AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **39** (2007), Pages 141–154

A generalized switching method for combinatorial estimation

Veerle Fack

Department of Applied Mathematics and Computer Science
Ghent University
B-9000 Ghent
Belgium
`Veerle.Fack@UGent.be`

Brendan D. McKay*

Department of Computer Science
Australian National University
Canberra, ACT 0200
Australia
`bdm@cs.anu.edu.au`

Abstract

The method of switchings is a standard tool for enumerative and probabilistic applications in combinatorics. In its simplest form, it analyses a relation between two sets to estimate the ratio of their sizes. Via a sequence of such pairwise ratios, the relative sizes of a larger family of sets can be estimated. However, in some situations, the available relations might not form a simple chain in this manner. For example, a relation might be defined by an operation that takes an object in one set and converts it to an object that is only known to lie in some subfamily of other sets (rather than in a single known other set). In this article, we describe this situation in a general setting and present it as an optimisation problem. Then we prove that its optimal solution can be assumed to have a very simple form. To illustrate this result we give two examples. First we extend a special case of a lemma of Greenhill and McKay (2006) that bounds the probability of large entries in certain integer matrices. Then we strengthen a lemma of McKay, Wormald and Wysocka (2004) that bounds the number of edges that lie in short cycles in a random regular graph.

* Research supported by the Australian Research Council

1 Introduction

The simplest example of the method of switchings involves two finite sets A, B, and a relation $R \subseteq A \times B$. If d_A is the average number of elements of B that are related to a uniformly chosen element of A, and d_B is the average number of elements of A that are related to a uniformly chosen element of B, then $d_A|A| = |R| = d_B|B|$. Thus, estimates of the relative values of d_A and d_B provide estimates of the relative sizes of A and B.

This idea is easily extended to a sequence of finite sets $A_1, A_2, \ldots, A_n$ and relations $R_i \subseteq A_i \times A_{i+1}$ for $1 \le i \le n-1$. Let a_i be the average number of elements of A_{i+1} related by R_i to a uniformly chosen element of A_i $(1 \le i \le n-1)$ and let b_i be the average number of elements of A_{i-1} related by R_{i-1} to a uniformly chosen element of A_i $(2 \le i \le n)$. Then, provided the denominators are nonzero,

$$\frac{|A_n|}{|A_1|} = \frac{a_1 a_2 \cdots a_{n-1}}{b_2 b_3 \cdots b_n},$$

and, indeed, the relative sizes of any of $A_1, \ldots, A_n$ can be determined.

There are many examples in the literature where classes of combinatorial objects are approximately enumerated by this technique. A few examples are [2, 3, 5, 8, 6, 7, 9, 10].

In this paper we generalise the switching method to the case where the relations form a graph other than a simple chain. This is most easily explained in terms of the elementary operations (called "switchings") that are used to define the relations in most published examples.

Suppose we have a directed graph G where each vertex v defines some finite class $\mathcal{C}(v)$ of objects, such classes being disjoint. We also have a family of operations ("switchings") that each take an object and convert it to a possibly-distinct object. For example, if the objects are graphs, an switching might involve altering some small subgraph. Formally, we have that for each object x, there is a multiset $\mathcal{S}(x)$ of the objects that result from applying a switching to x. The inverse of $\mathcal{S}$ is defined in the obvious fashion by

$$\mathcal{S}^{-1}(y) = \bigl\{\, x \bigm| y \in \mathcal{S}(x) \bigr\},$$

where $\mathcal{S}^{-1}(y)$ is again a multiset: the multiplicity of x in $\mathcal{S}^{-1}(y)$ is the same as the multiplicity of y in $\mathcal{S}(x)$. The (directed) edges of G indicate the possible trajectories of a switching:

$$E(G) = \Bigl\{\, (v, w) \in V(G) \times V(G) \Bigm| \bigcup_{x \in \mathcal{C}(v)} \mathcal{S}(x) \cap \mathcal{C}(w) \ne \emptyset \Bigr\}.$$

As defined, G has no multiple edges; this allows us to write vw for the unique edge, if any, from v to w.

Now suppose we have bounds on $\sum_{x\in\mathcal{C}(v)}|\mathcal{S}(x)|/|\mathcal{C}(v)|$ and $\sum_{x\in\mathcal{C}(v)}|\mathcal{S}^{-1}(x)|/|\mathcal{C}(v)|$ for each $v \in V(G)$. The problem is to infer bounds on the relative sizes of the sets $\mathcal{C}(v)$.

The presence of directed cycles in G greatly complicates the analysis so we will assume there are none except that we allow loops.

2 The task as an optimisation problem

Let $G = (V, E)$ be a directed graph without multiple edges, acyclic except that loops are permitted. For a vertex $v \in V$, let $G^-(v)$ and $G^+(v)$ be the set of (directed) edges entering and leaving v, respectively.

The input to our problem consists of G together with a pair of functions $a : V \to \mathbb{R}_{>0}$ and $b : V \to \mathbb{R}_{>0}$, where $\mathbb{R}_{>0}$ is the set of positive real numbers, and two nonempty subsets $X, Y \subseteq V$. Note that X and Y may overlap.

A *feasible solution* is a pair of functions $S = (s, N)$, where $s : E \to \mathbb{R}_{\geq 0}$ and $N : V \to \mathbb{R}_{\geq 0}$, such that

$$\sum_{e\in G^-(v)} s(e) \leq b(v)N(v) \quad \text{for all } v \text{ such that } G^-(v) \neq \emptyset, \tag{1}$$

$$\sum_{e\in G^+(v)} s(e) \geq a(v)N(v) \quad \text{for all } v \text{ such that } G^+(v) \neq \emptyset. \tag{2}$$

A *proper solution* is a feasible solution $S = (s, N)$ for which $N(v) > 0$ for some $v \in X \cup Y$. In this case, the *weight* of the solution is $f(S) = \infty$ if $N(X) = 0$ and

$$f(S) = \frac{N(Y)}{N(X)}$$

otherwise, where $N(W) = \sum_{w\in W} N(w)$ for any $W \subseteq V$. Note that any positive scalar multiple of a proper solution is a proper solution with the same weight.

Problem (G, a, b, X, Y) is to determine the maximum of $f(S)$ over proper solutions S, if proper solutions exist. In that case, a proper solution achieving the maximum is called an *optimal solution*. In the example shown in Figure 1, $f(S) = \frac{1}{5}$.

To see how this problem fits into the description in the Introduction, interpret $N(v)$ as $|\mathcal{C}(v)|$ and $s(vw)$ as $\sum_{x\in\mathcal{C}(v)}|\mathcal{S}(X) \cap \mathcal{C}(w)|$, where the intersection inherits the multiplicities of the multiset $\mathcal{S}(X)$. In this interpretation, $\sum_{e\in G^+(v)} s(e) = \sum_{x\in\mathcal{C}(v)}|\mathcal{S}(x)|$ and $\sum_{e\in G^-(v)} s(e) = \sum_{x\in\mathcal{C}(v)}|\mathcal{S}^{-1}(x)|$, so $a(v)$ and $b(v)$ correspond to bounds on those quantities relative to $|\mathcal{C}(v)|$.

Our analysis allows s and N to take non-integral values, potentially leading to an upper bound higher than necessary, but it will turn out that the optimum occurs for some rational solution. Since a scalar multiple of a feasible solution is also a feasible solution, the optimum occurs for an integral solution.

If every feasible solution (s, N) has $N(Y) = 0$, then either there are no proper solutions or every proper solution achieves the maximum weight of 0. Otherwise, if

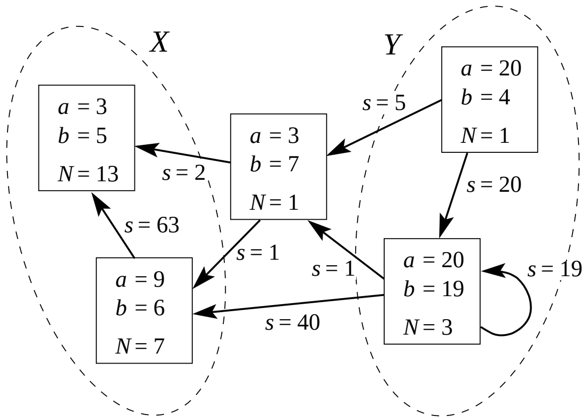

Figure 1: Example of a problem (G, a, b, X, Y) and a feasible solution (s, N)

we were to add the constraint $N(Y) = 1$, we would have a linear program of which the objective is to minimise $N(X)$. We know from the theory of linear programming that a finite optimum is achieved by some point on the boundary of the feasible region, so in this case we know that the minimum of $N(X)$ is achieved by some proper solution.

Thus, in all cases where there are proper solutions, the optimum is achieved. We could proceed with our investigation using the theory of linear programming, but our purpose is rather to show that the special structure of the problem implies a special structure for its solutions. In practical applications, this usually allows the solution to be obtained without recourse to the more general techniques of linear programming.

3 Path solutions

We next define five simple types of feasible solution that we call *path solutions*. Let $P = v_1, v_2, \ldots, v_k$ be a path in G from v_1 to v_k. Here, and throughout the paper, "path" means "simple directed path".

- A *type-0 path solution* has $k = 1$ and requires v_1 to have a loop with $a(v_1) \leq b(v_1)$. It has

$$N(v_1) = 1,$$
$$s(v_1 v_1) = a(v_1),$$

  and $s(e) = N(v) = 0$ otherwise. This solution satisfies (1)–(2) with equality except possibly for (1) at v_1.

- A *type-1 path solution* requires v_k to be a sink in G. Any $k \geq 1$ is acceptable.

It has

$$\begin{aligned} N(v_1) &= 1, \\ N(v_i) &= \frac{a(v_1)\cdots a(v_{i-1})}{b(v_2)\cdots b(v_i)} && (2 \le i \le k), \\ s(v_iv_{i+1}) &= a(v_i)N(v_i) && (1 \le i \le k-1), \end{aligned}$$

and $s(e) = N(v) = 0$ otherwise. This solution satisfies (1)–(2) with equality except possibly for (1) at v_1.

- A *type-2 path solution* requires v_1 to have a loop with $a(v_1) > b(v_1)$, and v_k to be a sink in G (so $k \ge 2$). It has

$$\begin{aligned} N(v_1) &= 1, \\ N(v_i) &= \bigl(a(v_1) - b(v_1)\bigr)\frac{a(v_2)\cdots a(v_{i-1})}{b(v_2)\cdots b(v_i)} && (2 \le i \le k), \\ s(v_1v_1) &= b(v_1), \\ s(v_1v_2) &= a(v_1) - b(v_1), \\ s(v_iv_{i+1}) &= a(v_i)N(v_i) && (2 \le i \le k-1), \end{aligned}$$

and $s(e) = N(v) = 0$ otherwise. This solution satisfies (1)–(2) with equality.

- A *type-3 path solution* requires $k \ge 2$ and v_k to have a loop with $a(v_k) < b(v_k)$. It has

$$\begin{aligned} N(v_1) &= 1, \\ N(v_i) &= \frac{a(v_1)\cdots a(v_{i-1})}{b(v_2)\cdots b(v_i)} && (2 \le i \le k-1), \\ N(v_k) &= \frac{a(v_{k-1})N(v_{k-1})}{b(v_k) - a(v_k)}, \\ s(v_iv_{i+1}) &= a(v_i)N(v_i) && (1 \le i \le k-1), \\ s(v_kv_k) &= a(v_k)N(v_k), \end{aligned}$$

and $s(e) = N(v) = 0$ otherwise. This solution satisfies (1)–(2) with equality except possibly for (1) at v_1.

- A *type-4 path solution* requires v_1 to have a loop with $a(v_1) > b(v_1)$, and v_k to have a loop with $a(v_k) < b(v_k)$ (so $k \ge 2$). It has

$$\begin{aligned} N(v_1) &= 1, \\ N(v_i) &= \bigl(a(v_1) - b(v_1)\bigr)\frac{a(v_2)\cdots a(v_{i-1})}{b(v_2)\cdots b(v_i)} && (2 \le i \le k-1), \\ N(v_k) &= \frac{a(v_{k-1})N(v_{k-1})}{b(v_k) - a(v_k)}, \end{aligned}$$

$$\begin{aligned} s(v_1v_1) &= b(v_1), \\ s(v_1v_2) &= a(v_1) - b(v_1), \\ s(v_iv_{i+1}) &= a(v_i)N(v_i) && (2 \le i \le k-1), \\ s(v_kv_k) &= a(v_k)N(v_k), \end{aligned}$$

and $s(e) = N(v) = 0$ otherwise. This solution satisfies (1)–(2) with equality.

In each case, we also regard nonzero scalar multiples of path solutions to be path solutions.

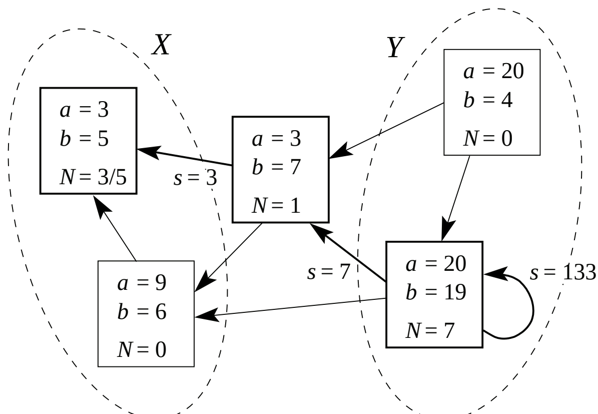

Figure 2: A type-2 path solution for the problem of Figure 1

4 Path solutions are enough

By a *nonzero vertex* we mean a vertex v with $N(v) > 0$, and by a *nonzero edge* we mean an edge e with $s(e) > 0$. A *nonzero path* means either a single nonzero vertex, or a path of one or more nonzero edges and their incident vertices (nonzero or not). In all cases we will write *S-nonzero* if it might not be clear that we are referring to solution S. By a *maximal* nonzero path we mean one that cannot be extended at the final vertex (extendibility backwards from the starting vertex is not important).

A path $P = v_1, \ldots, v_k$ will be called *well-terminated* (for G) if either v_k is a sink, or $k = 1$ and v_k has a loop with $a(v_k) \le b(v_k)$, or $k > 1$ and v_k has a loop with $a(v_k) < b(v_k)$.

Lemma 1. *Let $S = (s, N)$ be a feasible solution that is not everywhere zero. If P is a maximal nonzero path starting at a vertex v with $N(v) > 0$, then P is well-terminated. Conversely, if there is a well-terminated path starting at a vertex v, then there is a feasible solution (s, N) with $N(v) > 0$.*

Proof. If P consists only of v, we cannot satisfy (1)–(2) unless v is a sink or has a loop with $a(v) \leq b(v)$. Suppose instead that P ends at $w \neq v$, and that w is not a sink. Then we cannot satisfy (1)–(2) at w unless w has a loop. The fact that $\sum_{e \in G^-(w)} s(e) > s(ww) \geq 0$ implies by (1) that $N(w) > 0$, which implies by (2) that $s(ww) > 0$ or else P would not be maximal. Constraints (1)–(2) now imply that $a(w) < b(w)$.

For the converse, a path solution of type 0, 1 or 3 based on the well-terminated path provides the solution required. □

Lemma 1 provides the tools to classify problems according to whether their optima are zero, finite, or infinite.

Theorem 1. *Consider problem (G, a, b, X, Y).*

1. *If there is no well-terminated path starting in $X \cup Y$, then there is no proper solution.*
2. *If there is a well-terminated path starting in X, but not one starting in Y, then an optimal solution S^* has $f(S^*) = 0$.*
3. *If there is a well-terminated path starting in $Y \setminus X$ and avoiding X, then an optimal solution S^* has $f(S^*) = \infty$.*
4. *In all other cases, an optimal solution S^* has $0 < f(S^*) < \infty$.*

Proof. The first three of these claims follow immediately from Lemma 1. To see that the fourth also follows, suppose there are well-terminated paths starting in both X and Y, or a well-terminated path that intersects with both X and Y. The lemma tells us that there are proper solutions (s_1, N_1) and (s_2, N_2) with $N_1(X) > 0$ and $N_2(Y) > 0$. Then $S = (s_1 + s_2, N_1 + N_2)$ is a proper solution with $0 < f(S) < \infty$. This leaves the possibility that there is a sequence of proper solutions with arbitrarily large weights. This cannot happen unless the optimum is ∞, since, as we noted before, the optimum is always achieved. Therefore we have a feasible solution (s, N) with $N(X) = 0$ and $N(Y) > 0$. A path solution based on a nonzero maximal path beginning at some $v \in Y \setminus X$ with $N(v) > 0$ shows that this possibility belongs in case 3. □

Theorem 2. *Suppose problem (G, a, b, X, Y) has a finite nonzero optimum. Then there is an optimal solution of path type with the first vertex in Y.*

Proof. Let $S^* = (s^*, N^*)$ be an optimal solution with the least number of nonzero vertices and edges. If $N^*(x) = 0$ for any vertex x, then all edges e outgoing from x, including loops, have $s^*(e) = 0$, since they can otherwise be set to 0 without changing $f(S^*)$ or violating (1)–(2).

If there is only one S^*-nonzero vertex, it is easy to check that S^* is a scalar multiple of a path solution meeting the requirement of the theorem. So assume there are at least two S^*-nonzero vertices.

Since G is acyclic apart from loops, there is a vertex v such that $N^*(v) > 0$ and v has no incoming S^*-nonzero non-loop edges. We can assume that $v \in Y$, since otherwise we could set $N^*(v) = 0$ and $s^*(e) = 0$ for all edges e outgoing from v to obtain a feasible solution which violates our choice of S^*.

Let $P = (v, \ldots, w)$ be a maximal S^*-nonzero path starting at v.

(i) By Lemma 1, P is well-terminated.

(ii) Suppose $v \neq w$ and there is a loop at v with $a(v) \leq b(v)$. For any $q \geq 0$, define $S_q = (s_q, N_q)$ which is the same as S^* except that $N_q(v) = q$ and $s_q(vv) = a(v)q$. Then S_q is feasible. Recall that $v \in Y$. If $v \notin X$ then $f(S_q)$ can be made larger than $f(S^*)$ by choosing large q, contrary to our assumption that S^* is optimal, so $v \in X$. Therefore, $v \in X \cap Y$ and

$$f(S_q) = \frac{N^*(Y \setminus \{v\}) + q}{N^*(X \setminus \{v\}) + q},$$

which is either strictly monotonic in q, or independent of q. In the first case, S^* is not optimal. In the second case S_0 is optimal but has fewer nonzero vertices and edges than S^*. Both possibilities violate our conditions on S^*, so we must have $a(v) > b(v)$.

In all cases, we see that P satisfies enough requirements that there is a path solution $S_P = (s_P, N_P)$ whose nonzero edges are those of P together with any S^*-nonzero loops at v and w. If S_P is a scalar multiple of S^* then S_P satisfies the requirements of the theorem, so suppose this is not the case.

For any ε (positive or negative), let $S_\varepsilon = S^* + \varepsilon S_P$. If $\varepsilon \geq 0$, it is immediate that S_ε is feasible. If $\varepsilon < 0$ but ε is small enough to ensure that S_ε has no negative values, we claim that S_ε is still feasible. S_ε satisfies (2) for all v because S_P satisfies (2) with equality (by construction). For the same reason, S_ε satisfies (1) except possibly at v_1 in the cases where S_P has no nonzero loop at v_1 (i.e., types 0, 1 and 3). However, in those cases, (1) is satisfied at v_1 because $\sum_{e \in G^-(v_1)} \bigl(s^*(e) + \varepsilon s_P(e)\bigr) = 0$.

Now consider

$$f(S_\varepsilon) = \frac{N^*(Y) + \varepsilon N_P(Y)}{N^*(X) + \varepsilon N_P(X)},$$

which implies that

$$\frac{\partial f(N_\varepsilon)}{\partial \varepsilon} = \frac{N^*(X)N_P(Y) - N^*(Y)N_P(X)}{\bigl(N^*(X) + \varepsilon N_P(X)\bigr)^2}. \tag{3}$$

If the numerator is not zero, then by taking tiny ϵ (either positive or negative) we find feasible S_ε with $f(S_\varepsilon) > f(S^*)$. If the numerator of (3) is zero, $f(S_\varepsilon)$ is independent of ε. In that case, choose ε to be the negative value closest to 0 such that S_ε has fewer nonzero vertices and edges than S^*. Both possibilities contradict our choice of S^*. □

In many cases, the same nontrivial path in G can be the basis of either a type-1 or a type-2 path solution. Up to scaling, these solutions are the same except that

$N(v_1)$ is larger for the type-2 solution. If $v_1 \in Y \setminus X$ or $Y \subseteq X$, a type-2 path solution is preferred in that case. Similarly, a type-4 path solution is preferred over a type-3 path solution if both are available for the same path and $v_1 \in Y \setminus X$ or $Y \subseteq X$.

5 Example: Large entries in integer matrices

The first use of Theorem 2 was in a recent study of Greenhill and McKay [4] that determined the asymptotic number of matrices of nonnegative integers that have a specified sequence of row sums and column sums. A key step in the proof was to show that entries greater than 3 were improbable under the conditions that were imposed. Here we give a simple special case to illustrate the application.

Let $k = k(n)$ be a non-negative integer function satisfying $1 \le k \le (n/6)^{1/2}$. Let $\mathcal{M} = \mathcal{M}(n, k)$ be the set of all $n \times n$ matrices of nonnegative integers with all row sums and column sums equal to k. Write $[k]_t$ for $k(k-1)\cdots(k-t+1)$.

Theorem 3. *Let $\Delta = \Delta(n)$ be an integer with $\Delta \ge 3$. Then at most a fraction $2n^{2-\Delta}\Delta^{\Delta}[k]_{\Delta}$ of matrices in $\mathcal{M}$ have any entries equal to or greater than Δ.*

Proof. For $d, m \ge 0$, let $\mathcal{M}_d(m)$ be the set of matrices in $\mathcal{M}$ which have exactly m entries equal to d and no entries greater than d. Note that $\bigcup_{m\ge 0} \mathcal{M}_d(m) = \mathcal{M}_{d+1}(0)$ and $\mathcal{M}_{k+1}(0) = \mathcal{M}$. The fraction that the theorem requires us to bound is

$$\sum_{d=\Delta}^{k} \frac{|\mathcal{M}_d(>0)|}{|\mathcal{M}|}, \tag{4}$$

where $\mathcal{M}_d(>0) = \bigcup_{m\ge 1} \mathcal{M}_d(m)$.

Choose d with $3 \le d \le k$. Define the graph G with vertices $V = \{\mathcal{M}_d(0), \mathcal{M}_d(1), \ldots\}$. Let $X = V$ and $Y = V \setminus \{\mathcal{M}_d(0)\}$. The edges will be defined by means of the following operation that preserves row and column sums.

$$\begin{pmatrix} d & 0 & 0 & \cdots & 0 \\ 0 & x_1 & & & \\ 0 & & x_2 & & \\ \vdots & & & \ddots & \\ 0 & & & & x_d \end{pmatrix} \longmapsto \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & x_1{-}1 & & & \\ 1 & & x_2{-}1 & & \\ \vdots & & & \ddots & \\ 1 & & & & x_d{-}1 \end{pmatrix}, \tag{5}$$

where $x_1, x_2, \ldots, x_d \ge 1$ and these submatrices may occur in any position, not necessarily contiguous or in the order shown. Matrix entries not shown can have any values and are unchanged by the operation. The edges of G are those $\mathcal{M}_d(j)\mathcal{M}_d(i)$ such that some $A_j \in \mathcal{M}_d(j)$ is taken by this operation into some $A_i \in \mathcal{M}_d(i)$. Clearly $j - d - 1 \le i \le j - 1$, so G is acyclic and has no loops. We can interpret $N(\mathcal{M}_d(j))$ as $|\mathcal{M}_d(j)|$, and $s\big(\mathcal{M}_d(j)\mathcal{M}_d(i)\big)$ as the number of pairs (A_j, A_i) for $A_j \in \mathcal{M}_d(j), A_i \in \mathcal{M}_d(i)$ related by operation (5).

Now consider $A_j \in \mathcal{M}_d(j)$ for $j \geq 1$. We bound the number of submatrices (taken as ordered $(d{+}1)$-tuples of rows and columns) with the form of the left side of (5). We can choose one entry equal to d in j ways (call this "the d"), then choose the entries $x_1, x_2, \ldots, x_d$ one at a time avoiding choices that violate the pattern. The choice of entry x_d is the most restricted, so we bound that. Since the maximum entry in A_j is d, the number of nonzero entries is at least nk/d. Of those we must exclude the d (1 case), entries in the same column as a nonzero entry in the same row as the d other than the column of the d (at most $(k-d)k$ cases), entries in the same row as a nonzero entry in the same column as the d other than the row of the d (at most $(k-d)k$ cases) and entries in the same row or column as one of the entries we chose for $x_1, x_2, \ldots, x_{d-1}$ (at most $(d-1)(2k-1)$ choices). Overall, therefore, we can choose the entry x_d (and thus each of the less-restricted entries $x_1, x_2, \ldots, x_{d-1}$) in at least $nk/d - 1 - 2(k-d)k - (d-1)(2k-1) > nk/d - 2k^2$ ways. Therefore, define $a(\mathcal{M}_d(j)) = j(nk/d - 2k^2)^d$ for $j \geq 1$. Arbitrarily define $a(\mathcal{M}_d(0)) = 1$.

Next consider $A_i \in \mathcal{M}_d(i)$ for $i \geq 0$. We want to bound the number of submatrices of A_i matching the right side of (5). We can choose an ordered sequence of d ones in a column in at most $n[k]_d$ ways and similarly for d ones in a row. Some of these choices are not allowed, but we just need an upper bound. Therefore, define $b(\mathcal{M}_d(i)) = n^2\bigl([k]_d\bigr)^2$ for all i.

Now we can apply Theorem 2. Since there are no loops, the only path solutions are of type 1. The general case is described by a path $(v_1, v_2, \ldots, v_q)$ where $v_i = \mathcal{M}_d(t_i)$ for all i and $t_1 > t_2 > \cdots > t_q = 0$ with $1 \leq t_i - t_{i+1} \leq d+1$ for all i. Let $S = (s, N)$ be the type 1 path solution for this path, scaled so that $N(v_q) = 1$. Then $N(v_{i-1}) = b(v_i)N(v_i)/a(v_{i-1})$ for $1 \leq i \leq q$. Therefore,

$$\begin{aligned}
\frac{N(Y)}{N(X)} &= \frac{N(v_{q-1}) + N(v_{q-2}) + \cdots + N(v_1)}{N(v_q) + N(v_{q-1}) + \cdots + N(v_1)} \\
&\leq \max\nolimits_{i=2}^{q} \frac{N(v_{i-1})}{N(v_i)} \\
&= \max\nolimits_{i=2}^{q} \frac{n^2\bigl([k]_d\bigr)^2}{t_{i-1}(nk/d - 2k^2)^d} \\
&\leq \frac{n^2\bigl([k]_d\bigr)^2}{(nk/d - 2k^2)^d} \\
&\leq n^{2-d} d^d [k]_d. \qquad (6)
\end{aligned}$$

The last step can be proved by replacing one of the $[k]_d$ factors by its AM/GM bound $(k - (d-1)/2)^d$ and applying our assumptions that $n \geq 6k^2$ and $d \geq 3$. Bound (6) is independent of the path and so, by Theorem 2, (6) is a bound on $|\mathcal{M}_d({>}0)|/|\mathcal{M}_{d+1}(0)|$ and hence on $|\mathcal{M}_d({>}0)|/|\mathcal{M}|$. Recalling (4), the theorem now follows if we note that

$$\frac{n^{2-d-1}(d+1)^{d+1}[k]_{d+1}}{n^{2-d}d^d[k]_d} \leq \frac{ek^2}{n} \leq \frac{e}{6} < \frac{1}{2}$$

for $3 \leq d \leq k-1$. □

6 Example: Many short cycles in regular graphs

In the paper [9], McKay, Wormald and Wysocka studied the distribution of the numbers of short cycles in random regular graphs. At one point they needed to show that graphs with a very large number of short cycles were very rare, and applied a switching argument to that purpose.

Theorem 4 ([9]). *Let $k = k(n) \geq 3$ and $d = d(n) \geq 3$ satisfy $k(d-1)^{k-1} = o(n)$. Let $M = M(n) = 20Ak(d-1)^k$ with $A = A(n) > c$ for some constant $c > 1$. Then the probability that a random d-regular graph of order n has exactly M edges which lie on cycles of length at most k is less than*

$$\left(e^{5(A-1)}A^{-5A}\right)^{(d-1)^k} = e^{-5(d-1)^k}(e/A)^{M/4k}$$

for sufficiently large n. □

Now we will show that a wider and stronger result can be obtained from the same switching operation if it is analysed using Theorem 2.

We will use the factorial notation for arbitrary real arguments, with its usual meaning $z! = \Gamma(z+1)$. The following can be proved by case analysis for tiny x and Stirling's approximation for larger x.

Lemma 2. *For any real $x \geq \frac{10}{3}$,*

$$\frac{\lceil x^{1/2}\rceil x^{x+\lceil x^{1/2}\rceil}}{\left(x+\lceil x^{1/2}\rceil\right)!} > \frac{5}{28}\, e^{x}. \qquad \square$$

Theorem 5. *Let $k = k(n) \geq 3$ and $d = d(n) \geq 3$ satisfy $(d-1)^{k-1} = o(n)$. Let $M = M(n) \geq 6D$, where $D = (d-1)^k$. Then the probability that a random d-regular graph of order n has exactly M edges which lie on cycles of length at most k is less than*

$$7\,\exp\left(\frac{M-5D}{4k}\right)\left(\frac{5D}{M}\right)^{M/4k}$$

for sufficiently large n.

Proof. Let $\mathcal{H}(m)$ be the set of d-regular graphs of order n such that exactly m edges are on cycles of length k or less. In [9], a particular switching operation on d-regular graphs is defined. Here we will use the notation of our Introduction: for graph H, $\mathcal{S}(H)$ is the multiset of graphs that result from applying a switching to H. A property of this switching is that $\mathcal{S}(H) \subseteq \mathcal{H}(m-4k) \cup \cdots \cup \mathcal{H}(m)$ when $H \in \mathcal{H}(m)$. The precise definition of the switching also implies that $\mathcal{S}(H) = \emptyset$ if $H \in \mathcal{H}(0)$. As shown in [9], $|\mathcal{S}(H)| \geq \frac{1}{5}mnd$ for $H \in \mathcal{H}(m)$, and $|\mathcal{S}^{-1}(H)| \leq ndD$ for any H.

We can present the problem in the form required by Theorem 2 using a graph with vertices $V = \{\mathcal{H}(0), \mathcal{H}(1), \ldots\}$ and edges $\mathcal{H}(j)\mathcal{H}(i)$ for $j - 4k \leq i \leq j$ with $i \geq 0, j > 0$. Thus, every vertex except $\mathcal{H}(0)$ has a loop. Define $X = V$, $Y = \{\mathcal{H}(M)\}$. For each m, take $a(\mathcal{H}(m)) = \frac{1}{5}mnd$ and $b(\mathcal{H}(m)) = ndD$. The value $a(\mathcal{H}(0)) = 0$

violates the rules, but since the value of $a(\,)$ at a sink plays no part in the problem, we will keep it.

Bearing in mind the comment after the proof of Theorem 2, there is an optimal solution which is either a type-2 path solution starting at $\mathcal{H}(M)$ and finishing at $\mathcal{H}(0)$, or a type-4 path solution starting at $\mathcal{H}(M)$ and finishing at some $\mathcal{H}(m)$ with $m < 5D$. The last condition comes from the need to have $a(\mathcal{H}(m)) < b(\mathcal{H}(m))$.

Let $M = t_0 > t_1 > \cdots > t_q$ be a sequence with $0 \le t_{i+1} - t_i \le 4k$ and $0 \le t_q < 5D$. This defines a path solution $S = (s, N)$ for the path $(v_0, v_1, \ldots, v_q)$, where $v_i = \mathcal{H}(t_i)$ for all i. Scale the path solution so that $N(v_0) = a(v_0)/(a(v_0) - b(v_0))$. Then

$$N(Y) = \frac{\frac{1}{5}Mnd}{\frac{1}{5}Mnd - ndD} = \frac{M}{M - 5D} \le 6. \tag{7}$$

Moreover, for $0 \le i \le q$,

$$N(v_i) \ge N'(v_i) = \frac{t_0 t_1 \cdots t_{i-1}}{(5D)^i}$$

whether or not we have a type-2 or type-4 path solution. If $t_j, t_{j-1}, \ldots, t_q$ are decreased by 1, for any $j > 0$, then $\sum_{v_i \in X} N'(v_i)$ is decreased, and similarly this sum is decreased if v_q is removed from the sequence. Therefore, the least value of $\sum_{v_i \in X} N'(v_i)$ occurs when $t_i = M - 4ik$ for $0 \le i \le q$ and t_q is the first number in this sequence less than $5D$. In this case, we have

$$\begin{aligned} N'(v_i) &= \frac{M(M-4k)\cdots(M-4(i-1)k)}{(5D)^i} \\ &= \Big(\frac{4k}{5D}\Big)^i \Big[\frac{M}{4k}\Big]_i \\ &= \Big(\frac{4k}{5D}\Big)^{M/4k} \Big(\frac{M}{4k}\Big)!\, \Big(\frac{5D}{4k}\Big)^j \Big/ j!\,, \end{aligned} \tag{8}$$

where $j = M/4k - i$ is not necessarily an integer. Define $R = \lceil \sqrt{5D/4k}\, \rceil$. We will bound $\sum_{v_i \in X} N'(v_i)$ from below by including only the R terms with $5D < t_i \le 5D + 4kR$. These correspond to $5D/4k < j \le 5D/4k + R$. The value of (8) is decreasing with respect to j in that range, so we can apply Lemma 2 with $x = 5D/4k$ to obtain

$$N(X) \ge \tfrac{5}{28} \Big(\frac{4k}{5D}\Big)^{M/4k} \Big(\frac{M}{4k}\Big)!\, \exp\Big(\frac{5D}{4k}\Big).$$

Combining this with (7), we obtain

$$f(S) \le 34 \exp\Big(-\frac{5D}{4k}\Big) \Big(\frac{5D}{4k}\Big)^{M/4k} \Big/ \Big(\frac{M}{4k}\Big)!\,.$$

Noting that $M/k \ge 6D/k \ge 16$, the required result now follows from Stirling's formula. □

A related result for fixed d is obtained by different means in [1].

7 Generalizations

A number of generalizations of the problem can be considered. One of them results from replacing the functions $a, b : V \to \mathbb{R}_{>0}$ by functions $\alpha, \beta : E \to \mathbb{R}_{>0}$. Inequalities (1)–(2) are then replaced by

$$\sum_{uv \in G^-(v)} \beta(uv)s(uv) \le N(v) \quad \text{for all } v \text{ such that } G^-(v) \ne \emptyset,$$

$$\sum_{vw \in G^+(v)} \alpha(vw)s(vw) \ge N(v) \quad \text{for all } v \text{ such that } G^+(v) \ne \emptyset.$$

The analysis of this generalization is almost the same, with essentially the same result. $\alpha(vw)^{-1}$ and $\beta(uv)^{-1}$ play the roles of $a(v)$ and $b(v)$, respectively. The optimal solution can again be assumed to be a path solution, defined in the obvious manner. This seems very natural, but we don't have any practical examples.

Other generalizations involve allowing extra constraints. One possibility is to admit constraints of the form $s(vw) \le As(vw')$ for constant A. This can arise when one can prove that at most some fraction of switchings applied to $C(v)$ take us to $\mathcal{C}(w)$ rather than to $\mathcal{C}(w')$. Examples of where such a situation is analysed include [8, 11]. However, the optimal solution can now have a form more complicated than a path.

References

[1] P. Collet and J.-P. Eckmann, The number of large graphs with a positive density of triangles, *J. Statist. Phys.* **109** (2002), 923–943.

[2] C. Cooper, A. Frieze, B. Reed and O. Riordan, Random regular graphs of non-constant degree: independence and chromatic number, *Combin. Probab. Comput.* **11** (2002), 323–341.

[3] C. D. Godsil and B. D. McKay, Asymptotic enumeration of Latin rectangles, *J. Combin. Theory Ser. B* **48** (1990), 19–44.

[4] C. Greenhill and B. D. McKay, Asymptotic enumeration of sparse non-negative integer matrices with specified row and column sums, *in preparation.*

[5] M. Krivelevich, B. Sudakov, V. H. Vu and N. C. Wormald, Random regular graphs of high degree, *Random Structures Algorithms* **18** (2001), 346–362.

[6] B. D. McKay, Subgraphs of random graphs with specified degrees, *Congr. Numer.* **33** (1981), 213–223.

[7] B. D. McKay, Asymptotics for 0-1 matrices with prescribed line sums, *in* Enumeration and Design, (Academic Press, 1984), 225–238.

[8] B. D. McKay and I. Wanless, Most Latin squares have many subsquares, *J. Combin. Theory Ser A* **86** (1999), 323–347.

[9] B. D. McKay, N. C. Wormald and B. Wysocka, Short cycles in random regular graphs, *Electron. J. Combin.* **11** (2004), #R66, 12 pages.

[10] C. Stein, Asymptotic evaluation of the number of Latin rectangles, *J. Combin. Theory Ser. A* **25** (1978), 38–49.

[11] N. C. Wormald, The perturbation method and triangle-free random graphs, *Random Structures Algorithms* **9** (1996), 253–270.

(Received 2 Sep 2006)