

Anticoncentration for subgraph counts in random graphs

Jacob Fox*

Matthew Kwan†

Lisa Sauermann‡

Abstract

Fix a graph H and some $p \in (0, 1)$, and let X_H be the number of copies of H in a random graph $\mathbb{G}(n, p)$. Random variables of this form have been intensively studied since the foundational work of Erdős and Rényi. There has been a great deal of progress over the years on the large-scale behaviour of X_H , but the more challenging problem of understanding the small-ball probabilities has remained poorly understood until now. More precisely, how likely can it be that X_H falls in some small interval or is equal to some particular value? In this paper we prove the almost-optimal result that if H is connected then for any $x \in \mathbb{N}$ we have $\Pr(X_H = x) \leq n^{1-v(H)+o(1)}$. Our proof proceeds by iteratively breaking X_H into different components which fluctuate at “different scales”, and relies on a new anticoncentration inequality for random vectors that behave “almost linearly”.

1 Introduction

Let $\mathbb{G}(n, p)$ be the binomial random graph model, where we fix a set of n vertices and include each of the $\binom{n}{2}$ possible edges independently with probability p . For graphs H and G , let $X_H(G)$ be the number of subgraphs of G isomorphic to H , so that if $G \sim \mathbb{G}(n, p)$ then we can interpret X_H as the random variable that counts the number of copies of H in a random graph.

These subgraph-counting random variables and their distributions are central objects of study in the theory of random graphs, going back to the foundational work of Erdős and Rényi [17]. Early work [9, 17, 35] concerned existence of subgraphs: fixing H , for which n and p is it likely that $X_H > 0$, and for which n and p is it likely that $X_H = 0$? It turns out that there is a *threshold* value of p (as a decaying function of n) that cleanly separates these two behaviours, and further work [9, 10, 17, 25, 35, 37] focused on investigating the (Poisson-type) distribution of X_H near this threshold.

In this paper, we are more interested in the behaviour far above this existence threshold (when p is a constant independent of n). When appropriately normalised, X_H has an asymptotically¹ normal distribution (this was proved by Nowicki and Wierman [32] and Ruciński [34], following several results [3, 24, 25] pushing increasingly further past the existence threshold). Further work by Barbour, Karoński and Ruciński [4] provided quantitative bounds on the rate of convergence to the normal distribution. However, this asymptotic normality only characterises the “large-scale” behaviour of the distribution of X_H , and is basically due to the fact that X_H closely correlates with the number of edges in $\mathbb{G}(n, p)$.

A more challenging direction of research is to understand “local” aspects of the distributions of these subgraph-counting random variables². In the past decade, there have been a number of advances in this direction. Following work by Loeb, Matoušek and Pangrác [27] for the case where H is a triangle, it was proved by Kolaitis and Kopparty [26] (see also [14]) that if we fix some $p \in (0, 1)$, some prime $q \in \mathbb{N}$ and some connected graph H with at least one edge, then $X_H \bmod q$ has an asymptotically uniform distribution on $\{0, \dots, q-1\}$. More recently, *local central limit theorems* have begun to emerge, giving asymptotic formulas for the point probabilities $\Pr(X_H = x)$ in terms of a normal density function. Such

*Department of Mathematics, Stanford University, Stanford, CA 94305. Email: jacobfox@stanford.edu. Research supported by a Packard Fellowship and by NSF Career Award DMS-1352121.

†Department of Mathematics, Stanford University, Stanford, CA 94305. Email: mattkwan@stanford.edu. Research supported in part by SNSF project 178493.

‡School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540. Email: lsauerma@stanford.edu.

¹All asymptotics, here and for the rest of the paper, are as $n \rightarrow \infty$, while p is fixed.

²We remark that another completely different challenging direction of research is the study of *large deviations* of subgraph counts in random graphs. Recently, there has been a lot of progress in this area, see for example the monograph [11], the recent papers [2, 5, 8, 13, 22], and the references therein.

a theorem was first proved for the case where H is a triangle by Gilmer and Kopparty [20] (see also [6]), and this was extended by Berkowitz [7] to the case where H is any clique.

In this paper we are concerned with a somewhat looser question: what can be said about the *anticoncentration* behaviour of X_H ? Roughly speaking, this is asking for uniform upper bounds on the point probabilities $\Pr(X_H = x)$, or more generally on the small ball probabilities $\Pr(X_H \in I)$, where I is an interval of prescribed length. Meka, Nguyen and Vu [30] developed some general polynomial anticoncentration inequalities, and used the polynomial structure of X_H to prove the bound $\Pr(X_H = x) \leq n^{-1+o(1)}$ for constant $p \in (0, 1)$ and any H that contains at least one edge. In [19] we proposed the following conjecture.

Conjecture 1.1. *Fix $p \in (0, 1)$ and fix a graph H with no isolated vertices. Then*

$$\max_{x \in \mathbb{N}} \Pr(X_H = x) = O\left(n^{1-v(H)}\right),$$

where $v(H)$ is the number of vertices of H .

The motivation for Conjecture 1.1 is that if p is fixed then $\text{Var } X_H = \Theta(n^{2v(H)-2})$ (see for example [23, Lemma 3.5]), and the aforementioned asymptotic normality therefore implies that X_H is concentrated on an interval of length $\Theta(n^{v(H)-1})$. Provided that the distribution of X_H is sufficiently “smooth”, we should expect each value in this interval to have comparable probability. Note that this line of reasoning implies that Conjecture 1.1, if true, is best possible: any stronger bound would contradict Chebyshev’s inequality. Also, observe that the assumption that H has no isolated vertices is necessary: if H' is obtained from H by removing isolated vertices then X_H is a deterministic multiple of $X_{H'}$, so inherits its point probabilities.

We also remark that while Meka, Nguyen and Vu were the first to explicitly consider anticoncentration of subgraph counts in general, actually Pangrác, Matoušek and Loeb [27] considered anticoncentration of the triangle-count X_{K_3} more than ten years earlier: a primary motivation for their aforementioned work on triangle-counts mod q was to show that the point probabilities $\Pr(X_{K_3} = x)$ are small (they gave a bound of $O(1/\log n)$), which in turn implies that two independent copies of $\mathbb{G}(n, p)$ are unlikely to have the same Tutte polynomial. In addition, many of the other aforementioned results concerning the distribution of X_H imply anticoncentration bounds: any central limit theorem already implies that $\max_x \Pr(X_H = x) = o(1)$, and one can³ deduce from the quantitative central limit theorem of Barbour, Karoński and Ruciński [4] that $\Pr(X_H = x) \leq \Pr(|X - x| \leq n^{v(H)-3/2}) = O(1/\sqrt{n})$. The local central limit theorem of Berkowitz [7] definitively settles the matter in the case where $H = K_h$ is an h -vertex clique, in which case it actually gives the asymptotically optimal bound $\Pr(X_H = x) \leq (2\pi \text{Var } X_H)^{-1/2} + o(n^{1-h}) = O(n^{1-h})$.

In [19] we used ideas related to Erdős’ combinatorial proof of the Erdős–Littlewood–Offord theorem (see [18]) to give a simple proof of the general bound $\Pr(X_H = x) \leq \Pr(|X_H - x| \leq n^{v(H)-2}) = O(1/n)$, and we showed how to extend these methods to prove the sharper bound $\Pr(X_{K_h} = x) = n^{1-h+o(1)}$ in the case where $H = K_h$ is a clique. In this paper we develop these methods much further, proving an approximate version of Conjecture 1.1 for all connected H .

Theorem 1.2. *Fix $p \in (0, 1)$ and fix a connected graph H . Then*

$$\max_{x \in \mathbb{Z}} \Pr(X_H = x) = n^{1-v(H)+o(1)}.$$

Concerning the $o(1)$ -term in Theorem 1.2, the arguments in our proof yield a bound for the $o(1)$ -term that decays extremely slowly as n goes to infinity. In order to simplify the presentation of the proof and to avoid additional technical details, we decided to write the proof in a way that does not give any explicit bounds for the $o(1)$ -term.

The general idea for the proof of Theorem 1.2 is to break up X_H into different components that fluctuate at “different scales” and handle each component separately. In the case where H is a clique, this plan is relatively simple to execute, but in the more general setting of Theorem 1.2 there are a number of

³The central limit theorem of Barbour, Karoński and Ruciński was not stated in a way that allows one to directly read off estimates for probabilities regarding X_H . But, it is possible to deduce such an estimate with the method of [33, Proposition 1.2.2].

additional challenges that must be overcome. We discuss these in Section 2. Our proof has a number of new ingredients; one that is perhaps worth highlighting is a combinatorial anticoncentration inequality for vector-valued random variables that behave “almost linearly”, in the spirit of some anticoncentration theorems due to Halász. The details are in Section 3.

1.1 Basic definitions and notation

We use standard graph-theoretic notation throughout. In particular, the vertex and edge sets of a graph G are denoted by $V(G)$ and $E(G)$, and the sizes of these sets are denoted by $e(G)$ and $v(G)$. For disjoint vertex sets A, B in a graph G , we write $e_G(A)$ for the number of edges $e(G[A])$ inside A , and we write $e_G(A, B)$ for the number of edges between A and B . Abusing notation, for a vertex v we write $e_G(v, A)$ to mean $e_G(\{v\}, A)$, which is the size of the A -neighbourhood of v in G . We write $N(v)$ for the neighbourhood of v (that is, the set of vertices adjacent to v). A homomorphism ϕ from a multigraph H to a multigraph G consists of a map from the vertices of H to the vertices of G , and a map from the edges of H to the edges of G , such that whenever e is an edge of H between vertices x and y , the image $\phi(e)$ is an edge between the vertices $\phi(x)$ and $\phi(y)$.

We initially introduced X_H as the random variable that counts unlabelled copies of H (as is standard in this area), but for the proof it will be slightly more convenient to redefine X_H to count the number of *labelled* copies of H (injective homomorphisms from H into G). The labelled/unlabelled distinction is irrelevant for Theorem 1.2, because these two counts differ by a fixed multiplicative factor (the number of automorphisms of H).

We use standard asymptotic notation throughout. For functions $f = f(n)$ and $g = g(n)$ we write $f = O(g)$ to mean that there is a constant C such that $|f| \leq C|g|$, we write $f = \Omega(g)$ to mean there is a constant $c > 0$ such that $f \geq c|g|$ for sufficiently large n , we write $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $g = \omega(f)$ to mean that $f/g \rightarrow 0$ as $n \rightarrow \infty$. Unless stated otherwise, all asymptotics are as $n \rightarrow \infty$ (all other variables should be viewed as constant).

We will use notation of the form \mathbb{E}_G to indicate an expected value with respect to a random choice of G (if there are other sources of randomness, then formally this is a conditional expected value). We write $\text{Ber}(p)$ for the p -Bernoulli distribution, meaning that if $\xi \sim \text{Ber}(p)$ then $\Pr(\xi = 1) = p$ and $\Pr(\xi = 0) = 1 - p$. Finally, we write \mathbb{N} for the set of non-negative integers, we write $[n]$ for the set $\{1, \dots, n\}$, and all logarithms are to base e .

2 Discussion and main ideas of the proof

Before discussing the new ideas in the proof of Theorem 1.2, it is worth reviewing the proofs in [19] giving weaker anticoncentration bounds. We will build on these ideas to prove Theorem 1.2.

First, to prove the bound $\Pr(|X_H - x| \leq n^{v(H)-2}) = O(1/n)$, the key observation was that X_H is an “almost linear” function of the edges of the underlying random graph $G \sim \mathbb{G}(n, p)$. Specifically, for a pair of vertices $e = \{x, y\}$, the difference $\Delta X_H := X_H(G + e) - X_H(G - e)$ is tightly concentrated around its expectation $\mathbb{E}\Delta X_H = \Theta(n^{v(H)-2})$, meaning that adding or removing an edge typically causes X_H to increase or decrease by about this amount⁴. Using this observation, and some ideas from Erdős’ proof of the Erdős–Littlewood–Offord theorem [18] and Lubell’s proof of the LYM inequality [28], it is possible to prove that the anticoncentration behaviour of $X_H/\mathbb{E}\Delta X_H$ is about the same as the anticoncentration behaviour of the number of edges of G , which is a binomial random variable with parameters $\binom{n}{2}$ and p . This gives the desired bound, which in some sense gives “coarse scale” anticoncentration for X_H .

Second, to prove the bound $\Pr(X_{K_h} = x) \leq n^{1-h+o(1)}$ for cliques, the key idea was to fix a vertex v and write $X_{K_h}(G) = X_{K_h}(G - v) + X_{K_{h-1}}(G[N(v)])$. That is, the number of h -cliques in G is the same as the number of h -cliques in $G - v$, plus the number of $(h - 1)$ -cliques in the neighbourhood of v (which yield an h -clique when combined with v). Now, it is possible to use similar ideas as in the preceding paragraph to show that $X_{K_h}(G - v)$ is anticoncentrated at a coarse scale. On the other hand, $X_{K_{h-1}}(G[N(v)])$ has a much smaller order of magnitude, and it is actually concentrated on a relatively small interval around its expectation. Furthermore, we can bound the point probabilities for $X_{K_{h-1}}(G[N(v)])$ inductively.

⁴This observation is closely related to the fact that the *number* of edges in G is closely correlated with X_H . This fact can be used to prove a central limit theorem for X_H (see for example [23, Example 6.4]).

Then, roughly speaking, the idea was as follows: We want to bound the probability of having $X_{K_h}(G) = X_{K_h}(G - v) + X_{K_{h-1}}(G[N(v)]) = x$. Since $X_{K_{h-1}}(G[N(v)])$ is concentrated on a small interval around its expectation $\mathbb{E}X_{K_{h-1}}(G[N(v)])$, in order to have $X_{K_h}(G) = x$ the value of $X_{K_h}(G - v)$ must (typically) be reasonably close to $x - \mathbb{E}X_{K_{h-1}}(G[N(v)])$. Using the coarse scale anticoncentration bounds for $X_{K_h}(G - v)$, we can bound the probability that this happens. After knowing the value $X_{K_h}(G - v)$, we know what value $X_{K_{h-1}}(G[N(v)])$ needs to take to have $X_{K_h}(G) = x$. By induction, we can bound the probability that $X_{K_{h-1}}(G[N(v)])$ takes this particular value.

If $X_{K_h}(G - v)$ and $X_{K_{h-1}}(G[N(v)])$ were independent, it would be easy to conclude the desired anti-concentration bound $\Pr(X_{K_h} = x) \leq n^{1-h+o(1)}$; we would be able to simply multiply the above two probability estimates. Unfortunately, these random variables are not independent, so we need to rule out the possibility that the fluctuations in $X_{K_h}(G - v)$ “cancel out” the fluctuations in $X_{K_{h-1}}(G[N(v)])$ in a way that causes X_{K_h} to concentrate on a particular value. The approach we took was to show that actually $X_{K_{h-1}}(G[N(v)])$ is anticoncentrated even after conditioning on a typical outcome of $G - v$ (after which the only remaining randomness comes from the set of neighbours $N(v)$).

We proved a conditional anticoncentration bound of this type by induction, using a moment argument. To be more specific, we viewed the conditional probabilities $\Pr(X_{K_{h-1}}(G[N(v)]) = z \mid G - v)$ as random variables depending on $G - v$. To study these random variables we studied their high moments, which essentially comes down to considering collections of different candidates for the neighborhood $N(v)$, and bounding the probability that for each of these candidates we simultaneously have $X_{K_{h-1}}(G[N(v)]) = z$. We accomplished this with a multiple exposure argument: we iteratively went through our candidate sets for $N(v)$ and exposed the edges of $G - v$ inside each set which had not yet been exposed before. Using a suitable induction hypothesis, and the ideas sketched earlier, at each step we can bound the probability that the corresponding candidate set for $N(v)$ gives rise to the desired value of $X_{K_{h-1}}(G[N(v)])$. This way we obtained a suitable bound for the moment argument.

Now, there are several obstacles that need to be overcome to generalise the above argument beyond the case where H is a clique. First, the decomposition $X_{K_h} = X_{K_h}(G - v) + X_{K_{h-1}}(G[N(v)])$ was very convenient for us: it allowed us to consider two separate random variables, one of which can be studied on a “coarse” scale using our Littlewood–Offord type techniques, and the other of which can be studied inductively. Actually, it is not a huge problem to generalise this decomposition. In general, we have $X_H = X_H(G - v) + X_H^v$, where X_H^v is the number of copies of H in G which contain v . One can check that X_H^v is a certain sum of weighted subgraph counts in $G - v$, where each of the subgraphs has $v(H) - 1$ vertices, and the weight of a subgraph depends on its intersection with the neighbourhood $N(v)$ of v . So, if we generalise the induction hypothesis to certain weighted sums of subgraph counts, it is still possible to control the anticoncentration of X_H^v inductively.

The second main obstacle, which is more serious, concerns the multiple-exposure argument we used to show that $X_{K_{h-1}}(G[N(v)])$ is anticoncentrated even after conditioning on a typical outcome of $G - v$. This crucially depended on the fact that in order to know the value of $X_{K_{h-1}}(G[N(v)])$ given a particular candidate for $N(v)$, the only edges of $G - v$ that we need to expose are the edges inside $N(v)$ (leaving the remaining edges for future rounds of exposure). Unfortunately, in general one may need to examine all the edges of $G - v$ to determine the value of X_H^v , even if we fix a candidate for $N(v)$. Specifically, this is the case whenever H is not a complete multipartite graph (if H is complete multipartite, then we do not need to expose the edges which lie completely outside $N(v)$, and actually in this special case it is not too hard to extend the proof in [19] to prove Conjecture 1.1).

In the general case where H is not a complete multipartite graph, in order to use a moment argument as above, we need some other way to estimate the joint probability that many candidates for $N(v)$ each result in a specific outcome of X_H^v . Write $X_H^v(N(v), G - v)$ to indicate the dependence of X_H^v on both $N(v)$ and $G - v$. For a collection of sets A_1, \dots, A_t as candidates for $N(v)$, we want to control the joint probability that all of $X_H^v(A_1, G - v), \dots, X_H^v(A_t, G - v)$ are equal to a given value. Since we cannot consider these random variables separately anymore (as we did before with the multiple exposure argument), we need to somehow modify the induction hypothesis to handle this joint probability.

Specifically, we can generalise to a statement about joint anticoncentration probabilities of random variables of the following type (thus strengthening the induction hypothesis). Take a sequence of distinct vertices v_1, \dots, v_g and for each v_i , consider some collection of t_i different candidates for the neighbourhood of v_i . Then, consider the $T = t_1 \cdots t_g$ different random variables obtained by making different choices for the neighbourhoods for each of v_1, \dots, v_g , and considering the number of copies of H which contain all

of v_1, \dots, v_g , conditioned on v_1, \dots, v_g having these neighbourhoods. The idea is that at each step of the induction we introduce a new vertex v_j , and we consider many candidates for the neighbourhood of v_j for a moment argument. Given that we are considering joint probabilities of $T = t_1 \cdots t_g$ random variables, our induction hypothesis needs to give a bound of the form $n^{(g-h+1)T+o(1)}$ on the joint probability that all our random variables are equal to particular values. This can be viewed as an anticoncentration bound for a random *vector* \mathbf{X} .

To make the above ideas work, we need multivariate generalisations of some of the ideas described so far. For example, we need an anticoncentration inequality for random *vectors* that are almost-linear in the sense sketched earlier, having the property that adding or removing an edge typically causes a predictable change in their values. There are some classical anticoncentration inequalities by Halász [21] that give the kind of bounds we need, for random vectors that depend linearly (and “non-degenerately”) on a sequence of independent random choices. (Some kind of non-degeneracy assumption is necessary, to rule out situations where the random vector is always contained in a proper subspace of smaller dimension). The standard proofs of Halász’ inequalities are Fourier-analytic, and are not robust enough to apply to our almost-linear setting, but we were able to find some combinatorial arguments that apply to our setting, again inspired by proofs of Erdős [18] and Lubell [28]. More details are in Section 3.

In order to use the estimates in Section 3, we need to check a non-degeneracy condition: basically, we need to consider the effects of changing the status of various edges, and we need to show that the corresponding changes to \mathbf{X} are in “many different directions”, spanning \mathbb{R}^T . We also need to check a similar non-degeneracy condition for the effects of adding or removing vertices from the various candidate neighbourhoods. Unfortunately, these non-degeneracy conditions do not hold for an arbitrary connected graph H (they do hold, however, if H has a vertex with edges to all other vertices). Therefore, we actually need to further modify our approach.

Instead of considering a single vertex v_j in each step of the induction, we will consider a_j different vertices, having a diverse range of adjacencies to the vertices previously chosen. Then, our decomposition is that we split the copies of H into the copies that contain none of our a_j identified vertices, and the copies that contain at least one of them. With this modification, there is a much richer range of possibilities for the effect of changing the status of an edge, and this allows us to prove the desired non-degeneracy condition. Unfortunately, while this modification is conceptually rather simple, it complicates notation enormously. We now need to maintain a collection of sets of vertices, and a collection of possibilities for the neighbourhoods of these vertices. To encode all of these data we introduce the notion of a *colour system*: each step of the induction is associated with a different colour, and at each step there are multiple “shades” of each colour indicating the different possibilities for the neighbourhoods of the various vertices introduced at that step. We can then state our induction hypothesis for a class of random variables defined in terms of colour systems, and prove it using the ideas we discussed in this outline.

3 Anticoncentration for “almost-linear” random vectors

The Erdős–Littlewood–Offord theorem states that if ξ_1, \dots, ξ_n are independent Bernoulli random variables satisfying $\Pr(\xi_i = 0) = \Pr(\xi_i = 1) = 1/2$, and $X = a_1\xi_1 + \cdots + a_n\xi_n$ is a linear combination of these random variables (where each coefficient a_i has absolute value at least one) then for any $x \in \mathbb{R}$ we have $\Pr(|X - x| \leq 1) = O(1/\sqrt{n})$. As outlined in Section 2, in [19, Theorem 1.2] we adapted Erdős’ proof of this theorem to handle the case of “almost-linear” functions of ξ_1, \dots, ξ_n .

In [21], among other results, Halász gave a multivariate generalisation of the Erdős–Littlewood–Offord theorem, for sums of random vectors satisfying a certain non-degeneracy condition. Specifically, suppose that $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^d$ are d -dimensional vectors with the property that for every unit vector $\mathbf{e} \in \mathbb{R}^d$, there are $\Omega(n)$ indices i with $|\langle \mathbf{a}_i, \mathbf{e} \rangle| \geq 1$. Halász proved that, with $\mathbf{X} = \xi_1\mathbf{a}_1 + \cdots + \xi_n\mathbf{a}_n$, we have $\max_{\mathbf{x} \in \mathbb{R}^d} \Pr(\|\mathbf{X} - \mathbf{x}\|_2 \leq 1) = O(n^{-d/2})$. As outlined in Section 2, for the proof of Theorem 1.2 we will need a similar bound for almost-linear \mathbf{X} . Our non-degeneracy condition will be somewhat cruder than Halász’; we assume that there are vectors $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^d$, spanning \mathbb{R}^d , such that each of these vectors is represented $\Omega(n)$ times as the direction of the “typical effect” of changing the status of some ξ_i .

Theorem 3.1. *Fix real numbers $0 < p < 1$ and $0 < \varepsilon < 1$, integers $m \geq d \geq 1$, and vectors $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^d$ spanning \mathbb{R}^d . Then there is a constant $c > 0$, depending on p, ε, d and $\mathbf{v}_1, \dots, \mathbf{v}_m$, such that the following holds. For any positive integer n and any function $\mathbf{f} : \{0, 1\}^n \rightarrow \mathbb{R}^d$, let $\boldsymbol{\xi} \sim \text{Ber}(p)^n$ and for*

$i = 1, \dots, n$ define the random variables

$$\Delta_i \mathbf{f}(\boldsymbol{\xi}) = \mathbf{f}(\xi_1, \dots, \xi_{i-1}, 1, \xi_{i+1}, \dots, \xi_n) - \mathbf{f}(\xi_1, \dots, \xi_{i-1}, 0, \xi_{i+1}, \dots, \xi_n).$$

Suppose that for some positive real numbers r and s with $r\sqrt{n \log n} \geq s$ there are disjoint subsets $I_1, \dots, I_m \subseteq \{1, \dots, n\}$ of size at least εn such that for each $i \in I_j$ we have $\Pr(\|\Delta_i \mathbf{f}(\boldsymbol{\xi}) - s\mathbf{v}_j\|_\infty \geq r) \leq n^{-6d^2}$. Then for any $\mathbf{x} \in \mathbb{R}^d$ we have

$$\Pr(\|\mathbf{f}(\boldsymbol{\xi}) - \mathbf{x}\|_\infty < r\sqrt{n \log n}) \leq c \cdot \left(\frac{r\sqrt{\log n}}{s}\right)^d.$$

Roughly speaking, the assumption on the function \mathbf{f} in Theorem 3.1 means the following: For each of the vectors \mathbf{v}_j (with $1 \leq j \leq m$) there is a reasonably large subset $I_j \subseteq \{1, \dots, n\}$, such that for every $i \in I_j$ the following holds. When changing the i -th coordinate of $\boldsymbol{\xi}$ from 0 to 1 the corresponding vectors $\mathbf{f}(\boldsymbol{\xi})$ typically differ by roughly $s\mathbf{v}_j$ (more precisely, with high probability the difference of the corresponding vectors $\mathbf{f}(\boldsymbol{\xi})$ is close to the vector $s\mathbf{v}_j$). This condition can be seen as some sort of ‘‘almost-linearity’’ (at least with respect to certain coordinates of $\boldsymbol{\xi}$). Intuitively, this condition suggests that for a random vector $\boldsymbol{\xi} \sim \text{Ber}(p)^n$, the vector $\mathbf{f}(\boldsymbol{\xi})$ must be reasonably spread out and not too concentrated close to any given $\mathbf{x} \in \mathbb{R}^d$. Theorem 3.1 makes this intuition precise. The precise bound of n^{-6d^2} for the probability $\Pr(\|\Delta_i \mathbf{f}(\boldsymbol{\xi}) - s\mathbf{v}_j\|_\infty \geq r)$ in the assumptions of Theorem 3.1 was chosen for convenience in the proof of the theorem. The exponent $6d^2$ is certainly not sharp, but the precise value of this exponent is not relevant for the remainder of this paper, so we made no effort to optimise the exponent at the cost of complicating the proof of Theorem 3.1.

As mentioned above, Theorem 3.1 can be considered to be an analogue of Halász’ classical anticoncentration inequality for linear vector-valued functions [21] (satisfying a certain non-degeneracy condition), in the weaker setting of ‘‘almost-linear’’ functions. However, our non-degeneracy condition is somewhat more restrictive than the one in Halász’ original result. We also remark that an inequality very similar to Halász’ was also proved by Tao and Vu [39, Theorem 1.4], and that there is a large body of work proving similar results without a non-degeneracy condition (in which case the bounds are much weaker; see for example the survey in [31, Section 2]).

Before starting the proof of Theorem 3.1, we record the following basic fact about lattices.

Lemma 3.2. *Fix a basis $\mathbf{v}_1, \dots, \mathbf{v}_d \in \mathbb{R}^d$ of \mathbb{R}^d . There exists a constant c' , only depending on $\mathbf{v}_1, \dots, \mathbf{v}_d$, such that for any $\mathbf{x}' \in \mathbb{R}^d$ and any real number $z \geq 1$ there are at most $c' \cdot z^d$ different d -tuples of integers $(t_1, \dots, t_d) \in \mathbb{Z}^d$ with $\|(t_1\mathbf{v}_1 + \dots + t_d\mathbf{v}_d) - \mathbf{x}'\|_\infty < z$.*

We next prove Theorem 3.1, further developing the ideas in the proof of [19, Theorem 1.2]. As mentioned above, the assumption on \mathbf{f} in Theorem 3.1 means that when changing the i -th coordinate of $\boldsymbol{\xi}$ from zero to one for some $i \in I_j$, the vector $\mathbf{f}(\boldsymbol{\xi})$ typically changes by roughly $s\mathbf{v}_j$. This means that, if we successively change appropriately chosen zeros to ones in $\boldsymbol{\xi}$, we can (with sufficiently high probability) control the changes of the vector $\mathbf{f}(\boldsymbol{\xi})$, and show that $\mathbf{f}(\boldsymbol{\xi})$ cannot be too often close to any given vector $\mathbf{x} \in \mathbb{R}^d$. Indeed, if we change the coordinates of $\boldsymbol{\xi}$ with indices in $I_1 \cup \dots \cup I_m$, then the vector $\mathbf{f}(\boldsymbol{\xi})$ will typically move roughly along a lattice spanned by the vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$, and we can use Lemma 3.2 to show that there are not too many choices for $\boldsymbol{\xi}$ where $\|\mathbf{f}(\boldsymbol{\xi}) - \mathbf{x}\|_\infty < r\sqrt{n \log n}$.

Proof of Theorem 3.1. First, by relabelling the given vectors $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^d$, we may assume that $\mathbf{v}_1, \dots, \mathbf{v}_d$ form a basis of \mathbb{R}^d . We ignore the other vectors $\mathbf{v}_{d+1}, \dots, \mathbf{v}_m$ (in other words, we may assume that $m = d$).

Now, for $j = 1, \dots, d$, let $n_j = |I_j|$, so $\varepsilon n \leq n_j \leq n$. Let $\boldsymbol{\xi}^j = (\xi_i)_{i \in I_j}$ be the restriction of the random vector $\boldsymbol{\xi} \sim \text{Ber}(p)^n$ to the coordinates in I_j . Observe that the number $|\boldsymbol{\xi}^j|$ of ones in $\boldsymbol{\xi}^j$ is binomially distributed with parameters n_j and p . Therefore, for any $0 \leq t \leq n_j$ we have that

$$\Pr(|\boldsymbol{\xi}^j| = t) \leq c_p \cdot n_j^{-1/2} \leq (c_p/\sqrt{\varepsilon}) \cdot n^{-1/2} \quad (3.1)$$

for a constant $c_p > 0$ only depending on p .

Now, observe that $\Pr(|\xi^j| = t)$ is an increasing function in t for $t \leq pn_j$ and a decreasing function for $t \geq pn_j$. Thus, for each $j = 1, \dots, d$, there are integers $0 \leq a_j \leq b_j \leq n_j$ such that for any integer t we have

$$\Pr(|\xi^j| = t) > n^{-2d} \text{ if and only if } a_j \leq t \leq b_j. \quad (3.2)$$

That is to say, a_j and b_j are defined as the boundaries of the range of values that have probability at least n^{-2d} of occurring as $|\xi^j|$. We next bound the difference $b_j - a_j$.

The Chernoff bound (see for example [1, Theorem A.1.4]) yields $\Pr(|\xi^j| < pn_j - d\sqrt{n \log n}) \leq n^{-2d^2} \leq n^{-2d}$ and $\Pr(|\xi^j| > pn_j + d\sqrt{n \log n}) \leq n^{-2d^2} \leq n^{-2d}$. Thus, we must have

$$pn_j - d\sqrt{n \log n} \leq a_j \leq b_j \leq pn_j + d\sqrt{n \log n}$$

and in particular $b_j - a_j \leq 2d\sqrt{n \log n}$ for each $j = 1, \dots, d$. By the choice of a_j and b_j we have

$$\Pr(|\xi^j| < a_j \text{ or } |\xi^j| > b_j) = \sum_{t=0}^{a_j-1} \Pr(|\xi^j| = t) + \sum_{t=b_j+1}^{n_j} \Pr(|\xi^j| = t) \leq n \cdot n^{-2d} = n^{-2d+1} \quad (3.3)$$

for every $j = 1, \dots, d$.

Now, for each $j = 1, \dots, d$, let $\sigma_j : [n_j] \rightarrow I_j$ be a uniformly random bijection (independently chosen for each j). Also, independently sample $\chi_i \sim \text{Ber}(p)$ for each $i \in [n] \setminus (I_1 \cup \dots \cup I_d)$. For any integers $t_1, \dots, t_d \in [a_1, b_1] \times \dots \times [a_d, b_d]$, let the vector $\chi(t_1, \dots, t_d) \in \{0, 1\}^n$ be defined as follows. For $i \in [n] \setminus (I_1 \cup \dots \cup I_d)$, we already chose χ_i , the i -th entry of $\chi(t_1, \dots, t_d)$. If $i \in I_j$, then set $\chi_i = 1$ if and only if $i \in \sigma_j([t_j])$. In other words, among the entries χ_i for $i \in I_j$ there are precisely t_j ones and those are in positions $\sigma_j(1), \dots, \sigma_j(t_j)$.

For any given $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$, the random vector $\chi(t_1, \dots, t_d)$ depends on the choices of the bijections σ_j for $j = 1, \dots, d$ and the random entries $\chi_i \sim \text{Ber}(p)$ for $i \in [n] \setminus (I_1 \cup \dots \cup I_d)$. Very importantly, the distribution of $\chi(t_1, \dots, t_d)$ is the same as the distribution of the random vector $\xi \sim \text{Ber}(p)^n$ in the theorem statement conditioned on having $|\xi^j| = t_j$ for $j = 1, \dots, d$. In particular, fixing any $\mathbf{x} \in \mathbb{R}^d$, we have

$$\begin{aligned} \Pr\left(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{n \log n} \mid |\xi^j| = t_j \text{ for } j = 1, \dots, d\right) \\ = \Pr\left(\|\mathbf{f}(\chi(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right). \end{aligned}$$

Hence, using the independence of the random variables $|\xi^1|, \dots, |\xi^d|$ as well as (3.1), we obtain

$$\begin{aligned} \Pr\left(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{n \log n} \text{ and } |\xi^j| = t_j \text{ for } j = 1, \dots, d\right) \\ = \Pr\left(\|\mathbf{f}(\chi(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right) \cdot \prod_{j=1}^d \Pr(|\xi^j| = t_j) \\ \leq (c_p/\sqrt{\varepsilon})^d \cdot n^{-d/2} \cdot \Pr\left(\|\mathbf{f}(\chi(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right) \end{aligned}$$

for each $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$. On the other hand, (3.3) implies

$$\Pr(|\xi^j| < a_j \text{ or } |\xi^j| > b_j \text{ for some } 1 \leq j \leq d) \leq d \cdot n^{-2d+1} \leq n^{-d}.$$

(Note that to have $|I_j| \geq \varepsilon n > 0$ for $j = 1, \dots, d$, we must have $d \leq n$). Thus, we obtain

$$\begin{aligned} \Pr\left(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right) \\ \leq n^{-d} + \sum_{(t_1, \dots, t_d)} \Pr\left(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{n \log n} \text{ and } |\xi^j| = t_j \text{ for } j = 1, \dots, d\right) \\ \leq n^{-d} + (c_p/\sqrt{\varepsilon})^d \cdot n^{-d/2} \cdot \sum_{(t_1, \dots, t_d)} \Pr\left(\|\mathbf{f}(\chi(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right), \end{aligned}$$

where the sum is taken over all $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$. In other words,

$$\Pr\left(\|\mathbf{f}(\boldsymbol{\xi}) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right) \leq n^{-d} + (c_p/\sqrt{\varepsilon})^d \cdot n^{-d/2} \cdot \mathbb{E}Y, \quad (3.4)$$

where Y is the random variable counting the number of d -tuples $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ with $\|\mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}$. This random variable Y depends on the random choices of σ_j for $j = 1, \dots, d$ and on $\chi_i \sim \text{Ber}(p)$ for $i \in [n] \setminus (I_1 \cup \dots \cup I_d)$.

Note that we always have $Y \leq (n_1 + 1) \dots (n_d + 1) \leq (2n)^d$. Furthermore, recall that the distribution of $\boldsymbol{\chi}(t_1, \dots, t_d)$ is the same as the distribution of $\boldsymbol{\xi} \sim \text{Ber}(p)^n$ conditioned on having $|\boldsymbol{\xi}^j| = t_j$ for $j = 1, \dots, d$. This implies that for any $1 \leq j^* \leq d$, any $i \in I_{j^*}$ and any $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$, we have

$$\begin{aligned} \Pr(\|\Delta_i \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - s\mathbf{v}_{j^*}\|_\infty \geq r) &= \Pr\left(\|\Delta_i \mathbf{f}(\boldsymbol{\xi}) - s\mathbf{v}_{j^*}\|_\infty \geq r \mid |\boldsymbol{\xi}^j| = t_j \text{ for } j = 1, \dots, d\right) \\ &\leq \Pr(\|\Delta_i \mathbf{f}(\boldsymbol{\xi}) - s\mathbf{v}_{j^*}\|_\infty \geq r) \cdot \prod_{j=1}^d \Pr(|\boldsymbol{\xi}^j| = t_j)^{-1} \leq n^{-6d^2} \cdot \prod_{j=1}^d n^{2d} \leq n^{-4d^2} \leq n^{-2d-2}, \end{aligned}$$

where we used (3.2) and the assumption in the theorem that $\Pr(\|\Delta_i \mathbf{f}(\boldsymbol{\xi}) - s\mathbf{v}_{j^*}\|_\infty \geq r) \leq n^{-6d^2}$. Thus, by a union bound, the probability that there exist $1 \leq j^* \leq d$, $i \in I_{j^*}$ and $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ with $\|\Delta_i \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - s\mathbf{v}_{j^*}\|_\infty \geq r$, is at most $n^{-2d-2} \cdot d \cdot n \cdot n^d \leq n^{-d}$.

Now, fix $c' > 0$, only depending on $\mathbf{v}_1, \dots, \mathbf{v}_d \in \mathbb{R}^d$, as in Lemma 3.2.

Claim 3.3. *If $\|\Delta_i \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - s\mathbf{v}_{j^*}\|_\infty \leq r$ for all $1 \leq j^* \leq d$, all $i \in I_{j^*}$ and all $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$, then we have $Y \leq c' \cdot (2d^2 + 1)^d \cdot (r\sqrt{n \log n}/s)^d$*

Proof. Note that for any $1 \leq j^* \leq d$ and any $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ with $t_{j^*} < b_{j^*}$, the vectors $\boldsymbol{\chi}(t_1, \dots, t_{j^*-1}, t_{j^*} + 1, t_{j^*+1}, \dots, t_d)$ and $\boldsymbol{\chi}(t_1, \dots, t_d)$ only differ in that the first of these vectors has a one in position $\sigma_{j^*}(t_{j^*} + 1)$, while the second has a zero in that position. Hence

$$\mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_{j^*-1}, t_{j^*} + 1, t_{j^*+1}, \dots, t_d)) - \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) = \Delta_{\sigma_{j^*}(t_{j^*} + 1)} \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)).$$

As $\sigma_{j^*}(t_{j^*} + 1) \in I_{j^*}$, under the assumptions of the claim this implies

$$\|\mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_{j^*-1}, t_{j^*} + 1, t_{j^*+1}, \dots, t_d)) - \mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - s\mathbf{v}_{j^*}\|_\infty \leq r$$

for all $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ and all $1 \leq j^* \leq d$ with $t_{j^*} < b_{j^*}$. Adding this up for different values of $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ and using the triangle inequality, this implies that

$$\begin{aligned} \|\mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - \mathbf{f}(\boldsymbol{\chi}(a_1, \dots, a_d)) - (t_1 - a_1)s\mathbf{v}_1 - \dots - (t_d - a_d)s\mathbf{v}_d\|_\infty \\ \leq ((t_1 - a_1) + \dots + (t_d - a_d)) \cdot r \leq 2d^2 \sqrt{n \log n} \cdot r, \end{aligned}$$

for all $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$, where for the second inequality we used that $t_j - a_j \leq b_j - a_j \leq 2d\sqrt{n \log n}$ for each $1 \leq j \leq d$.

Recall that Y is the number of (integer) d -tuples $(t_1, \dots, t_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]$ which satisfy $\|\mathbf{f}(\boldsymbol{\chi}(t_1, \dots, t_d)) - \mathbf{x}\|_\infty < r\sqrt{n \log n}$. For each such d -tuple we then have (by the triangle inequality)

$$\|\mathbf{x} - \mathbf{f}(\boldsymbol{\chi}(a_1, \dots, a_d)) - (t_1 - a_1)s\mathbf{v}_1 - \dots - (t_d - a_d)s\mathbf{v}_d\|_\infty < (2d^2 + 1)\sqrt{n \log n} \cdot r,$$

and therefore

$$\left\| t_1 \mathbf{v}_1 + \dots + t_d \mathbf{v}_d - a_1 \mathbf{v}_1 - \dots - a_d \mathbf{v}_d + \frac{1}{s} \mathbf{f}(\boldsymbol{\chi}(a_1, \dots, a_d)) - \frac{1}{s} \mathbf{x} \right\|_\infty < (2d^2 + 1) \cdot \frac{\sqrt{n \log n} \cdot r}{s}.$$

Thus by Lemma 3.2 applied with $\mathbf{x}' = a_1 \mathbf{v}_1 + \dots + a_d \mathbf{v}_d - \frac{1}{s} \mathbf{f}(\boldsymbol{\chi}(a_1, \dots, a_d)) + \frac{1}{s} \mathbf{x}$ as well as $z = (2d^2 + 1)\sqrt{n \log n} \cdot r/s$ (note that $z \geq 1$ as $r\sqrt{n \log n} \geq s$ by the assumptions of the theorem), we obtain that

$$Y \leq c' \cdot (2d^2 + 1)^d \cdot \left(\frac{\sqrt{n \log n} \cdot r}{s} \right)^d,$$

as desired. \square

Just before Claim 3.3, we proved that its assumptions are satisfied with probability at least $1 - n^{-d}$. Thus, we obtain

$$\mathbb{E}Y \leq n^{-d} \cdot (2n)^d + c' \cdot (2d^2 + 1)^d \cdot \left(\frac{\sqrt{n \log n} \cdot r}{s} \right)^d \leq (c' + 1) \cdot (2d^2 + 1)^d \cdot \left(\frac{\sqrt{n \log n} \cdot r}{s} \right)^d,$$

using that $r\sqrt{n \log n} \geq s$. Plugging this into (3.4), we can conclude

$$\begin{aligned} \Pr\left(\|f(\boldsymbol{\xi}) - \mathbf{x}\|_\infty < r\sqrt{n \log n}\right) &\leq n^{-d} + (c_p/\sqrt{\varepsilon})^d \cdot n^{-d/2} \cdot (c' + 1) \cdot (2d^2 + 1)^d \cdot \left(\frac{\sqrt{n \log n} \cdot r}{s} \right)^d \\ &= n^{-d} + (c_p/\sqrt{\varepsilon})^d \cdot (c' + 1) \cdot (2d^2 + 1)^d \cdot \left(\frac{r\sqrt{\log n}}{s} \right)^d \leq ((c_p/\sqrt{\varepsilon})^d \cdot (c' + 1) \cdot (2d^2 + 1)^d + 1) \cdot \left(\frac{r\sqrt{\log n}}{s} \right)^d, \end{aligned}$$

where in the last inequality we used that $r\sqrt{\log n}/s \geq n^{-1/2} \geq n^{-1}$. This implies the statement of Theorem 3.1 with $c = (c_p/\sqrt{\varepsilon})^d \cdot (c' + 1) \cdot (2d^2 + 1)^d + 1$. \square

4 Colour systems and the induction hypothesis

As outlined in Section 2, our proof of Theorem 1.2 proceeds via induction over a class of random variables generalising subgraph counts. These random variables are defined via *colour systems*.

Definition 4.1. For integers $g \geq 0$ and $a_1, \dots, a_g, t_1, \dots, t_g \geq 1$, a *colour system* \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ is a multigraph without loops which is coloured according to the following rules.

- Each vertex has at most one colour and for each $1 \leq i \leq g$, there are exactly a_i vertices of colour i .
- Each edge is incident to at least one coloured vertex.
- Each edge has exactly one colour. If an edge is incident to exactly one coloured vertex, it receives the colour of that vertex. If an edge is incident to two coloured vertices, and these vertices have colours i_1 and i_2 , then the edge has colour $\min(i_1, i_2)$.
- Each edge of colour i is additionally labelled with an integer in $\{1, \dots, t_i\}$ (we say that there are t_i different *shades*) of colour i . We do not assign shades to vertices, only edges.
- Between any two vertices, there is at most one edge of each shade of each colour (but there can be multiple edges of different shades of the same colour).

The *order* of a colour system \mathcal{G} is its total number of vertices (both coloured and uncoloured).

For a colour system \mathcal{G} , let $U(\mathcal{G})$ be the set of uncoloured vertices of \mathcal{G} . In most of the statements throughout the paper, we will consider colour systems whose order n is large with respect to the parameters $g, a_1, \dots, a_g, t_1, \dots, t_g$ (in which case almost all vertices of \mathcal{G} are uncoloured). In all of our statements involving asymptotic notation, the colour system parameters $g, a_1, \dots, a_g, t_1, \dots, t_g$ will be treated as fixed constants for the asymptotic notation, while the order n tends to infinity.

To make some sense of the definition of a colour system, recall from the outline in Section 2 that our proof is inductive, and at each step of the induction we consider multiple possibilities for the neighbourhoods of certain vertices. The g colours in a colour system correspond to the vertices chosen at the g different steps of the induction, and the t_i different shades of colour i correspond to t_i different choices of neighbourhoods for the vertices of colour i .

Now, we will mostly want to consider colour systems which have “typical” structure, meaning that the sizes of intersections between neighbourhoods of vertices are about what one should expect if the neighbourhoods were chosen randomly. For this, we introduce a notion of “general position” for families of sets.

Definition 4.2. Consider subsets S_1, \dots, S_m of some ground set R . For any subset $I \subseteq \{1, \dots, m\}$, we write $S_I = R \cap \bigcap_{i \in I} S_i \cap \bigcap_{i \notin I} (R \setminus S_i)$. For an integer $K \geq 1$ and some $0 < p < 1$, we say that $S_1, \dots, S_m \subseteq R$ are in (p, K) -general position if for each $I \subseteq \{1, \dots, m\}$, we have

$$\left| |S_I| - p^{|I|} (1-p)^{m-|I|} |R| \right| \leq K \cdot |R|^{1/2} \log |R|.$$

Note that if $m = 0$, then $S_\emptyset = R$ and therefore in this case the empty collection of sets is in (p, K) -general position for every integer $K \geq 1$.

Definition 4.3. Say that a colour system \mathcal{G} is p -general if the following holds. If we define $S_1, \dots, S_m \subseteq U(\mathcal{G})$ to be the $U(\mathcal{G})$ -neighbourhoods of each of the coloured vertices of \mathcal{G} in each of the shades of the respective colour (so we have $m = a_1 \cdot t_1 + \dots + a_g \cdot t_g$ if \mathcal{G} has parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$), then the sets $S_1, \dots, S_m \subseteq U(\mathcal{G})$ are in $(p, 3^g)$ -general position.

Furthermore, say that the colour system \mathcal{G} is *weakly* p -general if the sets $S_1, \dots, S_m \subseteq U(\mathcal{G})$ are in $(p, 2 \cdot 3^g)$ -general position.

Note that every p -general colour system is in particular weakly p -general (the reason for having both these definitions is that when we make small changes to a collection of sets in (p, K) -general position, the parameter K changes slightly, and it is convenient to not have to explicitly keep track of this change). Also, note that for $g = 0$, every colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ is p -general (since $m = 0$ and the empty collection of sets is always in (p, K) -general position for all $K \geq 1$).

Recall from Section 2 that the whole point of introducing multiple vertices at each step of the induction is to allow for a richer range of possibilities for the effect of changing the status of an edge. In order to ensure the richest possible range of possibilities, we consider colour systems which are *complete* in the sense that we see essentially all the possible adjacencies between the coloured vertices, as follows.

Definition 4.4. Call a colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ complete, if for any $1 \leq i \leq g$ the following holds. Suppose for each $1 \leq j \leq i-1$ and each vertex v in \mathcal{G} of colour j we are given a subset $I_v \subseteq [t_j]$. Then there exists a vertex w of colour i such that for every $1 \leq j \leq i-1$ and every vertex v of colour j the vertices w and v are connected by edges of exactly those shades of colour j that belong to the set I_v .

Informally speaking, Definition 4.4 demands that for every colour i we can find a vertex with prescribed edges to all the vertices of the previous (smaller) colours. For $g = 0$ every colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ is complete, as the condition in Definition 4.4 is vacuous.

Note that whether a colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ is complete only depends on the edges between the coloured vertices, and it does not at all depend on the edges in colour g . In contrast, whether \mathcal{G} is p -general for given $0 < p < 1$ only depends on the edges between the coloured and the uncoloured vertices.

Now, to obtain a graph from a colour system, we choose a shade of each colour, and we choose a graph on the uncoloured vertices, as follows.

Definition 4.5. Let \mathcal{G} be a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$. Then, given a g -tuple $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_g]$, and a graph G_0 on the vertex set $U(\mathcal{G})$, define a graph $\mathcal{G}(G_0, j_1, \dots, j_g)$ by taking all vertices of \mathcal{G} together with all edges of G_0 and all edges of shade j_i of colour i for all $1 \leq i \leq g$. Furthermore, for a graph H let $\psi_H(\mathcal{G}, G_0, j_1, \dots, j_g)$ be the number of labelled copies of H in the graph $\mathcal{G}(G_0, j_1, \dots, j_g)$ which use at least one vertex of each of the g colours.

We will use notation such as $\psi_H(\mathcal{G}, G_0, \cdot)$ to denote the function $[t_1] \times \dots \times [t_g] \rightarrow \mathbb{Z}$ that maps (j_1, \dots, j_g) to $\psi_H(\mathcal{G}, G_0, j_1, \dots, j_g)$. Our goal for the rest of this paper will be to prove the following strengthening of Theorem 1.2.

Theorem 4.6. Fix some $0 < p < 1$, an h -vertex graph H , and integers $0 \leq g \leq h-1$ and $a_1, \dots, a_g, t_1, \dots, t_g \geq 1$. Let $T = t_1 \cdots t_g$. Then for any p -general complete colour system \mathcal{G} of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$ and for any function $\lambda : [t_1] \times \dots \times [t_g] \rightarrow \mathbb{Z}$ the following holds. If $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ is a random graph on the vertex set $U(\mathcal{G})$, then

$$\Pr\left(\psi_H(\mathcal{G}, G_0, \cdot) = \lambda\right) \leq n^{(g-h+1) \cdot T + o(1)}.$$

The $o(1)$ -term in Theorem 4.6 goes to zero as n tends to infinity, but it may depend on the choices for p , h , H , and $g, a_1, \dots, a_g, t_1, \dots, t_g$ fixed in the beginning of the theorem statement (in other words, these values are treated as constants in the asymptotics).

Note that for $g = 0$, in Theorem 4.6 we have $T = 1$ (using the convention that the empty product is equal to 1), and the colour system \mathcal{G} has no coloured vertices (so it consists of n uncoloured vertices and no edges). We already observed that such a colour system is always p -general and complete. Furthermore, $\psi_H(\mathcal{G}, G_0)$ is simply the number of labelled copies of H in $G_0 \sim \mathbb{G}(n, p)$. This quantity is precisely the random variable X_H in Theorem 1.2. Thus, Theorem 4.6 for $g = 0$ states that for all $\ell \in \mathbb{Z}$ we have $X_H = \ell$ with probability at most $n^{(1-h) \cdot 1 + o(1)} = n^{1-h+o(1)}$. This is precisely the statement of Theorem 1.2.

So, Theorem 1.2 corresponds to the case $g = 0$ in Theorem 4.6, and it therefore suffices to prove Theorem 4.6. We will use backwards induction starting from $g = h - 1$ and going down to $g = 0$. Note that the case $g = h - 1$ is trivial.

For the rest of the paper, we fix a particular graph H with h vertices and some $0 < p < 1$. Before concluding this section, we make a few more definitions and state an important intermediate result for the induction step (Proposition 4.10, to follow). Basically, at each step of the induction, we have a colour system \mathcal{G} with $g - 1$ colours, and we add vertices of a new colour with random neighbourhoods, obtaining a random colour system \mathcal{G}_S . We will use the “ g ” case of Theorem 4.6 and a moment argument to show that typically G_0 has the property that $\psi(\mathcal{G}_S, G_0, \cdot)$ is anticoncentrated, subject only to the randomness in \mathcal{G}_S . This will be the content of Proposition 4.10.

Definition 4.7. For integers $g \geq 1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$, define a *restricted colour system* with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ to be a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$ in which there are no edges in colour g between the coloured and the uncoloured vertices (so all edges of colour g are between the vertices of colour g , recalling that the colour of an edge is the minimum of the colours of its endpoints).

Call a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ *complete*, if it is complete when viewed as a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$. Call a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ *essentially p -general*, if the colour system with parameters $(g - 1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ obtained by ignoring colour g is p -general. Similarly, call \mathcal{G} *essentially weakly p -general*, if the colour system obtained by ignoring colour g is weakly p -general.

Definition 4.8. Consider a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. For each vertex v of colour g in \mathcal{G} , choose a random subset $S_v \subseteq U(\mathcal{G})$ by taking each element of $U(\mathcal{G})$ independently with probability p (and choose the different sets S_v all independent from each other). Let $\mathcal{S} = (S_v)_v$ be the collection of random sets chosen this way. Then we can obtain a random colour system \mathcal{G}_S with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$ by connecting each vertex v of colour g to all vertices in S_v and colouring all these new edges in the unique shade of colour g .

Definition 4.9. Consider $0 < q < 1$ and an essentially p -general complete restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. We say that a graph G_0 on the vertex set $U(\mathcal{G})$ is (p, q, \mathcal{G}) -*dispersed* if for all functions $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$ the following holds. When choosing \mathcal{S} randomly as in Definition 4.8, we have $\Pr(\psi_H(\mathcal{G}_S, G_0, \cdot) = \lambda) \leq q$.

Now we are ready to state Proposition 4.10, as announced above.

Proposition 4.10. Fix integers $1 \leq g \leq h - 1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$ and let $T = t_1 \cdots t_{g-1}$. Then there exists a function $\delta : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ with $\lim_{n \rightarrow \infty} \delta(n) = 0$, such that for any essentially p -general complete restricted colour system \mathcal{G} of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ the following holds. If $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ is a random graph on the vertex set $U(\mathcal{G})$, then with probability $1 - n^{-\omega(1)}$ the graph G_0 is (p, q, \mathcal{G}) -dispersed, where $q = n^{(g-h+(1/2)) \cdot T + \delta(n)}$.

We remark that in the case $g = 1$, we have $T = 1$ in Proposition 4.10 (again by the convention that the empty product is equal to 1).

The rest of the paper is organised as follows. In Section 5 we give some very straightforward lemmas about sets in general position (in particular, random collections of sets are very likely to be in general

position). In Section 6 we use a moment argument to prove that the “ g ” case of Theorem 4.6 implies the corresponding case of Proposition 4.10. In Section 7 we show how this case of Proposition 4.10 implies the “ $g - 1$ ” case of Theorem 4.6, completing the induction step. The contents of both of these sections consist mostly of calculations and putting various pieces together. However, in both these sections we omit the proofs of important anticoncentration lemmas (Lemmas 6.3 and 7.3). The rest of the paper is spent proving these lemmas via our new multivariate anticoncentration inequality in Theorem 3.1. In Section 8 we introduce the formalism of a “core”, which will be used for the proofs of Lemmas 6.3 and 7.3. To be specific, we prove a linear independence lemma for certain vectors defined in terms of cores, which we will use to check the linear independence condition in Theorem 3.1. Finally, in Section 9 we put everything together to prove Lemmas 6.3 and 7.3.

5 Sets in general position

In this section we record some simple lemmas regarding sets in general position. Recall that in the last section we fixed some $p \in (0, 1)$.

Lemma 5.1. *Suppose that $S_1, \dots, S_m \subseteq R$ are subsets of some ground set R which are in (p, K) -general position, for some integer $K \geq 1$. Then for every subset $I \subseteq \{1, \dots, m\}$, we have*

$$\left| \left| \bigcap_{i \in I} S_i \right| - p^{|I|} |R| \right| \leq 2^m \cdot K \cdot |R|^{1/2} \log |R|.$$

Proof. Fix some subset $I \subseteq \{1, \dots, m\}$. For every subset $J \subseteq \{1, \dots, m\}$ with $I \subseteq J$, let $S_J = R \cap \bigcap_{i \in J} S_i \cap \bigcap_{i \notin J} (R \setminus S_i)$ be as in Definition 4.2. Note that there are at most 2^m such subsets J . Also, note that the set $\bigcap_{i \in I} S_i$ is the union of all the set S_J for all $J \subseteq \{1, \dots, m\}$ with $I \subseteq J$, and all these sets S_J are disjoint. Hence

$$\left| \bigcap_{i \in I} S_i \right| = \sum_J |S_J|.$$

Here, and in the rest of this proof, the sum is over all subsets $J \subseteq \{1, \dots, m\}$ with $I \subseteq J$. Noting that

$$p^{|I|} |R| = p^{|I|} (p + (1 - p))^{m - |I|} |R| = \sum_J p^{|J|} (1 - p)^{m - |J|} |R|,$$

we obtain that

$$\begin{aligned} \left| \left| \bigcap_{i \in I} S_i \right| - p^{|I|} |R| \right| &= \left| \sum_J |S_J| - \sum_J p^{|J|} (1 - p)^{m - |J|} |R| \right| \leq \sum_J \left| |S_J| - p^{|J|} (1 - p)^{m - |J|} |R| \right| \\ &\leq \sum_J K \cdot |R|^{1/2} \log |R| \leq 2^m \cdot K \cdot |R|^{1/2} \log |R|, \end{aligned}$$

as desired. \square

Lemma 5.2. *Fix some $m \in \mathbb{N}$. Let R be an n -element set and let $S_1, \dots, S_m \subseteq R$ be in (p, K) -general position for some integer $K \geq 1$. Let S_{m+1} be a random set chosen by taking each element of R independently with probability p . Then with probability $1 - n^{-\omega(1)}$ the sets $S_1, \dots, S_m, S_{m+1} \subseteq R$ are in (p, K) -general position.*

Proof. For any subset $I \subseteq \{1, \dots, m\}$, let $S_I = R \cap \bigcap_{i \in I} S_i \cap \bigcap_{i \notin I} (R \setminus S_i)$ be as in Definition 4.2. We need to show that with probability $1 - n^{-\omega(1)}$ we have

$$\left| |S_I \cap S_{m+1}| - p^{|I|+1} (1 - p)^{m - |I|} n \right| \leq K \cdot n^{1/2} \log n. \quad (5.1)$$

and

$$\left| |S_I \cap (R \setminus S_{m+1})| - p^{|I|} (1 - p)^{m - |I|+1} n \right| \leq K \cdot n^{1/2} \log n. \quad (5.2)$$

for each $I \subseteq \{1, \dots, m\}$. By a union bound over all 2^m choices for I , it suffices that for each individual set $I \subseteq \{1, \dots, m\}$ each of these properties holds with probability $1 - n^{-\omega(1)}$.

Fix some $I \subseteq \{1, \dots, m\}$. By assumption, the set $S_I \subseteq R$ satisfies

$$\left| |S_I| - p^{|I|}(1-p)^{m-|I|}n \right| \leq K \cdot n^{1/2} \log n. \quad (5.3)$$

Note that $|S_I \cap S_{m+1}|$ is binomially distributed with parameters $(|S_I|, p)$. Thus, the probability to have

$$\left| |S_I \cap S_{m+1}| - p|S_I| \right| \leq (1-p) \cdot n^{1/2} \log n \quad (5.4)$$

is by the Chernoff bound at least

$$1 - 2 \exp\left(-\frac{((1-p) \cdot n^{1/2} \log n)^2}{|S_I|}\right) \geq 1 - \exp(-\Omega((\log n)^2)) = 1 - n^{-\omega(1)},$$

where we used that $|S_I| \leq n$ as $S_I \subseteq R$. Whenever (5.4) is satisfied, then together with (5.3) we obtain by the triangle inequality

$$\left| |S_I \cap S_{m+1}| - p^{|I|+1}(1-p)^{m-|I|}n \right| \leq (1-p) \cdot n^{1/2} \log n + p \cdot K \cdot n^{1/2} \log n \leq K \cdot n^{1/2} \log n,$$

as desired. Thus (5.1) indeed holds with probability $1 - n^{-\omega(1)}$. Similarly, we can show that (5.2) holds with probability $1 - n^{-\omega(1)}$ by considering the probability that

$$\left| |S_I \cap (R \setminus S_{m+1})| - (1-p)|S_I| \right| \leq p \cdot n^{1/2} \log n$$

holds, using that $|S_I \cap (R \setminus S_{m+1})|$ is binomially distributed with parameters $(|S_I|, 1-p)$. \square

Lemma 5.3. *Fix some positive integer ℓ . Then the following holds for sufficiently large n . Let R be an n -element set and let $S_1, \dots, S_m \subseteq R$ be in (p, K) -general position for some integer $K \geq 1$. Let K' be an integer with $K' > K$. Then for any subset $R' \subseteq R$ obtained by deleting ℓ elements from R , the sets $S_1 \cap R', \dots, S_m \cap R' \subseteq R'$ are in (p, K') -general position.*

Proof. We need to show that for every subset $I \subseteq \{1, \dots, m\}$ we have

$$\left| |S_I \cap R'| - p^{|I|}(1-p)^{m-|I|}|R'| \right| \leq K' \cdot |R'|^{1/2} \log |R'|.$$

So fix some $I \subseteq \{1, \dots, m\}$. By assumption, the set $S_I \subseteq R$ satisfies

$$\left| |S_I| - p^{|I|}(1-p)^{m-|I|}|R| \right| \leq K \cdot |R|^{1/2} \log |R|.$$

Now, $S_I \cap R'$ is obtained from S_I by deleting at most ℓ elements, and therefore

$$\left| |S_I \cap R'| - |S_I| \right| \leq \ell.$$

Furthermore, we have

$$\left| p^{|I|}(1-p)^{m-|I|}|R| - p^{|I|}(1-p)^{m-|I|}|R'| \right| = p^{|I|}(1-p)^{m-|I|} \left| |R| - |R'| \right| = p^{|I|}(1-p)^{m-|I|} \ell \leq \ell.$$

Thus, the triangle inequality yields (using $|R'| = |R| - \ell$)

$$\left| |S_I \cap R'| - p^{|I|}(1-p)^{m-|I|}|R'| \right| \leq K \cdot |R|^{1/2} \log |R| + 2\ell \leq K' \cdot |R'|^{1/2} \log |R'|.$$

as long as $n = |R|$ is sufficiently large with respect to ℓ . \square

Corollary 5.4. *Fix $g, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1}$, and suppose n is sufficiently large with respect to these values. Let \mathcal{G} be a p -general colour system of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$. Then, whenever we delete $2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ vertices in $U(\mathcal{G})$, the resulting colour system is weakly p -general.*

Proof. Let $S_1, \dots, S_m \subseteq U(\mathcal{G})$ be the $U(\mathcal{G})$ -neighbourhoods of each of the coloured vertices of \mathcal{G} in each shade of their respective colours. Then $S_1, \dots, S_m \subseteq U(\mathcal{G})$ are in $(p, 3^{g-1})$ -general position. Thus, by Lemma 5.3, whenever we delete $2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ vertices in $U(\mathcal{G})$, the resulting configuration of sets is in $(p, 2 \cdot 3^{g-1})$ -general position. \square

Lemma 5.5. *Fix $g, a_1, \dots, a_g, t_1, \dots, t_{g-1}$. If \mathcal{G} is an essentially weakly p -general restricted colour system of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, then the colour system $\mathcal{G}_{\mathcal{S}}$ in Definition 4.8 is p -general with probability $1 - n^{-\omega(1)}$.*

Proof. Let \mathcal{G}' be the colour system of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ obtained from the restricted colour system \mathcal{G} by ignoring colour g (so $U(\mathcal{G}')$ consists of the set $U(\mathcal{G})$, together with the a_g vertices of colour g in \mathcal{G}). By the assumption on \mathcal{G} , the colour system \mathcal{G}' is weakly p -general. Let $S_1, \dots, S_m \subseteq U(\mathcal{G}')$ be the $U(\mathcal{G}')$ -neighbourhoods of each of the coloured vertices of \mathcal{G}' in each shade of their respective colours (so we have $m = a_1 \cdot t_1 + \dots + a_{g-1} \cdot t_{g-1}$). Then the sets $S_1, \dots, S_m \subseteq U(\mathcal{G}')$ are in $(p, 2 \cdot 3^{g-1})$ -general position. Note that $|U(\mathcal{G})| = n - a_1 + \dots + a_g = n - O(1)$ and $|U(\mathcal{G}')| - |U(\mathcal{G})| = a_g = O(1)$, and that $3^g > 2 \cdot 3^{g-1}$, so by Lemma 5.3 (assuming n is large), the sets $S_1 \cap U(\mathcal{G}), \dots, S_m \cap U(\mathcal{G}) \subseteq U(\mathcal{G})$ are in $(p, 3^g)$ -general position. Applying Lemma 5.2 a_g times, we see that with probability at least $1 - a_g \cdot |U(\mathcal{G})|^{-\omega(1)} = 1 - n^{-\omega(1)}$ these sets together with the random sets in \mathcal{S} are in $(p, 3^g)$ -general position. This proves Lemma 5.5. \square

6 Random neighbourhoods: Theorem 4.6 implies Proposition 4.10

In this section we will prove that if Theorem 4.6 holds for some $1 \leq g \leq h-1$, then Proposition 4.10 also holds for this value of g .

Fix some $1 \leq g \leq h-1$ and assume that Theorem 4.6 holds for this value of g . Let $a_1, \dots, a_g, t_1, \dots, t_{g-1}$ be arbitrary positive integers. Our goal in this section is to prove Proposition 4.10 for these values of $g, a_1, \dots, a_g, t_1, \dots, t_{g-1}$.

For the entirety of this section, we fix these values of $g, a_1, \dots, a_g, t_1, \dots, t_{g-1}$, and define $T = t_1 \cdots t_{g-1}$. In all our asymptotics, these values will be treated as constants, while $n \rightarrow \infty$.

In Section 6.1 we will start with some preparations. First, we use a martingale concentration inequality to prove that $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot)$ is very likely to be close to its conditional expectation $\mu_{\mathcal{G}, \mathcal{S}}$ given \mathcal{S} (by symmetry, this conditional expectation actually only depends on the sizes of the intersections between the sets in \mathcal{S} and the neighbourhoods of the various coloured vertices). Second, we state (but do not yet prove) an anticoncentration lemma for $\mu_{\mathcal{G}, \mathcal{S}}$, subject to the randomness in \mathcal{S} .

In Section 6.2 we will use our preparatory lemmas and Theorem 4.6 to prove an anticoncentration bound for certain joint probabilities concerning the values of $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot)$ for different choices of \mathcal{S} . This will be the key input for a moment argument (as outlined in Section 2) with which we will deduce that G_0 is very likely to be dispersed, proving the desired case of Proposition 4.10. The details of this deduction will be presented in Section 6.3.

6.1 Preparations

Definition 6.1. For a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ and an outcome of the random collection of sets $\mathcal{S} = (S_v)_v$ in Definition 4.8, let $\mu_{\mathcal{G}, \mathcal{S}} : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$ be the function given by

$$\mu_{\mathcal{G}, \mathcal{S}}(j_1, \dots, j_g) = \mathbb{E}_{G_0}[\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)]$$

for all $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$. Here, $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ is a random graph on the vertex set $U(\mathcal{G})$.

We remark that $\mu_{\mathcal{G}, \mathcal{S}}$ only depends on \mathcal{G} and \mathcal{S} , and not G_0 .

Lemma 6.2. *For any essentially p -general complete restricted colour system \mathcal{G} which has parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, the following holds. If we choose a random graph $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ and independently choose \mathcal{S} randomly as in Definition 4.8, then*

$$\Pr(\|\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot) - \mu_{\mathcal{G}, \mathcal{S}}\|_{\infty} \leq n^{h-g-1} \cdot \log n) = 1 - n^{-\omega(1)}.$$

Proof. Condition on an arbitrary outcome of \mathcal{S} . By a union bound, it suffices to prove that for each $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$ the probability that

$$|\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g) - \mu_{\mathcal{G}, \mathcal{S}}(j_1, \dots, j_g)| > n^{h-g-1} \cdot \log n \quad (6.1)$$

is $n^{-\omega(1)}$. So fix some $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$. The expectation of $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ is precisely $\mu_{\mathcal{G}, \mathcal{S}}(j_1, \dots, j_g)$ (recall that we are conditioning on an outcome of \mathcal{S}).

Note that changing the status of an edge of G_0 changes $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ by at most $O(n^{h-g-2})$. This is because there are at most $h^{g+2} \cdot a_1 \dots a_g \cdot n^{h-g-2} = O(n^{h-g-2})$ different labelled copies of H in the n -vertex graph $\mathcal{G}_{\mathcal{S}}(G_0, j_1, \dots, j_g)$ which contain any particular pair of vertices of $\mathcal{U}(\mathcal{G})$, and contain at least one vertex of each of the g colours. Thus, by the Azuma–Hoeffding inequality (see for example [1, Theorem 7.2.1]) with the edge-exposure martingale, the probability that (6.1) occurs is at most

$$\exp\left(-\Omega\left(\frac{n^{2h-2g-2} \cdot (\log n)^2}{n^2 \cdot n^{2(h-g-2)}}\right)\right) = \exp\left(-\Omega((\log n)^2)\right) = n^{-\omega(1)}.$$

This finishes the proof of Lemma 6.2. \square

Recall that we fixed g, h, a_1, \dots, a_g and t_1, \dots, t_{g-1} throughout this section, and that we defined $T = t_1 \dots t_{g-1}$.

Lemma 6.3. *For any essentially p -general complete restricted colour system \mathcal{G} of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ and any $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$ the following holds. If we choose \mathcal{S} randomly as in Definition 4.8 then we have*

$$\Pr(\|\mu_{\mathcal{G}, \mathcal{S}} - \lambda\|_{\infty} \leq n^{h-g-1} \cdot \log n) \leq n^{-T/2+o(1)}.$$

We defer the proof of Lemma 6.3 to Section 9.

6.2 Joint anticoncentration

The following statement is the key lemma for the moment argument we will use to finish the proof of Proposition 4.10.

Lemma 6.4. *Fix $t \in \mathbb{N}$. Then for any essentially p -general complete restricted colour system \mathcal{G} of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, and any function $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$, the following holds. If we choose, all independently, a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ on the vertex set $\mathcal{U}(\mathcal{G})$ as well as t random collections $\mathcal{S}_1, \dots, \mathcal{S}_t$ chosen as in Definition 4.8, then*

$$\Pr(\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for all } i = 1, \dots, t) \leq n^{(g-h+(1/2)) \cdot T \cdot t + o(1)}.$$

Proof. Fix an essentially p -general complete restricted colour system \mathcal{G} of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, and a function $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$.

For an outcome of the random collections $\mathcal{S}_1, \dots, \mathcal{S}_t$, we obtain a colour system \mathcal{G}^* with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, t)$ as follows. Recall that by Definition 4.7, \mathcal{G} has only one shade of colour g and all edges of colour g are between the vertices of colour g . Replace each of these edges by t parallel edges with all shades $\{1, \dots, t\}$ of colour g . Also, for each $i = 1, \dots, t$ and each vertex v with colour g , add edges, in shade i of colour g , between v and all vertices in the set $S_{i,v}$ (where $S_{i,v}$ is the set corresponding to the vertex v in the collection \mathcal{S}_i). We emphasise that \mathcal{G}^* only depends on the random choices of $\mathcal{S}_1, \dots, \mathcal{S}_t$, and not on the random choice of $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$.

Note that $\mathcal{U}(\mathcal{G}^*) = \mathcal{U}(\mathcal{G})$ and that for any $i = 1, \dots, t$, deleting all edges of colour g except those of shade i yields precisely the colour system $\mathcal{G}_{\mathcal{S}_i}$ with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$. Thus, for any outcome of $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$, we have

$$\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, j_1, \dots, j_{g-1}, 1) = \psi_H(\mathcal{G}^*, G_0, j_1, \dots, j_{g-1}, i) \quad (6.2)$$

for all $(j_1, \dots, j_{g-1}) \in [t_1] \times \dots \times [t_{g-1}]$ and all $i = 1, \dots, t$. Also note that \mathcal{G}^* is always complete, because \mathcal{G} is complete and being complete does not depend on the edges of colour g .

Say that an outcome of $(G_0, \mathcal{S}_1, \dots, \mathcal{S}_t)$ is *common* if both of the following conditions hold.

(i) we have

$$\|\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) - \mu_{\mathcal{G}, \mathcal{S}_i}\|_\infty \leq n^{h-g-1} \cdot \log n$$

for all $i = 1, \dots, t$;

(ii) the colour system \mathcal{G}^* given by \mathcal{G} and $\mathcal{S}_1, \dots, \mathcal{S}_t$ is p -general.

Claim 6.5. $(G_0, \mathcal{S}_1, \dots, \mathcal{S}_t)$ is common with probability $1 - n^{-\omega(1)}$.

Proof. By Lemma 6.2, for each $i = 1, \dots, t$ the probability that (i) fails is at most $n^{-\omega(1)}$. Thus, (i) holds with probability at least $1 - t \cdot n^{-\omega(1)} = 1 - n^{-\omega(1)}$.

For (ii), note that by Lemma 5.5 the colour system $\mathcal{G}_{\mathcal{S}_i}$ is p -general with probability $1 - n^{-\omega(1)}$. We can then apply Lemma 5.2 $(t-1) \cdot a_g$ times for all the additional sets in $\mathcal{S}_2, \dots, \mathcal{S}_t$. \square

Now, in order to prove Lemma 6.4, we need to bound the probability that

$$\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for all } i \in [t]. \quad (6.3)$$

We remark that if we condition on an outcome for $\mathcal{S}_1, \dots, \mathcal{S}_t$ satisfying (ii), then one can apply Theorem 4.6 to get an upper bound on this probability (using only the randomness of G_0). However, this bound is weaker than the one claimed in Lemma 6.4. We will be able to obtain a stronger bound by using the randomness of both G_0 and of $\mathcal{S}_1, \dots, \mathcal{S}_t$. For the argument, both of the conditions (i) and (ii) will be relevant.

Whenever the outcome of $(G_0, \mathcal{S}_1, \dots, \mathcal{S}_t)$ is common (specifically, when (i) is satisfied), in order to satisfy (6.3) we must have

$$\|\mu_{\mathcal{G}, \mathcal{S}_i} - \lambda\|_\infty \leq n^{h-g-1} \cdot \log n \text{ for } i = 1, \dots, t. \quad (6.4)$$

Note that (6.4) does not depend on G_0 , only on the random sets in $\mathcal{S}_1, \dots, \mathcal{S}_t$. By Lemma 6.3, for each $i = 1, \dots, t$ the probability of having $\|\mu_{\mathcal{G}, \mathcal{S}_i} - \lambda\|_\infty \leq n^{h-g-1} \cdot \log n$ is at most $n^{-T/2+o(1)}$. So, by the independence of the \mathcal{S}_i , the probability that (6.4) holds is bounded as follows.

$$\Pr(\|\mu_{\mathcal{G}, \mathcal{S}_i} - \lambda\|_\infty \leq n^{h-g-1} \cdot \log n \text{ for } i = 1, \dots, t) \leq \left(n^{-T/2+o(1)}\right)^t = n^{-T \cdot t/2+o(1)}. \quad (6.5)$$

Now, fix any outcomes of $\mathcal{S}_1, \dots, \mathcal{S}_t$ satisfying (ii) and (6.4). By condition (ii), the colour system \mathcal{G}^* is p -general, and furthermore recall that \mathcal{G}^* is complete. Hence, applying Theorem 4.6 to the colour system \mathcal{G}^* (which has parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, t)$), conditioned on our outcomes of $\mathcal{S}_1, \dots, \mathcal{S}_t$, we have

$$\Pr(\psi_H(\mathcal{G}^*, G_0, \cdot, i) = \lambda(\cdot, 1) \text{ for } i = 1, \dots, t) \leq n^{(g-h+1) \cdot t_1 \cdots t_{g-1} \cdot t + o(1)} = n^{(g-h+1) \cdot T \cdot t + o(1)}.$$

(Recall that in this section we are assuming that Theorem 4.6 holds for g). In other words, recalling (6.2), if we condition on any outcomes of $\mathcal{S}_1, \dots, \mathcal{S}_t$ such that (ii) and (6.4) hold, then the random choice $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ satisfies (6.3) with probability at most $n^{(g-h+1) \cdot T \cdot t + o(1)}$.

Combining this with (6.5), we see that the probability that $(G_0, \mathcal{S}_1, \dots, \mathcal{S}_t)$ simultaneously satisfies (6.4), (ii) and (6.3) is at most

$$n^{-T \cdot t/2+o(1)} \cdot n^{(g-h+1) \cdot T \cdot t + o(1)} = n^{(g-h+(1/2)) \cdot T \cdot t + o(1)}.$$

Recall that any common outcome satisfying condition (6.3) also needs to satisfy (6.4), and recall from Claim 6.5 that $(G_0, \mathcal{S}_1, \dots, \mathcal{S}_t)$ is common with probability at least $1 - n^{-\omega(1)}$. Thus, the total probability that (6.3) holds is at most $n^{(g-h+(1/2)) \cdot T \cdot t + o(1)} + n^{-\omega(1)} = n^{(g-h+(1/2)) \cdot T \cdot t + o(1)}$. This finishes the proof of Lemma 6.4. \square

In the proof of Proposition 4.10, we will use the following corollary of Lemma 6.4. In contrast to the setting of Lemma 6.4, in Corollary 6.6 we will not require $t \in \mathbb{N}$ to be fixed, but just that n is sufficiently large with respect to t (and with respect to the values for g, h, a_1, \dots, a_g and t_1, \dots, t_{g-1} that we fixed for this entire section). Also note that the statement of Corollary 6.6 does not contain any asymptotic notation.

Corollary 6.6. *As before, let g, h, a_1, \dots, a_g and t_1, \dots, t_{g-1} be fixed. Then for every $t \in \mathbb{N}$, there exists $N(t) \in \mathbb{N}$ such that the following holds for any essentially p -general complete restricted colour system G of order $n \geq N(t)$ with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, and any function $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$. If we choose, all independently, a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ on the vertex set $\mathcal{U}(\mathcal{G})$ as well as t random collections $\mathcal{S}_1, \dots, \mathcal{S}_t$ chosen as in Definition 4.8, then*

$$\Pr(\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for all } i = 1, \dots, t) \leq n^{(g-h+(1/2)) \cdot T \cdot t+1}.$$

Proof. For each $t \in \mathbb{N}$, the $o(1)$ -term in the statement of Lemma 6.4 with this particular value of t converges to zero (as n goes to infinity). Hence there is some $N(t) \in \mathbb{N}$ such that this $o(1)$ -term is at most 1 for all $n \geq N(t)$. In other words, this means that whenever $n \geq N(t)$, the probability appearing in the statements of Lemma 6.4 and Corollary 6.6 is indeed at most $n^{(g-h+(1/2)) \cdot T \cdot t+1}$, as desired. \square

6.3 Completing the proof of Proposition 4.10

In this subsection, we will finally deduce Proposition 4.10 for our given value of g . In order to do so, we will apply Corollary 6.6 with a carefully chosen value of t . Recall that t is not required to be fixed in Corollary 6.6, it is only required that n is sufficiently large with respect to t (and the values of g, a_1, \dots, a_g and t_1, \dots, t_{g-1} that we fixed throughout this section). Let $N : \mathbb{N} \rightarrow \mathbb{N}$ be a function mapping each $t \in \mathbb{N}$ to some $N(t) \in \mathbb{N}$ such that Corollary 6.6 holds.

In order to prove Proposition 4.10, we can assume that $n \geq N(1)$. For any integer $n \geq N(1)$, let us define $t(n)$ to be the maximum $t \in \mathbb{N}$ with $N(t) \leq n$. Note that then we have $N(t(n)) \leq n$ for all $n \geq N(1)$, and that $t(n) = \omega(1)$.

Define the function δ by $\delta(n) = t(n)^{-1/2}$ for all $n \geq N(1)$, and observe that $\lim_{n \rightarrow \infty} \delta(n) = 0$. Let \mathcal{G} be any essentially p -general complete restricted colour system \mathcal{G} of order $n \geq N(1)$ with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. We need to show that a randomly chosen graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ on the vertex set $\mathcal{U}(\mathcal{G})$ is (p, q, \mathcal{G}) -dispersed with probability $1 - n^{-\omega(1)}$, where

$$q = n^{(g-h+(1/2)) \cdot T + \delta(n)}.$$

Recall from Definition 4.9 that a graph G_0 on the vertex set $\mathcal{U}(\mathcal{G})$ being (p, q, \mathcal{G}) -dispersed means that for all functions $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{Z}$ the following condition holds: when choosing \mathcal{S} randomly as in Definition 4.8, $\Pr(\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot) = \lambda) \leq q$. Let A_λ be the event that G_0 fails to satisfy this condition (i.e. that the probability is strictly larger than q) for some particular λ . So, we need to show that with probability $1 - n^{-\omega(1)}$, no A_λ occurs. Note that we always have $0 \leq \psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g) \leq n^h$, so we only need to consider $(n^h + 1)^T$ possibilities for λ . It therefore suffices to show that $\Pr(A_\lambda) = n^{-\omega(1)}$ for each λ : we then can take a union bound over all possibilities for λ . So, fix a function $\lambda : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \{0, \dots, n^h\}$.

Let $t = t(n)$. All independently, choose a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$, and choose t collections $\mathcal{S}_1, \dots, \mathcal{S}_t$ as in Definition 4.8. By Corollary 6.6 (using that $N(t) = N(t(n)) \leq n$), we have

$$\Pr(\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for all } i = 1, \dots, t) \leq n^{(g-h+(1/2)) \cdot T \cdot t+1}. \quad (6.6)$$

We remark that the probability on the left-hand side of (6.6) can be interpreted as the t -th moment of the random variable (depending on $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$) measuring the conditional probability that $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot) = \lambda$, for a random choice of \mathcal{S} as in Definition 4.8. The rest of the proof, to follow, can basically be interpreted as applying Markov's inequality with this t -th moment.

By the definition of A_λ , if we condition on an outcome of G_0 for which A_λ holds, then for each i , with probability at least q we have $\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda$. Since the \mathcal{S}_i are independent, we deduce that

$$\Pr(\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for } i = 1, \dots, t) \geq \Pr(A_\lambda) \cdot \Pr(\psi_H(\mathcal{G}_{\mathcal{S}_i}, G_0, \cdot) = \lambda \text{ for } i = 1, \dots, t \mid A_\lambda) \geq \Pr(A_\lambda) q^t.$$

Together with (6.6), this implies

$$\Pr(A_\lambda) \cdot q^t \leq n^{(g-h+(1/2)) \cdot T \cdot t+1},$$

and therefore, recalling $q = n^{(g-h+(1/2)) \cdot T + \delta(n)}$,

$$\Pr(A_\lambda) \leq \frac{n^{(g-h+(1/2)) \cdot T \cdot t+1}}{n^{(g-h+(1/2)) \cdot T \cdot t + \delta(n) \cdot t}} = n^{1-\delta(n) \cdot t} = n^{1-\sqrt{t}} = n^{-\omega(1)}.$$

(Recall that $\delta(n) = t(n)^{-1/2} = t^{-1/2}$ and that $t = t(n) = \omega(1)$). This concludes the proof.

7 Completing the induction step: Proposition 4.10 for g implies Theorem 4.6 for $g - 1$

For the duration of this section, we fix some $1 \leq g \leq h - 1$ and assume that Proposition 4.10 holds for this value of g . Our goal for this section is to prove Theorem 4.6 for $g - 1$. Our approach is as outlined in Section 2: we decompose $\psi_H(\mathcal{G}, G_0, \cdot)$ into two parts, use a ‘‘coarse-scale’’ anticoncentration lemma to handle the larger part, and use our assumed case of Proposition 4.10 to handle the smaller part.

Let $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1}$ be arbitrary positive integers. Our goal is to prove Theorem 4.6 for $g - 1$ and these values of $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1}$.

For this entire section, let us fix the values of $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1}$, and define $T = t_1 \cdots t_{g-1}$ and $a_g = 2^{a_1 t_1 + \cdots + a_{g-1} t_{g-1}}$. In all our asymptotics, these values will be treated as constants, while $n \rightarrow \infty$. In Section 7.1 we will prove a concentration lemma which we will use to control the fluctuation of the smaller of our two parts. We also state (but do not yet prove) an anticoncentration lemma which we will use for the larger of our two parts. In Section 7.2 we combine these lemmas with our assumed case of Proposition 4.10 to prove our desired case of Theorem 4.6.

7.1 Preparations

Definition 7.1. Consider an essentially p -general complete restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. Let $\mu_{\mathcal{G}} : [t_1] \times \cdots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$ be the function given by

$$\mu_{\mathcal{G}}(j_1, \dots, j_g) = \mathbb{E}_{G_0, \mathcal{S}}[\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)]$$

for all $(j_1, \dots, j_g) \in [t_1] \times \cdots \times [t_{g-1}] \times [1]$. Here, $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ is a random graph on the vertex set $\mathcal{U}(\mathcal{G})$ and $\mathcal{S} = (S_v)_v$ is a randomly chosen collection of sets as in Definition 4.8 (and G_0 and \mathcal{S} are chosen independently).

We remark that $\mu_{\mathcal{G}}$ only depends on \mathcal{G} .

Lemma 7.2. *The following holds for any essentially p -general complete restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. If we choose a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ and independently choose \mathcal{S} randomly as in Definition 4.8, then*

$$\Pr\left(\|\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot) - \mu_{\mathcal{G}}\|_{\infty} \leq n^{h-g-(1/2)} \cdot \log n\right) = 1 - n^{-\omega(1)}.$$

Proof. By a union bound, it suffices to prove that for each g -tuple $(j_1, \dots, j_g) \in [t_1] \times \cdots \times [t_{g-1}] \times [1]$ the probability to have

$$|\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g) - \mu_{\mathcal{G}}(j_1, \dots, j_g)| > n^{h-g-(1/2)} \cdot \log n \tag{7.1}$$

is $n^{-\omega(1)}$. So fix some g -tuple $(j_1, \dots, j_g) \in [t_1] \times \cdots \times [t_{g-1}] \times [1]$. Note that the expectation of $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ is precisely $\mu_{\mathcal{G}}(j_1, \dots, j_g)$.

Note that the random collection $\mathcal{S} = (S_v)_v$ as in Definition 4.8 consists of a_g different sets $S_v \subseteq \mathcal{U}(\mathcal{G})$, one for each vertex of colour g . For any vertex $u \in \mathcal{U}(\mathcal{G})$, changing which of the sets S_v the vertex u belongs to, changes the edges in the graph $\mathcal{G}_{\mathcal{S}}(G_0, j_1, \dots, j_g)$ between u and the vertices of colour g . Hence it changes the value of $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ by at most $h^{g+1} \cdot a_1 \cdots a_g \cdot n^{h-g-1} = O(n^{h-g-1})$ (since there are at most that many labelled copies of H in the n -vertex graph $\mathcal{G}_{\mathcal{S}}(G_0, j_1, \dots, j_g)$ which contain u as well as at least one vertex of each of the g colours).

Consider the Doob martingale with respect to $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ obtained by first exposing the graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ one edge at a time, and then exposing the random collection \mathcal{S} in the following way. In each step, for one vertex $u \in \mathcal{U}(\mathcal{G})$, we expose which of the sets S_v of the collection \mathcal{S} contain the vertex u . As in the proof of Lemma 6.2, changing the status of an edge of G_0 changes $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g)$ by at most $h^{g+2} \cdot a_1 \cdots a_g \cdot n^{h-g-2} = O(n^{h-g-2})$. So, by the Azuma–Hoeffding inequality the probability that (7.1) occurs is at most

$$\exp\left(-\Omega\left(\frac{n^{2(h-g-1/2)} \cdot (\log n)^2}{n^2 \cdot n^{2(h-g-2)} + n \cdot n^{2(h-g-1)}}\right)\right) \leq \exp(-\Omega((\log n)^2)) = n^{-\omega(1)}.$$

This finishes the proof of Lemma 7.2. □

Lemma 7.3. *The following holds for any weakly p -general complete colour system \mathcal{G} of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ and any $\lambda : [t_1] \times \dots \times [t_{g-1}] \rightarrow \mathbb{Z}$. If we choose a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ on the vertex set $\mathcal{U}(\mathcal{G})$ then*

$$\Pr\left(\|\psi_H(\mathcal{G}, G_0, \cdot) - \lambda\|_\infty \leq n^{h-g-(1/2)} \cdot \log n\right) \leq n^{-T/2+o(1)}.$$

We defer the proof of Lemma 7.3 to Section 9.

7.2 Proof of Theorem 4.6 for $g-1$

In this subsection, we deduce Theorem 4.6 for $g-1$ from Proposition 4.10 for g . Consider any p -general complete colour system \mathcal{G} of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ and any function $\lambda : [t_1] \times \dots \times [t_{g-1}] \rightarrow \mathbb{Z}$. We may assume that n is sufficiently large with respect to the parameters $g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1}$ that were fixed throughout this section.

Since \mathcal{G} is a p -general colour system, the $\mathcal{U}(\mathcal{G})$ -neighbourhoods of each of the coloured vertices of \mathcal{G} in each of the shades of the respective colour are in p -general position. This implies that there exist vertices in $\mathcal{U}(\mathcal{G})$ representing all possible ways to be adjacent to the coloured vertices. To be specific, for each $1 \leq j \leq g-1$ and each vertex v in \mathcal{G} of colour j , consider any subset $I_v \subseteq [t_j]$. There is a vertex $w \in \mathcal{U}(\mathcal{G})$ such that for every $1 \leq j \leq g-1$ and every vertex v of colour j , between w and v there are edges of exactly those shades of colour j that belong to the set I_v . For each choice of the subsets I_v , fix a particular such vertex w , and let W be the set of all these fixed vertices w . Then, W has size

$$(2^{t_1})^{a_1} \dots (2^{t_{g-1}})^{a_{g-1}} = 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}} = a_g.$$

To prove Theorem 4.6, we need to show that for a random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$, we have

$$\Pr(\psi_H(\mathcal{G}, G_0, \cdot) = \lambda) \leq n^{((g-1)-h+1) \cdot T + o(1)} = n^{(g-h) \cdot T + o(1)}.$$

Fix any possible outcome of the edges of the graph $G_0[W]$ between the vertices in W . We will prove that the event $\psi_H(\mathcal{G}, G_0, \cdot) = \lambda$ occurs with probability at most $n^{(g-h) \cdot T + o(1)}$ conditioned on this outcome.

We can obtain a colour system \mathcal{G}^- of order $n - a_g$ with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ from the colour system \mathcal{G} by deleting the a_g vertices in $W \subseteq \mathcal{U}(\mathcal{G})$. By Corollary 5.4, since \mathcal{G} is p -general, \mathcal{G}^- is weakly p -general. Also, \mathcal{G}^- is complete, because \mathcal{G} is complete and being complete does not depend on any of the uncoloured vertices.

Furthermore, from \mathcal{G} and our fixed outcome of $G_0[W]$, we can obtain a restricted colour system \mathcal{G}' of order n with parameters $(g, a_1, \dots, a_{g-1}, a_g, t_1, \dots, t_{g-1})$ in the following way. First, colour all the a_g vertices in $W \subseteq \mathcal{U}(\mathcal{G})$ with colour g . Then, take a single shade of colour g and colour all edges in $G_0[W]$ with this single shade of colour g . The restricted colour system \mathcal{G}' so obtained is complete and essentially p -general, by our choice of W and the assumptions that \mathcal{G} is complete and p -general.

Given our fixed outcome of $G_0[W]$, we can choose the rest of the random graph $G_0 \sim \mathbb{G}(\mathcal{U}(\mathcal{G}), p)$ in two steps as follows. First, we choose a random graph $G_0^- \sim \mathbb{G}(\mathcal{U}(\mathcal{G}) \setminus W, p)$ on the vertex set $\mathcal{U}(\mathcal{G}) \setminus W$. Then, choose random subsets $S_v \subseteq \mathcal{U}(\mathcal{G}) \setminus W$ for each vertex $v \in W$ by taking each element of $\mathcal{U}(\mathcal{G}) \setminus W$ independently with probability p (and choose all these sets S_v independently of each other and independent of G_0^-). Now, take G_0 to be the graph on the vertex set $\mathcal{U}(\mathcal{G})$ obtained by starting with the union of the edges of the graphs G_0^- and $G_0[W]$, and adding edges between each vertex $v \in W$ and all the vertices in $S_v \subseteq \mathcal{U}(\mathcal{G}) \setminus W$.

Let $\mathcal{S} = (S_v)_{v \in W}$ be the collection of random sets chosen above. Note that the random choice of \mathcal{S} is precisely what is described in Definition 4.8 for the restricted colour system \mathcal{G}' (recall that W is the set of vertices of colour g in the restricted colour system \mathcal{G}').

Then, for any outcome of \mathcal{S} and $G_0^- \sim \mathbb{G}(\mathcal{U}(\mathcal{G}) \setminus W, p)$, and any $(j_1, \dots, j_{g-1}) \in [t_1] \times \dots \times [t_{g-1}]$, the graph $\mathcal{G}(G_0, j_1, \dots, j_{g-1})$ is the same as the graph $\mathcal{G}'_{\mathcal{S}}(G_0^-, j_1, \dots, j_{g-1}, 1)$. (Here $\mathcal{G}'_{\mathcal{S}}$ is the colour system obtained from the restricted colour system \mathcal{G}' as in Definition 4.8.) Recall that by definition $\psi_H(\mathcal{G}, G_0, j_1, \dots, j_{g-1})$ is the number of labelled copies of H in the graph $\mathcal{G}(G_0, j_1, \dots, j_{g-1})$ which use at least one vertex of each of the colours $1, \dots, g-1$. If such a copy of H contains at least one vertex in W , then it is one of the $\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1)$ labelled copies of H in the graph $\mathcal{G}'_{\mathcal{S}}(G_0^-, j_1, \dots, j_{g-1}, 1) =$

$\mathcal{G}(G_0, j_1, \dots, j_{g-1})$ which use at least one vertex of each of the g colours of \mathcal{G}' . On the other hand, if it does not contain a vertex of W , then it is one of the $\psi_H(\mathcal{G}^-, G_0^-, j_1, \dots, j_{g-1})$ labelled copies of H in the graph $\mathcal{G}^-(G_0^-, j_1, \dots, j_{g-1}) = \mathcal{G}(G_0, j_1, \dots, j_{g-1}) \setminus W$ which use at least one vertex of each of the colours $1, \dots, g-1$. Thus, for any outcome of \mathcal{S} and G_0^- , we have

$$\psi_H(\mathcal{G}, G_0, \cdot) = \psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot, 1) + \psi_H(\mathcal{G}^-, G_0^-, \cdot).$$

For random $\mathcal{S} = (S_v)_{v \in W}$ and $G_0^- \sim \mathbb{G}(\mathbb{U}(\mathcal{G}) \setminus W, p)$ as above, it therefore suffices to prove that

$$\Pr(\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot, 1) + \psi_H(\mathcal{G}^-, G_0^-, \cdot) = \lambda) \leq n^{(g-h) \cdot T + o(1)}. \quad (7.2)$$

Let $\delta : \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}$ be the function obtained by applying Proposition 4.10 with parameters $g, a_1, \dots, a_g, t_1, \dots, t_{g-1}$ (recall that we are assuming that Proposition 4.10 holds for g). So $\delta(n) = o(1)$. Call an outcome of (\mathcal{S}, G_0^-) *typical* if the following two conditions hold.

- (a) For $q = n^{(g-h+(1/2)) \cdot T + \delta(n)}$, the graph G_0^- is (p, q, \mathcal{G}') -dispersed;
- (b) we have $\|\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot) - \mu_{\mathcal{G}'}\|_{\infty} \leq n^{h-g-(1/2)} \cdot \log n$.

Claim 7.4. (\mathcal{S}, G_0^-) is typical with probability $1 - n^{-\omega(1)}$.

Proof. Recall that \mathcal{G}' is an essentially p -general complete restricted colour system of order n with parameters $(g, a_1, \dots, a_{g-1}, a_g, t_1, \dots, t_{g-1})$. Its set of uncoloured vertices is $\mathbb{U}(\mathcal{G}) \setminus W$. Thus, by Proposition 4.10 (which we assumed to hold for g), the random graph $G_0^- \sim \mathbb{G}(\mathbb{U}(\mathcal{G}) \setminus W, p)$ is (p, q, \mathcal{G}') -dispersed with probability $1 - n^{-\omega(1)}$. This shows that (a) holds with probability $1 - n^{-\omega(1)}$.

On the other hand, (b) holds with probability $1 - n^{-\omega(1)}$ by Lemma 7.2 applied to the restricted colour system \mathcal{G}' . \square

Note that whenever an outcome of (\mathcal{S}, G_0^-) is typical (specifically, whenever (b) holds), we cannot have $\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot, 1) + \psi_H(\mathcal{G}^-, G_0^-, \cdot) = \lambda$ unless

$$\|\psi_H(\mathcal{G}^-, G_0^-, \cdot) + \mu_{\mathcal{G}'}(\cdot, 1) - \lambda\|_{\infty} \leq n^{h-g-(1/2)} \cdot \log n \leq 2 \cdot (n - a_g)^{h-g-(1/2)} \cdot \log(n - a_g). \quad (7.3)$$

(here we used that n is sufficiently large with respect to $a_g = 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$).

Consider the function $\lambda' : [t_1] \times \dots \times [t_{g-1}] \rightarrow \mathbb{Z}$ defined by $\lambda'(j_1, \dots, j_{g-1}) = \lambda(j_1, \dots, j_{g-1}) - \mu_{\mathcal{G}'}(j_1, \dots, j_{g-1}, 1)$. By Lemma 7.3 applied with the function λ' and the weakly p -general complete colour system \mathcal{G}^- of order $n - a_g$ with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$, we see that (7.3) holds with probability at most $(n - a_g)^{-T/2 + o(1)} = n^{-T/2 + o(1)}$.

Note that both (7.3) and (a) only depend on the random choice of G_0^- and not on the random choice of \mathcal{S} . For any outcome of G_0^- such that (7.3) and (a) hold, the conditional probability of the event that $\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot, 1) = \lambda - \psi_H(\mathcal{G}^-, G_0^-, \cdot)$ is at most $q = n^{(g-h+(1/2)) \cdot T + \delta(n)}$. (This follows directly from the definition of being (p, q, \mathcal{G}') -dispersed; see Definition 4.9).

Thus, the total probability that (G_0^-, \mathcal{S}) is typical and satisfies $\psi_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, \cdot, 1) + \psi_H(\mathcal{G}^-, G_0^-, \cdot) = \lambda$ is at most

$$n^{-T/2 + o(1)} \cdot n^{(g-h+(1/2)) \cdot T + \delta(n)} = n^{(g-h) \cdot T + o(1)},$$

recalling that $\delta(n) = o(1)$. By Claim 7.4, the probability that (G_0^-, \mathcal{S}) is not typical is $n^{-\omega(1)}$, so the probability in (7.2) is at most $n^{(g-h) \cdot T + o(1)} + n^{-\omega(1)} = n^{(g-h) \cdot T + o(1)}$. This finishes the proof of Theorem 4.6 for $g-1$.

8 Cores and non-degeneracy

It remains to prove Lemmas 6.3 and 7.3. Both of these lemmas will be proved using our new multivariate anticoncentration inequality (Theorem 3.1), which requires us to check a non-degeneracy condition for a certain collection of vectors. In this section we introduce the formalism of *cores*, which are special types of colour systems of bounded size. For a core \mathcal{C} , we will define certain functions $\Gamma_{\mathcal{C}, e}$ (which may be

interpreted as belonging to a vector space of functions), and we show that under certain conditions these functions span the entire space.

The point of this section is that in the settings of both Lemma 6.3 and Lemma 7.3, the collections of vectors that we need to study to apply Theorem 3.1 are in correspondence with collections of functions $\Gamma_{\mathcal{C},e}$, for appropriately chosen cores \mathcal{C} . In the next section (Section 9) we will specify how to actually choose the cores for these correspondences. Throughout this section, fix any $0 \leq g \leq h - 1$ (recall that h is the number of vertices of our fixed graph H).

Definition 8.1. For integers $0 \leq g \leq h - 1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$, a *core* \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ is a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$ which has exactly one uncoloured vertex, such that the uncoloured vertex has edges to all coloured vertices in all possible shades (meaning that for each $1 \leq i \leq g$, the uncoloured vertex has edges in all t_i shades of colour i to all vertices of colour i).

Note that for fixed parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ there are only finitely many different cores \mathcal{C} with these parameters.

A core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ is called *complete* if it is complete when viewed as a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1}, 1)$.

Definition 8.2. A *partial copy* of H in a core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ is a labelled copy of a $(g + 1)$ -vertex induced subgraph of H which uses one vertex of each colour in \mathcal{C} and the uncoloured vertex. More formally, a partial copy of H in a core \mathcal{C} is given by a graph homomorphism $\phi : H[V'] \rightarrow \mathcal{C}$ for some subset $V' \subseteq V(H)$ of size $g + 1$ such that the image of ϕ contains one vertex of each colour and the uncoloured vertex.

Note that the homomorphism $\phi : H[V'] \rightarrow \mathcal{C}$ in Definition 8.2 is automatically injective on vertices (and therefore also on edges).

Definition 8.3. Consider a core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. For a subset $V' \subseteq V(H)$ of size $g + 1$, define the *weight* of a partial copy $\phi : H[V'] \rightarrow \mathcal{C}$ of H to be $p^{e(H) - e_H(V')}$. Also, for a g -tuple $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$, say that a partial copy $\phi : H[V'] \rightarrow \mathcal{C}$ of H in \mathcal{C} is (j_1, \dots, j_g) -*coloured* if for each $1 \leq i \leq g$ all the edges of colour i in the image of ϕ have shade j_i .

Note that if a partial copy $\phi : H[V'] \rightarrow \mathcal{C}$ of H in \mathcal{C} does not contain any edges of colour i in its image, then it can be (j_1, \dots, j_g) -coloured for different values of j_i . However, if ϕ has at least one edge of colour i in its image, then it can only be (j_1, \dots, j_g) -coloured if j_i is the shade of the edges of colour i in the image of ϕ (if there are different edges of colour i with different shades in the image of ϕ , then ϕ is not (j_1, \dots, j_g) -coloured for any (j_1, \dots, j_g)).

Definition 8.4. For a core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ and a collection of edges $F \subseteq E(\mathcal{C})$, let $\Gamma_{\mathcal{C},F} : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$ be the function defined as follows. For all $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$, let $\Gamma_{\mathcal{C},F}(j_1, \dots, j_g)$ be the sum of the weights of all (j_1, \dots, j_g) -coloured partial copies of H in \mathcal{C} whose image contains all edges in F . If F just consists of one edge e , we write $\Gamma_{\mathcal{C},e}$ instead of $\Gamma_{\mathcal{C},\{e\}}$.

Note that we have $\Gamma_{\mathcal{C},F}(j_1, \dots, j_g) = 0$ if for some $1 \leq i \leq g$ the set F contains an edge of colour i with a shade distinct from j_i , because then there are no (j_1, \dots, j_g) -coloured partial copies of H in \mathcal{C} whose image contains this edge.

Now, the main result of this section is as follows, showing that the functions $\Gamma_{\mathcal{C},e}$ satisfy a non-degeneracy condition.

Lemma 8.5. *Let \mathcal{C} be a complete core with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. Consider the functions $\Gamma_{\mathcal{C},e}$, where e ranges over all edges between the uncoloured vertex of \mathcal{C} and a vertex of colour g . Then these functions span the real vector space of all functions $[t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$.*

8.1 Proof of Lemma 8.5

For this subsection, fix a core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. Let L be the span of the functions $\Gamma_{\mathcal{C},e}$, for all edges e between the uncoloured vertex of \mathcal{C} and a vertex of colour g . Note that L

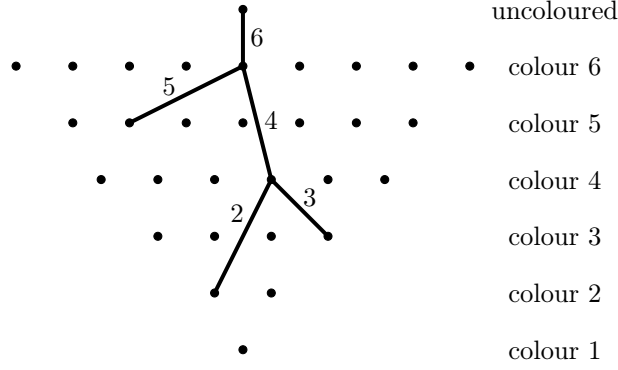


Figure 1: An example of a downward tree, where $g = 6$ and $b = 5$. The numbers on the edges indicate the respective colours.

is a subspace of the real vector space of all functions $[t_1] \times \cdots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$. Our goal is to show that L is actually the entire space of functions $[t_1] \times \cdots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$.

First, let us make some more definitions.

Definition 8.6. For an integer $1 \leq b \leq g$, a *downward tree* of size b in \mathcal{C} is a collection $F \subseteq E(\mathcal{C})$ of b edges of colours $g - b + 1, \dots, g$ which form a tree containing the uncoloured vertex of \mathcal{C} as a leaf.

Although formally a downward tree F is just a collection of edges, we say that F contains a vertex of \mathcal{C} if this vertex is part of the tree formed by the edges in F . For an example of a downward tree, see Figure 1.

Lemma 8.7. For any integer $1 \leq b \leq g$, every downward tree F of size b in \mathcal{C} contains the uncoloured vertex and exactly one vertex in each of the colours $g - b + 1, \dots, g$ (and no vertices in the colours $1, \dots, g - b$). Furthermore, the vertex of colour $g - b + 1$ is a leaf in the tree F . Finally, if $b \geq 2$ and e is the unique edge of F incident to the vertex of colour $g - b + 1$, then e is not incident to the uncoloured vertex and $F \setminus \{e\}$ is a downward tree of size $b - 1$.

Proof. Since F has an edge in each of the colours $g - b + 1, \dots, g$, it must also have a vertex in each of these colours (recall that in a colour system an edge can only have colour i if it is incident to a vertex of colour i). Thus, F contains at least one vertex in each of the colours $g - b + 1, \dots, g$ and by definition it also contains the uncoloured vertex. As F has only $b + 1$ vertices, this establishes the first part of the lemma.

For the second part we need to show that there is only one edge of F incident to the vertex of colour $g - b + 1$. However, note that each edge of F incident to the vertex of colour $g - b + 1$ has colour $g - b + 1$ (because the other vertex of each such edge is either uncoloured or has a colour with a number larger than $g - b + 1$). As F has only one edge of colour $g - b + 1$, there is indeed only one edge of F incident to the vertex of colour $g - b + 1$.

Finally, for the third part, note that the edge e has colour $g - b + 1$ (by the argument above), so $F \setminus \{e\}$ is a tree with $b - 1$ edges of colours $g - b + 2, \dots, g$. The edge e is not incident to the uncoloured vertex, as otherwise both vertices of e would be leaves in F which would contradict $b \geq 2$. In particular, the uncoloured vertex is still a leaf of the tree $F \setminus \{e\}$. Hence $F \setminus \{e\}$ is indeed a downward tree of size $b - 1$. \square

Lemma 8.8. For every downward tree $F \subseteq E(\mathcal{C})$, we have $\Gamma_{\mathcal{C}, F} \in L$.

Proof. We prove the lemma by induction on b .

If $b = 1$, then $F = \{e\}$ for a single edge e . By definition, the uncoloured vertex is a leaf of F , so e is incident to the uncoloured vertex. Furthermore, e has colour g , so it is also incident to a vertex of colour g . We therefore trivially have $\Gamma_{\mathcal{C}, e} \in L$, by the definition of L .

Now, let us assume that $b \geq 2$ and that the lemma statement holds for $b - 1$. Let F be a downward tree of size b . Recall that by Lemma 8.7, F contains precisely one vertex v^* of colour $g - b + 1$ and this vertex is a leaf in F . So let e^* be the unique edge in F incident to v^* . Then by the last part of Lemma 8.7, $F \setminus \{e^*\}$ is a downward tree of size $b - 1$. Finally, let w be the other vertex of e^* , so that (again by Lemma 8.7) w is coloured with one of the colours $g - b + 2, \dots, g$.

Applying Lemma 8.7 to the downward tree $F \setminus \{e^*\}$, we can label the vertices of $F \setminus \{e^*\}$ as v_{g-b+2}, \dots, v_g, u , such that u is the uncoloured vertex of \mathcal{C} and such that each vertex v_i has colour i . Note that w is one of the vertices v_{g-b+2}, \dots, v_g .

Now, since \mathcal{C} is complete (recall Definition 4.4), we can recursively define vertices v'_{g-b+2}, \dots, v'_g in \mathcal{C} , with colours $g - b + 2, \dots, g$ respectively, such that the following three conditions are satisfied.

- For all $g - b + 2 \leq j < i \leq g$, the vertices v'_i and v'_j are connected by edges of exactly the same shades of colour j as the vertices v_i and v_j .
- For all $g - b + 2 \leq i \leq g$ and all vertices v of \mathcal{C} of any of the colours $1, \dots, g - b + 1$ such that $(v_i, v) \neq (w, v^*)$, the vertices v'_i and v are connected by edges of exactly the same shades of the colour of v as the vertices v_i and v .
- For the index $g - b + 2 \leq i \leq g$ such that $v_i = w$, the vertices v'_i and v^* are connected by edges of exactly the same shades of the of colour $g - b + 1$ as the vertices v_i and v^* except that there is no edge between v'_i and v^* with the shade of e^* .

Informally speaking, these conditions are saying that the shade of the edges between any two of the vertices v'_{g-b+2}, \dots, v'_g and the shades of the edges between the vertices v'_{g-b+2}, \dots, v'_g and the vertices of colours $1, \dots, g - b + 1$ are the same as the corresponding shades for v_{g-b+2}, \dots, v_g except that between the vertices w' and v^* the shade of the edge e^* between w and v^* is missing. Here, w' denotes the vertex v'_i for the index $g - b + 2 \leq i \leq g$ such that $v_i = w$.

Now, for each edge $e \in F \setminus \{e^*\}$ with endpoints v_i and v_j , for $g - b + 2 \leq j < i \leq g$, there exists an edge $e' \in E(\mathcal{C})$ between v'_i and v'_j such that e' has the same shade of colour j as e . Let F' be the collection of all these edges together with the unique edge between v'_g and u (recall that there is only one shade of colour g and that the edge between v_g and u is the only edge in $F \setminus \{e^*\}$ incident to u). Then F' forms a tree which is isomorphic to $F \setminus \{e^*\}$ and the corresponding edges are coloured the same way. As $F \setminus \{e^*\}$ is a downward tree of size $b - 1$, this implies that F' is also a downward tree of size $b - 1$. Furthermore, between any two vertices of F' there exist edges in exactly the same shades as between the corresponding vertices of $F \setminus \{e^*\}$. Finally, every vertex in one of the colours $1, \dots, g - b + 1$ has edges in the same shades to the vertices of $F \setminus \{e^*\}$ as to the corresponding vertices of F' except that the shade of the edge e^* is missing between the vertices v^* and w' (recall that $e^* \in F$ is an edge between v^* and w).

Thus, for every partial copy of H in \mathcal{C} containing F' we can form a corresponding partial copy of H in \mathcal{C} containing $F \setminus \{e^*\}$ but not containing e^* , by simply replacing each of the vertices v'_{g-b+2}, \dots, v'_g in the image of the partial copy by v_{g-b+2}, \dots, v_g (and by replacing each edge in the image of the partial copy by an edge of the same shade between the corresponding vertices). This process is bijective and does not change which shades of which colours occur among the edges in the image of the partial copy. It also does not change the weight of the partial copy. Thus, for every $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$, the quantity $\Gamma_{\mathcal{C}, F'}(j_1, \dots, j_g)$ is equal to the sum of the weights of the (j_1, \dots, j_g) -coloured partial copies of H in \mathcal{C} that contain $F \setminus \{e^*\}$ but not e^* . In other words, we have $\Gamma_{\mathcal{C}, F'}(j_1, \dots, j_g) = \Gamma_{\mathcal{C}, F \setminus \{e^*\}}(j_1, \dots, j_g) - \Gamma_{\mathcal{C}, F}(j_1, \dots, j_g)$ for all $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$.

Since F' and $F \setminus \{e^*\}$ are downward trees of size $b - 1$, we have $\Gamma_{\mathcal{C}, F'}, \Gamma_{\mathcal{C}, F \setminus \{e^*\}} \in L$ by the induction assumption. Hence $\Gamma_{\mathcal{C}, F} = \Gamma_{\mathcal{C}, F \setminus \{e^*\}} - \Gamma_{\mathcal{C}, F'} \in L$, as desired. \square

Now, note that each downward tree $F \subseteq E(\mathcal{C})$ of size g contains exactly one edge in each of the colours $1, \dots, g$. For each $1 \leq i \leq g$, let $j_i \in [t_i]$ be the shade of the edge of colour i in F . Then for any tuple $(j'_1, \dots, j'_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$ with $(j'_1, \dots, j'_g) \neq (j_1, \dots, j_g)$ we have $\Gamma_{\mathcal{C}, F}(j'_1, \dots, j'_g) = 0$, since by definition there cannot be any (j'_1, \dots, j'_g) -coloured partial copies of H containing F . That is to say, $\Gamma_{\mathcal{C}, F} \in L$ is a (possibly zero) multiple of the indicator function of (j_1, \dots, j_g) . Since the set of indicator functions of each (j_1, \dots, j_g) span the space of all functions $[t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$, it suffices to prove the following lemma in order to finish the proof of Lemma 8.5.

Lemma 8.9. *For every $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$ there exists a downward tree F of size g in \mathcal{C} such that $\Gamma_{\mathcal{C},F}(j_1, \dots, j_g) > 0$.*

Proof. Fix $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$. Recall that $\Gamma_{\mathcal{C},F}(j_1, \dots, j_g)$ is the sum of the weights of all (j_1, \dots, j_g) -coloured partial copies of H in \mathcal{C} whose image contains all edges in F . Since each partial copy of H in \mathcal{C} has positive weight, it suffices to show that there exists a (j_1, \dots, j_g) -coloured partial copy of H in \mathcal{C} which contains some downward tree F of size g .

Since \mathcal{C} is complete, we can recursively choose vertices v_1, \dots, v_g in \mathcal{C} such that for each $1 \leq i \leq n$, the vertex v_i has colour i and for any $1 \leq j < i \leq g$ there are edges of all t_j shades of colour j between v_j and v_i . Also, since H is connected, and has $h \geq g + 1$ vertices, we can choose a g -edge subtree F_H of H . Let $V' \subseteq V(H)$ be the set of vertices of this subtree F_H and note that $|V'| = g + 1$. Choose one leaf of F_H and call it u . Now, define $\phi : V' \rightarrow V(\mathcal{C})$ by mapping u to the uncoloured vertex of \mathcal{C} and mapping the remaining vertices of F_H to v_1, \dots, v_g in order of decreasing distance to u in the tree F_H (this means the vertex of F_H with maximum distance from u in the tree F_H will be mapped to v_1 , the vertex with second-largest distance to v_2 , and so on), where we break ties arbitrarily.

Note that the image of $\phi : V' \rightarrow V(\mathcal{C})$ contains one vertex of each colour (the vertices v_1, \dots, v_g) and the uncoloured vertex. By the choice of v_1, \dots, v_g , we can extend ϕ to a (j_1, \dots, j_g) -coloured partial copy of H in \mathcal{C} (we have already defined the way ϕ maps the relevant vertices of H , we just need to define the way it maps edges). Let $F = \phi(F_H)$ be the image of our subtree F_H of H . We can check that F is a downward tree of size g . \square

9 Proofs of Lemmas 6.3 and 7.3

In this section we finally prove Lemmas 6.3 and 7.3, using our new anticoncentration inequality in Section 3 and the functions $\Gamma_{\mathcal{C},e}$ defined in Section 8. In Section 9.1 we make some definitions and state some auxiliary lemmas. Most importantly, we explain how to define cores \mathcal{C} in such a way that the functions $\Gamma_{\mathcal{C},e}$ represent the typical effects of changing the status of certain edges. In Sections 9.2 and 9.3 we prove the auxiliary lemmas, and we put everything together in Section 9.4.

9.1 Preparations

First, we define functions $\kappa_H(\mathcal{G}, G_0, \cdot, u, v)$, measuring the change to $\psi_H(\mathcal{G}, G_0, \cdot)$ that results from changing an edge uv in G_0 . In our proof of Lemma 7.3, taking $u, v \in U(\mathcal{G})$, these functions will correspond to the $\Delta_i \mathbf{f}$ that appear in the statement of Theorem 3.1. Recall the definition of the graph $\mathcal{G}(G_0, j_1, \dots, j_g)$ from Definition 4.5.

Definition 9.1. Let \mathcal{G} be a colour system with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_g)$. Then, given two distinct vertices u and v in \mathcal{G} , a g -tuple $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_g]$, and a graph G_0 on the vertex set $U(\mathcal{G})$, let $\kappa_H(\mathcal{G}, G_0, j_1, \dots, j_g, u, v)$ be the number of labelled copies of H in the graph $\mathcal{G}(G_0, j_1, \dots, j_g) + \{uv\}$ which use the edge uv as well as at least one vertex of each of the g colours. (Here, $\mathcal{G}(G_0, j_1, \dots, j_g) + \{uv\}$ is the graph obtained from $\mathcal{G}(G_0, j_1, \dots, j_g)$ by adding the edge uv if this edge is not already present.)

Next, we need a similar definition for the proof of Lemma 6.3. Recall that Lemma 6.3 concerns functions $\mu_{\mathcal{G},\mathcal{S}}$, which are obtained by averaging functions of the form $\psi_H(\mathcal{G}_{\mathcal{S}}, G_0, \cdot)$ over G_0 . We will be interested in the effects on $\mu_{\mathcal{G},\mathcal{S}}$ of adding or removing vertices from the “neighbourhood” sets in \mathcal{S} , which is equivalent to changing the status of edges incident to one of the a_g vertices of colour g . We define functions $\nu_{\mathcal{G},\mathcal{S},u,v}$ (where u is an uncoloured vertex and v is a vertex of colour g) to measure the average effects of such changes.

Definition 9.2. Given a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ where $1 \leq g \leq h - 1$, an outcome of the random collection of sets \mathcal{S} in Definition 4.8, a vertex $u \in U(\mathcal{G})$, and a vertex v of colour g in \mathcal{G} , let $\nu_{\mathcal{G},\mathcal{S},u,v} : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$ be the function given by

$$\nu_{\mathcal{G},\mathcal{S},u,v}(j_1, \dots, j_g) = \mathbb{E}_{G_0}[\kappa_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g, u, v)]$$

for all $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$. Here, $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ is a random graph on the vertex set $U(\mathcal{G})$.

Now, we want to show that the typical values of the $\nu_{\mathcal{G},\mathcal{S},u,v}$ and $\kappa_H(\mathcal{G}, G_0, \cdot, u, v)$ can be expressed in terms of functions $\Gamma_{\mathcal{C},e}$. First, we consider $\nu_{\mathcal{G},\mathcal{S},u,v}$, for the proof of Lemma 6.3. It will suffice to restrict our attention to the cases where u comes from a very ‘‘rich’’ subset of the uncoloured vertices.

Definition 9.3. For a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, let $U^*(\mathcal{G})$ be the set of all uncoloured vertices in \mathcal{G} which are connected to all vertices of the colours $1, \dots, g-1$ in all possible shades of these colours.

When a general position assumption is satisfied, $U^*(\mathcal{G})$ has linear size (by Lemma 5.1), as follows.

Fact 9.4. Fix integers $g, a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$. If \mathcal{G} is an essentially p -general restricted colour system of order n which has parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ then $|U^*(\mathcal{G})| = \Omega(n)$.

We remind the reader that (as in the rest of this paper) the asymptotics in Fact 9.4 are as $n \rightarrow \infty$, while p and the parameters of the restricted colour system are treated as fixed constants for all asymptotic notation. Now, the relevant core for Lemma 6.3 is as follows.

Definition 9.5. Given a restricted colour system \mathcal{G} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$, we can obtain a core \mathcal{C} with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$ as follows. Consider all coloured vertices of \mathcal{G} together with one additional uncoloured vertex which we connect to all coloured vertices by edges in all possible shades. We call \mathcal{C} the *core of the restricted colour system* \mathcal{G} .

Note that if a restricted colour system \mathcal{G} is complete, then its core is also complete (recall that being complete only depends on the edges between the coloured vertices). The following lemma gives the connection between the functions $\nu_{\mathcal{G},\mathcal{S},u,v}$ and the functions $\Gamma_{\mathcal{C},e}$. It will be proved in Section 9.2.

Lemma 9.6. Fix integers $1 \leq g \leq h-1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$. Let \mathcal{G} be an essentially weakly p -general restricted colour system of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. Furthermore, let \mathcal{S} be an outcome of the random collection of sets \mathcal{S} in Definition 4.8, let $u \in U^*(\mathcal{G})$ and let v be a vertex of colour g in \mathcal{G} . Finally, let \mathcal{C} be the core of the restricted colour system \mathcal{G} , and let e be the unique edge from v to the uncoloured vertex in \mathcal{C} (recall that v is a vertex of colour g in \mathcal{C} , and that the colour g only has one shade). Then, if the colour system $\mathcal{G}_{\mathcal{S}}$ is p -general, we have

$$\|\nu_{\mathcal{G},\mathcal{S},u,v} - n^{h-g-1}\Gamma_{\mathcal{C},e}\|_{\infty} = O\left(n^{h-g-(3/2)} \log n\right).$$

Next, we turn to the functions $\kappa_H(\mathcal{G}, G_0, \cdot, u, v)$, for the proof of Lemma 7.3. For this, we consider cores of a different type.

Definition 9.7. Given a colour system \mathcal{G} with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$, define its *extended core* to be the core with parameters $(g, a_1, \dots, a_{g-1}, 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}, t_1, \dots, t_{g-1})$ obtained as follows. Start with all coloured vertices of \mathcal{G} . Now, for each possible choice of subsets $I_v \subseteq [t_j]$ for each $1 \leq j \leq g-1$ and each vertex v of colour j , add a vertex of colour g which is connected to all the vertices v of colours $1, \dots, g-1$ with edges of exactly the shades given by the set I_v . Finally, add one uncoloured vertex and connect it to all coloured vertices by edges in all possible shades (including exactly one shade of colour g).

Note that in Definition 9.7, there are precisely $(2^{t_1})^{a_1} \dots (2^{t_{g-1}})^{a_{g-1}} = 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ different choices for all the subsets $I_v \subseteq [t_j]$. Thus, $2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ vertices of colour g get added and the resulting core indeed has parameters $(g, a_1, \dots, a_{g-1}, 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}, t_1, \dots, t_{g-1})$. Furthermore, note that if the colour system \mathcal{G} is complete, then its extended core is a complete core.

In a similar way to Lemma 9.6, for the proof of Lemma 7.3 it will suffice to restrict our attention to those $\kappa_H(\mathcal{G}, G_0, \cdot, u, v)$ where u and v belong to certain special sets of uncoloured vertices.

Definition 9.8. Let \mathcal{G} be a colour system with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$ and let \mathcal{C} be the extended core of \mathcal{G} . Let $E^g(\mathcal{C})$ be the set of edges e connecting the uncoloured vertex in \mathcal{C} to some vertex w of colour g in \mathcal{C} (such an edge is uniquely determined by w because colour g only has one shade). For each such $e \in E^g(\mathcal{C})$, we define the subset $U_e \subseteq U(\mathcal{G})$ as follows. Let U_e consist of all those uncoloured vertices v in \mathcal{G} such that v is connected to all vertices of \mathcal{G} of colours $1, \dots, g-1$ in precisely the same shades of the corresponding colours in which the vertex w is connected to these vertices in \mathcal{C} . Also, let $U_* \subseteq U(\mathcal{G})$ be the set of all uncoloured vertices v in \mathcal{G} which are connected to all vertices of \mathcal{G} of colours $1, \dots, g-1$ in all possible shades of these colours.

Note that U_* is a special case of U_e , for the edge $e \in E^g(\mathcal{C})$ connecting the uncoloured vertex of \mathcal{C} to the unique vertex of colour g in \mathcal{C} which is connected to all vertices of colours $1, \dots, g-1$ in all possible shades of these colours. Also note that the $2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ sets U_e , for $e \in E^g(\mathcal{C})$, form a partition of $U(\mathcal{G})$. We will need a counterpart of Fact 9.4 (again a simple consequence of Lemma 5.1): when a general position assumption is satisfied, each $U_e(\mathcal{G})$ has linear size, as follows.

Fact 9.9. *Fix integers $g, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1} \geq 1$. Let \mathcal{G} be a weakly p -general colour system of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$, and let \mathcal{C} be the extended core of \mathcal{G} . Then for every $e \in E^g(\mathcal{C})$, we have $|U_e| = \Omega(n)$. In particular, $|U_*| = \Omega(n)$.*

The next lemma will be proved in Section 9.3.

Lemma 9.10. *Fix integers $1 \leq g \leq h-1$ and $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1} \geq 1$. Let \mathcal{G} be a weakly p -general colour system of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$, and let \mathcal{C} be the extended core of \mathcal{G} . Consider an edge $e \in E^g(\mathcal{C})$, and consider the sets $U_e \subseteq U(\mathcal{G})$ and $U_* \subseteq U(\mathcal{G})$ as defined as in Definition 9.8. Then for any distinct vertices $u \in U_*$ and $v \in U_e$ the following holds. If we choose a random graph $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ on the vertex set $U(\mathcal{G})$, then with probability $1 - n^{-\omega(1)}$ we have*

$$\left\| \kappa_H(\mathcal{G}, G_0, \cdot, u, v) - n^{h-g-1} \Gamma_{\mathcal{C}, e}(\cdot, 1) \right\|_\infty = O\left(n^{h-g-(3/2)} \log n\right).$$

9.2 Proof of Lemma 9.6

In this subsection we prove Lemma 9.6. Throughout the subsection, fix integers $1 \leq g \leq h-1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$ (in order to prove Lemma 9.6 for these values). For all asymptotic notation, these fixed values are treated as constants.

As in the statement of Lemma 9.6, let \mathcal{G} be an essentially weakly p -general restricted colour system of order n with parameters $(g, a_1, \dots, a_g, t_1, \dots, t_{g-1})$. Let \mathcal{S} be an outcome of the random collection of sets \mathcal{S} in Definition 4.8 such that $\mathcal{G}_{\mathcal{S}}$ is p -general, let $u \in U^*(\mathcal{G})$ and let v be a vertex of colour g in \mathcal{G} . Finally, let \mathcal{C} be the core of the restricted colour system \mathcal{G} , and let $e \in E^g(\mathcal{C})$ be the unique edge from v to the uncoloured vertex in \mathcal{C} . We need to prove that

$$\left\| \nu_{\mathcal{G}, \mathcal{S}, u, v} - n^{h-g-1} \Gamma_{\mathcal{C}, e} \right\|_\infty = O\left(\log n \cdot n^{h-g-(3/2)}\right).$$

Now, let us define slightly modified versions of κ_H and the $\nu_{\mathcal{G}, \mathcal{S}, u, v}$, that are easier to work with. For any outcome of the random graph $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ and any $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$, let $\kappa'_H(\mathcal{G}, G_0, j_1, \dots, j_g, u, v)$ be the number of labelled copies of H in the graph $\mathcal{G}(G_0, j_1, \dots, j_g) + \{uv\}$ which use the edge uv and which use *exactly* one vertex of each of the g colours (for κ_H we considered the number of copies of H which use *at least* one vertex of each colour). Then, we define $\nu'_{\mathcal{G}, \mathcal{S}, u, v} : [t_1] \times \dots \times [t_{g-1}] \times [1] \rightarrow \mathbb{R}$ analogously to $\nu_{\mathcal{G}, \mathcal{S}, u, v}$:

$$\nu'_{\mathcal{G}, \mathcal{S}, u, v}(j_1, \dots, j_g) = \mathbb{E}_{G_0}[\kappa'_H(\mathcal{G}_{\mathcal{S}}, G_0, j_1, \dots, j_g, u, v)]$$

for all $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$.

Note that for any outcome of the random graph $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ and any $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$ the difference $\kappa_H(\mathcal{G}, G_0, j_1, \dots, j_g, u, v) - \kappa'_H(\mathcal{G}, G_0, j_1, \dots, j_g, u, v)$ is precisely the number of labelled copies of H in the graph $\mathcal{G}_{\mathcal{S}}(G_0, j_1, \dots, j_g) + \{uv\}$ which use the edge uv as well as at least one vertex of each of the g colours and which use at least two vertices of the same colour. Each such labelled copy has to use at least $g+1$ of the coloured vertices and also the vertex $u \in U^*(\mathcal{G}) \subseteq U(\mathcal{G})$. Hence it can use at most $h-g-2$ vertices in $U(\mathcal{G}) \setminus \{u\}$. Thus, the number of such labelled copies is always at most $h^{g+2} \cdot (a_1 + \dots + a_g + 1)^{g+2} \cdot n^{h-g-2} = O(n^{h-g-2})$. It follows that

$$\left\| \nu_{\mathcal{G}, \mathcal{S}, u, v} - \nu'_{\mathcal{G}, \mathcal{S}, u, v} \right\|_\infty = O(n^{h-g-2}).$$

So, to prove Lemma 9.6 it suffices to prove the following lemma.

Lemma 9.11. *We have*

$$\left\| \nu'_{\mathcal{G}, \mathcal{S}, u, v} - n^{h-g-1} \Gamma_{\mathcal{C}, e} \right\|_\infty = O\left(\log n \cdot n^{h-g-(3/2)}\right).$$

Proof. We need to show that for every $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$ we have

$$|\nu'_{\mathcal{G}, \mathcal{S}, u, v}(j_1, \dots, j_g) - n^{h-g-1} \Gamma_{\mathcal{C}, e}(j_1, \dots, j_g)| = O\left(\log n \cdot n^{h-g-(3/2)}\right). \quad (9.1)$$

So let us fix some $(j_1, \dots, j_g) \in [t_1] \times \dots \times [t_{g-1}] \times [1]$. For every $i = 1, \dots, g$ let us refer to shade j_i of colour i as the *desired* shade of colour i .

Recall that $\nu'_{\mathcal{G}, \mathcal{S}, u, v}(j_1, \dots, j_g) = \mathbb{E}_{G_0}[\kappa'_H(\mathcal{G}_S, G_0, j_1, \dots, j_g, u, v)]$ is the expected number of labelled copies of H in the graph $\mathcal{G}_S(G_0, j_1, \dots, j_g) + \{uv\}$ which use exactly one vertex of each of the g colours and use the edge uv . We organise these copies by how they interact with the coloured vertices, as follows.

Consider any graph homomorphism $\phi : H[V_\phi] \rightarrow \mathcal{G}_S(G_0, j_1, \dots, j_g) + \{uv\}$, such that V_ϕ is a subset of $g+1$ vertices of H and such that the image of ϕ contains exactly one vertex of each of the g colours and the edge uv (let Φ be the set of all such homomorphisms). Let E_ϕ be the expected number of labelled copies of H in the graph $\mathcal{G}_S(G_0, j_1, \dots, j_g) + \{uv\}$ that extend ϕ by mapping the vertices in $V \setminus V_\phi$ into $U(\mathcal{G}) \setminus \{u\}$ (where the expectation is taken over the random choice of $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$). Then we have

$$\nu'_{\mathcal{G}, \mathcal{S}, u, v}(j_1, \dots, j_g) = \sum_{\phi \in \Phi} E_\phi. \quad (9.2)$$

Now, since \mathcal{G}_S is p -general, we can estimate the E_ϕ as follows.

Claim 9.12. *For each $\phi \in \Phi$ as above, we have*

$$\left| E_\phi - p^{e_H - e_H(V_\phi)} n^{h-g-1} \right| \leq O\left(n^{h-g-(3/2)} \log n\right).$$

Proof. Let V_{col} be the set of those vertices $x \in V_\phi$ such that $\phi(x)$ is a coloured vertex (i.e. $\phi(x) \neq u$). So, $|V_{\text{col}}| = g$. Recall that E_ϕ is the expected number of labelled copies of H in the graph $\mathcal{G}_S(G_0, j_1, \dots, j_g) + \{uv\}$ that extend ϕ by mapping the vertices in $V \setminus V_\phi$ into $U(\mathcal{G}) \setminus \{u\}$. For every vertex $y \in V(H) \setminus V_\phi$, let $M_\phi(y)$ be the set of possible choices for the image of y that are compatible with the map ϕ on $H[V_\phi]$. More precisely, $M_\phi(y)$ is the set of vertices $w \in U(\mathcal{G}) \setminus \{u\}$ such that for every neighbour $x \in V_{\text{col}}$ of y , the vertex w is connected to $\phi(x)$ in the desired shade of the colour of the vertex $\phi(x)$. Let N be the number of $(h-g-1)$ -tuples in $\prod_{y \in V(H) \setminus V_\phi} M_\phi(y)$ whose vertices are distinct (that is, the number of ways to choose a distinct vertex from each $M_\phi(y)$). Then $N = \prod_{y \in V(H) \setminus V_\phi} |M_\phi(y)| + O(n^{h-g-2})$, and

$$E_\phi = p^{e_H(V(H) \setminus V_{\text{col}})} \cdot N = p^{e_H(V(H) \setminus V_{\text{col}})} \cdot \prod_{y \in V(H) \setminus V_\phi} |M_\phi(y)| + O(n^{h-g-2}). \quad (9.3)$$

Indeed, if we choose possible images for all the vertices $y \in V(H) \setminus V_\phi$ (there are N such choices), then each of the $e_H(V(H) \setminus V_{\text{col}})$ edges of H inside $V(H) \setminus V_{\text{col}}$ needs to be mapped to an edge of $G_0 \sim \mathbb{G}(|U(\mathcal{G})|, p)$ and the probability for this to happen is $p^{e_H(V(H) \setminus V_{\text{col}})}$.

Now, the sizes of the $M_\phi(y)$ are dictated by our assumption that \mathcal{G}_S is p -general. Fix a vertex $y \in V(H) \setminus V_\phi$ and let x_1, \dots, x_k be its neighbours in V_{col} . Let $N_1, \dots, N_k \subseteq U(\mathcal{G})$ be the neighbourhoods of the vertices $\phi(x_1), \dots, \phi(x_k)$ in $U(\mathcal{G})$ in the desired shades of the colours of $\phi(x_1), \dots, \phi(x_k)$, respectively. Then the set of possible choices for the image of y is $N_1 \cap \dots \cap N_k \setminus \{u\}$. So, $|M_\phi(y)|$ differs from $|N_1 \cap \dots \cap N_k|$ by at most 1. Now, if we consider all the neighbourhoods in $U(\mathcal{G})$ of all vertices of colours $1, \dots, g$ in \mathcal{G}_S in all the respective shades of these colours, then these are $a_1 t_1 + \dots + a_{g-1} t_{g-1} + a_g$ subsets of the set $U(\mathcal{G})$ in $(p, 3^g)$ -general position (this is because \mathcal{G}_S is by assumption a p -general colour system). Thus, by Lemma 5.1 we have

$$\left| |N_1 \cap \dots \cap N_k| - p^k |U(\mathcal{G})| \right| = O\left(|U(\mathcal{G})|^{1/2} \log |U(\mathcal{G})|\right).$$

Recall that $|U(\mathcal{G})| = n - a_1 - \dots - a_g = n - O(1)$, and that k was the number of neighbours in of y in V_{col} , so

$$\left| |M_\phi(y)| - p^{e_H(y, V_{\text{col}})} n \right| = \left| |M_\phi(y)| - p^k n \right| \leq \left| |N_1 \cap \dots \cap N_k| - p^k n \right| + 1 = O\left(n^{1/2} \log n\right),$$

or equivalently

$$|M_\phi(y)| = \left(1 + O\left(\frac{\log n}{\sqrt{n}}\right)\right) p^{e_H(y, V_{\text{col}})} n.$$

Finally, observe that $e_H(V(H)\setminus V_{\text{col}})$, plus the sum of all the $e_H(y, V_{\text{col}})$, for $y \in V(H)\setminus V_\phi$, is equal to $e(H) - e_H(V_\phi)$. Indeed, since V_ϕ and $V(H)\setminus V_{\text{col}}$ intersect in only one vertex z , every edge of H is either between two vertices of V_ϕ , two vertices of $V(H)\setminus V_{\text{col}}$, or between a vertex of $V_\phi\setminus\{z\} = V_{\text{col}}$ and a vertex of $(V(H)\setminus V_{\text{col}})\setminus\{z\} = V(H)\setminus V_\phi$. From (9.3) we therefore conclude that

$$\begin{aligned} E_\phi &= p^{e_H(V(H)\setminus V_{\text{col}})} \cdot \prod_{y \in V(H)\setminus V_\phi} \left((1 + O(\log n/\sqrt{n})) p^{e_H(y, V_{\text{col}})} n \right) + O(n^{h-g-2}) \\ &= (1 + O(\log n/\sqrt{n})) p^{e(H)-e_H(V_\phi)} n^{h-(g+1)} = p^{e(H)-e_H(V_\phi)} n^{h-g-1} + O(n^{h-g-(3/2)} \log n), \end{aligned}$$

which is equivalent to the desired bound. \square

Now, the sum in (9.2) is only over $|\Phi| \leq h^{g+1} a_1 \dots a_g = O(1)$ choices of ϕ , so Claim 9.12 implies that

$$\left| \nu'_{\mathcal{G}, \mathcal{S}, u, v}(j_1, \dots, j_g) - \sum_{\phi \in \Phi} \left(p^{e(H)-e_H(V_\phi)} n^{h-g-1} \right) \right| = O\left(n^{h-g-(3/2)} \log n \right).$$

Finally, recall that Φ is the set of homomorphisms of the form $\phi : H[V_\phi] \rightarrow \mathcal{G}_{\mathcal{S}}(G_0, j_1, \dots, j_g) + \{uv\}$, with $|V_\phi| = g + 1$, such that the image of ϕ contains exactly one vertex of each of the g colours and the edge uv . The core \mathcal{C} of the restricted colour system \mathcal{G} was defined (in Definition 9.5) in such a way that there is a bijective correspondence between Φ and the set of (j_1, \dots, j_g) -coloured partial copies ϕ^* of H in \mathcal{C} which contain the edge e . Recall that $\Gamma_{\mathcal{C}, e}(j_1, \dots, j_g)$ was defined (in Definition 8.4) as the sum of the weights of all (j_1, \dots, j_g) -coloured partial copies whose image contains e , and the weight of a partial copy $\phi^* : H[V'] \rightarrow \mathcal{C}$ was defined to be $p^{e(H)-e_H(V')}$. So,

$$\Gamma_{\mathcal{C}, e}(j_1, \dots, j_g) = \sum_{\phi \in \Phi} p^{e(H)-e_H(V_\phi)}.$$

The desired bound (9.1) follows. \square

9.3 Proof of Lemma 9.10

In this subsection we deduce Lemma 9.10 from Lemma 9.6. Throughout the subsection, fix integers $1 \leq g \leq h-1$ and $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1} \geq 1$ (in order to prove Lemma 9.10 for these values). For all asymptotic notation, these fixed values are treated as constants.

Let \mathcal{G} be a weakly p -general colour system of order n with parameters $(g-1, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$, let \mathcal{C} be the extended core of \mathcal{G} , and consider some $e \in E^g(\mathcal{C})$. Fix distinct vertices $u \in U_*$ and $v \in U_e$. We need to show that for a random graph $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$, with probability $1 - n^{-\omega(1)}$ we have

$$\left| \kappa_H(\mathcal{G}, G_0, j_1, \dots, j_{g-1}, u, v) - n^{h-g-1} \Gamma_{\mathcal{C}, e}(j_1, \dots, j_{g-1}, 1) \right| = O\left(\log n \cdot n^{h-g-(3/2)} \right). \quad (9.4)$$

for all $(j_1, \dots, j_{g-1}) \in [t_1] \times \dots \times [t_{g-1}]$. For the rest of the proof fix some $(j_1, \dots, j_{g-1}) \in [t_1] \times \dots \times [t_{g-1}]$: we will show that (9.4) holds with probability $1 - n^{-\omega(1)}$ (then the desired result will follow, taking a union bound over all choices of (j_1, \dots, j_{g-1})).

By Fact 9.9, each of the $2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ disjoint sets U_f has size $\Omega(n) \geq 2$. Let Z be a set containing one representative from each U_f , taking $v \in U_e$, but taking some vertex other than u in U_* . If we imagine that the vertices of Z are coloured with colour g , then, by our choice of Z , the coloured vertices of \mathcal{G} together with Z form a colour system which looks the same as the extended core \mathcal{C} of \mathcal{G} except that the uncoloured vertex of \mathcal{C} is missing.

Now, for the rest of the proof, we condition on some outcome of the induced subgraph $G_0[Z]$ on the vertices in Z . To apply Lemma 9.6, we define a restricted colour system \mathcal{G}' of order n with parameters $(g, a_1, \dots, a_{g-1}, 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}, t_1, \dots, t_{g-1})$ by starting with \mathcal{G} , colouring the vertices in Z with colour g , and including all edges of our conditioned outcome of $G_0[Z]$ in a single shade of colour g . By construction, the core \mathcal{C}' of the restricted colour system \mathcal{G}' is almost isomorphic to \mathcal{C} ; the only difference is that \mathcal{C}' has the edges of $G_0[Z]$ between the vertices of colour g , whereas \mathcal{C} has no edges of colour g . (To be precise, there is a colour/shade-preserving graph isomorphism between \mathcal{C} and $\mathcal{C}' - E(G_0[Z])$).

Since we chose Z such that $v \in Z$ and $u \notin Z$, we have $u \in U(\mathcal{G}')$ (that is, u is uncoloured in \mathcal{G}'), and v has colour g in \mathcal{G}' . Recall that $u \in U_*$, meaning that u is connected to all vertices of colours $1, \dots, g-1$ in all possible shades of these colours. This implies $u \in U^*(\mathcal{G}')$. Now, let $e' \in E^g(\mathcal{C}')$ be the unique edge in the core \mathcal{C}' between v and the uncoloured vertex of \mathcal{C}' . As $v \in U_e$, this edge e' corresponds to the edge e in \mathcal{C} under the isomorphism in the previous paragraph. Note that the functions $\Gamma_{\mathcal{C},e}$ do not actually depend on the edges between the vertices of colour g in \mathcal{C} (since the partial copies of H in \mathcal{C} use exactly one vertex of colour g). Thus, we have $\Gamma_{\mathcal{C},e} = \Gamma_{\mathcal{C}',e'}$.

We have conditioned on an outcome of $G_0[Z]$. For each $z \in Z$, let $S_z = N_{G_0}(z) \cap U(\mathcal{G}')$ be the (random) set of neighbours of z in $U(\mathcal{G}') = U(\mathcal{G}) \setminus Z$, and let $G_0^- = G_0[U(\mathcal{G}')] \sim \mathbb{G}(U(\mathcal{G}'), p)$ be the induced subgraph on $U(\mathcal{G}')$. Then, with $\mathcal{S} = (S_z)_{z \in Z}$ we have $\mathcal{G}(G_0, j_1, \dots, j_{g-1}) = \mathcal{G}'_{\mathcal{S}}(G_0^-, j_1, \dots, j_{g-1}, 1)$ in the notation of Definition 4.8. It follows that $\kappa_H(\mathcal{G}, G_0, j_1, \dots, j_{g-1}, u, v) = \kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v)$, because every labelled copy of H using the edge uv as well as at least one vertex of each of the colours $1, \dots, g-1$ automatically also uses a vertex of colour g (namely v). Thus, (9.4) is equivalent to the inequality

$$\left| \kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v) - n^{h-g-1} \Gamma_{\mathcal{C}',e'}(j_1, \dots, j_{g-1}, 1) \right| = O\left(\log n \cdot n^{h-g-(3/2)}\right), \quad (9.5)$$

where $G_0^- \sim \mathbb{G}(U(\mathcal{G}'), p)$, and $\mathcal{S} = (S_z)_{z \in Z}$ is a collection of random sets with respect to the restricted colour system \mathcal{G}' as in Definition 4.8.

Note that \mathcal{G}' is essentially weakly p -general, due to the way it was defined in terms of \mathcal{G} . So, by Lemma 5.5, $\mathcal{G}'_{\mathcal{S}}$ is p -general with probability at least $1 - n^{-\omega(1)}$, and if $\mathcal{G}'_{\mathcal{S}}$ is p -general then by Lemma 9.6 we have $\Gamma_{\mathcal{C}',e'}(j_1, \dots, j_{g-1}, 1) n^{h-g-1} = \nu_{\mathcal{G}',\mathcal{S},u,v}(j_1, \dots, j_{g-1}, 1) + O(\log n \cdot n^{h-g-(3/2)})$. So, to conclude the proof of Lemma 9.10, it suffices to prove the following claim.

Claim 9.13. *With probability $1 - n^{-\omega(1)}$ we have*

$$\left| \kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v) - \nu_{\mathcal{G}',\mathcal{S},u,v}(j_1, \dots, j_{g-1}, 1) \right| \leq \log n \cdot n^{h-g-(3/2)}. \quad (9.6)$$

Proof. Condition on a fixed outcome of \mathcal{S} (so that only G_0^- remains random). Recall that by Definition 9.2 we have $\nu_{\mathcal{G}',\mathcal{S},u,v}(j_1, \dots, j_{g-1}, 1) = \mathbb{E}_{G_0^-}[\kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v)]$. Consider the vertex-exposure martingale for G_0^- (with respect to $\kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v)$), where we fix an ordering of $U(\mathcal{G}')$ (ending with u) and at each step we consider the next vertex in our ordering and expose all the edges of G_0^- incident to that vertex which have not yet been exposed. Changing the status of edges adjacent to a single vertex in $U(\mathcal{G}') \setminus \{u\}$ changes the value of $\kappa_H(\mathcal{G}'_{\mathcal{S}}, G_0^-, j_1, \dots, j_{g-1}, 1, u, v)$ by at most $h^{g+2} \cdot a_1 \dots a_{g-1} \cdot 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}} \cdot n^{h-g-2} = O(n^{h-g-2})$. This is due to the fact that there can be at most $h^{g+2} \cdot a_1 \dots a_{g-1} \cdot 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}} \cdot n^{h-g-2}$ different labelled copies of H in the n -vertex graph $\mathcal{G}_{\mathcal{S}}(G_0^-, j_1, \dots, j_{g-1}) + \{uv\}$ which use the edge uv as well as at least one vertex of each of the g colours and the exposed vertex (as both u and the exposed vertex are uncoloured). Thus, by the Azuma–Hoeffding inequality the probability that (9.6) fails to hold is at most

$$\exp\left(-\Omega\left(\frac{(\log n)^2 \cdot n^{2(h-g-3/2)}}{n \cdot n^{2(h-g-2)}}\right)\right) = \exp(-\Omega((\log n)^2)) = n^{-\omega(1)},$$

as desired. \square

9.4 Putting everything together

Proof of Lemma 7.3. Recall that the statement of Lemma 7.3 is for fixed integers $1 \leq g \leq h-1$ and $a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1} \geq 1$ (which were fixed throughout Section 7), and recall that $T = t_1 \dots t_{g-1}$. Let \mathcal{G} and λ be as in the statement of the lemma, and let \mathcal{C} be the extended core of \mathcal{G} (which is a complete core). Let $N = \binom{|U(\mathcal{G})|}{2} = \Theta(n^2)$, so that the random choice of $G_0 \sim \mathbb{G}(U(\mathcal{G}), p)$ can be encoded by a Bernoulli sequence $\xi \sim \text{Ber}(p)^N$, with one random bit for each of the N possible edges of G_0 . Abusing notation slightly, we identify the integers $1, \dots, N$ with pairs of vertices in $U(\mathcal{G})$, so that we may write $\xi_{\{u,v\}}$ to indicate the random bit that encodes the presence of the edge $\{u,v\}$.

Now, abusing notation, we index the coordinates of \mathbb{R}^T by tuples in $[t_1] \times \dots \times [t_{g-1}]$ (so that we may talk about the (j_1, \dots, j_{g-1}) -coordinate of a vector in \mathbb{R}^T). Let $\mathbf{f} : \{0, 1\}^N \rightarrow \mathbb{R}^T$ be the vector-valued

function defined such that, for $\xi \in \{0, 1\}^N$ corresponding to a graph G_0 , the (j_1, \dots, j_{g-1}) -coordinate of $\mathbf{f}(\xi)$ is $\psi_H(\mathcal{G}, G_0, j_1, \dots, j_{g-1})$. With this definition, and the notation of Theorem 3.1, the random vector $\Delta_{\{u,v\}}\mathbf{f}(\xi)$ corresponds to the function $\kappa_H(\mathcal{G}, G_0, \cdot, 1, u, v)$.

The plan is to now apply Theorem 3.1 with $d = T$ and $m = 2^{a_1 t_1 + \dots + a_{g-1} t_{g-1}}$ and with the $\Gamma_{\mathcal{C}, e}$ taking the role of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$. For each edge $e \in E^g(\mathcal{C})$, let $\gamma_e \in \mathbb{R}^T$ be the vector corresponding to the function $\Gamma_{\mathcal{C}, e}(\cdot, 1)$, and let $\mathbf{x} \in \mathbb{R}^T$ be the vector corresponding to the function $\lambda(\cdot, 1)$. By Fact 9.9, there is some $\varepsilon = \Omega(1)$ such that for each edge $e \in E^g(\mathcal{C})$, there are $\Omega(n^2) \geq \varepsilon N$ pairs of vertices $\{u, v\}$ with $u \in U_*, v \in U_e$. Let I_e be the set of these pairs $\{u, v\}$ (observe that all the I_e are disjoint). By Lemma 9.10, for each $\{u, v\} \in I_e$ we have $\Pr(\|\Delta_{\{u,v\}}\mathbf{f}(\xi) - s\gamma_e\|_\infty \geq r) \leq n^{-\omega(1)}$ for $s = n^{h-g-1}$ and some $r = \Theta(n^{h-g-3/2} \log n)$. Note that $r\sqrt{N \log N} \geq s$, and by Lemma 8.5, the vectors γ_e span \mathbb{R}^T . We can now apply Theorem 3.1 to obtain

$$\Pr(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{N \log N}) = O\left(\left(\frac{r\sqrt{\log N}}{s}\right)^T\right) = O\left((\log n)^{3T/2} n^{-T/2}\right) \leq n^{-T/2+o(1)}.$$

(The implied constants in the above asymptotic notation may *a priori* depend on \mathcal{C} , but note that there are only finitely many possibilities for a core with parameters $(g, a_1, \dots, a_{g-1}, t_1, \dots, t_{g-1})$). Finally, to conclude the proof we recall that $\|\psi_H(\mathcal{G}, G_0, \cdot) - \lambda\|_\infty = \|\mathbf{f}(\xi) - \mathbf{x}\|_\infty$ and observe that $r\sqrt{N \log N} = \Theta(n^{h-g-(1/2)}(\log n)^{3/2}) \geq n^{h-g-(1/2)} \log n$ for large n . \square

Proof of Lemma 6.3. This proof is very similar to the proof of Lemma 7.3. Recall that the statement of Lemma 6.3 is for fixed integers $1 \leq g \leq h-1$ and $a_1, \dots, a_g, t_1, \dots, t_{g-1} \geq 1$ (which were fixed throughout Section 6), and recall that $T = t_1 \cdots t_{g-1}$. Let \mathcal{G} and λ be as in the statement of the lemma, and let \mathcal{C} be the core of \mathcal{G} . Let $N = a_g \cdot |\mathcal{U}(\mathcal{G})| = \Theta(n)$, so that a random choice of \mathcal{S} as in Definition 4.8 can be encoded by a Bernoulli sequence $\xi \in \text{Ber}(p)^N$, with one random bit for each potential element in each $S_v \in \mathcal{S}$. Abusing notation slightly, we identify $1, \dots, N$ with ordered pairs of vertices: for $u \in \mathcal{U}(\mathcal{G})$ and a vertex v of colour g we write $\xi_{(u,v)}$ for the random bit that encodes the presence of u in S_v .

Let $\mathbf{f} : \{0, 1\}^N \rightarrow \mathbb{R}^T$ be the vector-valued function (with coordinates indexed by $[t_1] \times \dots \times [t_{g-1}]$) defined such that, for $\xi \in \{0, 1\}^N$ corresponding to an outcome of \mathcal{S} , the (j_1, \dots, j_{g-1}) -coordinate of $\mathbf{f}(\xi)$ is $\mu_{\mathcal{G}, \mathcal{S}}(j_1, \dots, j_{g-1}, 1)$. Then, $\Delta_{(u,v)}\mathbf{f}(\xi)$ corresponds to the function $\nu_{\mathcal{G}, \mathcal{S}, u, v}(\cdot, 1)$. For each $e \in E^g(\mathcal{C})$, let $\gamma_e \in \mathbb{R}^T$ be the vector corresponding to $\Gamma_{\mathcal{C}, e}(\cdot, 1)$, and let $\mathbf{x} \in \mathbb{R}^T$ be the vector corresponding to $\lambda(\cdot, 1)$.

By Fact 9.4, there is some $\varepsilon = \Omega(1)$ such that $|\mathcal{U}^*(\mathcal{G})| \geq \varepsilon n$. For each edge $e \in E^g(\mathcal{C})$ between the uncoloured vertex of \mathcal{C} and some vertex v of colour g , let $I_e = \mathcal{U}^*(\mathcal{G}) \times \{v\}$. By Lemma 5.5 the colour system $\mathcal{G}_{\mathcal{S}}$ is p -general with probability $1 - n^{-\omega(1)}$, in which case, by Lemma 9.6, for each $(u, v) \in I_e$ we have $\|\Delta_{(u,v)}\mathbf{f}(\xi) - s\gamma_e\|_\infty \leq r$ for $s = n^{h-g-1}$ and some $r = \Theta(n^{h-g-(3/2)} \log n)$. Also, by Lemma 8.5, the vectors γ_e span \mathbb{R}^T .

We can now apply Theorem 3.1 to obtain $\Pr(\|\mathbf{f}(\xi) - \mathbf{x}\|_\infty < r\sqrt{N \log N}) \leq n^{-T/2+o(1)}$. Finally, to conclude the proof we observe that $r\sqrt{N \log N} = \Theta(n^{h-g-1}(\log n)^{3/2}) \geq n^{h-g-1} \log n$ for large n , and recall that $\|\mu_{\mathcal{G}, \mathcal{S}} - \lambda\|_\infty = \|\mathbf{f}(\xi) - \mathbf{x}\|_\infty$. \square

10 Concluding remarks

We have proved that for connected H and constant $p \in (0, 1)$, we have $\max_x \Pr(X_H = x) \leq n^{1-v(H)+o(1)}$. There are several interesting directions for future research. Most obviously, Conjecture 1.1 remains open: for connected H we are still a factor of $n^{o(1)}$ away from an optimal bound, and for disconnected H we do not even have a bound that improves as H grows (the best general bound is $\Pr(X_H = x) \leq \Pr(|X_H - x| \leq n^{v(H)-2}) = O(1/n)$, as we mentioned in the introduction). It seems that the ideas in this paper are robust enough to give certain nontrivial bounds (in terms of the size of the largest component of H) even in the disconnected case, but we have not explored this further.

For certain graphs H , a possible route to a proof of Conjecture 1.1 might be via a local central limit theorem, which one might hope to prove by extending the methods of Gilmer and Kopparty [20], and Berkowitz [6, 7]. Basically, this involves carefully estimating the characteristic function $\varphi(t) = \mathbb{E}e^{itX_H}$, using different arguments for different ranges of t . We remark that $\varphi(1/k)$ is small if the distribution

of X_H is not too biased mod k , which seems comparable to anticoncentration of X_H at “scale” k . So, we wonder whether the ideas in this paper might be helpful for estimating φ : recall that our argument proceeds by breaking up X_H into a sum of many random variables that fluctuate at different scales. However, we emphasise that local central limit theorems do not seem to be the right path to a proof of Conjecture 1.1 in its full generality: for example, if H is a disjoint union of an edge and a 2-edge path, then the probability that X_H is odd is substantially different from the probability that it is even (see [16]), meaning that X_H does not obey a local central limit theorem.

Also, let X_H^{hom} be the number of (possibly non-injective) homomorphisms from H into $G \sim \mathbb{G}(n, p)$. This random variable is very closely related to X_H , and we remark that with very minimal changes, one can modify our proof of Theorem 1.2 to prove the corresponding theorem for X_H^{hom} , when H is connected. Interestingly, the homomorphism-counting analogue of Conjecture 1.1 fails dramatically in general: if H is the disjoint union of two copies of a graph H' , then $X_H^{\text{hom}} = (X_{H'}^{\text{hom}})^2$, meaning that X_H has the same point probabilities as $X_{H'}$. This means that any proof of Conjecture 1.1 must be sensitive to the difference between subgraph counts and homomorphism counts.

It would also be interesting to consider the “sparse” regime where p is allowed to decay with n . For example, it is known that if H is *strictly balanced* (see for example [23, Section 3.2]) and $p = p(n)$ is such that $\text{Var} X_H = o((\mathbb{E}X_H)^2)$, then $\Pr(X_H > 0) = 1 - o(1)$. Could it be that under these conditions we also have that $\max_{x \in \mathbb{Z}} \Pr(X_H = x) = O(1/\sqrt{\text{Var} X_H})$?

As mentioned in [19] it may also be interesting to study anticoncentration of the number of *induced* copies X'_H of a subgraph H in a random graph $\mathbb{G}(n, p)$. (This question was also raised by Meka, Nguyen and Vu [30]). The natural analogue of Conjecture 1.1 is that for a fixed graph H and fixed $p \in (0, 1)$, we have

$$\max_{x \in \mathbb{N}} \Pr(X'_H = x) = O\left(\frac{1}{\sqrt{\text{Var}(X'_H)}}\right).$$

We remark that the behaviour of $\sqrt{\text{Var}(X'_H)}$ is not entirely trivial: for most values of p it has order $\Theta(n^{h-1})$, but when p is exactly equal to the edge-density of H it may have order $\Theta(n^{h-3/2})$ or $\Theta(n^{h-2})$ (see [23, Theorem 6.42]).

Finally, it would be interesting to prove similar anticoncentration results in other combinatorial settings. One important example is random subsets of the integers (or other groups): for instance, what can we say about anticoncentration of the number of k -term arithmetic progressions in a random subset of $\{1, \dots, n\}$? Arithmetic configuration counts have been an interesting analogue to subgraph counts in a number of other settings, for example in the study of large deviations (both fall in the framework of *nonlinear large deviations* initiated by Chatterjee and Dembo [12]; see for example [22] and the references therein). Another interesting direction of research would be to consider subgraph counts in random k -uniform hypergraphs, or for other random graph models (for example, the uniform distribution $\mathbb{G}(n, m)$ on graphs with a fixed set of n vertices and exactly m edges).

Remark added in proof. While this paper was under review, Sah and Sawhney [36] proved Conjecture 1.1 for connected H (via a local limit theorem) and disproved Conjecture 1.1 in general.

Acknowledgements. We thank the referees for their careful reading, and for many helpful comments.

References

- [1] N. Alon and J. H. Spencer, *The probabilistic method*, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.
- [2] F. Augeri, *Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős-Renyi graphs*, Ann. Probab., to appear.
- [3] A. D. Barbour, *Poisson convergence and random graphs*, Math. Proc. Cambridge Philos. Soc. **92** (1982), no. 2, 349–359.
- [4] A. D. Barbour, M. Karoński, and A. Ruciński, *A central limit theorem for decomposable random variables with applications to random graphs*, J. Combin. Theory Ser. B **47** (1989), no. 2, 125–145.

- [5] A. Basak and R. Basu, *Upper tail large deviations of regular subgraph counts in Erdős-Rényi graphs in the full localized regime*, arXiv preprint arXiv:1912.11410 (2019).
- [6] R. Berkowitz, *A quantitative local limit theorem for triangles in random graphs*, arXiv preprint arXiv:1610.01281 (2016).
- [7] R. Berkowitz, *A local limit theorem for cliques in $G(n, p)$* , arXiv preprint arXiv:1811.03527 (2018).
- [8] B. B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao, *Upper tails and independence polynomials in random graphs*, *Adv. Math.* **319** (2017), 313–347.
- [9] B. Bollobás, *Threshold functions for small subgraphs*, *Math. Proc. Cambridge Philos. Soc.* **90** (1981), no. 2, 197–206.
- [10] B. Bollobás and J. C. Wierman, *Subgraph counts and containment probabilities of balanced and unbalanced subgraphs in a large random graph*, *Graph theory and its applications: East and West* (Jinan, 1986), *Ann. New York Acad. Sci.*, vol. 576, New York Acad. Sci., New York, 1989, pp. 63–70.
- [11] S. Chatterjee, *Large deviations for random graphs*, *Lecture Notes in Mathematics*, vol. 2197, Springer, Cham, 2017, Lecture notes from the 45th Probability Summer School held in Saint-Flour, June 2015, École d’Été de Probabilités de Saint-Flour.
- [12] S. Chatterjee and A. Dembo, *Nonlinear large deviations*, *Adv. Math.* **299** (2016), 396–450.
- [13] N. Cook and A. Dembo, *Large deviations of subgraph counts for sparse Erdős-Rényi graphs*, *Adv. Math.* **373** (2020).
- [14] B. DeMarco, J. Kahn, and A. Redlich, *Modular statistics for subgraph counts in sparse random graphs*, *Electron. J. Combin.* **22** (2015), no. 1, Paper 1.37, 6 pages.
- [15] R. Eldan, *Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations*, *Geom. Funct. Anal.* **28** (2018), no. 6, 1548–1596.
- [16] B. DeMarco and A. Redlich, *Graph decomposition and parity*, *J. Graph Theory* **82** (2016), no. 4, 374–386.
- [17] P. Erdős and A. Rényi, *On the evolution of random graphs*, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.
- [18] P. Erdős, *On a lemma of Littlewood and Offord*, *Bull. Amer. Math. Soc.* **51** (1945), 898–902.
- [19] J. Fox, M. Kwan, and L. Saueremann, *Combinatorial anti-concentration inequalities, with applications*, *Math. Proc. Cambridge Philos. Soc.*, to appear.
- [20] J. Gilmer and S. Kopparty, *A local central limit theorem for triangles in a random graph*, *Random Structures Algorithms* **48** (2016), no. 4, 732–750.
- [21] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, *Period. Math. Hungar.* **8** (1977), no. 3-4, 197–211.
- [22] M. Harel, F. Mousset, and W. Samotij, *Upper tails via high moments and entropic stability*, arXiv preprint arXiv:1904.08212 (2019).
- [23] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Cambridge University Press, 2000.
- [24] M. Karoński, *Balanced subgraphs of large random graphs*, *Seria Matematyka [Mathematics Series]*, vol. 7.
- [25] M. Karoński and A. Ruciński, *On the number of strictly balanced subgraphs of a random graph*, *Graph theory* (Łagów, 1981), *Lecture Notes in Math.*, vol. 1018, Springer, Berlin, 1983, pp. 79–83.
- [26] P. G. Kolaitis and S. Kopparty, *Random graphs and the parity quantifier*, *J. ACM* **60** (2013), no. 5, Art. 37, 34.
- [27] M. Loeb, J. Matoušek, and O. Pangrác, *Triangles in random graphs*, *Discrete Math.* **289** (2004), no. 1-3, 181–185.

- [28] D. Lubell, *A short proof of Sperner's lemma*, J. Combinatorial Theory **1** (1966), 299.
- [29] E. Lubetzky and Y. Zhao, *On the variational problem for upper tails in sparse random graphs*, Random Structures Algorithms **50** (2017), no. 3, 420–436.
- [30] R. Meka, O. Nguyen, and V. Vu, *Anti-concentration for polynomials of independent random variables*, Theory Comput. **12** (2016), Paper No. 11, 16 pages.
- [31] H. H. Nguyen and V. H. Vu, *Small ball probability, inverse theorems, and applications*, Erdős centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 409–463.
- [32] K. Nowicki and J. C. Wierman, *Subgraph counts in random graphs using incomplete U -statistics methods*, Proceedings of the First Japan Conference on Graph Theory and Applications (Hakone, 1986), vol. 72, 1988, pp. 299–310.
- [33] N. Ross, *Fundamentals of Stein's method*, Probab. Surv. **8** (2011), 210–293.
- [34] A. Ruciński, *When are small subgraphs of a random graph normally distributed?*, Probab. Theory Related Fields **78** (1988), no. 1, 1–10.
- [35] A. Ruciński and A. Vince, *Balanced graphs and the problem of subgraphs of random graphs*, Proceedings of the sixteenth Southeastern international conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1985), vol. 49, 1985, pp. 181–190.
- [36] A. Sah, M. Sawhney, *Local limit theorems for subgraph counts*, arXiv preprint arXiv:2006.11369 (2020).
- [37] K. Schürger, *Limit theorems for complete subgraphs of random graphs*, Period. Math. Hungar. **10** (1979), no. 1, 47–53.
- [38] T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2010.
- [39] T. Tao and V. H. Vu, *The Littlewood-Offord problem in high dimensions and a conjecture of Frankl and Füredi*, Combinatorica **32** (2012), 363–372.