Combinatorics, Probability and Computing (2018) 27, 21–43.
doi:10.1017/S0963548317000244

A Stronger Bound for the Strong Chromatic Index

HENNING BRUHN and FELIX JOOS

Institut für Optimierung und Operations Research, Universität Ulm, Ulm, Germany
(e-mail: henning.bruhn@uni-ulm.de, felix.joos@uni-ulm.de)

Received 30 October 2015; revised 5 June 2017; first published online 19 July 2017

We prove $\chi'_s(G) \leqslant 1.93\Delta(G)^2$ for graphs of sufficiently large maximum degree where $\chi'_s(G)$ is the strong chromatic index of G. This improves an old bound of Molloy and Reed. As a by-product, we present a Talagrand-type inequality where we are allowed to exclude unlikely bad outcomes that would otherwise render the inequality unusable.

2010 *Mathematics subject classification*: 05C15

1. Introduction

Edge colourings are well understood, but for strong edge colourings this is much less the case. An edge colouring can be viewed as a partition of the edge set of a graph G into matchings; the smallest such number of partition classes is the *chromatic index* of G. If we consider the natural stronger notion of a partition into *induced* (or *strong*) matchings, we arrive at the *strong chromatic index* $\chi'_s(G)$ of G, the minimal number of induced matchings needed.

The classic result of Vizing, and independently Gupta, constrains the chromatic index of a (simple) graph G to a narrow range: it is either equal to the trivial lower bound of the maximum degree $\Delta(G)$, or one more than that. The strong chromatic index, in contrast, can vary much more. The trivial lower bound and a straightforward greedy argument give a range of $\Delta(G) \leqslant \chi'_s(G) \leqslant 2\Delta(G)^2 - 2\Delta(G) + 1$ for all graphs G. Erdős and Nešetřil [8] conjectured a much stricter upper bound.

Strong edge colouring conjecture. For all graphs G, we have $\chi'_s(G) \leqslant \frac{5}{4}\Delta(G)^2$.

If true, the conjecture would be optimal, because any blow-up of the 5-cycle as in Figure 1 attains equality. For odd maximum degree, Erdős and Nešetřil conjectured that

$$\chi'_s(G) \leqslant \frac{5}{4}\Delta(G)^2 - \frac{1}{2}\Delta(G) + 1,$$

which again would be tight.

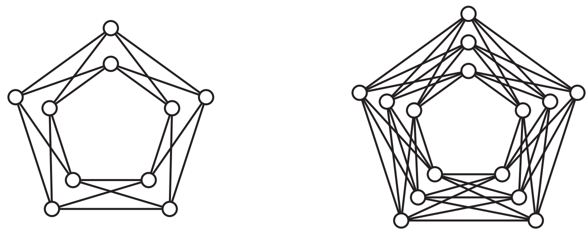

Figure 1. Two blow-ups of the 5-cycle.

In a breakthrough article of 1997, Molloy and Reed [22] demonstrated how probabilistic colouring methods could be used to beat the trivial greedy bound

$$\chi_s'(G) \leqslant 1.998\Delta(G)^2$$

for graphs G with $\Delta(G)$ sufficiently large.

We improve this bound as follows.

Theorem 1.1. *If G is a graph of sufficiently large maximum degree Δ, then*

$$\chi_s'(G) \leqslant 1.93\Delta^2.$$

A strong edge colouring of a graph G may be viewed as an ordinary vertex colouring in the square $L^2(G)$ of the line graph of G. (The *square* of any graph is obtained by adding edges between any vertices of distance of most 2.) Working in $L^2(G)$ permitted Molloy and Reed to split the strong edge colouring problem into two weaker subproblems. First, they showed that the neighbourhood of any vertex in $L^2(G)$ is somewhat sparse. Second, based on a probabilistic colouring method, they proved a colouring result for graphs with sparse neighbourhoods that holds for general graphs, not only for squares of line graphs.

We follow these same steps but make a marked improvement in each subproblem: our sparsity result is asymptotically best possible; and our colouring lemma needs fewer colours. We discuss the differences between our approach and that of Molloy and Reed in detail in Sections 2 and 5.

As a tool for our colouring lemma, we develop in Section 7 a version of Talagrand's inequality that excludes exceptional outcomes. Talagrand's inequality is used to verify that random variables on product spaces are tightly concentrated around their expected value. It is particularly suited for random variables that change only a little when a single coordinate is modified. This will not be the case in our application: in some very rare events a single change might result in a very large effect. To cope with this, we formulate a version of Talagrand's inequality in which such large effects of tiny probability can be ignored. We take some effort to make the application of the inequality as simple as possible, as we have some hopes that it might be useful elsewhere.

A weakening of the strong edge colouring conjecture yields a statement on *strong* cliques, the cliques of the square of the line graph.

Conjecture 1.2. *Any strong clique in a graph G has size at most $\omega(L^2(G)) \leqslant \frac{5}{4}\Delta^2(G)$.*

Note that the edge set of any blow-up of the 5-cycle is a strong clique, so the conjecture would be tight. Even this seemingly easier conjecture is still open. Chung, Gyárfás, Tuza and

Trotter [5] showed that any graph G that is $2K_2$-free has at most $\frac{5}{4}\Delta^2(G)$ edges. In such a graph, the whole edge set forms a strong clique. Faudree, Schelp, Gyárfás and Tuza [9] found an upper bound of $(2-\varepsilon)\Delta^2(G)$ for the size of any strong clique, for some small ε. Bipartite graphs are easier to handle in this respect: the same authors proved that a strong clique can never have size larger than $\Delta(G)^2$. Again, this is tight, as balanced complete bipartite graphs attain that bound. A consequence of our approach is the following bound.

Theorem 1.3. *If G is a graph with maximum degree $\Delta \geqslant 400$, then any strong clique in G has size at most $\omega(L^2(G)) \leqslant 1.74\Delta^2$.*

After our result was published on arXiv, Śleszyńska-Nowak [26] found a neat elementary proof that the size of a strong clique is always at most $\frac{3}{2}\Delta(G)^2$ for any graph G.

Coming back to strong edge colourings, let us note that there are a number of results for special graph classes. The conjecture was verified for maximum degree 3 by Andersen [2], and independently by Horek, Qing and Trotter [11]. For $\Delta(G) = 4$, Cranston [6] achieves a bound of $\chi'_s(G) \leqslant 22$, which is off by 2. Mahdian [19] proved that

$$\chi'_s(G) \leqslant \frac{2\Delta(G)^2}{\log \Delta(G)}$$

if G does not contain 4-cycles by quite involved probabilistic methods. Finally, a number of works concern d-degenerate graphs, the earliest of which is by Faudree, Schelp, Gyárfás and Tuza [9], who established the bound $\chi'_s(G) \leqslant 4\Delta(G)+4$ for planar graphs G. Kaiser and Kang [15] consider a generalization of strong edge colourings, where alike coloured edges have to be even farther apart.

The use of probabilistic methods to colour graphs is explored in depth in the book by Molloy and Reed [24], where many more references can be found. We only mention additionally the article by Alon, Krivelevich and Sudakov [1] on colouring graphs with sparse neighbourhoods. However, their result has non-trivial implications only if the neighbourhoods are much sparser than we can expect in squares of line graphs.

Strong edge colourings seem much more difficult than edge colourings. This is because induced matchings are much harder to handle than ordinary matchings. While the size of a largest matching can be determined quite precisely, it is hard even to obtain good bounds for induced matchings; see for instance [12, 13, 14].

In 2016, after a draft of this article became publicly available, Bonamy, Perrett and Postle announced a further and substantial improvement of the bound on the strong chromatic number (see *e.g.* [4]). The precise obtained bound is unclear to us since, so far, no preprint or draft seems to have been made available in the usual repositories.

All our graphs are simple and finite. We use standard graph-theoretic notation and concepts that can be found, for instance, in the book by Diestel [7].

2. Outline and proof of Theorem 1.1

A strong edge colouring of a graph G is simply an ordinary vertex colouring in $L^2(G)$, the square of the line graph of G. Molloy and Reed [22] use this simple observation to prove their bound on the strong chromatic index in two steps.

First, they show that neighbourhoods in $L^2(G)$ cannot be too dense. To formulate this more precisely, for any edge e of G, let N_e^{s} denote the set of edges of distance at most 1 to e, which is equivalent to saying that N_e^{s} is the neighbourhod of e in $L^2(G)$. We will often call N_e^{s} the *strong* neighbourhood of e. Molloy and Reed show that, for every edge e,

$$N_e^{\mathrm{s}} \text{ induces in } L^2(G) \text{ at most } \left(1-\frac{1}{36}\right)\binom{2\Delta^2}{2} \text{edges,} \tag{2.1}$$

where Δ is the maximum degree of G.

In the second step, Molloy and Reed show that any graph with sparse neighbourhoods, such as $L^2(G)$, can by coloured with a probabilistic procedure.

Following this strategy, we also prove a sparsity result and a colouring lemma.

Lemma 2.1. *Let G be a graph of maximum degree $\Delta \geqslant 1$, and let e be an edge of G. Then the neighbourhood N_e^{s} of e induces in $L^2(G)$ a graph with at most $\frac{3}{2}\Delta^4 + 5\Delta^3$ edges.*

Lemma 2.2. *Let $\gamma, \delta \in (0,1)$ be so that*

$$\gamma < \frac{\delta}{2(1-\gamma)} e^{-1/(1-\gamma)} - \frac{\delta^{3/2}}{6(1-\gamma)^2} e^{-7/(8(1-\gamma))}. \tag{2.2}$$

Then there is an integer R with the following property. If G is any graph of maximum degree at most $r \geqslant R$ in which, for every vertex v, the neighbourhood $N(v)$ induces a subgraph with at most $(1-\delta)\binom{r}{2}$ edges, then $\chi(G) \leqslant (1-\gamma)r$.

Condition (2.2) might be slightly hard to parse. Therefore, let us remark that, for $\delta \in (0, 0.9]$,

$$\gamma = 0.1827 \cdot \delta - 0.0778 \cdot \delta^{3/2}$$

satisfies the condition and is not too far away from the best possible γ.

Our main theorem is a direct consequence of the two lemmas.

Proof of Theorem 1.1. Let G be a graph with maximum degree Δ sufficiently large. By Lemma 2.1, the neighbourhood of any vertex v of $L^2(G)$ induces a subgraph with at most $(3/4 + o(1))\binom{2\Delta^2}{2}$ edges. Therefore, we can apply Lemma 2.2 with $r = 2\Delta^2$, $\delta = 0.24$, and $\gamma = 0.035$. □

While Molloy and Reed developed this very neat proof technique, our contribution consists in improving its two constituent steps. In particular, in the sparsity lemma we improve Molloy and Reed's $1/36$ in (2.1) to roughly $1/4$. This is almost best possible: in Section 4 we construct graphs that asymptotically reach the upper density bound of Lemma 2.1. Our colouring lemma also yields a γ that is somewhat smaller than the corresponding γ of Molloy and Reed. We discuss the differences between their colouring lemma and ours in more detail in Section 5.

Let us have a look at some concrete numbers. There is a small oversight in the proof of Molloy and Reed (a lost 2) that results in the actual bound of $\chi_s'(G) \leqslant 1.9987\Delta(G)^2$ instead of the claimed $\chi_s'(G) \leqslant 1.998\Delta(G)^2$ (from now on always provided $\Delta(G)$ is sufficiently large).

What improvements follow from our lemmas? With our sparsity lemma and the colouring lemma of Molloy and Reed it is possible to obtain $\chi_s'(G) \leqslant 1.99\Delta(G)^2$. Our colouring lemma then leads to the factor 1.93 in our main theorem. Viewing all statements as bounds of the form $\chi_s'(G) \leqslant (2-\varepsilon)\Delta(G)^2$, we improve ε by a factor of 53.

Finally, even assuming that the colouring lemma could be vastly improved (doubtful), this method can never go beyond $1.73\Delta(G)^2$. The sparsity bound in Lemma 2.1, which, again, is asymptotically best possible, does not exclude a clique of size $1.73\Delta(G)^2$ in the neighbourhood N_e^{s} in $L^2(G)$. This means that just considering edge densities will never yield a factor smaller than 1.73 (and probably not even close to that number).

For the reader's convenience, we summarize the structure of the paper. In Section 3 we prove Lemma 2.1, and in Section 4 we provide a family of graphs which verify that the bound in Lemma 2.1 is best possible. In Section 5 we give a proof of Lemma 2.2 whereby we rely on Lemma 5.1, which states that certain parameters in the colouring procedure are strongly concentrated around their mean. The proof of Lemma 5.1 is presented in Section 8, because we first need to introduce and prove our version of Talagrand's inequality (Section 7). In Section 6 we discuss how one may improve the bound in Theorem 1.1. We close in Section 9 with some discussion about further applications of our version of Talagrand's inequality.

3. Density of the strong neighbourhood

In this section we prove the sparsity lemma, Lemma 2.1. For any edge e in a graph G, we write $d^{\mathsf{s}}(e)$ for the degree in $L^2(G)$. Then $d^{\mathsf{s}}(e) = |N_e^{\mathsf{s}}|$.

Lemma 3.1. *Let G be a graph of maximum degree Δ, and let e be an edge of G. Then*

$$d^{\mathsf{s}}(e) \leqslant (2-\alpha-\beta)\Delta^2 - 2\Delta,$$

where $\alpha\Delta$ is the number of triangles in G containing e, and where $\beta\Delta^2$ is the number of 4-cycles containing e plus the number of triangles incident with exactly one end-vertex of e.

Proof. Let $e = uv$. Then

$$|N(u) \cup N(v) \setminus \{u,v\}| \leqslant 2(\Delta-1) - |N(u) \cap N(v)| = (2-\alpha)\Delta - 2.$$

Every edge in N_e^{s} has at least one of its end-vertices in $N(u) \cup N(v)$. Edges with both end-vertices in $N(u) \cup N(v)$ lie in a common 4-cycle with e or form a triangle with either u or v, and thus count towards $\beta\Delta^2$. In total we obtain

$$d^{\mathsf{s}}(e) \leqslant |N(u) \cup N(v) \setminus \{u,v\}| \cdot \Delta - \beta\Delta^2 \leqslant ((2-\alpha)\Delta - 2)\Delta - \beta\Delta^2,$$

which gives the claimed bound. □

We remark that the inequality given in Lemma 3.1 is even an equality if G is Δ-regular. We will frequently use the observation that the containment of an edge f in k 4-cycles implies $d^{\mathsf{s}}(f) \leqslant 2\Delta^2 - k$.

We will occasionally need to measure how many neighbours a vertex has in some subset of the vertex set. So, for a vertex v in a graph G and some set $X \subseteq V(G)$, we write $d_X(v)$ for $|N_G(v) \cap X|$. We will similarly use the notation $d_F^{\mathsf{s}}(e)$ for the degree of an edge e in $L^2(G)$ restricted to some edge set F.

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

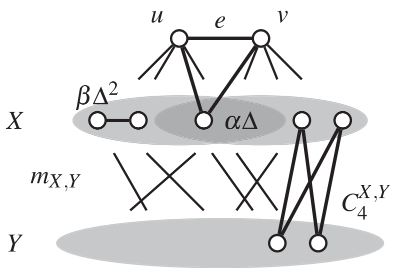

Figure 2. Some of the parameters used in the proof of Lemma 2.1.

Proof of Lemma 2.1. Without loss of generality, we may assume G to be Δ-regular. Indeed, if G is not then it may be embedded into a Δ-regular graph, in which the number of edges induced by N_e^{s} will only be larger. (The embedding is a standard technique: iteratively make a second copy of G and connect copies of vertices of too low degree by an edge.)

Let $e = uv$ and set $X = N_G(u) \cup N_G(v) \setminus \{u,v\}$, while $Y = N_G(X) \setminus (X \cup \{u,v\})$. Then, using the notation of Lemma 3.1, we set $\alpha\Delta = |N_G(u) \cap N_G(v)|$ to be the number of triangles containing e, and let $\beta\Delta^2$ be the number of edges with both end-vertices in X. Observe that $|X| = (2-\alpha)\Delta - 2$. We furthermore note that $0 \leqslant \alpha, \beta \leqslant 1$; for β this follows from $2\beta\Delta^2 \leqslant |X| \cdot \Delta$. Let m_e denote the number of edges induced by N_e^{s} in $L^2(G)$. Our objective is to bound m_e from above.

Let us start with the estimation of m_e. Writing $\overline{N_e^{\mathsf{s}}}$ for $E(G) \setminus N_e^{\mathsf{s}}$, we obtain

$$2m_e = \sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{N_e^{\mathsf{s}}}(f) = \sum_{f \in N_e^{\mathsf{s}}} (d^{\mathsf{s}}(f) - d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f))$$

To get an upper bound on $\sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}(f)$ we use Lemma 3.1, where we omit the triangles and only count those 4-cycles that consist entirely of edges between X and Y. Let $C_4^{X,Y}$ be the number of those 4-cycles, that is, the number of those 4-cycles that consist only of edges with one end-vertex in X and the other in Y. Observe that each cycle counted by $C_4^{X,Y}$ reduces the degree of four edges in N_e^{s}. Thus,

$$\sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}(f) \leqslant \left(\sum_{f \in N_e^{\mathsf{s}}} 2\Delta^2 \right) - 4C_4^{X,Y},$$

by Lemma 3.1. Using the lemma again, this time to bound $|N_e^{\mathsf{s}}|$ from above, and substituting in our above estimation of m_e, we get

$$\begin{aligned} 2m_e &\leqslant 2\Delta^2 |N_e^{\mathsf{s}}| - 4C_4^{X,Y} - \sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f) \\ &\leqslant 2\Delta^2((2-\alpha-\beta)\Delta^2 - 2\Delta) - 4C_4^{X,Y} - \sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f) \\ &= 2(2-\alpha-\beta)\Delta^4 - 4\Delta^3 - 4C_4^{X,Y} - \sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f) \end{aligned}$$

In order to obtain a lower bound on $\sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f)$, we consider paths of the form $pxyq$ in G, where $x \in X$, $y \in Y$ and $q \notin X$. The first edge px then is in N_e^{s}, while the last edge yq is outside. So, each such path $pxyq$ contributes 1 to $\sum_{f \in N_e^{\mathsf{s}}} d^{\mathsf{s}}_{\overline{N_e^{\mathsf{s}}}}(f)$. Since G is Δ-regular, there are $\Delta - 1$ such paths for each fixed xyq. Counting the number of such xyq, we arrive at $\sum_{y \in Y} d_X(y)(\Delta - d_X(y))$.

For later use, we give this parameter a name,

$$\gamma\Delta^3 := \sum_{y\in Y} d_X(y)(\Delta - d_X(y)),$$

and observe that $0 \leqslant \gamma \leqslant 1/2$. Indeed, for any $y \in Y$, we have

$$d_X(y)(\Delta - d_X(y)) \leqslant d_X(y)^2/4 \leqslant \frac{\Delta}{4} d_X(y)$$

and consequently

$$\gamma\Delta^3 \leqslant \sum_{y\in Y} \frac{\Delta}{4}\cdot d_X(y) \leqslant \frac{\Delta}{4}\cdot 2\Delta^2,$$

as there can be at most $2\Delta^2$ edges between X and Y.

Summing up our discussion,

$$\sum_{f\in N_e^{\mathsf{s}}} d_{N_e^{\mathsf{s}}}^{\mathsf{s}}(f) \geqslant (\Delta-1)\gamma\Delta^3,$$

which leads to

$$m_e \leqslant \left(2-\alpha-\beta-\frac{\gamma}{2}\right)\Delta^4 - 2C_4^{X,Y} + \left(\frac{\gamma}{2}-2\right)\Delta^3. \tag{3.1}$$

It remains to estimate $C_4^{X,Y}$. For this, let us first compute $m_{X,Y}$, the number of edges with one end-vertex in X and the other in Y. On the one hand, we have

$$\sum_{x\in X} d_G(x) = m_{X,Y} + 2\beta\Delta^2 + |X|,$$

while Δ-regularity, on the other hand, gives us $\sum_{x\in X} d_G(x) = \Delta\cdot|X|$. Together with $|X| = (2-\alpha)\Delta - 2$, this implies

$$m_{X,Y} = (2-\alpha-2\beta)\Delta^2 - (4-\alpha)\Delta + 2. \tag{3.2}$$

For $x_1, x_2 \in X$, define $c(x_1,x_2)$ to be the number of common neighbours of x_1 and x_2 in Y, where we put $c(x_1,x_2) = 0$ if $x_1 = x_2$. Then

$$C_4^{X,Y} = \sum_{x_1,x_2\in X} \binom{c(x_1,x_2)}{2}.$$

We compute

$$\sum_{x_1,x_2\in X} c(x_1,x_2) = \sum_{y\in Y}\binom{d_X(y)}{2} = \frac{1}{2}\sum_{y\in Y}(d_X(y))^2 - \frac{1}{2}m_{X,Y}.$$

Using the definition of γ, this gives

$$\begin{aligned}\sum_{x_1,x_2\in X} c(x_1,x_2) &= \frac{1}{2}\left(\Delta\sum_{y\in Y} d_X(y) - \gamma\Delta^3\right) - \frac{1}{2}m_{X,Y}\\ &= \frac{1}{2}((\Delta-1)m_{X,Y} - \gamma\Delta^3).\end{aligned}$$

Using $m_{X,Y} \leqslant 2\Delta^2$ as well as (3.2), we obtain a lower and an upper bound:

$$\frac{1}{2}((2-\alpha-2\beta-\gamma)\Delta^3-6\Delta^2) \leqslant \sum_{x_1,x_2\in X} c(x_1,x_2) \leqslant \left(1-\frac{\gamma}{2}\right)\Delta^3. \tag{3.3}$$

We come back to the calculation of $C_4^{X,Y}$:

$$\begin{aligned} C_4^{X,Y} &= \sum_{x_1,x_2\in X} \binom{c(x_1,x_2)}{2} = \frac{1}{2}\sum_{x_1,x_2\in X}(c(x_1,x_2)^2 - c(x_1,x_2)) \\ &\geqslant \frac{1}{2}\sum_{x_1,x_2\in X} c(x_1,x_2)^2 - \frac{1}{2}\left(1-\frac{\gamma}{2}\right)\Delta^3, \end{aligned} \tag{3.4}$$

where we used (3.3). We use the Cauchy–Schwarz inequality,

$$\begin{aligned} \sum_{x_1,x_2\in X} c(x_1,x_2)^2 &\geqslant \binom{|X|}{2}\left(\frac{\sum_{x_1,x_2\in X} c(x_1,x_2)}{\binom{|X|}{2}}\right)^2 \\ &\geqslant 2\cdot\frac{(\sum_{x_1,x_2\in X} c(x_1,x_2))^2}{|X|^2} \\ &\geqslant 2\cdot\frac{(\frac{1}{2}((2-\alpha-2\beta-\gamma)\Delta^3-6\Delta^2))^2}{(2-\alpha)^2\Delta^2}, \end{aligned}$$

where the last inequality is because of $|X| = (2-\alpha)\Delta - 2$ and (3.3). We continue:

$$\begin{aligned} \sum_{x_1,x_2\in X} c(x_1,x_2)^2 &\geqslant \frac{1}{2}\left(\frac{(2-\alpha-2\beta-\gamma)^2\Delta^4}{(2-\alpha)^2} - 12\cdot\frac{(2-\alpha)\Delta^3}{(2-\alpha)^2}\right) \\ &\geqslant \frac{(2-\alpha-2\beta-\gamma)^2}{2(2-\alpha)^2}\Delta^4 - 6\Delta^3. \end{aligned}$$

We substitute this in our estimation (3.4) of $C_4^{X,Y}$:

$$\begin{aligned} C_4^{X,Y} &\geqslant \frac{1}{2}\left(\frac{(2-\alpha-2\beta-\gamma)^2\Delta^4}{2(2-\alpha)^2} - 6\Delta^3\right) - \frac{1}{2}\left(1-\frac{\gamma}{2}\right)\Delta^3 \\ &= \frac{1}{2}\left(\frac{(2-\alpha-2\beta-\gamma)^2\Delta^4}{2(2-\alpha)^2} - \left(7-\frac{\gamma}{2}\right)\Delta^3\right). \end{aligned}$$

We can finally complete our estimation (3.1) of m_e, obtaining

$$m_e \leqslant \left(2-\alpha-\beta-\frac{\gamma}{2}\right)\Delta^4 - \frac{(2-\alpha-2\beta-\gamma)^2}{2(2-\alpha)^2}\Delta^4 + 5\Delta^3.$$

To see that this gives

$$m_e \leqslant \frac{3}{2}\Delta^4 + 5\Delta^3,$$

observe first that we may assume that $0 \leqslant \alpha \leqslant 1/2$ or we are already done. Setting $x = \beta + \gamma/2$, we may also assume $x \leqslant 1/2$. Let

$$f(\alpha,x) = 2-\alpha-x-\frac{(2-\alpha-2x)^2}{2(2-\alpha)^2}.$$

Note that

$$\frac{\partial f(\alpha,x)}{\partial \alpha} = -1 + \frac{2x(2-\alpha-2x)}{(2-\alpha)^3} < 0$$

for $0 \leqslant x \leqslant 1/2$. Thus we may assume that $\alpha = 0$. Consequently, it remains to verify that

$$2 - x - \frac{(2-2x)^2}{8} \leqslant \frac{3}{2},$$

which is an elementary task. □

We note that a good number of elements used in the proof appear already in the article of Molloy and Reed [22]: triangles through e, edges in X, paths xyq and 4-cycles between X and Y. Our contribution consists in parametrizing these elements and then combining the parameters in the right way to give a nearly tight bound.

Lemma 2.1 naturally gives an upper bound on the size of the largest clique in $L^2(G)$. Although cliques have a very simple structure, we have not been able to push the bound given in Theorem 1.3 significantly further.

Proof of Theorem 1.3. Let K be a largest strong clique, that is, a largest clique in $L^2(G)$, and let $\kappa\Delta^2$ be its size. If $e \in K$ is an edge of G, then its neighbourhood induces by Lemma 2.1 a graph of at most $\frac{3}{2}\Delta^4 + 5\Delta^3$ edges. Thus

$$\binom{\kappa\Delta^2 - 1}{2} \leqslant \frac{3}{2}\Delta^4 + 5\Delta^3,$$

which implies

$$\kappa \leqslant \frac{3}{2\Delta^2} + \sqrt{3 + \frac{10}{\Delta} + \frac{9}{4\Delta^4}} < 1.74$$

for $\Delta \geqslant 400$. □

4. The sparsity lemma is best possible

In this section we describe a family of graphs with strong neighbourhoods that almost reach the upper density bound of Lemma 2.1. To show this, we turn to *Hadamard codes*. For every $k \geqslant 2$, the corresponding Hadamard code consists of $2 \cdot 2^k$ 0,1-strings of length 2^k each with certain properties. For instance, for $k = 2$, the Hadamard code is

$$\{1111, 0000, 1100, 0011, 0110, 1001, 1010, 0101\}$$

Notably, the code always contains the all-0-string and the all-1-string. We drop these and interpret the remaining code words as subsets of some fixed ground set of $n = 2^k$ elements $\{x_1, \dots, x_n\}$. Then we can see the Hadamard code as a set $\mathcal{H}_k$ of subsets of $\{x_1, \dots, x_n\}$ with the following properties:

(i) $\mathcal{H}_k$ contains $2(n-1)$ sets, each of which has cardinality $n/2$
(ii) every element x_i lies in precisely $n-1$ sets in $\mathcal{H}_k$

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

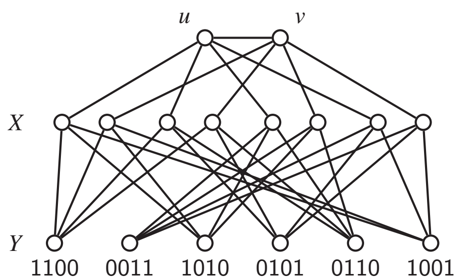

Figure 3. The graph G for $k=2$.

(iii) if $S, S' \in \mathcal{H}_k$ are distinct, then $|S \cap S'| \in \{0, n/4\}$.

For a proof and more details, see for example van Lint [28].

Based on $\mathcal{H}_k = \{S_1, \ldots, S_{2n-2}\}$, we construct a graph G. Make a copy x_i' for each x_i, and put $X = \{x_1, \ldots, x_n, x_1', \ldots, x_n'\}$. Let $Y = \{y_1, \ldots, y_{2n-2}\}$ be a set of $2n-2$ vertices that is disjoint from X. We define a graph G on $X \cup Y \cup \{u, v\}$, where u, v are two vertices outside $X \cup Y$. First, make u complete to $x_1, \ldots, x_n$ and v complete to $x_1', \ldots, x_n'$, and make u adjacent to v. Now, let us add a bipartite graph on $X \cup Y$. For this, we make each y_j adjacent to every vertex in S_j and to the set S_j' of their copies. See Figure 3 for an illustration. With this, each vertex in X has $(n-1)+1$ neighbours, by property (ii) of $\mathcal{H}_k$, and each y_j has degree $2 \cdot n/2 = n$, by property (i). Each of u and v has degree $n+1$. Thus the maximum degree is $\Delta = n+1$. We note, furthermore, that there are $(2n-2)n = 2\Delta^2 - O(\Delta)$ edges between X and Y.

We calculate the number of 4-cycles in $G-u-v$. Observe that for all $i<j$ the number $c(y_i, y_j)$ of common neighbours of y_i and y_j is at most $2 \cdot n/4$ (by property (iii)). Thus, the number of 4-cycles in $G-u-v$ is

$$\sum_{i<j} \binom{c(y_i, y_j)}{2} \leqslant \binom{2n-2}{2} \cdot \binom{n/2}{2} = \frac{n^4}{4} + O(n^3) = \frac{\Delta^4}{4} + O(\Delta^3).$$

Since there are only $2\Delta - 1$ edges incident with u or v, deleting all these edges will only result in a loss of $O(\Delta^3)$ edges in $L^2(G)$. In the square of the line graph of $G-u-v$, however, only the 4-cycles counted above reduce the degree $d^{\mathsf{s}}(e)$ away from the maximal possible value $2\Delta^2$ (recall the argumentation in the proof of Lemma 3.1). Thus, recalling that there are $2\Delta^2 - O(\Delta)$ edges between X and Y, we obtain that the number m_{uv} of edges induced by N_{uv}^{s} in $L^2(G)$ satisfies

$$2m_{uv} \geqslant 2\Delta^2 \cdot (2\Delta^2 - O(\Delta)) - 4\left(\frac{\Delta^4}{4} + O(\Delta^3)\right) - O(\Delta^3)$$

(Note that each 4-cycle reduces the degree of four edges in $L^2(G)$.) Therefore

$$m_{uv} \geqslant \frac{3}{2}\Delta^4 - O(\Delta^3),$$

which asymptotically coincides with the bound of Lemma 2.1.

5. The colouring procedure

Molloy and Reed [22] described a probabilistic method, called the *naive colouring procedure* in [24], to partially colour a sparse graph. In the procedure, every vertex first receives a colour

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

independently and uniformly at random. Then follows a conflict resolution step: whenever two adjacent vertices receives the same colour, both become uncoloured. The brilliant insight here is that, with a non-negligible probability, in every neighbourhood a good number of colours appear at least twice at the end of this procedure. In this way, the partial colouring *saves* colours in comparison to a $(\Delta(G)+1)$-colouring. This then makes it possible to complete the partial colouring greedily to a full colouring.

In the easiest manifestation of the procedure, whenever there is a conflict, that is, whenever two adjacent vertices receive the same colour, both vertices lose their colour. Molloy and Reed hint at a possible improvement that is less wasteful: if additionally random weights on the vertices are chosen, only the vertex of lower weight need to be uncoloured. However, the analysis becomes much more tedious. We propose a different conflict resolution mechanism that allows for a simpler analysis. In addition, we not only count colours that are used exactly twice in the neighbourhood of a fixed vertex, but also take into account multiple occurrences. Molloy and Reed mentioned such a modification, too, but did not discuss it in detail.

Why do we use naive colouring instead of a more sophisticated technique such as the *semi-random* colouring method? Introduced in a seminal article by Kim [16], the semi-random colouring method is an iterative procedure, in which in every round only a small fraction of the vertices are additionally coloured. For the method to work several parameters have to be tightly controlled so that every round is essentially the same.

We have tried, unsuccessfully, to employ semi-random colouring. Normally the method is applied to graphs in which all neighbourhoods are fairly similar (*e.g.* stable sets), such as triangle-free or otherwise sparse graphs.[1] We, on the other hand, have to cope with heterogeneous neighbourhoods and, consequently, found it very hard to control the necessary parameters. This is particularly the case for the number of non-edges between pairs of uncoloured vertices (within neighbourhoods), a parameter that is essential as such vertices may receive the same colour.

We now describe the modified Molloy and Reed colouring procedure. For this, assume G to be an r-regular graph, which we will colour with $C = \lceil (1-\gamma)r \rceil$ colours, where $\gamma \in (0,1)$ is a constant. The following random experiment is performed.

(1) Colour every vertex uniformly and independently at random from the set $\{1,\dots,C\}$.
(2) For every edge uv, choose an orientation d_{uv} independently and uniformly at random.
(3) If uv is an edge and u and v received the same colour, then uncolour u if d_{uv} points towards u and uncolour v otherwise.

The key parameter is the number of colours that are saved in the neighbourhood of an arbitrary vertex u, that is, the number of vertices that are assigned a colour that is already used by some other vertex in the neighbourhood. To estimate this quantity, we introduce the random variable P_u that counts the number of pairs of non-adjacent vertices in $N(u)$ that have the same colour after the uncolouring step. To control overcounting, we also define T_u, the number of triples of distinct non-adjacent vertices v,w,x in $N(u)$ that have the same colour at the end of the colouring procedure.

[1] While this is not the case in Molloy and Reed's [23] breakthrough result on the total chromatic number, the setting there is arguably quite different due to the nature of total colouring.

Most of the effort in the proof of Theorem 2.2 will be spent on estimating the expected values of P_u and T_u. Later, in Section 8, we will show that the random variables are strongly concentrated around their expectation.

Lemma 5.1. *Let $\gamma \in (0,1)$. There is an R such that, for every r-regular graph G with $r \geqslant R$, we have*

$$\mathbb{P}\big[|P_u - \mathbb{E}[P_u]| \geqslant \sqrt{r}\log^3 r\big] \leqslant r^{-\frac{1}{2}\log\log r} \quad \text{and} \quad \mathbb{P}\big[|T_u - \mathbb{E}[T_u]| \geqslant \sqrt{r}\log^4 r\big] \leqslant r^{-\frac{1}{2}\log\log r}$$

when the colouring procedure is performed with $C = \lceil (1-\gamma)r \rceil$ colours.

In order to prove the existence of a global partial colouring with certain nice properties, we use the Lovász Local Lemma to deduce this from our local structure.

Lemma 5.2 (Lovász Local Lemma [24]). *Let $p \in [0,1)$, and let $\mathcal{A}$ be a finite set of events such that, for every $A \in \mathcal{A}$,*

(i) *$\mathbb{P}[A] \leqslant p$, and*
(ii) *A is mutually independent of a set of all but at most d of the other events in $\mathcal{A}$.*

If $4pd \leqslant 1$, then the probability that none of the events in $\mathcal{A}$ occurs is strictly positive.

Proof of Lemma 2.2. How large R has to be will become clear at the end of the proof. Consider some graph G with maximum degree at most r satisfying the density condition on the neighbourhoods. We may assume that G is r-regular; otherwise we embed G, as in Lemma 2.1, in an r-regular graph while keeping the local density condition.

We will colour G with $C = \lceil (1-\gamma)r \rceil$ colours, which may be slightly more than the claimed bound of $\chi(G) \leqslant (1-\gamma)r$. However, by choosing R and thus r large enough we can find a γ' that still satisfies (2.2) and for which $\lceil (1-\gamma')r \rceil \leqslant (1-\gamma)r$ for every $r \geqslant R$. Thus, if necessary, we may replace γ by γ' in what follows.

Pick some vertex u of G. We start with the estimation of P_u, the number of non-adjacent neighbours of u with the same (final) colour. To this end, consider two non-adjacent neighbours v and w of u. Note first that the probability that v and w receive the same colour in step (1) is equal to $1/C$. Assuming that this is the case, we observe that in order for v and w to keep their colour in step (3), all the edges between $\{v,w\}$ and the neighbours of the same colour as v,w have to be chosen in step (2) so that they point away from $\{v,w\}$. Thus, if v and w have k common neighbours, the probability that they both keep their (common) colour is

$$\left(1 - \frac{1}{2C}\right)^{2r-2k}\left(1 - \frac{3}{4}\cdot\frac{1}{C}\right)^{k},$$

where the factor $3/4$ stems from the fact that, out of four orientations of the two edges between v,w and a common neighbour, only one of these allows both v,w to keep their colour.

In total, we get

$$\begin{aligned}&\mathbb{P}[v,w \text{ received the same colour and both stay coloured}]\\ &\quad = \frac{1}{C}\left(1 - \frac{1}{2C}\right)^{2r-2k}\left(1 - \frac{3}{4}\cdot\frac{1}{C}\right)^{k} = \frac{1}{C}\left(1 - \frac{1}{2C}\right)^{2r}\left(\frac{1-3/(4C)}{1-1/C+1/(4C^2)}\right)^{k}.\end{aligned}$$

Note that, for any $C \geqslant 1$, we have

$$\frac{1-3/(4C)}{1-1/C+1/(4C^2)} \geqslant 1.$$

Thus, if r is large enough, then

$$\begin{aligned}
&\mathbb{P}[v,w \text{ received the same colour and both stay coloured}] \\
&\quad \geqslant \frac{1}{C}\left(1-\frac{1}{2C}\right)^{2r} = \frac{1}{\lceil(1-\gamma)r\rceil}\left(1-\frac{1}{2\lceil(1-\gamma)r\rceil}\right)^{2r} \\
&\quad \geqslant (1+o(1))\frac{1}{(1-\gamma)r}e^{-1/(1-\gamma)}
\end{aligned}$$

Let δ_u be such that there are exactly $\delta_u\binom{r}{2}$ pairs of non-adjacent vertices in $N(u)$ (observe that $\delta_u \geqslant \delta$). Thus the calculation above implies

$$\begin{aligned}
\mathbb{E}[P_u] &\geqslant (1+o(1))\delta_u\binom{r}{2}\frac{1}{(1-\gamma)r}e^{-1/(1-\gamma)} \\
&= (1+o(1))\frac{\delta_u r}{2(1-\gamma)}e^{-1/(1-\gamma)}.
\end{aligned}$$

We will need a similar estimate for triples of non-adjacent neighbours of u. So, assume v,w,x to be three vertices in $N(u)$ that are pairwise non-adjacent. For $1 \leqslant i \leqslant 3$, let k_i be the number of vertices with i neighbours in $\{v,w,x\}$. Since G is r-regular, $k_1+2k_2+3k_3 = 3r$.

The probability that all three of v,w,x receive the same colour is $1/C^2$. The probability that all three retain their colour is computed in a similar way as above, where it should be noted that there are now eight possibilities for the orientations of the three edges between v,w,x and a neighbour common to all of them.

Using the binomial series

$$(1+z)^\alpha = \sum_{k=0}^{\infty}\binom{\alpha}{k}z^k \quad \text{and} \quad \frac{1}{C^\alpha} = o(1/r) \quad \text{for } \alpha > 1,$$

we obtain

$$\begin{aligned}
&\mathbb{P}[v,w,x \text{ received the same colour and stay coloured}] \\
&\quad = \frac{1}{C^2}\left(1-\frac{1}{2C}\right)^{k_1}\left(1-\frac{3}{4C}\right)^{k_2}\left(1-\frac{7}{8C}\right)^{k_3} \\
&\quad = \frac{1}{C^2}\left(1-\frac{3}{2C}+o(1/r)\right)^{k_1/3}\left(1-\frac{9}{8C}+o(1/r)\right)^{2k_2/3}\left(1-\frac{7}{8C}\right)^{k_3} \\
&\quad \leqslant (1+o(1))\frac{1}{(1-\gamma)^2r^2}\left(1-\frac{7}{8(1-\gamma)r}\right)^{k_1/3+2k_2/3+k_3} \\
&\quad = (1+o(1))\frac{1}{(1-\gamma)^2r^2}e^{-7/(8(1-\gamma))}.
\end{aligned}$$

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

By using that every graph with $\delta_u\binom{r}{2}$ edges contains at most $(\delta_u^{3/2}r^3)/6$ distinct triangles (see, for instance, Rivin [25]), we obtain

$$\begin{aligned}\mathbb{E}[T_u] &\leqslant (1+o(1))\frac{\delta_u^{3/2}r^3}{6}\frac{1}{(1-\gamma)^2r^2}e^{-7/(8(1-\gamma))}\\ &= (1+o(1))\frac{\delta_u^{3/2}r}{6(1-\gamma)^2}e^{-7/(8(1-\gamma))}.\end{aligned}$$

By the inclusion–exclusion principle, we save at least $P_u - T_u$ colours in $N(u)$, that is, the number of coloured vertices minus the number of colours actually used in $N(u)$ is at least $P_u - T_u$. Now, if $P_u - T_u \geqslant \gamma r$, then the number of uncoloured vertices is smaller than the number of unused colours in $N(u)$. Thus, if we can show that, with positive probability, $P_u - T_u \geqslant \gamma r$ for all $u \in V(G)$, then we can colour all remaining uncoloured vertices greedily, which then concludes the proof.

Next, we want to point out that we may assume that $\delta_u = \delta$; that is,

$$\mathbb{E}[P_u - T_u] \geqslant (1+o(1))\frac{\delta r}{2(1-\gamma)}e^{-1/(1-\gamma)} - (1+o(1))\frac{\delta^{3/2}r}{6(1-\gamma)^2}e^{-7/(8(1-\gamma))} \tag{5.1}$$

for all $u \in V(G)$. Observe first that if $\delta, \gamma \in (0,1)$ and (2.2) holds, then in particular,

$$\gamma < \frac{1}{2(1-\gamma)}e^{-1/(1-\gamma)}$$

and hence $\gamma < 0.2$. To prove (5.1), we show that the function

$$f(x) = \frac{xr}{2(1-\gamma)}e^{-1/(1-\gamma)} - \frac{x^{3/2}r}{6(1-\gamma)^2}e^{-7/(8(1-\gamma))}$$

is monotonically increasing in the appropriate interval.

Put

$$a = \frac{r}{2(1-\gamma)}e^{-1/(1-\gamma)}.$$

Then

$$f(x) = ax - \frac{e^{1/(8(1-\gamma))}}{3(1-\gamma)}ax^{3/2}$$

and

$$f'(x) = a - \frac{3e^{1/(8(1-\gamma))}}{6(1-\gamma)}a\sqrt{x}.$$

Clearly, $f'(0) > 0$. For us it suffices to check when f' becomes 0:

$$f'(x) = 0 \Leftrightarrow \sqrt{x} = \frac{2(1-\gamma)}{e^{1/(8(1-\gamma))}} \Rightarrow x = 4(1-\gamma)^2e^{-1/(4(1-\gamma))}.$$

As long as $\gamma \leqslant 0.3$, this implies $x > 1$. Thus, we obtain (5.1) for $\delta \in (0,1)$.

To show that $P_u - T_u \geqslant \gamma r$ for all $u \in V(G)$ holds with positive probability, let A_u be the event that

$$P_u - T_u \leqslant \left(1 - \frac{1}{\log r}\right)\left(\frac{\delta}{2(1-\gamma)}e^{-1/(1-\gamma)} - \frac{\delta^{3/2}}{6(1-\gamma)^2}e^{-7/(8(1-\gamma))}\right)r.$$

Since both random variables P_u and T_u are highly concentrated (see Lemma 5.1), it follows for large r that

$$\mathbb{P}[A_u] = O(r^{-\frac{1}{2}\log\log r}).$$

Note that the event A_u is independent of any collection of events $\{A_{u'}\}_{u'\in Z}$ if $Z \subseteq V(G)$ contains only vertices u' in distance at least, say, 10 from u. We deduce, therefore, from the Lovász Local Lemma that there is a colouring of the vertices of G such that no A_u holds (choose d to be r^{10} in the statement of the Lovász Local Lemma). This, however, implies together with (2.2) that $P_u - T_u \geqslant \gamma r$, provided that r is large enough. □

6. How to possibly save more colours

The factor of 1.93 of Theorem 1.1 is still very far from the conjectured factor of 1.25. While it seems doubtful that probabilistic colouring can ever get very close to 1.25, there is still some hope that with more sophisticated arguments the factor can be improved. For us, this hope is fuelled by two observations.

First, the two steps, the sparsity lemma and the colouring lemma, are completely dissociated. That is, the colouring lemma only exploits the sparsity of strong neighbourhoods but uses no structural information whatsoever. Surely it should help to recall that the task consists in colouring edges!

Second, while the sparsity lemma, Lemma 2.1, is asymptotically tight, the conjectured extreme example for the strong colouring conjecture, the blow-up of the 5-cycle, has much sparser strong neighbourhoods.

To be more concrete, consider the blow-up in which every vertex of the 5-cycle is replaced by a stable set of size k, so that the maximum degree becomes $\Delta = 2k$. Then every edge e has as strong neighbourhood all of the rest of the graph, and consequently the strong neighbourhood induces about $\frac{25}{32}\Delta^4$ edges in the square of the line graph, which is much less than the $\frac{3}{2}\Delta^4$ of Lemma 2.1. In some sense, this is not surprising because already the degree $d^{\mathrm{s}}(e)$ of e in the square of the line graph is much smaller than the $2\Delta^2$ we are working with, namely it is only about $\frac{5}{4}\Delta^2$. In conclusion, the blow-up of the 5-cycle is quite different from what we assume in the sparsity lemma.

What kind of effects could be at work that explain this difference? In Lemma 2.1 we assume that the edge e has degree $d^{\mathrm{s}}(e)$ close to $2\Delta^2$ and at the same time a very dense neighbourhood. While it is possible to have such edges, this cannot be the case for all edges. Indeed, for the strong neighbourhood N^{s}_e to induce many edges, there have to be many 4-cycles in N^{s}_e. (This is simply because X and Y, the first and second neighbourhoods of the end-vertices of e, cannot be very large, so that the bipartite graph between X and Y is very dense.) However, every 4-cycle reduces the degree of any of its edges in $L^2(G)$ by 1, and consequently there should be many edges f in N^{s}_e of lower degree $d^{\mathrm{s}}(f)$ than $2\Delta^2$.

Obviously, an edge f of low degree $d^{\mathrm{s}}(f)$ is to our advantage, as we can always colour f at the very end when all high degree edges are already coloured. But if we defer colouring of low degree edges, then any high degree edge e with many low degree edges in its strong neighbourhood has, morally, low degree as well. Unfortunately, we did not manage to exploit these observations in such a way that they result in a substantial improvement.

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

7. Talagrand's inequality and exceptional outcomes

To finish the proof of Lemma 2.2, we still need to show that the probability of P_u or T_u deviating significantly from their expected values is very small. To prove that a random variable on a product probability space is strongly concentrated around its expectation is a very common task, and consequently, a number of powerful tools have been developed for this, among them McDiarmid's, Azuma's and Talagrand's inequalities [20, 3, 27]. All of these tools have in common that they require the random variable to be fairly smooth.

Consider a family of probability spaces $((\Omega_i, \Sigma_i, \mathbb{P}_i))_{i=1}^n$, and let $(\Omega, \Sigma, \mathbb{P})$ be their product. One common smoothness assumption for a random variable $X : \Omega \to \mathbb{R}$ is that *each coordinate has effect at most c*: whenever any two $\omega, \omega' \in \Omega$ differ in exactly one coordinate, then $|X(\omega) - X(\omega')| \leqslant c$.

For McDiarmid's or Azuma's inequality to give strong concentration in this situation, the effect c has to be small compared to n. If that is not the case, then Talagrand's inequality might still be useful. We describe a weaker version that is easier to apply than the full inequality.

We say that X *has certificates of size s for exceeding value k* if for any $\omega \in \Omega$ with $X(\omega) \geqslant k$, there is a set I of at most s coordinates such that also $X(\omega') \geqslant k$ for any $\omega' \in \Omega$ with $\omega|_I = \omega'|_I$. The following version of Talagrand's inequality appears in Molloy and Reed [24, page 234].

Theorem 7.1 (Talagrand). *Let $((\Omega_i, \Sigma_i, \mathbb{P}_i))_{i=1}^n$ be probability spaces, and let $(\Omega, \Sigma, \mathbb{P})$ be their product space. Let $X : \Omega \to \mathbb{R}$ be a non-negative random variable with $X \neq 0$ such that each coordinate has effect at most c, and assume X to have, for any k, certificates of size at most $k\ell$ for exceeding k. Then, for any $0 \leqslant t \leqslant \mathbb{E}[X]$,*

$$\mathbb{P}\left[|X - \mathbb{E}[X]| > t + 60c\sqrt{\ell \mathbb{E}[X]}\right] \leqslant 4e^{-t^2/(8c^2\ell\mathbb{E}[X])}.$$

Unfortunately, for the random variables P_u and T_u that are of interest to us, the effect c is too large, so Talagrand's inequality becomes useless. Indeed, in an extreme case, changing the colour of a single vertex $v \in N(u)$ to a new colour λ might result in all other vertices in $N(u)$ of colour λ losing their colour. This happens when all these vertices have an edge pointing away from v. Then, all pairs of colour λ counted in P_u are lost, and these might be up to r^2. (Recall that our graph G is r-regular.) Consequently, the effect c has to be at least r^2; however, even an effect of $c \approx r$ would necessitate a deviation of $t \approx r \approx \mathbb{E}[P_u]$ for the probability to become small. But for Talagrand's inequality to be useful for us, we need a vanishing probability for deviations t that are small compared to $\mathbb{E}[P_u]$.

Changing a single colour might have a very large effect, but this is a rare exception. Normally, few vertices in $N(u)$ have the same colour, so that also few are affected by any colour change. That is, very large effects only occur with a very tiny probability. It seems unreasonable that exceedingly unlikely events should have a serious impact on whether or not a random variable is concentrated.

What we need, therefore, is a version of Talagrand's inequality that excludes a very unlikely set Ω^* of *exceptional* outcomes that nevertheless spoils smoothness. Warnke [29] (but see also Kutin [18]) extended McDiarmid's inequality in a similar direction by considering a sort of *typical* effect. However, Warnke's inequality is still too weak for us. Grable [10] also presents

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

a concentration inequality that excludes exceptional outcomes, which would be suitable for our purposes were it not for the fact that there is a serious error in its proof. McDiarmid [21], too, describes a Talagrand-type inequality that excludes an exceptional set. (Its main feature, though, is to allow permutations as coordinates.) The inequality, however, does not seem to be of much use to us either.

We mention that the powerful, but technical, method of Kim and Vu [17] can also handle large but unlikely effects.

To deal with exceptional outcomes, we modify the definition of certificates. Given an exceptional set $\Omega^* \subseteq \Omega$ and $s, c > 0$, we say that X *has upward* (s,c)*-certificates* if, for every $t > 0$ and for every $\omega \in \Omega \setminus \Omega^*$, there is an index set I of size at most s so that $X(\omega') > X(\omega) - t$ for any $\omega' \in \Omega \setminus \Omega^*$ for which the restrictions $\omega|_I$ and $\omega'|_I$ differ in fewer than t/c coordinates.

Directly, Talagrand's inequality does not give concentration around the expectation but around the *median* $\mathrm{med}(X)$ of X, that is, around

$$\mathrm{med}(X) = \sup\left\{t \in \mathbb{R} : \mathbb{P}[X \leqslant t] \leqslant \frac{1}{2}\right\}.$$

However, in typical applications the median is very close to the expected value. We will deal with this technicality later.

Lemma 7.2. *Let* $((\Omega_i, \Sigma_i, \mathbb{P}_i))_{i=1}^n$ *be probability spaces, and let* $(\Omega, \Sigma, \mathbb{P})$ *be their product space. Let* $\Omega^* \subseteq \Omega$ *be a set exceptional events. Let* $X : \Omega \to \mathbb{R}$ *be a random variable and let* $t \geqslant 0$.

If X *has upward* (s,c)*-certificates then*

$$\mathbb{P}[|X - \mathrm{med}(X)| \geqslant t] \leqslant 4e^{-t^2/(4c^2 s)} + 4\mathbb{P}[\Omega^*]. \tag{7.1}$$

We prove Lemma 7.2 with the original, full version of Talagrand's inequality. Recall that the *Hamming distance* of two points $\omega, \omega' \in \Omega$ is defined by the number of non-common coordinates or equally $\sum_{i:\omega_i \neq \omega'_i} 1$. For a weighted version (together with a normalization), let $\alpha \in \mathbb{R}^n$ be a unit vector with $\alpha_i \geqslant 0$. We define the α*-Hamming distance* between ω and ω' by $\sum_{i:\omega_i \neq \omega'_i} \alpha_i$.

For a set $A \subseteq \Omega$ and a point $\omega \in \Omega$, let

$$d(\omega, A) = \sup_{\alpha}\left\{\tau : \sum_{i:\omega_i \neq \omega'_i} \alpha_i \geqslant \tau \text{ for all } \omega' \in A\right\},$$

which is equivalent to the largest value τ such that all points in A have α-Hamming distance at least τ from ω (for a best possible choice of α).

Theorem 7.3 (Talagrand [27]). *Let* $((\Omega_i, \Sigma_i, \mathbb{P}_i))_{i=1}^n$ *be probability spaces, and let* $(\Omega, \Sigma, \mathbb{P})$ *be their product space. If* $A, B \subseteq \Omega$ *are two (measurable) sets such that* $d(\omega, A) \geqslant \tau$ *for all* $\omega \in B$*, then*

$$\mathbb{P}[A]\mathbb{P}[B] \leqslant e^{-\tau^2/4}.$$

Proof of Lemma 7.2. The two-sided estimation (7.1) follows from the two one-sided estimations

$$\mathbb{P}[X \leqslant \mathrm{med}(X) - t] \leqslant 2e^{-t^2/(4c^2 s)} + 2\mathbb{P}[\Omega^*], \quad \mathbb{P}[X \geqslant \mathrm{med}(X) + t] \leqslant 2e^{-t^2/(4c^2 s)} + 2\mathbb{P}[\Omega^*],$$

of which we only show the first; the argumentation for the second is analogous. Let

$$A = \{\omega \in \Omega \setminus \Omega^* : X(\omega) \leqslant \mathrm{med}(X) - t\} \quad \text{and} \quad B = \{\omega \in \Omega \setminus \Omega^* : X(\omega) \geqslant \mathrm{med}(X)\}.$$

Pick an arbitrary $\omega \in B$. By assumption, X has upward (s,c)-certificates, which means, in particular, that there is an index set I of at most s indices such that $\omega|_I$ and $\omega'|_I$ differ in at least t/c coordinates for every $\omega' \in A$. Consequently, if we set $\alpha = 1/\sqrt{|I|} \cdot \mathbb{1}_I$, where $\mathbb{1}_I$ is the characteristic vector of I, then ω and ω' have α-Hamming distance at least $t/(c\sqrt{s})$. Hence $d(\omega, A) \geqslant t/(c\sqrt{s})$.

Using Theorem 7.3, we obtain $\mathbb{P}[A]\mathbb{P}[B] \leqslant e^{-t^2/(4c^2 s)}$. We conclude that

$$\begin{aligned} \mathbb{P}[X \leqslant \mathrm{med}(X) - t] \cdot \frac{1}{2} &\leqslant \mathbb{P}[X \leqslant \mathrm{med}(X) - t]\mathbb{P}[X \geqslant \mathrm{med}(X)] \\ &\leqslant \mathbb{P}[A]\mathbb{P}[B] + \mathbb{P}[\Omega^*] \\ &\leqslant e^{-t^2/(4c^2 s)} + \mathbb{P}[\Omega^*]. \end{aligned}$$

This completes the proof. □

Next, we prove that usually the median is close to the expected value.

Lemma 7.4. *Let* $((\Omega_i, \Sigma_i, \mathbb{P}_i))_{i=1}^n$ *be probability spaces, and let* $(\Omega, \Sigma, \mathbb{P})$ *be their product space. Let* $\Omega^* \subseteq \Omega$. *Let* $X : \Omega \to \mathbb{R}$ *be a random variable, let* $M = \max\{\sup|X|, 1\}$, *and let* $c \geqslant 1$. *If* X *has upward* (s,c)*-certificates, then*

$$|\mathbb{E}[X] - \mathrm{med}(X)| \leqslant 20c\sqrt{s} + 20M^2\mathbb{P}[\Omega^*].$$

We note that the proof technique is not new. A similar proof appears, for instance, in Molloy and Reed [24, Chapter 20].

Proof. Clearly,

$$|\mathbb{E}[X] - \mathrm{med}(X)| \leqslant \mathbb{E}[|X - \mathrm{med}(X)|].$$

Put $K = 2M/(c\sqrt{s})$, and observe that $|X - \mathrm{med}(X)| < (K+1)c\sqrt{s}$. By splitting the possible values of $|X - \mathrm{med}(X)|$ into intervals of length $c\sqrt{s}$, we find an upper bound as follows:

$$\mathbb{E}[|X - \mathrm{med}(X)|] \leqslant \sum_{k=0}^{K} c\sqrt{s}(k+1)\mathbb{P}[|X - \mathrm{med}(X)| \geqslant kc\sqrt{s}\}].$$

We apply Lemma 7.2 for each summand with $t = kc\sqrt{s}$:

$$\begin{aligned} \mathbb{E}[|X - \mathrm{med}(X)|] &\leqslant c\sqrt{s}\sum_{k=0}^{K} 4(k+1)(e^{-k^2/4} + \mathbb{P}[\Omega^*]) \\ &\leqslant 20c\sqrt{s} + 20M^2\mathbb{P}[\Omega^*], \end{aligned}$$

as

$$\sum_{k=0}^{\infty}(k+1)e^{-k^2/4} \approx 4.1869 < 5. \qquad \square$$

We conclude under the assumptions of Lemma 7.4 by combining Lemmas 7.2 and 7.4:

$$\mathbb{P}\big[|X-\mathbb{E}[X]| \geqslant t+20c\sqrt{s}+20M^2\mathbb{P}[\Omega^*]\big] \leqslant 4e^{-t^2/(4c^2s)}+4\mathbb{P}[\Omega^*]. \tag{7.2}$$

Let us return to the certificates. As defined, they are witnesses for large values of X. Sometimes it is easier to certify smaller values. To capture such situations we say that X *has downward* (s,c)-*certificates* if, for every $t>0$, and for every $\omega \in \Omega\setminus\Omega^*$, there is an index set I of size at most s such that $X(\omega') < X(\omega)+t$ for every $\omega' \in \Omega\setminus\Omega^*$ for which the restrictions $\omega|_I$ and $\omega'|_I$ differ in less than t/c coordinates. By replacing X with $-X$, we observe that Lemmas 7.2 and 7.4, and thus (7.2), remain valid for downward certificates.

We simplify (7.2) a bit more. If $t \geqslant 50c\sqrt{s}$ and $\mathbb{P}[\Omega^*] \leqslant M^{-2}$, then

$$t \geqslant t/2+(20c\sqrt{s}+20M^2\mathbb{P}[\Omega^*]).$$

Thus we obtain the following result.

Theorem 7.5. *Let* $((\Omega_i,\Sigma_i,\mathbb{P}_i))_{i=1}^n$ *be probability spaces, let* $(\Omega,\Sigma,\mathbb{P})$ *be their product, and let* $\Omega^* \subset \Omega$ *be a set of exceptional outcomes. Let* $X:\Omega\to\mathbb{R}$ *be a random variable, let* $M=\max\{\sup|X|,1\}$*, and let* $c \geqslant 1$*. Suppose* $\mathbb{P}[\Omega^*] \leqslant M^{-2}$ *and* X *has upward* (s,c)*-certificates or downward* (s,c)*-certificates. Then, for* $t > 50c\sqrt{s}$*,*

$$\mathbb{P}[|X-\mathbb{E}[X]| \geqslant t] \leqslant 4e^{-t^2/(16c^2s)}+4\mathbb{P}[\Omega^*].$$

8. Concentration for colouring procedure

We finally complete the proof of Lemma 2.2 by proving that for every vertex u the random variables P_u and T_u are tightly concentrated around their expected values. Recall that P_u counts the number of vertex pairs in $N(u)$ that have the same colour at the end of the colouring procedure, and recall that T_u counts the number of such vertex triples.

Proof of Lemma 5.1. We show the statement for P_u; the proof for T_u is almost the same.

Recall that the colouring procedure that defines P_u (and T_u) is based on a random experiment on a product space $\Omega=\prod\Omega_i$, where the product ranges over the vertices (that receive a colour) and the edges (that receive a direction) of the graph. Thus, every event $\omega\in\Omega$ is a vector that is indexed by $V(G)\cup E(G)$.

To apply Theorem 7.5, we define the set Ω^* of exceptional events as the space of events that assign some colour to more than $\log r$ vertices in $N(u)$ (before uncolouring). Recall that we work with $C=\lceil(1-\gamma)r\rceil$ colours. To estimate $\mathbb{P}[\Omega^*]$, observe first that the probability that a particular colour appears more than $\log r$ times in $N(u)$ is at most

$$\sum_{i=\lfloor\log r\rfloor+1}^{r}\binom{r}{i}\frac{1}{C^i} \leqslant \sum_{i=\lfloor\log r\rfloor+1}\left(\frac{er}{i}\right)^i\frac{1}{(1-\gamma)^i r^i} \leqslant r\cdot\left(\frac{e}{(1-\gamma)\log r}\right)^{\log r}$$

Thus, we get

$$\mathbb{P}[\Omega^*] \leqslant r^2 \cdot \left(\frac{e}{(1-\gamma)\log r} \right)^{\log r} \leqslant r^{-\frac{2}{3}\log\log r},$$

for large enough r.

Setting $s = 3r$ and $c = \log^2 r$, let us check that P_u has downward (s,c)-certificates. So, let $t > 0$ be given and consider an event $\omega \notin \Omega^*$. We have to define an index set I of size at most s. We start by including all vertices of $N(u)$ in I. Next, consider any vertex $v \in N(u)$ that becomes uncoloured in step (3) of the colouring procedure. This is only the case if there is a neighbour v' of v that receives the same colour under ω and if, again subject to ω, the edge vv' points towards v. We add v' and vv' to I for each such vertex v. In total we have $|I| \leqslant 3r = s$, as for every vertex $v \in N(u)$ we add at most two additional indices to I.

Now let $\omega' \notin \Omega^*$ be an event with $P_u(\omega') \geqslant P_u(\omega) + t$. As $\omega' \notin \Omega^*$, every colour may contribute at most $\binom{\log r}{2} \leqslant \frac{1}{2}\log^2 r$ pairs to P_u.

For every colour λ for which there are more pairs of vertices coloured with λ contributing to $P_u(\omega')$ than to $P_u(\omega)$, there is a vertex v in $N(u)$ of colour λ under ω' but that, under ω, is either uncoloured or coloured with a different colour to λ.

In the latter case, ω and ω' differ in the coordinate $v \in I$ (as v receives different colours under ω and ω'). For the former case, observe that we had added a vertex v' and the edge vv' to I to witness v being uncoloured under ω. Either the colour of v' or the direction of vv' must be different in ω', that is, ω and ω' differ again in at least one coordinate of I. In both cases, call such a coordinate a *λ-difference*.

Can a coordinate in I be a λ-difference and a μ-difference for two different colours λ and μ? Yes, but only if it is a vertex $v' \in I$ in $N(u)$ that satisfies two conditions: first, under ω it serves as a witness for one of its neighbours $v \in N(u)$ losing its colour λ, say, in the conflict resolution step of the colouring procedure; and second, by flipping from colour λ under ω to colour μ in ω', the vertex v' contributes new pairs of colour μ in P_u. This then immediately shows that no coordinate can be a λ-difference for three (or more) colours.

As a consequence, ω' and ω must differ in at least $t/\log^2 r = t/c$ coordinates in I as $P_u(\omega') \geqslant P_u(\omega) + t$. (Recall that under ω' no colour can contribute more than $\frac{1}{2}\log^2 r$ pairs to P_u.) This proves that P_u has downward (s,c)-certificates.

For Theorem 7.5, we set $M = \sup P_u \leqslant r^2$. With $t = \log^3 r\sqrt{r}$, Theorem 7.5 implies for large r that

$$\begin{aligned}\mathbb{P}[|P_u - \mathbb{E}[P_u]| \geqslant \sqrt{r}\log^3 r] &\leqslant 4\exp\left(-\frac{r\log^6 r}{16\log^4 r \cdot 3r}\right) + 4r^{-\frac{2}{3}\log\log r}\\ &\leqslant r^{-\frac{1}{2}\log\log r}\end{aligned}$$

The only difference of the proof for T_u lies in the fact that, outside Ω^*, every colour can contribute up to $\log^3 r/6$ triples to T_u; for P_u this was at most $\log^2 r/2$ pairs. The resulting higher value $\log^3 r$ for c can easily be compensated for by increasing t to $\log^4 r\sqrt{r}$ in the application of Theorem 7.5. □

9. Exceptional outcomes spoil triangle counting

We close this article by arguing that Theorem 7.5 has potential applications beyond our colouring lemma. To make this case, we discuss the number of triangles in a random graph. We note that this problem also serves as a motivating example for Kim and Vu [17], and for Warnke [29].

Consider the random graphs $\mathcal{G}(n,p)$ that are obtained from K_n by retaining each edge with probability p, independently of the others, where $p = p(n)$ may be a function in n. The threshold probability for the triangles is $p = 1/n$: below that threshold there is with high probability no triangle; above it there is with high probability at least one. Let us now examine the number T of triangles in $\mathcal{G}(n,p)$, or rather the expected number of triangles

$$\mathbb{E}[T] = \binom{n}{3} p^3 \approx \frac{1}{6} n^3 p^3.$$

Is T concentrated around its expected value whenever $np \to \infty$? We consider here a relatively mild notion of concentration, where we allow deviations from the expected value of up to $t = \varepsilon n^3 p^3$ for small but fixed $\varepsilon > 0$. For simpler notation, set $N = \binom{n}{2}$.

For p relatively large, that is, for $p \geqslant n^{-1/3+\gamma}$ for any $\gamma > 0$, McDiarmid's inequality [20] is strong enough to show concentration. Indeed, changing any coordinate, that is, any edge, may result in at most n new triangles (or at most n triangles fewer), so that the effect c of each coordinate is bounded by n. Consequently, for McDiarmid's inequality to show that $|T - \mathbb{E}[T]| \leqslant t$ for $t = \varepsilon n^3 p^3$, it is necessary that $t^2/(N \cdot n^2)$ tends to infinity. This is the case if $p \geqslant n^{-1/3+\gamma}$. A version of McDiarmid's inequality for binary random variables (see again [20]) allows us to drop this threshold to $p \geqslant n^{-2/5+\gamma}$.

To go below this threshold, Warnke [29] (but also others) observed that, while changing a single edge may create up to $n-2$ new triangles (or destroy that many), this is exceedingly unlikely. Indeed, we only expect a particular edge to be in roughly np^2 many triangles. Thus, by the standard Chernoff bound, for any $\delta > 0$ it is extremely unlikely that an edge is contained in more than $\max\{2np^2, n^\delta\}$ triangles.

Exploiting the fact that, typically, the effect of changing a single edge is much smaller, Warnke could verify concentration as long as $t^2/(pN \cdot \max\{2np^2, n^\delta\})$ tends to infinity, which is the case when $p \geqslant n^{-4/5+\gamma}$.

To reduce beyond such p, Kim and Vu [17] developed a powerful method that evidently has a very wide scope of application. Usually, great power does not come for free, and this is also the case here: Kim and Vu's inequality is rather technical and not easy to use.

Let us now apply Theorem 7.5 and show the strong concentration of T also for values of p smaller than $n^{-4/5+\gamma}$. We exclude all outcomes where at least one edge is contained in more than n^δ many triangles. As seen above, this is an event of very small probability. Moreover, we may use Warnke's result (or a previous application of Theorem 7.5) to observe that it is extremely unlikely that

$$T \geqslant \frac{2}{6} n^3 \cdot n^{3(-4/5+\gamma)} = \frac{1}{3} n^{3/5+3\gamma}.$$

(We have used here that T is monotone in p.) We add all these outcomes to our exceptional set.

Next, let us check that T has upward (s,c)-certificates, where $s = n^{3/5+3\gamma}$ and $c = n^\delta$. For a non-exceptional event ω, we use as index set I the set of all edges lying in any triangle. As there

https://doi.org/10.1017/S0963548317000244 Published online by Cambridge University Press

are at most $\frac{1}{3}n^{3/5+3\gamma}$ triangles, it follows that $|I| \leqslant s$. Now, consider some non-exceptional event ω' such that ω and ω' differ in less than t'/c coordinates of I. Then, any edge in I that is present in ω but lost in ω' may only result in ω' having at most n^{δ} fewer triangles than ω. Moreover, edges outside I obviously cannot result in the loss of triangles. Therefore, $T(\omega') > T(\omega) - t'$, and we see that T has upward (s,c)-certificates.

With these values of s and c, we deduce from Theorem 7.5 that T is strongly concentrated if $p \geqslant n^{-9/10+\gamma}$.

For even smaller values of p, we may apply Theorem 7.5 once again, and set $s \approx n^{-9/10+\gamma}$, which will then yield concentration for $p \geqslant n^{-19/20+\gamma}$. Of course, this can be iterated several times, so that we obtain strong concentration for $p \geqslant n^{-1+\beta}$ for every $\beta > 0$.

Let us finally point out that we even have fairly tight concentration around the expected value, namely, T is very likely within the range $\mathbb{E}[T] \pm n^{\delta}\sqrt{\mathbb{E}[T]}$.

Acknowledgement

We would both like to thank Ross Kang for insightful discussions.

References

[1] Alon, N., Krivelevich, M. and Sudakov, B. (1999) Coloring graphs with sparse neighborhoods. *J. Combin. Theory Ser. B* **77** 73–82.

[2] Andersen, L. D. (1992) The strong chromatic index of a cubic graph is at most 10. *Discrete Math.* **108** 231–252.

[3] Azuma, K. (1967) Weighted sums of certain dependent random variables. *Tôhoku Math. J.* (2) **19** 357–367.

[4] Bonamy, M., Perrett, T. and Postle, L. (2016) Reed's conjecture and strong edge coloring. Paper given at Scottish Combinatorics Meeting, Glasgow.

[5] Chung, F. R. K., Gyárfás, A., Tuza, Z. and Trotter, W. T. (1990) The maximum number of edges in 2K2-free graphs of bounded degree. *Discrete Math.* **81** 129–135.

[6] Cranston, D. W. (2006) Strong edge-coloring of graphs with maximum degree 4 using 22 colors. *Discrete Math.* **306** 2772–2778.

[7] Diestel, R. (2010) *Graph Theory*, fourth edition, Springer.

[8] Faudree, R. J., Schelp, R. H., Gyárfás, A. and Tuza, Z. (1989) Induced matchings in bipartite graphs. *Discrete Math.* **78** 83–87.

[9] Faudree, R. J., Schelp, R. H., Gyárfás, A. and Tuza, Z. (1990) The strong chromatic index of graphs. *Ars Combin.* **29** 205–211.

[10] Grable, D. A. (1998) A large deviation inequality for functions of independent, multi-way choices. *Combin. Probab. Comput.* **7** 57–63.

[11] Horák, P., Qing, H. and Trotter, W. T. (1993) Induced matchings in cubic graphs. *J. Graph Theory* **17** 151–160.

[12] Joos, F. (2016) Induced matchings in graphs of bounded maximum degree. *SIAM J. Discrete Math.* **30** 1876–1882.

[13] Joos, F. and Nguyen, V. H. (2016) Induced matchings in graphs of degree at most 4. *SIAM J. Discrete Math.* **30** 154–165.

[14] Joos, F., Rautenbach, D. and Sasse, T. (2014) Induced matchings in subcubic graphs. *SIAM J. Discrete Math.* **28** 468–473.

[15] Kaiser, T. and Kang, R. (2014) The distance-t chromatic index of graphs. *Combin. Probab. Comput.* **23** 90–101.

[16] Kim, J. H. (1995) On Brooks' theorem for sparse graphs. *Combin. Probab. Comput.* **4** 97–132.

[17] Kim, J. H. and Vu, V. H. (2000) Concentration of multivariate polynomials and its applications. *Combinatorica* **20** 417–434.

[18] Kutin, S. (2002) Extensions to McDiarmid's inequality when differences are bounded with high probability. Technical Report TR-2002-04, University of Chicago.

[19] Mahdian, M. (2000) The strong chromatic index of C_4-free graphs. *Random Struct. Alg.* **17** 357–375.

[20] McDiarmid, C. (1989) On the method of bounded differences. In *Surveys in Combinatorics*, Vol. 141 of London Mathematical Society Lecture Note series, Cambridge University Press, pp. 148–188.

[21] McDiarmid, C. (2002) Concentration for independent permutations. *Combin. Probab. Comput.* **11** 163–178.

[22] Molloy, M. and Reed, B. (1997) A bound on the strong chromatic index of a graph. *J. Combin. Theory Ser. B* **69** 103–109.

[23] Molloy, M. and Reed, B. (1998) A bound on the total chromatic number. *Combinatorica* **18** 241–280.

[24] Molloy, M. and Reed, B. (2002) *Graph Coloring and the Probabilistic Method*, Springer.

[25] Rivin, I. (2002) Counting cycles and finite dimensional L^p norms. *Adv. Appl. Math.* **29** 647–662.

[26] Śleszyńska-Nowak, M. (2016) Clique number of the square of a line graph. *Discrete Math.* **339** 1551–1556.

[27] Talagrand, M. (1995) Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81** 73–205.

[28] van Lint, J. H. (1999) *Introduction to Coding Theory*, Graduate Texts in Mathematics, New York University Press.

[29] Warnke, L. (2016) On the method of typical bounded differences. *Combin. Probab. Comput.* **25** 269–299.