Combinatorics, Probability and Computing: page 1 of 9. © Cambridge University Press 2011
doi:10.1017/S096354831100006X

Covering Complete r-Graphs with Spanning Complete r-Partite r-Graphs

SEBASTIAN M. CIOABĂ[1,†] ANDRÉ KÜNDGEN[2],
CRAIG M. TIMMONS[3] and VLADISLAV V. VYSOTSKY[4]

[1]Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
(e-mail: cioaba@math.udel.edu)
[2]Department of Mathematics, California State University San Marcos, San Marcos, CA 92096, USA
(e-mail: akundgen@csusm.edu)
[3]Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA
(e-mail: ctimmons@ucsd.edu)
[4]Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
(e-mail: vysotsky@math.udel.edu)

Received 27 June 2010; revised 26 December 2010

An r-cut of the complete r-uniform hypergraph K_n^r is obtained by partitioning its vertex set into r parts and taking all edges that meet every part in exactly one vertex. In other words it is the edge set of a spanning complete r-partite subhypergraph of K_n^r. An r-cut cover is a collection of r-cuts such that each edge of K_n^r is in at least one of the cuts. While in the graph case $r = 2$ any 2-cut cover on average covers each edge at least $2 - o(1)$ times, when r is odd we exhibit an r-cut cover in which each edge is covered exactly once. When r is even no such decomposition can exist, but we can bound the average number of times an edge is cut in an r-cut cover between $1 + \frac{1}{r+1}$ and $1 + \frac{1+o(1)}{\log r}$. The upper bound construction can be reformulated in terms of a natural polyhedral problem or as a probability problem, and we solve the latter asymptotically.

1. Introduction

A cut in a graph G is the set of all edges that meet both sets in a partition of its vertex set $V(G)$ into two parts X_1, X_2. A cut cover of G is a collection of cuts whose union is its edge set $E(G)$. Cut covers in graphs have been studied at least since the 1970s, when several authors [7, 10, 17] showed that the minimum number of cuts in a cut cover of G is $\lceil \log_2 \chi(G) \rceil$, where $\chi(G)$ is the chromatic number of G. A more challenging problem is determining how often an edge must be cut on average in a cut cover. In a cut cover with

† Research supported by a start-up grant from the Department of Mathematical Sciences of the University of Delaware.

star cuts ($|X_1| = 1$) each edge is cut at most twice, so that for non-bipartite graphs this number is strictly between 1 and 2. (For more detailed results see [6], [16].) The complete graph K_n is the n-vertex graph that is the hardest to cut, with each edge cut at least $2 - \frac{1}{n}$ times on average (when $n \geqslant 8$) and this value can only be achieved by using $n-1$ star cuts, a result that was discovered by several authors in different contexts [11, 12, 14, 15].

The goal of this paper is to extend the results for K_n to the complete r-uniform hypergraph on n vertices, K_n^r. One way to generalize the notion of a cut to a hypergraph $\mathcal{H}$ is to partition $V(\mathcal{H})$ into two parts X_1, X_2 and take every edge that meets both parts. In this way the proof for graphs can be easily adapted to yield that the minimum number of cuts in a cover of $E(\mathcal{H})$ is $\lceil \log_2 \chi(\mathcal{H}) \rceil$. Again, on average K_n^r is hardest to cover, and in [3] it is shown that for n sufficiently large the most efficient way is to take $\lfloor \frac{n-1}{r-1} \rfloor$ such cuts with $|X_1| \leqslant r-1$, covering each edge $r - O\left(\frac{1}{n}\right)$ times on average.

In this paper we focus on another way to generalize the notion of a cut to r-uniform hypergraphs: partition the vertex set into r parts $X_1, X_2, \ldots, X_r$ and let the *r-cut* $[X_1, X_2, \ldots, X_r]$ of $\mathcal{H}$ consist of all edges $\{x_1, \ldots, x_r\}$ such that $x_i \in X_i$ for all i with $1 \leqslant i \leqslant r$. So an r-cut in K_n^r is a spanning complete r-partite subhypergraph of K_n^r. In Section 2 we will see that every r-uniform hypergraph on n vertices can be covered with $O(\log n)$ r-cuts (when r is fixed), and K_n^r shows that this is best possible. Turning our focus to the average number of times an edge must be cut in an r-cut cover, we show in Section 3 that the answer is always 1 (that is, we can decompose $\mathcal{H}$ into r-cuts) when r is odd. We also show that this is not possible when r is even, and in fact each edge of K_n^r is cut on average at least $1 + \frac{1}{r+1}$ times. In Section 4 we propose an efficient r-cut cover of K_n^r for even r, and we show in Section 5 that in this cover on average each edge is cut $1 + \frac{1+o(1)}{\log r}$ times.

2. Minimizing the number of r-cuts in a cover

In this section we will see that the minimum number of r-cuts needed to cover K_n^r is of size $c \log n$, where c depends only on r. To our knowledge, the best estimates for the constant c are $\Omega\left(\frac{e^r}{r\sqrt{r}}\right) = c = O(re^r)$ (see [21]). This problem has been well studied and is related to interesting questions in data storage, computer science and information theory regarding families of perfect hash functions, separating systems and graph entropy (see [5], [18], [19], [21]). The results contained in this section are due to Mehlhorn [18] and Graham (*cf.* Fredman and Komlós [5]; see also [19], [21]), and we include some discussion here for the sake of completeness.

The following simple lower bound is a straightforward adaptation of the graph case $r = 2$.

Theorem 2.1. *Every r-cut cover of K_n^r has size at least $\frac{1}{\log r} \log n$.*

Proof. Suppose we have a cover of K_n^r with m r-cuts. With each vertex v we associate an m-tuple $(v_1, v_2, \ldots, v_m)$ where $v_i = j$ precisely when $v \in X_j$ in the ith r-cut. Different vertices must receive different labels, since otherwise none of the edges containing both are covered in an r-cut. Thus $n \leqslant r^m$. □

The observation that vertices of the same label must form an independent set easily yields that an r-cut cover of the edge set of any r-uniform hypergraph $\mathcal{H}$ of order n must contain at least $\Omega(\log \chi(\mathcal{H}))$ cuts, where $\chi(\mathcal{H})$ is the chromatic number of $\mathcal{H}$ (see also [21]). A more careful argument can be used to improve the $1/\log r$ factor in Theorem 2.1 to $\Omega\left(\frac{e^r}{r\sqrt{r}}\right)$; see, for example, [19] with $b = k = r = j + 2$ for a simple probabilistic proof.

The upper bound complementing Theorem 2.1 can also be easily proved probabilistically.

Theorem 2.2. *For every r there is a c such that K_n^r can be covered with $c \log n$ r-cuts.*

Proof. Consider a random cut in which each vertex is put into the parts with equal probability. The probability that a given edge is not cut by such an r-cut is $P = \frac{r^r - r!}{r^r} < 1$. Thus in a cover with $c \log n$ random cuts the expected number of uncut edges is $\binom{n}{r} P^{c \log n} < n^{r + c \log P}$. This number is less than 1 for $c > -r/\log P$, so that some such cover will have no uncut edges. The value $c = O(re^r)$ can be obtained by instead using the alteration method in the same set-up (see [2], [19] for more details). □

A very important problem is to explicitly construct covers of K_n^r with $O(\log n)$ r-cuts. Some explicit constructions are known; see [5], [21] for more details on explicit constructions and related questions.

3. Decomposing K_n^r into r-cuts for odd r

In this section we give a simple construction to show that when r is odd we can actually partition the edges of any r-uniform hypergraph into r-cuts, and we also show that this is not possible when r is even. It suffices to prove both statements for complete hypergraphs, since the r-cover we propose in the next proposition in fact yields a decomposition for all r-uniform hypergraphs.

For a given sequence $1 \leqslant a_1 < a_2 < \cdots < a_k \leqslant n$ we let $A_i = \{a_{i-1} + 1, a_{i-1} + 2, \ldots, a_i - 1\}$ for $1 \leqslant i \leqslant k + 1$, where we set $a_0 = 0$ and $a_{k+1} = n + 1$. The *sequential cut* of K_n^{2k+1} induced by $a_1, \ldots, a_k$ is $[A_1, \{a_1\}, A_2, \{a_2\}, \ldots, A_k, \{a_k\}, A_{k+1}]$.

Proposition 3.1. *The collection of all sequential cuts of K_n^{2k+1} forms a partition of $E(K_n^{2k+1})$.*

Proof. An edge $\{x_1, \ldots, x_{2k+1}\}$ is covered by the sequential cut induced by $a_1, a_2, \ldots, a_k$ if and only if $a_1 = x_2, a_2 = x_4, \ldots, a_k = x_{2k}$. □

Observe that a sequential cut is empty if $a_i = a_{i-1} + 1$ for some $1 \leqslant i \leqslant k + 1$. Thus the number of non-empty sequential cuts in K_n^{2k+1} is $\binom{n-k-1}{k}$, and this construction was used in [4] to give an upper bound on the number of complete r-partite r-graphs needed to partition the edges of K_n^r. This question, studied first by Alon [1], is a hypergraph version of a well-known problem of partitioning the edge set of a graph into the minimum number of complete bipartite subgraphs that was investigated by Graham and Pollak, and Tverberg, among others (see [8], [9], [20], [22], [23] for more details).

We now turn our attention to the case when r is even. We start by observing that the minimum number of times an edge in K_n^r is cut on average, $c_r(n)$, is monotone in n. To that end we define $t(n,r)$ as the minimum total size (sum of the sizes of the r-cuts) of an r-cut cover of K_n^r, *i.e.*, $t(n,r) = c_r(n)\binom{n}{r}$.

Lemma 3.2. *For $m \leqslant n$ we have $c_r(m) \leqslant c_r(n)$.*

Proof. Consider an optimal r-cut cover $\mathcal{C}$ of K_n^r. For each m-set M of vertices of K_n^r, $\mathcal{C}$ induces an r-cut cover $\mathcal{C}|_M$ of K_m^r by restricting each X_i to $X_i \cap M$. Since each edge is covered in exactly $\binom{n-r}{m-r}$ covers $\mathcal{C}|_M$, and the total size of each $\mathcal{C}|_M$ is at least $t(m,r)$, the result follows from $\binom{n}{m}t(m,r) \leqslant \sum_{|M|=m}|\mathcal{C}|_M| \leqslant \binom{n-r}{m-r}t(n,r)$. □

When r is even, K_{r+1}^r has an odd number of edges, but every r-cut has size 2, so it is easy to see that $t(r+1,r) = r+2$. Thus it follows from Lemma 3.2 that K_n^r has no partition into r-cuts when $n \geqslant r+1$, since

$$c_r(n) \geqslant c_r(r+1) = \frac{r+2}{r+1} = 1 + \frac{1}{r+1}. \tag{3.1}$$

Determining $c_r(n)$ is a non-trivial problem, but Lemma 3.2 implies that $c_r = \lim_{n\to\infty} c_r(n)$ exists, with $c_r \geqslant 1 + \frac{1}{r+1}$ for r even, since we will show at the start of the next section that $c_r(n) \leqslant 2$. In fact, the rest of this paper is devoted to proving the upper bound of our main result.

Theorem 3.3. $1 + \frac{1}{r+1} \leqslant c_r \leqslant 1 + \frac{1+o(1)}{\log r}$ *when r is even.*

4. Constructions for even r

In this section we give an explicit cover of K_n^r for r even with very small covering multiplicity, and we start analysing the efficiency of the cover by reformulating it as a probability problem and a polyhedral problem that could be of independent interest.

Since sequential covers are so efficient in the odd case it seems sensible to obtain a modified cover for K_n^{2k} from the sequential cover for K_n^{2k+1}. The simplest way to do this is to always merge A_1 and A_{k+1} into one part, since then each edge $\{x_1, x_2, \ldots, x_{2k}\}$ of K_n^{2k} is covered exactly twice: once in the $2k$-cut derived from the sequential $(2k+1)$-cut induced by $x_2, x_4, \ldots, x_{2k}$ and once in the $2k$-cut derived from $x_1, x_3, \ldots, x_{2k-1}$. Thus $c_r(n) \leqslant 2$ and $c_r \leqslant 2$.

However, we can do better. In fact we will give a construction that shows that $c_r \to 1$ as $r \to \infty$. Let $[a_1, \ldots, a_k]$ denote the $2k$-cut obtained by merging A_{k+1} with the *largest* of the remaining parts (call it A_m) in the sequential $(2k+1)$-cut derived from $a_1, \ldots, a_k$. (If there are several largest parts, then let A_m be any one of these.) More formally, we let $B_m = A_m \cup A_{k+1}$ and $B_i = A_i$ for $i \neq m$, and let $[a_1, \ldots, a_k] = [B_1, \{a_1\}, B_2, \{a_2\}, \ldots, B_k, \{a_k\}]$ be the *modified* cut of K_n^{2k} derived from the sequential cut $[A_1, \{a_1\}, A_2, \{a_2\}, \ldots, A_k, \{a_k\}, A_{k+1}]$ of K_n^{2k+1}.

Definition. An edge $E = \{x_1, x_2, \dots, x_{2k}\}$ of K_n^{2k} with $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} \leqslant n$ is called *exceptional* if there is an even number q with $0 \leqslant q \leqslant 2k-2$ and

$$x_{q+1} - x_q \geqslant x_2 - x_0, x_4 - x_2, \dots, x_q - x_{q-2}, x_{q+3} - x_{q+1}, \dots, x_{2k-1} - x_{2k-3}. \tag{4.1}$$

Let $M_k(n)$ denote the number of exceptional edges.

Observe that for any exceptional edge q is unique, since $x_{q+1} - x_q$ must be the unique maximum among all values $x_i - x_{i-1}$ for $1 \leqslant i \leqslant 2k-1$.

Proposition 4.1. *The collection of modified cuts of K_n^{2k} is a cover of $E(K_n^{2k})$ of total size $\binom{n}{2k} + M_k(n)$.*

Proof. Consider an edge $E = \{x_1, \dots, x_{2k}\}$ of K_n^{2k}, where $x_1 < x_2 < \cdots < x_{2k}$. A modified cut $[a_1, \dots, a_k]$ covering E must have that for all i, $a_i = x_{f(i)}$, where f is an increasing function with $1 \leqslant f(1) \leqslant 2$, $2k-1 \leqslant f(k) \leqslant 2k$, and $f(i+1) \leqslant f(i) + 2$. Thus it follows that there is a j such that $[a_1, \dots, a_k] = [x_2, x_4, \dots, x_{2j-2}, x_{2j-1}, x_{2j+1}, \dots, x_{2k-1}] = C_j$, where $1 \leqslant j \leqslant k+1$. So $C_1 = [x_1, x_3, \dots, x_{2k-1}]$, $C_{k+1} = [x_2, x_4, \dots, x_{2k}]$ and all other C_j contain exactly two consecutive x_i, namely x_{2j-2}, x_{2j-1}.

The cut C_{k+1} clearly contains E, so the modified cuts form a cover of K_n^{2k}. For any other C_j to cover E we need $x_{2k} \in A_{k+1} \subseteq B_{m(C_j)}$, and in fact $m(C_j) = j$. The definition of B_m implies that $|A_m| \geqslant |A_1|, |A_2|, \dots, |A_k|$, and thus E satisfies (4.1) with $q = 2j-2$.

So for E to be covered at least twice it must be exceptional. In that case, since q is unique for any edge satisfying (4.1), at most one other cut covers E, namely C_j. In fact, when E is exceptional it is easy to see that C_j indeed covers E. Thus every exceptional edge is covered twice and every other edge is covered once, and the result follows. □

To see how good the modified cover is, it remains to count exceptional edges.

Definition. A sequence $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} < x_{2k+1}$ is called *special* if

$$x_{2k+1} - x_{2k} \geqslant x_2 - x_0, x_4 - x_2, \dots, x_{2k} - x_{2k-2}. \tag{4.2}$$

Let $Q_k(N)$ be the number of special sequences of $2k+1$ integers with $x_{2k+1} = N$.

Lemma 4.2.

$$M_k(n) = k \sum_{N=1}^{n} Q_{k-1}(N)(n-N).$$

Proof. If we fix $x_{2k-1} = N$, then $Q_{k-1}(N)(n-N)$ counts exactly the number of those sequences of length $2k$ satisfying (4.1) for $q = 2k-2$. So to prove the equation it suffices to see that the number of such sequences is the same for any other choice of an even number q with $0 \leqslant q \leqslant 2k-4$. We do so by giving a bijection from the sequences $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} \leqslant n$ satisfying (4.1) for fixed $q \neq 2k-2$ to those for $q = 2k-2$.

From a given $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} \leqslant n$, we obtain $0 = y_0 < y_1 < y_2 < \cdots < y_{2k} \leqslant n$, by letting

$$y_i = \begin{cases} x_i & \text{for } 0 \leqslant i \leqslant q, \\ x_{i+1} - x_{q+1} + x_q & \text{for } q \leqslant i \leqslant 2k-2, \\ x_i & \text{for } i = 2k-1, 2k. \end{cases}$$

This shifted sequence is increasing and satisfies (4.1) with $q = 2k - 2$, since

$$y_{2k-1} - y_{2k-2} = x_{2k-1} - (x_{2k-1} - x_{q+1} + x_q) = x_{q+1} - x_q. \qquad \square$$

Observe that

$$\sum_{N=1}^{n} \binom{N-1}{2k-2}(n-N) = \sum_{N=1}^{n} \left[n\binom{N-1}{2k-2} - (2k-1)\binom{N}{2k-1} \right]$$
$$= n\binom{n}{2k-1} - (2k-1)\binom{n+1}{2k} = \binom{n}{2k}.$$

Thus, if $Q_{k-1}(N)/\binom{N-1}{2k-2} \to q_{k-1}$ as $N \to \infty$, then Lemma 4.2 implies that for $n \to \infty$ we have $M_k(n)/\binom{n}{2k} \to kq_{k-1}$, and it follows from Proposition 4.1 that

$$c_{2k} \leqslant 1 + kq_{k-1}. \tag{4.3}$$

Obviously $Q_k(N)/\binom{N-1}{2k}$ is the probability that a randomly chosen sequence of distinct integers $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} < x_{2k+1} = N$ is special. By standard probability theory arguments, as $N \to \infty$ this sequence of probabilities converges to the probability that a randomly chosen sequence of distinct numbers in the unit interval $0 = x_0 < x_1 < x_2 < \cdots < x_{2k} < x_{2k+1} = 1$ is special. We devote the next section to the computation of this limit probability q_k.

We want to briefly mention a natural *polyhedral reformulation* suggested to us by Zoltán Füredi. With the change of variables $y_i = x_i - x_{i-1}$, q_k is the fraction of the simplex bounded by $y_1 + \cdots + y_{2k} \leqslant 1$ and $y_i \geqslant 0$ that remains when we cut off the vertices in pairs (y_{2i-1}, y_{2i}) by including the stronger constraints $y_1 + y_2 + \cdots + y_{2i-2} + 2y_{2i-1} + 2y_{2i} + y_{2i+1} + \cdots + y_{2k} \leqslant 1$ for each i with $1 \leqslant i \leqslant k$.

5. Probability that a random sequence in [0, 1] is special

In this section we determine the exact value of q_k as a simple integral, and then we solve this integral asymptotically.

Let $X_1, X_2, \ldots, X_n$ be independent random variables uniformly distributed over $[0, 1]$. We arrange this sequence in ascending order and get the sequence $X_{1,n} \leqslant X_{2,n} \leqslant \ldots \leqslant X_{n,n}$ of the so-called order statistics. Is is known (see [13, Section 13.1]) that the random vector $(X_{1,n}, \ldots, X_{n,n})$ admits the following representation. Let $Y_1, Y_2, \ldots, Y_{n+1}$ be independent identically distributed random variables with a standard exponential distribution, that is, $\mathbb{P}(Y_i \in dy) = e^{-y}dy$ for any $y \geqslant 0$, and let $S_n := Y_1 + \cdots + Y_n$ be the corresponding random walk. Then the random vector $(X_{1,n}, \ldots, X_{n,n})$ has the same distribution as $\left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}\right)$.

This representation provides a simple way of computing the probability that the sequence $X_{1,2k}, \ldots, X_{2k,2k}$ is special. Indeed, defining $X_{0,n} := 0$, we have

$$\begin{aligned} q_k &= \mathbb{P}\Big(\max_{1\leqslant i\leqslant k}(X_{2i} - X_{2i-2}) \leqslant 1 - X_{2k,2k}\Big) \\ &= \mathbb{P}\left(\max_{1\leqslant i\leqslant k}\left(\frac{S_{2i}}{S_{2k+1}} - \frac{S_{2i-2}}{S_{2k+1}}\right) \leqslant 1 - \frac{S_{2k}}{S_{2k+1}}\right) \\ &= \mathbb{P}\Big(\max_{1\leqslant i\leqslant k}(Y_{2i-1} + Y_{2i}) \leqslant Y_{2k+1}\Big) \\ &= \int_0^\infty \mathbb{P}^k\big(Y_1 + Y_2 \leqslant y\big)e^{-y}dy, \end{aligned}$$

where we used independence of Y_n to get the last line. It is well known that $Y_1 + Y_2$ has a gamma(2) distribution whose density is given by $\mathbb{P}(Y_1 + Y_2 \in dy) = ye^{-y}dy$ for $y \geqslant 0$. Then $\mathbb{P}(Y_1 + Y_2 \leqslant y) = 1 - (1 + y)e^{-y}$ for $y \geqslant 0$, and we get

$$q_k = \int_0^\infty \big(1 - (1 + y)e^{-y}\big)^k e^{-y}dy.$$

The change of variables $x = e^{-y}$ now yields the following result.

Theorem 5.1.

$$q_k = \int_0^1 \big(1 - (1 - \log x)x\big)^k dx.$$

Theorem 5.1 implies immediately that q_k is decreasing in k, whereas we still have no proof that c_{2k} is decreasing. While it seems difficult to solve this integral explicitly, it is now easy to compute small values using a computer algebra system: $q_1 = \frac{1}{4}$, $q_2 = 7/54$, $q_3 = \frac{97}{1152}$. Thus, combining (3.1) with (4.3) we obtain the following estimates for c_{2r}: $1.2 \leqslant c_4 \leqslant 1.5$, $1.14 \leqslant c_6 \leqslant 1.39$, $1.11 \leqslant c_8 \leqslant 1.34$. While we have little reason to believe that the upper bounds obtained this way are optimal, the lower bounds are certainly weak. In fact for the graph case we know that $c_2 = 2$, which is exactly what the construction in Proposition 4.1 yields, whereas the lower bound from (3.1) is only 4/3.

The upper bound for our main result, Theorem 3.3, immediately follows from the asymptotics of q_k given below.

Theorem 5.2. $q_k \sim \frac{1}{k \log k}$.

Proof. With the definition $f(0) = 0$, the function $f(x) := (1 - \log x)x$ is continuous and increasing on $[0, 1]$, with $f'(x) = -\log x$ and $f(1) = 1$. Thus

$$q_k = \int_0^1 (1 - f(x))^k dx = \int_0^{f^{-1}(\frac{2\log k}{k})} (1 - f(x))^k\, dx + \int_{f^{-1}(\frac{2\log k}{k})}^1 (1 - f(x))^k\, dx.$$

Observe that using $y = -\frac{u}{k}$ in the inequality $e^y \geqslant 1 + y$ we obtain $(1 - \frac{u}{k})^k \leqslant e^{-u}$ when $u \leqslant k$, and thus the second integral is bounded from above by $e^{-2\log k} = 1/k^2$. The change of variables $u = kf(x)$ now yields

$$q_k = \frac{1}{k}\int_0^{2\log k} \frac{(1-\frac{u}{k})^k}{-\log f^{-1}(\frac{u}{k})}\,du + O\left(\frac{1}{k^2}\right).$$

Note that $\log f^{-1}(z) \sim \log z$, as $z \to 0$ now follows by L'Hôpital's rule:

$$\lim_{z\to 0}\frac{\log f^{-1}(z)}{\log z} = \lim_{z\to 0}\frac{zf^{-1}(z)'}{f^{-1}(z)} = \lim_{z\to 0}\frac{z}{f^{-1}(z)f'(f^{-1}(z))} = \lim_{y\to 0}\frac{f(y)}{yf'(y)} = 1.$$

Hence

$$\begin{aligned} q_k &= \frac{1}{k}\int_0^{2\log k} \frac{(1-\frac{u}{k})^k(1+o(\frac{u}{k}))}{\log k - \log u}\,du + O\left(\frac{1}{k^2}\right) \\ &= \frac{1}{k\log k}\int_0^{\infty} \frac{(1-\frac{u}{k})^k(1+o(\frac{u}{k}))}{1-\frac{\log u}{\log k}}\mathbb{1}_{[0,2\log k]}(u)\,du + O\left(\frac{1}{k^2}\right). \end{aligned}$$

For k large enough, the integrand does not exceed $2e^{-u}$, which is integrable on $[0,\infty)$. Since for fixed u the integrand converges to e^{-u} as $k \to \infty$, it now follows from the dominated convergence theorem that the integral converges to $\int_0^\infty e^{-u}du = 1$. □

Acknowledgements

We thank the referees for their useful suggestions.

References

[1] Alon, N. (1986) Decomposition of the complete r-graph into complete r-partite r-graphs. *Graphs Combin.* **2** 95–100.
[2] Alon, N. and Spencer, J. (2008) *The Probabilistic Method*, third edition, Wiley Interscience Series in Discrete Mathematics and Optimization.
[3] Cioabă, S. M. and Kündgen, A. Covering hypergraphs with cuts of minimum total size. *Graphs Combin.* (to appear).
[4] Cioabă, S. M., Kündgen, A. and Verstraëte, J. (2009) On decompositions of complete hypergraphs. *J. Combin. Theory Ser. A* **116** 1232–1234.
[5] Fredman, M. K. and Komlós, J. (1984) On the size of separating systems and families of perfect hash functions. *SIAM J. Alg. Discrete Methods* **5** 61–68.
[6] Füredi, Z. and Kündgen, A. (2001) Covering a graph with cuts of minimum total size. *Discrete Math.* **237** 129–148.
[7] Garey, M., Johnson, D. and So, H. (1976) An application of graph coloring to printed circuit testing. *IEEE Trans. Circuits and Systems* **CAS-23** 591–599.
[8] Graham, R. L. and Pollak, H. O. (1971) On the addressing problem for loop switching. *Bell System Tech. J.* **50** 2495–2519.
[9] Graham, R. L. and Pollak, H. O. (1972) On embedding graphs in squashed cubes. In *Graph Theory and Applications*, Vol. 303 of *Lecture Notes in Mathematics*, Springer, pp. 99–110.
[10] Harary, F., Hsu, D. and Miller, Z. (1977) The biparticity of a graph. *J. Graph Theory* **1** 131–133.
[11] Jaeger, F., Khelladi, A. and Mollard, M. (1985) On shortest cocycle covers of graphs. *J. Combin. Theory Ser. B* **39** 153–163.
[12] Jamshy, U. and Tarsi, M. (1989) Cycle covering of binary matroids. *J. Combin. Theory Ser. B* **46** 154–161.
[13] Karlin, S. and Taylor, H. M. (1981) *A Second Course in Stochastic Processes*, Academic Press.

[14] Klugerman, M., Russell, A. and Sundaram, R. (1998) A note on embedding complete graphs into hypercubes. *Discrete Math.* **186** 289–293.
[15] Kündgen, A. (1998) Covering cliques with spanning bicliques. *J. Graph Theory* **27** 223–227.
[16] Kündgen, A. and Spangler, M. (2005) A bound on the total size of a cut cover. *Discrete Math.* **296** 121–128.
[17] Matula, D. W. (1972) k-components, clusters and slicings in graphs. *SIAM J. Appl. Math.* **22** 459–480.
[18] Mehlhorn, K. (1982) On the program size of perfect functions and universal hash functions. In *Proc. 23rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 170–175.
[19] Nilli, A. (1994) Perfect hashing and probability. *Combin. Probab. Comput.* **3** 407–409.
[20] Peck, G. W. (1984) A new proof of a theorem of Graham and Pollak. *Discrete Math.* **49** 327–328.
[21] Radhakrishnan, J. (2001) Entropy and counting. In IIT Kharagpur Golden Jubilee Volume on *Computational Mathematics, Modelling and Algorithms* (J. C. Mishra, ed.), Narosa, New Delhi.
[22] Tverberg, H. (1982) On the decomposition of K_n into complete bipartite graphs. *J. Graph Theory* **6** 493–494.
[23] Vishwanathan, S. (2008) A polynomial space proof of the Graham–Pollak theorem. *J. Combin. Theory Ser. A* **115** 674–676.