 2012. doi:10.1017/S0963548312000016.
- 14 A. Nicholas Day, Victor Falgas-Ravry, and Robert Hancock. Long paths and connectivity in 1-independent random graphs. *Random Structures & Algorithms*, 57(4):1007–1049, 2020. doi:10.1002/rsa.20972.
- 15 Maria Deijfen, Remco van der Hofstad, and Gerard Hooghiemstra. Scale-free percolation. *Annales de l’IHP Probabilités et statistiques*, 49(3):817–838, 2013. doi:10.1214/12-AIHP480.
- 16 Paul Erdős. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959. doi:10.4153/CJM-1959-003-9.
- 17 Victor Falgas-Ravry and Vincent Pfenninger. 1-independent percolation on  $\mathbb{Z}^2 \times K_n$ . *Random Structures & Algorithms*, 2022. doi:10.1002/rsa.21129.
- 18 Heidi Gebauer, Robin A. Moser, Dominik Scheder, and Emo Welzl. The Lovász Local Lemma and satisfiability. In Susanne Albers, Helmut Alt, and Stefan Näher, editors, *Efficient Algorithms: Essays Dedicated to Kurt Mehlhorn on the Occasion of His 60th Birthday*, pages 30–54. Springer, Berlin, Heidelberg, 2009. doi:10.1007/978-3-642-03456-5\_3.



- 19 Alexander E. Holroyd and Thomas M. Liggett. Finitely dependent coloring. In *Forum of Mathematics, Pi*, volume 4, page e9. Cambridge University Press, 2016. doi:10.1017/fmp.2016.7.
- 20 Anders Johansson. Asymptotic choice number for triangle free graphs. *Technical Report, DIMACS*, pages 91–95, 1996.
- 21 Brian Karrer and Mark E. J. Newman. Stochastic blockmodels and community structure in networks. *Physical review E*, 83(1):016107, 2011. doi:10.1103/PhysRevE.83.016107.
- 22 Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010. doi:10.1103/PhysRevE.82.036106.
- 23 Thomas M. Liggett, Roberto H. Schonmann, and Alan M. Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997. doi:10.1214/aop/1024404279.
- 24 Michael Molloy. The list chromatic number of graphs with small clique number. *Journal of Combinatorial Theory, Series B*, 134(1):264–284, 2019. doi:10.1016/j.jctb.2018.06.007.
- 25 Zoltán L. Nagy. A multipartite version of the Turán problem – density conditions and eigenvalues. *The Electronic Journal of Combinatorics*, 18(1):P46, 2011. doi:10.37236/533.
- 26 Mathew Penrose. *Random Geometric Graphs*. Oxford University Press, 2003. doi:10.1093/acprof:oso/9780198506263.001.0001.
- 27 Hannu Reittu and Ilkka Norros. Random graph models of communication network topologies. In *Unifying Themes in Complex Systems VII: Proceedings of the Seventh International Conference on Complex Systems*, pages 214–221. Springer, 2012. doi:10.1007/978-3-642-18003-3\_21.
- 28 James B. Shearer. On a problem of Spencer. *Combinatorica*, 5(3):241–245, 1985. doi:10.1007/BF02579368.
- 29 Joel Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Mathematics*, 20:69–76, 1977. doi:10.1016/0012-365X(77)90044-9.

## A A Strict Hierarchy

This section is dedicated to proving Theorem 6:

► **Theorem 6.**  $\mathcal{G}_{col} \subsetneq \mathcal{G}_{sub} \subsetneq \mathcal{G}_{esub} \subsetneq \mathcal{G}_{mat} \subsetneq \mathcal{G}_{pw}$ .

We will prove each strict inclusion as its own lemma, from left to right. The first among these lemmas shows that each coloring model is subgraph independent but some subgraph independent distributions are not coloring models.

► **Lemma 36.**  $\mathcal{G}_{col} \subsetneq \mathcal{G}_{sub}$ , i.e., every coloring model is subgraph independent, but there are some subgraph independent distributions that are not coloring models.

**Proof.** We first prove  $\mathcal{G}_{col} \subseteq \mathcal{G}_{sub}$ . Consider some graph distribution  $D \in \mathcal{G}_{col}$  and consider any two disjoint subsets of vertices  $V, W$ . The resulting graph within  $V$  depends only on elementary experiments on  $V$ , and the resulting graph within  $W$  depends only on elementary experiments on  $W$ . Since the elementary experiments are mutually independent, we have independence of the resulting graphs within  $V$  and  $W$ . Thus,  $D \in \mathcal{G}_{sub}$ .

The fact that  $\mathcal{G}_{col} \neq \mathcal{G}_{sub}$  is well-known from prior work [1, 10, 14, 19]. ◀

► **Lemma 37.**  $\mathcal{G}_{sub} \subsetneq \mathcal{G}_{esub}$ , i.e., every subgraph independent distribution is edge-subgraph independent, but there are some edge-subgraph independent distributions that are not subgraph independent.

**Proof.** Clearly  $\mathcal{G}_{sub} \subseteq \mathcal{G}_{esub}$  as subgraph independence implies edge-subgraph independence by definition, since a single edge is also a subgraph.

### 30:18 On Connectivity in Random Graph Models with Limited Dependencies

To prove  $\mathcal{G}_{\text{sub}} \neq \mathcal{G}_{\text{esub}}$ , we consider the following graph distribution  $CC(n)$  on  $n \geq 6$  vertices: with probability  $1/2$ , the distribution returns a graph drawn from the Erdős-Rényi distribution  $G_{n,1/2}$ . Otherwise, the distribution picks a uniformly random decomposition of the vertex set  $[n]$  into two sets  $A$  and  $B$ , and returns the graph consisting of the union of cliques on  $A$  and  $B$ . We first show that  $CC(n)$  is edge-subgraph independent. In both the Erdős-Rényi regime as well as the two-cliques regime, for any edge  $e$  we have  $Pr[X_e] = 1/2$ . This holds even when we condition on the outcome within any subgraph disjoint from  $e$ , thus  $CC(n)$  is edge-subgraph independent. On the other hand, we show that  $CC(n)$  is not subgraph independent. To this end, we decompose the vertex set  $[n]$  into the sets  $P = \{u, v, w\}$  and  $Q = [n] \setminus P$ . We consider the two events  $\mathcal{P} :=$  “There are exactly two edges within  $P$ ” and  $\mathcal{Q} :=$  “ $Q$  is a clique”. Clearly, we have

$$Pr[\mathcal{Q}] = \frac{1}{2} \cdot 2^{-(n-4)} + \frac{1}{2} \cdot 2^{-\binom{n-3}{2}} > 2^{-\binom{n-3}{2}}.$$

On the other hand,  $\mathcal{P}$  implies that we are in the Erdős-Rényi regime, and thus

$$Pr[\mathcal{Q}|\mathcal{P}] = 2^{-\binom{n-3}{2}}.$$

We conclude that  $Pr[\mathcal{Q}] > Pr[\mathcal{Q}|\mathcal{P}]$  and thus  $\mathcal{P}$  and  $\mathcal{Q}$  are not independent, showing that  $CC(n)$  is not subgraph independent.  $\blacktriangleleft$

The graph distribution  $CC(n)$  that we built in the proof above is a *convex combination* of its two regimes, as defined in the full version of this paper.

► **Lemma 38.**  $\mathcal{G}_{\text{esub}} \subsetneq \mathcal{G}_{\text{mat}}$ , i.e., every edge-subgraph independent distribution is matching independent, but there are some matching independent distributions that are not edge-subgraph independent.

**Proof.** To prove  $\mathcal{G}_{\text{esub}} \subseteq \mathcal{G}_{\text{mat}}$ , consider some edge-subgraph independent distribution. To prove that it is also matching independent, let  $M = \{e_1, \dots, e_k\}$  be some matching. We show for all  $i \leq k$  that  $Pr[X_{e_i} | X_{e_1}, \dots, X_{e_{i-1}}] = Pr[X_{e_i}]$ . This implies

$$Pr[X_{e_1} \text{ and } X_{e_2} \text{ and } \dots \text{ and } X_{e_k}] = \prod_{i=1}^k Pr[X_{e_i}],$$

and thus implies matching independence. Let  $V'$  be the set of endpoints of the edges  $e_1, \dots, e_{i-1}$ . Furthermore, let  $\mathcal{G}(V')$  be the set of all possible graphs on  $V'$ . Then, by the law of total probability, we have

$$Pr[X_{e_i} | X_{e_1}, \dots, X_{e_{i-1}}] = \sum_{G \in \mathcal{G}(V')} Pr[X_{e_i} | G] Pr[G | X_{e_1}, \dots, X_{e_{i-1}}].$$

Due to edge-subgraph independence,  $Pr[X_{e_i} | G] = Pr[X_{e_i}]$ , and we thus have

$$Pr[X_{e_i} | X_{e_1}, \dots, X_{e_{i-1}}] = Pr[X_{e_i}] \sum_{G \in \mathcal{G}(V')} Pr[G | X_{e_1}, \dots, X_{e_{i-1}}] = Pr[X_{e_i}],$$

proving the desired claim.

For the second part of the statement,  $\mathcal{G}_{\text{esub}} \neq \mathcal{G}_{\text{mat}}$ , we consider the following graph distribution  $SC(n)$  on  $n$  vertices for odd  $n \geq 5$ : every vertex independently and uniformly picks from the two colors red and blue. Since  $n$  is odd, one color was picked by an even number of vertices. The vertices with that color form a clique, and no other edges are present.

We first show that  $SC(n)$  is matching independent: every edge individually occurs with probability  $1/4$ , since first both endpoints need to have the same color, and second this color must be the color picked by an even number of vertices. We prove that a matching of  $k$  edges occurs with probability  $(1/4)^k$ . For the matching to occur, all  $2k$  vertices must pick the same color, which happens with probability  $(1/2)^{2k-1}$ . Second, that color must end up to be the color picked by an even number of vertices, which has probability  $1/2$ . Altogether, we have probability  $(1/2)^{2k-1+1} = (1/4)^k$ .

Finally, we show that  $SC(n)$  is not edge-subgraph independent. Let  $W$  be a set of  $n - 2$  vertices, and  $e$  be the edge between the remaining 2 vertices. If  $W$  is a clique,  $e$  cannot be present. Since both  $W$  being a clique and  $e$  being present have non-zero marginal probabilities, edge-subgraph independence cannot hold. ◀

► **Lemma 39.**  $\mathcal{G}_{mat} \subsetneq \mathcal{G}_{pw}$ , i.e., every matching independent distribution is pairwise independent, but there are some pairwise independent distributions that are not matching independent.

To prove this final strict inclusion lemma, we define the following distribution. This distribution will also be useful later to prove a lower bound on  $\rho_{pw}$ .

► **Definition 40.**  $CM(n, q)$  for  $n$  even is the following distribution over the graphs on  $n$  vertices: to sample  $G \sim CM(n, q)$ , with probability  $q$  we sample from the clique regime, and otherwise from the matching regime. In the clique regime, we pick a vertex  $x \in [n]$  uniformly at random and add all edges  $e$  with  $x \notin e$  to  $G$ . In the matching regime, we pick a perfect matching on  $[n]$  uniformly at random and add its edges to  $G$ .

▷ **Claim 41.** There exists  $0 \leq q(n) \leq 1$  with

$$q(n) = 1 - \Theta\left(\frac{1}{n^2}\right)$$

such that  $CM(n, q(n))$  for  $n$  even is pairwise independent but not matching independent, and the probability of each edge is

$$p(n) = 1 - \frac{2}{n} - \Theta\left(\frac{1}{n^2}\right).$$

The proof of this claim is straightforward and just requires some calculations. It can be found in the full version.

**Proof of Lemma 39.** Clearly  $\mathcal{G}_{mat} \subseteq \mathcal{G}_{pw}$  as matching independence implies pairwise independence by definition, since two vertex-disjoint edges  $e, f$  are a matching.

To see that  $\mathcal{G}_{mat} \neq \mathcal{G}_{pw}$ , recall that by Claim 41, there exists some  $q(n)$  such that  $CM(n, q)$  is pairwise independent but not matching independent. ◀

**Proof of Theorem 6.** Theorem 6 now follows directly from Lemmas 36–39. ◀

## **B** Omitted Proofs from Section 3.5

**Proof of Lemma 18.** We first prove  $\rho_{col,2}(n) \geq 1/4$  by defining the following coloring model which is always disconnected. Decompose the vertex set into three non-empty sets  $A, B, C$ . Every vertex chooses a color uniformly among the colors red and blue. Two vertices in the same set are connected if they are both red. A vertex  $a \in A$  is connected to  $b \in B$ , if  $a$  is red and  $b$  is blue. The same goes for  $b \in B$  and  $c \in C$ , as well as  $c \in C$  and  $a \in A$ . Clearly,

the marginal edge probability of every edge is  $1/4$ . We now show that every connected component of a graph sampled from this distribution is a subset of the union of at most two of the sets  $A, B, C$ . To see this, we pick some edge  $\{a, b\}$  for  $a \in A$  and  $b \in B$  to be present, and try to grow its connected component. Since the edge  $\{a, b\}$  is present,  $a$  must be red and  $b$  blue. Now,  $a$  can only be connected to other red vertices in  $A$ , and to blue vertices in  $B$ . Similarly,  $b$  can only be connected to other red vertices in  $A$ . No red vertex in  $A$  or blue vertex in  $B$  can be connected to a vertex in  $C$ , thus the connected component containing the edge  $\{a, b\}$  is contained in  $A \cup B$ . Symmetrically this holds for any pair of sets, and we conclude that the graph must be disconnected, proving the lower bound.

Next, we prove  $\rho_{col,2}(n) \leq 1/4$  by showing that if every marginal edge probability is strictly larger than  $1/4$ , we can always find a coloring that connects the graph. We say that a vertex  $v$  covers vertex  $w$ , if for both colors at vertex  $w$ , there exists a color at  $v$ , such that the edge  $\{v, w\}$  is present under this coloring. We now remove vertices one by one, by repeatedly removing a vertex which covers some remaining vertex, until no more such vertices exist. If the graph on the remaining vertices  $V'$  can be connected using some coloring, we can add back the removed vertices in reverse order, and connect them to the vertex they cover, thus connecting the whole graph. We thus only have to show that the graph on  $V'$  can be connected. Since in this graph no vertex covers any other, each edge is only present under exactly one of the four possible color combinations of its endpoints.

We pick a vertex  $v' \in V'$  which maximizes  $\max(Pr[v \text{ is red}], Pr[v \text{ is blue}])$  among all  $v \in V'$ . We give this vertex  $v'$  the color which is more likely, let  $p$  be the probability of this color. One can see that for any other vertex  $w \in V' \setminus \{v'\}$ , there must be a color such that the edge  $\{v', w\}$  is present. Otherwise, the color combination making  $\{v', w\}$  present would have probability of at most  $p \cdot (1 - p)$ , which is at most  $1/4$ . Thus, we can connect  $V'$  by simply coloring each vertex with the correct color to connect to  $v'$ . This concludes the proof of the upper bound, and thus the whole lemma.  $\blacktriangleleft$

**Proof of Lemma 35.** We consider the potential edges between  $OR$  and  $OB$ . Each such edge  $r'b'$  can be oriented in some direction, by considering the following two probabilities:

$$p_{r' \rightarrow b'} := Pr_{c_{r'}}[\exists c_{b'} \text{ such that } r'b' \in E(G) \text{ in coloring } c(r') = c_{r'}, c(b') = c_{b'}]$$

$$p_{b' \rightarrow r'} := Pr_{c_{b'}}[\exists c_{r'} \text{ such that } r'b' \in E(G) \text{ in coloring } c(r') = c_{r'}, c(b') = c_{b'}]$$

Note that  $p_{r' \rightarrow b'} \cdot p_{b' \rightarrow r'} \geq Pr[r'b' \in E(G)] \geq p'$ , and thus  $\max(p_{r' \rightarrow b'}, p_{b' \rightarrow r'}) \geq \sqrt{p'}$ . We now direct the potential edge  $r'b'$  from  $r'$  to  $b'$  if  $p_{r' \rightarrow b'} > p_{b' \rightarrow r'}$ , and from  $b'$  to  $r'$  otherwise. We pick an arbitrary perfect matching  $M$  among the potential edges between  $OR$  and  $OB$ , using  $|OR| = |OB|$ .

**Case 1:** There exists an arc  $(b', r') \in M$  directed towards  $r' \in OR$ . In this case we will connect the graph as shown in Figure 5a. Every obligate red vertex is connected to its matching obligate blue vertex. The vertex  $b'$  must have a large neighborhood size that can be guaranteed with probability  $\geq 1 - \sqrt{p}$ , so that it can simultaneously be connected to  $r'$ . Every obligate blue and every non-obligate vertex connects to this neighborhood.

By Lemma 26, part 3, we have

$$Pr[|N(b', c)| \geq (1 - p)(1 - \beta)n + \varepsilon n/2] \geq 1 - \sqrt{p}.$$

Thus, with positive probability over the choice of color for  $b'$ , the size of the neighborhood of  $b'$  is at least  $(1 - p)(1 - \beta)n + \varepsilon n/2$  and there is a choice of color for  $r'$  such that  $r'b'$  is an edge, since  $1 - \sqrt{p} + \sqrt{p'} > 0$ .

Similarly, for all other obligate blue vertices  $u$ , by Lemma 26, part 1, we have

$$Pr[|N(u, c)| \geq \frac{p + p\beta - \beta}{p}n + \varepsilon n/2] \geq 1 - p,$$

so with positive probability there is an edge between  $u$  and its matching vertex in  $OR$  and  $u$ 's neighborhood has size at least  $\frac{p + p\beta - \beta}{p}n + \varepsilon n/2$ .

We need to prove that these guaranteed sizes of neighborhoods intersect outside of  $OB \cup OR$ . Let  $S := (1-p)(1-\beta)n + \varepsilon n/2$  and  $T := \frac{p + p\beta - \beta}{p}n + \varepsilon n/2$ . Since both neighborhoods can intersect  $R$  in at most  $C$  vertices, we need to prove that  $(S - 3C) + (T - 3C) > n - |R|$ , i.e.,

$$S + T + \alpha n - n - 6C \stackrel{!}{>} 0. \tag{1}$$

Additionally, to guarantee that each non-obligate vertex  $u$  outside the neighborhood of  $b'$  can connect to that neighborhood, we need to guarantee that  $u$  has a neighborhood which intersects that of  $b'$  outside  $OB \cup OR$ . Thus, we check that  $(S - 3C) + (p_u^b - 3C) > n - |R|$ , i.e.,

$$S + p_u^b + \alpha n - n - 6C \stackrel{!}{>} 0. \tag{2}$$

Since our lower bound on  $p_u^b$  is smaller than  $T$  (as  $\alpha \geq \beta$ ), (2) implies (1), so it suffices to show (2). The proof of (2) can be found in the full version.

**Case 2:** All edges in  $M$  are oriented towards  $OB$ . In this case we will connect the graph as shown in Figure 5b. We give each obligate red vertex a large neighborhood such that we can still guarantee to be able to connect the obligate blue vertices to their matching partner. Then, we give each non-obligate vertex which is not yet connected directly to any of the obligate vertices a large neighborhood not intersecting  $B$  (as guaranteed by Lemma 34). These neighborhoods must intersect all neighborhoods of the obligate red vertices due to their sizes, thus the graph is connected.

Let  $Y := (1-p)(1-\alpha)n + \varepsilon n/2$ . For any obligate red vertex  $v$ , by Lemma 26, part 4, we have  $Pr[|N(v, c_v)| \geq Y] \geq 1 - \sqrt{p}$ , so with probability  $1 - \sqrt{p} + \sqrt{p'} > 0$  there is a color for  $v$ 's obligate blue partner vertex so that they are connected by an edge and also  $|N(v, c_v)| \geq Y$ .

We now check that every non-obligate vertex  $u$  has a neighborhood which intersects with every neighborhood of size at least  $Y$  of the obligate red vertices. Recall that by Lemma 34, every  $u$  has a neighborhood not robustly intersecting  $B$  of size  $p_u^r$ . Note that we need to show that these neighborhoods intersect outside of  $OR \cup OB$  since these vertices may change their colors. To this end, it suffices to show that  $(p_u^r - 3C) + (Y - 3C) > n - |B|$ , i.e.,

$$p_u^r + Y + \beta n - n - 6C \stackrel{!}{>} 0,$$

which is easy to verify as can be seen in the full version. ◀

## C Finitely Many Colors Suffice for the Coloring Model

**Proof of Lemma 8.** Let  $\mathcal{D}$  be a random graph distribution given as a coloring model. We will reduce the number of colors for each vertex one by one, preserving the probability distribution  $\mathcal{D}$ . By a slight abuse of notation, here we refer to any outcome of the experiment

## 30:22 On Connectivity in Random Graph Models with Limited Dependencies

at some vertex as its *color*, even if there are infinitely many of them. Suppose we're considering vertex  $v$  and for each graph  $H$  on  $n$  vertices and color  $c$  of  $v$ , let  $p_{H,c}$  be the probability that  $H$  is sampled from  $\mathcal{D}$  conditioned on vertex  $v$  having color  $c$ , and let  $p_H$  be the unconditional probability of  $H$  being sampled from  $\mathcal{D}$ . Now for each color  $c$  of  $v$ , consider the vector  $\vec{p}_{*,c}$  in  $2^{\binom{n}{2}}$ -dimensional space that consists of all the  $p_{H,c}$  for all possible graphs  $H$  on  $n$  vertices in some canonical order. Consider also the vector  $\vec{p}_*$  of the same dimension with the  $p_H$  for all possible graphs  $H$  on  $n$  vertices as entries in the same canonical order. By the law of total probability,  $\vec{p}_*$  is contained in the convex hull of all the  $\vec{p}_{*,c}$  vectors (of which there may be infinitely many). It follows by Carathéodory's theorem [11] that  $\vec{p}_*$  is a convex combination of some  $2^{\binom{n}{2}} + 1$  many vectors among  $\{\vec{p}_{*,c} \mid c \text{ is a color of vertex } v\}$ . Thus we can pick the corresponding  $2^{\binom{n}{2}} + 1$  colors for  $v$ , use the coefficients given by that convex combination as probabilities for these colors in the experiment at vertex  $v$ , and we end up with the same distribution  $\mathcal{D}$ . We repeat this process for each vertex, completing the proof. ◀