On the chromatic number of a random hypergraph

Martin Dyer*   Alan Frieze [†]   Catherine Greenhill[‡]

December 22, 2014

Abstract

We consider the problem of k-colouring a random r-uniform hypergraph with n vertices and cn edges, where k, r, c remain constant as $n \to \infty$. Achlioptas and Naor showed that the chromatic number of a random graph in this setting, the case $r = 2$, must have one of two easily computable values as $n \to \infty$. We give a complete generalisation of this result to random uniform hypergraphs.

1 Introduction

We study the problem of k-colouring a random r-uniform hypergraph with n vertices and cn edges, where k, r and c are considered to be constant as $n \to \infty$. We generalise a theorem of Achlioptas and Naor [4] for k-colouring a random graph (2-uniform hypergraph) on n vertices.

Their theorem specifies the two possible values for the chromatic number of the random graph as $n \to \infty$. We give a complete generalisation of the result of [4]. We broadly follow the approach of Achlioptas and Naor [4], although they rely on simplifications which are available only in the case $r = 2$. We show that these simplifications can be replaced by more general techniques, valid for all k, $r \geq 2$ except $k = r = 2$.

There is an extensive literature on this problem in the case $r = 2$, colouring random graphs. In the setting we consider here, this culminates with the results of Achlioptas and Naor [4], though these do not give a complete answer to the problem. Our results here include those of [4].

There is also a literature for the case $k = 2$, random hypergraph 2-colouring. Achlioptas, Kim, Krivelevich and Tetali [2] gave a constructive approach, but their results were substantially improved by Achlioptas and Moore [3], using non-constructive methods. The results of [3] are asymptotic in r. Our results here include those of [3], but we also give a non-asymptotic treatment. Recently, Coja-Oghlan and Zdeborová [8] have given a small qualitative improvement of the result of [3], which goes beyond what can be proved here. See these papers, and their references, for further information.

Finally, we note that Krivelevich and Sudakov [14] studied a wide range of random hypergraph colouring problems, and some of their results were recently improved by Kupavskii and Shabanov [15]. But, in the setting of this paper, these results are much less precise than those we establish here.

Remark 1.1. *After preparing this paper, we learnt of related work by Coja-Oghlan and his coauthors in the case $r = 2$. Coja-Oghlan and Vilenchik [7] improved the upper bound on the k-colourability threshold, restricting the sharp threshold for k-colourability to an interval of constant width, compared with logarithmic width in [4]. (See also Remark 3.7 below.) A small improvement in the lower bound was obtained by Coja-Oghlan [5]. Additionally, Coja-Oghlan, Efthymiou and Hetterich [6] adapted the methods from [7] to study k-colourability of random regular graphs.*

1.1 Hypergraphs

Let $[n] = \{1, 2, \ldots, n\}$. Unless otherwise stated, the asymptotic results in this paper are as $n \to \infty$. Consider the set $\Omega(n, r, m)$ of r-uniform hypergraphs on the vertex set $[n]$ with m edges. Such a hypergraph is defined by its edge set $\mathcal{E}$, which consists of m distinct r-subsets of n. Let $N = \binom{n}{r}$ denote the total number of r-subsets.

Now let $\mathcal{G}(n, r, m)$ denote the uniform model of a random r-regular hypergraph with m edges. So $\mathcal{G}(n, r, m)$ consists of the set $\Omega(n, r, m)$ equipped with the uniform probability distribution. We write $G \in \mathcal{G}(n, r, m)$ for a random hypergraph chosen uniformly from $\Omega(n, r, m)$. The edge set $\mathcal{E}$ of this random hypergraph may be viewed as a sample of size m chosen uniformly, without replacement, from the set of N possible edges.

Although our main focus is the uniform model $\mathcal{G}$, it is simpler for many calculations to work with an alternative model. Let $\Omega^*(n, r, m)$ denote the set of all r-uniform multi-hypergraphs on $[n]$, defined as follows: each element of $\Omega^*(n, r, m)$ consists of vertex set $[n]$ and a multiset of edges, where each edge is now a multiset of r vertices (not necessarily distinct). We can generate a random element of G of $\Omega^*(n, r, m)$ using the following simple procedure: choose $\boldsymbol{v} = (v_1, v_2, \ldots, v_{rm}) \in [n]^{rm}$ uniformly at random and let the edge multiset of G be $\{e_1, \ldots, e_m\}$, where $e_i = \{v_{r(i-1)+1}, \ldots, v_{ri}\}$ for $i \in [m]$. Let $\mathcal{G}^*(n, r, m)$ denote the probability space on $\Omega^*(n, r, m)$ which arises from this procedure, and write $G \in \mathcal{G}^*(n, r, m)$ for a hypergraph G generated in this fashion.

Observe that an element $G \in \Omega^*(n, r, m)$ may not satisfy the definition of r-uniform hypergraph given above, for two reasons. First, an edge of G may contain repeated vertices, which $\Omega(n, r, m)$ does not permit. We call such an edge *defective*. Second, an edge of $G \in \Omega^*(n, r, m)$ may be identical to some other edge, which again $\Omega(n, r, m)$ does not permit. We call such an edge a *duplicate*.

Say an edge is *bad* if it is a defective or duplicate edge. Note that $\mathcal{G}^*(n, r, m)$ is not the uniform probability space over $\Omega^*(n, r, m)$, but that all G without bad edges are equiprobable. Thus $\mathcal{G}^*(n, r, m)$, conditional on there being no bad edges, is identical to $\mathcal{G}(n, r, m)$.

If E_n is a sequence of events, we say that E_n occurs "asymptotically almost surely" (a.a.s.) if $\Pr(E_n) \to 1$ as $n \to \infty$. In this paper, the event E_n usually concerns $G \in \mathcal{G}^*(n, r, m)$, where $m(n) = \lfloor cn \rfloor$, for some constant c. The difference between cn and $\lfloor cn \rfloor$ is usually negligible, and we follow [4] in disregarding it unless the distinction is important. Thus we will write the model simply as $G \in \mathcal{G}^*(n, r, cn)$, and similarly for the other models we consider.

Lemma 1.1. *Let c be a positive constant. For $G \in \mathcal{G}^*(n, r, cn)$,*

$$\Pr\left(G \text{ has no bad edge}\right) \sim \begin{cases} e^{-c(c+1)} & \text{if } r = 2, \\ e^{-cr(r-1)/2} & \text{if } r > 2. \end{cases}$$

Furthermore, for $G \in \mathcal{G}^*(n, r, cn)$, a.a.s. G has at most $2 \ln n$ bad edges.

Proof. Throughout this proof, all probabilities are calculated in $\mathcal{G}^*(n, r, cn)$. For any edge $e \in \mathcal{E}$,

$$\Pr(e \text{ is defective}) = 1 - \frac{n(n-1)\cdots(n-r+1)}{n^r} = 1 - \exp\left(-\frac{r(r-1)}{2n} + O\left(\frac{1}{n^2}\right)\right) \sim \frac{r(r-1)}{2n}.$$

Since this is true independently for each $e \in \mathcal{E}$, we have

$$\Pr(\text{no defective edge}) \sim \exp\left(-\frac{r(r-1)m}{2n}\right) \sim e^{-cr(r-1)/2} \tag{1}$$

as $m \sim cn$. Next note that, conditional on there being no defective edges, $\mathcal{E}$ is a uniform sample of size m chosen, with replacement, from the N possible r-subsets of $[n]$. Thus

$$\Pr\left(\text{no duplicate edge} \mid \text{no defective edge}\right) = \frac{N(N-1)\cdots(N-m+1)}{N^m}$$

$$\sim \exp\left(-\frac{m(m-1)}{2N}\right) \sim \begin{cases} e^{-c^2} & \text{if } r = 2, \\ 1 & \text{if } r > 2. \end{cases} \tag{2}$$

Combining (1) and (2) proves the first statement.

Now let m_{def} (m_{dup}, m_{bad}, respectively) denote the number of defective edges (duplicate edges, bad edges, respectively) in $G \in \mathcal{G}^*(n, r, m)$ (counting multiplicities). For the second statement, note that m_{def} has distribution $\text{Bin}(m, p_{\text{def}})$, and so $\mathbf{E}[m_{\text{def}}] \sim cr(r-1)/2$ as $n \to \infty$. Hence Chernoff's bound [13, Corollary 2.4] gives, for large enough n,

$$\Pr(m_{\text{def}} \geq \ln n) \leq e^{-\ln n} = 1/n. \tag{3}$$

Therefore a.a.s. $G \in \mathcal{G}^*(n, r, cn)$ has at most $\ln n$ defective edges. Next, note that each edge in $\mathcal{E}$ has at most $(m-1)/n^r$ duplicates in expectation, and so

$$\mathbf{E}[m_{\text{dup}}] \leq \frac{m(m-1)}{n^r} \leq c^2$$

for large n. (Indeed, if $r > 2$ then $\mathbf{E}[m_{\text{dup}}] = o(1)$, but we do not exploit this.) Thus, using Markov's inequality [13, (1.3)],

$$\Pr(m_{\text{dup}} \geq \ln n) \leq \frac{c^2}{\ln n}. \tag{4}$$

Combining this with (3) proves the second statement, since $m_{\text{bad}} \leq m_{\text{def}} + m_{\text{dup}}$. $\qquad\square$

As already stated, conditional on there being no bad edges, $\mathcal{G}^*(n, r, cn)$ is identical to $\mathcal{G}(n, r, cn)$. By the first statement of Lemma 1.1, G has no bad edges with probability $\Omega(1)$ as $n \to \infty$. This implies that any event occurring a.a.s. in $\mathcal{G}^*(n, r, cn)$ occurs a.a.s. in $\mathcal{G}(n, r, cn)$. In Lemma 1.4 we use the second statement of Lemma 1.1 to show that $\mathcal{G}(n, r, cn)$ and $\mathcal{G}^*(n, r, cn)$ are essentially equivalent, for our purposes.

We also make use of the following simple property of $\mathcal{G}^*(n, r, m)$. A vertex $i \in [n]$ of $G \in \Omega^*(n, r, m)$ is *isolated* if it appears in no edge. Note that a vertex is isolated if and only if it is absent from the vector $\boldsymbol{v} \in [n]^{rm}$ defined above. The following simply restates this property.

3

Observation 1.1. *For $S \subseteq [n]$, let $\mathcal{I}_S$ be the event that all vertices in S are isolated in $G \in \mathcal{G}^*(n,r,m)$. Let G' be G conditional on $\mathcal{I}_S$ and let G'' be obtained from G' by deleting all vertices in S and relabelling the remaining vertices by $[n - |S|]$, respecting the original ordering. Then $G'' \in \mathcal{G}^*(n - |S|, r, m)$.*

We show, in the proof of Lemma 4.1, that $G \in \mathcal{G}^*(n,r,cn)$ has $\Omega(n)$ isolated vertices a.a.s. and hence G has many disconnected components.

A further model of random hypergraphs is often used, which we will denote by $\widehat{\mathcal{G}}(n,r,p)$. In this, the edge set $\mathcal{E}$ of G is chosen by Bernoulli sampling. Each of the N possible r-subsets of $[n]$ is included in $\mathcal{E}$ independently with probability p. Essentially, this is $\mathcal{G}(n,r,m)$ where m is a binomial random variable $\mathrm{Bin}(N,p)$. We show in Section 1.2 below that $\widehat{\mathcal{G}}(n,r,cn/N)$ and $\mathcal{G}(n,r,cn)$ are equivalent for our problem.

1.2   Hypergraph colouring

Let $\mathbb{N}$ denote the set of positive integers and define $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A function $\sigma : [n] \to [k]$ is called a k-partition of $[n]$, the blocks of the partition being the sets $\sigma^{-1}(i)$, with sizes $n_i = |\sigma^{-1}(i)|$ ($i \in [k]$). Let Π_k denote the set of k-partitions of $[n]$, so $|\Pi_k| = k^n$. A k-colouring of a hypergraph $H = ([n], \mathcal{E})$ is a k-partition σ such that for each edge $e \in \mathcal{E}$, the set $\sigma(e)$ satisfies $|\sigma(e)| > 1$. (We use the notation H for fixed hypergraphs and G for random hypergraphs.) We say an edge $e \in \mathcal{E}$ is *monochromatic* in σ if $|\sigma(e)| = 1$, so a k-partition is a colouring if no edge is monochromatic. The *chromatic number* $\chi(H)$ is the smallest k such that there exists a k-colouring of H.

Note that what we study here is sometimes called the *weak* chromatic number of the hypergraph. The *strong* chromatic number is defined similarly in terms of strong colourings, which are k-partitions σ such that $|\sigma(e)| = |e|$ for each edge $e \in \mathcal{E}$. Even more general notions of colouring may be defined. See, for example, [14]. We will not consider this further here, though it seems probable that the methods we use would be applicable.

The principal objective of the paper will be to prove the following result.

Theorem 1.1. *Define $u_{r,k} = k^{r-1} \ln k$ for integers $r \geq 2$ and $k \geq 1$. Suppose that $r \geq 2$, $k \geq 1$, and let c be a positive constant. Then for $G \in \mathcal{G}(n,r,cn)$,*

(a) If $c \geq u_{r,k}$ then a.a.s. $\chi(G) > k$.

(b) If $k \geq 2$ and $\max\{r,k\} \geq 3$ then there exists a constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ such that if $c < c_{r,k}$ is a positive constant then a.a.s. $\chi(G) \leq k$.

Now the following theorem, which is a complete generalisation of the result of [4] to uniform hypergraphs, follows easily. Note that the lower bound $u_{r,k-1} \leq c$ from Theorem 1.2 is trivial when $k = 2$, since $u_{r,1} = 0$ for all $r \geq 2$.

Theorem 1.2. *For all r, $k \geq 2$, if $c \in [u_{r,k-1}, u_{r,k})$ is a positive constant then a.a.s. the chromatic number of $G \in \mathcal{G}(n,r,cn)$ is either k or $k+1$. Indeed, if $\max\{r,k\} \geq 3$ and $c \in [u_{r,k-1}, c_{r,k})$, where $c_{r,k}$ is a constant satisfying the conditions of Theorem 1.1(b), then a.a.s. $\chi(G) = k$ for $G \in \mathcal{G}(n,r,cn)$.*

Proof. Let $G \in \mathcal{G}(n,r,cn)$ and suppose that $u_{r,k-1} \leq c < u_{r,k}$. By Theorem 1.1(a), we know that $\chi(G) \geq k$ a.a.s., and by Theorem 1.1(b) we know that $\chi(G) \leq k+1$ a.a.s., since $c < u_{r,k} < c_{r,k+1}$. This proves the first statement. Furthermore, if $\max\{r,k\} \geq 3$ and $c < c_{r,k}$ then $\chi(G) \leq k$ a.a.s., by Theorem 1.1(b), proving the final statement. $\square$

For all but a few small values of (r, k) we will see that $c_{r,k}$ is much closer to $u_{r,k}$ than to $u_{r,k-1}$, so that for most values of c, the chromatic number of $G \in \mathcal{G}(n, r, cn)$ is a.a.s. uniquely determined. For more detail see Remark 3.7.

Part (a) of Theorem 1.1 is easy, and is proved in Lemma 2.1. As in [4], part (b) will be proved using the second moment method [13, p.54]. If Z is a random variable defined on $\mathbb{N}_0$, this method applies the inequalities

$$\frac{\mathbf{E}[Z]^2}{\mathbf{E}[Z^2]} \leq \Pr(Z > 0) \leq \mathbf{E}[Z]. \tag{5}$$

Although based on a rather simple idea, the second moment method is often very laborious to apply, and our analysis will be no exception.

A k-partition is called *balanced* if $\lfloor n/k \rfloor \leq n_i \leq \lceil n/k \rceil$ for $i = 1, \ldots, k$. A *balanced k-colouring* of a H is a balanced k-partition which is also a k-colouring of H. For convenience, we will assume that k divides n, so in a balanced colouring, each colour class has precisely n/k vertices. Since we suppose k to be constant, the effects of this assumption are asymptotically negligible as $n \to \infty$. (This is proved in Lemma 1.4 below.) Following [4], our analysis will be carried out mainly in terms of balanced colourings. Indeed, we will apply (5) to the random variable Z which is the number of balanced k-colourings (defined formally in Section 2.1).

Clearly, if $Z > 0$ then a k-colouring exists. However, the analysis in Section 2 will only allow us to conclude that $c < c_{r,k}$ implies that $\liminf_{n\to\infty} \Pr(Z > 0) > 0$. Thus, we first prove a weaker statement about $G \in \mathcal{G}(n, r, cn)$:

(b') *If $r, k \geq 2$ then there exists a constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ such that if $c < c_{r,k}$ is a positive constant then*

$$\liminf_{n\to\infty} \Pr(\chi(G) \leq k) > 0.$$

Then part (b) of Theorem 1.1 will follow from the fact that there is a *sharp threshold* for k-colourability of a random hypergraph (see Lemma 1.3, below). Achlioptas and Naor [4] used a result of Achlioptas and Friedgut [1] which established that random graph k-colourability has a sharp threshold. We will use instead the following, more general, result.

Hatami and Molloy [12] studied the problem of the existence of a homomorphism from a random hypergraph to a fixed hypergraph H. They used the Bernoulli random hypergraph model $\widehat{\mathcal{G}}(n, r, p)$, defined at the end of Section 1.1.

Given a fixed hypergraph $H = ([\nu], \mathcal{E}_H) \in \Omega^*(\nu, r, \mu)$, Hatami and Molloy considered the threshold p for the existence of a homomorphism from $G = ([n], \mathcal{E}_G) \in \widehat{\mathcal{G}}(n, r, p)$ to H. A homomorphism from G to H is a function $\sigma : [n] \to [\nu]$ such that $\sigma(e) \in \mathcal{E}_H$ for all $e \in \mathcal{E}_G$. If H' is formed from H by deleting duplicate edges then the homomorphisms from G to H' are identical to those from G to H, so we may assume that H has no duplicate edges. A *loop* in H is an edge $e \in \mathcal{E}_H$ for which the underlying set is a singleton. A triangle in H is a sequence $(v_1, e_1, v_2, e_2, v_3, e_3)$ of distinct vertices $v_i \in [\nu]$ and edges $e_i \in \mathcal{E}_H$ $(i \in [3])$, such that $v_1, v_2 \in e_1$, $v_2, v_3 \in e_2$ and $v_1, v_3 \in e_3$. The following was proved in [12] (with minor changes of notation):

Theorem 1.3 (Hatami and Molloy)**.** *Let H be a connected undirected loopless r-uniform hypergraph with at least one edge. Then the H-homomorphism problem has a sharp threshold iff (i) $r \geq 3$ or (ii) $r = 2$ and H contains a triangle.*

5

Here a sharp threshold means that there exists a function $p(n)$ taking values in $[0,1]$ for all sufficiently large n such that, for all $0 < \varepsilon < 1$, $G \in \widehat{\mathcal{G}}(n, r, (1-\varepsilon)p)$ has a homomorphism to H a.a.s., and $G \in \widehat{\mathcal{G}}(n, r, (1+\varepsilon)p)$ has no homomorphism to H a.a.s.

Observation 1.2. *The property of having an H-homomorphism is a* monotone decreasing *property of G, that is, an H-homomorphism cannot be destroyed by deleting arbitrary edges of G. This fact will be used later.*

Observation 1.3. *A random hypergraph in $G \in \widehat{\mathcal{G}}(n, r, cn/\binom{n}{r})$ a.a.s. has $cn\left(1 + \Theta\left(n^{-1/4}\right)\right)$ edges (see (7), and $G \in \widehat{\mathcal{G}}(n, r, cn/\binom{n}{r})$ is uniformly random conditioned on the number of edges it contains. Hence if an existence problem has a sharp threshold (with respect to p) for $\widehat{\mathcal{G}}(n, r, p)$ then it has a sharp threshold (with respect to c) for $\mathcal{G}(n, r, cn)$. In this setting, existence of a sharp threshold means that there exists a function $c(n) = \Theta(1)$ such that, for all $0 < \varepsilon < 1$, $G \in \mathcal{G}(n, r, (1-\varepsilon)cn)$ has a homomorphism to H a.a.s., and $G \in \mathcal{G}(n, r, (1+\varepsilon)cn)$ has no homomorphism to H a.a.s.*

Lemma 1.2. *Suppose that $r, k \geq 2$ with $\max\{k, r\} \geq 3$, and let c be a positive constant. Then the problem of k-colouring $G \in \mathcal{G}(n, r, cn)$ has a sharp threshold.*

Proof. Take $K = ([k], \mathcal{E}_K) \in \Omega^*(k, r, \mu)$ to be such that $\mathcal{E}_K$ contains all r-multisets with elements in $[k]$, except for the k possible loops. Then $\mu = \binom{k+r-1}{r} - k$. It is easy to see that the homomorphisms from a graph G to K are precisely the k-colourings of G. If $r = 2$ and $k \geq 3$ then K contains a triangle. (We may take $v_i = i \pmod 3 + 1$ and e_i to be an edge with underlying set $[3] \setminus \{i\}$, for $i \in [3]$.) Thus it follows from Theorem 1.3 that the problem of k-colouring $G \in \widehat{\mathcal{G}}(n, r, p)$ has a sharp threshold unless $k = r = 2$. Hence, by Observation 1.3, the problem of k-colouring $G \in \mathcal{G}(n, r, cn)$ has a sharp threshold unless $k = r = 2$. $\square$

In the excluded case, which is the question of whether a random graph is 2-colourable, it is known that there is no sharp threshold (see [9, Corollary 7]).

We now use Lemma 1.2 to prove the following.

Lemma 1.3. *Suppose that $k \geq 2$ and $\max\{r, k\} \geq 3$. Then (b$'$) implies (b).*

Proof. From part (b$'$) of Theorem 1.1, we have a constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ such that for $G \in \mathcal{G}(n, r, cn)$,
$$\liminf_{n \to \infty} \Pr(\chi(G) \leq k) > 0$$
whenever $c < c_{r,k}$ is a positive constant. Then Lemma 1.2 implies that the threshold function $c(n)$ satisfies $\liminf_{n \to \infty} c(n) \geq c_{r,k}$. Thus for any $c < c_{r,k}$ we have a.a.s. $\chi(G) \leq k$, proving part (b) of Theorem 1.1. $\square$

In fact, we will prove an even weaker statement than (b$'$).

(b$''$) *If $r, k \geq 2$ then there exists a constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ such that for any positive constant $c < c_{r,k}$, the random hypergraph $G \in \mathcal{G}^*(kt, r, ckt)$ satisfies $\liminf_{t \to \infty} \Pr(\chi(G) \leq k) > 0$.*

Observe that, in addition to restricting n to multiples of k, the random hypergraph model for (b$''$) is different from that used in (b$'$). We now show why (b$''$) is sufficient.

Lemma 1.4. *If $r, k \geq 2$ then (b$''$) implies (b$'$).*

6

Proof. Let $\mathrm{P}^*(n,m) = \Pr(\chi(G) \le k)$, where $G \in \mathcal{G}^*(n,r,m)$, and let $\delta(c) = \liminf_{t\to\infty} \mathrm{P}^*(kt,ckt)$. Then (b″) is the statement that there exists a constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ such that $\delta(c) > 0$ for all positive $c < c_{r,k}$. Assume that (b″) holds.

Given n and $c < c_{r,k}$, let $t = \lfloor n/k \rfloor$ and let c' be such that $c < c' < c_{r,k}$. We show in Lemma 4.1 that $G \in \mathcal{G}^*(n,r,cn)$ has at least $k-1$ isolated vertices a.a.s.. Let I be a set of $n - kt \le k-1$ isolated vertices in G, chosen randomly from the set of isolated vertices in G. Form G' from G by deleting the set I of isolated vertices and relabelling the vertices in G' with $[kt]$, respecting the relative ordering. By symmetry, each set of size $n - kt$ is equally likely to be the chosen set I. Hence $G' \in \mathcal{G}^*(kt,r,cn)$, by Observation 1.1, since G can be uniquely reconstructed from G' and I. So $\mathrm{P}^*(n,cn) = \mathrm{P}^*(kt,cn) - o(1)$. Next, if $n \ge c'k/(c'-c)$ then $c'kt > c'(n-k) \ge cn$. Therefore, since k-colourability is a monotone decreasing property (Observation 1.2), it follows that $\mathrm{P}^*(kt,cn) \ge \mathrm{P}^*(kt,c'kt)$.

Finally, since $c' < c_{r,k}$, (b″) implies that $\mathrm{P}^*(kt,c'kt) > \delta(c') - o(1)$, with $\delta(c') > 0$. Hence we have

$$\mathrm{P}^*(n,cn) \ge \mathrm{P}^*(kt,cn) - o(1) \ge \mathrm{P}^*(kt,c'kt) - o(1) \ge \delta(c') - o(1),$$

which implies that

$$\liminf_{n\to\infty} \mathrm{P}^*(n,cn) \ge \delta(c') > 0. \tag{6}$$

By Lemma 1.1, a.a.s. $G' \in \mathcal{G}^*(n,r,c'n)$ has at most $2\ln n$ bad edges. Denote the set of bad edges in G' by $B(G')$. Let G' be a uniformly chosen element of $\Omega^*(n,r,c'n)$ with at most $2\ln n$ bad edges, and form the random hypergraph $\varphi(G')$ as follows: delete $B(G')$ and a set of $(c'-c)n - |B(G')|$ randomly chosen good edges from G'. (If n is sufficiently large then $2\ln n \le (c'-c)n$, making this procedure possible.) The resulting hypergraph $\varphi(G')$ belongs to $\Omega(n,r,cn)$, and, by symmetry, it is a uniformly random element of $\Omega(n,r,cn)$. That is, that $\varphi(G')$ has the same distribution as $G \in \mathcal{G}(n,r,cn)$ when G' is chosen uniformly from those elements of $\Omega^*(n,r,c'n)$ with at most $2\ln n$ bad edges.

Now choose a constant c'' with $c' < c'' < c_{r,k}$. Then by (6) applied to c', we have $\mathrm{P}^*(n,c'n) \ge \delta(c'') > 0$. It follows that for $G' \in \mathcal{G}^*(n,r,c'n)$,

$$\Pr(\chi(G') \le k \text{ and } G' \text{ has at most } 2\ln n \text{ bad edges}) \ge \delta(c'') - o(1).$$

By monotonicity (Observation 1.2), since $\varphi(G')$ has fewer edges than G', we conclude that

$$\Pr(\chi(\varphi(G')) \le k \mid G' \text{ has at most } 2\ln n \text{ bad edges}) \ge \delta(c'') - o(1).$$

Hence using the second statement of Lemma 1.1, $\Pr(\chi(G) \le k) \ge \delta(c'') - o(1)$ for $G \in \mathcal{G}(n,r,cn)$. This shows that (b′) holds, completing the proof. □

The remainder of the paper will be devoted to proving Theorem 1.1, with part (b) weakened to (b″). First we obtain expressions for $\mathbf{E}[Z]$ and $\mathbf{E}[Z^2]$ in Sections 2.1 and 2.2, respectively. The expression for $\mathbf{E}[Z^2]$ is analysed using Laplace's method, under the assumption that constants $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ exist which satisfy some other useful conditions (see Lemma 2.2). This is established in Section 3, completing the proof. Some remarks about asymptotics are made in Section 3.9.

The analysis of Section 3 will require many technical lemmas, some merely verifying inequalities. These inequalities are obvious for large r and k but, since r and k are constants, we need to establish precise conditions under which they are true. We relegate the proofs of most technical lemmas to the

appendix, since they complicate what are fairly natural and straightforward arguments. Therefore, whenever we use a lemma without proof, the proof can be found in the appendix.

To complete this section, we prove the result corresponding to Theorem 1.2 for the Bernoulli random hypergraph model $\widehat{\mathcal{G}}(n, r, p)$. Recall that $u_{r,k} = k^{r-1} \ln k$ and $N = \binom{n}{r}$.

Corollary 1.1. *Let $r \geq 2$. Given a positive constant c, let $k(c, r)$ be the smallest integer k such that $c \leq u_{r,k}$. (Note, $u_{r,k} > 0$ by definition.) If $G \in \widehat{\mathcal{G}}(n, r, cn/N)$ then $\chi(G) \in \{k(c, r), k(c, r) + 1\}$ a.a.s.*

Proof. Let $G \in \widehat{\mathcal{G}}(n, r, cn/N)$, and let m be its (random) number of edges. Then Chernoff's bound [13, Corollary 2.3] gives

$$\Pr\left(|m - cn| \geq cn^{3/4}\right) \leq 2e^{-c\sqrt{n}/3}. \tag{7}$$

Therefore $cn(1 - n^{-1/4}) \leq m \leq cn(1 + n^{-1/4})$ a.a.s., and hence $c'n < m < c''n$ a.a.s. for any positive constants c', c'' such that $c' < c < c''$.

Let $k = k(r, c)$, so $u_{r,k-1} < c \leq u_{r,k}$. Choose $c' \in (u_{r,k-1}, c)$, so $m > c'n$ a.a.s. Now, conditional on $m > c'n$, $c' > u_{r,k-1}$ implies $\chi(G) \geq k$ a.a.s., by Theorem 1.2 and monotonicity (Observation 1.2).

Similarly, choose $c'' \in (c, c_{r,k+1})$, so $m < c''n$ a.a.s. Then, conditional on $m < c''n$, $c'' < c_{r,k+1}$ implies $\chi(G) \leq k+1$ a.a.s., by Theorem 1.2 and Observation 1.2. Thus $\chi(G) \in \{k, k+1\}$ a.a.s. $\square$

Remark 1.2. *We have shown the equivalence of various models for our problem when $\max\{k, r\} \geq 3$. We note that this equivalence does not hold for the case $k = r = 2$, where the non-existence of a 2-colouring is equivalent to the appearance of an odd cycle in a random graph. This is due to the absence of a sharp threshold for this appearance [9, Corollary 7]. Fortunately, this has little impact on our results.*

2 Moment calculations

2.1 First moment

Lemma 2.1. *Let $r \geq 2$, $k \geq 1$ and recall that $u_{r,k} = k^{r-1} \ln k$. Suppose that $c \geq u_{r,k}$ is a positive constant and let $G \in \mathcal{G}^*(n, r, cn)$. Then a.a.s. $\chi(G) > k$.*

Proof. First suppose that $k = 1$. Since $c > 0$, the hypergraph G has at least one edge, so $\chi(G) > 1$ with probability 1.

For the rest of the proof, assume that $k \geq 2$. Consider any k-partition $\sigma \in \Pi_k$ with block sizes n_i ($i \in [k]$). Given σ, a random edge $e \in \mathcal{E}$ is monochromatic with probability

$$\sum_{i=1}^{k} (n_i/n)^r \geq k(1/k)^r = 1/k^{r-1},$$

using Jensen's inequality [11] with the convex function x^r. Since the edges in $\mathcal{E}$ are chosen independently, the probability that σ is a k-colouring of G is at most $(1 - 1/k^{r-1})^{cn}$. Let X be the number of k-colourings of G. Using (5) and the fact that $|\Pi_k| = k^n$, we conclude that $\Pr(X > 0) \leq \mathbf{E}[X] \leq \left(k\left(1 - 1/k^{r-1}\right)^c\right)^n$. If $c \geq u_{r,k}$ then $c > (k^{r-1} - 1/2) \ln k$, and hence

$$k\left(1 - \frac{1}{k^{r-1}}\right)^c = \exp\left(\ln k + c \ln\left(1 - \frac{1}{k^{r-1}}\right)\right) \leq \exp\left(\ln k - \frac{c}{k^{r-1} - 1/2}\right) < 1,$$

8

where we have used Lemma 4.7 in the penultimate inequality. It follows that $\Pr(X > 0) \to 0$ as $n \to \infty$ when $c > u_{r,k}$. $\qquad\square$

Remark 2.1. *We have proved the slightly stronger bound $(k^{r-1} - 1/2) \ln k$. This is used in [3], and noted, but not used, in [4]. Since the difference is small, we mainly use the simpler bound $k^{r-1} \ln k$.*

In the remainder of the paper, we will assume that k divides n, unless stated otherwise. Recall that Z is the number of balanced colourings of $G \in \mathcal{G}^*(n, r, cn)$. Let Ξ_k denote the set of all balanced k-partitions of $[n]$. For any balanced partition $\sigma \in \Xi_k$ and any $e \subseteq [n]$, let $M_e(\sigma)$ be the event that $|\sigma(e)| = 1$. If e is an edge of $G \in \mathcal{G}^*(n, r, cn)$ then clearly $\Pr(\overline{M_e(\sigma)}) = 1 - 1/k^{r-1}$, and these events are independent for $e \in \mathcal{E}$. Thus, since $|\Xi_k| = n!/\big((n/k)!\big)^k$,

$$\mathbf{E}[Z] \;=\; \frac{n!}{\big((n/k)!\big)^k}\Big(1 - \frac{1}{k^{r-1}}\Big)^{cn} \;\sim\; \frac{k^{k/2}}{(2\pi n)^{(k-1)/2}}\Big(k\Big(1 - \frac{1}{k^{r-1}}\Big)^c\Big)^n. \tag{8}$$

We have suppressed the discretisation error $cn - \lfloor cn \rfloor$. This would apparently give an additional $O(1)$ factor in $\mathbf{E}[Z]$ here, and in $\mathbf{E}[Z^2]$ below. This is of no consequence for two reasons:

(i) We need only prove that $\liminf_{n\to\infty} \mathbf{E}[Z^2]/\mathbf{E}[Z] = \Omega(1)$, so the correction is unimportant.

(ii) The asymptotic value for $\mathbf{E}[Z^2]/\mathbf{E}[Z]$ we obtain is independent of n, so using the sequence $c_n = \lfloor cn \rfloor/n$ gives the same asymptotic approximation as that given by using c.

2.2 Second moment

Using the notation of Section 2.1, let $\sigma, \tau \in \Xi_k$ be balanced partitions. Then $\overline{M_e(\sigma)} \cap \overline{M_e(\tau)}$ is the event that the edge e is not monochromatic in either σ or τ. For $i, j \in [k]$, define

$$\ell_{ij} \;=\; |\{v \in [n] : \sigma(v) = i,\ \tau(v) = j\}|.$$

Let $\boldsymbol{L}$ be the $k \times k$ matrix (ℓ_{ij}). Then $\boldsymbol{L} \in \mathcal{D}$, where

$$\mathcal{D} \;=\; \{\boldsymbol{L} \in \mathbb{N}_0^{k \times k} \;:\; \textstyle\sum_{i=1}^{k} \ell_{ij} = \sum_{j=1}^{k} \ell_{ij} = n/k\}.$$

There are exactly $n!/\big(\prod_{i,j=1}^{k} \ell_{ij}!\big)$ pairs $\sigma,\ \tau \in \Xi_k$ which share the same matrix $\boldsymbol{L} \in \mathcal{D}$.

Now $\Pr(M_e(\sigma)) = \Pr(M_e(\tau)) = 1/k^{r-1}$, and

$$\Pr(M_e(\sigma) \cap M_e(\tau)) \;=\; \sum_{i=1}^{k}\sum_{j=1}^{k} \Big(\frac{\ell_{ij}}{n}\Big)^{r}.$$

Thus by inclusion-exclusion,

$$\Pr(\overline{M_e(\sigma)} \cap \overline{M_e(\tau)}) \;=\; 1 - \Pr(M_e(\sigma)) - \Pr(M_e(\tau)) + \Pr(M_e(\sigma) \cap M_e(\tau)) \tag{9}$$

$$= 1 - \frac{2}{k^{r-1}} + \sum_{i=1}^{k}\sum_{j=1}^{k}\Big(\frac{\ell_{ij}}{n}\Big)^{r}.$$

Therefore

$$\mathbf{E}[Z^2] \;=\; \sum_{\sigma,\tau\in\Xi_k}\Big(1 - \frac{2}{k^{r-1}} + \sum_{i,j=1}^{k}\Big(\frac{\ell_{ij}}{n}\Big)^{r}\Big)^{cn}$$

9

$$= \sum_{\boldsymbol{L} \in \mathcal{D}} \frac{n!}{\prod_{i,j=1}^{k} \ell_{ij}!} \left(1 - \frac{2}{k^{r-1}} + \sum_{i,j=1}^{k} \left(\frac{\ell_{ij}}{n} \right)^r \right)^{cn}. \tag{10}$$

Let $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$ and $\overline{\mathbb{R}}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Then, for $\boldsymbol{X} = (x_{ij}) \in \mathbb{R}_+^{k \times k}$, define the functions

$$F(\boldsymbol{X}) = -\sum_{i=1}^{k} \sum_{j=1}^{k} x_{ij} \ln x_{ij} + c \ln \left(1 - \frac{2}{k^{r-1}} + \sum_{i=1}^{k} \sum_{j=1}^{k} x_{ij}^r \right), \tag{11}$$

$$G(\boldsymbol{X}) = (2\pi n)^{-(k^2-1)/2} \left(\prod_{i,j=1}^{k} x_{ij} \right)^{-1/2}.$$

We can extend F to $\overline{\mathbb{R}}_+^{k \times k}$ by continuity, setting $x \ln x = 0$ when $x = 0$.

We now apply Stirling's inequality in the form

$$p! = \sqrt{2\pi(p \wedge 1)} \left(\frac{p}{e} \right)^p \left(1 + O\left(\frac{1}{p+1} \right) \right),$$

valid for all integers $p \geq 0$, where $p \wedge 1 = \max\{p, 1\}$. If all the ℓ_{ij} are positive then the summand of (10) becomes

$$\frac{n!}{\prod_{i,j=1}^{k} \ell_{ij}!} \left(1 - \frac{2}{k^{r-1}} + \sum_{i,j=1}^{k} \left(\frac{\ell_{ij}}{n} \right)^r \right)^{cn} = G(\boldsymbol{L}/n) \, e^{nF(\boldsymbol{L}/n)} \left(1 + O\left(\frac{1}{\min \ell_{ij} + 1} \right) \right). \tag{12}$$

If any of the ℓ_{ij} equal zero then the above expression still holds with the corresponding argument x_{ij} of G replaced by $1/n$, for all such i, j (and treating n as fixed).

Let $\boldsymbol{J}_0$ be the $k \times k$ matrix with all entries equal to $1/k^2$. Then

$$F(\boldsymbol{J}_0) = 2 \ln k + 2c \ln \left(1 - 1/k^{r-1} \right) = \ln \left(k \left(1 - \frac{1}{k^{r-1}} \right)^c \right)^2, \tag{13}$$

$$G(\boldsymbol{J}_0) = (2\pi n)^{-(k^2-1)/2} k^{k^2}. \tag{14}$$

Hence the term of (10) corresponding to $\boldsymbol{L} = n\boldsymbol{J}_0$ is asymptotically equal to

$$\frac{k^{k^2}}{(2\pi n)^{(k^2-1)/2}} \left(k \left(1 - \frac{1}{k^{r-1}} \right)^c \right)^{2n}.$$

Observe from (8) that this term is smaller than $\mathbf{E}[Z]^2$ by a factor which is polynomial in n. We will find a positive constant $c_{r,k}$ such that when $c \in (0, c_{r,k})$, the function $F(\boldsymbol{X})$ has a unique maximum at $\boldsymbol{X} = \boldsymbol{J}_0$. This will allow us to apply the following theorem of Greenhill, Janson and Ruciński [10] to estimate $\mathbf{E}[Z^2]$ in the region where $c < c_{r,k}$. (See that paper for background and definitions.)

Theorem 2.1 (Greenhill, Janson and Ruciński [10]). *Suppose the following:*

(i) $\mathcal{L} \subset \mathbb{R}^{\mathsf{N}}$ *is a lattice with rank* r.

(ii) $\mathsf{V} \subseteq \mathbb{R}^{\mathsf{N}}$ *is the* r-*dimensional subspace spanned by* $\mathcal{L}$.

(iii) $\mathsf{W} = \mathsf{V} + \mathsf{w}$ *is an affine subspace parallel to* V, *for some* $\mathsf{w} \in \mathbb{R}^{\mathsf{N}}$.

10

(iv) $\mathsf{K} \subset \mathbb{R}^{\mathsf{N}}$ *is a compact convex set with non-empty interior* K°.

(v) $\varphi : \mathsf{K} \to \mathbb{R}$ *is a continuous function and the restriction of* φ *to* $\mathsf{K} \cap \mathsf{W}$ *has a unique maximum at some point* $x_0 \in \mathsf{K}^{\circ} \cap \mathsf{W}$.

(vi) φ *is twice continuously differentiable in a neighbourhood of* x_0 *and* $\mathsf{H} := \mathsf{D}^2\varphi(x_0)$ *is its Hessian at* x_0.

(vii) $\psi : \mathsf{K}_1 \to \mathbb{R}$ *is a continuous function on some neighbourhood* $\mathsf{K}_1 \subseteq \mathsf{K}$ *of* x_0 *with* $\psi(x_0) > 0$.

(viii) *For each positive integer* n *there is a vector* $\ell_n \in \mathbb{R}^{\mathsf{N}}$ *with* $\ell_n/n \in \mathsf{W}$,

(ix) *For each positive integer* n, *there is a positive real number* b_n, *and a function*
$$a_n : (\mathcal{L} + \ell_n) \cap n\mathsf{K} \to \mathbb{R} \text{ such that, as } n \to \infty,$$
$$a_n(\ell) = O\big(b_n e^{n\varphi(\ell/n)+o(n)}\big), \qquad\qquad \ell \in (\mathcal{L} + \ell_n) \cap n\mathsf{K},$$
$$\text{and} \quad a_n(\ell) = b_n\big(\psi(\ell/n) + o(1)\big)e^{n\varphi(\ell/n)}, \qquad \ell \in (\mathcal{L} + \ell_n) \cap n\mathsf{K}_1,$$

uniformly for ℓ *in the indicated sets.*

Then provided $\det(-\mathsf{H}|_{\mathsf{V}}) \neq 0$, *as* $n \to \infty$,
$$\sum_{\ell \in (\mathcal{L}+\ell_n) \cap n\mathsf{K}} a_n(\ell) \sim \frac{(2\pi)^{r/2}\psi(x_0)}{\det(\mathcal{L})\det(-\mathsf{H}|_{\mathsf{V}})^{1/2}} b_n n^{r/2} e^{n\varphi(x_0)}. \qquad\qquad \square$$

As remarked in [10], the asymptotic approximation given by this theorem remains valid for $n \in I$, where $I \subset \mathbb{N}$ is infinite, provided (viii) and (ix) hold for all $n \in I$. The conclusion of the theorem then holds for $n \in I$ as $n \to \infty$.

We will use this observation with $I = \{kt : t \in \mathbb{N}\}$, since we require only the weaker statement (b″) in Theorem 1.1.

We must relate the quantities in Theorem 2.1 to our notation and analysis. We let n be n, restricted to positive integers divisible by k. Denote by $\mathbb{R}^{k \times k}$ the set of real $k \times k$ matrices, which we will view as k^2-vectors in the space $\mathbb{R}^{k^2}$. Then $\mathsf{N} = k^2$ in Theorem 2.1. Next, V in Theorem 2.1 will be the subspace $\mathcal{M}$ of $\mathbb{R}^{k \times k}$ containing all matrices $\boldsymbol{X}$ such that all row and column sums are zero, i.e.
$$\sum_{i=1}^{k} x_{ij} = \sum_{i=1}^{k} x_{ji} = 0 \qquad (j \in [k]),$$
and the affine subspace W will consist of the matrices $\boldsymbol{X}$ such that all row and column sums are $1/k$, i.e.
$$\sum_{i=1}^{k} x_{ij} = \sum_{i=1}^{k} x_{ji} = 1/k \qquad (j \in [k]).$$
The point $w \in \mathsf{W}$ will be $\boldsymbol{J}_0$.

The lattice $\mathcal{L}$ in Theorem 2.1 will be the set of integer matrices in $\mathcal{M}$: that is, the set of all $k \times k$ integer matrices $\boldsymbol{L} = (\ell_{ij})$ such that
$$\sum_{i=1}^{k} \ell_{ij} = \sum_{i=1}^{k} \ell_{ji} = 0 \qquad (j \in [k]).$$
Let ℓ_n equal the $k \times k$ diagonal matrix with all diagonal entries equal to n/k. Then $\ell_n/n \in \mathsf{W}$ and ℓ_n is an integer matrix, since we assume that n is divisible by k.

The compact convex set K will be the subset of $\mathbb{R}^{k \times k}$ such that $0 \leq x_{ij} \leq 1/k$ $(i, j \in [k])$, which has non-empty interior $\mathsf{K}_0 = \{x_{ij} : 0 < x_{ij} < 1/k\}$. Define $a_n(\boldsymbol{L})$ to be the summand of (10); that is,
$$a_n(\boldsymbol{L}) = \frac{n!}{\prod_{i,j=1}^{k} \ell_{ij}!} \left(1 - \frac{2}{k^{r-1}} + \sum_{i,j=1}^{k} \left(\frac{\ell_{ij}}{n}\right)^r\right)^{cn}.$$

11

We wish to calculate $\mathbf{E}[Z^2]$, which by (10) equals

$$\sum_{\boldsymbol{L} \in (\mathcal{L}+\ell_n) \cap nK} \mathsf{a}_\mathsf{n}(\boldsymbol{L}). \tag{15}$$

In Section 3 we will prove the following result.

Lemma 2.2. *Recall that $u_{r,k} = k^{r-1} \ln k$ for $r \geq 2, k \geq 1$. Now fix $r, k \geq 2$. There exists a positive constant $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ which satisfies*

$$c_{r,k} \leq \frac{(k^{r-1}-1)^2}{r(r-1)},$$

such that F has a unique maximum in $\mathsf{K} \cap \mathsf{W}$ at the point $\boldsymbol{J}_0 \in \mathsf{K}_0 \cap \mathsf{W}$ whenever $c \in (0, c_{r,k})$.

Throughout this section we assume that Lemma 2.2 holds. Then $\boldsymbol{J}_0$ is the unique maximum of F within $\mathsf{K} \cap \mathsf{W}$, so we set $\phi := F$ and $\mathsf{x}_0 := \boldsymbol{J}_0$. Note that F is analytic in a neighbourhood of $\boldsymbol{J}_0$.

Let K_1 be any neighbourhood of $\boldsymbol{J}_0$ whose closure is contained within K_0. The function ψ in Theorem 2.1 will be defined by $\psi(\boldsymbol{X}) = \prod_{i,j=1}^{k} x_{ij}^{-1/2}$. So ψ is positive and analytic on K_1. We let b_n equal $(2\pi n)^{-(k^2-1)/2}$. By (12), the quality of approximations required by (ix) of Theorem 2.1 hold. To see this, observe that the relative error in (12) is always $O(1)$ and that $G(\boldsymbol{L}/n) = e^{o(n)}$ for all $\boldsymbol{L} \in (\mathcal{L} + \ell_n) \cap nK$. This proves the first statement in (ix). However, if $\boldsymbol{L} \in (\mathcal{L} + \ell_n) \cap n\mathsf{K}_1$ then all $\ell_{ij} = \Theta(n)$ and hence the relative error in (12) is $1 + O(1/n)$. Since then $\psi(\boldsymbol{L}/n)(1 + O(1/n)) = \psi(\boldsymbol{L}/n) + o(1)$, the second statement in (ix) holds.

Next, observe that $\mathcal{L}$ and $\mathcal{M}$ respectively have rank and dimension $(k-1)^2$, since we may specify ℓ_{ij} or x_{ij} $(i, j \in [k-1])$ arbitrarily, and then all ℓ_{ij} or x_{ij} $(i, j \in [k])$ are determined. Thus $\mathsf{r} = (k-1)^2$ in Theorem 2.1.

We now calculate the determinants required in Theorem 2.1. Let H be the Hessian of F at the point $\boldsymbol{J}_0$. This matrix can be regarded as a quadratic form on $\mathbb{R}^{k \times k}$. In Theorem 2.1 we need the determinant of $-H|_{\mathcal{M}}$, which denotes the quadratic form $-H$ restricted to the subspace $\mathcal{M}$ of $\mathbb{R}^{k \times k}$. This can be calculated by

$$\det(-H|_{\mathcal{M}}) = \frac{\det U^T(-H)U}{\det U^T U} \tag{16}$$

for any $k^2 \times (k-1)^2$ matrix U whose columns form a basis of $\mathcal{M}$.

Lemma 2.3. *Suppose that $r, k \geq 2$ and $0 < c < c_{r,k}$, where $c_{r,k}$ satisfies Lemma 2.2. Then the determinant of $\mathcal{L}$ is $\det \mathcal{L} = k^{k-1}$ and the determinant of $-H|_{\mathcal{M}}$ is $(k^2\alpha)^{(k-1)^2}$, where*

$$\alpha = 1 - \frac{cr(r-1)}{(k^{r-1}-1)^2}.$$

Proof. Let δ_{ij} be the Kronecker delta, and define the matrices $\boldsymbol{E}_{ij}$ by $(\boldsymbol{E}_{ij})_{i'j'} = \delta_{ii'}\delta_{jj'}$. Then $\{\boldsymbol{E}_{ij} : i, j \in [k]\}$ forms a basis for $\mathbb{R}^{k \times k}$. Let $\boldsymbol{E}_{i*}$ be such that $(\boldsymbol{E}_{i*})_{i'j'} = \delta_{ii'}$, and $\boldsymbol{E}_{*j}$ be such that $(\boldsymbol{E}_{*j})_{i'j'} = \delta_{jj'}$. Then $\mathcal{M}$ is the subspace of $\mathbb{R}^{k \times k}$ which is orthogonal to $\{\boldsymbol{E}_{i*}, \boldsymbol{E}_{*i} : i \in [k]\}$. We claim that the vectors $\boldsymbol{U}_{ij} = \boldsymbol{E}_{ij} - \boldsymbol{E}_{ik} - \boldsymbol{E}_{kj} + \boldsymbol{E}_{kk}$ $(i, j \in [k-1])$ form a basis for $\mathcal{M}$. To show this, consider elements of $\mathbb{R}^{k \times k}$ as vectors in $\mathbb{R}^{k^2}$ (under the lexicographical ordering of the

12

indices (i, j), say). Then for $i' \in [k]$, $i, j \in [k-1]$, taking dot products in $\mathbb{R}^{k^2}$ gives

$$
\begin{aligned}
\boldsymbol{E}_{i'*} \cdot \boldsymbol{U}_{ij} &= \sum_{\ell=1}^{k}\sum_{\ell'=1}^{k} (\boldsymbol{E}_{i'*})_{\ell\ell'} \, (\boldsymbol{U}_{ij})_{\ell\ell'} \\
&= \sum_{\ell=1}^{k}\sum_{\ell'=1}^{k} \delta_{i'\ell}\big(\delta_{i\ell}\delta_{j\ell'} - \delta_{i\ell}\delta_{k\ell'} - \delta_{k\ell}\delta_{j\ell'} + \delta_{k\ell}\delta_{k\ell'}\big) = \delta_{ii'} - \delta_{ii'} - \delta_{ki'} + \delta_{ki'} = 0, \quad (17)
\end{aligned}
$$

and, similarly,

$$
\boldsymbol{E}_{*i'} \cdot \boldsymbol{U}_{ij} = \sum_{\ell=1}^{k}\sum_{\ell'=1}^{k} \delta_{i'\ell'}\big(\delta_{i\ell}\delta_{j\ell'} - \delta_{i\ell}\delta_{k\ell'} - \delta_{k\ell}\delta_{j\ell'} + \delta_{k\ell}\delta_{k\ell'}\big) = \delta_{i'j} - \delta_{i'k} - \delta_{i'j} + \delta_{i'k} = 0. \quad (18)
$$

Thus, from (17) and (18), the $(k-1)^2$ vectors $\boldsymbol{U}_{ij}$ lie in $\mathcal{M}$, so we need only show that they are linearly independent. We will do this by computing the determinant of the corresponding $(k-1)^2 \times (k-1)^2$ Gram matrix M. Let U be the $k^2 \times (k-1)^2$ matrix with columns $\boldsymbol{U}_{ij}$ $(i, j \in [k-1])$. Then $M = (m_{ij,i'j'}) = U^T U$, and we calculate (taking dot products in $\mathbb{R}^{k^2}$),

$$
\begin{aligned}
m_{ij,i'j'} &= \boldsymbol{U}_{ij} \cdot \boldsymbol{U}_{i'j'} \\
&= \sum_{\ell=1}^{k}\sum_{\ell'=1}^{k} \big(\delta_{i\ell}\delta_{j\ell'} - \delta_{i\ell}\delta_{k\ell'} - \delta_{k\ell}\delta_{j\ell'} + \delta_{k\ell}\delta_{k\ell'}\big)\big(\delta_{i'\ell}\delta_{j'\ell'} - \delta_{i'\ell}\delta_{k\ell'} - \delta_{k\ell}\delta_{j'\ell'} + \delta_{k\ell}\delta_{k\ell'}\big) \\
&= \delta_{ii'}\delta_{jj'} + \delta_{ii'} + \delta_{jj'} + 1.
\end{aligned}
$$

It follows that M is a $(k-1) \times (k-1)$ block matrix, with blocks of size $(k-1) \times (k-1)$, such that

$$
M = \begin{bmatrix} 2B & B & \cdots & B \\ B & 2B & \cdots & B \\ B & B & \ddots & B \\ B & B & \cdots & 2B \end{bmatrix}, \quad \text{where} \quad B = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ 1 & 1 & \ddots & 1 \\ 1 & 1 & \cdots & 2 \end{bmatrix}.
$$

We compute the determinant of matrices of this form in Lemma 4.2. Taking $p = q = k - 1$ in Lemma 4.2, we have $\det M = k^{k-1} k^{k-1} = k^{2(k-1)}$. In particular, since the determinant is nonzero, it follows that the $\boldsymbol{U}_{ij}$ $(i, j \in [k-1])$ give a basis for $\mathcal{M}$. Also note that, after permuting its rows,

$$
U = \begin{bmatrix} I_{k-1} \\ \hline U' \end{bmatrix},
$$

where I_{k-1} is the $(k-1)^2 \times (k-1)^2$ identity matrix, and U' is a $(2k-1) \times (k-1)^2$ integer matrix with entries in $\{-1, 0, 1\}$. Therefore for $\boldsymbol{X} \in \mathbb{R}^{k^2}$ and $\boldsymbol{Y} \in \mathbb{R}^{(k-1)^2}$, we have $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{Y}$ if and only if $\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}$ for $i, j \in [k-1]$. It follows that $\{\boldsymbol{U}_{ij} : i, j \in [k-1]\}$ is a basis for the lattice $\mathcal{L}$ and hence the determinant of $\mathcal{L}$ is $\det \mathcal{L} = \sqrt{\det M} = k^{k-1}$.

We also require the determinant of $-H|_{\mathcal{M}}$. For $\boldsymbol{X} \in \mathcal{M}$, let

$$
F_1(\boldsymbol{X}) = -\sum_{i=1}^{k}\sum_{j=1}^{k} x_{ij} \ln x_{ij}, \qquad F_2(\boldsymbol{X}) = 1 - \frac{2}{k^{r-1}} + \sum_{i=1}^{k}\sum_{j=1}^{k} x_{ij}^r.
$$

Then $H = \left(h_{ij,i'j'}\right)$ has entries

$$h_{ij,i'j'} = \left[\frac{\partial^2 F_1}{\partial x_{ij}\partial x_{i'j'}}\right]_{\boldsymbol{J}_0} + c\left[\frac{\partial^2 \ln F_2}{\partial x_{ij}\partial x_{i'j'}}\right]_{\boldsymbol{J}_0}.$$

Now

$$\left[\frac{\partial^2 F_1}{\partial x_{ij}\partial x_{i'j'}}\right]_{\boldsymbol{J}_0} = \left[-\frac{1}{x_{ij}}\right]_{\boldsymbol{J}_0}\delta_{ii'}\delta_{jj'} = -k^2\,\delta_{ii'}\delta_{jj'}.$$

Next,

$$\frac{\partial \ln F_2}{\partial x_{ij}} = \frac{1}{F_2}\frac{\partial F_2}{\partial x_{ij}} \quad \text{and} \quad \frac{\partial F_2}{\partial x_{ij}} = rx_{ij}^{r-1}.$$

Hence

$$
\begin{aligned}
\left[\frac{\partial^2 \ln F_2}{\partial x_{ij}\partial x_{i'j'}}\right]_{\boldsymbol{J}_0} &= \left[\frac{1}{F_2}\frac{\partial^2 F_2}{\partial x_{ij}\partial x_{i'j'}}\delta_{ii'}\delta_{jj'} - \frac{1}{F_2^2}\frac{\partial F_2}{\partial x_{ij}}\frac{\partial F_2}{\partial x_{i'j'}}\right]_{\boldsymbol{J}_0} \\
&= \left[\frac{r(r-1)x_{ij}^{r-2}}{F_2}\delta_{ii'}\delta_{jj'} - \frac{r^2 x_{ij}^{r-1}x_{i'j'}^{r-1}}{F_2^2}\right]_{\boldsymbol{J}_0} \\
&= \frac{r(r-1)}{k^{2(r-2)}(1-1/k^{r-1})^2}\,\delta_{ii'}\delta_{jj'} - \frac{r^2}{k^{4(r-1)}(1-1/k^{r-1})^4} \\
&= \frac{k^2 r(r-1)}{(k^{r-1}-1)^2}\,\delta_{ii'}\delta_{jj'} - \frac{r^2}{(k^{r-1}-1)^4}.
\end{aligned}
$$

Here we have used the fact that $F_2(\boldsymbol{J}_0) = (1-1/k^{r-1})^2$.

These calculations show that

$$-H = k^2\,\alpha I_k + \beta J$$

where I_k is the $k^2 \times k^2$ identity matrix, J is the $k^2 \times k^2$ matrix with all entries equal to 1,

$$\alpha = 1 - \frac{cr(r-1)}{(k^{r-1}-1)^2}\,, \qquad \beta = \frac{cr^2}{(k^{r-1}-1)^4}.$$

By (16), the determinant of $-H|_{\mathcal{M}}$ equals

$$\frac{\det U^T(-H)U}{\det U^T U} = \frac{\det U^T(k^2\alpha I_k + \beta J)U}{\det U^T U} = \frac{\det(k^2\alpha)U^T U}{\det U^T U} = \frac{(k^2\alpha)^{(k-1)^2}\det U^T U}{\det U^T U} = (k^2\alpha)^{(k-1)^2}.$$

Here we have used the fact that $JU = 0$, which follows since every column of U is an element of $\mathcal{M}$ and hence has zero sum. This completes the proof. $\qquad\square$

Note that $\psi(\boldsymbol{J}_0) = k^{k^2}$, while (13) gives $\phi(\boldsymbol{J}_0) = F(\boldsymbol{J}_0) = 2\ln\left(k(1-1/k^{r-1})^c\right)$. Now α is positive when $c \in (0, c_{r,k})$, using Lemma 2.2. Hence Lemma 2.3 guarantees that $\det(-H|_{\mathcal{M}}) \neq 0$. Therefore we can apply Theorem 2.1 to (10), giving

$$\mathbf{E}[Z^2] \sim \frac{k^k}{(2\pi n)^{k-1}\alpha^{(k-1)^2/2}}\left(k\left(1-\frac{1}{k^{r-1}}\right)^c\right)^{2n}.$$

Thus, from (8), for all $r, k \geq 2$ we have

$$\Pr(Z > 0) \geq \frac{\mathbf{E}[Z]^2}{\mathbf{E}[Z^2]} \sim \alpha^{(k-1)^2/2},$$

which is a positive constant. So $\liminf_{n\to\infty} \Pr(Z > 0) > 0$ and we have established part (b$''$) of Theorem 1.1, under the assumption that Lemma 2.2 holds.

It remains to prove Lemma 2.2, which is the focus of the next section.

14

3 Optimisation

We now consider maximising the function F in (11), and develop conditions under which this function has a unique maximum at $\boldsymbol{J}_0$. In doing so, we will determine suitable constants $c_{r,k}$ and prove that Lemma 2.2 holds. This will complete the proof of Theorem 1.1.

Our initial goal will be to reduce the maximisation of F to a univariate optimisation problem. This reduction is performed in several stages, presented in Sections 3.2–3.4. We analyse the resulting univariate problem in Sections 3.5–3.8. For more detail on our optimisation strategy, see Section 3.1. Finally, we consider a simplified asymptotic treatment of the univariate optimisation problem in Section 3.9.

As is common when working with convex functions, we define $x \ln x = +\infty$ for all $x < 0$.

It will be convenient to rescale the variables, letting $\boldsymbol{A} = (a_{ij})$ be the $k \times k$ matrix defined by $\boldsymbol{A} = k\boldsymbol{X}$, so $a_{ij} = kx_{ij}$ for all $i, j \in [k]$. Substituting into (11), we can write

$$F(\boldsymbol{X}) = \ln k - \frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} \ln a_{ij} + c \ln \left(1 - \frac{2}{k^{r-1}} + \frac{\rho}{k^{2r-2}} \right)$$

where

$$\rho = k^{r-2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^r.$$

Letting $z = F(\boldsymbol{X}) - \ln k$, we consider the optimisation problem

$$\text{maximise} \quad z = -\frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} \ln a_{ij} + c \ln \left(1 - \frac{2}{k^{r-1}} + \frac{\rho}{k^{2r-2}} \right) \tag{19a}$$

$$\text{subject to} \quad \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^r = \frac{\rho}{k^{r-2}}, \tag{19b}$$

$$\sum_{j=1}^{k} a_{ij} = 1 \quad (i \in [k]), \tag{19c}$$

$$\sum_{i=1}^{k} a_{ij} = 1 \quad (j \in [k]), \tag{19d}$$

$$a_{ij} \geq 0 \quad (i, j \in [k]). \tag{19e}$$

In any feasible solution to (19) we have

$$\rho = k^{r-2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^r \leq k^{r-2} \sum_{i=1}^{k} \left(\sum_{j=1}^{k} a_{ij} \right)^r = k^{r-1} \tag{20}$$

and

$$\rho = k^{r-2} \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^r \geq \frac{k^{r-2} \left(\sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} \right)^r}{\left(\sum_{i=1}^{k} \sum_{j=1}^{k} 1 \right)^{r-1}} = \frac{k^{r-2} k^r}{k^{2(r-1)}} = 1, \tag{21}$$

where we have used Hölder's inequality [11] in (21). Hence the system (19b)–(19e) is infeasible if $\rho \notin [1, k^{r-1}]$, in which case we set $\max z = -\infty$. Conversely, it is easy to show that the system (19b)–(19e) is feasible for all $\rho \in [1, k^{r-1}]$. A formal proof of this is given in Lemma 4.3.

We wish to determine the structure of the maximising solutions in the optimisation problem (19). Following [4], we will relax the constraints (19d), and write (19b) as

$$\sum_{j=1}^{k} a_{ij}^r = \frac{\varrho_i}{k^{r-2}}, \qquad \sum_{i=1}^{k} \varrho_i = \rho. \tag{22}$$

By the same method as for Lemma 4.3, we can show that the system given by (19c), (19e) and (22) is feasible if and only if $\sum_{j=1}^{k} \varrho_i = \rho$ and $\varrho_i \in [1/k, k^{r-2}]$ for $i \in [k]$. Note that (20) and (21) assume $a_{ij} \geq 0$, but the relaxation of (19e) will be unimportant. Since $z = -\infty$ whenever some $a_{ij} < 0$, these conditions must be satisfied automatically at any finite optimum.

Remark 3.1. *For a fixed value of $\rho \in [1, k^{r-1}]$, we can ignore the term $c \ln(1 - 2/k^{r-1} + \rho/k^{2r-2})$ in the objective function z in (19a), leading to the maximisation problem considered in Lemma 4.4 with $\ell = k$ and $\widehat{\rho} = \rho/k^{r-2}$. Then Lemma 4.4 shows that this objective function cannot be maximised on the boundary of the region determined by (19c) and (19e), unless $\rho = k^{r-1}$. Thus we can also omit the constraints (19e) in the optimisation, and consider the reduced system (19a)–(19c). Allowing ρ to vary clearly cannot introduce a local maximum on the boundary, since Lemma 4.4 holds for all $\rho \in [1, k^{r-1}]$. In fact, it is not difficult to see that there can be no local maximum on the boundary even for the complete system (19). This can be proved formally in a similar way to Lemma 4.4. We omit the proof since we do not use this observation.*

3.1 Outline of the optimisation strategy

Having dropped the constraints (19d), we can break (19) up into k independent simpler one-row subproblems, if we specify the values $(\varrho_1, \varrho_2, \ldots, \varrho_k)$ such that $\rho = \varrho_1 + \varrho_2 + \cdots + \varrho_k$. This is done in Section 3.2. Later in Section 3.3 we will determine the optimal values of $\varrho_1, \varrho_2, \ldots, \varrho_k$ for a fixed value of ρ. Then in Section 3.4 we will allow ρ to vary.

Section 3.2 reduces the one-row problem (23) to an essentially one-variable problem (30). For this problem there is a unique optimum value of $\beta = \beta(\varrho)$ that is given in (29). This value of β maximises the objective function, now expressed as $-\mathsf{f}(\beta)$: see (32). The nonlinear constraint (23b) will now have been replaced by an equation $\mathsf{g}(\beta) = k^{2-r}\varrho - k^{1-r}$, see (32).

In Section 3.3 we try to find values $\varrho_1, \varrho_2, \ldots, \varrho_k$ that sum to a fixed ρ and and associated values $\beta_1, \beta_2, \ldots, \beta_k$ that minimise $\mathsf{f}(\beta_1) + \mathsf{f}(\beta_2) + \cdots + \mathsf{f}(\beta_k)$. The constraints become $\mathsf{g}(\beta_1) + \mathsf{g}(\beta_2) + \cdots + \mathsf{g}(\beta_k) = k^{2-r}(\rho - 1)$, see (33). We show that the β_i take one of at most two values, $\gamma_1 \leq \gamma_2$. These values are the solutions to $\mathsf{f}'(\beta)/\mathsf{g}'(\beta) = \lambda$ where λ is a Lagrange multiplier, to be optimised over. This reduces the optimisation to (42). Here t_j is the number of β_i taking the value γ_j, and $h_j = \mathsf{f}(\gamma_j) - \lambda \mathsf{g}(\gamma_j)$ for $j = 1, 2$. We relax the equation in (42) to an inequality and argue that $t_2 = 0$ in an optimal solution. At this point we can think of the optimisation as being over λ or, equivalently, over $\beta_* = \gamma_1$. Choosing the latter we end up with the optimisation problem (44).

We now have to optimise over ρ and find a bound on c that ensures that $a_{ij} = 1/k$ optimises the relaxed problem (19a)–(19c). Using the relaxation (44), we see that it is sufficient to satisfy (46). This leads to an inequality (48) for c. Making the right hand side of this inequality as small as possible leads to a univariate optimisation problem that is dealt with in Sections 3.5–3.8.

3.2 The subproblem corresponding to one row

Now, consider any *fixed* feasible values of the ϱ_i $(i \in [k])$ such that $\sum_{j=1}^{k} \varrho_i = \rho$. Then the problem decomposes into k independent maximisation subproblems. In this subsection we use Lagrange multipliers to perform the optimisation on these subproblems.

As already mentioned, when the ϱ_i are fixed (and hence ρ is fixed), the term $c \ln \left(1 - 2/k^{r-1} + \rho/k^{2r-2}\right)$ in the objective function z of (19) is constant. Hence we omit this term from the optimisation problems we consider until we once again allow ρ to vary, in Section 3.4.

We temporarily suppress the subscript i, to write $\boldsymbol{a} = (a_1, a_2, \ldots, a_k)$ for the ith row of A. For a fixed value of ϱ, the subproblem is then

$$\text{maximise } z_1^{(\varrho)} = -\sum_{j=1}^{k} a_j \ln a_j \tag{23a}$$

$$\text{subject to} \qquad \sum_{j=1}^{k} a_j^r = \frac{\varrho}{k^{r-2}}, \tag{23b}$$

$$\sum_{j=1}^{k} a_j = 1, \tag{23c}$$

$$a_j \geq 0. \tag{23d}$$

We assume that $1/k \leq \varrho \leq k^{r-2}$, so that the problem is feasible.

Remark 3.2. *When $\varrho = 1/k$ or $\varrho = k^{r-2}$ the optimization is trivial. If $\varrho = 1/k$ then there is a unique optimal solution, which satisfies $a_j = 1/k$ for all $j \in [k]$ and gives $z_1^{(\varrho)} = \ln k$. This is the value of ϱ that gives the global optimum to our problem. If $\varrho = k^{r-2}$ then there are k distinct optimal solutions, each with $a_j = 1$ for exactly one value of j, and $a_j = 0$ otherwise, each giving $z_1^{(\varrho)} = 0$. For ease of exposition, we include these cases in our argument below, though the analysis is unnecessary in these cases.*

Remark 3.3. *Applying Lemma 4.4 with $\ell = 1$ and $\widehat{\rho} = \varrho/k^{r-2}$ shows that $z_1^{(\varrho)}$ cannot be maximised at any point which lies in the boundary of the region determined by (23c)–(23d), unless $\varrho = k^{r-2}$. Thus we can omit the constraints (23d) in the optimisation of (23).*

Introducing the multiplier λ for (23b) and μ for (23c), the Lagrangian is

$$L(\boldsymbol{a}, \lambda, \mu) \;=\; -\sum_{j=1}^{k} a_j \ln a_j + \lambda \left(\sum_{j=1}^{k} a_j^r - \frac{\varrho}{k^{r-2}} \right) + \mu \left(\sum_{j=1}^{k} a_j - 1 \right). \tag{24}$$

The maximisation of L gives (23b) and (23c), together with the equations

$$\varphi(a_j) = 0 \qquad (j \in [k]), \quad \text{where} \quad \varphi(x) = \varphi_{\lambda,\mu}(x) = -1 - \ln x + \lambda r x^{r-1} + \mu. \tag{25}$$

If the equation $\varphi(x) = 0$ has only one root then all the a_j equal this root and hence, from (23c), $a_j = 1/k$ for all $j \in [k]$. In this case $\varrho = 1/k$. It follows from Lemma 4.3 that φ has at least one root.

Now suppose that the equation $\varphi(x) = 0$ has more than one root (that is, $1/k < \varrho \le k^{r-2}$), and let α be the largest. If $\boldsymbol{a}$ satisfies $a_j \ne \alpha$ for some $j \in [k]$ then subtracting the corresponding equations in (25) gives

$$\ln \alpha - \ln a_j - \lambda r(\alpha^{r-1} - a_j^{r-1}) = 0.$$

That is,

$$\lambda = \frac{\ln \alpha - \ln a_j}{r(\alpha^{r-1} - a_j^{r-1})} > 0. \tag{26}$$

Hence, since $-\ln x$ and x^{r-1} are both convex on $x > 0$ and λ is positive, $\varphi(x)$ is a strictly convex function. It follows that the equation $\varphi(x) = 0$ has at most two roots in $(0, \infty)$. Let the roots of $\varphi(x) = 0$ be α and β, where we assume that $\alpha > \beta$. We have $a_j \in \{\alpha, \beta\}$ for all $j \in [k]$. But we still need to determine how many of the a_j equal α and how many equal β.

Consider any stationary point $(\boldsymbol{a}_*, \lambda_*, \mu_*)$ of L. Then $\boldsymbol{a}_*$ and λ_* satisfy (26). Suppose without loss of generality that for some $1 \le t \le k - 1$ we have $a_1, \ldots, a_t = \alpha$, $a_{t+1}, \ldots, a_k = \beta$, where $\boldsymbol{a}_* = (a_1, \ldots, a_k)$. The Hessian $\boldsymbol{H} = \boldsymbol{H}_{\lambda_*, \mu_*}$ of the Lagrangian $L_{\lambda_*, \mu_*} = L(\,\cdot\,, \lambda_*, \mu_*)$, considered as a function of $\boldsymbol{a}$ only, is a $k \times k$ diagonal matrix with diagonal entries

$$h_{jj} = \begin{cases} \varphi'(\alpha) = -\frac{1}{\alpha} + \lambda r(r-1)\alpha^{r-2} & (j = 1, \ldots, t), \\ \varphi'(\beta) = -\frac{1}{\beta} + \lambda r(r-1)\beta^{r-2} & (j = t+1, \ldots, k). \end{cases}$$

Since φ is strictly convex with zeros $\beta < \alpha$, we know that $\varphi'(\beta) < 0 < \varphi'(\alpha)$. The quadratic form determined by the Hessian at $\boldsymbol{a}_*$ is

$$\boldsymbol{x}^T \boldsymbol{H} \boldsymbol{x} = \varphi'(\alpha) \sum_{j=1}^{t} x_j^2 + \varphi'(\beta) \sum_{j=t+1}^{k} x_j^2. \tag{27}$$

To determine the nature of the stationary point $\boldsymbol{a}_*$, we restrict the quadratic form to $\boldsymbol{x}$ lying in the tangent space at $\boldsymbol{a}_*$. This means that $\boldsymbol{x}$ satisfies linear equations determined by the gradient vectors of the constraint functions at $\boldsymbol{a}_*$. See, for example, [16]. In our case, these equations are

$$\alpha^{r-1} \sum_{j=1}^{t} x_j + \beta^{r-1} \sum_{j=t+1}^{k} x_j = 0,$$

$$\sum_{j=1}^{t} x_j + \sum_{j=t+1}^{k} x_j = 0.$$

These equations are linearly independent since $\alpha > \beta$. Rearranging these equations allows us to express x_1 and x_k in terms of $x_2, \ldots, x_{k-1}$ as follows:

$$x_1 = -\sum_{j=2}^{t} x_j, \qquad x_k = -\sum_{j=t+1}^{k-1} x_j.$$

Substituting these into (27) gives

$$\boldsymbol{x}^T \boldsymbol{H} \boldsymbol{x} = \varphi'(\alpha) \Big(\sum_{j=2}^{t} x_j \Big)^2 + \varphi'(\alpha) \sum_{j=2}^{t} x_j^2 + \varphi'(\beta) \sum_{j=t+1}^{k-1} x_j^2 + \varphi'(\beta) \Big(\sum_{j=t+1}^{k-1} x_j \Big)^2. \tag{28}$$

18

For $\boldsymbol{a}_*$ to be a strict local maximum, the right hand side of (28) must be negative for all x_2, x_3, ..., x_{k-1} such that $\boldsymbol{x} \neq \boldsymbol{0}$. Since $\varphi'(\alpha) > 0$, $\varphi'(\beta) < 0$, this will be true if and only if $t = 1$, when the terms with coefficient $\varphi'(\alpha)$ in (28) are absent. This local maximum is clearly unique up to the choice of $j \in [k]$ such that $a_j = \alpha$. Hence it is global, since $z_1^{(\varrho)}$ is bounded on the compact region determined by (23b)–(23d), and has no local maxima on the boundary when $\varrho < k^{r-2}$ (see Remark 3.3). Thus there are k global maxima, given by choosing $p \in [k]$ and setting $a_p = \alpha$, $a_j = \beta$ $(j \in [k], j \neq p)$, where (α, β) is the unique solution such that $\alpha \geq \beta \geq 0$ to the equations

$$\alpha^r + (k-1)\beta^r = k^{2-r}\varrho, \qquad \alpha + (k-1)\beta = 1. \tag{29}$$

The fact that there is at least one solution to these equations follows from Lemma 4.3. Next, note that the derivative of the function

$$(1 - (k-1)\beta)^r + (k-1)\beta^r$$

is zero at $\beta = 1/k$ and negative for $\beta \in (0, 1/k)$. Hence there can be at most one solution to these equations which satisfies $0 \leq \beta \leq 1/k$, or equivalently, $0 \leq \beta \leq \alpha$.

Note that the relaxation of the constraints (19e) proves to be unimportant, since the optimised values of the $a_{ij} \in \{\alpha, \beta\}$ are positive. Thus the optimisation (23) results in the system

$$\text{maximise } z_1^{(\varrho)} = -\alpha \ln \alpha - (k-1)\beta \ln \beta \tag{30a}$$
$$\text{subject to } \alpha^r + (k-1)\beta^r = k^{2-r}\varrho, \tag{30b}$$
$$\alpha + (k-1)\beta = 1, \tag{30c}$$
$$\beta \leq 1/k. \tag{30d}$$

We have omitted the constraint $0 \leq \beta$ here, but this will be enforced in any optimal solution since $z_1^{(\varrho)} = -\infty$ if $\beta < 0$. The maximisation problem is trivial since there is only one feasible solution which satisfies $0 \leq \beta \leq 1/k$, and no other feasible solution can be a maximum.

When $\varrho = 1/k$ we have $\alpha = \beta = 1/k$, while if $\varrho = k^{r-2}$ then $\alpha = 1$ and $\beta = 0$. When $1/k < \varrho < k^{r-2}$ we have $0 < \beta < 1/k < \alpha < 1$.

3.3 The combined problem, for a fixed value of ρ

We now combine these subproblems (one for each row) to give an optimisation problem corresponding to a *fixed* value of $\rho \in [1, k^{r-1}]$, as follows:

$$\text{maximise } z_2^{(\rho)} = -\sum_{i=1}^{k} \big(\alpha_i \ln \alpha_i + (k-1)\beta_i \ln \beta_i\big) \tag{31a}$$
$$\text{subject to } \sum_{i=1}^{k} \big(\alpha_i^r + (k-1)\beta_i^r\big) = k^{2-r}\rho, \tag{31b}$$
$$\alpha_i + (k-1)\beta_i = 1 \quad (i \in [k]), \tag{31c}$$
$$\beta_i \leq 1/k \quad (i \in [k]).$$

As before, the objective function ensures that $\beta_i \geq 0$ for $i \in [k]$ at any finite optimum. Recall from Remark 3.1 that $z_2^{(\rho)}$ has no maximum on the boundary when $\rho < k^{r-1}$.

For $\beta \in \mathbb{R}$, write

$$\mathsf{f}(\beta) \;=\; \ln k + \alpha \ln \alpha + (k-1)\beta \ln \beta\,, \quad \mathsf{g}(\beta) \;=\; \alpha^r + (k-1)\beta^r - \frac{1}{k^{r-1}}\,, \tag{32}$$

where α is defined as $1 - (k-1)\beta$ and hence $\mathrm{d}\alpha/\mathrm{d}\beta = -(k-1)$. We use the notation f and g here, and reserve the symbols f and g for transformed versions of these functions which will be introduced in Section 3.5.

Now $\mathsf{f}(\beta) = +\infty$ if $\beta < 0$ or $\beta > 1/(k-1)$. Also

$$\mathsf{f}(0) = \ln k, \qquad \mathsf{g}(0) = 1 - 1/k^{r-1} \quad \text{and} \quad \mathsf{f}(1/k) = \mathsf{g}(1/k) = 0.$$

Note further that

$$\mathsf{f}'(\beta) \;=\; -(k-1)(\ln \alpha - \ln \beta) < 0\,, \quad \mathsf{g}'(\beta) \;=\; -r(k-1)(\alpha^{r-1} - \beta^{r-1}) < 0 \quad (\beta \in [0, 1/k))\,,$$

so both $\mathsf{f}(\beta)$ and $\mathsf{g}(\beta)$ are positive and decreasing for $\beta \in [0, 1/k)$.

Letting $\widehat{z}_2^{(\rho)} = k \ln k - z_2^{(\rho)}$, (31) can now be rewritten as

$$\text{minimise } \widehat{z}_2^{(\rho)}(\boldsymbol{\beta}) \;=\; \sum_{i=1}^{k} \mathsf{f}(\beta_i) \tag{33a}$$

$$\text{subject to } \sum_{i=1}^{k} \mathsf{g}(\beta_i) \;=\; k^{2-r}(\rho - 1)\,, \tag{33b}$$

$$\beta_i \;\leq\; 1/k \qquad (i \in [k])\,. \tag{33c}$$

We remark that

$$\widehat{z}_2^{(\rho)}(\mathbf{0}) = k \ln k > 0 = \widehat{z}_2^{(\rho)}(1/k, \dots, 1/k)\,. \tag{34}$$

We therefore ignore $\boldsymbol{\beta} = \mathbf{0}$ in our search for the minimum in (33): see Remark 3.2.

We also ignore (33c) and apply the Lagrangian method to (33a) and (33b), using the multiplier $-\lambda$ for (33b). The Lagrangian optimisation will be to minimise the function

$$\psi^{(\rho)}(\boldsymbol{\beta}, \lambda) = \sum_{i=1}^{k} \mathsf{f}(\beta_i) - \lambda\left(\left(\sum_{i=1}^{k} \mathsf{g}(\beta_i)\right) - k^{2-r}(\rho - 1)\right). \tag{35}$$

The stationary points of the Lagrangian $\psi^{(\rho)}$ are given by (31b), (31c) and the equations

$$\frac{-1}{k-1}\frac{\partial \psi^{(\rho)}}{\partial \beta_i} \;=\; \frac{-1}{k-1}\big(\mathsf{f}'(\beta_i) - \lambda \mathsf{g}'(\beta_i)\big) \;=\; (\ln \alpha_i - \ln \beta_i) - \lambda r(\alpha_i^{r-1} - \beta_i^{r-1}) \;=\; 0 \qquad (i \in [k])\,. \tag{36}$$

We will concentrate on those stationary points of the Lagrangian that can give us the optimum for (33).

Let $B = [0, 1/k]$. We define

$$\eta(\beta) \;=\; \frac{\mathsf{f}(\beta)}{\mathsf{g}(\beta)}, \qquad \omega(\beta) \;=\; \frac{\mathsf{f}'(\beta)}{\mathsf{g}'(\beta)}$$

for $\beta \in [0, 1/k)$, and extend by continuity to give

$$\eta(1/k) = \omega(1/k) = \frac{k^{r-1}}{r(r-1)}\,.$$

20

Again, we reserve the notation η and ω for transformed versions of these functions, introduced in Section 3.5. (The values of $\eta(1/k)$ and $\omega(1/k)$ are established in Lemma 4.10, in terms of the transformed functions.)

Now suppose that $(\boldsymbol{\beta}_*, \lambda_*)$ is a stationary point of the Lagrangian $\psi^{(\rho)}$ which satisfies $\boldsymbol{\beta}_* = (\beta_{1*}, \ldots, \beta_{k*}) \in B^k$. Then $\widehat{z}_2^{(\rho)}(\boldsymbol{\beta}_*) = \psi^{(\rho)}(\boldsymbol{\beta}_*, \lambda_*)$.

Suppose that there exists i such that $\mathbf{0} < \beta_{i*} < 1/k$. Then $\beta_{i*} < 1/k < \alpha_{i*} = 1 - (k-1)\beta_{i*}$, and $(\beta, \lambda) = (\beta_{i*}, \lambda_*)$ must satisfy the equation

$$\lambda = \omega(\beta) = \frac{f'(\beta)}{g'(\beta)} = \frac{\ln \alpha - \ln \beta}{r(\alpha^{r-1} - \beta^{r-1})} > 0, \tag{37}$$

where $\alpha = 1 - (k-1)\beta$. This shows that $0 < \lambda_* = \omega(\boldsymbol{\beta}_*) < \infty$. Furthermore, (37) implies that $\beta_{j*} > 0$ for all j, since $\omega(0) = \infty$. Thus in any stationary point $(\boldsymbol{\beta}_*, \lambda_*)$ of $\psi^{(\rho)}$ with $\mathbf{0} \neq \boldsymbol{\beta}_* \in B^k$, for each $i \in [k]$, either $\beta_{i*} = 1/k$ (in which case $\alpha_{i*} = 1/k$ and (36) holds), or $\beta_{i*} \in (0, 1/k)$ and (β_{i*}, λ_*) is a solution to (37).

We now assume that $\boldsymbol{\beta}_* \neq (1/k, \ldots, 1/k)$ (see Remark 3.2) and rewrite (35) as

$$\psi^{(\rho)}(\boldsymbol{\beta}, \lambda) = \lambda k^{2-r}(\rho - 1) + \sum_{\beta_i \neq 1/k} g(\beta_i)(\eta(\beta_i) - \lambda). \tag{38}$$

First suppose that $\lambda_* > \max_B \eta(\beta)$. Since

$$\eta'(\beta) = \frac{g'(\beta)}{g(\beta)} \left(\omega(\beta) - \eta(\beta)\right)$$

for all $\beta \in (0, 1/k)$, it follows from (37) that $\eta'(\beta_{i*}) < 0$ for any $i \in [k]$ with $\beta_{i*} \neq 1/k$. But this shows that $(\boldsymbol{\beta}_*, \lambda_*)$ is not a local minimum of $\psi^{(\rho)}$, as we can decrease the value of $\psi^{(\rho)}$ by increasing β_{i*} infinitesimally, while holding all other values of β_{j*} and λ_* steady. Since we wish to minimise $\psi^{(\rho)}$ (and hence $\widehat{z}_2^{(\rho)}$), we now assume that $\lambda_* \leq \max_B \eta(\beta)$.

Next, suppose that $\lambda_* < \min_B \eta(\beta)$. As $\boldsymbol{\beta}_* \neq (1/k, \ldots, 1/k)$, by (38) we conclude that

$$\widehat{z}_2^{(\rho)}(\boldsymbol{\beta}_*) = \psi^{(\rho)}(\boldsymbol{\beta}_*, \lambda_*) > \lambda_* k^{2-r}(\rho - 1) \geq 0.$$

Hence $(\boldsymbol{\beta}_*, \lambda_*)$ cannot minimise $\widehat{z}_2^{(\rho)}$ if $\lambda_* < \min_{\beta \in B} \eta(\beta)$.

Therefore (see Remark 3.2) we may now assume that $(\boldsymbol{\beta}_*, \lambda_*)$ is a stationary point of (33) with

$$\boldsymbol{\beta}_* \in B^k \setminus \{\mathbf{0}, (1/k, \ldots, 1/k)\},$$

where $\lambda_* = \omega(\beta_{i*})$ for any i such that $\beta_{i*} \neq 1/k$, and such that $\lambda = \lambda_*$ satisfies

$$\min_{\beta \in B} \eta(\beta) \leq \lambda \leq \max_{\beta \in B} \eta(\beta). \tag{39}$$

We will prove the following in Section 3.5 below.

Lemma 3.1. *The function $\omega(\beta)$ has a unique minimum in $(0, 1/k)$. Furthermore, if (39) holds for some $\lambda > 0$ then the equation $\omega(\beta) = \lambda$ has at most two distinct solutions $\beta \in B$.*

21

Now consider the case that the equation $\omega(\beta) = \lambda$ has exactly two distinct roots $\gamma_1 > \gamma_2$. Define $\gamma_0 = 1/k$. Let t_i $(i = 1, 2)$ be the multiplicity of γ_i amongst the β_j $(j \in [k])$. We write f_i for $\mathsf{f}(\gamma_i)$ $(i = 0, 1, 2)$, and similarly for g, η and ω. For $i = 0, 1, 2$ we define $h_i = \mathsf{f}_i - \lambda \mathsf{g}_i$. Since $\gamma_1 > \gamma_2$ and $\mathsf{g}'(\beta) < 0$ for $\beta \in [0, 1/k)$, we have $\mathsf{g}_1 < \mathsf{g}_2$. Now ω is continuous on B and has a unique minimum in $(0, 1/k)$, by Lemma 3.1. Hence, for any β strictly between the two solutions γ_1, γ_2 of $\omega(\beta) = \lambda$, it follows that $\omega(\beta) < \lambda$. Therefore

$$h_1 - h_2 \;=\; (\mathsf{f}_1 - \lambda \mathsf{g}_1) - (\mathsf{f}_2 - \lambda \mathsf{g}_2) \;=\; \int_{\gamma_2}^{\gamma_1} \big(\mathsf{f}'(\beta) - \lambda \mathsf{g}'(\beta)\big)\mathrm{d}\beta \;<\; 0. \tag{40}$$

Hence $h_1 < h_2$. Also, as $\mathsf{f}_0 = \mathsf{g}_0 = 0$ we have

$$-h_1 \;=\; (\mathsf{f}_0 - \lambda \mathsf{g}_0) - (\mathsf{f}_1 - \lambda \mathsf{g}_1) \;=\; \int_{\gamma_1}^{1/k} \big(\mathsf{f}'(\beta) - \lambda \mathsf{g}'(\beta)\big)\mathrm{d}\beta \;>\; 0, \tag{41}$$

where the final inequality holds since $\omega(\beta) > \lambda$ for $\gamma_1 < \beta < 1/k$, by Lemma 3.1 (noting that γ_1 is the larger of the two solutions of $\omega(\beta) = \lambda$). Hence $h_1 < 0$. Now the minimum of (33) is bounded below by the solution of the following problem:

$$\begin{aligned} \text{minimise} \quad & t_1 h_1 + t_2 h_2 \\ \text{where} \quad & t_1 \mathsf{g}_1 + t_2 \mathsf{g}_2 = k^{2-r}(\rho - 1), \ \ t_1 + t_2 \le k, \ \ t_1, t_2 \in \mathbb{N}_0. \end{aligned} \tag{42}$$

We relax the equality constraint in (42) to give

$$\begin{aligned} \text{minimise} \quad & t_1 h_1 + t_2 h_2 \\ \text{where} \quad & t_1 \mathsf{g}_1 + t_2 \mathsf{g}_2 \le k^{2-r}(\rho - 1), \ \ t_1 + t_2 \le k, \ \ t_1, t_2 \in \mathbb{N}_0. \end{aligned} \tag{43}$$

It follows that we must have $t_2 = 0$ in the optimal solution to (43). To see this, suppose the optimal solution is $t_1 = \tau_1$, $t_2 = \tau_2 > 0$. Consider the solution $t_1 = \tau_1 + \tau_2$, $t_2 = 0$. This clearly satisfies the second and third constraint of (43). Since g is decreasing on $[0, 1/k]$ we have $\mathsf{g}_1 < \mathsf{g}_2$, so

$$(\tau_1 + \tau_2)\mathsf{g}_1 < \tau_1 \mathsf{g}_1 + \tau_2 \mathsf{g}_2 \le k^{2-r}(\rho - 1).$$

Hence the solution $t_1 = \tau_1 + \tau_2$, $t_2 = 0$ also satisfies the first constraint. Now $(\tau_1 + \tau_2)h_1 < \tau_1 h_1 + \tau_2 h_2$ by (40), contradicting the optimality of $t_1 = \tau_1, t_2 = \tau_2$. Therefore, we will simply write β_* for γ_1 and t for t_1 from this point.

The rest of the argument also holds when $\omega(\beta) = \lambda$ has only one solution β_*, so this case re-enters the argument now. By (41), we must choose t to be as large as possible subject to the constraints $t \le k$ and $t\,\mathsf{g}_1 \le k^{2-r}(\rho - 1)$. Therefore t must be the smaller of $\lfloor k^{2-r}(\rho - 1)/\mathsf{g}(\beta_*)\rfloor$ and k. We will usually relax the constraint $t \le k$ below, since we are mainly interested in small values of t. In any case, this relaxation can only worsen the objective function. Recalling that $z_2^{(\rho)} = k \ln k - \widehat{z}_2^{(\rho)}$, the objective function of the system (31) can be bounded above by

$$\max \ z_2^{(\rho)} \;\le\; k \ln k - t\,\mathsf{f}(\beta_*), \quad \text{where } t = \lfloor k^{2-r}(\rho - 1)/\mathsf{g}(\beta_*)\rfloor. \tag{44}$$

Note that λ has now been removed from consideration, as this upper bound on $z_2^{(\rho)}$ depends on $\beta_* = \beta_*(\rho)$ only.

3.4 Allowing ρ to vary

Here we return to the relaxed problem given by (19–19b), now allowing the value of ρ to vary.

As ρ increases from 1 to k^{r-1}, the bound in (44) changes only at integral values of $k^{2-r}(\rho-1)/\mathsf{g}(\beta_*)$. Thus the only relevant values of ρ of are those for which $k^{2-r}(\rho-1)/\mathsf{g}(\beta_*)$ is an integer. Then we may write (44) simply as

$$\max z_2^{(\rho)} \leq k \ln k - t\,\mathsf{f}(\beta_*), \quad \text{where } t = k^{2-r}(\rho-1)/\mathsf{g}(\beta_*), \; t \in \mathbb{N}_0. \tag{45}$$

Let $\boldsymbol{J}$ be the $k \times k$ matrix with all entries $1/k$, and note that $\boldsymbol{J}/k = \boldsymbol{J}_0$. We wish to find conditions on c which guarantee that $F(\boldsymbol{A}/k) < F(\boldsymbol{J}/k)$ for all $\boldsymbol{A} \neq \boldsymbol{J}$ which satisfy (19b), (19c). From the above, and (19), this will be true when

$$\frac{k \ln k - t\mathsf{f}(\beta_*)}{k} + c \ln\left(1 - \frac{2}{k^{r-1}} + \frac{\rho}{k^{2r-2}}\right) < \ln k + 2c \ln\left(1 - \frac{1}{k^{r-1}}\right), \tag{46}$$

that is, when

$$c \ln\left(1 + \frac{\rho - 1}{(k^{r-1} - 1)^2}\right) < \frac{t}{k}\,\mathsf{f}(\beta_*). \tag{47}$$

Next, from (45) we have $(\rho - 1) = t\,\mathsf{g}(\beta_*)k^{r-2}$. Substituting this into (47) gives

$$c \ln\left(1 + \frac{k^{r-1}}{(k^{r-1} - 1)^2}\frac{t\mathsf{g}(\beta_*)}{k}\right) < \frac{t\mathsf{g}(\beta_*)}{k}\,\eta(\beta_*). \tag{48}$$

Define

$$\vartheta(\beta) = \ln\left(1 + \frac{k^{r-1}}{(k^{r-1} - 1)^2}\frac{t\mathsf{g}(\beta)}{k}\right), \quad \text{and} \quad C(\beta) = \frac{(k^{r-1} - 1)^2}{k^{r-1}}\eta(\beta). \tag{49}$$

Then (48) can be written as

$$c < C(\beta_*)\frac{e^\vartheta - 1}{\vartheta} = C(\beta_*)\sum_{i=0}^{\infty}\frac{\vartheta^i}{(i+1)!}. \tag{50}$$

Now the right side of (50) is clearly minimised when ϑ is as small as possible. From (49), this is when t is as small as possible. If we relax the integrality constraint on t and allow $t \to 0$, then $\vartheta \to 0$ and (50) becomes $c < C(\beta_*)$. Thus we can estimate $c_{r,k}$ by minimising $C(\beta)$ over $\beta \in [0, 1/k]$. Then the computation of $c_{r,k}$ reduces to minimising the function

$$\eta(\beta) = \frac{\ln k + (1 - (k-1)\beta)\ln(1 - (k-1)\beta) + (k-1)\beta\ln\beta}{(1 - (k-1)\beta)^r + (k-1)\beta^r - 1/k^{r-1}} \quad (0 \leq \beta \leq 1/k). \tag{51}$$

Therefore we may take

$$c_{r,k} = \frac{(k^{r-1} - 1)^2}{k^{r-1}}\min_{\beta \in B}\eta(\beta). \tag{52}$$

Then, whenever $c \in (0, c_{r,k})$, we know that (50) holds, and hence that $\boldsymbol{J}$ is the unique maximum of F over all doubly stochastic matrices.

Remark 3.4. *We have taken $\vartheta = 0$ in (50), when the smallest value possible for ϑ is clearly larger. In (48), $t \in \mathbb{N}_0$ is the number of rows of $\boldsymbol{A}$ whose entries are not all $1/k$. We wish to estimate $c_{r,k}$, which is the largest value of c such that $\boldsymbol{A} \neq \boldsymbol{J}$, so we must clearly have $t \geq 1$. However, we*

23

cannot have $t = 1$, since (19c)–(19e) *imply that* $\boldsymbol{A}$ *cannot have a single row whose entries are not all* $1/k$. *Thus we may assume that* $t \geq 2$. *Since t should be as small as possible, we may take $t = 2$. Then* (48) *becomes*

$$c \leq \frac{2\mathsf{f}(\beta)}{k \ln\left(1 + 2k^{r-2}\mathsf{g}(\beta)/(k^{r-1} - 1)^2\right)}. \tag{53}$$

We could use (53) *directly to improve the estimate of $c_{r,k}$. This is done in [3] for $k = 2$, giving a small improvement in $c_{r,2}$, though [4] uses only* (52) *for $r = 2$. In the main, we will also use* (52), *which corresponds to allowing $t \to 0$. However, we show in Section 3.9 that the increment in $c_{r,k}$ which results from using* (53) *is small, and can be obtained indirectly from* (50).

Remark 3.5. *We might improve the estimate of $c_{r,k}$ further by avoiding the relaxation of* (19d) *in the optimisation. We note that taking $t = k$ in* (48) *results in a local maximum of* (19), *as follows. Let p be any permutation of $[k]$, and set $a_{ip(i)} = \alpha$, $a_{ij} = \beta$ $(j \neq p(i),\ i \in [k])$. This gives $k!$ local maxima of* (19). *We conjecture that these solutions are the global maxima, but we are unable to prove this. The inclusion of* (19d) *gives conditions for the local maxima which may have solutions yielding larger values of z in* (19). *The local maxima seem rather difficult to describe explicitly, so we leave this as an open question. However, we show in Section 3.9 that including* (19d) *cannot result in a large improvement in $c_{r,k}$.*

3.5 The univariate optimisation

We have now achieved the objective of reducing the problem to a univariate optimisation, namely, minimising the function η. To carry out this minimisation, we will first make a substitution $x = (k-1)\beta$ in (51), so that

$$\eta(x) = \eta\left(\frac{x}{k-1}\right) = \frac{\ln k - x\ln(k-1) + (1-x)\ln(1-x) + x\ln x}{(1-x)^r + x^r/(k-1)^{r-1} - 1/k^{r-1}} = \frac{f(x)}{g(x)} \quad (x \in [0, 1 - 1/k]), \tag{54}$$

where, using (32), we let

$$f(x) = \mathsf{f}\left(\frac{x}{k-1}\right) = \ln k - x\ln(k-1) + (1-x)\ln(1-x) + x\ln x,$$

$$g(x) = \mathsf{g}\left(\frac{x}{k-1}\right) = (1-x)^r + x^r/(k-1)^{r-1} - 1/k^{r-1}.$$

Figure 3.5 gives a plot of the function η when $k = 4$ and $r = 3$. Now

$$f(x) = \ln\left(k(1-x)\right) - x\ln\left((k-1)(1-x)/x\right)$$

for $x \in (0, 1 - 1/k)$, and at the boundaries we have $f(0) = \ln k$ and $f(1 - 1/k) = 0$. Differentiating gives

$$f'(x) = -\ln(k-1) - \ln(1-x) + \ln x \tag{55}$$
$$= -\ln\left((k-1)(1-x)/x\right) < 0 \quad \text{for } x \in (0, 1 - 1/k).$$

Therefore $f(x) > f(1 - 1/k) = 0$ for all $x \in (0, 1 - 1/k)$. Also $\lim_{x \to 0} f'(x) = -\infty$ while $f'(1 - 1/k) = 0$. Note, using (55), that

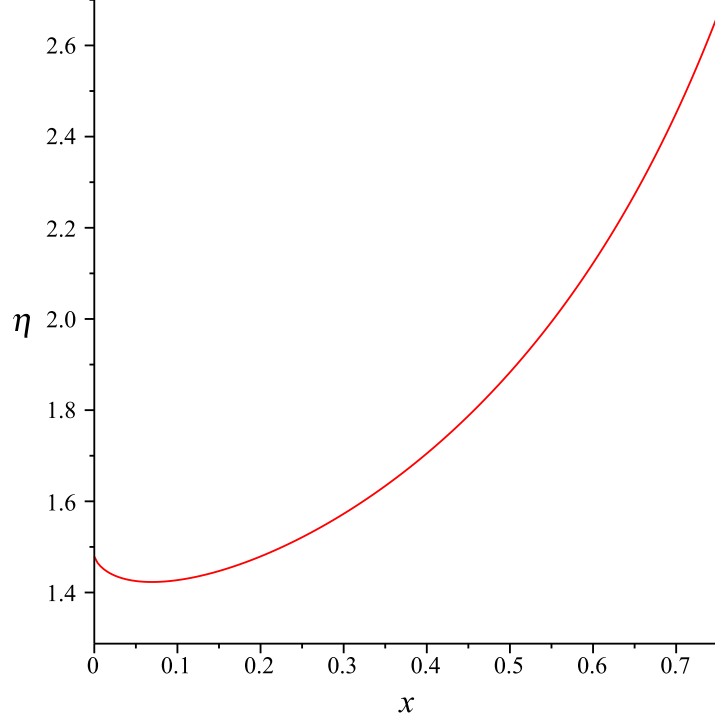$$f(x) = \ln\left(k(1-x)\right) + xf'(x). \tag{56}$$

Figure 1: The function η when $k = 4$ and $r = 3$.

Also,

$$f''(x) = \frac{1}{1-x} + \frac{1}{x} = \frac{1}{x(1-x)} > 0 \quad \text{for } x \in (0, 1 - 1/k], \tag{57}$$

$$f'''(x) = \frac{1}{(1-x)^2} - \frac{1}{x^2}. \tag{58}$$

We note that $f''(1 - 1/k) = k^2/(k-1)$ and $f'''(1 - 1/k) = k^3(k-2)/(k-1)^2$. Now we turn our attention to the function g, which satisfies $g(0) = 1 - 1/k^{r-1}$ and $g(1 - 1/k) = 0$. Differentiating gives

$$g'(x) = -r\big((1-x)^{r-1} - x^{r-1}/(k-1)^{r-1}\big) < 0 \quad \text{for } x \in (0, 1 - 1/k),$$

which shows that $g(x) > g(1 - 1/k) = 0$ for $x \in (0, 1 - 1/k)$. Also $g'(0) = -r$ and $g'(1 - 1/k) = 0$. Finally,

$$g''(x) = r(r-1)\big((1-x)^{r-2} + x^{r-2}/(k-1)^{r-1}\big) > 0 \quad \text{for } x \in (0, 1 - 1/k], \tag{59}$$

$$g'''(x) = -r(r-1)(r-2)\big((1-x)^{r-3} - x^{r-3}/(k-1)^{r-1}\big).$$

Note that, when $r = 2$, g'' is constant and g''' is identically zero. Also, in particular,

$$g''(1 - 1/k) = \frac{r(r-1)}{(k-1)k^{r-3}}, \quad g'''(1 - 1/k) = -\frac{r(r-1)(r-2)(k-2)}{(k-1)^2 k^{r-4}}.$$

Hence $f(x)$ and $g(x)$ are positive, strictly decreasing and strictly convex functions on $(0, 1 - 1/k)$.

25

Returning to the function η defined in (54), in Lemma 4.10 we show that

$$\lim_{x \to 1-1/k} \eta(x) = \frac{k^{r-1}}{r(r-1)}, \quad \lim_{x \to 0} \eta'(x) = -\infty, \quad \lim_{x \to 1-1/k} \eta'(x) = \frac{(k-2)k^r}{r(k-1)} \geq 0, \quad (60)$$

and we will take these limits as defining $\eta(1 - 1/k)$, $\eta'(0)$ and $\eta'(1 - 1/k)$, respectively. Note also that $\eta(0) = k^{r-1} \ln k / (k^{r-1} - 1)$.

If $k = 2$ then η has a stationary point at $x = 1 - 1/k = 1/2$. Otherwise, η has an interior minimum in $(0, 1 - 1/k)$, since $\eta'(0) < 0$ and $\eta'(1 - 1/k) > 0$. We first show that this is the unique stationary point of η in $(0, 1 - 1/k)$. This is not straightforward, since η is not convex, as observed in [4] for the case $r = 2$. Furthermore, the approach of [4], making a nonlinear substitution in η, does not generalise beyond $r = 2$. Hence our arguments here are very different from those in [4].

To determine the nature of the stationary points of η, we consider the function $h(x) = f(x) - \lambda g(x)$ on $(0, 1 - 1/k]$, for fixed $\lambda > 0$. Then h is analytic, and its zeros contain the points at which $\eta(x) = \lambda$ in $(0, 1 - 1/k)$. We will apply Rolle's Theorem [17] to h. The zeros of h are separated by zeros of h', and these are separated by zeros of h''. Since $f(1 - 1/k) = g(1 - 1/k) = 0$ and $f'(1 - 1/k) = g'(1 - 1/k) = 0$, we conclude that h' has a zero at $x = 1 - 1/k$ for all λ, and h has a double zero at $x = 1 - 1/k$. Now, from (57) and (59), the zeros of $h''(x) = f''(x) - \lambda g''(x)$ in $(0, 1 - 1/k]$ are the solutions of

$$x(1-x)^{r-1} + \frac{(1-x)x^{r-1}}{(k-1)^{r-1}} = \frac{1}{\lambda r(r-1)}. \quad (61)$$

In Lemma 4.11 we show that if $r \leq 2k$ then (61) has at most two solutions in $[0, 1]$, while if $r \geq 2k + 1$ then (61) has at most two solutions in $[0, 1 - 1/k]$ whenever $\lambda < \lambda_0$, where

$$\frac{1}{\lambda_0} = r(r-1) \left(\frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r} \right).$$

(Here, as elsewhere in the paper, we have $r, k \geq 2$.) For uniformity, we set $\lambda_0 = \infty$ if $r \leq 2k$ and define

$$\Lambda = \{x \in [0, 1 - 1/k] : \eta(x) < \lambda_0\}, \quad \Lambda' = \Lambda \cap (0, 1 - 1/k).$$

Then Λ' is a union of open intervals. We show in Lemma 4.13 that $\eta(0) < \eta(1 - 1/k) < \lambda_0$, which implies that $0, 1 - 1/k \in \Lambda$. Hence $\Lambda = \Lambda' \cup \{0, 1 - 1/k\}$, which shows that Λ is nonempty. Now $\eta'(0) < 0$, $\eta'(1 - 1/k) \geq 0$ imply that Λ' is nonempty. Our search for a value of x making η small will be restricted to Λ'. We have shown that h'' has at most two zeros in Λ, and hence h has at most four zeros in Λ. Since there is a double zero of h at $x = 1 - 1/k \in \Lambda \setminus \Lambda'$, it follows that there are at most two zeros of h in Λ'. Thus $\eta(x) = \lambda$ at most twice in Λ'. Since $\eta'(0) < 0$, $\eta'(1 - 1/k) \geq 0$, we know that η has a local minimum in Λ'. Then η has at most one local minimum $\xi \in \Lambda'$. To see this, suppose there are two local minima $\xi_1, \xi_2 \in \Lambda'$ with $\eta(\xi_1) \leq \eta(\xi_2) = \lambda < \lambda_0$. If $\eta(\xi_1) = \lambda$ then $\eta(x) = \lambda$ has at least four roots in Λ', with double roots at both ξ_1 and ξ_2. If $\eta(\xi_1) < \lambda$ then $\eta(x) = \lambda$ has at least three roots in Λ', with a double root at ξ_2 and, by continuity, a root strictly between ξ_1 and ξ_2. In either case, we have a contradiction. It also follows that Λ is connected. Otherwise, since $\eta'(0) < 0$, $\eta'(1 - 1/k) \geq 0$, each maximal interval of Λ' must contain a local minimum, a contradiction. Thus $\Lambda = [0, 1 - 1/k]$. In other words,

$$\eta(x) < \lambda_0 \quad \text{for all } x \in [0, 1 - 1/k]. \quad (62)$$

We have proved that η has exactly one local minimum in $(0, 1 - 1/k)$, and we will denote this minimum point by $\xi \in (0, 1 - 1/k)$. It also follows that there are no local maxima of η in $[0, 1 - 1/k]$,

as we now prove. If there were a local maximum $\xi' \in [0, \xi)$ then $\eta'(0) < 0$ would imply that there is a local minimum in $(0, \xi')$, a contradiction. The same argument applies to the interval $(\xi, 1 - 1/k]$, for $k > 2$. If $k = 2$, it is possible that $x = 1/2$ is a local maximum, but it still follows that there can be no local maximum in $(\xi, 1/2)$.

To summarise: if $k > 2$ then η has exactly one stationary point $\xi \in (0, 1 - 1/k)$, a local minimum. If $k = 2$ then there is a unique local minimum $\xi \in (0, 1/2]$ but, if $\xi \neq 1/2$, then $1/2$ may be a local maximum. In either case, ξ is the global minimum.

We now prove Lemma 3.1, using the same method but working with the transformed function ω defined by

$$\omega(x) = \omega\left(\frac{x}{k-1}\right) = \frac{\ln(k-1) + \ln(1-x) - \ln(x)}{r\left((1-x)^{r-1} - x^{r-1}/(k-1)^{r-1}\right)} = \frac{f'(x)}{g'(x)}.$$

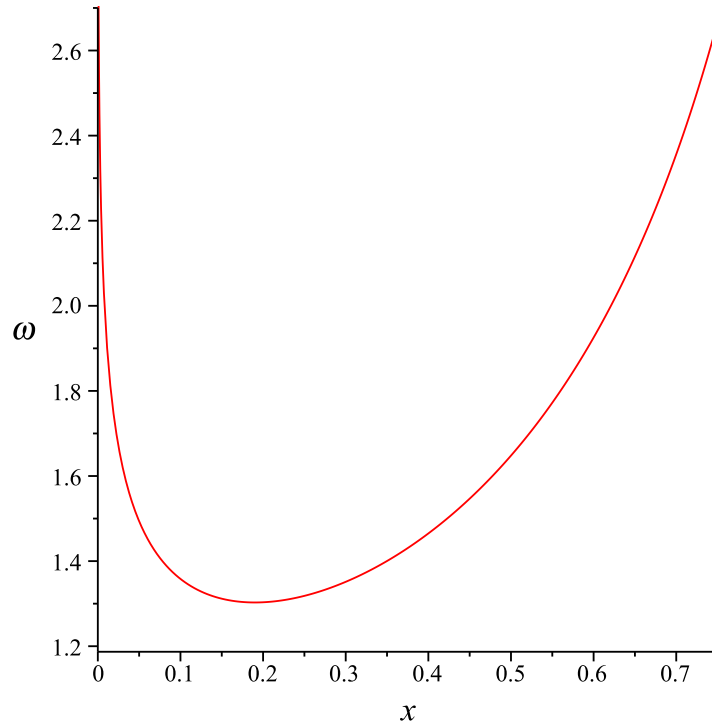Figure 3.5 gives a plot of the function ω when $k = 4$ and $r = 3$.

Figure 2: The function ω when $k = 4$ and $r = 3$.

Proof of Lemma 3.1. Let λ be a real number which satisfies (39). Note that the solutions to $\omega(\beta) = \lambda$ in (37) correspond to the zeros of $h'(x)$, where $h(x)$ is the function defined above. Combining (39) and (62), we see that $\lambda < \lambda_0$. Therefore by Lemma 4.11 and Lemma 4.13, we may conclude that $h''(x)$ has at most two zeros in $[0, 1 - 1/k]$. Hence $h'(x)$ has at most three zeros, and we know that $h'(1 - 1/k) = 0$. Thus there can be at most two zeros of $h'(x)$ in $[0, 1 - 1/k)$. Therefore $\omega(x) = f'(x)/g'(x)$ can take the value λ at most twice in $[0, 1 - 1/k)$. Since ω is analytic on $(0, 1 - 1/k)$, by the arguments above, ω can have at most one stationary point in $(0, 1 - 1/k)$.

Now $\omega(0) = +\infty$ since $f'(0) = -\infty$ and $g'(0) = -r$. By Lemma 4.10 $\omega(1 - 1/k) = \eta(1 - 1/k) = k^{r-1}/r(r-1)$ and that $\omega(\xi) = \eta(\xi)$, where ξ denotes the point which minimises η. Since $\omega(\xi) =$

$\eta(\xi) < +\infty = \omega(0)$ and $\omega(\xi) = \eta(\xi) < \eta(1 - 1/k) = \omega(1 - 1/k)$, ω must have a unique minimum in $(0, 1 - 1/k)$, completing the proof. $\qquad\square$

It remains to identify the local minimum ξ of η to a close enough approximation. Using (56), the condition that $\eta'(x) \le 0$ is

$$g'(x)\big(\ln(k(1-x)) + f'(x)(x - g(x)/g'(x))\big) \ge 0, \qquad x \in (0, 1 - 1/k).$$

We have shown that $f'(x) < 0$ and $g'(x) < 0$ for $x \in (0, 1 - 1/k)$, so the condition $\eta'(x) \le 0$ is equivalent to

$$x - \frac{g(x)}{g'(x)} \ge \frac{\ln\big(k(1-x)\big)}{-f'(x)} = \frac{\ln\big(k(1-x)\big)}{\ln\big((k-1)(1-x)/x\big)}, \qquad x \in (0, 1 - 1/k). \tag{63}$$

We will now use (63) to show that ξ is approximately $1/k^{r-1}$, except for the cases $k = 2$, $r = 3$, 4. (If $r = 2$ then $\xi = 1/k^{r-1}$ exactly.) This will enable us to determine the value of $c_{r,k}$ and establish that Lemma 2.2 holds.

3.6   The case $k = 2$

We will first examine the case $k = 2$ in more detail. We must determine whether $x = 1/2$ is a local minimum or maximum of η. If it is a local minimum, then it is the global minimum. Otherwise, there is a unique local minimum $\xi \in (0, 1/2)$. To resolve this, we must examine η in the neighbourhood of $x = 1/2$. We show in Lemma 4.14 that $1/2$ is a local minimum of η for $2 \le r \le 4$, but is a local maximum if $r \ge 5$. Thus, for $r = 2$, 3, 4, the global minimum is $\xi = 1/2$. (Note that we include the case $r = k = 2$ here, though ultimately it plays no part in our analysis.) Hence from (52) and (60) we have that for $r = 2$, 3, 4,

$$c_{r,2} = \frac{(2^{r-1} - 1)^2}{2^{r-1}} \frac{2^{r-1}}{r(r-1)} = \frac{(2^{r-1} - 1)^2}{r(r-1)}. \tag{64}$$

Specifically,
$$c_{2,2} = 1/2 = 0.5, \quad c_{3,2} = 3/2 = 1.5, \quad c_{4,2} = 49/12 \simeq 4.0833. \tag{65}$$

Now $u_{r,1} = 0$ for all r, and $u_{r,2} = 2^{r-1} \ln 2$, so

$$u_{2,2} \simeq 1.3863, \quad u_{3,2} \simeq 2.7726, \quad u_{4,2} \simeq 5.5452. \tag{66}$$

It follows that
$$u_{r,1} < c_{r,2} < u_{r,2} \quad \text{for } r = 2, 3, 4, \tag{67}$$

as required. (We cannot use this result in Theorem 1.1 when $k = r = 2$, since there is no sharp threshold in this case.)

In the cases $k = 2$, $r \ge 5$, there is a local minimum $\xi \in (0, 1/2)$, so the optimisation has similar characteristics to $k \ge 3$. We consider these cases in Section 3.8 below.

28

3.7 The case $r = 2$

We will consider the case $r = 2$ separately, since η can be minimised exactly in this case. The results given in this section were obtained by Achlioptas and Naor in [4], by making a nonlinear substitution in η. We can derive their results more simply, since we know that η has a unique minimum. We have

$$g(x) = (1-x)^2 + \frac{x^2}{k-1} - \frac{1}{k} = \frac{k}{k-1}\left(x - \frac{k-1}{k}\right)^2,$$

$$\text{so } g'(x) = \frac{2k}{k-1}\left(x - \frac{k-1}{k}\right).$$

It follows that

$$x - \frac{g(x)}{g'(x)} = \tfrac{1}{2}\big(x + (k-1)/k\big).$$

Hence (63) implies that x minimises η if and only if

$$x + \frac{k-1}{k} = \frac{2\ln\big(k(1-x)\big)}{\ln\big((k-1)(1-x)/x\big)}.$$

It is easily verified that $x = 1/k$ satisfies this equation, and hence is the unique minimum of η in $[0, 1 - 1/k]$.

We have dealt with the case $k = 2$ in the previous section, so we now assume that $k \geq 3$. Then

$$\min_{x \in [0, 1 - 1/k]} \eta(x) = \frac{f(1/k)}{g(1/k)} = \frac{(k-1)\ln(k-1)}{k-2}$$

and hence

$$c_{2,k} = \frac{(k-1)^2}{k}\eta(1/k) = \frac{(k-1)^3 \ln(k-1)}{k(k-2)}. \tag{68}$$

Now $(k-1)^3/k(k-2) = (k-1)(1 + 1/k(k-2))$ which lies strictly between $k-1$ and k, for $k \geq 3$. Thus

$$u_{2,k-1} = (k-1)\ln(k-1) < c_{2,k} < k\ln k = u_{2,k} \tag{69}$$

and moreover

$$c_{2,k} \leq \frac{(k-1)^2}{2}, \tag{70}$$

as required for Lemma 2.2.

3.8 The general case

We now consider the remaining cases $k \geq 3$ or $k = 2$, $r \geq 5$. We will do this by finding values $w, y \in (0, 1 - 1/k)$ such that $\eta'(w) \leq 0$ and $\eta'(y) > 0$. That is, w satisfies (63), but y does not. The uniqueness of ξ then implies that $w \leq \xi < y$, and we will use this to place a lower bound on $\eta(\xi)$. We will achieve this for all pairs r, k except for a small number, and we will solve these few remaining cases numerically.

To simplify the analysis, we will exclude some cases initially. Thus we assume below that

$$k = 2, \; r \geq 9 \quad \text{or} \quad k = 3, \; r \geq 4 \quad \text{or} \quad k \geq 4, \; r \geq 3. \tag{71}$$

By Lemma 4.15, the inequality

$$r^2(k+2)/k^r < 1 \tag{72}$$

holds whenever (71) holds.

First we set $x = w$ in (63), where $w = (k-1)/k^r$. Note that $w < 1/r^2$, from (72). Using Lemmas 4.5, 4.6 and 4.8, we have

$$\frac{r \ln\left(k(1-w)\right)}{\ln\left((k-1)(1-w)/w\right)} < \frac{r \ln k - rw}{r \ln k - 3w/2} = \frac{1 - w/\ln k}{1 - 3w/(2r \ln k)}$$

$$< \left(1 - \frac{w}{\ln k}\right)\left(1 + \frac{3w}{r \ln k}\right) < 1 - \left(1 - \frac{3}{r}\right)\frac{w}{\ln k} \le 1, \tag{73}$$

since $r \ge 3$.

Using Lemma 4.9, we have

$$g(w) = (1-w)^r + \frac{w^r}{(k-1)^{r-1}} - \frac{1}{k^{r-1}} \ge 1 - rw - \frac{kw}{k-1} = 1 - \frac{(k-1)r+k}{k-1}w,$$

$$\frac{-g'(w)}{r} = (1-w)^{r-1} - \frac{w^{r-1}}{(k-1)^{r-1}} \le (1-w)^{r-1} \le \frac{1}{1 + (r-1)w}.$$

So we have

$$rw - \frac{rg(w)}{g'(w)} \ge rw + \left(1 - \frac{(k-1)r+k}{k-1}w\right)\left(1 + (r-1)w\right)$$

$$> 1 + \left(r - \frac{2k-1}{k-1} - (r-1)\frac{k(r+1)}{k-1}\frac{k-1}{k^r}\right)w$$

$$> 1 + \left(r - \frac{2k-1}{k-1} - \frac{r^2}{k^{r-1}}\right)w,$$

$$= 1 + \left(r - 2 - \frac{1}{k-1} - \frac{r^2}{k^{r-1}}\right)w,$$

and the right hand side is bounded below by 1 whenever

$$\frac{1}{k-1} + \frac{r^2}{k^{r-1}} \le r - 2. \tag{74}$$

We may easily show that the left hand side of (74) is decreasing with r for $r \ge 3$, and it is clearly decreasing with $k \ge 2$. The right hand side is independent of k and increasing with r. Now (74) holds by calculation when $(k,r) \in \{(2,5), (3,4), (4,3)\}$. Therefore (74) holds for all (k,r) which satisfy (71), and combining this with (73) shows that w satisfies (63), as desired.

We now set $x = y$ in (63), where $y = (k+2)/k^r$. We have $ry < 1/r$ from (72). Then, using Lemmas 4.5 and 4.6, we have

$$\frac{r \ln\left(k(1-y)\right)}{\ln\left((k-1)(1-y)/y\right)} > \frac{\ln k^r - 3ry/2}{\ln k^r + \ln\left(1 - 3/(k+2)\right) + \ln(1-y)} > \frac{\ln k^r - 3ry/2}{\ln k^r - 3/(k+2)}$$

$$= 1 + \frac{3/(k+2) - 3ry/2}{\ln k^r - 3/(k+2)} > 1 + \frac{3/(k+2) - 3ry/2}{r \ln k}.$$

Using Lemma 4.9,

$$g(y) = (1-y)^r + \frac{y^r}{(k-1)^{r-1}} - \frac{1}{k^{r-1}} \le 1 - ry + \tfrac{1}{2}(ry)^2 + \frac{y^r}{(k-1)^{r-1}} - \frac{ky}{k+2}$$

30

$$= 1 - \left(r - \tfrac{1}{2}r^2 y - \frac{y^{r-1}}{(k-1)^{r-1}} + \frac{k}{k+2}\right)y,$$

$$\frac{-g'(y)}{r} = (1-y)^{r-1} - \frac{y^{r-1}}{(k-1)^{r-1}} \geq 1 - (r-1)y - \frac{y^{r-1}}{(k-1)^{r-1}}$$

$$= 1 - \left(r - 1 + \frac{y^{r-2}}{(k-1)^{r-1}}\right)y.$$

Now $y^{r-2}/(k-1)^{r-1} < 1$ for $r, k \geq 2$ and $ry < 1/r < 1/2$, using Lemma 4.8. Therefore

$$\frac{r}{-g'(y)} \leq 1 + \left(r - 1 + \frac{y^{r-2}}{(k-1)^{r-1}}\right)y + 2\left(r - 1 + \frac{y^{r-2}}{(k-1)^{r-1}}\right)^2 y^2$$

$$< 1 + \left(r - 1 + \frac{y^{r-2}}{(k-1)^{r-1}} + 2r^2 y\right)y.$$

Thus

$$ry - \frac{rg(y)}{g'(y)} \leq ry + \left(1 - \left(r - \tfrac{1}{2}r^2 y - \frac{y^{r-1}}{(k-1)^{r-1}} + \frac{k}{k+2}\right)y\right)\left(1 + \left(r - 1 + \frac{y^{r-2}}{(k-1)^{r-1}} + 2r^2 y\right)y\right)$$

$$\leq 1 + \left(r - 1 + \frac{5r^2 y}{2} + \frac{(1+y)y^{r-2}}{(k-1)^{r-1}} - \frac{k}{k+2}\right)y.$$

So $p'(y) > 0$ if y does not satisfy (63); that is, if

$$\frac{3/(k+2) - 3ry/2}{r \ln k} > \left(r - 1 + \frac{5r^2 y}{2} + \frac{(1+y)y^{r-2}}{(k-1)^{r-1}} - \frac{k}{k+2}\right)y.$$

Dividing by y and rearranging gives the equivalent condition

$$\frac{3k^r}{r(k+2)^2 \ln k} > r - 2 + \frac{2}{k+2} + \frac{3}{2\ln k} + \frac{5r^2 y}{2} + \frac{(1+y)y^{r-2}}{(k-1)^{r-1}}. \tag{75}$$

From Lemma 4.15, we have $r^2 y \leq 1$ and that $y = (r^2 y)/r^2$ is decreasing with both r and k. Since $y < 1$, it follows easily that $(1+y)y^{r-2}/(k-1)^{r-1}$ is decreasing with r and k. We may now check numerically that $(1+y)y^{r-2}/(k-1)^{r-1} \leq 1/50$ for all k, r satisfying (71). It follows that (75) is implied by the inequality

$$\frac{3k^r}{r^2(k+2)^2 \ln k} \geq 1 + \frac{0.52}{r} + \frac{2}{r(k+2)} + \frac{3}{2r \ln k}. \tag{76}$$

We show in Lemma 4.16 that, if (76) holds for some $r \geq 3$, $k \geq 2$, then it holds for any r', k' such that $r' \geq r$, $k' \geq k$. We may verify numerically that (76) holds for the following pairs r, k.

$$k = 2,\ r = 9, \quad k = 3,\ r = 6, \quad k = 4,\ r = 5, \quad k = 5,\ r = 4, \quad k = 15,\ r = 3.$$

Thus it holds for all pairs r, k such that

$$k = 2,\ r \geq 9, \quad k = 3,\ r \geq 6, \quad k = 4,\ r \geq 5, \quad k \in \{5, \dots, 14\},\ r \geq 4, \quad k \geq 15,\ r \geq 3.$$

Let us call these the pairs (k, r) *regular*, with the remaining nineteen pairs being *irregular*. We deal with the irregular pairs below by numerical methods.

31

First we continue our focus on regular pairs. For such pairs we have argued that $(k-1)/k^r \leq \xi < (k+2)/k^r$ and hence, using Lemmas 4.7 and 4.9,

$$f(\xi) = \ln k - \xi \ln(k-1) + \xi \ln \xi + (1-\xi)\ln(1-\xi) > \ln k - (r \ln k + 1)\xi,$$
$$g(\xi) = (1-\xi)^r + \xi^r/(k-1)^{r-1} - 1/k^{r-1} < 1/(1+r\xi).$$

Hence, using Lemma 4.17,

$$\eta(\xi) > \big(\ln k - (r \ln k + 1)\xi\big)(1 + r\xi) = \ln k - \xi - r(r \ln k + 1)\xi^2 \geq \ln k - 2\xi.$$

From (52), we can now determine

$$c_{r,k} \geq \frac{(k^{r-1}-1)^2}{k^{r-1}} \left(\ln k - \frac{2(k+2)}{k^r} \right) > (k^{r-1}-2)\ln(k-1),$$

using Lemma 4.18. Now $k^{r-1} > (k-1)^{r-1} + (r-1)(k-1)^{r-2} \geq (k-1)^{r-1} + 2$ for $r \geq 3$, $k \geq 2$, which shows that

$$u_{r,k-1} = (k-1)^{r-1}\ln(k-1) < c_{r,k} \tag{77}$$

for all regular pairs.

We also have

$$c_{r,k} < \frac{(k^{r-1}-1)^2}{k^{r-1}}\eta(0) = \frac{(k^{r-1}-1)^2}{k^{r-1}}\frac{k^{r-1}\ln k}{k^{r-1}-1} = (k^{r-1}-1)\ln k < k^{r-1}\ln k = u_{r,k}, \tag{78}$$

as required, and this holds for all $r, k \geq 2$.

Next we consider irregular pairs and use (63) to bound ξ numerically, by bisection. This is quite straightforward, since we know that $\xi \in (0, 1 - 1/k)$ is unique. The resulting values of $c_{r,k}$ are shown below, along with the corresponding values of $u_{r,k-1}$ and $u_{r,k}$.

k	r	$u_{r,k-1}$	$c_{r,k}$	$u_{r,k}$
2	5	0.0000	9.8771	11.0904
	6	0.0000	21.2990	22.1807
	7	0.0000	43.7678	44.3614
	8	0.0000	88.3486	88.7228
3	3	2.7726	8.1566	9.8875
	4	5.5452	27.9595	29.6625
	5	11.0904	87.4703	88.9876
4	3	9.8875	20.0491	22.1807
	4	29.6625	86.6829	88.7228
5	3	22.1807	37.8417	40.2359
6	3	40.2359	61.8958	64.5033
7	3	64.5033	92.5637	95.3496
8	3	95.3496	130.1457	133.0843
9	3	133.0843	174.9034	177.9752
10	3	177.9752	227.0688	230.2585
11	3	230.2585	286.8499	290.1453
12	3	290.1453	354.4353	357.8266
13	3	357.8266	429.9977	433.4764
14	3	433.4764	513.6960	517.2552

By inspection, $u_{r,k-1} < c_{r,k} < u_{r,k}$ for all irregular pairs.

We have already proved most of Lemma 2.2, and we complete the task below.

Proof of Lemma 2.2. The above analysis shows the existence of constants $c_{r,k}$ for all $r, k \geq 2$ such that F has a unique maximum at $\boldsymbol{J}$ whenever $c < c_{r,k}$. Combining the numerical results for irregular pairs with (67), (69), (77) and (78) shows that $c_{r,k} \in (u_{r,k-1}, u_{r,k})$ for all $r, k \geq 2$.

It remains to prove that for all $r, k \geq 2$ we have

$$c_{r,k} \leq \frac{(k^{r-1} - 1)^2}{r(r-1)}.$$

This follows from (64) if $k = 2$ and $r = 2, 3, 4$, or from (70) if $r = 2$ and $k \geq 3$. In all other cases we have $c_{r,k} < (k^{r-1} - 1) \ln k$, from (78). Furthermore, it follows from Lemma 4.20 that $r(r-1) \ln k/(k^{r-1} - 1) < 1$ whenever $k \geq 3$, $r \geq 2$, or $k = 2$, $r \geq 5$. This completes the proof of Lemma 2.2. $\qquad\square$

Combining this result with the conclusion of Section 2, we see that Theorem 1.1 is established.

3.9 Asymptotics

We have given precise bounds on $c_{r,k}$, but if we require only asymptotic estimates as $r \to \infty$ and/or $k \to \infty$ then the following simplified analysis suffices.

Remark 3.6. *When we write "$r \to \infty$ and/or $k \to \infty$", this is not to be interpreted as "$r(n) \to \infty$ and/or $k(n) \to \infty$", but merely as "r and/or k are arbitrarily large constants". Otherwise, we cannot use Theorem 1.3 to establish the existence of a sharp threshold between $c_{r,k}$ and $u_{r,k}$. This is the approach to asymptotic estimates taken, for example, in [3].*

We will use (48) to improve the estimate of $c_{r,k}$ asymptotically, as discussed in Remarks 3.4 and 3.5. First let us consider the maximum possible improvement that we might be able to achieve.

From Remark 3.5, we know that the maximum value of z in (19) cannot be smaller than that given by taking $t = k$ in (48). Thus we may bound the possible increase in $c_{r,k}$ as follows. Since $g(\beta) \leq 1 - 1/k^{r-1}$ and $t \leq k$, it follows from (49), using Lemma 4.5, that $\vartheta \leq 1/(k^{r-1} - 1)$ in (50). Thus $\sum_{i=0}^{\infty} \vartheta^i/(i+1)! \leq 1 + \vartheta/2 + O(\vartheta^2)$ in (50). Therefore we can increase $c_{r,k}$ asymptotically by a factor at most $1 + 1/(2k^{r-1}) + O(1/k^{2r-2})$. Since $c_{r,k} < u_{r,k} = k^{r-1} \ln k$, the additive improvement to $c_{r,k}$ from fully optimising (19) is at most $1/2 \ln k + O(\ln k/k^{r-1})$. Hence we cannot improve $c_{r,k}$ asymptotically by more than an additive term $1/2 \ln k$.

Now let us consider what improvement we can rigorously justify. From Remark 3.4, we know that we can take $t = 2$ in (48). Let $\kappa = 4(k-1)/k^r$, and $\mathcal{R} = \{x \in \mathbb{R} : 1/k^r \leq x \leq \kappa\}$. We proved in Sections 3.6 – 3.8 that the minimum of $\eta(x)$ for $x \in [0, 1 - 1/k]$ lies in $[(k-1)/k^r, (k+2)/k^r] \subset \mathcal{R}$ for all $r, k \geq 2$. However, all we require here is the fact that η has a unique minimum in $[0, 1 - 1/k]$, as shown in Section 3.5. Now we may approximate

$$\begin{aligned}
f(x) &= \ln k - x \ln(k-1) + x \ln x - x + O(x^2), \\
g(x) &= 1 - rx - 1/k^{r-1} + O(r^2 x^2).
\end{aligned}$$

Hence, using Lemma 4.8, and noting that $-\ln x = O(r \ln k)$ since $x \geq 1/k^r$, in $\mathcal{R}$ we have

$$
\begin{aligned}
\eta(x) &= \big(\ln k - x \ln(k-1) + x \ln x - x + O(x^2) \big)\big(1 + rx + 1/k^{r-1} + O(r^2 x^2) \big) \\
&= (1 + 1/k^{r-1}) \ln k + x(r \ln k - \ln(k-1) - 1) + x \ln x + O(r^2 x^2 \ln k) \,.
\end{aligned}
$$

Therefore, let φ be the function defined by

$$
\varphi(x) = (1 + 1/k^{r-1}) \ln k + x(r \ln k - \ln(k-1) - 1) + x \ln x \,.
$$

We have seen that φ approximates η. Now

$$
\begin{aligned}
\varphi'(x) &= (r \ln k - \ln(k-1) - 1) + 1 + \ln x = \ln x - \ln\big((k-1)/k^r\big) \,, \\
\varphi''(x) &= 1/x > 0 \,.
\end{aligned}
$$

Thus $\varphi(x)$ is minimised at $\hat{\xi} = (k-1)/k^r \in \mathcal{R}$, as expected. We can write

$$
\varphi(x) = (1 + 1/k^{r-1}) \ln k - x + x \ln(x/\hat{\xi}) \,. \tag{79}
$$

In particular,

$$
\varphi(\hat{\xi}) = (1 + 1/k^{r-1}) \ln k - \hat{\xi} \,.
$$

Hence, reinstating the error term in (51), we may take

$$
\begin{aligned}
c_{r,k} &= \frac{(k^{r-1} - 1)^2}{k^{r-1}} \left(\left(1 + \frac{1}{k^{r-1}}\right) \ln k - \frac{k-1}{k^r} - O\Big(\frac{r^2 \ln k}{k^{2r-2}}\Big) \right) \\
&= (k^{r-1} - 1) \ln k - \frac{k-1}{k} - O\Big(\frac{r^2 \ln k}{k^{r-1}}\Big) \,. \tag{80}
\end{aligned}
$$

Since $\kappa = 4\hat{\xi}$, using (79) we have,

$$
\varphi(\kappa) - \varphi(\hat{\xi}) = (\hat{\xi} - \kappa) + \kappa \ln(\kappa/\hat{\xi}) = -3\hat{\xi} + 4\hat{\xi} \ln 4 > 2.5\,\hat{\xi} = 2.5\,(k-1)/k^r \,.
$$

Therefore, since η has a unique minimum in $[0, 1 - 1/k]$, we have

$$
\begin{aligned}
\eta(x) &\geq \varphi(\hat{\xi}) - O\big(r^2 \ln k/k^{2r-2}\big) && (x \leq \kappa) \\
\omega(x) &\geq \varphi(\hat{\xi}) + 2.5(k-1)/k^r - O\big(r^2 \ln k/k^{2r-2}\big) && (x \geq \kappa) \,.
\end{aligned}
$$

We have $g(x) = 1 - O(r/k^{r-1})$ when $x \leq \kappa$, and hence $\vartheta = 2/k^r - O(r/k^{2r-1})$, taking $t = 2$ in (49). Thus the factor $(e^\vartheta - 1)/\vartheta$ in (50) is $1 + 1/k^r - O(r/k^{2r-1})$. This is effectively the maximum value of $(e^\vartheta - 1)/\vartheta$ for $x \in [0, 1 - 1/k]$ and $(e^\vartheta - 1)/\vartheta$ is effectively constant for $x \leq \kappa$. Thus

$$
\begin{aligned}
\min_{x \leq \kappa} \left(\eta(x) \frac{e^\vartheta - 1}{\vartheta} \right) &\geq \varphi(\hat{\xi})(1 + 1/k^r) - O(r^2 \ln k/k^{2r-2}) \\
&= \varphi(\hat{\xi}) + \ln k/k^r - O(r^2 \ln k/k^{2r-2}) \,,
\end{aligned}
$$

since $\varphi(\hat{\xi}) = \ln k + O(\ln k/k^{r-1})$. Also, $\vartheta > 0$ in $[0, 1 - 1/k]$, so

$$
\min_{x > \kappa} \left(\eta(x) \frac{e^\vartheta - 1}{\vartheta} \right) \geq \varphi(\hat{\xi}) + 2.5(k-1)/k^r - O\big(r^2 \ln k/k^{2r-2}\big) \tag{81}
$$

34

$$> \varphi(\hat{\xi}) + \ln k / k^r - O(r^2 \ln k / k^{2r-2}) \, ,$$

for any $k, r \geq 2$, provided k^r is large enough. Thus, after multiplying the right side of (81) by $(k^{r-1} - 1)^2 / k^{r-1}$, the additive improvement in $c_{r,k}$ is $\ln k / k - O(r^2 \ln k / k^{r-1})$. Applying this to (80), we have

$$c_{r,k} = k^{r-1} \ln k - \frac{k-1}{k}(1 + \ln k) - O\Big(\frac{r^2 \ln k}{k^{r-1}}\Big). \tag{82}$$

Substituting $k = 2$ in (82),

$$c_{r,2} = 2^{r-1} \ln 2 - 1/2(1 + \ln 2) - O(r^2 / 2^r) \, ,$$

the result obtained by Achlioptas and Moore [3] for 2-colouring r-uniform hypergraphs. The case $r = 2$ (colouring random graphs), studied by Achlioptas and Naor [4], is discussed further below.

Remark 3.7. *The best lower bound on $u_{r,k}$ is $\widetilde{u}_{r,k} = u_{r,k} - 1/2 \ln k$ from Remark 2.1, so there is a gap*

$$\widetilde{u}_{r,k} - c_{r,k} = \frac{k-1}{k} + \frac{k-2}{2k} \ln k + O\Big(\frac{r^2 \ln k}{k^{r-1}}\Big).$$

Asymptotically, this gap is always nonzero, though extremely small compared to $c_{r,k}$ or $u_{r,k}$. It is independent of r (up to the error term), and grows slowly with k. It is minimised when $k = 2$ and $r \to \infty$. The existence of this gap merely indicates that the second moment method is not powerful enough to pinpoint the sharp threshold. We know from Theorem 1.3 that the threshold lies in $[c_{r,k}, \widetilde{u}_{r,k}]$, although it is possible that it does not converge to a constant as $n \to \infty$. Note that if we could obtain the maximum possible correction $1/2 \ln k$, as discussed above, then the gap would be approximately $(k-1)/k$, and hence uniformly bounded for all $k, r \geq 2$ except $k = r = 2$.

Observe that the asymptotic estimate of $c_{r,k}$ given in (82) is not sharp in one case, namely when $r = 2$ and $k \to \infty$. Here the error in (82) is $O(\ln k / k)$, so we have not improved (80). Since this is the important case of colouring random graphs, we will examine it separately.

From (68), we know that the bound on $c_{2,k}$ from minimising η is precisely

$$\frac{(k-1)^3}{k(k-2)} \ln(k-1) = k \ln k - \frac{k-1}{k}(1 + \ln k) - \frac{1}{2k} - O\Big(\frac{\ln k}{k^2}\Big). \tag{83}$$

The right side of (82) is $\varphi(\hat{\xi}) + O(\ln k / k)$, so (81) still implies that, when k is large enough, we need only consider $\vartheta(x)$ for $x \in \mathcal{R}$. It follows, as above, that the factor $(e^\vartheta - 1)/\vartheta = 1 + 1/k^2 - O(1/k^3)$. Thus the additive improvement in $c_{2,k}$ is $\ln k / k - O(\ln k / k^2)$. Adding this to (83), we have

$$c_{2,k} = k \ln k - \frac{k-2}{k} \ln k + \frac{2k-1}{2k} - O\Big(\frac{\ln k}{k^2}\Big), \tag{84}$$

which marginally improves (69) asymptotically. Note that, taken together, (82) and (84) exhaust the possibilities for the manner in which r and/or k can grow large.

References

[1] D. ACHLIOPTAS AND E. FRIEDGUT, *A sharp threshold for k-colorability*, Random Struct. Algorithms, 14 (1999), pp. 63–70.

[2] D. Achlioptas, J. H. Kim, M. Krivelevich, and P. Tetali, *Two-coloring random hypergraphs*, Random Struct. Algorithms, 20 (2002), pp. 249–259.

[3] D. Achlioptas and C. Moore, *Random k-SAT: Two moments suffice to cross a sharp threshold*, SIAM J. Comput., 36 (2006), pp. 740–762.

[4] D. Achlioptas and A. Naor, *The two possible values of the chromatic number of a random graph*, Ann. of Math., 162 (2005), pp. 1333–1349.

[5] A. Coja-Oghlan, *Upper-bounding the k-colorability threshold by counting covers*, Electron. J. Combin., 20(3) (2013), P32.

[6] A. Coja-Oghlan, C. Efthymiou and S. Hetterich, *On the chromatic number of random regular graphs*, arXiv preprint, `arxiv:1308:4287`, 2013.

[7] A. Coja-Oghlan and D. Vilenchik, *Chasing the k-colorability threshold*, in Proc. 54th Annual IEEE Symposium on Foundations of Computer Science, IEEE, 2013, pp. 380–389. Full preprint at `arxiv:1304:1063`, April 2013.

[8] A. Coja-Oghlan and L. Zdeborová, *The condensation transition in random hypergraph 2-coloring*, in Proc. 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2012, pp. 241–250.

[9] P. Flajolet, D. E. Knuth and B. Pittel, *The first cycles in an evolving graph*, Discrete Math., 75 (1989), pp. 167–215.

[10] C. Greenhill, S. Janson, and A. Ruciński, *On the number of perfect matchings in random lifts*, Comb. Probab. Comput., 19 (2010), pp. 791–817.

[11] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 2nd ed., 1988.

[12] H. Hatami and M. Molloy, *Sharp thresholds for constraint satisfaction problems and homomorphisms*, Random Struct. Algorithms, 33 (2008), pp. 310–332.

[13] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.

[14] M. Krivelevich and B. Sudakov, *The chromatic numbers of random hypergraphs*, Random Struct. Algorithms, 12 (1998), pp. 381–403.

[15] A. Kupavskii and D. Shabanov, *On r-colorability of random hypergraphs*, arXiv preprint, `arXiv:1110.1249`, 2011.

[16] P. Shutler, *Constrained critical points*, Amer. Math. Monthly, 102 (1995), pp. 49–52.

[17] M. Spivak, *Calculus*, Cambridge University Press, 3rd ed., 2006.

4 Appendix: Technical lemmas

Lemma 4.1. $G \in \mathcal{G}^*(n, r, cn)$ *has at least* $(k-1)$ *isolated vertices* a.a.s..

Proof. Define $m = \lfloor cn \rfloor$ and let $Y(\boldsymbol{v})$ be the number of isolated vertices in G, determined by $\boldsymbol{v}$. The mr entries of $\boldsymbol{v}$ are uniform on $[n]$, from which it follows that $\mathbf{E}[Y] = n(1 - 1/n)^{mr} \sim ne^{-cr}$. Also, the entries of $\boldsymbol{v}$ are independent, and arbitrarily changing any single entry can only change $Y(\boldsymbol{v})$ by ± 1. Thus we may apply a standard martingale inequality [13, Corollary 2.27] to give

$$\Pr\left(Y \leq \tfrac{1}{2}ne^{-cr}\right) \leq e^{-ne^{-2cr}/12cr},$$

for large n. Thus G has $\Omega(n)$ isolated vertices a.a.s., from which the result follows easily. □

Lemma 4.2. *Suppose that M is a $p \times p$ matrix of $q \times q$ blocks, such that*

$$M = \begin{bmatrix} 2B & B & \cdots & B \\ B & 2B & \cdots & B \\ B & B & \ddots & B \\ B & B & \cdots & 2B \end{bmatrix} \in \mathbb{R}^{pq \times pq}, \quad \text{where} \quad B = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ 1 & 1 & \ddots & 1 \\ 1 & 1 & \cdots & 2 \end{bmatrix} \in \mathbb{R}^{q \times q}.$$

Then $\det(M) = (p+1)^q(q+1)^p$.

Proof. We have, by adding and subtracting rows and columns of M,

$$\det M = \det \begin{bmatrix} 2B & -B & \cdots & -B \\ B & B & \cdots & 0 \\ B & 0 & \ddots & 0 \\ B & 0 & \cdots & B \end{bmatrix} = \det \begin{bmatrix} (p+1)B & 0 & \cdots & 0 \\ B & B & \cdots & 0 \\ B & 0 & \ddots & 0 \\ B & 0 & \cdots & B \end{bmatrix}$$

$$= \det\big((p+1)B\big)(\det B)^{p-1} = (p+1)^q(\det B)^q.$$

We can use the same transformations to compute $\det B$, replacing B by the 1×1 unit matrix in the argument. We obtain $\det B = (q+1)\, 1^{q-1} = q+1$. Hence $\det M = (p+1)^q(q+1)^p$. □

(We are grateful to Brendan McKay for pointing out that $(p+1)^q(q+1)^p$ is the number of spanning trees in the complete bipartite graph $K_{p+1,q+1}$. This suggests that an alternative proof of the above lemma may be possible using Kirchhoff's Matrix Tree Theorem, but we do not explore this here.)

Lemma 4.3. *If $\rho \in [1, k^{r-1}]$ then the system defined by (19b)–(19e) is feasible.*

Proof. Firstly, note that the system (19c)–(19e) defines a convex set. The $k \times k$ matrix $\boldsymbol{J}$ with all entries equal to $1/k$ is feasible when $\rho = 1$, while any $k \times k$ permutation matrix $\boldsymbol{\Delta}$ is feasible when $\rho = k^{r-1}$. Now define the $k \times k$ matrices $\boldsymbol{A}(\epsilon) = (1-\epsilon)\boldsymbol{J}_0 + \epsilon\boldsymbol{\Delta}$ for all $\epsilon \in [0,1]$. Then $\boldsymbol{A}(\epsilon)$ satisfies (19c)–(19e) by convexity, while (19b) becomes

$$\Psi(\epsilon) = k^{r-2}\sum_{i=1}^{k}\sum_{j=1}^{k}\big((1-\epsilon)\boldsymbol{J}_{ij} + \epsilon\boldsymbol{\Delta}_{ij}\big)^r = \rho.$$

Now $\Psi(\epsilon)$ is a polynomial function of ϵ, and hence continuous. Also $\Psi(0) = 1$ and $\Psi(1) = k^{r-1}$. Therefore, by the Intermediate Value Theorem, for any $\rho \in [1, k^{r-1}]$ there is some $\epsilon^* \in [0,1]$ such that $\Psi(\epsilon^*) = \rho$, and hence $\boldsymbol{A}(\epsilon^*)$ is a feasible solution. □

Lemma 4.4. *Let $\ell \in \{1, 2, \ldots, k\}$ and let $\widehat{\rho} \le \ell$ be a fixed positive constant. Consider the maximisation problem*

$$\text{maximise } z(\boldsymbol{a}) = -\frac{1}{k} \sum_{i=1}^{\ell} \sum_{j=1}^{k} a_{ij} \ln a_{ij} \tag{85a}$$

$$\text{subject to} \qquad \sum_{i=1}^{\ell} \sum_{j=1}^{k} a_{ij}^r = \widehat{\rho}, \tag{85b}$$

$$\sum_{j=1}^{k} a_{ij} = 1 \qquad (i \in [\ell]), \tag{85c}$$

$$a_{ij} \ge 0 \qquad (i \in [\ell], j \in [k]). \tag{85d}$$

If $\widehat{\rho} < \ell$ then no boundary point of (85c)–(85d) *can be a local maximum of z.*

Proof. Suppose that $\boldsymbol{b}$ is a local optimum to (85) on the boundary of (85c)–(85d). Without loss of generality, we may assume that $1 \ge b_{i1} \ge b_{i2} \ge \cdots \ge b_{ik} \ge 0$ for all $i \in [\ell]$. If $\widehat{\rho} < \ell$ then there exist $i \in [\ell]$, $j_1, j_2 \in [k]$ such that $j_1 \ne j_2$ and $0 < b_{ij_1}, b_{ij_2} < 1$. Without loss of generality, suppose that $i = 1$, $j_1 = 1, j_2 = 2$. Since $\boldsymbol{b}$ is on the boundary, there exist $t \in [\ell], j \in [k]$ such that $b_{tj} = 0$. Without loss of generality, we may assume that $t \in \{1, 2\}$ and $j = k$.

Let $\mathcal{S}$ denote the region given by (85c)–(85d), such that $a_{ij} = b_{ij}$ for $(i, j) \notin \{(1, 1), (1, 2), (1, k)\}$ if $t = 1$, and for $(i, j) \notin \{(1, 1), (1, 2), (2, 1), (2, k)\}$ if $t = 2$. Then let $\mathcal{S}'$ be the subset of $\mathcal{S}$ determined by (85b), and let $\mathcal{S}^o$ denote the interior of $\mathcal{S}$. Let $\Phi(\boldsymbol{a}) = \sum_{i=1}^{\ell} \sum_{j=1}^{k} a_{ij}^r$. Thus (85b) is equivalent to $\Phi(\boldsymbol{a}) = \Phi(\boldsymbol{b}) = \widehat{\rho}$.

At the point $\boldsymbol{b}$, note that $\partial z / \partial a_{ij}$ is finite for all $b_{ij} > 0$ and $+\infty$ for all $b_{ij} = 0$. Thus, for all small enough $\delta > 0$, there is a ball B with centre $\boldsymbol{b}$ and radius δ, such that $z(\boldsymbol{a}) > z(\boldsymbol{b})$ for every point $\boldsymbol{a} \in B'$, where $B' = B \cap \mathcal{S}^o$. Note that B' is a convex set. So, we need only show that there is a point in $S' \cap B'$, since this will contradict the assumption that $\boldsymbol{b}$ is a local maximum of z.

Let us write the points in $\mathcal{S}$ as $\boldsymbol{u} = (a_{11}, a_{12}, a_{1k})$ if $t = 1$, or as $\boldsymbol{u} = (a_{11}, a_{12}, a_{21}, a_{2k})$ if $t = 2$. Let

$$\boldsymbol{u}_0 = (b_{11} + \theta - \theta^3, b_{12} - \theta, \theta^3), \qquad \boldsymbol{u}_1 = (b_{11} - \theta, b_{12}, \theta) \qquad \text{if } t = 1,$$
$$\boldsymbol{u}_0 = (b_{11} + \theta, b_{12} - \theta, b_{21} - \theta^3, \theta^3), \qquad \boldsymbol{u}_1 = (b_{11}, b_{12}, b_{21} - \theta, \theta) \qquad \text{if } t = 2.$$

Then $\boldsymbol{u}_i \in \mathcal{S}^o$ and $\|\boldsymbol{u}_i - \boldsymbol{b}\| \le 3\theta$ for $\theta \in (0, 1)$ and $i = 1, 2$. Thus, for small enough θ, $\boldsymbol{u}_i \in B'$ ($i = 1, 2$). Also

$$\Phi(\boldsymbol{u}_0) - \Phi(\boldsymbol{b}) = r(b_{11}^{r-1} - b_{12}^{r-1})\theta + \tfrac{1}{2}r(r-1)(b_{11}^{r-2} + b_{12}^{r-2})\theta^2 + O(\theta^3) > 0 \quad (t = 1, 2)$$

for small enough (positive) θ, since $b_{11} \ge b_{12} > 0$ and $r \ge 2$. Thus $\Phi(\boldsymbol{u}_0) > \Phi(\boldsymbol{b}) = \widehat{\rho}$. Similarly

$$\Phi(\boldsymbol{u}_1) - \Phi(\boldsymbol{b}) = -rb_{t1}^{r-1}\theta + O(\theta^2) < 0 \quad (t = 1, 2),$$

for small enough θ, since $b_{t1} > 0$ ($t = 1, 2$) and $r \ge 2$. Thus $\Phi(\boldsymbol{u}_1) < \Phi(\boldsymbol{b}) = \widehat{\rho}$.

Now consider the points $\boldsymbol{u}_\epsilon = (1 - \epsilon)\boldsymbol{u}_0 + \epsilon\boldsymbol{u}_1$, for $\epsilon \in [0, 1]$. By convexity, $\boldsymbol{u}_\epsilon \in B'$ for all $\epsilon \in [0, 1]$. Also $\Phi(\boldsymbol{u}_\epsilon)$ is a polynomial function of ϵ with $\Phi(\boldsymbol{u}_0) > \widehat{\rho}$ and $\Phi(\boldsymbol{u}_1) < \widehat{\rho}$. Hence, by the Intermediate Value Theorem, there exists $\epsilon^* \in [0, 1]$ such that $\Phi(\boldsymbol{u}_{\epsilon^*}) = \widehat{\rho}$. Then $\boldsymbol{u}_{\epsilon^*}$ is the required point in $S' \cap B'$. $\square$

Lemma 4.5. $\ln(1+z) \le z$ for all $z > -1$.

Proof. Let $\phi(z) = z - \ln(1+z)$, which is strictly convex on $z > -1$, since $\ln(1+z)$ is strictly concave. Also $\phi'(z) = 1 - 1/(1+z)$, so ϕ is stationary at $z = 0$, and this must be its unique minimum. Since $\phi(0) = 0$, we have $\phi(z) \ge 0$ for all $z > -1$, and $\phi(z) > 0$ if $z \ne 0$. $\qquad\square$

Lemma 4.6. $\ln(1-z) \ge -3z/2$ for all $0 \le z \le 1/2$.

Proof. Let $\phi(z) = \ln(1-z) + 3z/2$. Then ϕ is strictly concave on $[0, 1)$, since $\ln(1-z)$ is strictly concave. Also $\phi'(z) = -1/(1-z) + 3/2$, so ϕ is stationary at $z = 1/3$, and this must be its unique maximum. Now $\phi(0) = 0$, and we may calculate $\phi(1/2) > 0$, so $\phi(z) > 0$ for $0 < z \le 1/2$. $\qquad\square$

Lemma 4.7. For all $z \in (0, 1)$, $(1-z)\ln(1-z) > -z$ and $(1-\frac{1}{2}z)\ln(1-z) < -z$.

Proof. We have

$$(1-z)\ln(1-z) = -z + \sum_{i=2}^{\infty} \frac{z^i}{i(i-1)} > -z,$$

$$(1-\tfrac{1}{2}z)\ln(1-z) = -z - \sum_{i=3}^{\infty} \frac{(i-2)z^i}{2i(i-1)} < -z. \qquad\square$$

Lemma 4.8. $1 + z \le 1/(1-z) \le 1 + z + 2z^2 \le 1 + 2z$ for all $0 \le z \le 1/2$.

Proof. The first inequality is equivalent to $z^2 \ge 0$ if $z < 1$. The second inequality is equivalent to $z \le 1/2$. The third follows trivially from the second. $\qquad\square$

Lemma 4.9. For $p \in \mathbb{N}$, $z \in [0, 1]$, $1 - pz \le (1-z)^p \le 1 - pz + \frac{1}{2}(pz)^2$. Also $(1-z)^p \le 1/(1+pz)$.

Proof. Let $\phi_1(z) = (1-z)^p - 1 + pz$. Then $\phi_1(0) = 0$ and $\phi_1'(z) = p(1 - (1-z)^{p-1}) \ge 0$ if $z \in [0, 1]$, giving the first inequality. Let $\phi_2(z) = 1 - pz + \frac{1}{2}(pz)^2 - (1-z)^p$. Then $\phi_2(0) = 0$ and $\phi_2'(z) = -p + p^2 z + p(1-z)^{p-1} \ge -p + p^2 z + p(1-(p-1)z) = pz \ge 0$, by the first inequality, giving the second. For the third inequality, using Lemma 4.5, we have $(1-z)^p \le e^{-pz} = 1/e^{pz} \le 1/(1+pz)$. $\qquad\square$

Lemma 4.10. Let $\eta(x) = f(x)/g(x)$ for $x \in [0, 1 - 1/k)$, where

$$f(x) = \ln k - x\ln(k-1) + (1-x)\ln(1-x) + x\ln x, \qquad g(x) = (1-x)^r + x^r/(k-1)^{r-1} - 1/k^{r-1}.$$

Then

$$\eta(0) = \frac{k^{r-1}\ln k}{k^{r-1} - 1}, \qquad \lim_{x\to 1-1/k}\eta(x) = \lim_{x\to 1-1/k}\omega(x) = \frac{k^{r-1}}{r(r-1)},$$

$$\lim_{x\to 0}\eta'(x) = -\infty, \qquad \lim_{x\to 1-1/k}\eta'(x) = \frac{(k-2)k^r}{r(k-1)}.$$

Furthermore, if $\eta'(x) = 0$ and $g(x) \ne 0$ then $\omega(x) = \eta(x)$.

Proof. The stated value of $\eta(0)$ follows from the definition. Recall the calculations of Section 3.5. Using L'Hôpital's rule [17],

$$\lim_{x\to 1-1/k}\eta(x) = \left[\frac{f''(x)}{g''(x)}\right]_{1-1/k} = \frac{k^{r-1}}{r(r-1)}. \tag{86}$$

39

The same calculations prove that $\lim_{x\to 1-1/k} \omega(x)$ also takes this value. Next,

$$\eta'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{g(x)^2} = \frac{f'(x) - \eta(x)\,g'(x)}{g(x)}. \tag{87}$$

As $x \to 0$, all quantities in (87) are finite, except $f'(x) \to -\infty$. Since $g(0) > 0$, we have $\eta'(x) \to -\infty$ as $x \to 0$. Note also that the last statement of the lemma follows from (87).

For the final calculation note that when $x = 1 - 1/k$, the numerator and denominator in the expression for $\eta'(x)$ are both zero. Hence applying L'Hôpital's rule again gives

$$\begin{aligned}
\lim_{x\to 1-1/k} \eta'(x) &= \lim_{x\to 1-1/k} \frac{f'(x) - \eta(x)\,g'(x)}{g(x)} = \lim_{x\to 1-1/k} \frac{f'(x) - \eta''(1 - 1/k)\,g'(x)}{g(x)} \\
&= \lim_{x\to 1-1/k} \frac{f'''(x) - \eta(1 - 1/k)\,g'''(x)}{g''(x)} \\
&= \frac{(k-2)k^r}{r(k-1)},
\end{aligned}$$

using the values of $f''(1 - 1/k)$, $f'''(1 - 1/k)$, $g''(1 - 1/k)$ and $g'''(1 - 1/k)$ calculated in Section 3.5. $\qquad\square$

Lemma 4.11. *Let $k \geq 2$ and $r \geq 2$ be integers, and let λ be a positive real number. If $r \leq 2k$ then the equation*

$$x(1-x)^{r-1} + \frac{(1-x)x^{r-1}}{(k-1)^{r-1}} = \frac{1}{\lambda r(r-1)}$$

has at most two solutions for x in $[0, 1 - 1/k]$. Otherwise $r \geq 2k + 1$ and the above equation has at most two solutions for x in $[0, 1 - 1/k]$ whenever $\lambda < \lambda_0$, where

$$\frac{1}{\lambda_0} = r(r-1)\left(\frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r}\right).$$

Proof. Let $\theta(x) = x(1-x)^{r-1}$ and define κ, ℓ by $1/\kappa = (k-1)^{r-1}$ and $1/\ell = \lambda r(r-1)$. We wish to investigate the number of solutions of $\phi(x) = \ell$, where

$$\phi(x) = \theta(x) + \kappa\theta(1-x).$$

Differentiating gives

$$\phi'(x) = \theta'(x) - \kappa\theta'(1-x), \qquad\qquad \phi''(x) = \theta''(x) + \kappa\theta''(1-x).$$

Thus the stationary points of ϕ are the solutions of $\theta'(x) = \kappa\theta'(1-x)$. We may calculate

$$\begin{aligned}
\theta'(x) &= (1-x)^{r-2}(1-rx), & \theta'(1-x) &= -x^{r-2}\big((r-1) - rx\big), \\
\theta''(x) &= -(r-1)(1-x)^{r-3}(2-rx), & \theta''(1-x) &= (r-1)x^{r-3}\big((r-2) - rx\big).
\end{aligned}$$

We summarise the behaviour of ϕ in $[0, 1]$ in the following table. Here $\downarrow$ means "decreasing", $\uparrow$ means "increasing". The final column gives the maximum number of stationary points of ϕ in the corresponding subinterval of $[0, 1]$.

$x \in [0, 1/r)$	$\theta'(x) > 0, \ \theta'(1-x) < 0$	$\phi(x) \uparrow$	0
$x \in [1/r, 2/r)$	$\theta''(x) \le 0, \ \theta''(1-x) > 0$	$\theta'(x) \downarrow \ \kappa\theta'(1-x) \uparrow$	1
$x \in (2/r, 1 - 2/r]$	$\theta''(x) > 0, \ \theta''(1-x) \ge 0$	$\phi(x)$ strictly convex	1
$x \in (1 - 2/r, 1 - 1/r]$	$\theta''(x) > 0, \ \theta''(1-x) \le 0$	$\theta'(x) \uparrow \ \kappa\theta'(1-x) \downarrow$	1
$x \in (1 - 1/r, 1]$	$\theta'(x) < 0, \ \theta'(1-x) > 0$	$\phi(x) \downarrow$	0

Now ϕ is analytic on $[0,1]$, with $\phi'(0) = 1$ and $\phi'(1) = -\kappa$. Therefore ϕ' changes sign an odd number of times in $[0,1]$, which implies that ϕ has an odd number of stationary points in $[0,1]$. From the table it follows that ϕ has either one or three stationary points. Hence $\phi(x) = \ell$ has at most four solutions in $[0,1]$, for any ℓ.

We first consider small values of r. When $r = 2$ the union of the first and last subinterval is $[0,1] \setminus \{1/2\}$, which contains no stationary point. Hence ϕ has at most one stationary point in $[0,1]$ (and it can only occur at $x = 1/2$). When $r = 3$ the union of the first, second and last subinterval equals $[0,1] \setminus \{2/3\}$ and contains at most one stationary point. Hence ϕ has at most two stationary points in $[0,1]$.

When $r = 4$, the central subinterval is empty, so ϕ has at most two stationary points in $[0,1]$. However, we know that an even number of stationary points is impossible, from above. Therefore when $r = 2, 3, 4$ the function ϕ has at most one stationary point in $[0,1]$, and hence at most two solutions to $\phi(x) = \ell$ in $[0,1]$, for any fixed ℓ.

Next we assume that $r \ge 5$, which implies that all five subintervals are nonempty. Either ϕ has one stationary point which is a local maximum, or it has three stationary points: a local maximum μ_1, a local minimum μ_2, and a local maximum μ_3, with $\mu_1 < \mu_2 < \mu_3$.

Let
$$L_1 = \sup\{\phi(y) : y \in [1/r, 2/r)\}, \qquad L_2 = \sup\{\phi(z) : z \in (1 - 2/r, 1 - 1/k].$$

(We take $L_2 = -\infty$ if there is only one stationary point.) First we show that
$$L_1 \ge L_2. \tag{88}$$

We readily see that
$$L_1 \ge \phi(1/r) > \theta(1/r) = \frac{(r-1)^{r-1}}{r^r}.$$

Next we calculate an upper bound on L_2 by considering two cases. First, if $2 \le r \le k$ then
$$L_2 \le \theta(1 - 2/r) + \kappa\theta(1/r) = \frac{1}{r^r}\left((r-2)2^{r-1} + \left(\frac{r-1}{k-1}\right)^{r-1}\right)$$
$$\le \frac{(r-2)2^{r-1} + 1}{r^r} \le \frac{(r-2)2^{r-1} + (r/2)^r}{r^r},$$

since $\theta(1-x)$ is maximised at $x = 1 - 1/r$ in $[0,1]$. Next, if $r > k \ge 2$ then $\theta(1-x)$ is maximised when $x = 1/k$ in $[0, 1 - 1/k]$. Therefore
$$L_2 \le \theta(1 - 2/r) + \kappa\theta(1/k) = \frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r} \le \frac{(r-2)2^{r-1} + (r/2)^r}{r^r}. \tag{89}$$

Thus (88) holds if $(r-2)2^{r-1} + (r/2)^r < (r-1)^{r-1}$. We show in Lemma 4.12 that this is true for all $r \ge 5$, so (88) holds for $r \ge 5$. Now ϕ has at least one local maximum, so we have established that ϕ has a local maximum $\mu_1 \in [1/r, 2/r)$ whenever $r \ge 5$.

41

We now consider whether ϕ has a local minimum $\mu_2 \in (2/r, 1-2/r)$. Since there is a local maximum $\mu_1 \in [1/r, 2/r)$ we know that $\phi'(2/r) < 0$. Thus ϕ has a local minimum $\mu_2 \in (2/r, 1 - 2/r]$ if and only if $\phi'(1 - 2/r) > 0$. Now

$$\phi'(1 - 2/r) = \theta'(1 - 2/r) - \kappa\theta'(2/r) = -\frac{r-3}{r^{r-2}}\left(2^{r-2} - \frac{1}{(r-3)(k-1)}\left(\frac{r-2}{k-1}\right)^{r-2}\right).$$

This expression is certainly nonpositive if $2 \geq (r-2)/(k-1)$; that is, if $r \leq 2k$. So, if $r \leq 2k$, there is no local minimum in $(2/r, 1 - 2/r)$ and it follows that μ_1 is the only stationary point of ϕ. In this case, the equation $\phi(x) = \ell$ has at most two solutions on $[0, 1]$.

When $r \geq 2k + 1 \geq 5$ we know that $\phi(x) = \ell$ has at most two solutions for all $\ell > L_2$, using (88). From (89) we have

$$L_2 < \frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r}.$$

Substituting $\ell = (\lambda r(r-1))^{-1}$ we find that $\phi(x) = \ell$ has at most two solutions in $[0, 1 - 1/k]$ so long as

$$\lambda > \frac{1}{r(r-1)L_2} > \frac{1}{r(r-1)}\left(\frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r}\right)^{-1} = \lambda_0. \qquad \square$$

Lemma 4.12. *For all $r \geq 5$ the inequality $(r-2)2^{r-1} + (r/2)^r < (r-1)^{r-1}$ holds.*

Proof. We will show $(r-2)2^{r-1} < (r-1)^{r-1}/2$ and $(r/2)^r < (r-1)^{r-1}/2$.

To show $(r-2)2^{r-1} < (r-1)^{r-1}/2$, let $\gamma_1(r) = 2(r-2)2^{r-1}/(r-1)^{r-1}$. Then

$$\frac{\gamma_1(r+1)}{\gamma_1(r)} = \frac{2}{r-2}\left(\frac{r-1}{r}\right)^r < 1$$

if $r \geq 4$. Thus $\gamma_1(r)$ is decreasing for $r \geq 4$. Since $\gamma_1(5) = 3/8 < 1$, the inequality follows.

To show $(r/2)^r < (r-1)^{r-1}/2$, let $\gamma_2(r) = r^r/(2r-2)^{r-1}$. Then

$$\frac{\gamma_2(r+1)}{\gamma_2(r)} = \frac{r+1}{2r-2}\left(\frac{r^2-1}{r^2}\right)^r \leq \left(\frac{r^2-1}{r^2}\right)^r < 1$$

if $r \geq 4$. Thus $\gamma_2(r)$ is decreasing for $r \geq 4$. Since $\gamma_2(5) = 5^5/2^{12} < 1$, the inequality follows. $\square$

Lemma 4.13. *For $k \geq 2$ and $r \geq 2k + 1$, we have*

$$\eta(0) < \eta(1 - 1/k) < \lambda_0.$$

(The values of $\eta(0)$ and $\eta(1 - 1/k)$ are stated in Lemma 4.10 while λ_0 is defined in Lemma 4.11.)

Proof. The left hand inequality reduces to $r(r-1)\ln k/(k^{r-1} - 1) < 1$. In Lemma 4.20 we show that $r(r-1)\ln k/(k^{r-1} - 1) < 1$ for all $k \geq 3$, $r \geq 2$, or $k = 2$, $r \geq 5$. Clearly this includes all $r \geq 2k + 1$, and so establishes the left hand inequality.

The right hand inequality is

$$\frac{(r-2)2^{r-1}}{r^r} + \frac{1}{k^r} < \frac{1}{k^{r-1}},$$

which is equivalent to $\gamma(r, k) < 1$, where

$$\gamma(r, k) = \frac{r-2}{2k-2}\left(\frac{2k}{r}\right)^r.$$

For fixed $k \geq 2$, if $r > 2k$ then

$$\frac{\gamma(r+1,k)}{\gamma(r,k)} = \frac{2k(r-1)r^r}{(r-2)(r+1)^{r+1}} \leq \frac{2kr(r-1)}{(r-2)(r+1)^2} \leq \frac{2k}{r} < 1,$$

if $r^2(r-1) \leq (r-2)(r+1)^2$. This is equivalent to $r^2 - 3r - 2 \geq 0$, which is true for all $r \geq 4$. Thus $\gamma(r,k)$ is decreasing in r, so we need only establish the critical case $r = 2k+1$. We have

$$\gamma(2k+1,k) = \frac{2k-1}{2k-2}\left(\frac{2k}{2k+1}\right)^{2k+1} \leq \frac{2k-1}{2k-2}\left(\frac{2k}{2k+1}\right)^2 < 1,$$

if $(2k-2)(2k+1)^2 - (2k-1)(2k)^2 > 0$, which is $2k^2 - 3k - 1 > 0$. This holds for all $k \geq 2$. $\quad\square$

Lemma 4.14. *If $k = 2$ then $x = 1/2$ is a local minimum of η for $r = 2, 3, 4$, and a local maximum if $r \geq 5$.*

Proof. We have

$$\eta(x) = \frac{\ln 2 + x \ln x + (1-x)\ln(1-x)}{(1-x)^r + x^r - 1/2^{r-1}}.$$

Substituting $x = (1-z)/2$, we find

$$\frac{\eta(z)}{2^{r-1}} = \frac{(1-z)\ln(1-z) + (1+z)\ln(1+z)}{(1-z)^r + (1+z)^r - 2}.$$

We may compute Taylor expansions, giving

$$\frac{r(r-1)\eta(z)}{2^{r-1}} = \frac{z^2 + z^4/6 + O(z^6)}{z^2 + (r-2)(r-3)z^4/12 + O(z^6)}$$

$$= 1 + \frac{2 - (r-2)(r-3)}{12}z^2 + O(z^4).$$

If $r = 2, 3$ then the coefficient of z^2 is positive, so $z = 0$ is a local minimum. If $r \geq 5$ then the coefficient of z^2 is negative, so $z = 0$ is a local maximum. However, if $r = 4$, the coefficient of z^2 is zero, so we need a higher order approximation. We compute

$$\frac{3\eta(z)}{2} = \frac{z^2 + z^4/6 + z^6/15 + O(z^8)}{z^2 + z^4/6} = \frac{1 + z^2/6 + z^4/15 + O(z^6)}{1 + z^2/6} = 1 + \frac{z^4}{15} + O(z^6).$$

The coefficient of z^4 is positive, and hence $z = 0$ is a local minimum. $\quad\square$

Lemma 4.15. *The function $r^2(k+2)/k^r$ is decreasing in both r and k for all $r \geq 3$, $k \geq 2$. Hence $r^2(k+2)/k^r < 1$ if*

$$k = 2, \; r \geq 9, \quad k = 3, \; r \geq 4, \quad k \geq 4, \; r \geq 3.$$

Proof. Let $\phi(r,k) = r^2(k+2)/k^r$. Then

$$\frac{\phi(r+1,k)}{\phi(r,k)} = \frac{(r+1)^2}{kr^2} < 1,$$

if $k \geq (1 + 1/r)^2$. Since $(1 + 1/r)^2 \leq 16/9$ for $r \geq 3$, this is satisfied for all $k \geq 2$. Also

$$\frac{\phi(r,k+1)}{\phi(r,k)} = \frac{(k+3)k^r}{(k+2)(k+1)^r} < \frac{(k+3)k}{(k+2)(k+1)} = \frac{k^2 + 3k}{k^2 + 3k + 2} < 1.$$

We can now check numerically that $r^2(k+2)/k^r < 1$ for $(k,r) \in \{(2,9), (3,4), (4,3)\}$. $\quad\square$

Lemma 4.16. *If the inequality*

$$\frac{3k^r}{r^2(k+2)^2\ln k} \geq 1 + \frac{0.52}{r} + \frac{2}{r(k+2)} + \frac{3}{2r\ln k}$$

holds for some (r,k) with $r \geq 3$, $k \geq 2$, then it holds for all (r',k') such that $r' \geq r$, $k' \geq k$.

Proof. The right side of this inequality is decreasing with r and k, so it suffices to show that the function $\phi(r,k)$ on the left side is increasing. This follows since, if $k \geq 2$, $r \geq 3$,

$$\frac{\phi(r+1,k)}{\phi(r,k)} = \frac{kr^2}{(r+1)^2} \geq 1.$$

Also, if $r \geq 3$ then

$$\frac{\phi(r,k+1)}{\phi(r,k)} = \frac{(k+1)^r(k+2)^2\ln k}{k^r(k+3)^2\ln(k+1)} \geq \frac{(k+1)^3(k+2)^2\ln k}{k^3(k+3)^2\ln(k+1)} > \frac{(k+1)\ln k}{k\ln(k+1)},$$

since $(k+1)(k+2) > k(k+3)$ for all $k \geq 0$. Now we will have

$$\frac{(k+1)\ln k}{k\ln(k+1)} = \frac{(k+1)/\ln(k+1)}{k/\ln k} > 1$$

if the function $\gamma(x) = x/\ln x$ is increasing for $x \geq k$. Since $\gamma'(x) = (\ln x - 1)/(\ln x)^2 > 0$ for $x > e$, we have $\phi(r,k+1)/\phi(r,k) > 1$ for $k \geq 3$. For $k = 2$ and $r \geq 3$, we may verify that

$$\frac{\phi(r,3)}{\phi(r,2)} = \frac{16\ln 2}{25\ln 3}\left(\frac{3}{2}\right)^r \geq \frac{54\ln 2}{25\ln 3} > 1.$$

Thus $\phi(r,k)$ is increasing in k and r for all $k \geq 2$, $r \geq 3$, and the conclusion follows. $\square$

Lemma 4.17. *For all regular pairs, $r(r\ln k + 1)\xi \leq 1$.*

Proof. We have $\xi \leq (k+2)/k^r$ for all regular pairs. Thus the inequality is true if $\phi(r,k) \leq 1$, where $\phi(r,k) = r(r\ln k + 1)(k+2)/k^r$. Now

$$\phi(r,k) = \frac{r(r\ln k + 1)}{k^2} \cdot \frac{k+2}{k^{r-2}} = r\left(\frac{r\ln k}{k^2} + \frac{1}{k^2}\right)\left(\frac{1}{k^{r-3}} + \frac{2}{k^{r-2}}\right)$$

is decreasing with $k \geq 2$ for all $r \geq 3$, since $\ln k/k^2$ is decreasing for $k \geq 2$. Also

$$\frac{\phi(r+1,k)}{\phi(r,k)} = \frac{1}{k}\left(1 + \frac{1}{r}\right)\left(1 + \frac{\ln k}{r\ln k + 1}\right) < \frac{1}{k}\left(1 + \frac{1}{r}\right)^2 \leq \frac{8}{9} < 1$$

if $r \geq 3$, $k \geq 2$. Thus $\phi(r,k)$ is decreasing with $r \geq 3$ for $k \geq 2$. Direct calculation shows that $\phi(9,2)$, $\phi(6,3)$, $\phi(5,4)$, $\phi(4,5)$, $\phi(3,15)$ are all less than 1. Thus $r(r\ln k + 1)(k+2)/k^r < 1$ for all regular pairs. $\square$

Lemma 4.18. *For all regular pairs, $\ln k - 2(k+2)/k^r > \ln(k-1)$.*

Proof. Using Lemma 4.5, $\ln k - \ln(k-1) = -\ln(1-1/k) > 1/k > 2(k+2)/k^r$, provided $2+4/k < k^{r-2}$. The left hand side of $2+4/k < k^{r-2}$ is decreasing, and the right hand side increasing, for all r, k. Thus we need only determine the smallest pairs $r \geq 3$, $k \geq 2$ which satisfy it. These are $k = 2$, $r = 5$, $k = 3$, $r = 4$ and $k = 4$, $r = 3$, which are not regular pairs. $\square$

Lemma 4.19. *For all $k \geq 1$, $4(k - 1) \geq \sqrt{k} \ln k$.*

Proof. Using Lemma 4.5, $\ln k = 2 \ln \sqrt{k} \leq 2(\sqrt{k} - 1)$. So the conclusion is implied by $2(k - 1) \geq k - \sqrt{k}$, which follows from $2(k - 1) \geq (k - 1)$ for all $k \geq 1$. $\square$

Lemma 4.20. $r(r - 1) \ln k / (k^{r-1} - 1) < 1$ *for all $k \geq 3$, $r \geq 2$, or $k = 2$, $r \geq 5$.*

Proof. Let $\phi(r, k) = r(r - 1) \ln k / (k^{r-1} - 1)$. Then, for $k \geq 3$, $r \geq 2$, or $k = 2$, $r \geq 5$ we have

$$\frac{\phi(r + 1, k)}{\phi(r, k)} = \frac{(r + 1)(k^{r-1} - 1)}{(r - 1)(k^r - 1)} < \frac{r + 1}{(r - 1)k} \leq 1.$$

Furthermore

$$\frac{\phi(r, k + 1)}{\phi(r, k)} = \frac{(k^{r-1} - 1) \ln(k + 1)}{((k + 1)^{r-1} - 1) \ln k} < \frac{\ln(k + 1)}{\ln k} \left(\frac{k}{k + 1} \right)^{r-1} \leq \frac{k / \ln k}{(k + 1) / \ln(k + 1)} < 1$$

for $k \geq 3$, from the proof of Lemma 4.16. If $k = 2$, $r \geq 5$ then

$$\frac{\phi(r, 3)}{\phi(r, 2)} = \frac{(2^{r-1} - 1) \ln 3}{(3^{r-1} - 1) \ln 2} < \frac{16 \ln 3}{81 \ln 2} < 1.$$

So ϕ is decreasing with both r and k. Now we may calculate that

$$\phi(3, 3) = 3 \ln 3 / 4 < 1, \qquad \phi(5, 2) = 4 \ln 2 / 3 < 1. \qquad \square$$