

# Coloring Locally Sparse Graphs

James Anderson<sup>a</sup>      Abhishek Dhawan<sup>b</sup>      Aiya Kuchukova<sup>a</sup>

Submitted: Nov 26, 2024; Accepted: Dec 11, 2025; Published: Feb 13, 2026

© The authors. Released under the CC BY-ND license (International 4.0).

## Abstract

A graph  $G$  is  $k$ -*locally sparse* if for each vertex  $v \in V(G)$ , the subgraph induced by its neighborhood contains at most  $k$  edges. Alon, Krivelevich, and Sudakov showed that for  $f > 0$  if a graph  $G$  of maximum degree  $\Delta$  is  $\Delta^2/f$ -locally-sparse, then  $\chi(G) = O(\Delta/\log f)$ . We introduce a more general notion of local sparsity by defining graphs  $G$  to be  $(k, F)$ -*locally-sparse* for some graph  $F$  if for each vertex  $v \in V(G)$  the subgraph induced by the neighborhood of  $v$  contains at most  $k$  copies of  $F$ . Employing the Rödl nibble method, we prove the following generalization of the above result: for every bipartite graph  $F$ , if  $G$  is  $(k, F)$ -locally-sparse, then  $\chi(G) = O(\Delta/\log(\Delta k^{-1/|V(F)|}))$ . This improves upon results of Davies, Kang, Pirot, and Sereni who consider the case when  $F$  is a path. Our results also recover the best known bound on  $\chi(G)$  when  $G$  is  $K_{1,t,t}$ -free for  $t \geq 4$ , and hold for list and correspondence coloring in the more general so-called “color-degree” setting.

**Mathematics Subject Classifications:** 05C88, 05C89

## 1 Introduction

All graphs considered are finite, undirected, and simple. A *coloring* of a graph  $G = (V(G), E(G))$  is a function  $\varphi : V(G) \rightarrow \mathbb{N}$ , and it is *proper* if  $\varphi(x) \neq \varphi(y)$  whenever  $xy \in E(G)$ . The *chromatic number* of a graph  $G$ , denoted  $\chi(G)$ , is the smallest size of a set  $C \subseteq \mathbb{N}$  such that  $G$  has a proper coloring  $\varphi : V(G) \rightarrow C$ . Determining  $\chi(G)$  for various classes of graphs has been a central topic in graph theory. In this paper, we generalize and extend several results which bound  $\chi(G)$  when  $G$  satisfies various local sparsity conditions.

### 1.1 Preliminaries

For  $k \in \mathbb{R}$ , we say  $G$  is  **$k$ -locally-sparse** if for each  $v \in V(G)$ , the subgraph induced by  $N(v)$  contains at most  $\lfloor k \rfloor$  edges (letting  $k \in \mathbb{R}$  as opposed to  $k \in \mathbb{N}$  permits us to state

---

<sup>a</sup>School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA  
(james.anderson@math.gatech.edu, aiya.kuchukova@math.gatech.edu).

<sup>b</sup>Department of Mathematics, University of Illinois Urbana-Champaign, Urbana, IL, USA  
(adhawan2@illinois.edu).

results more cleanly by allowing us to avoid taking floors). Molloy and Reed showed that if a graph  $G$  is locally sparse, then  $\chi(G)$  can be bounded from above as follows:

**Theorem 1** ([32, Lemma 2]). *For every  $\delta > 0$ , there is  $\varepsilon > 0$  such that for  $\Delta \in \mathbb{N}$  sufficiently large (in terms of  $\delta$  and  $\varepsilon$ ), if  $G$  is a  $(1-\delta)\binom{\Delta}{2}$ -locally-sparse graph of maximum degree  $\Delta$ , then  $\chi(G) \leq (1-\varepsilon)(\Delta+1)$ .*

Alon, Krivelevich, and Sudakov generalized this result, giving an asymptotic improvement when the parameter  $f$  below is large (for example, when  $f = \Theta(\Delta)^1$ ).

**Theorem 2** ([4, Theorem 1.1]). *There exists a constant  $C > 0$  such that for  $\Delta \in \mathbb{N}$  sufficiently large and all  $f > 1$ , if  $G$  is a  $\frac{\Delta^2}{f}$ -locally-sparse graph of maximum degree  $\Delta$ , then  $\chi(G) \leq C \Delta / \log f$ .*

In the same paper, they show the above is tight up to the value of  $C$  [4, Proposition 1.2]. Davies, Kang, Pirot, and Sereni proved Theorem 2 holds with  $C = 1/2 + o(1)$ , which remains the best known bound for this problem [14, Theorem 5]; we note that this result holds in the more general setting of correspondence coloring (see §1.2 for a description of this setting).

For graphs  $F$  and  $G$ , a **copy** of  $F$  in  $G$  is a subgraph  $H \subseteq G$  (not necessarily induced) which is isomorphic to  $F$ . We say a graph  $G$  is  **$F$ -free** if  $G$  contains no copies of  $F$ . As a corollary to Theorem 2, Alon, Krivelevich, and Sudakov showed the following, where for  $t \geq 1$ ,  $K_{1,t,t}$  is the complete tripartite graph with set sizes 1,  $t$ ,  $t$ :

**Corollary 3** ([4]). *Let  $t \geq 1$ . Fix an arbitrary graph  $F \subseteq K_{1,t,t}$ . Then there exists a real value  $c_F > 0$  such that if  $G$  is an  $F$ -free graph of maximum degree  $\Delta$ , then*

$$\chi(G) \leq (c_F + o_\Delta(1)) \frac{\Delta}{\log \Delta}.$$

An outline of their proof goes as follows: if  $F \subseteq K_{1,t,t}$  and  $G$  is  $F$ -free, then the neighborhood of each vertex in  $G$  does not contain  $K_{t,t}$  as a subgraph. The celebrated Kővári–Sós–Turán Theorem below then implies the neighborhood of each vertex contains relatively few edges:

**Theorem 4** (Kővári–Sós–Turán [28]; see also Hyltén-Cavallius [21]). *Let  $G$  be a bipartite graph with a bipartition  $V(G) = X \sqcup Y$ , where  $|X| = m$ ,  $|Y| = n$ , and  $m \geq n$ . Suppose that  $G$  does not contain a complete bipartite subgraph with  $s$  vertices in  $X$  and  $t$  vertices in  $Y$ . Then*

$$|E(G)| \leq s^{1/t} m^{1-1/t} n + tm.$$

Thus when  $G$  is  $F$ -free,  $G$  is  $k$ -locally-sparse for  $k = \Theta_t(\Delta^{2-1/t})$ , allowing Theorem 2 to be applied, completing the proof. This result led Alon, Krivelevich, and Sudakov to conjecture that Corollary 3 holds for all graphs  $F$ , not just  $F \subseteq K_{1,t,t}$ .

---

<sup>1</sup>Throughout this work, we use the standard asymptotic notation  $O(\cdot)$ ,  $\Omega(\cdot)$ ,  $o(\cdot)$ , etc.

**Conjecture 5** ([4, Conjecture 3.1]). Fix an arbitrary graph  $F$ . Then there exists a real value  $c_F > 0$  such that if  $G$  is  $F$ -free and has maximum degree  $\Delta$ , then

$$\chi(G) \leq (c_F + o_\Delta(1)) \frac{\Delta}{\log \Delta}.$$

This conjecture has spurred a large research effort; we summarize the progress toward it in §2.1. In this paper, we introduce a more general notion of local sparsity:

**Definition 6.** Let  $F$  be a graph and let  $k \in \mathbb{R}$ . A graph  $G$  is  **$(k, F)$ -sparse** if  $G$  contains at most  $\lfloor k \rfloor$  copies of  $F$  (not necessarily vertex-disjoint). A graph  $G$  is  **$(k, F)$ -locally-sparse** if, for every  $v \in V(G)$ , the induced subgraph  $G[N(v)]$  is  $(k, F)$ -sparse.

As a graph  $G$  is  $k$ -locally-sparse if and only if it is  $(k, K_2)$ -locally-sparse, this definition generalizes the usual notion of local sparsity. The case of  $(k, P_t)$ -locally-sparse graphs (where  $P_t$  is the path on  $t$  vertices) was investigated by Davies, Kang, Pirot, and Sereni. While this result was not stated explicitly in their paper, they discuss the bound in section 4.

**Theorem 7** ([13]). *For every  $t \geq 2$ ,  $\varepsilon > 0$ , there exist  $\Delta_0 \in \mathbb{N}$  and  $C > 0$  such that whenever  $\Delta \geq \Delta_0$ , the following holds: Let  $k \geq 1/2$  and let  $G$  be a  $(k, P_t)$ -locally-sparse graph of maximum degree  $\Delta$ . If  $k \leq \Delta^2/C$ , then*

$$\chi(G) \leq (1 + \varepsilon) \frac{\Delta}{\log(\Delta/\sqrt{k})}.$$

By taking  $t = 2$ , this bound matches their earlier result [14, Theorem 5] (which improves upon the constant factor of Theorem 2). However, for constant  $t > 2$ , the neighborhood of a vertex of a graph with maximum degree  $\Delta$  can contain  $\Theta_t(\Delta^t)$  copies of  $P_t$ . As  $\Delta^2 \ll \Delta^t$  for large  $t$ , there is wide room to improve the range of  $k$  in Theorem 7.

Our main theorem improves upon Theorem 7 in three ways: first, we consider all bipartite graphs  $F$  rather than just paths  $P_t$ ; second, we consider the range  $k \leq \Delta^{|V(F)|/10}$ , which is an improvement for  $|V(F)| \geq 20$ ; finally, we provide an asymptotic improvement on the dependence of  $\chi(G)$  on  $k$ . This improvement is stated below:

**Theorem 8.** *For every  $\varepsilon > 0$  and every bipartite graph  $F$ , the following holds for  $\Delta$  large enough (in terms of  $\varepsilon$  and  $F$ ). Let  $1/2 \leq k \leq \Delta^{|V(F)|/10}$  and let  $C$  be defined as follows:*

$$C := \begin{cases} 4 + \varepsilon, & \text{if } k \leq \Delta^{\varepsilon|V(F)|/200}, \\ 8, & \text{otherwise.} \end{cases}$$

*Let  $G$  be a  $(k, F)$ -locally-sparse graph of maximum degree  $\Delta$ . Then,*

$$\chi(G) \leq \frac{C \Delta}{\log(\Delta k^{-1/|V(F)|})}.$$

We note that Theorem 8 holds in the more general so-called “color-degree” setting for *list coloring* and *correspondence coloring*, which is new even for  $F = K_2$ ; see Theorem 13.

## 1.2 List coloring, correspondence coloring, and the color-degree setting

Introduced independently by Vizing [39] and Erdős, Rubin, and Taylor [19], *list coloring* is a generalization of graph coloring in which each vertex is assigned a color from its own predetermined list of colors. Formally,  $L : V(G) \rightarrow 2^{\mathbb{N}}$  is a *list assignment* for  $G$ , and an  $L$ -coloring of  $G$  is a proper coloring  $\varphi : V(G) \rightarrow \mathbb{N}$  such that  $\varphi(v) \in L(v)$  for each  $v \in V(G)$ . When  $|L(v)| \geq q$  for each  $v \in V(G)$ , where  $q \in \mathbb{N}$ , we say  $L$  is  $q$ -fold. The *list chromatic number* of  $G$ , denoted  $\chi_\ell(G)$ , is the smallest  $q$  such that  $G$  has an  $L$ -coloring for every  $q$ -fold list assignment  $L$  for  $G$ .

It is often convenient to view list coloring from a different perspective. Given a graph  $G$  and a list assignment  $L$  for  $G$ , we create an auxiliary graph  $H$  as follows:

$$\begin{aligned} V(H) &:= \{(v, c) \in V(G) \times \mathbb{N} : c \in L(v)\} \\ E(H) &:= \{\{(v, c), (u, d)\} : vu \in E(G), c = d\}. \end{aligned}$$

We call  $H$  a *cover graph* of  $G$ , and the pair  $(L, H)$  a *list cover* of  $G$ . An  $L$ -coloring of  $G$  is then an independent set  $I$  in  $H$  which satisfies  $|I| = |V(G)|$ , and, for each vertex  $v \in V(G)$ , there exists  $c \in \mathbb{N}$  such that  $(v, c) \in I$ . Thus  $I$  selects exactly one vertex of the form  $(v, c)$  for each  $v \in V(G)$ .

**Correspondence coloring** (also known as **DP-coloring**) is a generalization of list coloring introduced by Dvořák and Postle [17] in order to solve a question of Borodin. Just as in list coloring, each vertex is assigned a list of colors,  $L(v)$ ; in contrast to list coloring, though, the identifications between the colors in the lists are allowed to vary from edge to edge. That is, each edge  $uv \in E(G)$  is assigned a matching  $M_{uv}$  (not necessarily perfect and possibly empty) from  $L(u)$  to  $L(v)$ . A proper correspondence coloring is a mapping  $\varphi : V(G) \rightarrow \mathbb{N}$  satisfying  $\varphi(v) \in L(v)$  for each  $v \in V(G)$  and  $\varphi(u)\varphi(v) \notin M_{uv}$  for each  $uv \in E(G)$ . Formally, correspondence colorings are defined in terms of an auxiliary graph known as a **correspondence cover** of  $G$ .

**Definition 9** (Correspondence Cover). A **correspondence cover** of a graph  $G$  is a pair  $\mathcal{H} = (L, H)$ , where  $H$  is a graph and  $L : V(G) \rightarrow 2^{V(H)}$  such that:

- (CC1) The set  $\{L(v) : v \in V(G)\}$  forms a partition of  $V(H)$ ,
- (CC2) For each  $v \in V(G)$ ,  $L(v)$  is an independent set in  $H$ , and
- (CC3) For each  $u, v \in V(G)$ , the edge set of  $H[L(u) \cup L(v)]$  forms a matching, which is empty if  $uv \notin E(G)$ .

We call the vertices of  $H$  **colors**. For  $c \in V(H)$ , we let  $L^{-1}(c)$  denote the **underlying vertex** of  $c$  in  $G$ , i.e., the unique vertex  $v \in V(G)$  such that  $c \in L(v)$ . If two colors  $c, c' \in V(H)$  are adjacent in  $H$ , we say that they **correspond** to each other and write  $c \sim c'$ .

An  **$\mathcal{H}$ -coloring** is a mapping  $\varphi : V(G) \rightarrow V(H)$  such that  $\varphi(v) \in L(v)$  for all  $v \in V(G)$ . Similarly, a **partial  $\mathcal{H}$ -coloring** is a partial mapping  $\varphi : V(G) \dashrightarrow V(H)$  such that  $\varphi(v) \in L(v)$  whenever  $\varphi(v)$  is defined. A (partial)  $\mathcal{H}$ -coloring  $\varphi$  is **proper** if the

image of  $\varphi$  is an independent set in  $H$ , i.e., if  $\varphi(u) \not\sim \varphi(v)$  for all  $u, v \in V(G)$  such that  $\varphi(u)$  and  $\varphi(v)$  are both defined. Notice, then, that the image of a proper  $\mathcal{H}$ -coloring of  $G$  is exactly an independent set  $I \subseteq V(H)$  with  $|I \cap L(v)| = 1$  for each  $v \in V(G)$ .

A correspondence cover  $\mathcal{H} = (L, H)$  is  **$q$ -fold** if  $|L(v)| \geq q$  for all  $v \in V(G)$ . The **correspondence chromatic number** of  $G$ , denoted by  $\chi_c(G)$ , is the smallest  $q$  such that  $G$  admits a proper  $\mathcal{H}$ -coloring for every  $q$ -fold correspondence cover  $\mathcal{H}$ .

Note that a list cover  $(L, H)$  of  $G$  is a correspondence cover of  $G$ , where each matching  $M_{uv}$  is such that  $(u, c)$  matches with  $(v, c)$  for each  $c \in \mathbb{N}$ . As classical coloring is the special case of list coloring in which all lists are identical, it follows that  $\chi(G) \leq \chi_\ell(G) \leq \chi_c(G)$ .

When considering locally sparse graphs, a correspondence coloring version of Theorem 7 was proven by Davies, Kang, Pirot, and Sereni, albeit for a smaller range of  $k$ . Once again, while they do not explicitly state this result, they discuss it in section 4 of [13].

**Theorem 10** (Correspondence Coloring version of Theorem 7, [13]). *For every  $t \geq 2$ ,  $\varepsilon > 0$ , there exists  $\Delta_0 \in \mathbb{N}$  such that whenever  $\Delta > \Delta_0$ , the following holds: let  $k \geq 1/2$  and let  $G$  be a  $(k, P_t)$ -locally-sparse graph of maximum degree  $\Delta$ . If  $k \leq \Delta^2/(\log \Delta)^{2/\varepsilon}$ , then*

$$\chi_c(G) \leq (1 + \varepsilon) \frac{\Delta}{\log(\Delta/\sqrt{k})}.$$

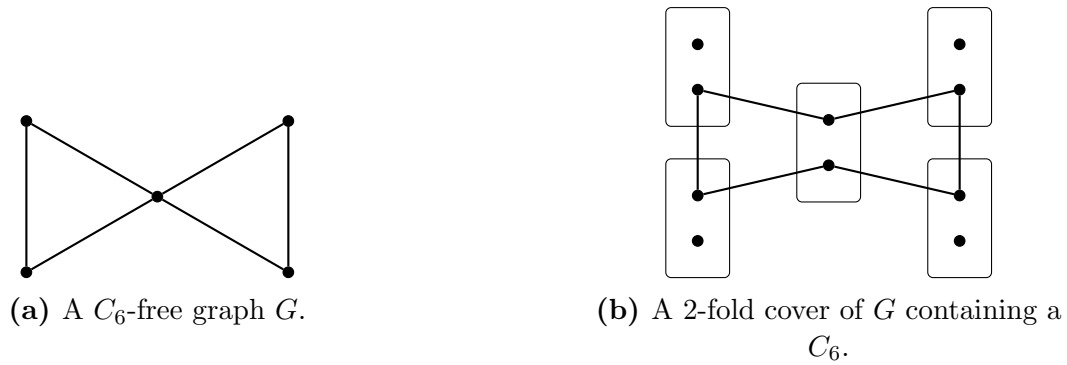
A curious feature of correspondence coloring is that structural constraints can be placed on the cover graph  $H$  instead of on the underlying graph  $G$ . For instance, if  $\mathcal{H} = (L, H)$  is a correspondence cover of a graph  $G$ , then  $\Delta(H) \leq \Delta(G)$ , so an upper bound on  $\Delta(H)$  is a weaker assumption than the same upper bound on  $\Delta(G)$ , and there exist a number of results in list and correspondence coloring in which the number of available colors given to each vertex is a function of  $\Delta(H)$  as opposed to  $\Delta(G)$ . This framework, often referred to as the **color-degree setting**, was pioneered by Kahn [25], Kim [27], Johansson [24, 23], and Reed [35], among others. For a selection of a few more recent examples, see [8, 36, 29, 2, 12, 26, 20, 6, 5] (see, also, §2.2). Similarly, as a result of condition (CC3), if  $G$  is  $(k, F)$ -locally-sparse, then so is  $H$ . However, the same does not hold for  $(k, F)$ -sparsity. For example, consider a  $q$ -fold list cover where all the lists are the same. Then for any graph  $F$ , each copy of  $F$  in  $G$  corresponds to at least  $q$  copies of  $F$  in  $H$ . Thus  $G$  may contain just one copy of  $F$ , while  $H$  contains many. Even for  $F$ -free graphs  $G$ , there can exist a cover  $H$  of  $G$  which contains  $F$ . For instance, in Fig. 1, the graph  $G$  is  $C_6$ -free, though its cover contains a  $C_6$ . Nevertheless, appropriate conditions exist that ensure whenever  $G$  is  $F$ -free, then so is any cover  $H$  of  $G$ . We summarize our observations in the following proposition (see §A for the proof):

**Proposition 11.** *Let  $F$  and  $G$  be graphs and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$ . For  $k \in \mathbb{R}$ , the following holds:*

(S1) *If  $G$  is  $(k, F)$ -locally-sparse, then so is  $H$ .*

(S2) *If  $F$  satisfies*

$$\forall u, v \in V(F), \quad uv \notin E(F) \implies N_F(u) \cap N_F(v) \neq \emptyset,$$



**Figure 1:** A  $C_6$ -free graph with a 2-fold cover containing a  $C_6$ .

then if  $G$  is  $F$ -free,  $H$  is as well.

For any graph  $F$ , there is a complete  $\chi(F)$ -partite graph  $F'$  such that  $F \subseteq F'$  and  $|V(F)| = |V(F')|$ . If  $G$  is  $F$ -free, then it is  $F'$ -free as well. Note that any complete  $r$ -partite graph  $F$  for  $r \geq 2$  satisfies the conditions of (S2). Therefore, we obtain the following corollary:

**Corollary 12.** *Let  $G$  be an  $F$ -free graph for some  $F$  satisfying  $\chi(F) = q \geq 2$  and let  $\varphi$  be a proper  $q$ -coloring of  $F$ . Define  $F'$  to be the complete  $q$ -partite graph with partitions  $V_1, \dots, V_q$  satisfying  $|V_i| = |\varphi^{-1}(i)|$ . For any correspondence cover  $\mathcal{H} = (L, H)$  of  $G$ , the cover graph  $H$  is  $F'$ -free.*

### 1.3 Main results

We are ready to present our main results. We begin with a color-degree version of Theorem 8 for correspondence coloring.

**Theorem 13.** *There exists a constant  $\alpha > 0$  such that for every  $\varepsilon > 0$ , there is  $d^* \in \mathbb{N}$  such that the following holds. Suppose that  $d, s, t \in \mathbb{N}$ , and  $k \in \mathbb{R}$  satisfy*

$$d \geq d^*, \quad 1 \leq t \leq s, \quad st \leq \frac{\alpha \varepsilon \log d}{\log \log d}, \quad \text{and} \quad 1/2 \leq k \leq d^{(s+t)/10}.$$

Define  $C$  as follows:

$$C := \begin{cases} 4 + \varepsilon & \text{if } k \leq d^{\varepsilon(s+t)/200}, \\ 8 & \text{otherwise.} \end{cases}$$

If  $G$  is a graph and  $\mathcal{H} = (L, H)$  is a correspondence cover of  $G$  such that:

- (i)  $H$  is  $(k, K_{s,t})$ -locally-sparse,
- (ii)  $\Delta(H) \leq d$ , and
- (iii)  $|L(v)| \geq C d / \log(dk^{-1/(s+t)})$  for all  $v \in V(G)$ ,

then  $G$  admits a proper  $\mathcal{H}$ -coloring.

We note that both the local sparsity and the degree constraints are on the cover graph  $H$  as opposed to  $G$ , which is a weaker assumption as a result of Proposition 11. This is new even for  $s = t = 1$ .<sup>2</sup> We also remark that the upper bound  $k \leq d^{(s+t)/10}$  can be relaxed slightly while increasing the value  $C$ , however, our approach does not work for  $k = \Theta(d^{s+t})$ . We discuss this further in §3.

Note that for any bipartite graph  $F$ , a  $(k, F)$ -locally-sparse graph is  $(k, K_{s,t})$ -locally-sparse for appropriate  $s, t \in \mathbb{N}$  satisfying  $s + t = |V(F)|$ . Therefore, by treating  $s$  and  $t$  as constants we immediately obtain the following corollary:

**Corollary 14.** *For every  $\varepsilon > 0$  and bipartite graph  $F$ , the following holds for  $d$  large enough. Let  $1/2 \leq k \leq d^{|V(F)|/10}$  and let  $C$  be defined as follows:*

$$C := \begin{cases} 4 + \varepsilon & \text{if } k \leq d^{\varepsilon|V(F)|/200}, \\ 8 & \text{otherwise.} \end{cases}$$

Let  $G$  be a graph and  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  such that:

1.  $H$  is  $(k, F)$ -locally-sparse,
2.  $\Delta(H) \leq d$ , and
3.  $|L(v)| \geq C d / \log(dk^{-1/|V(F)|})$  for all  $v \in V(G)$ ,

then  $G$  admits a proper  $\mathcal{H}$ -coloring.

Setting  $d = \Delta(G)$  together with Proposition 11 yields the following corollary:

**Corollary 15.** *For every  $\varepsilon > 0$  and bipartite graph  $F$ , the following holds for  $\Delta$  large enough. Let  $1/2 \leq k \leq \Delta^{|V(F)|/10}$  and let  $C$  be defined as follows:*

$$C := \begin{cases} 4 + \varepsilon, & \text{if } k \leq \Delta^{\varepsilon|V(F)|/200}, \\ 8, & \text{otherwise.} \end{cases}$$

Let  $G$  be a  $(k, F)$ -locally-sparse graph of maximum degree  $\Delta$ . Then,

$$\chi_c(G) \leq \frac{C \Delta}{\log(\Delta k^{-1/|V(F)|})}.$$

As  $\chi(G) \leq \chi_c(G)$ , this clearly implies Theorem 8, and, for  $F = P_2$ , recovers Theorem 10 up to a constant (albeit with a stricter bound on  $k$ ). More importantly, for  $t \geq 3$ , our results asymptotically improve the bound in Theorem 10 (i.e.,  $k^{-1/t}$  as opposed to  $k^{-1/2}$ ). Observe as well that for  $t \geq 20$ , our results hold for an asymptotically larger range of  $k$  than Theorem 10.

---

<sup>2</sup>This special case was further investigated in follow-up work of the second named author of this manuscript with a focus on algorithmic implications [16].

We conclude this section with a summary of the results contained in this manuscript. Primarily, we introduce a generalized notion of local sparsity, allowing us to rephrase the results of Alon, Krivelevich, and Sudakov [4] and Davies, Kang, Pirot, and Sereni [14] as theorems about  $(k, F)$ -locally-sparse graphs for when  $F$  is an edge or a path, respectively. Our theorem considers arbitrary bipartite graphs  $F$ , thereby recovering and improving upon both of these results. In addition, we consider the more general setting of placing sparsity and degree constraints on the cover graph  $H$  rather than on the underlying graph  $G$ . This is the first color-degree version of Theorem 2, even for list coloring. We also recover a result of [6] (which we discuss further in subsequent sections; see Theorem 19), which to date is the best known bound on  $\chi(G)$  for the largest class of graphs known to satisfy Conjecture 5.

## Structure of the paper

The rest of the paper is structured as follows. In §2, we discuss related works in the graph coloring literature to better place our results in context. In §3, we provide an informal overview of our proof techniques. In §4, we will formally describe a coloring procedure we employ in our proof and prove a key lemma regarding its output. In §5, we apply this key lemma iteratively to prove Theorem 13. Finally, in §6, we discuss potential avenues for future research.

## 2 Discussion of previous work

In this section, we discuss related works in the graph coloring literature. Specifically, we survey results related to Conjecture 5 and the color-degree setting, and compare our proof techniques to other approaches in the area.

### 2.1 History of the Alon–Krivelevich–Sudakov conjecture

A trivial upper bound on  $\chi(G)$  comes from a greedy coloring, which shows  $\chi(G) \leq \Delta(G) + 1$ , where  $\Delta(G)$  is the maximum degree of  $G$ . Brooks improved this bound to  $\chi(G) \leq \Delta(G)$  when  $\Delta(G) \geq 3$  and  $G$  contains no cliques of size  $\Delta(G) + 1$  [10]. Reed improved upon this, showing that  $\chi(G) \leq \Delta(G) - 1$  for sufficiently large  $\Delta(G)$  when  $G$  contains no cliques of size  $\Delta(G)$  [34]. Conjecture 5 is a natural extension of these results, i.e., can the theorems of Brooks and Reed be improved by forbidding subgraphs  $F$  other than cliques? Johansson [23] showed that for any graph  $F$ , if  $G$  is  $F$ -free and has maximum degree  $\Delta$ , then

$$\chi(G) = O\left(\frac{\Delta \log \log \Delta}{\log \Delta}\right),$$

where the  $O(\cdot)$  is with respect to  $\Delta$  and hides constants that may depend on  $F$ . This result was never published, however, employing a result of Shearer [38], both Molloy [30] and Bernshteyn [7] provide a proof of this result in the more general settings of list coloring and correspondence coloring, respectively. In [24], Johansson removed the  $\log \log \Delta$  factor for  $F = K_3$  (see Molloy and Reed’s book [31, Chapter 13] for a textbook presentation of



this proof), and in 1999, Alon, Krivelevich, and Sudakov extended this result to ***almost-bipartite graphs***, i.e., subgraphs of  $K_{1,t,t}$  for some  $t$  (see Corollary 3), leading them to pose Conjecture 5.

To date, almost-bipartite graphs remain the largest class of connected graphs for which Conjecture 5 is known to hold, and generalizing this result to a larger class of graphs remains a tantalizing open problem. However, as long as  $F$  contains a cycle, the bound in Conjecture 5 is best possible up to the value of  $c_F$ , since there exist  $\Delta$ -regular graphs  $G$  of arbitrarily high girth with  $\chi(G) \geq (1/2)\Delta/\log \Delta$  [9] (when  $F$  is acyclic, a simple degeneracy argument shows  $\chi(G) = O(1)$ ). Therefore, much research has focused on improving upper bounds on the value of  $c_F$  in an attempt to try to close the gap between the known upper and lower bounds. We highlight progress in Table 1 for the interested reader. (Note that analogues of all these results hold for *list* and *correspondence* coloring.)

$F$	$c_F$	References
forest	0	Follows from degeneracy
not forest	$\geq 1/2$	Bollobás [9]
$K_3$	finite	Johansson [24]
	$\leq 4$	Pettie–Su [33]
	$\leq 1$	Molloy [30]
cycle/fan	$\leq 1$	Davies–Kang–Pirot–Sereni [14, §5.2, §5.5]
bipartite	$\leq 1$	Anderson–Bernshteyn–Dhawan [5]
$K_{1,t,t}$	$O(t)$	Alon–Krivelevich–Sudakov [4]
	$\leq t$	Davies–Kang–Pirot–Sereni [14, §5.6]
	$\leq 4$	Anderson–Bernshteyn–Dhawan [6]

**Table 1:** Known bounds on  $c_F$ .

In particular, Davies, Kang, Pirot, and Sereni proved  $c_F \leq 1$  for cycles and fans [14, §5.2, §5.5], and in the same paper, improved upon the result of Alon, Krivelevich, and Sudakov by showing  $c_F \leq t$  for  $F = K_{1,t,t}$  [14, §5.6]. The first two authors of this paper, together with Bernshteyn, showed  $c_F \leq 1$  when  $F$  is bipartite [5]; and, in addition, showed  $c_F \leq 4$  when  $F \subseteq K_{1,t,t}$ , thereby making the bound for  $c_F$  independent of  $t$  [6]. This improves the bounds of Davies, Kang, Pirot, and Sereni when  $t \geq 5$ . We note that when we set  $k < 1$ , Theorem 8 recovers the aforementioned result of [6].

## 2.2 The color-degree setting

Recall from §2.1 the result of Johansson that  $\chi(G) = O(\Delta/\log \Delta)$  for  $K_3$ -free graphs  $G$  having maximum degree  $\Delta$  sufficiently large. Alon and Assadi extended Johansson’s result in the list coloring setting by placing degree restrictions on the cover graph  $H$  as opposed to  $G$ .

**Theorem 16** ([2, Proposition 3.2]). *The following holds for  $d$  sufficiently large. Let  $G$  be a  $K_3$ -free graph and let  $\mathcal{H} = (L, H)$  be a list cover of  $G$  satisfying the following:*

$$|L(v)| \geq \frac{8d}{\log d}, \quad \Delta(H) \leq d.$$

*Then,  $G$  admits a proper  $\mathcal{H}$ -coloring.*

Alon, Cambie, and Kang were able to prove a similar bound in the color-degree setting for bipartite graphs  $G$  with an improved constant of  $1 + o(1)$  as opposed to 8 [3]. In subsequent work, Cambie and Kang extended these results on bipartite graphs to correspondence coloring, where they show the following:

**Theorem 17** ([12, Corollary 1.3]). *For all  $\varepsilon > 0$ , the following holds for  $d$  sufficiently large. Let  $G$  be a bipartite graph and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  satisfying the following:*

$$|L(v)| \geq (1 + \varepsilon) \frac{d}{\log d}, \quad \Delta(H) \leq d.$$

*Then,  $G$  admits a proper  $\mathcal{H}$ -coloring.*

Cambie and Kang conjectured that Theorem 17 holds for  $K_3$ -free graphs as well, i.e., a color-degree version of Bernshteyn's result mentioned earlier.

**Conjecture 18** ([12, Conjecture 1.4]). *For all  $\varepsilon > 0$ , the following holds for  $d$  sufficiently large. Let  $G$  be a  $K_3$ -free graph and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  satisfying the following:*

$$|L(v)| \geq (1 + \varepsilon) \frac{d}{\log d}, \quad \Delta(H) \leq d.$$

*Then,  $G$  admits a proper  $\mathcal{H}$ -coloring.*

This is still an open problem, however, the first two authors of this manuscript along with Bernshteyn made progress towards it in [6], achieving a bound of  $(4 + \varepsilon)d/\log d$ . They also generalized the result from triangle-free graphs  $G$  to  $F$ -free covers  $H$  when  $F$  is almost-bipartite:

**Theorem 19** ([6, Corollary 1.11]). *Let  $F$  be an almost-bipartite graph. For all  $\varepsilon > 0$ , the following holds for  $d$  sufficiently large. Let  $G$  be a graph and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  satisfying the following:*

- $H$  is  $F$ -free,
- $\Delta(H) \leq d$ , and
- $|L(v)| \geq (4 + \varepsilon)d/\log d$ .

*Then  $G$  admits a proper  $\mathcal{H}$ -coloring.*

As  $K_3$  is an almost-bipartite graph, Theorem 19 along with Corollary 12 constitutes progress toward Conjecture 18. We remark that it is not known if Conjecture 18 holds even for list covers.

## 2.3 Comparison of proof techniques with prior work

In this section, we compare our techniques to earlier approaches in related works. A core component of our proof is a result showing that  $(k, K_{s,t})$ -locally-sparse graphs are  $k'$ -locally-sparse for appropriate  $k'$  (see Proposition 22). We prove Theorem 13 by combining this result with the “Rödl nibble method” for graph coloring, an iterative procedure to construct the desired coloring. We provide a detailed overview in §3.

The astute reader may wonder why Proposition 22 together with Theorem 10 for  $t = 2$  cannot be used to prove Corollary 15. The bound obtained by this method would be as follows for  $k^\star = \frac{(s+t)^2 \Delta^{s+t-1}}{2 s^s t^t}$ :

$$\chi_c(G) = \begin{cases} O\left(\frac{st\Delta}{\log \Delta}\right) & \text{if } k \leq k^\star; \\ O\left(\frac{st\Delta}{(s+t)\log(\Delta k^{-1/(s+t)})}\right) & \text{if } k > k^\star. \end{cases} \quad (2.1)$$

Note that our result falls in the setting where  $k \leq k^\star$ . Our approach not only yields an improved dependence on  $k$ , but also achieves a constant factor independent of  $s$  and  $t$ . Additionally, our result is significantly stronger in the regime  $st = \omega(1)$ . Indeed, while Corollary 15 is not stated as such, one can deduce from Theorem 13 that the result holds for all  $s, t$  satisfying  $st = O(\log \Delta / \log \log \Delta)$ . Clearly, Corollary 15 outperforms the bound from (2.1) in this regime. Finally, we recall that our result yields the first color-degree proof of Theorem 2 as a special case. In fact, the proof strategies of both Theorem 2 and Theorem 10 fail in this setting, which is considered the benchmark in the graph coloring literature.

We remark that a similar distinction in approach is evident when comparing the works [14] and [6] regarding the value of  $c_{K_{1,t,t}}$ . In [14], the authors observe that every  $K_{1,t,t}$ -free graph satisfies  $e(G[N(v)]) = O_t(\Delta^{2-1/t})$  to deduce their bound  $c_{K_{1,t,t}} \leq t$  from Theorem 10. In contrast, the first two authors of this manuscript along with Bernshteyn prove the bound  $c_{K_{1,t,t}} \leq 4$  in [6] (see Theorem 19) through a stronger observation: every  $K_{1,t,t}$ -free graph satisfies  $e(G[S]) = O_t(|S|^{2-1/t})$  for any  $v \in V(G)$  and  $S \subseteq N(v)$ . This allows them to take advantage of the local sparsity through the iterations of the nibble method and improve the bound of  $c_{K_{1,t,t}}$  for  $t > 4$ .

We remark that we employ a similar observation in our proof: every  $(k, K_{s,t})$ -locally-sparse graph satisfies  $G[S]$  is  $(k, K_{s,t})$ -sparse for any  $v \in V(G)$  and  $S \subseteq N(v)$ . In fact, our approach is inspired by that of [6] and we apply a number of their arguments in a black box manner (we discuss this further in §3).

We conclude this section with a brief discussion of the bipartiteness of  $F$ . The version of the nibble method that we employ can be adapted to any  $F$ -free graph where  $F$  satisfies  $ex(n, F) = o(n^2)$ . Indeed, by the celebrated Erdős–Stone Theorem [18], bipartite graphs are the only graphs for which this is true. It is, therefore, unclear how to extend our result to arbitrary graphs  $F$ .<sup>3</sup>

<sup>3</sup>In follow-up work, the second named author of this manuscript considers the case  $F = K_r$  for  $r \geq 3$ ; the results obtained are far weaker than those in this paper [15].

### 3 Proof overview

Let  $G$  be a graph and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  satisfying the conditions of Theorem 13. In order to find a proper  $\mathcal{H}$ -coloring of  $G$ , we use a variant of the well-known “Rödl Nibble method”, in which we randomly color a small fraction of vertices and repeat the procedure on the uncolored vertices. In particular, we use a coloring procedure developed by Kim [27], based on a technique of Rödl in [37]. A version of Kim’s procedure was further developed by Molloy and Reed [31], which they call the Wasteful Coloring Procedure. We specifically employ the version generalized to correspondence coloring by the first two authors and Bernshteyn in [6] (see §4.1 for a detailed description). At each step of the procedure, we start with a graph  $G$  and a correspondence cover  $\mathcal{H}$  of  $G$ , and randomly construct a partial  $\mathcal{H}$ -coloring  $\varphi$  of  $G$  and lists  $L'(v) \subseteq L(v)$  satisfying certain properties. In particular, let  $G_\varphi$  be the graph induced by the uncolored vertices under  $\varphi$ . For a vertex  $v \in V(G_\varphi)$  and a color  $c \in L(v)$ , it is possible  $\varphi$  cannot be extended by coloring  $v$  with  $c$ , as there may be a vertex  $u \in N_G(v)$  such that  $c \sim \varphi(u)$ . We therefore define  $L_\varphi(v)$  to be the set of *available colors* for  $v$ , i.e.,

$$L_\varphi(v) := \{c \in L(v) : N_H(c) \cap \text{im}(\varphi) = \emptyset\}.$$

It seems natural that the lists  $L'(\cdot)$  produced by the Wasteful Coloring Procedure should satisfy  $L'(v) = L_\varphi(v)$ ; however, in actuality  $L'(v)$  is potentially a strict subset of  $L_\varphi(v)$  (hence the “wasteful” in the name; see the discussion in [31, Chapter 12.2.1] for the utility of such “wastefulness”).

Let us define a cover  $\mathcal{H}'$  for the graph  $G_\varphi$  by letting

$$H' := H \left[ \bigcup_{v \in V(G_\varphi)} L'(v) \right].$$

Let  $\ell := \min |L(v)|$ ,  $\ell' := \min |L'(v)|$ ,  $d := \Delta(H)$ , and  $d' := \Delta(H')$ . We say  $\varphi$  is “good” if the ratio  $d'/\ell'$  is considerably smaller than  $d/\ell$ . If indeed the output of the Wasteful Coloring Procedure produces a “good” coloring, then we may repeatedly apply the procedure until we are left with a cover  $(\tilde{L}, \tilde{H})$  of the uncolored vertices such that  $\min |\tilde{L}(v)| \geq 8\Delta(\tilde{H})$ . At this point, we can complete the coloring by applying the following proposition:

**Proposition 20.** *Let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$ . If there is an integer  $\ell$  such that  $|L(v)| \geq \ell$  and  $\deg_H(c) \leq \ell/8$  for each  $v \in V(G)$  and  $c \in V(H)$ , then  $G$  admits a proper  $\mathcal{H}$ -coloring.*

This proposition is a rather standard application of the Lovász Local Lemma (Theorem 23 below), and one can find the first use of this proposition (for list colorings) in [34], and its correspondence coloring version in [7, Appendix].

Therefore, we spend considerable effort showing that the output of the Wasteful Coloring Procedure is “good”. The first step in showing this is to compute bounds on the

expected values of the list sizes  $|L'(v)|$  and degrees of colors  $\deg_{H'}(c)$ . In fact, we can show that

$$\frac{\mathbb{E}[\deg_{H'}(c)]}{\mathbb{E}[|L'(v)|]} \approx \text{uncolor} \frac{d}{\ell},$$

where  $\text{uncolor} < 1$  is a certain quantity defined precisely in §4.1. With this in hand, the next step would be to show that the above expression holds with high probability as we may then apply the Lovász Local Lemma to conclude there exists a “good” partial coloring  $\varphi$ . Unfortunately, the random variable  $\deg_{H'}(c)$  may not be concentrated about its expected value if the neighbors of  $c$  have many common neighbors. In particular, the events  $\{c' \in V(H')\}$  for  $c' \in N_H(c)$  may be strongly correlated, increasing the variance of  $\deg_{H'}(c)$ . An idea of Jamall [22], which was developed further by Pettie and Su [33], Alon and Assadi [2], and the first two authors and Bernshteyn [6] allows one to overcome this issue. Specifically, instead of considering the *maximum color-degree* we consider the *average color-degree* for each  $v \in V(G)$ :

$$\overline{\deg}_H(v) := \frac{1}{|L(v)|} \sum_{c \in L(v)} \deg_H(c).$$

Roughly speaking, the average color-degree is easier to concentrate than the maximum color-degree because the probability that the event described above occurs for *all*  $c$  in a list is relatively small. For a further discussion of why considering the average color-degree is useful, see the discussion by Pettie and Su in [33, §1]. Once we show  $\deg_{H'}(v)$  is concentrated, we remove any color  $c \in L'(v)$  satisfying  $\deg_{H'}(c) \geq 2\overline{\deg}_{H'}(v)$ . In this way, we have an upper bound on  $\Delta(H')$  as desired. We remark that at this step we remove possibly up to half the vertices in  $L'(v)$ , which is why we require  $C > 4$  in Theorem 13.

A curious aspect of the analysis of this coloring procedure is that the local sparsity is only employed in the computation of  $\mathbb{E}[\overline{\deg}_{H'}(v)]$ . In particular, as  $H$  is  $(k, K_{s,t})$ -locally-sparse, the subgraph induced by  $N_H(c)$  contains “few” edges. As a result, the event  $\{c \in V(H')\}$  and the random variable  $\deg_{H'}(c)$  are “nearly independent.” This is somewhat surprising as for most applications of the Wasteful Coloring Procedure, the concentration is considered to be the “harder” part of the analysis.

In the proof of Theorem 7, Davies, Kang, Pirot, and Sereni obtained an upper bound on the number of edges in  $G[N_G(v)]$  for a  $(k, P_t)$ -locally-sparse graph  $G$  and a vertex  $v \in V(G)$ . In fact, the bound follows from the following more general statement:

**Fact 21.** *Let  $G$  be a  $(k, F)$ -locally-sparse graph. For any  $v \in V(G)$ , we have*

$$e(G[N(v)]) \leq k + \text{ex}(\deg_G(v), F),$$

where  $\text{ex}(n, F)$  is the maximum number of edges in an  $n$ -vertex  $F$ -free graph.

The proof of the above statement follows from the fact that one can make the subgraph  $G[N(v)]$   $F$ -free by removing at most  $k$  edges. Note that the bound provided by Fact 21 is only useful when  $k \ll \deg(v)^2$ . This contributes to the necessary upper bound on  $k$  in Theorems 7 and 10. We remark that Fact 21 in conjunction with Theorem 4 is sufficient

to prove Theorem 13 when  $C = 4 + \varepsilon$  and  $k \leq d^{\varepsilon/200}$ . The following result allows us to provide a sufficient bound on  $e(G[N(v)])$  for larger values of  $k$  when  $C = 8$ .

**Proposition 22.** *Let  $G$  be an  $n$ -vertex  $(k, K_{s,t})$ -sparse graph. For  $k^* = \frac{(s+t)^2 n^{s+t-1}}{2 s^s t^t}$ , we have*

$$|E(G)| \leq \begin{cases} 4 n^{2-1/(st)} & \text{if } k \leq k^*; \\ 2 s^{1/t} t^{1/s} k^{1/st} n^{2-1/s-1/t} & \text{if } k > k^*. \end{cases}$$

We remark that we always apply the above proposition with  $k \leq k^*$ , however, we include the case  $k > k^*$  for completeness. Alon considered the case that  $s, t = \Theta(1)$  in order to prove a supersaturation-type result. In fact, for this range of  $s$  and  $t$ , the contrapositive of [1, Corollary 2.1] implies Proposition 22 for  $k = \Theta(n^{s+t})$ . Our proof follows a similar strategy to that of [1, Corollary 2.1] (see §B for the details).

Once we are able to prove concentration, it remains to show that a “good” partial coloring  $\varphi$  exists. To do so, we employ the symmetric version of the Lovász Local Lemma.

**Theorem 23** (Lovász Local Lemma; [31, §4]). *Let  $A_1, A_2, \dots, A_n$  be events in a probability space. Suppose there exists  $p \in [0, 1)$  such that for all  $1 \leq i \leq n$  we have  $\mathbb{P}[A_i] \leq p$ . Further suppose that each  $A_i$  is mutually independent from all but at most  $\mathfrak{d}$  other events  $A_j$ ,  $j \neq i$  for some  $\mathfrak{d} \in \mathbb{N}$ . If  $4p\mathfrak{d} \leq 1$ , then with positive probability none of the events  $A_1, \dots, A_n$  occur.*

In their proof of Theorem 19, the first two authors and Bernshteyn provided a framework for applying this version of the Wasteful Coloring Procedure, i.e., the version for correspondence coloring in the color-degree setting. Many of the arguments, particularly certain computational ones, can therefore be applied in a “black box” manner. The main differences in the statements of Theorems 19 and 13 are the structural constraints on the cover graph  $H$ , and the sizes of the lists defined by  $L$ . It turns out these assumptions appear in two key steps of the proof: first, the computation of  $\mathbb{E}[\overline{\deg}_{\mathcal{H}'}(v)]$ ; second, the inductive application of the procedure, i.e., ensuring the list sizes are not too small when we reach the stage where we may apply Proposition 20. The heart of our argument lies in the proofs of these two steps, which appear in §4.3 and §5, respectively.

We conclude this subsection with a discussion on the bound  $k \leq d^{(s+t)/10}$  in Theorem 13. When applying Proposition 22, we require  $k$  to be asymptotically smaller than  $d^{s+t}$ . If not, it could be the case that  $N_H(c)$  induces a very dense graph in  $H$ . The events  $\{c' \in V(H')\}$ ,  $\{c \in V(H')\}$  for  $c' \in N_H(c)$  would then be strongly correlated. As a result, the bound on  $\mathbb{E}[\overline{\deg}_{H'}(v)]$  would not be sufficient to apply the nibble method with the desired list size  $|L(v)|$ . We note that we require  $k \ll d_i^{s+t}$ , where  $d_i$  is the relevant parameter during the  $i$ -th application of the method. For  $C > 4$ , we are able to compute a lower bound  $d_i \geq d^\gamma$  for some appropriate  $\gamma := \gamma(C) > 0$ . As a result, it is enough to have  $k \leq d^{r(s+t)}$  for some positive  $r := r(\gamma) < 1$ . We set  $r = 1/10$  and  $C = 8$  purely for computational ease. We remark that we may decrease  $C$  by decreasing  $r$  to get a better bound on the chromatic number, however, we must have  $C > 4$ .

## 4 The Wasteful Coloring Procedure

This section will be split into three subsections. In the first, we provide a description of our coloring procedure and state a key lemma at the heart of our proof of Theorem 13. In the second, we prove this result modulo a technical lemma, which we prove in the final subsection.

### 4.1 Algorithm overview

In this section, we formally describe our coloring procedure and introduce a key lemma at the heart of our proof of Theorem 13. Roughly speaking, the lemma below will allow us to, in §5, inductively color our graph until a point at which Proposition 20 can be applied. As mentioned earlier, we will utilize the Wasteful Coloring Procedure (described formally in Algorithm 1) in order to prove the lemma. We first introduce some notation.

Recall from §3 for a fixed vertex  $v \in V(G)$ , its average color-degree is

$$\overline{\deg_{\mathcal{H}}}(v) = \frac{1}{|L(v)|} \sum_{c \in L(v)} \deg_H(c),$$

and the list of available colors for  $v$  with respect to a partial coloring  $\varphi$  is as follows:

$$L_{\varphi}(v) = \{c \in L(v) : N_H(c) \cap \text{im}(\varphi) = \emptyset\},$$

where  $\text{im}(\varphi)$  is the image of  $\varphi$ , i.e., the colors assigned to colored vertices in  $G$ . The main component of the proof of Theorem 13 is the following lemma, which shows that under certain conditions on  $G$  and its correspondence cover, there exists a partial coloring such that the uncolored graph has desirable properties.

**Lemma 24.** *There are  $\tilde{d} \in \mathbb{N}$ ,  $\tilde{\alpha} > 0$  such that the following holds. Suppose  $\eta \in \mathbb{R}$ ,  $d, \ell, s, t \in \mathbb{N}$ , and  $k \in \mathbb{R}$  satisfy:*

(L1)  *$d$  is sufficiently large:  $d \geq \tilde{d}$ ,*

(L2)  *$k$  is not too large:  $1/2 \leq k \leq d^{(s+t)/5}$ ,*

(L3)  *$\ell$  is bounded below and above in terms of  $d$ :  $4\eta d < \ell < 100d$ ,*

(L4)  *$s$  and  $t$  are bounded in terms of  $d$ :  $1 \leq t \leq s$  and  $st \leq \frac{\tilde{\alpha} \log d}{\log \log d}$ , and*

(L5)  *$\eta$  is bounded below and above in terms of  $d, k, s$ , and  $t$ :  $\frac{1}{\log^5(d)} < \eta < \frac{1}{\log(dk^{-1/(s+t)})}$ .*

*Let  $G$  be a graph with a correspondence cover  $\mathcal{H} = (L, H)$  such that for some  $\beta$  satisfying  $d^{-1/(200st)} \leq \beta \leq 1/10$ ,*

(L6)  *$H$  is  $(k, K_{s,t})$ -locally-sparse,*

(L7)  $\Delta(H) \leq 2d$ ,

(L8) the list sizes are roughly between  $\ell/2$  and  $\ell$ :  $(1 - \beta)\ell/2 \leq |L(v)| \leq (1 + \beta)\ell$  for all  $v \in V(G)$ ,

(L9) average color-degrees are smaller for vertices with smaller lists of colors:

$$\overline{\deg}_{\mathcal{H}}(v) \leq \left(2 - (1 - \beta)\frac{\ell}{|L(v)|}\right) d \quad \text{for all } v \in V(G).$$

Then there exists a proper partial  $\mathcal{H}$ -coloring  $\varphi$  of  $G$  and an assignment of subsets  $L'(v) \subseteq L_{\varphi}(v)$  to each  $v \in V(G) \setminus \text{dom}(\varphi)$  with the following properties. Let

$$G' := G[V(G) \setminus \text{dom}(\varphi)] \quad \text{and} \quad H' := H\left[\bigcup_{v \in V(G')} L'(v)\right].$$

Define the following quantities:

$$\begin{aligned} \text{keep} &:= \left(1 - \frac{\eta}{\ell}\right)^{2d}, & \ell' &:= \text{keep } \ell, & \beta' &:= (1 + 36\eta)\beta. \\ \text{uncolor} &:= \left(1 - \frac{\eta}{\ell}\right)^{\text{keep } \ell/2}, & d' &:= \text{keep uncolor } d, \end{aligned}$$

Let  $\mathcal{H}' := (L', H')$ , so  $\mathcal{H}'$  is a correspondence cover of  $G'$ . Then for all  $v \in V(G')$ :

- (i)  $|L'(v)| \leq (1 + \beta')\ell'$ ,
- (ii)  $|L'(v)| \geq (1 - \beta')\ell'/2$ ,
- (iii)  $\Delta(H') \leq 2d'$ ,
- (iv)  $\overline{\deg}_{\mathcal{H}'}(v) \leq \left(2 - (1 - \beta')\frac{\ell'}{|L'(v)|}\right) d'$ .
- (v)  $H'$  is  $(k, K_{s,t})$ -locally-sparse.

Note that condition (v) holds automatically as  $H' \subseteq H$ . Also note that conditions (i)–(iv) are similar to the conditions (L6)–(L9), except that the former uses  $\beta', \eta', \ell'$ . This will help us to apply Lemma 24 iteratively in §5 to prove Theorem 13. For the rest of this section, we define **keep**, **uncolor**,  $d'$ ,  $\ell'$ ,  $\beta'$  as above.

Let us now describe the Wasteful Coloring Procedure as laid out by the first two authors and Bernshteyn in [6] for correspondence colorings (the color-degree version for list coloring appeared in [2]).

---

**Algorithm 1** Wasteful Coloring Procedure

---

**Input:** A graph  $G$  with a correspondence cover  $\mathcal{H} = (L, H)$  and parameters  $\eta \in [0, 1]$  and  $d, \ell > 0$ .

**Output:** A proper partial  $\mathcal{H}$ -coloring  $\varphi$  and subsets  $L'(v) \subseteq L_{\varphi}(v)$  for all  $v \in V(G)$ .



1. Sample  $A \subseteq V(H)$  as follows: for each  $c \in V(H)$ , include  $c \in A$  independently with probability  $\eta/\ell$ . Call  $A$  the set of **activated** colors, and let  $A(v) := L(v) \cap A$  for each  $v \in V(G)$ .
2. Let  $\{\text{eq}(c) : c \in V(H)\}$  be a family of independent random variables with distribution

$$\text{eq}(c) \sim \text{Bernoulli} \left( \frac{\text{keep}}{(1 - \eta/\ell)^{\deg_H(c)}} \right).$$

(We discuss why this is well defined below). Call  $\text{eq}(c)$  the **equalizing coin flip** for  $c$ .

3. Sample  $K \subseteq V(H)$  as follows: for each  $c \in V(H)$ , include  $c \in K$  if  $\text{eq}(c) = 1$  and  $N_H(c) \cap A = \emptyset$ . Call  $K$  the set of **kept** colors, and  $V(H) \setminus K$  the **removed** colors. For each  $v \in V(G)$ , let  $K(v) := L(v) \cap K$ .
4. Construct  $\varphi : V(G) \dashrightarrow V(H)$  as follows: if  $A(v) \cap K(v) \neq \emptyset$ , set  $\varphi(v)$  to any color in  $A(v) \cap K(v)$ . Otherwise set  $\varphi(v) = \text{blank}$ .
5. Call  $v \in V(G)$  **uncolored** if  $\varphi(v) = \text{blank}$ , and define

$$U := \{c \in V(H) : \varphi(L^{-1}(c)) = \text{blank}\}.$$

(Recall that  $L^{-1}(c)$  denotes the underlying vertex of  $c$  in  $G$ .)

6. For each vertex  $v \in V(G)$ , let

$$L'(v) := \{c \in K(v) : |N_H(c) \cap K \cap U| \leq 2d'\}.$$

Note that if  $G$  and  $H$  satisfy assumption (L7) of Lemma 24, i.e.  $\deg_H(c) \leq 2d$ , then

$$0 \leq \frac{\text{keep}}{(1 - \eta/\ell)^{\deg_H(c)}} = \left(1 - \frac{\eta}{\ell}\right)^{2d - \deg_H(c)} \leq 1$$

and hence the equalizing coin flips are well defined. Furthermore, if we assume that the assumptions of Lemma 24 hold on the input graph of the Wasteful Coloring Procedure, then **keep** is precisely the probability that a color  $c \in V(H)$  is kept, and **uncolor** is roughly an upper bound on the probability that a vertex  $v \in V(G)$  is uncolored.

## 4.2 Proof of Lemma 24

In this section, we prove Lemma 24 under the assumption that a subsequently introduced lemma, Lemma 27, is true. Indeed, Lemma 27 is the key lemma in our argument; it is the part of the proof that relies on the local sparsity condition; and the bulk of the work in proving Lemma 24 comes from proving Lemma 27, which we will do in §4.3.

The rest of the proof of Lemma 24 follows the strategy employed by the first two authors and Bernshteyn to prove [6, Lemma 3.1], which is similar to Lemma 24. Indeed, the assumptions of [6, Lemma 3.1] and Lemma 24 differ in only four ways:

1. the assumption that  $H$  is  $K_{1,s,t}$ -free is replaced by the weaker assumption that  $H$  is  $(k, K_{s,t})$ -locally-sparse,
2. the assumption that  $\eta < \frac{1}{\log d}$  is replaced by the assumption that  $\eta < \frac{1}{\log(dk^{-1/(s+t)})}$ ,
3. the bounds  $s \leq d^{1/10}$  and  $t \leq \frac{\tilde{\alpha} \log d}{\log \log d}$  are replaced by  $st \leq \frac{\tilde{\alpha} \log d}{\log \log d}$ , and
4. the assumption  $d^{-1/(200t)} \leq \beta \leq 1/10$  is replaced by  $d^{-1/(200st)} \leq \beta \leq 1/10$ .

However, several of the lemmas used in the proof of [6, Lemma 3.1] do not rely on either of these assumptions, and thus the same proofs can be used for our lemmas here. Furthermore, as our upper bound on  $st$  matches that of  $t$  in [6, Lemma 3.1], the proofs requiring the lower bound on  $\beta$  follow by replacing  $t$  with  $st$  at the relevant steps. This essentially allows us to use these proofs in a “black box” manner towards proving Lemma 24. We therefore follow the strategy of the proof of [6, Lemma 3.1], and indeed, this section follows similar notation to that of [6, §4].

Assume parameters  $\eta, d, \ell, s, t, k, \beta$ , and a graph  $G$  with correspondence cover  $\mathcal{H} = (L, H)$  satisfy the assumptions of Lemma 24. We may assume that  $\Delta(G) \leq 2(1+\beta)\ell d$  by removing the edges of  $G$  whose corresponding matchings in  $H$  are empty. Let the quantities **keep**, **uncolor**,  $d', \ell'$ , and  $\beta'$  be defined as in the statement of Lemma 24. Suppose we have carried out the Wasteful Coloring Procedure with these  $G$  and  $\mathcal{H}$ . As in the statement of Lemma 24, we let

$$G' := G[V(G) \setminus \text{dom}(\varphi)], \quad H' := H\left[\bigcup_{v \in V(G')} L'(v)\right], \quad \text{and} \quad \mathcal{H}' := (L', H').$$

For each  $v \in V(G)$ , we define the following quantities:

$$\ell(v) := |L(v)| \quad \text{and} \quad \overline{\deg}(v) := \overline{\deg}_{\mathcal{H}}(v),$$

as well as the following random variables:

$$k(v) := |K(v)|, \quad \ell'(v) := |L'(v)|, \quad \text{and} \quad \bar{d}(v) := \frac{1}{\ell'(v)} \sum_{c \in L'(v)} |N_H(c) \cap V(H')|.$$

By definition, if  $v \in V(G')$ , then  $\bar{d}(v) = \overline{\deg}_{\mathcal{H}'}(v)$ . Our goal is to verify that statements (i)–(iv) in Lemma 24 hold for every  $v \in V(G')$  with positive probability. We follow an idea of Pettie and Su [33] (which was also used in [2, 6]) by defining the following auxiliary quantities:

$$\begin{aligned} \lambda(v) &:= \frac{\ell(v)}{\ell}, & \lambda'(v) &:= \frac{\ell'(v)}{\ell'}, \\ \delta(v) &:= \lambda(v) \overline{\deg}_{\mathcal{H}}(v) + (1 - \lambda(v)) 2d, & \delta'(v) &:= \lambda'(v) \bar{d}(v) + (1 - \lambda'(v)) 2d'. \end{aligned}$$

Note that, by (L8), we have  $(1 - \beta)/2 \leq \lambda(v) \leq 1 + \beta$ . When  $\lambda(v) \leq 1$ , we can think of  $\delta(v)$  as what the average color-degree of  $v$  would become if we added  $\ell - \ell(v)$  colors

of degree  $2d$  to  $L(v)$ . (We remark that in both [33] and [2], the value  $\lambda(v)$  is artificially capped at 1. However, as noted by the first two authors and Bernshteyn in [6], it turns out that there is no harm in allowing  $\lambda(v)$  to exceed 1, which moreover makes the analysis simpler.) The upper bound on  $\overline{\deg}(v)$  given by (L9) implies that

$$\delta(v) \leq (1 + \beta)d. \quad (4.1)$$

As demonstrated in [6, Lemma 4.1], an upper bound on  $\delta'(v)$  suffices to derive statements (ii)–(iv) in Lemma 24:

**Lemma 25.** *If  $\delta'(v) \leq (1 + \beta')d'$ , then conditions (ii)–(iv) of Lemma 24 are satisfied.*

*Proof.* The proof is the same as the proof of [6, Lemma 4.1].  $\square$

Therefore, to prove Lemma 24, it suffices to show that, with positive probability, the outcome of the Wasteful Coloring Procedure satisfies  $\delta'(v) \leq (1 + \beta')d'$  and  $\ell'(v) \leq (1 + \beta')\ell'$  for all  $v \in V(G)$ . We shall now prove some intermediate results. Before we do so, consider the following inequality, which will be useful for proving certain bounds and follows for small enough  $\tilde{\alpha}$ :

$$\eta\beta \geq d^{-1/(200st)} / \log^5 d \geq d^{-1/(100st)}. \quad (4.2)$$

In the next lemma, we show that  $k(v)$  is concentrated around its expected value, which we then use to show condition (i) in Lemma 24 is satisfied with high probability. The proof follows by a rather standard application of Talagrand’s inequality (a powerful concentration tool). We omit the proof here as it is identical to the one of [6, Lemma 4.2].

**Lemma 26.**  $\mathbb{P}[|k(v) - \text{keep } \ell(v)| \geq \eta\beta \text{ keep } \ell(v)] \leq \exp(-d^{1/10})$ .

*Proof.* The proof is the same as the proof of [6, Lemma 4.2].  $\square$

Since  $\ell'(v) \leq k(v)$ , we have the following with probability at least  $1 - \exp(-d^{1/10})$ :

$$\begin{aligned} \ell'(v) &\leq (1 + \eta\beta) \text{keep } \ell(v) \\ &\leq (1 + \eta\beta)(1 + \beta) \text{keep } \ell \\ &\leq (1 + \beta') \ell'. \end{aligned}$$

This implies that condition (i) is met with probability at least  $1 - \exp(-d^{1/10})$ .

In order to analyze the average color-degrees in the correspondence cover  $\mathcal{H}'$ , we define:

$$\begin{aligned} E_K(v) &:= \{cc' \in E(H) : c \in L(v), c' \in N_H(c), \text{ and } c, c' \in K\}, \\ E_U(v) &:= \{cc' \in E(H) : c \in L(v), c' \in N_H(c), \text{ and } c' \in U\}, \\ d_{K \cap U}(v) &:= \frac{|E_K(v) \cap E_U(v)|}{k(v)}. \end{aligned}$$

Note that  $d_{K \cap U}(v)$  is what the average color-degree of  $v$  would be if instead of removing colors with too many neighbors on step 6 of the Wasteful Coloring Procedure, we had just set  $L'(v) = K(v)$ . The only place in our proof that relies on  $H$  being  $(k, K_{s,t})$ -locally-sparse is the following lemma, which gives a bound on the expected value of  $|E_K(v) \cap E_U(v)|$ . As mentioned in §3, this lemma is one of the major differences between our proof and that of [6].

**Lemma 27.**  $\mathbb{E}[|E_K(v) \cap E_U(v)|] \leq \text{keep}^2 \text{uncolor} \ell(v) \overline{\deg_{\mathcal{H}}}(v)(1 + 6\eta\beta)$ .

*Proof.* See §4.3. □

The next lemma establishes that with high probability the quantity  $|E_K(v) \cap E_U(v)|$  is not too large. The proof is identical to that of [6, Lemma 4.4], and thus we omit the technical details here. However, we describe the main idea of the proof below for the unfamiliar reader.

The simplest approach to concentrate  $|E_K(v) \cap E_U(v)|$  is to use a suitable version of Chernoff's inequality (see [2, Claims A.6, A.7], [33, Lemmas 6, 7]). This approach fails for correspondence colorings as in this setting two colors from the same list can be connected by a path, and thus events defined on those colors are no longer independent. To deal with this, one can apply Talagrand's inequality, a powerful concentration tool for random variables satisfying certain Lipschitz-like constraints. Unfortunately, the Lipschitz parameter of the random variable  $|E_K(v) \cap E_U(v)|$  is too large to apply Talagrand's inequality directly. Instead, the authors show that with sufficiently high probability  $|E_U(v)|$  is not much larger than its expected value, and  $|E_U(v) \setminus E_K(v)|$  is not much smaller than its expected value. The identity  $|E_K(v) \cap E_U(v)| = |E_U(v)| - |E_U(v) \setminus E_K(v)|$  then implies the desired concentration bound.

While considering the random variable  $|E_U(v)|$ , it turns out that the Lipschitz parameter is still too large to apply Talagrand's inequality. However, the events that drive up the value of this parameter occur with very low probability. Concentration can therefore be achieved by employing a variation of Talagrand's Inequality, which can handle such exceptional events, developed by Bruhn and Joos [11, Theorem 3.1].

For the quantity  $|E_U(v) \setminus E_K(v)|$ , we cannot apply this result of Bruhn and Joos directly, i.e., the Lipschitz parameter is still too large. To circumvent this problem, the authors employ a *random partitioning technique* developed in their earlier work [5]. The idea is to find a partition  $\{X_1, \dots, X_\tau\}$  of the edge set  $\cup_{c \in L(v)} E_H(c)$  such that  $|(E_U(v) \setminus E_K(v)) \cap X_i|$  has a sufficiently small Lipschitz parameter for each  $i$ . It turns out a random partition satisfies this condition with high probability for an appropriate choice of  $\tau$ . The proof now follows by applying the exceptional version of Talagrand's inequality to each part separately and then taking a union bound.

The result of the lemma is stated below, where  $d_{\max} := \max\{d^{7/8}, \Delta(H)\}$ .

**Lemma 28.**  $\mathbb{P}[|E_K(v) \cap E_U(v)| \geq \text{keep}^2 \text{uncolor} \ell(v)(\overline{\deg_{\mathcal{H}}}(v) + 8\eta\beta d_{\max})] \leq d^{-100}$ .

*Proof.* The proof is the same as the proof of [6, Lemma 4.4]. □

The next two lemmas are purely computational, using the results of Lemmas 26 and 28. As the calculations are identical to those in [6, Section 4], we omit them here. First, we show that  $d_{K \cap U}(v)$  is not too large with high probability.

**Lemma 29.**  $\mathbb{P}[d_{K \cap U}(v) > \text{keep uncolor } \overline{\deg_{\mathcal{H}}}(v) + 15\eta\beta \text{ keep uncolor } d_{\max}] \leq d^{-75}.$

*Proof.* The proof is the same as the proof of [6, Lemma 4.5].  $\square$

We can now combine the results of this section to prove the desired bound on  $\delta'(v)$ . Again, the proof is purely computational, and so we refer to [6, Lemma 4.6].

**Lemma 30.**  $\mathbb{P}[\delta'(v) \leq (1 + \beta')d'] \geq 1 - d^{-50}.$

*Proof.* The proof is the same as the proof of [6, Lemma 4.6].  $\square$

We are now ready to finish the proof of Lemma 24 (modulo Lemma 27).

*Proof of Lemma 24.* This proof is the same as the proof of [6, Lemma 3.1], but is included here for completeness. We perform the Wasteful Coloring Procedure on  $G$  and  $\mathcal{H}$  and define the following random events for each  $v \in V(G)$ :

1.  $A_v := \{\ell'(v) \leq (1 + \beta')\},$
2.  $B_v := \{\delta'(v) \geq (1 + \beta')d'\}.$

We now use the Lovász Local Lemma (Theorem 23). By Lemmas 26 and 30, we have:

$$\mathbb{P}[A_v] \leq \exp(-d^{1/10}) \leq d^{-50}, \quad \mathbb{P}[B_v] \leq d^{-50}.$$

Let  $p := d^{-50}$ . Note that the events  $A_v, B_v$  are mutually independent from the events of the form  $A_u, B_u$ , where  $u \in V(G)$  is at distance more than 4 from  $v$ . Since  $\Delta(G) \leq 2(1 + \beta)\ell d$ , there are at most  $2(2(1 + \beta)\ell d)^4 \leq d^{10}$  events corresponding to the vertices at distance at most 4 from  $v$ . So we let  $\mathfrak{d} := d^{10}$  and observe that  $4p\mathfrak{d} = 4d^{-40} < 1$ . By the Lovász Local Lemma, with positive probability none of the events  $A_v, B_v$  occur. By Lemma 25, this implies that, with positive probability, the output of the Wasteful Coloring Procedure satisfies the conclusion of Lemma 24.  $\square$

### 4.3 Proof of Lemma 27

This is the only part of the proof that requires  $(k, K_{s,t})$ -local-sparsity. We will first provide an upper bound for  $\mathbb{E}[|E_K(v)|]$ .

**Lemma 31.**  $\mathbb{E}[|E_K(v)|] \leq \text{keep}^2 \ell(v) \overline{\deg}(v) (1 + \eta\beta)$

*Proof.* Note that

$$|E_K(v)| = \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \mathbb{1}_{\{c, c' \in K\}}.$$

If  $c \in K$ , then two events occur: no color in  $N_H(c)$  is activated, and  $c$  survives its equalizing coin flip. Since activations and equalizing coin flips occur independently, we have:

$$\begin{aligned}
\mathbb{E}[|E_K(v)|] &= \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \mathbb{P}[c, c' \in K] \\
&= \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \mathbb{P}[\mathbf{eq}(c) = 1] \mathbb{P}[\mathbf{eq}(c') = 1] \left(1 - \frac{\eta}{\ell}\right)^{|N_H(c) \cup N_H(c')|} \\
&= \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \mathbb{P}[\mathbf{eq}(c) = 1] \mathbb{P}[\mathbf{eq}(c') = 1] \left(1 - \frac{\eta}{\ell}\right)^{|N_H(c)| + |N_H(c')| - |N_H(c) \cap N_H(c')|} \\
&= \text{keep}^2 \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|}.
\end{aligned}$$

For  $c \in H$ , we will consider two cases based on the value of  $\deg_H(c)$  (we will abuse notation and say  $c \in \text{C1}$  and  $c \in \text{C2}$  when  $c$  satisfies the conditions to be in Case 1 and Case 2, respectively.)

(Case 1)  $\deg_H(c) < d^{1-1/(5st)}$ . Then

$$|N_H(c) \cap N_H(c')| \leq |N_H(c)| = \deg_H(c) \leq d^{1-1/(5st)}.$$

Thus, using the inequality  $(1-x)^n \geq 1-nx$  for  $n \geq 1$  and  $x \leq 1$ , and the lower bound  $\ell \geq 4\eta d$  from assumption (L3), we have:

$$\begin{aligned}
\sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|} &\leq \sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-d^{1-1/(5st)}} \quad (\text{as } 1 - \eta/\ell < 1) \\
&\leq \deg_H(c) \left(1 - d^{1-1/(5st)} \frac{\eta}{\ell}\right)^{-1} \\
&\leq \deg_H(c) \left(1 - \frac{d^{-1/(5st)}}{4}\right)^{-1} \\
&\leq \deg_H(c) (1 + d^{-1/(5st)}), \tag{4.3}
\end{aligned}$$

where the last inequality holds since  $d^{-1/(5st)} \leq 1$ .

(Case 2)  $\deg_H(c) \geq d^{1-1/(5st)}$ . We note that this is the only place where we use  $(k, K_{s,t})$ -local-sparsity. Define the following sets:

$$\begin{aligned}
\text{Bad}(c) &:= \{c' \in N_H(c) : |N_H(c') \cap N_H(c)| \geq \deg_H(c)^{1-1/(5st)}\}, \\
\text{Good}(c) &:= N_H(c) \setminus \text{Bad}(c).
\end{aligned}$$

Then:

$$\sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|}$$

$$\begin{aligned}
&= \sum_{c' \in \text{Bad}(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|} + \sum_{c' \in \text{Good}(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|} \\
&\leq \sum_{c' \in \text{Bad}(c)} \left(1 - \frac{\eta}{\ell}\right)^{-\deg_H(c)} + \sum_{c' \in \text{Good}(c)} \left(1 - \frac{\eta}{\ell}\right)^{-\deg_H(c)^{1-1/5st}} \\
&\leq |\text{Bad}(c)| \left(1 - \frac{\eta}{\ell}\right)^{-2d} + (\deg_H(c) - |\text{Bad}(c)|) \left(1 - \frac{\eta}{\ell}\right)^{-\deg_H(c)^{1-1/5st}}. \quad (4.4)
\end{aligned}$$

Since  $\deg_H(c)^{1-1/5st} \leq 2d$ , the last expression is increasing in terms of  $|\text{Bad}(c)|$ . The following claim provides an upper bound on the size of  $\text{Bad}(c)$ .

**Claim 32.**  $|\text{Bad}(c)| \leq \deg_H(c)^{1-1/(5st)}$ .

*Proof.* By the Handshaking Lemma,

$$|\text{Bad}(c)| \leq \frac{2|E(H[N_H(c)])|}{\deg_H(c)^{1-1/(5st)}}.$$

Since  $H$  is  $(k, K_{s,t})$ -locally-sparse,  $H[N_H(c)]$  is  $(k, K_{s,t})$ -sparse, and so by Proposition 22 it follows that if for  $k^* := \frac{(s+t)^2 \deg_H(c)^{s+t-1}}{2s^s t^t}$  we have  $k \leq k^*$ , then

$$|E(H[N_H(c)])| \leq 4 \deg_H(c)^{2-1/(st)}.$$

Let us first show  $k \leq k^*$ . As we are in (Case 2), we have  $\deg_H(c) \geq d^{1-1/(5st)}$ . It follows that

$$k^* \geq \frac{(s+t)^2 d^{(1-1/(5st))(s+t-1)}}{2s^s t^t} = \frac{(s+t)^2 d^{s+t-1-1/(5t)-1/(5s)+1/(5st)}}{2s^s t^t}$$

Note that

$$\frac{s^s t^t}{(s+t)^2} < (st)^{st} \leq \left( \frac{\tilde{\alpha} \log d}{\log \log d} \right)^{\frac{\tilde{\alpha} \log d}{\log \log d}} < \log d^{\frac{\tilde{\alpha} \log d}{\log \log d}} = d^{\tilde{\alpha}},$$

for  $\tilde{\alpha}$  sufficiently small, where the second inequality holds by assumption (L4). Thus,

$$\frac{(s+t)^2}{s^s t^t} > d^{-\tilde{\alpha}}.$$

Additionally, we claim that

$$s+t-1-1/(5t)-1/(5s)+1/(5st) \geq (s+t)/5+2/5, \quad \text{for all } s \geq t \geq 1.$$

One can manually verify the above for  $s=t=1$  and  $s=2, t=1$ . For  $s \geq t \geq 2$ , we have

$$4(s+t)/5+1/(5st) > 3 > 1/(5t)+1/(5s)+1+2/5,$$

which implies the desired inequality.

Thus, for  $d$  large enough and  $\tilde{\alpha}$  sufficiently small, we have  $k^* > d^{(s+t)/5} \geq k$  by assumption (L2). Hence, by Proposition 22 we have

$$\begin{aligned} |\text{Bad}(c)| &\leq \frac{2|E(H[N_H(c)])|}{\deg_H(c)^{1-1/(5st)}} \leq \frac{2 \cdot 4 \deg_H(c)^{2-1/(st)}}{\deg_H(c)^{1-1/(5st)}} \\ &= 8 \deg_H(c)^{1-4/(5st)} \\ &\leq \deg_H(c)^{1-1/(5st)}, \end{aligned}$$

as desired.  $\square$

With Claim 32 in hand, and as  $(1 - \frac{\eta}{\ell})^{-2d} = \frac{1}{\text{keep}} \leq 2$  due to assumption (L3), we have the following as a result of (4.4):

$$\begin{aligned} &\sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|} \\ &\leq 2 \deg_H(c)^{1-1/(5st)} + (\deg_H(c) - \deg_H(c)^{1-1/(5st)}) \left(1 - \frac{\eta}{\ell}\right)^{-\deg_H(c)^{1-1/5st}} \\ &\leq 2 \deg_H(c)^{1-1/(5st)} + \frac{(\deg_H(c) - \deg_H(c)^{1-1/(5st)})}{(1 - \deg_H(c)^{1-1/5st} \frac{\eta}{\ell})} \\ &= \deg_H(c) \left(2 \deg_H(c)^{-1/(5st)} + \frac{1 - \deg_H(c)^{-1/(5st)}}{(1 - \frac{\eta}{\ell} \deg_H(c)^{1-1/5st})}\right) \\ &\leq \deg_H(c) \left(2 \deg_H(c)^{-1/5st} + \frac{1 - \deg_H(c)^{-1/(5st)}}{(1 - \frac{\eta}{\ell} 2d \deg_H(c)^{-1/5st})}\right) \quad (\text{since } \deg_H(c) \leq 2d) \\ &\leq \deg_H(c) \left(2 \deg_H(c)^{-1/5st} + \frac{1 - \deg_H(c)^{-1/5st}}{(1 - \frac{1}{2} \deg_H(c)^{-1/5st})}\right) \quad (\text{since } 2d \frac{\eta}{\ell} \leq \frac{1}{2}) \\ &\leq \deg_H(c) (\deg_H(c)^{-1/10st} + 1), \end{aligned} \tag{4.5}$$

completing this case.

By (4.2), we have  $\eta\beta \geq d^{-1/(100st)} \geq d^{-1/(10st)}$ . Therefore, putting together the bounds (4.3) from (Case 1) and (4.5) from (Case 2), we have the following:

$$\begin{aligned} \mathbb{E}[|E_K(v)|] &= \text{keep}^2 \sum_{c \in L(v)} \sum_{c' \in N_H(c)} \left(1 - \frac{\eta}{\ell}\right)^{-|N_H(c) \cap N_H(c')|} \\ &\leq \text{keep}^2 \left( \sum_{c \in C1} \deg_H(c) (1 + d^{-1/(5st)}) + \sum_{c \in C2} \deg_H(c) (\deg_H(c)^{-1/10st} + 1) \right) \\ &\leq \text{keep}^2 \left( \sum_{c \in C1} \deg_H(c) (\eta\beta + 1) + \sum_{c \in C2} \deg_H(c) (\eta\beta + 1) \right) \end{aligned}$$



$$\begin{aligned}
&= \text{keep}^2 \sum_{c \in L(v)} \deg_H(c) (\eta\beta + 1) \\
&= \text{keep}^2 \ell(v) \overline{\deg}_{\mathcal{H}}(v) (\eta\beta + 1),
\end{aligned}$$

completing the proof.  $\square$

The following lemma bounds  $\mathbb{E}[|E_K(v) \cap E_U(v)|]$  in terms of  $\mathbb{E}[|E_K(v)|]$ . The proof involves a special case of the FKG inequality, dating back to Harris and Kleitman. As the argument is identical to that of [6, Lemma 5.5], we omit the details here.

**Lemma 33.**  $\mathbb{E}[|E_K(v) \cap E_U(v)|] \leq \text{uncolor} (1 + 4\eta\beta) \mathbb{E}[|E_K(v)|]$ .

*Proof.* The proof is the same as that in [6, Lemma 5.2].  $\square$

Lemmas 31 and 33 together imply that

$$\begin{aligned}
\mathbb{E}[|E_K(v) \cap E_U(v)|] &\leq \text{keep}^2 \text{uncolor} \ell(v) \overline{\deg}_{\mathcal{H}}(v) (1 + 4\eta\beta) (1 + \eta\beta) \\
&\leq \text{keep}^2 \text{uncolor} \ell(v) \overline{\deg}_{\mathcal{H}}(v) (1 + 6\eta\beta),
\end{aligned}$$

which completes the proof of Lemma 27.

## 5 Proof of Theorem 13

To prove Theorem 13, we start by defining several parameters: For a given  $\varepsilon > 0$ ,  $d, k \in \mathbb{R}$ ,  $s, t \in \mathbb{N} \setminus \{0\}$ , and  $i \in \mathbb{N}$ , we define the following:

$$\begin{aligned}
C &:= \begin{cases} 4 + \varepsilon & \text{when } k \leq d^{\varepsilon(s+t)/200}, \\ 8 & \text{else.} \end{cases} \\
\mu &:= \frac{C - \varepsilon}{2} \log \left( 1 + \frac{\varepsilon}{8C} \right) & \eta &:= \mu / \log (dk^{-1/(s+t)}) \\
\ell_0 &:= C d / \log (dk^{-1/(s+t)}) & d_0 &:= d \\
k_i &:= \left( 1 - \frac{\eta}{\ell_i} \right)^{2d_i} & u_i &:= \left( 1 - \frac{\eta}{\ell_i} \right)^{k_i \ell_i / 2} \\
\ell_{i+1} &:= k_i \ell_i & d_{i+1} &:= k_i u_i d_i \\
\beta_0 &:= d_0^{-1/(200st)} & \beta_{i+1} &:= \max \left\{ (1 + 36\eta) \beta_i, d_{i+1}^{-1/(200st)} \right\}.
\end{aligned}$$

We will also assume  $\varepsilon$  is sufficiently small such that  $\mu < 1$ . The reader may be familiar with similar arguments parameterized by  $\text{keep}_i$  and  $\text{uncolor}_i$ . For brevity, we use  $k_i$  and  $u_i$  instead.

**Lemma 34.** *For all  $\varepsilon > 0$  sufficiently small, there exists  $d^*$  such that whenever  $d \geq d^*$  and*

$$\frac{1}{2} < k \leq d^{(s+t)/10},$$

*the following hold:*

(I1) For  $j \in \mathbb{N}$ ,  $d_{j+1}/\ell_{j+1} \leq d_j/\ell_j \leq d_0/\ell_0 = \frac{1}{C} \log(dk^{-1/(s+t)})$ .

(I2) For  $j \in \mathbb{N}$ ,  $\ell_j \geq d(dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)}$ .

(I3) There exists a minimum integer  $i^* \leq \left\lceil \frac{16}{\mu} \log(dk^{-1/(s+t)}) \log \log(dk^{-1/(s+t)}) \right\rceil$  such that  $d_{i^*} \leq \ell_{i^*}/100$ .

*Proof.* We will not explicitly compute the value  $d^*$ , but instead simply state that we let  $d$  be sufficiently large when needed.

By definition,  $d_{j+1}/\ell_{j+1}$  is smaller than  $d_j/\ell_j$  by a factor of  $u_j$  and so (I1) holds. Note the following:

$$0 < 4/(C - 7\varepsilon/8) < 1, \quad \text{and} \quad d^{9/10} \leq dk^{-1/(s+t)} \leq 2d. \quad (5.1)$$

We now use strong induction on  $j \in \mathbb{N}$  to prove (I2). As  $d \rightarrow \infty$ , it follows  $dk^{-1/(s+t)} \rightarrow \infty$ . Thus for  $d$  large enough, we have

$$\frac{\log(dk^{-1/(s+t)})}{C} \leq (dk^{-1/(s+t)})^{4/(C-7\varepsilon/8)},$$

from where it follows that

$$\ell_0 = \frac{Cd}{\log(dk^{-1/(s+t)})} \geq d(dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)}.$$

This completes the proof of the base case. Now assume for some  $j' \in \mathbb{N}$ , (I2) holds for all  $0 \leq j \leq j'$ . We will show  $\ell_{j'+1} \geq d(dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)}$ . To do so, we will compute a sufficient lower bound on  $k_j$  for each  $0 \leq j \leq j'$ , which will allow us to prove the desired bound on  $\ell_{j'+1}$ . We start by showing  $k_j$  is bounded from below by a constant (independent of  $j$ ); we then use this to compute an upper bound on  $u_j$ , which we then use to prove a stronger lower bound on  $k_j$ .

Let  $0 \leq j \leq j'$ . First note

$$\begin{aligned} \frac{\eta}{\ell_j} &= \frac{\mu}{\log(dk^{-1/(s+t)})} \frac{1}{\ell_j} \\ &\leq \frac{\mu}{\log(dk^{-1/(s+t)})} \frac{(dk^{-1/(s+t)})^{4/(C-7\varepsilon/8)}}{d} \\ &\leq \frac{10}{9 \log(d)} \frac{(2d)^{4/(4+\varepsilon/8)}}{d} \\ &\leq \frac{1}{d^{\varepsilon/50}}, \end{aligned} \quad (5.2)$$

where the inequalities follow by (5.1), the bound on  $\ell_j$  from the induction hypothesis, and for  $d$  large enough. Now recall for any  $\varepsilon > 0$ , there exists  $x_0 > 0$  such that  $1 - x \geq$

$\exp\left(-\frac{x}{1-\varepsilon}\right)$  whenever  $0 < x < x_0$ . Thus for  $d$  sufficiently large, by (5.2) it follows  $\eta/\ell_j$  is sufficiently small for the following to hold:

$$k_j = \left(1 - \frac{\eta}{\ell_j}\right)^{2d_j} \geq \exp\left(-\frac{\eta 2d_j}{\ell_j(1-\varepsilon/C)}\right) \geq \exp\left(-\frac{2\eta d_0}{\ell_0(1-\varepsilon/C)}\right),$$

where the second inequality follows from (I1).

We now have:

$$\exp\left(-\frac{2\eta d_0}{\ell_0(1-\varepsilon/C)}\right) = \exp\left(-\frac{2\mu}{(1-\varepsilon/C)C}\right) = \exp\left(-\frac{2\mu}{C-\varepsilon}\right).$$

Thus

$$k_j \geq \exp\left(-\frac{2\mu}{C-\varepsilon}\right).$$

Now we use this lower bound on  $k_j$  to derive an upper bound on  $u_j$  as follows:

$$\begin{aligned} u_j &= \left(1 - \frac{\eta}{\ell_j}\right)^{k_j \ell_j / 2} \leq \exp\left(-\frac{\eta k_j \ell_j}{2\ell_j}\right) = \exp\left(-\frac{\eta k_j}{2}\right) \leq \exp\left(-\frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right)\right) \\ &\leq 1 - (1 - \varepsilon/(4C)) \frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right), \end{aligned}$$

where the last inequality follows by the fact that for any  $\varepsilon > 0$ , there exists  $x_0 > 0$  such that  $1 - x \geq \exp\left(-\frac{x}{1-\varepsilon}\right)$  whenever  $0 < x < x_0$  (we assume  $d$  is large enough so that we may apply this inequality).

As the expression above is independent of  $j$ , we may use it to compute an upper bound on  $\frac{d_j}{\ell_j}$  as follows:

$$\frac{d_j}{\ell_j} = \frac{d_0}{\ell_0} \prod_{m=0}^{j-1} u_m \leq \frac{d_0}{\ell_0} \left(1 - (1 - \varepsilon/(4C)) \frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right)\right)^j. \quad (5.3)$$

We use this to improve our lower bound on  $k_j$ . Assuming  $d$  is large enough, by (5.2) we have  $\eta/\ell_j$  is small enough such that:

$$\begin{aligned} k_j &= \left(1 - \frac{\eta}{\ell_j}\right)^{2d_j} \geq \exp\left(-\frac{\eta}{\ell_j} \frac{2d_j}{(1-\varepsilon/(2C))}\right) \\ &\geq \exp\left(-\frac{2\eta}{(1-\varepsilon/(2C))} \frac{d_0}{\ell_0} \left(1 - (1 - \varepsilon/(4C)) \frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right)\right)^j\right) \\ &= \exp\left(-\frac{2\mu}{(1-\varepsilon/(2C))C} \left(1 - (1 - \varepsilon/(4C)) \frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right)\right)^j\right) \\ &= \exp\left(-\frac{2\mu}{C-\varepsilon/2} \left(1 - (1 - \varepsilon/(4C)) \frac{\eta}{2} \exp\left(-\frac{2\mu}{C-\varepsilon}\right)\right)^j\right), \end{aligned}$$

This bound holds for all  $0 \leq j \leq j'$ . We may now lower bound  $\ell_{j'+1}$  as follows:

$$\begin{aligned}\ell_{j'+1} &= \ell_0 \prod_{m=0}^{j'} k_m \geq \ell_0 \prod_{m=0}^{j'} \exp \left( -\frac{2\mu}{C-\varepsilon/2} \left( 1 - (1-\varepsilon/(4C)) \frac{\eta}{2} \exp \left( -\frac{2\mu}{C-\varepsilon} \right) \right)^m \right) \\ &= \ell_0 \exp \left( -\frac{2\mu}{C-\varepsilon/2} \sum_{m=0}^{j'} \left( 1 - (1-\varepsilon/(4C)) \frac{\eta}{2} \exp \left( -\frac{2\mu}{C-\varepsilon} \right) \right)^m \right) \\ &\geq \ell_0 \exp \left( -\frac{2\mu}{(C-\varepsilon/2) \left( 1 - (1 - (1-\varepsilon/(4C)) \frac{\eta}{2} \exp(-\frac{2\mu}{C-\varepsilon})) \right)} \right),\end{aligned}$$

where the last inequality follows by bounding the summation as a partial sum of a geometric series (we may assume that  $d$  is large enough implying  $\eta$  is small enough so this is indeed true). From this it follows:

$$\begin{aligned}\ell_{j'+1} &\geq \ell_0 \exp \left( -\frac{2\mu}{(C-\varepsilon/2) \left( 1 - \varepsilon/(4C) \right) \eta \exp(-\frac{2\mu}{C-\varepsilon})} \right) \\ &= \ell_0 \exp \left( \frac{-4 \log(dk^{-1/(s+t)}) \exp(\frac{2\mu}{C-\varepsilon})}{(C-\varepsilon/2)(1-\varepsilon/(4C))} \right) \\ &\geq \ell_0 \exp \left( \frac{-4 \log(dk^{-1/(s+t)}) \exp(\frac{2\mu}{C-\varepsilon})}{C-3\varepsilon/4} \right) \quad (\text{for } \varepsilon \text{ sufficiently small}) \\ &= \ell_0 (dk^{-1/(s+t)})^\xi \quad \left( \text{where } \xi := \frac{-4 \exp(\frac{2\mu}{C-\varepsilon})}{C-3\varepsilon/4} = \frac{-4(1+\varepsilon/(8C))}{C-3\varepsilon/4} \right) \\ &= \frac{C d}{\log(dk^{-1/(s+t)})} (dk^{-1/(s+t)})^\xi \\ &\geq d (dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)},\end{aligned}$$

where the last inequality follows for  $d$  large enough as  $\xi > -4/(C-7\varepsilon/8)$ . This completes the induction and finishes the proof of (I2).

For (I3), note that by (5.3), it follows for all  $i \in \mathbb{N}$ ,

$$\begin{aligned}\frac{d_i}{\ell_i} &\leq \frac{d_0}{\ell_0} \left( 1 - \left( 1 - \frac{\varepsilon}{4C} \right) \frac{\eta}{2} \exp \left( -\frac{2\mu}{C-\varepsilon} \right) \right)^i \\ &\leq \frac{d_0}{\ell_0} \exp \left( -i \left( 1 - \frac{\varepsilon}{4C} \right) \frac{\eta}{2} \exp \left( -\frac{2\mu}{C-\varepsilon} \right) \right) \\ &= \frac{\log(dk^{-1/(s+t)})}{C} \exp \left( -i \left( 1 - \frac{\varepsilon}{4C} \right) \frac{\eta}{2} \frac{1}{1+\varepsilon/(8C)} \right) \\ &\leq \frac{\log(dk^{-1/(s+t)})}{C} \exp \left( -\frac{i \mu}{8 \log(dk^{-1/(s+t)})} \right),\end{aligned}$$

for  $\varepsilon$  small enough. Now, for

$$j^* := \left\lceil \frac{16}{\mu} \log(dk^{-1/(s+t)}) \log \log(dk^{-1/(s+t)}) \right\rceil,$$

we have:

$$\begin{aligned} \frac{d_{j^*}}{\ell_{j^*}} &\leq \frac{\log(dk^{-1/(s+t)})}{C} \exp\left(-\frac{j^* \mu}{8 \log(dk^{-1/(s+t)})}\right) \\ &\leq \frac{\log(dk^{-1/(s+t)})}{C} \exp(-2 \log \log(dk^{-1/(s+t)})) \\ &\leq \frac{1}{C \log(dk^{-1/(s+t)})} \\ &\leq \frac{1}{100} \end{aligned}$$

for  $d$  large enough by (5.1). Thus a minimum integer  $i^* \leq j^*$  such that  $d_{i^*}/\ell_{i^*} \leq 1/100$  exists.  $\square$

**Lemma 35.** *There exists  $\alpha > 0$  such that the following holds: Let  $\varepsilon > 0$  be sufficiently small. Then there exists  $d^\#$  such that whenever*

$$d \geq d^\#, \quad 1 \leq t \leq s, \quad st \leq \frac{\alpha \varepsilon \log d}{\log \log d}, \quad \text{and} \quad 1/2 < k \leq d^{(s+t)/10},$$

*the following hold for all  $i \in \mathbb{N}$  with  $0 \leq i < i^*$ , where  $i^*$  is defined (and guaranteed to exist) by Lemma 34 item (I3).*

(R1)  $d_i$  is sufficiently large:  $d_i \geq \tilde{d}$ , where  $\tilde{d}$  is from Lemma 24,

(R2)  $k$  is not too large:  $k \leq d_i^{(s+t)/5}$ .

(R3)  $\ell_i$  is bounded below and above in terms of  $d_i$ :  $4\eta d_i < \ell_i < 100d_i$ ,

(R4)  $s$  and  $t$  are bounded in terms of  $d_i$ :  $1 \leq t \leq s$  and  $st \leq \frac{\tilde{\alpha} \log d_i}{\log \log d_i}$ ,

(R5)  $\eta$  is sufficiently small:  $\frac{1}{\log^5 d_i} < \eta < \frac{1}{\log(d_i k^{-1/(s+t)})}$ ,

(R6)  $\beta_i$  is small:  $\beta_i \leq 1/10$ . (Note that the bound  $\beta_i \geq d_i^{-1/(200st)}$  holds by definition.)

*Proof.* Once again, we will not explicitly compute  $d^\#$ , but instead simply take  $d$  to be sufficiently large when needed. For all  $i < i^*$ , we have  $d_i \geq \ell_i/100$ , and by (I2) of Lemma 34, it follows for  $d$  large enough that  $\ell_i \geq d(dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)}$ . Note the following as a result of (5.1):

$$d_i \geq \frac{1}{100} d (dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)} \geq d^{\varepsilon/40}. \quad (5.4)$$

Using the above, note for  $d^\#$  large enough, namely  $d^\# \geq \tilde{d}^{40/\varepsilon}$ , it follows

$$d_i \geq d^{\varepsilon/40} \geq \tilde{d},$$

showing (R1).

From (5.4) it also follows that when  $k \leq d^{\varepsilon(s+t)/200}$ , we have:

$$k \leq d^{\varepsilon(s+t)/200} \leq d_i^{40(s+t)/200} = d_i^{(s+t)/5}.$$

Otherwise, when  $k > d^{\varepsilon(s+t)/200}$ , we have  $C = 8$ . Using the first inequality in (5.4), it follows that:

$$\begin{aligned} d_i^{(s+t)/5} &\geq \left( \frac{1}{100} d (dk^{-1/(s+t)})^{-4/(C-7\varepsilon/8)} \right)^{(s+t)/5} \\ &= \left( \frac{1}{100} d^{(1-\frac{4}{8-7\varepsilon/8})} k^{(\frac{4}{(s+t)(8-7\varepsilon/8)})} \right)^{(s+t)/5} \\ &\geq \left( d^{(1-\frac{4}{8-7\varepsilon/4})} k^{(\frac{4}{(s+t)(8-7\varepsilon/8)})} \right)^{(s+t)/5} \\ &\geq \left( k^{\frac{10}{s+t}(1-\frac{4}{8-7\varepsilon/4}) + \frac{4}{(s+t)(8-7\varepsilon/8)}} \right)^{(s+t)/5} \quad (\text{as } k \leq d^{(s+t)/10}) \\ &= k^{2(1-\frac{4}{8-7\varepsilon/4}) + \frac{4}{5(8-7\varepsilon/8)}}. \end{aligned}$$

Let us consider the exponent above. We have

$$\begin{aligned} 2 \left( 1 - \frac{4}{8-7\varepsilon/4} \right) + \frac{4}{5(8-7\varepsilon/8)} &= 2 - \frac{4}{5} \left( \frac{10}{8-7\varepsilon/4} - \frac{1}{8-7\varepsilon/8} \right) \\ &= 2 - \frac{4(10(8-7\varepsilon/8) - (8-7\varepsilon/4))}{5(8-7\varepsilon/4)(8-7\varepsilon/8)} \\ &= 2 - \frac{4}{5} \left( \frac{72-7\varepsilon}{64-21\varepsilon+49\varepsilon^2/32} \right) \\ &\geq 1, \end{aligned}$$

for  $\varepsilon$  sufficiently small. Therefore,  $k \leq d_i^{(s+t)/5}$  completing the proof of (R2).

We also have

$$\frac{\ell_i}{d_i} \geq \frac{\ell_0}{d_0} = \frac{C\eta}{\mu}.$$

Since  $\mu < 1$  for  $\varepsilon$  sufficiently small, this shows

$$\ell_i \geq \frac{C\eta}{\mu} d_i \geq C\eta d_i \geq 4\eta d_i$$

proving (R3) as the upper bound always holds for  $i < i^*$ .

As  $\log(x)/\log\log(x)$  is increasing for  $x$  large enough, it follows we can take  $\tilde{d}$  sufficiently large such that

$$st \leq \frac{\alpha\varepsilon \log d}{\log\log d} \leq \frac{\alpha\varepsilon \log(d_i^{40/\varepsilon})}{\log\log(d_i^{40/\varepsilon})} \leq \frac{40\alpha \log d_i}{\log\log d_i} \leq \frac{\tilde{\alpha} \log d_i}{\log\log d_i},$$

when  $\alpha \leq \tilde{\alpha}/40$ . This proves (R4).

Again using the fact that  $dk^{-1/(s+t)} \leq 2d$ , we see for  $d$  large enough that

$$\frac{1}{\log^5(d_i)} \leq \frac{1}{\log^5(d^{\varepsilon/40})} = \frac{(40/\varepsilon)^5}{\log^5(d)} < \frac{\mu}{\log(2d)} \leq \eta.$$

Furthermore, for  $\varepsilon$  sufficiently small, we have  $\mu < 1$ . Thus, as  $d_i \leq d$ , we may conclude that

$$\eta \leq \frac{1}{\log(dk^{-1/(s+t)})} \leq \frac{1}{\log(d_i k^{-1/(s+t)})}.$$

This shows (R5). Let  $0 \leq i' \leq i^* - 1$  be the largest integer such that  $\beta_{i'} = d_i^{-1/(200st)}$ . Then

$$\begin{aligned} \beta_{i^*-1} &= (1 + 36\eta)^{i^*-1-i'} d_i^{-1/(200st)} \\ &\leq (1 + 36\eta)^{i^*} (d^{\varepsilon/40})^{-1/(200st)} \\ &\leq \exp(36\eta i^*) d^{-\varepsilon/(800st)}. \end{aligned}$$

By Lemma 34 item (I3), we may conclude the following:

$$\begin{aligned} \beta_{i^*-1} &\leq \exp(36\eta i^*) d^{-\varepsilon/(800st)} \\ &\leq \exp\left(576 \frac{\eta}{\mu} \log(dk^{-1/(s+t)}) \log\log(dk^{-1/(s+t)})\right) d^{-\varepsilon/(800st)} \\ &= (\log(dk^{-1/(s+t)}))^{576} d^{-\varepsilon/(800st)} \\ &\leq (\log(2d))^{576} d^{-\varepsilon/(800st)} \\ &\leq (\log(2d))^{576} d^{-\frac{\varepsilon \log\log d}{800\alpha\varepsilon \log d}} \\ &\leq \log(d)^{1000} \log(d)^{-\frac{1}{800\alpha}} \\ &= \log(d)^{1000-\frac{1}{800\alpha}} \\ &\leq \frac{1}{10}, \end{aligned}$$

for  $d$  large enough and  $\alpha$  small enough. This shows (R6) and completes the proof. □

We are now ready to prove the main theorem:

**Theorem 36** (Restatement of Theorem 13). *There exists a constant  $\alpha > 0$  such that for every  $\varepsilon > 0$ , there is  $d^* \in \mathbb{N}$  such that the following holds. Suppose that  $d, s, t \in \mathbb{N}$ , and  $k \in \mathbb{R}$  satisfy*

$$d \geq d^*, \quad 1 \leq t \leq s, \quad st \leq \frac{\alpha \varepsilon \log d}{\log \log d}, \quad \text{and} \quad 1/2 \leq k \leq d^{(s+t)/10}.$$

Define  $C$  as follows:

$$C := \begin{cases} 4 + \varepsilon & k \leq d^{\varepsilon(s+t)/200}, \\ 8 & \text{otherwise.} \end{cases}$$

If  $G$  is a graph and  $\mathcal{H} = (L, H)$  is a correspondence cover of  $G$  such that:

- (i)  $H$  is  $(k, K_{s,t})$ -locally-sparse,
- (ii)  $\Delta(H) \leq d$ , and
- (iii)  $|L(v)| \geq C d / \log(dk^{-1/(s+t)})$  for all  $v \in V(G)$ .

Then,  $G$  admits a proper  $\mathcal{H}$ -coloring.

*Proof.* Let  $\alpha$  be as defined in Lemma 35. Let  $\varepsilon > 0$  be sufficiently small. We will not explicitly compute  $d^*$  but simply take  $d$  large enough when needed. Thus let  $d$  be sufficiently large so that we may apply Lemma 35.

Let  $G$  be a graph and  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$  satisfying the hypotheses of the theorem. Set

$$G_0 := G, \quad \mathcal{H}_0 = (L_0, H_0) := \mathcal{H}$$

By slightly modifying  $\varepsilon$  if necessary, and taking  $d$  sufficiently large, we may assume  $\ell_0$  is an integer. By removing some of the vertices from  $H$  if necessary, we may assume that  $|L(v)| = \ell_0$  for all  $v \in V(G)$ .

At this point, we recursively define a sequence of graphs  $G_i$  with correspondence cover  $\mathcal{H}_i = (L_i, H_i)$  for  $0 \leq i \leq i^*$  by iteratively applying Lemma 24 as follows: Lemma 35 shows conditions (L1)–(L5) of Lemma 24 (with  $d_i$  in place of  $d$ ,  $\ell_i$  in place of  $\ell$ ) hold for each  $0 \leq i \leq i^* - 1$ . Therefore, if  $G_i$  is a graph with correspondence cover  $\mathcal{H}_i = (L_i, H_i)$  such that for  $\beta_i$  as defined at the start of the section,

- (S1)  $H_i$  is  $(k, K_{s,t})$ -locally-sparse,
- (S2)  $\Delta(H_i) \leq 2d_i$ ,
- (S3) the list sizes are roughly between  $\ell_i/2$  and  $\ell_i$ :

$$(1 - \beta_i)\ell_i/2 \leq |L_i(v)| \leq (1 + \beta_i)\ell_i \quad \text{for all } v \in V(G_i),$$

- (S4) average color-degrees are smaller for vertices with smaller lists of colors:

$$\overline{\deg}_{\mathcal{H}_i}(v) \leq \left(2 - (1 - \beta_i) \frac{\ell_i}{|L_i(v)|}\right) d_i \quad \text{for all } v \in V(G_i),$$



then  $G_i$ , and  $\mathcal{H}_i$  satisfy the required hypotheses of Lemma 24. In particular, there exists a partial  $\mathcal{H}_i$ -coloring  $\varphi_i$  of  $G_i$  and an assignment of subsets  $L_{i+1}(v) \subseteq (L_i)_{\varphi_i}(v)$  to each vertex  $v \in V(G_i) \setminus \text{dom}(\varphi_i)$  such that, setting

$$G_{i+1} := G_i[V(G_i) \setminus \text{dom}(\varphi_i)], \quad H_{i+1} := H_i \left[ \bigcup_{v \in V(G_{i+1})} L_{i+1}(v) \right],$$

$$\text{and} \quad \mathcal{H}_{i+1} := (L_{i+1}, H_{i+1}),$$

we have items (S1)–(S4) hold for  $G_{i+1}$  and  $\mathcal{H}_{i+1}$  with parameter  $\beta_{i+1}$ . Thus we may iteratively apply Lemma 24  $i^*$  times, starting with  $G_0$  and  $\mathcal{H}_0$ , to define  $G_i$  and  $\mathcal{H}_i$  for  $0 \leq i \leq i^*$ .

We now show  $G_{i^*}$  and  $\mathcal{H}_{i^*}$  satisfy the hypotheses of Proposition 20. By (S3), it follows for each  $v \in V(G_{i^*})$ ,

$$|L_{i^*}(v)| \geq (1 - \beta_{i^*})\ell_{i^*}/2 \geq (1 - (1 + 36\eta)\beta_{i^*-1})\ell_{i^*}/2.$$

Since  $\beta_{i^*-1} \leq \frac{1}{10}$  and  $\eta \leq \frac{1}{100}$  for  $d$  large enough, it follows

$$(1 - (1 + 36\eta)\beta_{i^*-1})\ell_{i^*}/2 \geq \left(1 - \frac{1 + 36/100}{10}\right)\ell_{i^*}/2 \geq \frac{1}{4}\ell_{i^*}$$

for  $d$  large enough. As  $\ell_{i^*} \geq 100d_{i^*}$ , we have:

$$\frac{1}{4}\ell_{i^*} \geq 25d_{i^*}.$$

We have by (S2) that  $\Delta(H_{i^*}) \leq 2d_{i^*}$ , thus for all  $c \in V(H_{i^*})$ , we have

$$25d_{i^*} \geq 10 \deg_{H_{i^*}}(c).$$

Putting this chain of inequalities together yields  $|L_{i^*}(v)| \geq 8 \deg_{H_{i^*}}(c)$ , as desired.

Therefore by Proposition 20 there exists an  $\mathcal{H}_{i^*}$ -coloring of  $G_{i^*}$ , say  $\varphi_{i^*}$ . Letting  $\varphi_i$  be the partial  $\mathcal{H}_i$ -colorings of  $G_i$  for  $0 \leq i \leq i^* - 1$  guaranteed at each application of Lemma 24, it follows

$$\bigcup_{i=0}^{i^*} \varphi_i$$

is a proper  $\mathcal{H}$ -coloring of  $G$ , as desired. □

## 6 Concluding remarks

We conclude with a description of potential future directions of inquiry. In this paper, we show an asymptotically improved bound on the (correspondence) chromatic number of a graph  $G$  of maximum degree  $\Delta$  when  $G$  has at most  $k \ll \Delta^{s+t}$  copies of  $K_{s,t}$  in the neighborhood of any vertex. It is natural to ask if we can extend the range of  $k$  to

$k = \Theta(\Delta^{s+t})$ . We note that for  $k = \binom{\Delta}{s+t} \binom{s+t}{t}$ , in the worst case the graph is complete, implying

$$\chi_c(G) = \Delta + 1 = O\left(\frac{\Delta}{\log(\Delta k^{-1/(s+t)})}\right).$$

Therefore, we suspect our results should extend to all values of  $k$  in this range (although the proof would require some new ideas; recall the discussion at the end of §3). Furthermore, in light of our results and Conjecture 5, we make the following more general conjecture (a correspondence coloring version of Conjecture 5 appeared in [5]).

**Conjecture 37.** For every graph  $F$ , the following holds for  $\Delta \in \mathbb{N}$  large enough. Let

$$1/2 \leq k < \Delta^{|V(F)|},$$

and let  $G$  be a  $(k, F)$ -locally-sparse graph of maximum degree  $\Delta$ . Then,

$$\chi_c(G) = O\left(\frac{\Delta}{\log(\Delta k^{-1/|V(F)|})}\right),$$

where the constant factor in the  $O(\cdot)$  may depend on  $F$ .

We note that a simple application of Theorem 10 together with Proposition 22 proves Conjecture 37 for bipartite graphs  $F$  and the following range of  $k$ :

$$|V(F)|^2 \Delta^{|V(F)|-1} \leq k \leq \frac{\Delta^{|V(F)|}}{(\log \Delta)^c},$$

for some constant  $c := c(F) > 0$ . As a result of Corollary 15, it remains to consider the range

$$\Delta^{|V(F)|/10} < k < |V(F)|^2 \Delta^{|V(F)|-1}, \quad \text{and} \quad k > \frac{\Delta^{|V(F)|}}{(\log \Delta)^c},$$

in order to verify the conjecture for bipartite graphs.

We remark that the constant  $C$  in Corollary 15 is independent of  $F$ . This is in line with several recent results surrounding Conjecture 5 which prove the constant in the  $O(\cdot)$  is independent of  $F$  (such as [33, 30] for  $F = K_3$ , [14] for fans and cycles, [5] for  $F$  bipartite, and [6] for  $F$  almost-bipartite; see Table 1). To this end, we conjecture the following, which is a strengthening of [6, Conjecture 1.7]:

**Conjecture 38.** There exists a universal constant  $C \in \mathbb{R}$  such that for every graph  $F$ , the following holds for  $\Delta \in \mathbb{N}$  large enough. Let

$$1/2 \leq k < \Delta^{|V(F)|},$$

and let  $G$  be a  $(k, F)$ -locally-sparse graph of maximum degree  $\Delta$ . Then,

$$\chi_c(G) \leq \frac{C \Delta}{\log(\Delta k^{-1/|V(F)|})}.$$

We note that for an arbitrary graph, a simple counting argument shows that

$$k \leq \binom{\Delta}{|V(F)|} \frac{|V(F)|!}{|\text{Aut}(F)|},$$

where the bound is tight when considering the number of copies of  $F$  in  $K_\Delta$ . As the above is strictly less than  $\Delta^{|V(F)|}$ , the bounds on  $k$  in Conjecture 37 and Conjecture 38 cover all possible cases.

We also note that Corollary 15 verifies this conjecture in the case when  $F$  is bipartite for a smaller range of  $k$  than stated. Additionally, we remark that Conjecture 38 is quite challenging. Indeed, Conjecture 5 itself is considered ambitious (see the discussion at the end of [6, §1.1]). As such, falsifying Conjecture 38 may be a much more feasible task than disproving Conjecture 5 and so we pose it as an interesting question.

As mentioned earlier, the best known upper bound for Conjecture 5 and arbitrary  $F$  is  $O(\Delta \log \log \Delta / \log \Delta)$  due to Johansson [23]. In fact, his proof holds in the list coloring setting as well. The constant factor was improved by Molloy [30], and Bernshteyn extended the result to correspondence coloring [7]. A natural question to ask is whether we can prove a similar bound for  $(k, F)$ -locally-sparse graphs. This would be the first step toward proving Conjecture 37 and Conjecture 38.

## Acknowledgements

We thank József Balogh, Peter Bradshaw, and Ethan White for pointing out an error in an earlier version of this manuscript, and Anton Bernshteyn for his helpful comments. We also thank the anonymous referee for their helpful suggestions.

James Anderson’s research was partially supported by NSF CAREER grant DMS-2239187 (PI: Anton Bernshteyn). Abhishek Dhawan’s research was partially supported by the Georgia Tech ARC-ACO Fellowship, NSF grant DMS-2053333 (PI: Cheng Mao), NSF CAREER grant DMS-2239187 (PI: Anton Bernshteyn), and NSF RTG grant DMS-1937241. Aiya Kuchukova’s research was partially supported by the Georgia Tech ARC-ACO Fellowship.

## References

- [1] N. Alon. Testing subgraphs in large graphs. *Random Structures & Algorithms*, 21(3–4): 359–370, 2002.
- [2] N. Alon and S. Assadi. Palette sparsification beyond  $(\Delta + 1)$  vertex coloring. In *Proceedings of the International Conference on Approximation, Randomization, and Combinatorial Optimization*, 2020. doi:[10.4230/LIPIcs.APPROX/RANDOM.2020.6](https://doi.org/10.4230/LIPIcs.APPROX/RANDOM.2020.6)
- [3] N. Alon, S. Cambie, and R. J. Kang. Asymmetric list sizes in bipartite graphs. *Annals of Combinatorics*, 25(4): 913–933, 2021.
- [4] N. Alon, M. Krivelevich, and B. Sudakov. Coloring graphs with sparse neighborhoods. *Journal of Combinatorial Theory, Series B*, 77: 73–82, 1999.

- [5] J. Anderson, A. Bernshteyn, and A. Dhawan. Colouring graphs with forbidden bipartite subgraphs. *Combinatorics, Probability and Computing*, 32(1): 45–67, 2023.
- [6] J. Anderson, A. Bernshteyn, and A. Dhawan. Coloring graphs with forbidden almost bipartite subgraphs. Preprint, 2022. [arXiv:2203.07222](https://arxiv.org/abs/2203.07222).
- [7] A. Bernshteyn. The Johansson–Molloy theorem for DP-coloring. *Random Structures & Algorithms*, 54(4): 653–664, 2019.
- [8] T. Bohman and R. Holzman. On a list coloring conjecture of Reed. *Journal of Graph Theory*, 41(2): 106–109, 2002.
- [9] B. Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, 83(2): 433–436, 1981.
- [10] R. L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2): 194–197, 1941.
- [11] H. Bruhn and F. Joos. A stronger bound for the strong chromatic index. *Electronic Notes in Discrete Mathematics*, 49: 277–284, 2015.
- [12] S. Cambie and R. J. Kang. Independent transversals in bipartite correspondence-covers. *Canadian Mathematical Bulletin*, 65(4): 882–894, 2022.
- [13] E. Davies, R. J. Kang, F. Pirot, and J.-S. Sereni. An algorithmic framework for colouring locally sparse graphs. *Theory of Computing*, 2024.
- [14] E. Davies, R. J. Kang, F. Pirot, and J.-S. Sereni. Graph structure via local occupancy. Preprint, 2020. [arXiv:2003.14361](https://arxiv.org/abs/2003.14361).
- [15] A. Dhawan. Bounds for the independence and chromatic numbers of locally sparse graphs. *Annals of Combinatorics*, 2025. [doi:10.1007/s00026-025-00778-7](https://doi.org/10.1007/s00026-025-00778-7).
- [16] A. Dhawan. Palette sparsification for graphs with sparse neighborhoods. Preprint, 2024. [arXiv:2408.08256](https://arxiv.org/abs/2408.08256).
- [17] Z. Dvořák and L. Postle. Correspondence coloring and its application to list-coloring planar graphs without cycles of lengths 4 to 8. *Journal of Combinatorial Theory, Series B*, 2018.
- [18] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(12): 1087–1091, 1946.
- [19] P. Erdős, A. L. Rubin, and H. Taylor. Choosability in graphs. *Congressus Numerantium*, 26: 125–157, 1979.
- [20] S. Glock and B. Sudakov. An average degree condition for independent transversals. *Journal of Combinatorial Theory, Series B*, 154: 370–391, 2022.
- [21] C. Hyltén-Cavallius. On a combinatorial problem. *Colloquium Mathematicum*, 6: 61–65, 1958.
- [22] M. S. Jamall. *Coloring Triangle-Free Graphs and Network Games*. Ph.D. Thesis, University of California, San Diego, 2011.
- [23] A. Johansson. The choice number of sparse graphs. Preprint, 1996.

- [24] A. Johansson. Asymptotic choice number for triangle-free graphs. DIMACS Technical Report 91–95, 1996.
- [25] J. Kahn. Asymptotically good list-colorings. *Journal of Combinatorial Theory, Series A*, 73: 1–59, 1996.
- [26] R. J. Kang and T. Kelly. Colourings, transversals and local sparsity. *Random Structures & Algorithms*, 61(1): 173–192, 2022.
- [27] J. H. Kim. On Brooks’ theorem for sparse graphs. *Combinatorics, Probability and Computing*, 4: 97–132, 1995.
- [28] B. Kővári, V. T. Sós, and P. Turán. On a problem of K. Zarankiewicz. *Colloquium Mathematicum*, 3: 50–57, 1954.
- [29] P.-S. Loh and B. Sudakov. Independent transversals in locally sparse graphs. *Journal of Combinatorial Theory, Series B*, 97: 904–918, 2007.
- [30] M. Molloy. The list chromatic number of graphs with small clique number. *Journal of Combinatorial Theory, Series B*, 134: 264–284, 2019.
- [31] M. Molloy and B. Reed. *Graph Colourings and the Probabilistic Method*. Springer, 2002.
- [32] M. Molloy and B. Reed. A bound on the strong chromatic index of a graph. *Journal of Combinatorial Theory, Series B*, 69(2): 103–109, 1997.
- [33] S. Pettie and H.-H. Su. Distributed coloring algorithms for triangle-free graphs. *Information and Computation*, 243: 263–280, 2015.
- [34] B. Reed. A strengthening of Brooks’ theorem. *Journal of Combinatorial Theory, Series B*, 76(2): 136–149, 1999.
- [35] B. Reed. The list colouring constants. *Journal of Graph Theory*, 31: 149–153, 1999.
- [36] B. Reed and B. Sudakov. Asymptotically the list colouring constants are 1. *Journal of Combinatorial Theory, Series B*, 86: 27–37, 2002.
- [37] V. Rödl. On a packing and covering problem. *European Journal of Combinatorics*, 6(1): 69–78, 1985.
- [38] J. B. Shearer. On the independence number of sparse graphs. *Random Structures & Algorithms*, 7(3): 269–271, 1995.
- [39] V. G. Vizing. Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz*, 29: 10, 1976.

## A Proof of Proposition 11

We restate the proposition here for convenience.

**Proposition 39** (11). *Let  $F$  and  $G$  be graphs and let  $\mathcal{H} = (L, H)$  be a correspondence cover of  $G$ . For  $k \in \mathbb{R}$ , the following holds:*

(S1) *If  $G$  is  $(k, F)$ -locally-sparse, then so is  $H$ .*

(S2) If  $F$  satisfies

$$\forall u, v \in V(F), \quad uv \notin E(F) \implies N_F(u) \cap N_F(v) \neq \emptyset,$$

then if  $G$  is  $F$ -free,  $H$  is as well.

*Proof.* We first prove (S1). Let  $v \in V(G)$  and  $c \in L(v)$ . From (CC3) it follows  $|N_H(c) \cap L(u)| \leq 1$  for each  $u \in V(G)$ . Suppose  $H[N_H(c)]$  contains at least  $\lfloor k \rfloor + 1$  copies of  $F$ . Let  $X$  be one such copy. By the earlier observation, the map  $c' \rightarrow L^{-1}(c')$  is injective for  $c' \in X$ , i.e., no two vertices of  $X$  are in  $L(u)$  for any  $u \in V(G)$ . Thus the map  $\psi : V(X) \rightarrow V(G)$  defined by  $\psi(c') = L^{-1}(c')$  is an isomorphism from  $X$  to a subgraph of  $G$ . In addition, as  $V(X) \subseteq N_H(c)$ , it follows by the second clause in (CC3) that  $L^{-1}(c') \in N_G(v)$  for each  $c' \in V(X)$ . Therefore,  $\psi(X) \subseteq G[N_G(v)]$ . Thus if  $H$  contains a copy of  $F$  in a neighborhood, then so does  $G$ .

Now, let  $X \neq X'$  be distinct copies of  $F$  in  $H[N_H(c)]$ . As shown earlier, there are subgraphs  $\psi(X), \psi(X')$  in  $G[N_G(v)]$  isomorphic to  $X, X'$  respectively. It is now enough to show that  $\psi(X) \neq \psi(X')$ . If  $V(X) \neq V(X')$ , then by the fact that  $c \rightarrow L^{-1}(c)$  is injective for  $c \in X$ , we have  $V(\psi(X)) \neq V(\psi(X'))$  and thus  $\psi(X) \neq \psi(X')$ . If  $V(X) = V(X')$ , then as  $X \neq X'$ , it follows there is some edge  $c_1c_2 \in E(X) \setminus E(X')$ . Since  $\psi$  is an isomorphism, the edge  $L^{-1}(c_1)L^{-1}(c_2) \in E(\psi(X)) \setminus E(\psi(X'))$ , and  $\psi(X) \neq \psi(X')$ . Thus  $\psi(X)$  and  $\psi(X')$  are two different copies of  $F$  in  $G[N_G(v)]$ . It follows that if  $H$  contains  $\lfloor k \rfloor + 1$  copies of  $F$  in  $H[N_H(c)]$ , then  $G$  contains  $\lfloor k \rfloor + 1$  copies of  $F$  in  $G[N_G(v)]$ , contradicting the local sparsity  $G$ . Thus  $H$  is  $(k, F)$ -locally-sparse.

We now prove (S2). Suppose  $X$  is a copy of  $F$  in  $H$ . From the argument in the previous paragraph, if the map  $\varphi : V(X) \rightarrow V(G)$  by  $c \rightarrow L^{-1}(c)$  is injective, then it follows that  $\varphi(X)$  is a copy of  $F$  in  $G$ . If  $F$  is a clique, and  $X$  is a copy of  $F$ , then clearly (CC2) implies  $\varphi$  is injective. Suppose  $F$  is not a clique and let  $X$  be a copy of  $F$  such that  $\varphi$  is not injective, i.e., there exist  $c, c' \in V(X)$  with  $c, c' \in L(u)$  for some  $u \in V(G)$ . Then, as (CC2) states  $L(u)$  is an independent set, it follows  $cc' \notin E(F)$ . Thus, by the assumption of (S2), it follows  $N_X(c) \cap N_X(c') \neq \emptyset$ . However, this violates (CC3). Thus  $\varphi$  is injective. Therefore, if  $H$  contains a copy of  $F$ , then so does  $G$ . As  $G$  is  $F$ -free, we conclude  $H$  is as well.  $\square$

## B Proof of Proposition 22

Let us first restate the result.

**Proposition 40** (Restatement of Proposition 22). *Let  $G$  be an  $n$ -vertex  $(k, K_{s,t})$ -sparse graph. For  $k^* = \frac{(s+t)^2 n^{s+t-1}}{2 s^s t^t}$ , we have*

$$|E(G)| \leq \begin{cases} 4 n^{2-1/(st)} & \text{if } k \leq k^*; \\ 2 s^{1/t} t^{1/s} k^{1/st} n^{2-1/s-1/t} & \text{if } k > k^*. \end{cases}$$

We will use the following result of Alon in our proof.

**Lemma 41** (Contrapositive of [1, Lemma 2.1]). *Let  $s \geq t \geq 1$  and let  $G = (V, E)$  be a graph such that there are at most  $k$  homomorphisms from a labelled copy of  $H = K_{s,t}$  into  $G$ . Then,  $|E| \leq k^{1/st} n^{2-1/s-1/t}/2$ .*

*Proof of Proposition 22.* It is easy to see that there are at most  $(s+t)^2 n^{s+t-1}$  non-injective homomorphisms from a labelled copy of  $K_{s,t}$  into  $G$ . As  $G$  contains at most  $k$  copies of  $K_{s,t}$ , the number of injective homomorphisms from a labelled copy of  $K_{s,t}$  into  $G$  is at most

$$2^{\mathbb{1}\{s=t\}} s! t! k \leq 2 s^s t^t k.$$

Therefore, the number of homomorphisms from  $K_{s,t}$  into  $G$  is at most

$$(s+t)^2 n^{s+t-1} + 2 s^s t^t k \leq \begin{cases} 2 (s+t)^2 n^{s+t-1} & \text{if } k \leq k^*; \\ 4 s^s t^t k & \text{if } k > k^*. \end{cases}$$

As  $s, t \in \mathbb{N}$  and  $s \geq t \geq 1$ , we have

$$(2(s+t)^2)^{1/(st)} \leq 8, \quad \text{and} \quad 4^{1/(st)} \leq 4.$$

The claim now follows by applying Lemma 41. □