11. Azuma's Inequality and a Strengthening of Brooks' Theorem

11.1 Azuma's Inequality

In this chapter, we introduce a new tool for proving bounds on concentration. It differs from the tools we have mentioned so far, in that it can be applied to a sequence of dependent trials. To see a concrete example of such a situation, imagine that we are colouring the vertices of a graph one by one, assigning to each vertex a colour chosen uniformly *from those not yet assigned to any of its coloured neighbours.* This ensures that the colouring obtained is indeed a proper colouring, and analyzing such a random process may yield good bounds on the minimum number of colours required to obtain vertex colourings with certain properties. However, our choices at each vertex are now no longer independent of those made at the other vertices.

In this chapter, we introduce a new concentration inequality which can be used to handle such situations, Azuma's Inequality. It applies to a random variable R which is determined by a sequence $X_1, \ldots, X_n$ of random trials. Our new tool exploits the ordering on the random trials; it obtains a bound on the concentration of R using bounds on the maximum amount by which we *expect* each trial to affect R when it is performed. This approach bears fruit even when we are considering a set of independent trials. To illustrate this point, we consider the following simple game.

A player rolls a fair six-sided die $n+1$ times, with outcomes $r_0, r_1, \ldots, r_n$. Roll 0 establishes a target. The players winnings are equal to X, the number of rolls $i \geq 1$ such that $r_i = r_0$.

It should be intuitively clear that X is highly concentrated. However, changing the outcome of r_0 can have a dramatic effect on X. For example, if our sequence is $1, 1, \ldots, 1$ then any change to r_0 will change X from n to 0. Thus, we cannot directly apply the Simple Concentration Bound or Talagrand's Inequality to this problem. It turns out that we can apply Azuma's Inequality because the conditional expected value of X after the first roll is $\frac{n}{6}$ regardless of the outcome of this trial. Thus, the first trial has no effect whatsoever on the conditional expected value of X.

Remark Of course, we do not need a bound as powerful as Azuma's Inequality, to prove that X is highly concentrated – we could prove this by first rolling the die to determine r_0 and then simply applying the Chernoff Bound to the sequence $r_1, \ldots, r_n$. However, the reader can easily imagine that she could contrive a similar but more complicated scenario where it is not so easy to apply our other bounds. For one such example, see Exercise 11.1.

Like Talagrand's Inequality, Azuma's Inequality can be viewed as a strengthening of the Simple Concentration Bound. There are three main differences between Azuma's Inequality and the Simple Concentration Bound.

(i) We must compute a bound on the amount by which the outcome of each trial can affect the *conditional* expected value of X,
(ii) Azuma's Inequality can be applied to sequences of dependent trials and is therefore much more widely applicable,
(iii) The concentration bound given by the Simple Concentration Bound is in terms of an upper bound c on the maximum amount by which changing the outcome of a trial can affect the value of X. To apply Azuma's Inequality we obtain distinct values $c_1, \ldots, c_n$ where c_i bounds the amount by which changing the outcome of T_i can affect the conditional expected value of X. We then express our concentration bound in terms of $c_1, \ldots, c_n$. This more refined approach often yields stronger results.

Azuma's Inequality [14] *Let X be a random variable determined by n trials $T_1, \ldots, T_n$, such that for each i, and any two possible sequences of outcomes $t_1, \ldots, t_i$ and $t_1, \ldots, t_{i-1}, t_i'$:*

$$|\,\mathbf{Exp}(X\,|\,T_1 = t_1, \ldots, T_i = t_i) - \mathbf{Exp}(X\,|\,T_1 = t_1, \ldots, T_i = t_i')\,| \leq c_i \quad (11.1)$$

then

$$\mathbf{Pr}(|X - \mathbf{E}(X)| > t) \leq 2\mathrm{e}^{-t^2/(2\sum c_i^2)}.$$

Condition (11.1) corresponds to condition (10.1) in the Simple Concentration Bound, however the two inequalities are very different. The following discussion underscores the difference between them. Suppose we have an adversary who is trying to make X as large as he can, and a second adversary who is trying to make X as small as she can. Either adversary is allowed to change the outcome of exactly one trial T_i. Condition (10.1) says that if the adversaries wait until all trials have been carried out, and then change the outcome of T_i, then their power is always limited. Condition (11.1) says that if they must make their changes as soon as T_i is carried out, without waiting for the outcomes of all future trials, then their power is limited.

The above discussion suggests that condition (10.1) is more restrictive than condition (11.1), and thus that Azuma's Inequality implies the Simple Concentration Bound. It is, in fact, straightforward to verify this implication (this is Exercise 11.2). As we will see, Azuma's Inequality is actually much more powerful than the Simple Concentration Bound. For example, in the game discussed earlier, we satisfy condition (11.1) with $c_0 = 0$ and $c_i = 1$ for $i > 0$ (by Linearity of Expectation) and so Azuma's Inequality implies that X is highly concentrated.

Azuma's Inequality is an example of a Martingale inequality. For further discussion of Martingale inequalities, we refer the reader to [10], [112] or [114].

If we apply Azuma's Inequality to a set of independent trials with each c_i equal to a small constant c, then the resulting bound is $e^{-\epsilon t^2/n}$ for a positive constant ϵ rather than the often more desirable $e^{-\epsilon t^2/\mathbf{E}(X)}$ which is typically obtained from Talagrand's Inequality. While the bound given by applying Azuma's inequality using such c_i is usually sufficient when $\mathbf{E}(X) = \beta n$ for some constant $\beta > 0$, it is often not strong enough when $\mathbf{E}(X) = o(n)$. However, in such situations, by taking at least some of the c_i to be very small, we can often apply Azuma's Inequality to get the desired bound of $e^{-(\epsilon t^2/\mathbf{E}(X))}$. We will see an example of this approach in the proof of Lemma 11.8 later in this chapter.

We now consider variables determined by sequences of dependent trials, where the change in the conditional expectation caused by each trial is bounded. Our discussion focuses on one commonly occurring situation. Suppose X is a random variable determined by a uniformly random permutation P of $\{1, \ldots, n\}$, with the property that interchanging any two values $P(i), P(j)$ can never affect X by more than c. Then, as we discuss below, we can apply Azuma's Inequality to show that X is concentrated.

For each $1 \leq i \leq n$, we let T_i be a uniformly random element of $\{1, \ldots, n\} - \{T_1, \ldots, T_{i-1}\}$. It is easy to see that $T_1, \ldots, T_n$ forms a uniformly random permutation. Furthermore, we will show that this experiment satisfies condition (11.1).

Consider any sequence of outcomes $T_1 = t_1, \ldots, T_{i-1} = t_{i-1}$, along with two possibilities for T_i, t_i, t_i'. For any permutation P satisfying $P(1) = t_1, \ldots, P(i) = t_i$ and $P(j) = t_i'$ for some $j > i$, we let P' be the permutation obtained by interchanging $P(i)$ and $P(j)$. Our hypotheses yield $|X(P) - X(P')| \leq c$. Furthermore, it is easy to see that

$$\mathbf{Pr}(P|T_1 = t_1, \ldots, T_i = t_i) = \mathbf{Pr}(P'|T_1 = t_1, \ldots, T_i = t_i') = \frac{1}{(n-i)!}.$$

It is straightforward to verify that these two facts ensure that condition (11.1) holds, and so we can apply Azuma's Inequality to show that X is highly concentrated. (Note that it is important that Azuma's Inequality does not require our random trails to be independent.) Of course, Azuma's Inequality also applies in a similar manner when X is determined by a sequence of several random permutations.

Remark As discussed in the last chapter, we can generate a uniformly random permutation by generating n independent random reals between 0 and 1. We applied Talagrand's Inequality to this model to prove that the length of the longest increasing subsequence is concentrated. We cannot use Talagrand's inequality in the same way to prove the above result as the condition that swapping two values $P(i)$ and $P(j)$ can affect X by at most c

does not guarantee that changing the value of one of the random reals affects the value of X by a bounded amount.

For a long time, Azuma's Inequality (or, more generally, the use of Martingale inequalities) was the best way to prove many of the difficult concentration bounds arising in probabilistic combinatorics. However, the conditions of Talagrand's inequality are often much easier to verify. Thus in situations where they both apply, Talagrand's Inequality has begun to establish itself as the "tool of choice".

It is worth noting, in this vein, that Talagrand showed that his inequality can also be applied to a single uniformly random permutation (see Theorem 5.1 of [148]). More recently, McDiarmid obtained a more general version which applies to sequences of several permutations, as we will discuss in Chap. 16. Thus, we can now prove concentration for variables which depend on such a set of random trials, using a Talagrand-like inequality rather than struggling with Azuma. To see the extent to which this simplifies our task, compare some of the lengthy concentration proofs in [132] and [119] (which predated McDiarmid's extension of Talagrand's Inequality) with the corresponding proofs in Chaps. 16 and 18 of this book.

Nevertheless, there are still many sequences of dependent trials to which Talagrand cannot be applied but Azuma's Inequality can (see for example [116]).

11.2 A Strengthening of Brooks' Theorem

Brooks' Theorem characterizes graphs for which $\chi \leq \Delta$. For Δ at least 3, they are those that contain no $\Delta + 1$ clique. Characterizing which graphs have $\chi \leq \Delta - 1$ seems to be more difficult, Maffray and Preissmann [110] have shown it is NP-complete to determine if a 4-regular graph has chromatic number at most three (if you do not know what *NP-complete* means, replace it by *hard*). However, Borodin and Kostochka [29] conjectured that if $\Delta(G) \geq 9$ then an analogue of Brooks' Theorem holds; i.e. $\chi(G) \leq \Delta(G) - 1$ precisely if $\omega(G) \leq \Delta - 1$ (this is Problem 4.8 in [85] to which we refer readers for more details). To see that 9 is best possible here, consider the graph G, depicted in Fig. 11.1, obtained from five disjoint triangles $T_1, \ldots, T_5$ by adding all edges between T_i and T_j if $|i - j| \equiv 1 \bmod 5$. It is easy to verify that $\Delta(G) = 8$, $\omega(G) = 6$, and $\chi(G) = 8$. Beutelspacher and Hering [22] independently posed the weaker conjecture that this analogue of Brooks' Theorem holds for sufficiently large Δ. We prove their conjecture. That is, we show:

Theorem 11.1 *There is a Δ_2 such that if $\Delta(G) \geq \Delta_2$ and $\omega(G) \leq \Delta(G) - 1$ then $\chi(G) \leq \Delta(G) - 1$.*

It would be natural to conjecture that Theorem 11.1 could be generalized as follows:

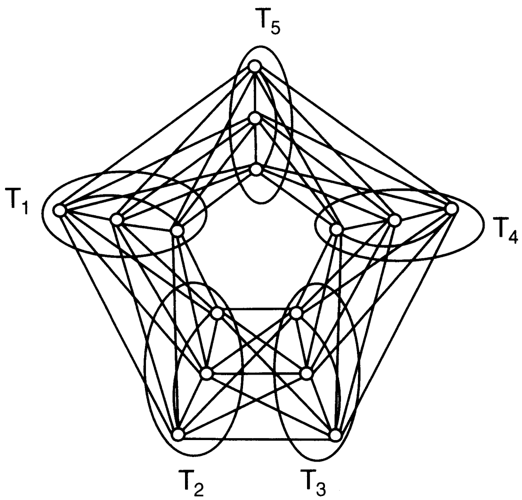

Fig. 11.1. G

For all k, there is a Δ_k such that if $\Delta(G) \geq \Delta_k$ and $\omega(G) \leq \Delta(G)+1-k$ then $\chi(G) \leq \Delta(G)+1-k$.

However this conjecture turns out to be false even for $k = 3$, as the following example shows: For $\Delta \geq 5$, let G_Δ be a graph obtained from a clique $K_{\Delta-4}$ with $\Delta - 4$ vertices and a chordless cycle C with 5 vertices by adding all edges between C and $K_{\Delta-4}$ (see Fig. 11.2). It is easy to verify that G_Δ has maximum degree Δ, clique number $\Delta - 2$, and chromatic number $\Delta - 1$.

Nevertheless, we can generalize Theorem 11.1 in two ways. Firstly, we can bound how quickly χ must decrease as ω moves away from $\Delta + 1$.

Theorem 11.2 *For all k, there is a Δ_k such that if $\Delta(G) \geq \Delta_k$ and $\omega(G) \leq \Delta(G) + 1 - 2k$ then $\chi(G) \leq \Delta(G) + 1 - k$.*

This result is a corollary of Theorem 16.4 discussed in Chap. 16. As pointed out in that chapter, the theorem is essentially best possible for large k.

Secondly, we can show that if χ is sufficiently near Δ then although we may not be able to determine χ precisely simply by considering the sizes of the cliques in G, we can determine it by considering only the chromatic numbers of a set of subgraphs of G which are very similar to cliques. For example, we have:

There is a Δ_0 such that for any $\Delta \geq \Delta_0$ and $k > \Delta - \sqrt{\Delta} + 2$,

> there is a collection of graphs $H_1, \ldots, H_t$, which are similar to k-cliques in that $\chi(H_i) = k$, $|V(H_i)| \leq \Delta + 1$ and $\delta(H_i) \geq k - 1$, such that the following holds:
> For any graph G with maximum degree Δ, $\chi(G) \geq k$ iff G contains at least one H_i as a subgraph.

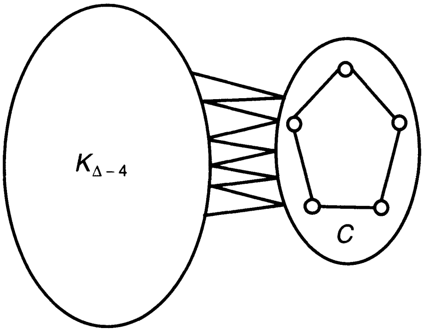

Fig. 11.2. G_Δ

We discuss a number of results of this type in Chap. 15. The proofs of these generalizations of Theorem 11.1 all use ideas introduced in its proof.

Proof of Theorem 11.1. We fix a Δ_2 which satisfies a number of implicit inequalities scattered throughout the proof and prove the theorem for this value of Δ_2. To this end, we assume the theorem is false and let G be a counter-example to it with the fewest number of vertices. Thus, G has maximum degree $\Delta \geq \Delta_2$, $\omega(G) \leq \Delta - 1$, and $\chi(G) = \Delta$.

Before presenting the key ideas of the proof, we make the following easy observations.

11.3 *Every subgraph H of G with $H \neq G$ has a $\Delta - 1$ colouring.*

Proof If $\Delta(H) = \Delta$ then the result follows by the minimality of G. Otherwise, the result follows from Brooks' Theorem because $\omega(H) \leq \omega(G) \leq \Delta - 1$. □

11.4 *Every vertex of G has degree at least $\Delta - 1$.*

Proof For any vertex v of G, by (11.3), $G - v$ has a $\Delta - 1$ colouring. If v had fewer than $\Delta - 1$ neighbours then we could extend this to a $\Delta - 1$ colouring of G. □

We have already seen in Sect. 10.3 that for Δ sufficiently large, if no vertex in G has more than $\binom{\Delta}{2} - (\log\ \Delta)^3 \Delta$ edges in its neighbourhood, then $\chi(G) \leq \Delta - \frac{(\log\ \Delta)^3}{e^6} \leq \Delta - 1$. Thus, the crux of the proof will be to deal with vertices which have such dense neighbourhoods. This motivates the following:

Definitions A vertex v of G is *dense* if $N(v)$ contains fewer than $\Delta(\log\ \Delta)^3$ non-adjacent pairs of vertices. Otherwise, it is *sparse*.

We need to investigate the structure surrounding such dense vertices.

Definition Set $r = (\log\ \Delta)^4$. We say a clique is *big* if $|C| \geq \Delta - r$.

We shall prove:

Lemma 11.5 *Every dense vertex is contained in a big clique.*

Lemma 11.6 *We can partition $V(G)$ into $D_1, \dots, D_l, S$ so that*

(i) *each D_i contains a big clique C_i Furthermore, either $D_i = C_i$ or $D_i = C_i + v_i$ for some vertex v_i which sees at least $\frac{3\Delta}{4}$ but not all of the vertices of C_i;*
(ii) *no vertex of $V - D_i$ sees more than $\frac{3\Delta}{4}$ vertices of D_i;*
(iii) *each vertex of S is sparse; and*
(iv) *each vertex v of C_i has at most one neighbour outside C_i which see more than $r+4$ vertices of C_i, furthermore if $|C_i| = \Delta - 1$ then v has no such neighbours.*

It is this decomposition of G into l dense sets and a set of sparse vertices which allows us to prove Theorem 11.1. For, having proved Lemma 11.6, to deal with the dense vertices we need only colour each D_i. This will be relatively easy, for these sets are disjoint and there are very few edges from D_i to $V - D_i$.

For ease of exposition, we consider the Δ-regular graph G' obtained from G by taking two copies of G and adding an edge between the two copies of each vertex of degree $\Delta - 1$. We note that applying Lemma 11.6 to both copies of G yields:

Corollary 11.7 *There is a decomposition of G' satisfying conditions (i)–(iv) of Lemma 11.6.*

Now, by taking advantage of this corollary, we can extend the proof technique of the last chapter to prove a useful lemma.

Definition Consider a decomposition as in Corollary 11.7. For $1 \leq i \leq l$, if D_i is the clique C_i set $K_i = C_i$ else set $K_i = C_i \cap N(v_i)$.

Lemma 11.8 *There is a partial $\Delta - 1$ colouring of G' satisfying the following two conditions.*

(a) *for every vertex $v \in S$ there are at least 2 colours appearing twice in the neighbourhood of v,*
(b) *every K_i contains two uncoloured vertices w_i and x_i whose neighbourhoods contain two repeated colours.*

To complete a partial $\Delta - 1$ colouring satisfying (a) and (b) to a $\Delta - 1$ colouring of G' and thereby obtain a $\Delta - 1$ colouring of G, we proceed as follows.

We let U_2 be the set of uncoloured vertices whose neighbourhoods contain at least two repeated colours and we let U_1 be the remaining uncoloured vertices. We complete the colouring greedily by colouring the uncoloured vertices one at a time. The only trick is to colour all the vertices of U_1 before colouring any vertex of U_2. Consider a vertex v of U_1. Since every sparse vertex is in U_2 by (a), v is in some D_i. By (b), the vertices x_i and w_i are in U_2 and hence are uncoloured when we come to colour v. Since v has these two uncoloured neighbours, it follows that there is a colour available with which to colour v. Thus, we can extend our partial colouring of G' to a $\Delta - 1$ colouring of $G' - U_2$. We can then complete the colouring because for each vertex u in U_2 there are two repeated colours in $N(u)$. □

Remark We note that in proving Theorem 11.1 from Lemma 11.8, we used a slight refinement of our greedy colouring procedure. We carefully chose the order in which we would complete the colouring, and then coloured each vertex greedily when we came to it. This idea, which we first encountered in the proof of Brooks' Theorem in Chap. 1, will prove crucial to many of the proofs to follow both in this chapter and future ones.

We have yet to prove Lemmas 11.5, 11.6, and 11.8. We prove the last of these in the next section, and prove the first two in the third and final section of the chapter.

11.3 The Probabilistic Analysis

Proof of Lemma 11.8. We find a partial $\Delta - 1$ colouring satisfying conditions (a) and (b) of Lemma 11.8 by analyzing our naive colouring procedure. In doing so, we take advantage of the partition given by Corollary 11.7. Once again, we will use the Local Lemma.

To do so, we need to define two kinds of events. For each $v \in S$, we let A_v be the event that (a) fails to hold for v, i.e., that there are fewer than 2 repeated colours on $N(v)$. For each D_i, we let A_i be the event that (b) fails to hold for D_i, i.e. that there do not exist two uncoloured vertices of K_i each of which has two repeated colours in its neighbourhood. We note that if none of the events in the set $\mathcal{E} = (\cup A_v) \cup (\cup A_i)$ hold then the random colouring satisfies (a) and (b) of Lemma 11.8. To finish the proof we use the Local Lemma to show that this occurs with positive probability.

We note that A_v depends only on the colours within distance two of v. Also each A_i depends only on the colour of the vertices in D_i or within distance two of D_i. It follows that each event in $\mathcal{E}$ is mutually independent of a set of all but at most Δ^5 other events. So, we need only show that each event in $\mathcal{E}$ holds with probability at most Δ^{-6}.

11.9 *Each A_v has probability at most Δ^{-6}.*

To prove this result we consider (as in the last chapter) the variable X_v which counts the number of colours assigned to at least two neighbours of v and retained by all such neighbours. We first obtain a bound on the expected value of X_v and then bound its concentration using Talagrand's Inequality. As the proof is almost identical to one in the last chapter, we omit the details.

11.10 *Each A_i has probability at most Δ^{-6}.*

To prove this result, we need the following simple corollary of Corollary 11.7.

Lemma 11.11 *For any D_i, there are at least $\frac{\Delta}{4r}$ disjoint triples each of which consists of a vertex v of K_i and two neighbours of v outside of K_i both of which have at most $r+4$ neighbours in K_i.*

Proof of Lemma 11.11. Consider any D_i, and corresponding K_i. By definition, $|K_i| \geq \frac{3\Delta}{4}$. Take a maximal set of disjoint triples each of which consists of a vertex in K_i and two neighbours of this vertex outside of D_i, each of which has at most $r+4$ neighbours in K_i. Suppose there are k triples in this set and let W be the $2k$ vertices in these triples which are not in D_i. By (iv) of Corollary 11.7 and the maximality of our set of triples, every vertex in K_i is a neighbour of some vertex in W. Hence, $(r+4)|W| \geq |K_i|$, which yields $k \geq \frac{\Delta}{4r}$ as required. □

To compute the probability bound on A_i, we consider the set $\mathcal{T}_i$ of $\frac{\Delta}{4r}$ disjoint triples guaranteed to exist by Lemma 11.11. We let T_i be the union of the vertex sets of these triples. We let M_i be the number of these triples for which (i) the vertex in K_i is uncoloured, (ii) both the other vertices are coloured with a colour which is also used to colour a vertex of K_i, and (iii) no vertex of the triple is assigned a colour assigned to any other vertex in T_i. This last condition is present to ensure that changing the colour of a vertex can only affect the value of M_i by two.

To begin, we compute the expected value of M_i. We note that M_i counts the number of triples (a,b,c) in $\mathcal{T}_i$ with $c \in K_i$ such that there are colours j,k,l and vertices x,y,z with $x \in K_i - T_i - N(a)$, $y \in K_i - T_i - N(b)$, $z \in N(c) - T_i$, such that

1. j is assigned to a and x but to none of the rest of $T_i \cup N(a) \cup N(x)$,
2. k is assigned to b and y but to none of the rest of $T_i \cup N(b) \cup N(y)$,
3. l is assigned to z and c but on none of the rest of T_i.

To begin, we fix a triple $\{a,b,c\}$ in $\mathcal{T}_i$. We let $A_{j,k,l,x,y,z}$ be the event that (1), (2), and (3) hold. Since $|T_i| \leq \frac{\Delta}{4}$, the probability of $A_{j,k,l,x,y,z}$ is at least $(\Delta-1)^{-6}\frac{1}{e^5}$. Furthermore, two such events with different sets of indices are disjoint. Now, there are at least $\frac{2\Delta}{3}$ choices for both x and y. There are at least $\frac{9\Delta}{10}$ choices for z and $(\Delta-1)(\Delta-2)(\Delta-3)$ choices for distinct j,k,l. So, a straightforward calculation shows that the probability that (1), (2), and (3) hold for some choice of $\{j,k,l,x,y,z\}$ is at least $(\Delta-1)^{-6}(\Delta-1)(\Delta-2)(\Delta-3)\frac{2\Delta}{3}^2\frac{9\Delta}{10}\frac{1}{e^5} \geq \frac{1}{e^6}$. Since, there are $\frac{\Delta}{4r}$ triples in $\mathcal{T}_i$, the expected value of M_i is at least $\frac{\Delta}{4r}\frac{1}{e^6} \geq \frac{\Delta}{r^2}$.

We now prove that M_i is concentrated around its mean, and hence at least two with high probability, by applying Azuma's Inequality. To apply Azuma's Inequality, we must be careful about the order in which we assign the random colours to $V(G)$. We will colour the vertices of $V - T_i - K_i$ first and then the vertices of $T_i \cup K_i$. So, we order the vertices of G as $w_1, \ldots, w_n$ where for some s, we have $\{w_1, \ldots, w_s\} = V - T_i - K_i$ and $\{w_{s+1} =, \ldots, w_n\} = T_i \cup K_i$. We then choose the random colour assignments for the vertices in the given order.

For each of these choices we now obtain our bound c_j on the effect of the choice on the conditional expected value of $\mathbf{E}(M_i)$. We note that changing the colour of any vertex can affect the conditional expected value of M_i by at most 2 since it affects the value of M_i by at most 2 for any given assignment of colours to the remaining vertices. So, $\sum_{n-s}^{n} c_j^2 \le 2^2|T_i \cup K_i|$ $\le 5\Delta$. Furthermore, changing the colour assigned to a vertex w_j of $V - T_i - K_i$ from α to β will only affect M_i if some neighbour of w_j in $T_i \cup K_i$ receives either α or β. This occurs with probability at most $\frac{2d_j}{\Delta}$ where d_j is the number of neighbours of w_j in $T_i \cup K_i$. Hence by colouring w_j we can change the conditional expected value of M_i by at most $c_j = \frac{4d_j}{\Delta}$. Since the d_j sum to at most Δ^2, $\sum_{i=1}^{n-s} c_j$ is at most 4Δ. As, each c_j is at most 4, we see that $\sum_{i=1}^{n-s} c_j^2 \le 16\Delta$. Thus, the sum of all the c_j^2 is at most 21Δ. Applying Azuma's Inequality with $t = \frac{\Delta}{r^2} - 2$ yields $\mathbf{Pr}(A_i) < \Delta^{-6}$, as desired. □

11.4 Constructing the Decomposition

In this section we prove our two lemmas on the local structure surrounding dense vertices, i.e. Lemmas 11.5 and 11.6. The proofs of these lemmas are not probabilistic. We include them for completeness. In these proofs, we repeatedly apply the refinement of the greedy colouring procedure discussed above. That is, we repeatedly find some partial $\Delta - 1$ colouring and complete it to a $\Delta - 1$ colouring by greedily colouring the uncoloured vertices in an appropriate order. Crucial to the proofs is the following:

Observation 11.12 *If H is a subgraph of G with at most $\Delta + 1$ vertices such that every vertex of H has at least $\frac{9\Delta}{10}$ neighbours in H then H is either a clique or is a clique and a vertex.*

Proof Let $X_1, \ldots, X_l$ be a maximum size family of disjoint stable sets of size two in H. If $l = 0$ then H is a clique and we are done, so we assume l is at least 1. Clearly $l \le \frac{\Delta+1}{2}$. We let $S = \cup_{i=1}^{l} X_i$. We can $\Delta - 1$ colour $G - H$ since it is a proper subgraph of G. We claim that if H is not a clique and a vertex then we can extend any such colouring to a $\Delta - 1$ colouring of G.

We will first extend our colouring to the vertices of S so that for each i, the vertices of X_i get the same colour. To do so, we colour the two vertices of X_i

at the same time. Between them, these vertices have at most $\frac{2\Delta}{10}$ neighbours outside of H. Our colouring procedure ensures that there are at this point at most l colours used on H. Thus, there are at least $\frac{3\Delta}{10} - 1$ colours which can be assigned to both vertices of X_i. So we can indeed extend the colouring to S so that each X_i is monochromatic.

Case 1: $l \geq \frac{\Delta}{10} + 2$.

By our degree condition on H, each vertex of $H - S$ misses at most $\frac{\Delta}{10}$ vertices of S and hence has two repeated colours in its neighbourhood. So we can complete our $\Delta - 1$ colouring greedily.

Case 2: $l < \frac{\Delta}{10} + 2$.

Note that $C = H - S$ is a clique with at least $\frac{7\Delta}{10} - 3$ vertices. By our degree and size conditions on H, there are at most $\frac{2\Delta}{10}$ vertices of C which miss a vertex in X_1. Thus, we can find vertices u and v of C both of which see both vertices of X_1. In fact, a similar argument allows us to insist that if $l \geq 2$ then u and v both see all of $X_1 \cup X_2$.

Now, we claim that if $l \geq 2$ then we can complete our colouring greedily provided we colour u and v last. When we colour a vertex of $C - u - v$ it has two uncoloured vertices (u and v). Both u and v have two repeated colours in their neighbourhoods. The claim follows.

Finally, if $l = 1$ then we let $X_1 = \{x, y\}$. Since H is not a clique and a vertex, both x and y miss a vertex of C. Since there are no two disjoint stable sets of size two in H, they both miss some vertex z of C and see all of $C - z$. In this case, we insist that when extending our colouring of $G - H$ to $G - (H - X_1)$ we actually extend it to a colouring of $G - (H - X_1 - z)$ by using one of the $\frac{7\Delta}{10} - 1$ colours which do not appear on $N(x) \cup N(y) \cup N(z) - H$ to colour $\{x, y, z\}$. Now we can complete the colouring greedily, as all the vertices of $C - z$ see the three vertices x, y, z on which we have used one colour. □

Proof of Lemma 11.5. Consider a dense vertex v. Define $S(v) = v + \{w | w \in N(v),\ |N(w) \cap N(v)| \geq \frac{9\Delta}{10} + r\}$. Since $|N(v)| \geq \Delta - 1$ and $|E(\overline{N(v)})| \leq \Delta(\log \Delta)^3$, it follows that $|N(v) - S(v)| < r$. This implies that $\Delta + 1 > |S(v)| > \Delta - r + 1$, and that every vertex of S_v has more than $\frac{9\Delta}{10}$ neighbours in S_v. By Observation 11.12, either S_v is a clique C_v or there is a vertex w of S_v such that $S_v - w$ is a clique C_v. In either case, C_v is the desired big clique containing v. □

A bit more work is required to prove Lemma 11.6.

Proof of Lemma 11.6. We have just proven that every dense vertex is in a big clique. We now examine these cliques more closely.

11.13 *If two maximal big cliques C_1 and C_2 with $|C_1| \leq |C_2|$ intersect, then $|C_1 - C_2| \leq 1$.*

Proof By considering a vertex in the intersection of the two cliques, we see that their union contains at most $\Delta + 1$ vertices. Hence, we can apply Observation 11.12 to the graph obtained from their union. □

We then obtain:

11.14 *No maximal big clique C intersects two other maximal big cliques.*

Proof By (11.13), the union S of these three big cliques would contain at most $|C| + 2 \leq \Delta + 1$ vertices. Applying Lemma 11.12 to S yields that S is a clique or a clique and a vertex. This contradicts our assumption that S contains three maximal big cliques. □

Now, (11.13), (11.14), and Lemma 11.5 imply that we can partition G up into sets $E_1, \ldots, E_l$ and T such that T is the set of vertices in no big clique and hence contains no dense vertices, and each E_i is either a maximal big clique C_i or consists of a maximal big clique C_i and a vertex v_i seeing at least $\Delta - r - 1$ but not all of the vertices of C_i.

To complete the proof of Lemma 11.6, we need the following result.

Observation 11.15 *For each vertex v of C_i, there is at most one neighbour of v in $G - C_i$ which sees more than $r + 4$ vertices of C_i. Furthermore, if $|C_i| = \Delta - 1$ then there is no such vertex.*

Proof For $|C_i| < \Delta - 1$, this proof is similar to the case $l = 2$ of Observation 11.12 and is left as an exercise. For $|C_i| = \Delta - 1$, the proof has the same flavour but is slightly more complicated. The reader may work through the details by solving Exercises 11.3 -11.5 □

Corollary 11.16 *For each C_i, there is at most one vertex in $G - C_i$ which sees at least $\frac{3\Delta}{4}$ vertices of C_i.*

This corollary ensures that we can obtain $D_1, \ldots, D_l$ and S satisfying (i),(ii), and (iii) of Lemma 11.6 by

1. Setting $D_i = E_i$ if E_i is not a clique,
2. Setting $D_i = E_i$ if E_i is C_i but no vertex v of $G - C_i$ satisfies $|N(v) \cap C_i| \geq \frac{3\Delta}{4}$,
3. Setting $D_i = E_i + v_i$ for the unique vertex v_i of $G - C_i$ satisfying $|N(v_i) \cap C_i| \geq \frac{3\Delta}{4}$ otherwise.

Now, Observation 11.15 implies (iv) of Lemma 11.6 holds as well, and the proof is complete. □

Exercises

Exercise 11.1 Consider the experiment in which we toss a fair coin once and then one of two biased coins $n-1$ times. If the ith flip came up heads then the coin we use for the $(i+1)$st flip will yield heads with probability $\frac{2}{3}$. If the ith flip came up tails then the coin we use for the $(i+1)$st flip will yield tails with probability $\frac{2}{3}$. Let X be the total number of flips which come up heads. Prove that each coin flip changes the conditional expected value of X by at most 3. Use Azuma's Inequality to prove that X is concentrated around its expected value. Can you apply Talagrand's Inequality to obtain this result?

Exercise 11.2 Show that Azuma's Inequality implies the Simple Concentration Bound.

In the following exercises, C is a clique of size $\Delta - 1$ in our minimal counterexample G, z is a vertex outside of C which sees at least $r+5$ vertices of C, and v is a neighbour of z in C.

Exercise 11.3 Mimic the proof of the case $l \leq 2$ of Observation 11.12 to show

(a) v has degree Δ, and
(b) the other external neighbour y of v has at most two other neighbours in C.

Exercise 11.4 Assume every neighbour of z in C has degree Δ and no vertex outside $C+z$ has more than three neighbours in $C \cap N(z)$. Let y be the other neighbour of v outside C. Show

(a) there is a vertex w in $N(z) \cap C - N(y)$ such that adding an edge between y and the neighbour of w in $V - C - z$ does not create a clique of size Δ,
(b) for any such w there is a colouring of $G - (C - w) - z$ in which y and w receive the same colour, and
(c) for any such w and vertex x in $C - N(z)$ there is a colouring of $G - (C - w - x)$ in which y and w receive the same colour and z and x receive the same colour.

Exercise 11.5 Combine the last two results to show that z does not exist.