# UPPER BOUNDS FOR MULTICOLOUR RAMSEY NUMBERS

PAUL BALISTER, BÉLA BOLLOBÁS, MARCELO CAMPOS, SIMON GRIFFITHS, EOIN HURLEY, ROBERT MORRIS, JULIAN SAHASRABUDHE, AND MARIUS TIBA

Abstract. The $r$-colour Ramsey number $R_r(k)$ is the minimum $n \in \mathbb{N}$ such that every $r$-colouring of the edges of the complete graph $K_n$ on $n$ vertices contains a monochromatic copy of $K_k$. We prove, for each fixed $r \geqslant 2$, that

$$R_r(k) \leqslant e^{-\delta k} r^{rk}$$

for some constant $\delta = \delta(r) > 0$ and all sufficiently large $k \in \mathbb{N}$. For each $r \geqslant 3$, this is the first exponential improvement over the upper bound of Erdős and Szekeres from 1935. In the case $r = 2$, it gives a different (and significantly shorter) proof of a recent result of Campos, Griffiths, Morris and Sahasrabudhe.

## 1. Introduction

One of the most classical (and notorious) problems in combinatorics is to estimate the $r$-colour Ramsey numbers $R_r(k)$, the minimum $n \in \mathbb{N}$ such that every $r$-colouring of the edges of the complete graph on $n$ vertices contains a monochromatic clique on $k$ vertices. Ramsey [10] proved in 1930 that $R_r(k)$ is finite for every $k, r \in \mathbb{N}$, and a few years later his theorem was rediscovered in a famous paper of Erdős and Szekeres [7], whose method gives the bound $R_r(k) \leqslant r^{rk}$. This bound turns out to be not too far from the truth: Erdős [6] showed in 1947 that a random colouring gives an exponential lower bound on $R_2(k)$, and his proof can easily be adapted to show that $R_r(k) \geqslant r^{k/2}$. When $r \geqslant 3$, this lower bound was significantly improved by Abbott [1], who combined Erdős' colouring with a simple product construction to show that

$$c^{rk} \leqslant R_r(k) \leqslant r^{rk} \tag{1}$$

for some constant $c > 1$. The value of the constant $c$ in (1) has been improved in recent years, first by Conlon and Ferber [5], and subsequently by Wigderson [15] and Sawin [12], though in the case $r = 2$ the best known bound (proved in [13]) still uses Erdős' random colouring, and only improves the bound from [6] by a factor of 2.

Progress on the upper bound has been much slower, except in the case $r = 2$, where the upper bound of Erdős and Szekeres was improved by Rödl (see [8]) and Thomason [14] in the 1980s, by Conlon [3] in 2009, and by Sah [11] in 2023, leading to a bound of the form

$$R_2(k) \leqslant e^{-c(\log k)^2} 4^k$$

SG was partially supported by FAPERJ (Proc. 201.194/2022) and by CNPq (Proc. 407970/2023-1); EH by the Gravitation Programme NETWORKS (024.002.003) of NWO; RM by FAPERJ (Proc. E-26/200.977/2021) and by CNPq (Procs 303681/2020-9 and 407970/2023-1); and JS by European Research Council (ERC) Starting Grant "High Dimensional Probability and Combinatorics", grant No. 101165900.

for some constant $c > 0$. An exponential improvement was finally obtained in 2023 by Campos, Griffiths, Morris and Sahasrabudhe [2], who showed that

$$R_2(k) \leqslant (4 - \varepsilon)^k$$

for some constant $\varepsilon > 0$ and all sufficiently large $k \in \mathbb{N}$. The method of [2] has since been adapted and optimised by Gupta, Ndiaye, Norin and Wei [9], who obtained the bound

$$R_2(k) \leqslant 3.8^k.$$

The authors of [9] also noted that the method of [2] can easily be extended to certain multicolour off-diagonal Ramsey numbers (see below), but when $r \geqslant 3$ it does not seem to be strong enough to give any improvement at all over the bound from [7] for $r$-colour Ramsey numbers that are close to the diagonal. In fact, we are not aware of *any* previous improvement[1] on the upper bound of Erdős and Szekeres for $R_r(k)$ when $r \geqslant 3$.

In this paper we will introduce a new approach to proving upper bounds on $R_r(k)$, and use it to give an exponential improvement over [7] for all fixed $r \geqslant 2$.

**Theorem 1.1.** *For each $r \geqslant 2$, there exists $\delta = \delta(r) > 0$ such that*

$$R_r(k) \leqslant e^{-\delta k} r^{rk}$$

*for all sufficiently large $k \in \mathbb{N}$.*

In particular, in the case $r = 2$ we will provide a different (and much shorter) proof of the main result from [2]. Moreover, the constant $\delta(r)$ given by our proof is polynomial in $r$, so we obtain an improvement over the bound of Erdős and Szekeres for all $r = k^{o(1)}$.

The main new ingredient in our proof of Theorem 1.1 is a geometric lemma (see Lemma 3.1, below) which (roughly speaking) says that if $r$ functions $f_1, \ldots, f_r \colon X \to \mathbb{R}^n$ defined on a finite set $X$ exhibit a large amount of negative correlation, in the sense that

$$\mathbb{P}\Big( \big\langle f_i(x), f_i(y) \big\rangle > -1 \text{ for all } i \in [r] \Big) \approx 0,$$

where $x$ and $y$ are independent random elements of $X$, then one of the functions $f_\ell$ must exhibit a significant amount of 'clustering', in the sense that it maps many pairs of elements of $X$ to pairs of points with large inner product. The second key new idea is to build $r$ monochromatic books (instead of only one, as in [2]), using the geometric lemma to significantly boost the density of edges of colour $\ell$ between our reservoir set and the book of colour $\ell$ whenever performing a 'normal' (Erdős–Szekeres) step in any of the colours would cause the density of edges in that colour between the reservoir and the corresponding book to decrease too much. As a consequence, we will be able to build a $(t, m)$-book (that is, a clique of size $t$ connected to $m$ extra vertices) with $t = \varepsilon k$ for some small constant $\varepsilon > 0$, and $m \approx n/r^t$ pages, avoiding the delicate tradeoffs and calculations of [2].

[1] Conlon suggested in his PhD thesis [4, page 46] that it may be possible to adapt his method to give a similar improvement for $R_r(k)$ when $r \geqslant 3$, but also noted several obstructions to obtaining such a result that would seem to require significant additional ideas to overcome.

The rest of this short paper is organised as follows. In Section 2 we will state our main technical result about the existence of monochromatic books in $r$-colourings, and describe the algorithm that we will use to prove it. In Section 3 we will prove the key lemma, and in Section 4 we will use it to deduce the book theorem. Finally, in Section 5, we will deduce our bound on $R_r(k)$ from the book theorem.

## 2. The Book Theorem

The book $(A, B)$ is the graph formed by removing the clique with vertex set $B$ from the clique with vertex set $A \cup B$; that is, it is the graph $H$ with vertex set $A \cup B$ and edge set

$$E(H) = \big\{ uv : \{u, v\} \not\subset B \big\},$$

We say that $H$ is a $(t, m)$-*book* if $|A| = t$ and $|B| = m$, and the sets $A$ and $B$ are disjoint. Our plan is to find a monochromatic copy of $K_k$ (in an arbitrary $r$-colouring $\chi$ of $E(K_n)$) by first finding a monochromatic $(t, m)$-book, where

$$t = \varepsilon k \qquad \text{and} \qquad m \geqslant e^{-t^2/8k}\, r^{-t} \cdot n \geqslant R_r(k, \ldots, k, k - t),$$

for some constant $\varepsilon = \varepsilon(r) > 0$, where $R_r(k_1, \ldots, k_r)$ is the minimum $N$ such that every $r$-colouring of $E(K_N)$ contains a monochromatic copy of $K_{k_i}$ in colour $i$ for some $i \in [r]$. Since we have

$$R_r(k, \ldots, k, k - t) \leqslant \binom{kr - t}{k, \ldots, k, k - t} \leqslant e^{-t^2/6k} r^{rk-t},$$

by the method of Erdős and Szekeres [7], this will suffice to prove Theorem 1.1.

We will find a monochromatic $(t, m)$-book by constructing $r$ monochromatic books, one in each colour. The first step is to find a large subset of the vertices in which the colouring is almost regular, and every colour has density close to $1/r$. To do so, we simply run the Erdős–Szekeres algorithm until we find such a subset: since we 'win' a factor of $1 + \varepsilon$ in each step (see Lemma 5.2), either we find a suitable set within $\varepsilon k$ steps, or we are already done. We then (randomly) partition[2] the remaining vertices into $r + 1$ sets: a 'reservoir' set $X$, and sets $Y_1, \ldots, Y_r$ that we will use to construct our books. We will find a set $T \subset X$ of size $t = \varepsilon k$ such that $T$ induces a monochromatic clique in some colour $i \in [r]$, and

$$\bigg| \bigcap_{u \in T} N_i(u) \cap Y_i \bigg| \geqslant m,$$

where we write $N_i(u)$ to denote the neighbourhood of $u$ in colour $i$.

We are now ready to state our main technical theorem about finding monochromatic books in $r$-colourings. With an eye to potential future applications, we will state a more general version than we need for Theorem 1.1. In particular, the parameter $\mu$ (which in our application will be $\Theta(r^3)$) allows us to control the loss in the size of the book (relative to $p^t|Y_i|$) in terms of the size of the reservoir set $X$. To avoid repetition, let us fix $n, r \in \mathbb{N}$.

[2] In fact, this is unnecessary: we can simply take $X = Y_1 = \cdots = Y_r$ all equal to the set of remaining vertices. However, the reader may find it easier to picture the sets as disjoint.

**Theorem 2.1.** *Let $\chi$ be an $r$-colouring of $E(K_n)$, and let $X, Y_1, \ldots, Y_r \subset V(K_n)$. For every $p > 0$ and $\mu \geqslant 2^{10} r^3$, and every $t, m \in \mathbb{N}$ with $t \geqslant \mu^5/p^2$, the following holds. If*

$$|N_i(x) \cap Y_i| \geqslant p|Y_i|$$

*for every $x \in X$ and $i \in [r]$, and moreover*

$$|X| \geqslant \left(\frac{\mu^2}{p}\right)^{\mu r t} \qquad \text{and} \qquad |Y_i| \geqslant \left(\frac{e^{2^{13} r^3/\mu^2}}{p}\right)^t m,$$

*then $\chi$ contains a monochromatic $(t, m)$-book.*

The algorithm that we will use to prove Theorem 2.1 is somewhat simpler than the one introduced in [2]. It is based on the following key lemma, which we will use in each step. Given sets $X, Y \subset V(K_n)$ and a colour $i \in [r]$, define

$$p_i(X, Y) = \min\left\{\frac{|N_i(x) \cap Y|}{|Y|} : x \in X\right\},$$

and note that $p_i(X', Y) \geqslant p_i(X, Y)$ for every subset $X' \subset X$. Set $\beta = 3^{-4r}$ and $C = 4r^{3/2}$.

**Lemma 2.2.** *Let $\chi$ be an $r$-colouring of $E(K_n)$, let $X, Y_1, \ldots, Y_r \subset V(K_n)$ be non-empty sets of vertices, and let $\alpha_1, \ldots, \alpha_r > 0$. There exists a vertex $x \in X$, a colour $\ell \in [r]$, sets $X' \subset X$ and $Y_1', \ldots, Y_r'$ with $Y_i' \subset N_i(x) \cap Y_i$ for each $i \in [r]$, and $\lambda \geqslant -1$, such that*

$$|X'| \geqslant \beta e^{-C\sqrt{\lambda+1}} |X| \qquad \text{and} \qquad p_\ell(X', Y_\ell') \geqslant p_\ell(X, Y_\ell) + \lambda \alpha_\ell, \tag{2}$$

*and moreover*

$$|Y_i'| = p_i(X, Y_i)|Y_i| \qquad \text{and} \qquad p_i(X', Y_i') \geqslant p_i(X, Y_i) - \alpha_i \tag{3}$$

*for every $i \in [r]$.*

Since the statement of the lemma might seem a little confusing at first sight, before continuing let us explain roughly how we use it in the algorithm below. We will build a collection of sets $T_1, \ldots, T_r$, with each $T_i$ being a monochromatic clique in colour $i$, and replace the sets $X$ and $Y_1, \ldots, Y_r$ by subsets satisfying

$$X \subset \bigcap_{j \in [r]} \bigcap_{u \in T_j} N_j(u) \qquad \text{and} \qquad Y_i \subset \bigcap_{u \in T_i} N_i(u)$$

for each $i \in [r]$, and such that $p_i(X, Y_i)$ does not decrease too much below its initial value. In each step we will apply Lemma 2.2 and then, depending on the value of $\lambda$, either add the vertex $x$ to one of the sets $T_j$ (a 'colour step'), or replace $X$ and $Y_\ell$ by $X'$ and $Y_\ell'$, and observe that $p_\ell(X, Y_\ell)$ increases significantly (a 'density-boost step').

To be more precise, if the $\lambda$ given by Lemma 2.2 is 'small', then we will choose a colour $j \in [r]$ such that the set $X'' = N_j(x) \cap X'$ is as large as possible, and update as follows:

$$X \to X'', \qquad Y_j \to Y_j' \qquad \text{and} \qquad T_j \to T_j \cup \{x\}.$$

This update does not cause $p_j(X, Y_j)$ to decrease too much (by (3)), and does not cause $p_i(X, Y_i)$ to decrease at all if $i \neq j$, since the set $Y_i$ does not change and $X'' \subset X$. Moreover, since $\lambda$ is small, the update does not cause the set $X$ to shrink too much.

If $\lambda$ is 'large', on the other hand, then we instead update the sets as follows:

$$X \to X' \qquad \text{and} \qquad Y_\ell \to Y'_\ell,$$

with all other sets staying the same. This time $X$ shrinks by a much larger factor, but to compensate, $p_\ell(X, Y_\ell)$ increases by $\lambda\alpha_\ell$. It will turn out that we are happy with this trade-off because the function $e^{-C\sqrt{\lambda+1}}$ that appears in our bound (2) on the size of $X'$ is sub-exponential in $\lambda$ (see below for discussion of why this is sufficient).

We are now ready to define the algorithm that we will use to prove Theorem 2.1. The inputs to the algorithm are an $r$-colouring $\chi$ of $E(K_n)$, sets $X, Y_1, \dots, Y_r \subset V(K_n)$, and parameters $t \in \mathbb{N}$ (the size of the book that we are trying to build), $\lambda_0 \geqslant -1$ (the cut-off between values of $\lambda$ that are 'small' and 'large'), and $\delta > 0$ (a small constant that bounds the total decrease in $p_i(X, Y_i)$). We also set

$$p_0 = \min\big\{p_i(X, Y_i) : i \in [r]\big\},$$

and emphasize that $t$, $\lambda_0$, $\delta$ and $p_0$ remain fixed throughout the algorithm.

**The Multicolour Book Algorithm.** Set $T_1 = \cdots = T_r = \emptyset$, and repeat the following steps until either $X = \emptyset$ or $\max\big\{|T_i| : i \in [r]\big\} = t$.

1. Applying the key lemma: let the vertex $x \in X$, the colour $\ell \in [r]$, the sets $X' \subset X$ and $Y'_1, \dots, Y'_r$, and $\lambda \geqslant -1$ be given by Lemma 2.2, applied with
$$\alpha_i = \frac{p_i(X, Y_i) - p_0 + \delta}{t} \tag{4}$$
for each $i \in [r]$, and go to Step 2.
2. Colour step: If $\lambda \leqslant \lambda_0$, then choose a colour $j \in [r]$ such that the set
$$X'' = N_j(x) \cap X'$$
has at least $(|X'| - 1)/r$ elements, and update the sets as follows:
$$X \to X'', \qquad Y_j \to Y'_j \qquad \text{and} \qquad T_j \to T_j \cup \{x\}$$
and go to Step 1. Otherwise go to Step 3.
3. Density-boost step: If $\lambda > \lambda_0$, then we update the sets as follows:
$$X \to X' \qquad \text{and} \qquad Y_\ell \to Y'_\ell,$$
and go to Step 1.

We remark that in our application of the algorithm, both $\lambda_0$ and $\delta$ will be polynomial functions of $r$, and our sets and colouring will satisfy $p_0 \geqslant 1/r - \delta$. Let us also emphasize once again that we only update one of the sets $Y_i$ in each round of the algorithm (either the set $Y_j$ in Step 2 or the set $Y_\ell$ in Step 3), and at most one of the sets $T_i$.

It remains to explain our choice of $\alpha_i$, and why it is important that the bound on the size of $X'$ given by Lemma 2.2 is sub-exponential in $\lambda$. The definition of $\alpha_i$ is just a continuous version of the definition of the parameter $\alpha$ from [2], with the parameter $\delta$ controlling the minimum value that it can take. The function (4) has two important properties: it will never be smaller than $\delta/4t$, and it increases linearly with $p_i(X,Y_i)$, which implies (see Lemma 4.3) that the total number of density-boost steps is much smaller than $t$, as long as $\lambda_0 \gg \log(1/\delta)$. This second property of $\alpha_i$ moreover allows us to bound the sum of the $\lambda$ that are used in density-boost steps by $O\big(\log(1/\delta)\cdot t\big)$, which in turn implies that our reservoir set $X$ does not shrink too much as a result of density-boost steps (see Lemmas 4.5 and 4.6). It is here that we require the factor by which $X$ shrinks to be a sub-exponential function of $\lambda$.

## 3. The Proof of the Key Lemma

In this section we will prove Lemma 2.2. The main step in the proof is the following geometric lemma, which is the most important new ingredient in the proof of Theorem 1.1. The proof of the geometric lemma is (to us, at least) surprisingly short and simple.

Recall that we fixed $n, r \in \mathbb{N}$, and defined $\beta = 3^{-4r}$ and $C = 4r^{3/2}$. In this section we will write $\langle \cdot,\cdot \rangle$ to denote the standard inner product on $\mathbb{R}^n$.

**Lemma 3.1.** *Let $U$ and $U'$ be i.i.d. random variables taking values in a finite set $X$, and let $\sigma_1,\ldots,\sigma_r \colon X \to \mathbb{R}^n$ be arbitrary functions. There exist $\lambda \geqslant -1$ and $i \in [r]$ such that*

$$\mathbb{P}\Big(\big\langle \sigma_i(U),\sigma_i(U')\big\rangle \geqslant \lambda \;\text{ and }\; \big\langle \sigma_j(U),\sigma_j(U')\big\rangle \geqslant -1 \;\text{ for all }\; j \neq i\Big) \geqslant \beta e^{-C\sqrt{\lambda+1}}.$$

The idea of the proof is as follows. We will first observe (see Lemma 3.2) that all of the moments of the inner products $\langle \sigma_i(U),\sigma_i(U')\rangle$ are positive. The key step is then to define a function $f\colon \mathbb{R}^r \to \mathbb{R}$ (see (5)) whose Taylor expansion has non-negative coefficients, that is only positive close to the positive quadrant, and that does not grow too fast. With this function in hand, the lemma will then follow from a simple calculation.

We begin with the following simple but key observation.

**Lemma 3.2.** *Let $U$ and $U'$ be i.i.d. random variables taking values in a finite set $X$, and let $\sigma_1,\ldots,\sigma_r \colon X \to \mathbb{R}^n$ be arbitrary functions. Then*

$$\mathbb{E}\Big[\big\langle \sigma_1(U),\sigma_1(U')\big\rangle^{\ell_1}\cdots\big\langle \sigma_r(U),\sigma_r(U')\big\rangle^{\ell_r}\Big] \geqslant 0$$

*for every $(\ell_1,\ldots,\ell_r) \in \mathbb{Z}^r$ with $\ell_1,\ldots,\ell_r \geqslant 0$.*

*Proof.* To simplify the notation, let us write

$$\big\langle \sigma_1(U),\sigma_1(U')\big\rangle^{\ell_1}\cdots\big\langle \sigma_r(U),\sigma_r(U')\big\rangle^{\ell_r} = \prod_{i=1}^{\ell}\big\langle \sigma_{a_i}(U),\sigma_{a_i}(U')\big\rangle$$

where $\ell = \ell_1+\cdots+\ell_r$ and $(a_1,\ldots,a_\ell)$ is such that $\big|\big\{i \in [\ell] : a_i = j\big\}\big| = \ell_j$ for each $j \in [r]$. Now set

$$Z = \sigma_{a_1}(U)\otimes\sigma_{a_2}(U)\otimes\cdots\otimes\sigma_{a_\ell}(U) \quad\text{and}\quad Z' = \sigma_{a_1}(U')\otimes\sigma_{a_2}(U')\otimes\cdots\otimes\sigma_{a_\ell}(U'),$$

and note that

$$\big\langle Z,Z'\big\rangle = \prod_{i=1}^{\ell}\big\langle \sigma_{a_i}(U),\sigma_{a_i}(U')\big\rangle,$$

since $\langle u_1\otimes\cdots\otimes u_r, v_1\otimes\cdots\otimes v_r\rangle = \langle u_1, v_1\rangle\cdots\langle v_r, u_r\rangle$ for any vectors $u_i, v_i \in \mathbb{R}^n$. Finally, note that $Z$ and $Z'$ are independent and identically distributed random vectors, and therefore

$$\mathbb{E}\big[\langle Z, Z'\rangle\big] = \mathbb{E}_{Z'}\big[\mathbb{E}_Z\big[\langle Z, Z'\rangle\big]\big] = \mathbb{E}_{Z'}\big[\big\langle \mathbb{E}[Z], Z'\big\rangle\big] = \big\langle \mathbb{E}[Z], \mathbb{E}[Z']\big\rangle \geqslant 0,$$

as required, where the final inequality holds because $\mathbb{E}[Z] = \mathbb{E}[Z']$. $\square$

Next, we will need the following special function:

$$f(x_1,\ldots,x_r) = \sum_{j=1}^{r} x_j \prod_{i\neq j}\big(2+\cosh\sqrt{x_i}\big), \tag{5}$$

where we define $\cosh\sqrt{x}$ via its Taylor expansion

$$\cosh\sqrt{x} = \sum_{n=0}^{\infty}\frac{x^n}{(2n)!}.$$

In particular, note that all of the coefficients of the Taylor expansion of $f$ are non-negative. The function $f$ also satisfies the following two inequalities.

**Lemma 3.3.** *Let $r \in \mathbb{N}$. The function $f\colon \mathbb{R}^r \to \mathbb{R}$ defined in* (5) *satisfies*

$$f(x_1,\ldots,x_r) \leqslant \begin{cases} 3^r r\exp\left(\displaystyle\sum_{i=1}^{r}\sqrt{x_i+3r}\right) & \text{if } \; x_i \geqslant -3r \; \text{ for all } \; i\in[r]; \\ -1 & \text{otherwise.} \end{cases}$$

*Proof.* Observe first that $x \leqslant 2+\cosh\sqrt{x} \leqslant 3e^{\sqrt{x}}$ for every $x \geqslant 0$, that

$$f(x_1,\ldots,x_r) = \left(\prod_{i=1}^{r}\big(2+\cosh\sqrt{x_i}\big)\right)\sum_{j=1}^{r}\frac{x_j}{2+\cosh\sqrt{x_j}},$$

and that $\cosh\sqrt{x} = \cos\sqrt{-x}$, and hence $-1 \leqslant \cosh\sqrt{x} \leqslant 1$ for all $x < 0$. It follows that

$$1 \leqslant 2+\cosh\sqrt{x} \leqslant 3e^{\sqrt{x+3r}} \qquad \text{and} \qquad \frac{x}{2+\cosh\sqrt{x}} \leqslant 1 \tag{6}$$

for all $x \geqslant -3r$, and hence

$$f(x_1,\ldots,x_r) \leqslant r\prod_{i=1}^{r}3e^{\sqrt{x_i+3r}} = 3^r r\exp\left(\sum_{i=1}^{r}\sqrt{x_i+3r}\right),$$

for every $x_1,\ldots,x_r \geqslant -3r$, as claimed. Moreover, if $x_i \leqslant -3r$, then, again using (6),

$$\sum_{j=1}^{r}\frac{x_j}{2+\cosh\sqrt{x_j}} \leqslant \frac{x_i}{3} + r - 1 \leqslant -1.$$

Since $2+\cosh\sqrt{x} \geqslant 1$ for every $x \in \mathbb{R}$, it follows that $f(x_1,\ldots,x_r) \leqslant -1$, as required. $\square$

*Remark* 3.4. The key idea in the lemma above was to find an entire function, $2 + \cosh\sqrt{x}$, which $(a)$ has a Taylor expansion at $x = 0$ with non-negative coefficients; $(b)$ is bounded on the negative real axis; and $(c)$ does not grow too quickly on the positive axis. The Phragmén–Lindelöf theorem implies that functions satisfying $(a)$ and $(b)$ must grow at least as fast as $\exp\big(x^{1/2+o(1)}\big)$ on the positive real axis. Thus the bound on the growth of $f$ given by the lemma is essentially best possible for constructions of this type.

We are now ready to prove our geometric lemma.

*Proof of Lemma 3.1.* For each $i \in [r]$, define $Z_i = 3r\big\langle \sigma_i(U), \sigma_i(U')\big\rangle$, and observe that, by Lemma 3.2 and linearity of expectation, we have

$$\mathbb{E}\big[f\big(Z_1,\dots,Z_r\big)\big] \geqslant 0, \tag{7}$$

since all of the coefficients of the Taylor expansion of $f$ are non-negative. By Lemma 3.3, it follows that if we define $E$ to be the event that $Z_i \geqslant -3r$ for every $i \in [r]$, then

$$3^r r \cdot \mathbb{E}\bigg[\exp\bigg(\sum_{i=1}^{r}\sqrt{Z_i + 3r}\bigg)\mathbf{1}_E\bigg] \geqslant 1 - \mathbb{P}(E), \tag{8}$$

where $\mathbf{1}_E \in \{0,1\}$ denotes the indicator of the event $E$, since the left-hand side bounds the contribution to the expectation on the left-hand side of (7) due to the event $E$, and the right-hand side bounds the contribution to the expectation due to the event $E^c$.

The result now follows from a simple calculation. First, note that

$$\mathbb{P}(E) = \mathbb{P}\Big(Z_i \geqslant -3r \ \text{ for all } \ i \in [r]\Big),$$

so if $\mathbb{P}(E) \geqslant \beta$, then the claimed inequality holds with $\lambda = -1$. We claim that if $\mathbb{P}(E) \leqslant \beta$, then there exists $\lambda \geqslant -1$ such that

$$\mathbb{P}\Big(\big\{M \geqslant \lambda\big\} \cap E\Big) \geqslant \beta r e^{-C\sqrt{\lambda+1}}, \tag{9}$$

where $M = \max\big\{\big\langle \sigma_i(U), \sigma_i(U')\big\rangle : i \in [r]\big\}$, and therefore (by the union bound) there exists an $i \in [r]$ as required. To see this, set $f(x) = e^{c\sqrt{x+1}}$ with $c = \sqrt{3}r^{3/2}$, and observe that

$$\mathbb{E}\big[f(M)\mathbf{1}_E\big] \;\leqslant\; \mathbb{P}(E) + \int_{-1}^{\infty} \mathbb{P}\Big(\big\{M \geqslant \lambda\big\} \cap E\Big) \cdot f'(\lambda)\,\mathrm{d}\lambda$$

since $f(-1) = 1$ and if $E$ holds then $M \geqslant -1$. It follows that if $\mathbb{P}(E) \leqslant \beta$ and (9) does not hold for any $\lambda \geqslant -1$, then

$$\begin{aligned}\mathbb{E}\Big[\exp\big(c\sqrt{M+1}\big)\mathbf{1}_E\Big] &\leqslant \mathbb{P}(E) + \int_{-1}^{\infty} \mathbb{P}\Big(\big\{M \geqslant \lambda\big\} \cap E\Big) \cdot \frac{c}{2\sqrt{\lambda+1}} \cdot e^{c\sqrt{\lambda+1}}\,\mathrm{d}\lambda \\ &\leqslant \beta + \beta r \int_{-1}^{\infty} \frac{c}{2\sqrt{\lambda+1}} \cdot e^{-\sqrt{\lambda+1}}\,\mathrm{d}\lambda \;\leqslant\; \beta(cr+1),\end{aligned}$$

where in the second step we used our choice of $C = 4r^{3/2} \geqslant c + 1$, and in the final step we used the fact that $\int_0^\infty \frac{1}{2\sqrt{x}} e^{-\sqrt{x}}\,\mathrm{d}x = 1$. In particular, recalling that $\beta = 3^{-4r}$, it follows that the left-hand side of (8) is at most $3^r r \cdot 3^{-4r}\big(\sqrt{3}r^{5/2} + 1\big) \leqslant 1/2$, which contradicts our assumption that $\mathbb{P}(E) \leqslant \beta$. Hence there exists $\lambda \geqslant -1$ such that (9) holds, as claimed. $\square$

It is now straightforward to deduce Lemma 2.2 from Lemma 3.1.

*Proof of Lemma 2.2.* For each colour $i \in [r]$, define a function $\sigma_i \colon X \to \mathbb{R}^{Y_i}$ as follows: for each $x \in X$, choose a set $N_i'(x) \subset N_i(x) \cap Y_i$ of size exactly $p_i|Y_i|$, where $p_i = p_i(X, Y_i)$, and set

$$\sigma_i(x) = \frac{\mathbb{1}_{N_i'(x)} - p_i \mathbb{1}_{Y_i}}{\sqrt{\alpha_i p_i |Y_i|}},$$

where $\mathbb{1}_S \in \{0,1\}^{Y_i}$ denotes the indicator function of the set $S$. Note that, for any $x, y \in X$,

$$\big\langle \sigma_i(x), \sigma_i(y) \big\rangle \geqslant \lambda \quad \Leftrightarrow \quad |N_i'(x) \cap N_i'(y)| \geqslant \big(p_i + \lambda \alpha_i\big) p_i |Y_i|.$$

Now, by Lemma 3.1, there exists $\lambda \geqslant -1$ and a colour $\ell \in [r]$ such that

$$\mathbb{P}\Big( \big\langle \sigma_\ell(U), \sigma_\ell(U') \big\rangle \geqslant \lambda \text{ and } \big\langle \sigma_i(U), \sigma_i(U') \big\rangle \geqslant -1 \text{ for all } i \neq \ell \Big) \geqslant \beta e^{-C\sqrt{\lambda+1}}.$$

where $U$, $U'$ are independent random variables distributed uniformly in the set $X$. Hence there exists a vertex $x \in X$ and a set $X' \subset X$ such that,

$$|X'| \geqslant \beta e^{-C\sqrt{\lambda+1}} |X| \qquad \text{and} \qquad |N_\ell'(x) \cap N_\ell'(y)| \geqslant \big(p_\ell + \lambda \alpha_\ell\big) p_\ell |Y_\ell|$$

for every $y \in X'$, and

$$|N_i'(x) \cap N_i'(y)| \geqslant \big(p_i - \alpha_i\big) p_i |Y_i|$$

for every $y \in X'$ and $i \in [r]$. Setting $Y_i' = N_i'(x)$ for each $i \in [r]$, it follows that

$$p_\ell\big(X', Y_\ell'\big) \geqslant p_\ell(X, Y_\ell) + \lambda \alpha_\ell \qquad \text{and} \qquad p_i\big(X', Y_i'\big) \geqslant p_i(X, Y_i) - \alpha_i$$

for every $i \in [r]$, as required. $\square$

## 4. Proof of the Book Theorem

In this section we will use the multicolour book algorithm to prove Theorem 2.1. To do so, we will first prove a few simple lemmas about the sets produced by the algorithm when applied with arbitrary inputs, and then apply these bounds to our setting. Recall that we are given an $r$-colouring $\chi$ of $E(K_n)$ and sets $X, Y_1, \ldots, Y_r \subset V(K_n)$. Define

$$p_0 = \min\big\{ p_i(X, Y_i) : i \in [r] \big\}, \tag{10}$$

and run the algorithm for some $t \in \mathbb{N}$, $\lambda_0 \geqslant 0$ and $\delta > 0$.

In order to simplify the statements of the lemmas below, let us write $X(s)$ and $Y_i(s)$ for the sets $X$ and $Y_i$ after $s$ steps of the algorithm, and set

$$p_i(s) = p_i\big(X(s), Y_i(s)\big) \qquad \text{and} \qquad \alpha_i(s) = \frac{p_i(s) - p_0 + \delta}{t}.$$

Let us also write $\ell(s) \in [r]$ and $\lambda(s) \geqslant -1$ for the colour and number given by the application of Lemma 2.2 to the sets $X(s)$ and $Y_1(s), \ldots, Y_r(s)$, and define

$$\mathcal{B}_i(s) = \big\{ 0 \leqslant j < s \,:\, \ell(j) = i \text{ and } \lambda(j) > \lambda_0 \big\},$$

for each $i \in [r]$ and $s \in \mathbb{N}$, to be the set of density-boost steps in colour $i$ during the first $s$ steps of the algorithm. We begin by noting the following lower bound on $p_i(s)$.

**Lemma 4.1.** *For each $i \in [r]$ and $s \in \mathbb{N}$,*

$$p_i(s) - p_0 + \delta \;\geqslant\; \delta \cdot \left(1 - \frac{1}{t}\right)^t \prod_{j \in \mathcal{B}_i(s)} \left(1 + \frac{\lambda(j)}{t}\right). \tag{11}$$

*Proof.* Note first that if $Y_i(s+1) = Y_i(s)$, then $p_i(s+1) \geqslant p_i(s)$, since the minimum degree does not decrease when we take a subset of $X(s)$. When we perform a colour step in colour $i$, we have $p_i(s+1) \geqslant p_i(s) - \alpha_i(s)$, by Lemma 2.2, and hence

$$p_i(s+1) - p_0 + \delta \geqslant \left(1 - \frac{1}{t}\right)\big(p_i(s) - p_0 + \delta\big),$$

by our choice of $\alpha_i(s)$. Similarly, when we perform a density-boost step in colour $i$ we have $p_i(s+1) \geqslant p_i(s) + \lambda(s)\alpha_i(s)$, by Lemma 2.2, and hence

$$p_i(s+1) - p_0 + \delta \geqslant \left(1 + \frac{\lambda(s)}{t}\right)\big(p_i(s) - p_0 + \delta\big).$$

Recalling that there are at most $t$ colour steps in colour $i$, and that $p_i(0) \geqslant p_0$, by the definition (10) of $p_0$, the claimed bound follows. $\square$

Before continuing, let's note a couple of important consequences of Lemma 4.1. First, it implies that neither $p_i(s)$ nor $\alpha_i(s)$ can get too small.

**Lemma 4.2.** *If $t \geqslant 2$, then*

$$p_i(s) \;\geqslant\; p_0 - \frac{3\delta}{4} \qquad \text{and} \qquad \alpha_i(s) \;\geqslant\; \frac{\delta}{4t}$$

*for every $i \in [r]$ and $s \in \mathbb{N}$.*

*Proof.* Both bounds follow immediately from (11) and the definition of $\alpha_i(s)$, since we chose $\lambda_0 \geqslant 0$ and using the fact that $(1 - 1/t)^t \geqslant 1/4$ for every $t \geqslant 2$. $\square$

It also implies the following bound on the number of density-boost steps.

**Lemma 4.3.** *If $t \geqslant \lambda_0 \geqslant 2$ and $\delta \leqslant 1/4$, then*

$$|\mathcal{B}_i(s)| \leqslant \frac{4\log(1/\delta)}{\lambda_0} \cdot t$$

*for every $i \in [r]$ and $s \in \mathbb{N}$.*

*Proof.* Since $\lambda(j) > \lambda_0$ for every $j \in \mathcal{B}_i(s)$, and $p_i(s) \leqslant 1$, it follows from (11) that

$$\frac{\delta}{4}\left(1 + \frac{\lambda_0}{t}\right)^{|\mathcal{B}_i(s)|} \leqslant 1 + \delta.$$

Since $t \geqslant \lambda_0$ and $\delta \leqslant 1/4$, the claimed bound follows. $\square$

Lemmas 4.2 and 4.3 together provide a lower bound on the size of the set $Y_i(s)$.

**Lemma 4.4.** *If $t \geqslant 2$, then*

$$|Y_i(s)| \geqslant \left(p_0 - \frac{3\delta}{4}\right)^{t+|\mathcal{B}_i(s)|} |Y_i(0)|$$

*for every $i \in [r]$ and $s \in \mathbb{N}$.*

*Proof.* Note that $Y_i(j+1) \neq Y_i(j)$ for at most $t + |\mathcal{B}_i(s)|$ of the first $s$ steps, and for those steps we have

$$|Y_i(j+1)| = p_i(j)|Y_i(j)| \geqslant \left(p_0 - \frac{3\delta}{4}\right)|Y_i(j)|,$$

by (3) and Lemma 4.2. $\square$

Finally, we need to bound the size of the set $X(s)$. To do so, set $\varepsilon = (\beta/r)e^{-C\sqrt{\lambda_0+1}}$, and define $\mathcal{B}(s) = \mathcal{B}_1(s) \cup \cdots \cup \mathcal{B}_r(s)$ to be the set of all density-boost steps.

**Lemma 4.5.** *For each $s \in \mathbb{N}$,*

$$|X(s)| \geqslant \varepsilon^{rt+|\mathcal{B}(s)|} \exp\bigg( - C \sum_{j \in \mathcal{B}(s)} \sqrt{\lambda(j)+1} \bigg)|X(0)| - rt. \tag{12}$$

*Proof.* If $\lambda(j) \leqslant \lambda_0$, then by (2) and Step 2 of the algorithm we have

$$|X(j+1)| \geqslant \frac{\beta e^{-C\sqrt{\lambda_0+1}}}{r} \cdot |X(j)| - 1 = \varepsilon|X(j)| - 1.$$

On the other hand, if $\lambda(j) > \lambda_0$, then $j \in \mathcal{B}(s)$, and we have

$$|X(j+1)| \geqslant \beta e^{-C\sqrt{\lambda(j)+1}}|X(j)|,$$

by (2) and Step 3 of the algorithm. Since there are at most $rt$ colour steps, and $\beta \geqslant \varepsilon$, the claimed bound follows. $\square$

We will use the following lemma to bound the right-hand side of (12).

**Lemma 4.6.** *If $t \geqslant \lambda_0/\delta > 0$ and $\delta \leqslant 1/4$, then*

$$\sum_{j \in \mathcal{B}(s)} \sqrt{\lambda(j)} \leqslant \frac{7r\log(1/\delta)}{\sqrt{\lambda_0}} \cdot t$$

*for every $s \in \mathbb{N}$.*

*Proof.* Observe first that, by Lemma 4.1, we have

$$\frac{\delta}{4} \prod_{j \in \mathcal{B}_i(s)} \left(1 + \frac{\lambda(j)}{t}\right) \leqslant 1 + \delta \tag{13}$$

for each $i \in [r]$. Note also that $\lambda_0 \leqslant \lambda(j) \leqslant 5t/\delta$ for every $j \in \mathcal{B}(s)$, where the lower bound holds by the definition of $\mathcal{B}(s)$, and the upper bound by (13) and since $\delta \leqslant 1/4$. Now, since $\log(1+x) \geqslant \min\{x/2, 1\}$ for all $x > 0$, it follows that

$$\frac{\sqrt{\lambda(j)}}{\log(1+\lambda(j)/t)} \leqslant \max\left\{\frac{2t}{\sqrt{\lambda_0}}, \sqrt{\frac{5t}{\delta}}\right\} \leqslant \frac{3t}{\sqrt{\lambda_0}}, \tag{14}$$

where in the second inequality we used our assumption that $t \geqslant \lambda_0/\delta$. It now follows that

$$\sum_{j \in \mathcal{B}_i(s)} \sqrt{\lambda(j)} \leqslant \frac{3t}{\sqrt{\lambda_0}} \sum_{j \in \mathcal{B}_i(s)} \log\left(1 + \frac{\lambda(j)}{t}\right) \leqslant \frac{7\log(1/\delta)}{\sqrt{\lambda_0}} \cdot t,$$

where the first inequality holds by (14), and the second follows from (13), since $\delta \leqslant 1/4$. Summing over $i \in [r]$, we obtain the claimed bound. $\square$

We are finally ready to prove the book theorem.

*Proof of Theorem 2.1.* Recall that we are given an $r$-colouring $\chi$ of $E(K_n)$ and a collection of sets $X, Y_1, \ldots, Y_r \subset V(K_n)$ with

$$|X| \geqslant \left(\frac{\mu^2}{p}\right)^{\mu r t} \qquad \text{and} \qquad |Y_i| \geqslant \left(\frac{e^{2^{13} r^3/\mu^2}}{p}\right)^{t} m \tag{15}$$

for each $i \in [r]$, for some $p \in (0,1]$, $\mu \geqslant 2^{10} r^3$ and $t, m \in \mathbb{N}$ with $t \geqslant \mu^5/p^2$, and moreover

$$|N_i(x) \cap Y_i| \geqslant p|Y_i| \tag{16}$$

for every $x \in X$ and $i \in [r]$. We will run the multicolour book algorithm with

$$\delta = \frac{p}{\mu^2} \qquad \text{and} \qquad \lambda_0 = \left(\frac{\mu \log(1/\delta)}{8C}\right)^2, \tag{17}$$

where $C = 4r^{3/2}$ (as before), and show that it ends with

$$\max\big\{|T_i(s)| : i \in [r]\big\} = t \qquad \text{and} \qquad \min\big\{|Y_i(s)| : i \in [r]\big\} \geqslant m,$$

and therefore produces a monochromatic $(t,m)$-book.

To do so, we just need to bound the sizes of the sets $X(s)$ and $Y_i(s)$ from below. Observe that $t \geqslant \lambda_0$ and $\delta \leqslant 1/4$, and therefore, by Lemma 4.3, that

$$|\mathcal{B}_i(s)| \,\leqslant\, \frac{4\log(1/\delta)}{\lambda_0} \cdot t \,=\, \frac{2^{12} r^3}{\mu^2 \log(1/\delta)} \cdot t \,\leqslant\, t \tag{18}$$

for every $i \in [r]$ and $s \in \mathbb{N}$. Since $p_0 \geqslant p$, by (16) and the definition of $p_0$, it follows that

$$|Y_i(s)| \,\geqslant\, \left(p - \frac{3\delta}{4}\right)^{t + |\mathcal{B}_i(s)|} |Y_i| \,\geqslant\, e^{-2\delta t/p}\, p^{|\mathcal{B}_i(s)|} \left(e^{2^{13} r^3/\mu^2}\right)^t m,$$

where the first inequality holds by Lemma 4.4, and the second by (15) and (18). Noting that $p^{|\mathcal{B}_i(s)|} \geqslant e^{-2^{12} r^3 t/\mu^2}$, by (18), and that $\delta/p = 1/\mu^2$, it follows that

$$|Y_i(s)| \,\geqslant\, e^{-2t/\mu^2} \left(e^{2^{12} r^3/\mu^2}\right)^t m \,\geqslant\, m,$$

as claimed. To bound $|X(s)|$, recall (12) and observe that

$$\varepsilon^{rt + |\mathcal{B}(s)|} \geqslant \left(\frac{\beta}{r} \cdot e^{-C\sqrt{\lambda_0 + 1}}\right)^{2rt} \geqslant \left(e^{-4C\sqrt{\lambda_0}}\right)^{rt} = \delta^{\mu r t/2} = \left(\frac{\mu^2}{p}\right)^{-\mu r t/2},$$

where the first step holds because $\varepsilon = (\beta/r)e^{-C\sqrt{\lambda_0+1}}$ and $|\mathcal{B}(s)| \leqslant rt$, by (18), the second step holds because $\beta = 3^{-4r}$ and $\lambda_0 \geqslant 2^{10}r^3$, and the third and fourth steps follow from our choice of $\lambda_0$ and $\delta$ in (17). Moreover, observe that $t \geqslant \lambda_0/\delta$, and therefore

$$\sum_{j\in\mathcal{B}(s)} \sqrt{\lambda(j)+1} \leqslant \frac{8r\log(1/\delta)}{\sqrt{\lambda_0}} \cdot t = \frac{2^6 Crt}{\mu},$$

by Lemma 4.6 and our choice of $\lambda_0$, and therefore

$$|X(s)| \geqslant \left(\frac{\mu^2}{p}\right)^{\mu rt/2} \exp\left(-\frac{2^6 C^2 rt}{\mu}\right) - rt > 0,$$

for every $s \in \mathbb{N}$, by Lemma 4.5 and since $\mu \geqslant 2^{10}r^3$ and $C = 4r^{3/2}$. It follows that the algorithm produces a monochromatic $(t,m)$-book, as claimed. □

## 5. Deducing the bound on $R_r(k)$

In this section we will deduce Theorem 1.1, our bound on the Ramsey numbers $R_r(k)$, from Theorem 2.1. We will in fact prove the following quantitative version of the theorem.

**Theorem 5.1.** *Let $r \geqslant 2$, and set $\delta = 2^{-160}r^{-12}$. Then*

$$R_r(k) \leqslant e^{-\delta k} r^{rk}$$

*for every $k \in \mathbb{N}$ with $k \geqslant 2^{200}r^{20}$.*

In order to apply Theorem 2.1, we need to find a suitable collection of sets $X, Y_1, \ldots, Y_r$ in an arbitrary $r$-colouring of $E(K_n)$. To do so, we simply run the Erdős–Szekeres process until we find a subset of the vertex set in which each colour has density close to $1/r$, and the colouring is close to regular in each colour.

**Lemma 5.2.** *Given $n, r \in \mathbb{N}$ and $\varepsilon > 0$, and an $r$-colouring of $E(K_n)$, the following holds. There exist disjoint sets of vertices $S_1, \ldots, S_r$ and $W$ such that, for every $i \in [r]$,*

$$|W| \geqslant \left(\frac{1+\varepsilon}{r}\right)^{\sum_{j=1}^r |S_j|} n \qquad \text{and} \qquad |N_i(w) \cap W| \geqslant \left(\frac{1}{r} - \varepsilon\right)|W| - 1$$

*for every $w \in W$, and $(S_i, W)$ is a monochromatic book in colour $i$.*

*Proof.* We use induction on $n$. Note that if $n \leqslant r$ or $r = 1$, then the sets $S_1 = \cdots = S_r = \emptyset$ and $W = V(K_n)$ have the required properties. We may therefore assume that $n > r \geqslant 2$, and that there exists a vertex $x \in V(K_n)$ and a colour $\ell \in [r]$ such that

$$|N_\ell(x)| < \left(\frac{1}{r} - \varepsilon\right)n - 1,$$

and hence there exists a different colour $j \neq \ell$ such that

$$|N_j(x)| \geqslant \frac{1}{r-1}\left(n - \left(\frac{1}{r} - \varepsilon\right)n\right) \geqslant \left(\frac{1+\varepsilon}{r}\right)n.$$

Applying the induction hypothesis to the colouring induced by the set $N_j(x)$, we obtain sets $S_1', \dots, S_r'$ and $W'$ satisfying the conclusion of the lemma with $n$ replaced by $|N_j(x)|$. Setting

$$W = W', \qquad S_j = S_j' \cup \{x\} \qquad \text{and} \qquad S_i = S_i' \qquad \text{for each } \; i \neq j,$$

we see that $W$ satisfies the minimum degree condition (by the induction hypothesis), and

$$|W| \geqslant \left( \frac{1+\varepsilon}{r} \right)^{\sum_{i=1}^r |S_i'|} |N_j(x)| \geqslant \left( \frac{1+\varepsilon}{r} \right)^{\sum_{i=1}^r |S_i|} n.$$

Since $(S_i, W)$ is a monochromatic book in colour $i$ for each $i \in [r]$, it follows that $S_1, \dots, S_r$ and $W$ are sets with the claimed properties. $\square$

The next lemma, which is an immediate consequence of Erdős and Szekeres' bound on the $r$-colour Ramsey numbers, implies (roughly speaking) that either the sets $S_1, \dots, S_r$ given by Lemma 5.2 satisfy $|S_i| = o(k)$ for each $i \in [r]$, or we are already done.

**Lemma 5.3.** *If $k, r \geqslant 2$ and $\varepsilon \in (0,1)$, then*

$$R_r\big(k - s_1, \dots, k - s_r\big) \leqslant e^{-\varepsilon^3 k/2} \left( \frac{1+\varepsilon}{r} \right)^s r^{rk}$$

*for every $s_1, \dots, s_r \in [k]$ with $s = s_1 + \dots + s_r \geqslant \varepsilon^2 k$.*

*Proof.* By the Erdős–Szekeres bound, we have

$$R_r\big(k - s_1, \dots, k - s_r\big) \leqslant r^{rk - s}.$$

The claimed bound follows, since $(1+\varepsilon)^{\varepsilon^2 k} \geqslant e^{\varepsilon^3 k/2}$ for every $\varepsilon \in (0,1)$. $\square$

Our bound on $R_r(k)$ now follows from a simple calculation.

*Proof of Theorem 5.1.* Given $r \geqslant 2$, set $\delta = 2^{-160} r^{-12}$, let $k \geqslant 2^{200} r^{20}$ and $n \geqslant e^{-\delta k} r^{rk}$, and let $\chi$ be an arbitrary $r$-colouring of $E(K_n)$. Set $\varepsilon = 2^{-50} r^{-4}$, and let $S_1, \dots, S_r$ and $W$ be the sets given by Lemma 5.2. Suppose first that $|S_1| + \dots + |S_r| \geqslant \varepsilon^2 k$, and observe that, by Lemma 5.3 and our lower bounds on $n$ and $|W|$, we have

$$|W| \geqslant \left( \frac{1+\varepsilon}{r} \right)^{\sum_{j=1}^r |S_j|} e^{-\delta k} r^{rk} \geqslant R_r\big(k - |S_1|, \dots, k - |S_r|\big).$$

Since $(S_i, W)$ is a monochromatic book in colour $i$ for each $i \in [r]$, it follows that there exists a monochromatic copy of $K_k$ in $\chi$, as required.

We may therefore assume from now on that $|S_1| + \dots + |S_r| \leqslant \varepsilon^2 k$. In this case we set $X = Y_1 = \dots = Y_r = W$, and apply Theorem 2.1 to the colouring restricted to $W$ with

$$p = \frac{1}{r} - 2\varepsilon, \qquad \mu = 2^{30} r^3, \qquad t = 2^{-40} r^{-3} k \qquad \text{and} \qquad m = R\big(k, \dots, k, k - t\big).$$

In order to apply Theorem 2.1, we need to check that $t \geqslant \mu^5 / p^2$, and that

$$|X| \geqslant \left( \frac{\mu^2}{p} \right)^{\mu r t} \qquad \text{and} \qquad |Y_i| \geqslant \left( \frac{e^{2^{13} r^3 / \mu^2}}{p} \right)^t m.$$

The first two of these inequalities are both easy: the first holds because $k \geqslant 2^{200} r^{20}$, and for the second note that since $n \geqslant r^{rk/2}$ and $\sum_{j=1}^{r} |S_j| \leqslant rk/4$, we have

$$|X| \geqslant \left(\frac{1+\varepsilon}{r}\right)^{rk/4} r^{rk/2} \geqslant r^{rk/4} \geqslant \left(2^{61} r^7\right)^{2^{-10} rk} \geqslant \left(\frac{\mu^2}{p}\right)^{\mu r t},$$

as required. The third inequality is more delicate: observe first[3] that

$$R\big(k, \ldots, k, k-t\big) \leqslant \binom{rk-t}{k, \ldots, k, k-t} \leqslant e^{-(r-1)t^2/3rk} r^{rk-t} \leqslant e^{-t^2/6k} r^{rk-t}.$$

and therefore, since $\delta \leqslant 2^{-10} t^2/k^2$ and $p = 1/r - 2\varepsilon \geqslant e^{-3\varepsilon r}/r$, we have

$$n \geqslant e^{-\delta k} r^{rk} \geqslant r^t e^{t^2/8k} \cdot R\big(k, \ldots, k, k-t\big) \geqslant \frac{m}{p^t} \cdot \exp\left(\frac{t^2}{8k} - 3\varepsilon r t\right).$$

Moreover, by our choice of $t$ and $\mu$, we have

$$\frac{t}{8k} = \frac{1}{2^{43} r^3} \geqslant \frac{1}{2^{47} r^3} + \frac{1}{2^{48} r^3} = \frac{2^{13} r^3}{\mu^2} + 4\varepsilon r.$$

Finally, since $\sum_{j=1}^{r} |S_j| \leqslant \varepsilon^2 k \leqslant \varepsilon t$, it follows that

$$|Y_i| = |W| \geqslant \left(\frac{1+\varepsilon}{r}\right)^{\varepsilon t} n \geqslant \frac{m}{p^t} \cdot \exp\left(\frac{t^2}{8k} - 4\varepsilon r t\right) \geqslant \left(\frac{e^{2^{13} r^3/\mu^2}}{p}\right)^t m,$$

as claimed. Thus $\chi$ contains a monochromatic $(t, m)$-book, and hence, by our choice of $m$, $\chi$ contains a monochromatic copy of $K_k$, as required. □


## Acknowledgements

Important steps in the work described in this paper were carried out during visits by various subsets of the authors to the University of Cambridge, IMPA, PUC-Rio, the IAS, and Sapienza Università di Roma. We thank each of these institutions for providing a wonderful working environment, and also Yuval Wigderson for pointing us to the paper [1].


## Appendix A. Bounding binomial coefficients

In this short appendix we will prove the following bound, which was used in Section 5.

**Lemma A.1.** *For each $k, t, r \in \mathbb{N}$ with $3 \leqslant t \leqslant k$,*

$$\binom{rk-t}{k, \ldots, k, k-t} \leqslant e^{-(r-1)t^2/3rk} r^{rk-t}.$$

*Proof.* To prove this, observe that

$$\binom{rk-t}{k, \ldots, k, k-t} \binom{rk}{k, \ldots, k}^{-1} = \prod_{i=0}^{t-1} \frac{k-i}{rk-i} = r^{-t} \prod_{i=0}^{t-1} \left(1 - \frac{(r-1)i}{rk-i}\right) \leqslant r^{-t} \cdot e^{-(r-1)t^2/3rk}.$$

Since $\binom{rk}{k, \ldots, k} \leqslant r^{rk}$, the claimed inequality follows. □

[3] See Appendix A for a proof of the second inequality.

## References


[1] H.L. Abbott, A note on Ramsey's theorem, *Canad. Math. Bull.*, **15** (1972), 9–10.

[2] M. Campos, S. Griffiths, R. Morris and J. Sahasrabudhe, An exponential improvement for diagonal Ramsey, *Ann. Math.*, to appear.

[3] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. Math.*, **170** (2009), 941–960.

[4] D. Conlon, Upper bounds for Ramsey numbers, PhD thesis, University of Cambridge, 2009.

[5] D. Conlon and A. Ferber, Lower bounds for multicolour Ramsey numbers, *Adv. Math.*, **378** (2021), Paper No. 107528, 5 pp.

[6] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.*, **53** (1947), 292–294,

[7] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compos. Math.*, **2** (1935), 463–470.

[8] R.L. Graham and V. Rödl, Numbers in Ramsey theory, Surveys in Combinatorics, London Math. Soc. Lecture Note Series, **123** (1987), 111-153.

[9] P. Gupta, N. Ndiaye, S. Norin and L. Wei, Optimizing the CGMS upper bound on Ramsey numbers, arXiv:2407.19026

[10] F.P. Ramsey, On a Problem of Formal Logic, *Proc. London Math. Soc.*, **30** (1930), 264–286.

[11] A. Sah, Diagonal Ramsey via effective quasirandomness, *Duke Math. J.*, **172** (2023), 545–567.

[12] W. Sawin, An improved lower bound for multicolour Ramsey numbers and a problem of Erdős, *J. Combin. Theory Ser. A*, **188** (2022), Paper No. 105579, 11 pp.

[13] J. Spencer, Asymptotic lower bounds for Ramsey functions, *Discrete Math.*, **20** (1977), 69–76.

[14] A. Thomason, An upper bound for some Ramsey numbers, *J. Graph Theory*, **12** (1988), 509–517.

[15] Y. Wigderson, An improved lower bound on multicolour Ramsey numbers, *Proc. Amer. Math. Soc.*, **149** (2021), 2371–2374.



Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

*Email address*: `Paul.Balister@maths.ox.ac.uk`

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge, CB3 0WA, UK, and Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA

*Email address*: `bb12@cam.ac.uk`

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brasil

*Email address*: `marcelo.campos@impa.br`

Departamento de Matemática, PUC-Rio, Rua Marquês de São Vicente 225, Gávea, 22451-900 Rio de Janeiro, Brasil

*Email address*: `simon@puc-rio.br`

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK

*Email address*: `hurley@maths.ox.ac.uk`

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, 22460-320, Brasil

*Email address*: `rob@impa.br`

Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge, CB3 0WA, UK

*Email address*: `jdrs2@cam.ac.uk`

Institute of Mathematics, EPFL SB MATH, MA A2 383 (Bâtiment MA), Station 8, CH-1015 Lausanne, Switzerland
*Email address*: marius.tiba@epfl.ch