J. Algebra Comb. Discrete Appl.
2(2) • 117-149
Received: 1 December 2014; Accepted: 22 April 2015
DOI 10.13069/jacodesmath.28239

Journal of Algebra Combinatorics Discrete Structures and Applications

Identifying long cycles in finite alternating and symmetric groups acting on subsets

Research Article

Steve Linton[1]*, Alice C. Niemeyer[2], Cheryl E. Praeger[3]§**

1. School of Computer Science, University of St. Andrews, North Haugh, St. Andrews, Fife, KY16 9SX, Scotland

2. Department of Mathematics and Statistics, Maynooth University, Co. Kildare, Ireland

3. Centre for the Mathematics of Symmetry and Computation, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia

Abstract: Let H be a permutation group on a set Λ, which is permutationally isomorphic to a finite alternating or symmetric group A_n or S_n acting on the k-element subsets of points from $\{1,\dots,n\}$, for some arbitrary but fixed k. Suppose moreover that no isomorphism with this action is known. We show that key elements of H needed to construct such an isomorphism φ, such as those whose image under φ is an n-cycle or $(n-1)$-cycle, can be recognised with high probability by the lengths of just four of their cycles in Λ.

2010 MSC: 20B30, 60C05, 20P05, 05A05

Keywords: Symmetric and alternating groups in subset actions, Large base permutation groups, Finding long cycles

1. Introduction

The second and third authors predicted in [9] that, for a permutation group H on a set Λ, which is permutationally isomorphic to a symmetric group S_n acting on the k-element subsets of points from $\{1,\dots,n\}$ (that is in its k-set action), for some arbitrary but fixed k, it should be possible to recognise an element in H corresponding to an n-cycle in S_n by the lengths of just four of its cycles in Λ. The purpose of this paper is to prove this result.

Theorem 1.1. *Let H be a permutation group on a set Λ, which is permutationally isomorphic, via an unknown isomorphism φ, to a finite symmetric group S_n in its k-set action, for some k. Let h be*

* *E-mail: sal@cs.st-andrews.ac.uk*
** *E-mail: alice.niemeyer@nuim.ie*
§ *E-mail: Cheryl.Praeger@uwa.edu.au*

a uniformly distributed random element of H and let $\lambda_1, \ldots, \lambda_4$ be independent, uniformly distributed random points of Λ. Then there exist positive constants N_0 and c such that, for $n \geq N_0$,

$$\text{Prob}\left(\begin{array}{l}\varphi(h) \text{ is an} \\ n\text{-cycle}\end{array} \;\middle|\; \begin{array}{l}\text{the } h\text{-cycle containing } \lambda_i \text{ has} \\ \text{length } n, \text{ for } i = 1, \ldots, 4\end{array}\right) > 1 - \frac{c}{n^{\frac{1}{6}}}.$$

Subset actions of S_n and the alternating group A_n play a crucial role in algorithms for permutation groups. They are examples of 'large-base' actions. Most primitive permutation groups are 'small-base' and very efficient algorithms are available to compute with them (for a detailed definition see [12, p. 51]). However these algorithms become prohibitively expensive when applied to large-base groups and, therefore, alternative means of handling large-base groups are essential (see [12, Chapter 10] for a discussion on currently available algorithms for this case). The large-base primitive permutation groups all contain in their socles alternating groups with associated subset actions. Hence finding efficient algorithms for these actions is important.

The probabilistic algorithm described in [6] recognises alternating and symmetric groups in their actions on k-sets constructively. It takes as input a group H and an integer n. Under the assumption that H is isomorphic to S_n, the algorithm examines a number of random elements of H seeking to find *key-elements*, namely elements for which good estimates for their proportions in S_n are known and which have particular properties. Should this search fail, the algorithm concludes that the assumption that H is isomorphic to S_n is incorrect and reports that H is not isomorphic to S_n. Otherwise, it attempts to construct an isomorphism from H to S_n. Since this algorithm is randomised, there is a possibility that the search appears to succeed but in fact has not found suitable elements. In most settings this will be detected as part of the larger algorithm, see [6].

The theoretical underpinning of the algorithm described above is to determine very good bounds for the probability of finding the key-elements in H under the assumption that H is isomorphic to S_n. This is the purpose of the current paper. For the connection between estimation results and probabilistic algorithms in the context of recognition algorithms for groups see [10].

The groups the algorithm can take as input need not be permutation groups. Rather, they may belong to a more general class of groups called *groups of black box permutations* (see [6]). This allows the algorithm to be employed in different computational models, for example, to deal with groups given by a set of generating matrices, which are permutation isomorphic to S_n acting on the underlying vector space, without converting such a group into a permutation group.

The crucial requirement is the ability to compute the image of a point under the action of a generator of the group. Suppose that t is an upper bound for the time taken to perform this action, which might be viewed as a black box procedure. Then the time required to determine that a word of length r in the group generators permutes a given point in a cycle of length m is at most mrt. In some contexts this time can even be very much less than the time required to find the product of two group generators.

For suitable n and k the algorithms in [6] have running time (excluding any time needed to read the input) growing significantly more slowly than $\binom{n}{k}$, implying that they can use the input black box permutations only through computing their action on a selection of the $\binom{n}{k}$ points, and not through examining any other aspect of their structure, or computing any other elements of the group they generate. For example, checking that the n-th power of an element is the identity may already be more expensive than our entire algorithm.

The key elements sought by the algorithm to recognise A_n or S_n in its k-set action are elements containing an m-cycle for large m, as described in Table 1. Our Theorem 1.1 follows from a more general theorem, Theorem 3.1, which determines the probability of finding elements of each of these types among a certain number of random elements of H. Once these key elements have been found, a permutational isomorphism from H on Λ to A_n or S_n on k-sets can be implemented using the methods described in [2, Section 4], especially Method B, or those described in [1, Sections 4 and 5], especially Lemmas 4.1, 4.3 and 5.5. Alternative methods, focussing in particular on the k-set action, are developed in [6].

1.1. Context of our results

In a seminal collection of papers, Erdös and Turán initiated the study of asymptotic behaviour of the proportions of various kinds of elements in permutation groups. For example, they showed [4, 5] that for n large enough, most elements in the symmetric group S_n of degree n have order $n^{(\frac{1}{2}+o(1))\log(n)}$. In the same vein Warlimont [13] proved that the conditional probability that a random element g in S_n is an n-cycle, given that $g^n = 1$, is $1 - O(n^{-1})$.

Applied algorithmically, Warlimont's result is used to conclude, from the fact that the nth power of a 'hidden' permutation $g \in S_n$ is the identity, that g is almost certainly an n-cycle. Finding an n-cycle is a key step in many algorithms that 'constructively recognise' S_n, so this is valuable. However, testing whether $g^n = 1$ requires $\log(n)$ multiplications of black box permutations. In computational models where a single multiplication costs about $\binom{n}{k}$ operations, employing this approach would not yield an algorithm whose running time grows significantly more slowly than $\binom{n}{k}$. The results of [9] provide a basis for extending this to a situation where we know only that $\langle g \rangle$ has an orbit of length n in some action. An extension of this nature to k-set actions of S_n and A_n is the subject of this paper. We refine and improve significantly the main result of [9]. For example, we employ a similar division of the elements of S_n into several families according to properties of points which lie in cycles of lengths dividing n. However, examining this subdivision alone is not sufficient to achieve the results in this paper. We need to study the probability that several k-element subsets of $\{1, \ldots, n\}$ have exactly n distinct images under $\langle g \rangle$ for g an element in one of the families. Moreover, in our algorithmic applications we also required analogous results for elements of S_n and A_n containing m-cycles, for $m \geq n - 6$.

Line	G	n	m	r	key-elements	$\rho(G, n, m)$
1			n	1	n-cycle	1
2	S_n	odd	$n-2$	2	2-cycle	1
3		even	$n-3$	2	2-cycle	2/3
4		odd	n	1	n-cycle	1
5		even	$n-1$	1	(n–1)-cycle	1
6	A_n	2 or 4 (mod 6)	$n-3$	3	3-cycle	1
7		3 or 5 (mod 6)	$n-4$	3	3-cycle	3/4
8		0 (mod 6)	$n-5$	3	3-cycle	7/20
9		1 (mod 6)	$n-6$	3	3-cycle	9/40

Table 1. Groups and types of elements

In Section 2 we briefly describe the algorithmic application, and in particular we explain the meaning of the parameters r and ρ in Table 1. In Section 3 we introduce the notation which we shall use throughout the paper and give the precise statement of the main result (Theorem 3.1). The proof of Theorem 3.1 (and hence of Theorem 1.1) is given in Section 4. In particular we exhibit an explicit value for the constant c of Theorem 3.1(a) and Theorem 1.1. We present some background material in Sections 5 and 6. Sections 7 - 11 contain the various parts which are pulled together in Section 4 for the proof of Theorem 3.1.

2. Algorithmic application

The results in this paper are motivated by algorithmic applications in [6] and [7]. In these applications, H is a permutation group acting on a set Λ of $\binom{n}{k}$ points. We wish to test whether H is permutation isomorphic to $G = A_n$ or $G = S_n$ acting on the set $\binom{\Omega}{k}$ of k-element subsets of $\Omega = \{1, \ldots, n\}$. That is

to say, whether there is a group isomorphism $\varphi : H \to G$ and a bijection $f : \Lambda \to \binom{\Omega}{k}$ such that, for each $h \in H$ and $\lambda \in \Lambda$, $(\lambda^h)f = (\lambda)f^{h\varphi}$. These isomorphisms will be constructed in the form of a computer program rather than listing the image of each element.

We say that an element $h \in H$ *corresponds* to an element $g \in G$ if the permutation isomorphism φ maps h to g. The algorithms construct a 'nice generating' set for H of size 2. In the case where H is permutation isomorphic to S_n in its action on $\binom{\Omega}{k}$, this generating set consists of elements that, in the natural representation of S_n on n points, correspond to an n-cycle and a 2-cycle interchanging two consecutive points of the n-cycle. In the case where H is permutation isomorphic to A_n in its action on $\binom{\Omega}{k}$ the nice generating set consists of elements that in A_n correspond to an n-cycle or $(n-1)$-cycle, and to a 3-cycle.

We wish to find these elements by selecting independent, uniformly distributed random elements from the group H. However, the proportion of 2-cycles in S_n or 3-cycles in A_n is too small to allow us to find such elements directly by random selection. Therefore, we seek elements in H which correspond to permutations containing a 2-cycle or a 3-cycle together with one long cycle of length m, say, where m is at least $n-6$ and m is coprime to 2 or 3, respectively. The algorithms in [6] and [7] seek elements $h \in H$ which correspond to the kinds of elements g listed in Table 1, where H is permutation isomorphic to $G = S_n$ or $G = A_n$, with G, n as in the second and third columns. The fourth column, labelled m, lists the length of the m-cycle which the element g contains. The fifth column, labelled r, lists an integer between 1 and 3. Ultimately we wish to find an element h in H which corresponds to an element in G with cycle type as recorded in the sixth column. This element is constructed as a power of the element h.

The first element in the nice generating set for H corresponds to an element satisfying the conditions of Line 1, 4, or 5, namely it corresponds to an n-cycle or an $(n-1)$-cycle. The second nice generator corresponds to a 2-cycle if $G = S_n$ and is constructed from an element $h \in H$ which corresponds to g as in Line 2 or 3. If $G = A_n$, the second nice generator corresponds to a 3-cycle and is constructed from $h \in H$ corresponding to an element g as in Line 6, 7, 8 or 9. The last column, labelled $\rho(G, n, m)$, records a rational number such that the proportion of elements h of H which correspond to elements of G containing an m-cycle and with order dividing rm is $\frac{\rho(G,n,m)}{m}$ (see (1)).

The group H acts on a set Λ of size $|\Lambda| = \binom{n}{k}$, and in the context of the algorithm m, n and k are so large that it is 'too expensive' to compute the full cycle structure of elements of H in their action on Λ. Instead we compute the cycle lengths of elements $h \in H$ on a handful of randomly chosen points of Λ, that is to say, we 'trace' these points under the action of $\langle h \rangle$.

In computer experiments in GAP [3], we discovered that if H is permutation isomorphic to $G = S_n$ or A_n on $\binom{\Omega}{k}$ then, for m, r as in one of the lines of Table 1, most elements of H which produced cycles of lengths a multiple of m and dividing rm, when we traced each of four or five independent random points of Λ, corresponded to elements of G containing an m-cycle. This computer experiment is formalised in procedures FindMCycle and TraceCycle. Our experimental observation turns out to be true in general, and is proved in Theorem 3.1 for sufficiently large n. Our experiments also suggest that the results hold for smaller values of n. For clarity of exposition the proofs of Theorem 3.1 are written in terms of the action of G on $\binom{\Omega}{k}$.

For n, m and r as in one of the lines of Table 1, define $\mathcal{N}(n, m)$ to be the set of all $g \in S_n$ that contain an m-cycle and $\mathcal{N}_{good}(G, n, m)$ to be the set of all $g \in \mathcal{N}(n, m) \cap G$ for which $o(g)$ divides rm. Note that, for given G, n, m, only one of the lines of Table 1 is satisfied, and hence r is determined by G, n, m. We define $\rho(G, n, m)$ to be the rational number satisfying

$$\frac{|\mathcal{N}_{good}(G, n, m)|}{|G|} = \frac{\rho(G, n, m)}{m}. \tag{1}$$

As an example of how to interpret this information, consider Line 3 of Table 1. The proportion of elements g of S_n containing an $(n-3)$-cycle is $\frac{1}{n-3}$, and $2/3$ of these elements contain also a 2-cycle or three 1-cycles on the remaining 3 points. Thus the proportion of elements of S_n containing an $(n-3)$-cycle and having order dividing $2(n-3)$ is $\frac{2/3}{n-3} = \frac{\rho(S_n,n,n-3)}{n-3}$. In order to construct a 2-cycle (the entry in column 6 for

this line), we raise the element g to the $(n-3)^{\text{rd}}$ power producing $x = g^{n-3}$. Since $n-3$ is odd, the element x is the identity if g has three fixed points, a 2-cycle if g contains a 2-cycle, or possibly a 3-cycle if g contains a 3-cycle and 3 does not divide n. Thus three quarters of the elements of $\mathcal{N}_{good}(S_n, n, n-3)$ yield a 2-cycle by powering. The algorithm FindMCycle can therefore easily be incorporated into a Monte Carlo algorithm to construct a transposition in this case: by repeating FindMCycle a number of times we will with high probability construct a transposition by powering the output of FindMCycle. The other lines have a similar interpretation for $\rho(G, n, m)$.

We now describe the two algorithms. Algorithm 1 assumes that we have a function RandomGrpElt which takes as input a generating set Y for a group H and returns independent, uniformly distributed random elements of H. Algorithm 2 assumes that we have a function RandomPoint which takes as input a finite set Λ and returns independent, uniformly distributed random points of Λ. Note that Algorithm 1 calls Algorithm 2 and that we assume that Algorithm 2 has access to the variables of Algorithm 1.

Algorithm 1: FindMCycle$(n, m, r, H, \Lambda, \varepsilon, M)$

Data: Let (n, m, r) be as in one of the lines of Table 1. Let H be a permutation group with a generating set Y acting on a finite set Λ. Let ε be a real number with $0 < \varepsilon < 1$ and let M be an integer with $M \geq 4$.
Result: An element $h \in H$ or `fail`;

This algorithm inspects up to $O(n \log(\varepsilon^{-1}))$ uniformly distributed independent random elements from H to find one which has orbits of length a multiple of m and dividing rm on each of M randomly selected points from Λ. If such an $h \in H$ is found it returns h, otherwise it returns `fail`.

Set $N := \lceil 5n \log(\frac{2}{\varepsilon}) \rceil$;
for $i = 1, \ldots, N$ **do**
  $h_i :=$ RandomGrpElt(Y);
  if TraceCycle$(h_i) =$ `true` **then**
    return h_i;
return `fail`;

Algorithm 2: TraceCycle(h)

Data: A permutation $h \in H$;
Result: A boolean '`true`' or '`false`'

This algorithm tests whether the permutation $h \in H$ has orbits of length a multiple of m and dividing rm on M randomly selected points from Λ. If this is the case it returns `true`*, otherwise it returns* `false`.

for $i = 1, .., M$ **do**
  $\lambda_i :=$ RandomPoint(Λ);
Put $\Gamma = \{\lambda_j\}_{j=1}^{M}$;
for $\lambda \in \Gamma$ **do**
  if $|\lambda^{\langle g \rangle}| \neq r_0 m$ for some $r_0 \mid r$ **then**
    return `false`;
return `true`;

Remark 2.1. *(a) The number M of random points of Λ tested in the algorithm* TraceCycle *is often a bounded constant (as, for example, in Theorem 3.1), but in our analysis we allow it to be as large as $O(n)$, see (2).*

(b) The algorithm TraceCycle *performs $O(n)$ image computations to check whether $|\lambda^{\langle g \rangle}| = r_0 m$, for each random point λ. Thus if ξ_{rp}, ξ_{rge}, ν_{im}, are upper bounds for the costs of producing a random point using* RandomPoint*, producing a random group element using* RandomGrpElt*, and computing the image of a point of Λ under an element of H, respectively, then the cost of* FindMCycle *is*

$$O(n \log(\varepsilon^{-1})(\xi_{rge} + M\xi_{rp} + Mn\nu_{im})).$$

This cost is modest when compared with the cost $\binom{n}{k}\nu_{im}$ *of computing the product of two permutations of* Λ *(especially when* $k = O(n)$*) or the cost of directly computing the order of any element.*

Our main result Theorem 3.1 shows that these simple and inexpensive procedures provide an effective way to find and identify elements of S_n and A_n containing m-cycles from their actions on k-element subsets.

3. Statement of the main theorem and notation

In order to state our main theorem we introduce several parameters that are used throughout the paper. Suppose that the triple (G, n, m) satisfies one of the lines of Table 1, and note that r is determined by G, n, m. The integer M used in the algorithm FindMCycle is assumed to satisfy

$$4 \le M \le \log\left(\frac{9}{8}\right)\frac{n-2}{2}. \tag{2}$$

Let $d(x)$ be the number of positive divisors of an integer x. By [11, pp. 395-396], $d(x) = x^{o(1)}$. In fact, for every $\delta > 0$, there is a positive constant c_δ such that

$$d(x) \le c_\delta x^\delta \tag{3}$$

for all x. Choose real numbers δ and s satisfying

$$0 < \delta < \min\{1-s, \frac{s}{3}, s-\frac{1}{2}\} \text{ and } \frac{1}{2} < s < \frac{M-1}{M}. \tag{4}$$

Further let

$$\ell = \min\left\{M(1-s), 3-2s-2\delta, 1+s-3\delta, 2s-2\delta\right\}. \tag{5}$$

By (4), all of $M(1-s) > 1$, $3-2s-2\delta > 1$, $1+s-3\delta > 1$ and $2s-2\delta > 1$ hold. Hence $\ell > 1$. Next we define the constant a_δ by

$$a_\delta := \frac{5}{4}\left(1 + 3\frac{c_\delta}{150^{s-\delta}} + \left(\frac{c_\delta}{150^{s-\delta}}\right)^2\right), \tag{6}$$

with c_δ as in (3), and the constant $b_{M,\delta,s}$, which we usually abbreviate to b_M, by

$$b_M = \left(\frac{33}{8}\right)^M + 72\, a_\delta c_\delta^2 r^{2s+2\delta} + 6.24\, a_\delta c_\delta^3 r^{3\delta} + \frac{c_\delta^2}{r^{2s-2\delta}} + \left(\frac{31}{r^{1-s}}\right)^M. \tag{7}$$

The theorem involves an 'error probability' ε, that is, a real number satisfying $0 < \varepsilon < 1$. We assume that the integer n satisfies the following inequalities:

$$n \ge \begin{cases} 12(rn)^s + 6 \\ (rn)^s \log n \\ \left(\frac{10 b_M}{\varepsilon}\right)^{1/(\ell-1)} \end{cases}. \tag{8}$$

Theorem 3.1. *Let* (G, n, m) *be as in one of the lines of Table 1, and let* k *be a positive integer satisfying* $2 \le k \le n/2$*. Suppose that* H *is a permutation group permutation isomorphic to* G *acting on* k*-element subsets of* $\{1, \dots, n\}$ *(via the unknown isomorphism* $\varphi : H \to G$*). Then the following hold*

(a) Let h *be a uniformly distributed random element of* H *and let* $\lambda_1, \dots, \lambda_4$ *be independent, uniformly distributed random points of* Λ*. Then there exist positive constants* N_0 *and* c *such that, for* $n \ge N_0$*,*

$$\text{Prob}\left(\begin{array}{l|l} \varphi(h) \text{ contains an} & \text{for } i = 1, \dots, 4 \text{ the} \\ m\text{-cycle} & h\text{-cycle on } \lambda_i \text{ has length} \\ & r_i m \text{ for some } r_i \mid r \end{array}\right) > 1 - \frac{c}{n^{\frac{1}{6}}}.$$

(b) Let M be an integer satisfying (2), and let s, δ be real numbers satisfying (4), and ℓ as in (5). Then FindMCycle *is a Monte Carlo Algorithm which, given as input the permutation group H, an error probability $\varepsilon > 0$ and the integer M, returns an output h such that, provided n satisfies (8),*

(i) the probability that $h \in H$ and $\varphi(h)$ contains an m-cycle is at least $1 - \varepsilon$,

(ii) the probability that $h \in H$ and $\varphi(h)$ does not contain an m-cycle is at most $\varepsilon/2$, and

(iii) the probability that $h =$ Fail is at most $\varepsilon/2$.

Notation 3.2. *For the rest of the paper we assume that n, m, r and G are as in one of the lines of Table 1, noting that r is determined by G, n, m. Let M be an integer satisfying (2), let s, δ be real numbers satisfying (4), and let ℓ, c_δ, a_δ and b_M be as in (5), (3), (6) and (7) respectively.*

Let S_n act naturally on $\Omega = \{1, 2, \ldots, n\}$. Let k and k_0 be positive integers satisfying $2 \le k \le n/2$, and $1 \le k_0 \le k$. A k_0-element subset of Ω is called a k_0-subset.

We use the notation in Table 2 to describe an element $g \in S_n$, where γ_0 is a k_0-subset of Ω. Here we identify a cycle of g with the subset of Ω it permutes.

$c_{k_0}(\gamma_0, g)$	length of the g-cycle containing γ_0 on k_0-subsets
s-small g-cycle	g-cycle in Ω of length less than $(rn)^s$
s-large g-cycle	g-cycle in Ω of length at least $(rn)^s$
$\Delta(g)$	union of g-cycles in Ω whose lengths divide rm
$\Sigma(g)$	$\Omega \setminus \Delta(g)$
v	cardinality of $\Delta(g)$
u	cardinality of $\Sigma(g)$

Table 2. Table for Notation 3.2

We define in Table 3 several classes of elements in G. We usually omit mentioning n and m in our notation. For example, we refer to $\mathcal{N}(n, m)$ (defined in Section 2) simply as $\mathcal{N}$ and to $\mathcal{N}_{good}(G, n, m)$ simply as $\mathcal{N}_{good}$.

$\mathcal{N}$	set of all $g \in S_n$ that contain an m-cycle
$\mathcal{N}_{good}$	set of all $g \in \mathcal{N} \cap G$ for which $o(g)$ divides rm
$\mathcal{F}$	set of all $g \in G \setminus \mathcal{N}$ such that $m \mid o(g)$
$\mathcal{R}$	set of all $g \in \mathcal{F}$ such that $\lvert\Delta(g)\rvert \le 4(rn)^s$
$\mathcal{S}_0$	set of all $g \in \mathcal{F}$ such that $\lvert\Delta(g)\rvert > 4(rn)^s$ and all g-cycles in $\Delta(g)$ are s-small
$\mathcal{S}_1^+$	set of all $g \in \mathcal{F}$ such that $\lvert\Delta(g)\rvert > 4(rn)^s$, exactly one g-cycle C in $\Delta(g)$ is s-large, and $\lvert\Delta(g) \setminus C\rvert > 3(rn)^s$
$\mathcal{S}_1^-$	set of all $g \in \mathcal{F}$ such that $\lvert\Delta(g)\rvert > 4(rn)^s$, exactly one g-cycle C in $\Delta(g)$ is s-large, and $\lvert\Delta(g) \setminus C\rvert \le 3(rn)^s$
$\mathcal{S}_{\ge 2}$	set of all $g \in \mathcal{F}$ such that $\lvert\Delta(g)\rvert > 4(rn)^s$ and at least two g-cycles in $\Delta(g)$ are s-large

Table 3. Families of Elements

Remark 3.3. *(a) The definition of a_δ is not too critical. We simply need a_δ to be greater than or equal to the right hand side of (6) for the values of rm we are considering, see Remark 8.2 and Lemma 8.3. For example, if $rm \geq c_\delta^{1/(s-\delta)}$ then we may take $a_\delta = 25/4$.*

(b) Currently Equation (8) limits the practical applicability of Theorem 3.1 severely, but we note that in our analysis we allow k to be as large as $n/2$. The first two inequalities of (8) imposed on n are due to the subdivision of the set of permutations of order divisible by m into disjoint subsets which depend on s. We give a uniform proof that holds for all values of k in the range $2 \leq k \neq n/2$. If, for example, k were bounded as n increases, then several of the arguments would be simpler and the constraints on n correspondingly less severe.

(c) The main constraint forcing n to be very large is the third inequality in (8). For example, for our parameter choice in Theorem 3.1, namely $M = 4, s = \frac{17}{24}$ and $\delta = \frac{1}{6}$, we have $c_\delta \leq 138.32$ and, for n large enough, $a_\delta = \frac{25}{4}$. In this case we find $b_M > 2 \cdot 10^8$ and the last inequality of (8) dictates $n > 3.3 \cdot 10^{112}/\varepsilon^{12}$. Moreover, even though a larger value of M allows us to choose a smaller value for c_δ, the choice might result in a smaller value for ℓ, which in turn has undesired consequences, making b_M larger, and hence requiring n to be larger.

4. Proof of the main theorem

The proof of the main theorem, Theorem 3.1, relies on many supporting results. In this section we subdivide the proof into various parts and show how these parts are then brought together to give a complete proof. The individual parts of the proof are proved in later sections. The main idea of the proof is to divide the elements of S_n that could possibly be returned by FindMCycle into disjoint families, and to compute the probability that TraceCycle returns `true` for an element of each of these families. The families of elements in this subdivision are defined in Table 3, namely $\mathcal{N}, \mathcal{R}, \mathcal{S}_0, \mathcal{S}_1^+, \mathcal{S}_1^-, \mathcal{S}_{\geq 2}$, and we use the notation introduced in this table throughout the paper.

Proof of Theorem 3.1(b). We prove this theorem by analysing the algorithm FindMCycle. Let $N = \lceil 5n \log(\frac{2}{\varepsilon}) \rceil$. A call to algorithm FindMCycle can terminate in one of three possible ways:

($\mathcal{G}$) For some i with $1 \leq i \leq N$ the i-th iteration of the **for**-loop returns an element in $\mathcal{N}$. We call this a *good* outcome.

($\mathcal{B}$) For some i with $1 \leq i \leq N$ the i-th iteration of the **for**-loop returns an element which is not in $\mathcal{N}$. We call this a *bad* outcome.

($\mathcal{U}$) The **for**-loop is executed N times and TraceCycle returns `false` for each of the selected random elements. In this case the algorithm returns Fail. We call this an *ugly* outcome.

Thus to prove the three parts of Theorem 3.1 we must prove

$$\text{Prob}(\mathcal{G}) \geq 1 - \varepsilon, \quad \text{Prob}(\mathcal{B}) \leq \varepsilon/2, \quad \text{Prob}(\mathcal{U}) \leq \varepsilon/2.$$

Clearly any two of these inequalities implies the third. We shall therefore prove only $\text{Prob}(\mathcal{B}) \leq \varepsilon/2$ and $\text{Prob}(\mathcal{U}) \leq \varepsilon/2$. To study these outcomes more closely we define the following events.

E_i the i-th iteration of the **for**-loop is executed. Let g_i denote the random element selected in the i-th iteration.
G_i event E_i occurs, $g_i \in \mathcal{N}$ and TraceCycle(g_i) = `true`
B_i event E_i occurs, $g_i \notin \mathcal{N}$ and TraceCycle(g_i) = `true`
U_i event E_i occurs and TraceCycle(g_i) = `false`

Note that $E_i = G_i \dot\cup B_i \dot\cup U_i$ and that $\mathrm{Prob}(E_1) = 1$. Further, for $i > 1$ we have that

$$E_i = U_1 \cap \ldots \cap U_{i-1} = U_{i-1}. \tag{9}$$

Thus

$$\begin{aligned} \mathcal{G} &= G_1 \vee G_2 \vee \ldots \vee G_N \\ \mathcal{B} &= B_1 \vee B_2 \vee \ldots \vee B_N \\ \mathcal{U} &= U_1 \wedge U_2 \wedge \ldots \wedge U_N = U_N. \end{aligned} \tag{10}$$

Proof that $\mathrm{Prob}(\mathcal{U}) \le \varepsilon/2$*:* For a uniformly distributed random element $g \in G$, let

$$\begin{aligned} p_1 &= \mathrm{Prob}(\textsc{TraceCycle}(g) = \texttt{false} \mid g \in \mathcal{N}_{good}) \\ p_2 &= \mathrm{Prob}(\textsc{TraceCycle}(g) = \texttt{false} \mid g \notin \mathcal{N}_{good}) \end{aligned}$$

and let $p = \frac{\rho}{m}p_1 + \frac{m-\rho}{m}p_2$, where $\rho := \rho(G, n, m)$ (see Table 1), the proportion of elements of G containing an m-cycle that have order dividing rm. Note that, since the proportion of elements containing an m-cycle in S_n is $1/m$, we have $\mathrm{Prob}(g \in \mathcal{N}_{good}) = \frac{\rho}{m}$.

Given E_i, the event U_i is the disjoint union of the events U_{i1}, that $g_i \in \mathcal{N}_{good}$ and $\textsc{TraceCycle}(g_i) = \texttt{false}$, and U_{i2}, that $g_i \notin \mathcal{N}_{good}$ and $\textsc{TraceCycle}(g_i) = \texttt{false}$. Thus

$$\begin{aligned} \mathrm{Prob}(U_i \mid E_i) &= \frac{\rho}{m}\mathrm{Prob}(\textsc{TraceCycle}(g_i) = \texttt{false} \mid g_i \in \mathcal{N}_{good}) \\ &+ \frac{m-\rho}{m}\mathrm{Prob}(\textsc{TraceCycle}(g_i) = \texttt{false} \mid g_i \notin \mathcal{N}_{good}) \\ &= \frac{\rho}{m}p_1 + \frac{m-\rho}{m}p_2 = p. \end{aligned}$$

Note, in particular, that this probability is independent of i. By (9) we have $E_i = U_{i-1}$, and hence $\mathrm{Prob}(U_i) = \mathrm{Prob}(E_i)\mathrm{Prob}(U_i \mid E_i) = \mathrm{Prob}(U_{i-1}) \cdot p$. As this is true for all i with $1 \le i \le N$, we have

$$\mathrm{Prob}(U_i) = p^i, \tag{11}$$

and in particular,

$$\mathrm{Prob}(\mathcal{U}) = \mathrm{Prob}(U_N) = p^N.$$

The required inequality $\mathrm{Prob}(\mathcal{U}) \le \varepsilon/2$ holds whenever $p^N \le \varepsilon/2$. We now prove the latter inequality. By Proposition 7.5 we have $1 - p_1 \ge \left(\frac{n-2}{n}\right)^M$. Therefore,

$$p \le \frac{\rho}{m}p_1 + \frac{m-\rho}{m} = 1 - \frac{\rho}{m}(1-p_1) \le 1 - \frac{\rho}{n}\left(\frac{n-2}{n}\right)^M. \tag{12}$$

Now $N = \left\lceil 5n\log(\frac{2}{\varepsilon})\right\rceil = \left\lceil \frac{\log((\varepsilon/2)^{-1})}{(5n)^{-1}}\right\rceil$, and so by Lemma 5.2, $(1 - \frac{1}{5n})^N \le \varepsilon/2$. Thus $p^N \le \varepsilon/2$ holds if $1 - \frac{\rho}{n}\left(\frac{n-2}{n}\right)^M \le 1 - \frac{1}{5n}$, or equivalently, if $\left(\frac{n}{n-2}\right)^M \le 5\rho$. Since $\rho \ge 9/40$ (see Table 1), it is sufficient to prove that $\left(\frac{n}{n-2}\right)^M \le \frac{9}{8}$. By our assumption, $M \le \log(\frac{9}{8})\frac{n-2}{2}$, and hence

$$M\log\left(\frac{n}{n-2}\right) = M\log\left(1 + \frac{2}{n-2}\right) \le M\frac{2}{n-2} \le \log\left(\frac{9}{8}\right)$$

and exponentiating both sides gives the required inequality. Thus $p^N \le \varepsilon/2$ and hence $\mathrm{Prob}(\mathcal{U}) \le \varepsilon/2$ is proved.

Proof that $\text{Prob}(\mathcal{B}) \le \varepsilon/2$*:* Recall the definition of $\mathcal{B}$ in (10). Note that, if $\textsc{TraceCycle}(g) = \texttt{true}$, then $o(g)$ is divisible by m. Thus, by the definition of $\mathcal{F}$, for a uniformly distributed, random element $g \in G$,

$$\begin{aligned} q &:= \text{Prob}(g \in \mathcal{F} \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \\ &= \text{Prob}(g \notin \mathcal{N} \text{ and } \textsc{TraceCycle}(g) = \texttt{true}). \end{aligned} \tag{13}$$

Now, for all i with $1 \le i \le N$, we have that

$$\text{Prob}(B_i \mid E_i) = \text{Prob}(g_i \notin \mathcal{N} \text{ and } \textsc{TraceCycle}(g_i) = \texttt{true}) = q.$$

Hence $\text{Prob}(B_i) = \text{Prob}(E_i)\text{Prob}(B_i \mid E_i) = \text{Prob}(E_i)\, q$. If $i \ge 2$ then $E_i = U_{i-1}$ by (9), and so by (11), $\text{Prob}(B_i) = p^{i-1}q$. Therefore,

$$\begin{aligned} \text{Prob}(\mathcal{B}) &= \sum_{i=1}^{N} \text{Prob}(B_i) = q \sum_{i=1}^{N} p^{i-1} \\ &= q\frac{1-p^N}{1-p} < \frac{q}{1-p}. \end{aligned} \tag{14}$$

The most substantial part of the paper is devoted to finding an upper bound for q. It follows from Table 3 that

$$\mathcal{F} = \mathcal{R} \,\dot\cup\, \mathcal{S}_0 \,\dot\cup\, \mathcal{S}_1^+ \,\dot\cup\, \mathcal{S}_1^- \,\dot\cup\, \mathcal{S}_{\ge 2}.$$

Hence

$$q = q(\mathcal{R}) + q(\mathcal{S}_0) + q(\mathcal{S}_1^+) + q(\mathcal{S}_1^-) + q(\mathcal{S}_{\ge 2}),$$

where We estimate these proportions in Sections 8 - 11. Recall the definition of ℓ in (5), and that $\ell > 1$.

$q(\mathcal{R})$	=	$\text{Prob}(g \in \mathcal{R} \text{ and } \textsc{TraceCycle}(g) = \texttt{true})$
$q(\mathcal{S}_0)$	=	$\text{Prob}(g \in \mathcal{S}_0 \text{ and } \textsc{TraceCycle}(g) = \texttt{true})$
$q(\mathcal{S}_1^+)$	=	$\text{Prob}(g \in \mathcal{S}_1^+ \text{ and } \textsc{TraceCycle}(g) = \texttt{true})$
$q(\mathcal{S}_1^-)$	=	$\text{Prob}(g \in \mathcal{S}_1^- \text{ and } \textsc{TraceCycle}(g) = \texttt{true})$
$q(\mathcal{S}_{\ge 2})$	=	$\text{Prob}(g \in \mathcal{S}_{\ge 2} \text{ and } \textsc{TraceCycle}(g) = \texttt{true})$.

Table 4. Subdivision of the probability q of (13).

Define $b_M(\mathcal{R}) = \left(\frac{33}{8}\right)^M$ and note that $q(\mathcal{R}) = \text{Prob}(g \in \mathcal{R}) \cdot \text{Prob}(\textsc{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{R}) \le \text{Prob}(\textsc{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{R})$. Then Proposition 7.3 gives

$$q(\mathcal{R}) \le \frac{b_M(\mathcal{R})}{n^{M(1-s)}} \le \frac{b_M(\mathcal{R})}{n^{\ell}}.$$

Define $b_M(\mathcal{S}_0) = a_\delta c_\delta^2 r^{2s+2\delta} 72$. Then Proposition 8.4 and (3) give

$$q(\mathcal{S}_0) \le \frac{b_M(\mathcal{S}_0)}{n^{3-2s-2\delta}} \le \frac{b_M(\mathcal{S}_0)}{n^{\ell}}.$$

Define $b_M(\mathcal{S}_1^+) = a_\delta c_\delta^3 r^{3\delta} 6.24$. Then Proposition 9.1 and (3) give

$$q(\mathcal{S}_1^+) \le \frac{b_M(\mathcal{S}_1^+)}{n^{1+s-3\delta}} \le \frac{b_M(\mathcal{S}_1^+)}{n^{\ell}}.$$

Define $b_M(\mathcal{S}_{\geq 2}) = c_\delta^2 r^{2\delta-2s}$. Then Proposition 10.1 gives

$$q(\mathcal{S}_{\geq 2}) \leq \frac{b_M(\mathcal{S}_{\geq 2})}{n^{2s-2\delta}} \leq \frac{b_M(\mathcal{S}_{\geq 2})}{n^\ell}.$$

Define $b_M(\mathcal{S}_1^-) = \left(\frac{31}{r^{1-s}}\right)^M$. Then Proposition 11.1(b) yields

$$q(\mathcal{S}_1^-) \leq \frac{b_M(\mathcal{S}_1^-)}{n^{M(1-s)}} \leq \frac{b_M(\mathcal{S}_1^-)}{n^\ell}.$$

Thus by (7),

$$\begin{aligned} & b_M(\mathcal{R}) + b_M(\mathcal{S}_0) + b_M(\mathcal{S}_1^+) + b_M(\mathcal{S}_{\geq 2}) + b_M(\mathcal{S}_1^-) \\ &\leq \left(\frac{33}{8}\right)^M + a_\delta c_\delta^2 r^{2s+2\delta} 72 + a_\delta c_\delta^3 r^{3\delta} 6.24 + \frac{c_\delta^2}{r^{2s-2\delta}} + \left(\frac{31}{r^{1-s}}\right)^M \\ &= b_M \end{aligned}$$

and

$$q \leq \frac{b_M}{n^\ell}. \tag{15}$$

Remark 4.1. *We make a critical observation that the argument up to this point relies only on the first two inequalities of (8), and does not depend on the third inequality of (8).*

By (15) and the inequalities (14) and (12), we have that

$$\begin{aligned} \text{Prob}(\mathcal{B}) &< \frac{q}{1-p} \\ &\leq \frac{b_M}{n^\ell \frac{\rho}{n} \left(\frac{n-2}{n}\right)^M} \\ &= \frac{b_M}{\rho} \left(\frac{n}{n-2}\right)^M \frac{1}{n^{\ell-1}}. \end{aligned}$$

We showed above that $\left(\frac{n}{n-2}\right)^M \leq \frac{9}{8} \leq 5\rho$. Thus $\text{Prob}(\mathcal{B}) < \frac{5b_M}{n^{\ell-1}}$. By assumption $n \geq \left(\frac{10b_M}{\varepsilon}\right)^{1/(\ell-1)}$ and so this is at most $\varepsilon/2$. Hence $\text{Prob}(\mathcal{B}) < \varepsilon/2$. □

The proof of Theorem 3.1(a) requires a short argument applying Theorem 3.1(b).

Proof of Theorem 3.1(a). We use the algorithm TraceCycle with $M = 4$. Note first that the probability that a random element $h \in H$ corresponds to an element $g \in G$ containing an m-cycle, given that the h-cycles containing four random k-subsets $\lambda_1, \ldots, \lambda_4$ all have lengths of the form $r_i m$ with $r_i \mid r$, is $\text{Prob}(g \in \mathcal{N} \mid \text{TraceCycle}(g) = \texttt{true})$. Recall the definition of q in (13). Then

$$\begin{aligned} & \text{Prob}(g \in \mathcal{N} \mid \text{TraceCycle}(g) = \texttt{true}) \\ &= \frac{\text{Prob}(g \in \mathcal{N} \text{ and } \text{TraceCycle}(g) = \texttt{true})}{\text{Prob}(\text{TraceCycle}(g) = \texttt{true})} \\ &= \frac{\text{Prob}(\text{TraceCycle}(g) = \texttt{true}) - q}{\text{Prob}(\text{TraceCycle}(g) = \texttt{true})} \\ &= 1 - \frac{q}{\text{Prob}(\text{TraceCycle}(g) = \texttt{true})} \\ &\geq 1 - \frac{q}{\text{Prob}(g \in \mathcal{N}_{good} \text{ and } \text{TraceCycle}(g) = \texttt{true})} \\ &= 1 - \frac{q}{\text{Prob}(\text{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{N}_{good}) \cdot \text{Prob}(g \in \mathcal{N}_{good})}. \end{aligned}$$

Set $s = \frac{5}{8}$, $\delta = \frac{1}{24}$ and let $\ell = 1 + \frac{1}{6}$. Note that $\ell = \min\{M(1-s),\ 3-2s-2\delta, 1+s-3\delta, 2s-2\delta\}$, so in particular the inequalities (4) and (5) all hold. We choose N_0 to be the least natural number for which inequality (8) holds. Hence the inequality (2) holds and in particular also $12(rn)^s + 6 \le n$ and $(rn)^s \log(n) \le n$.

Inequality (15) holds by Remark 4.1, so we have $q \le \frac{b_4}{n^\ell}$, where, since $M = 4$, the constant b_4 given by (7), satisfies

$$b_4 \ \le\ \left(\frac{33}{8}\right)^4 + 72\, a_\delta c_\delta^2 r^{5/3} + 6.24\, a_\delta c_\delta^3 r^{3/8} + \frac{c_\delta^2}{r^{7/6}} + \left(\frac{31}{r^{7/24}}\right)^4.$$

By Proposition 7.5 we have that $\mathrm{Prob}(\textsc{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{N}_{good}) \ge \left(\frac{n-2}{n}\right)^4$. Also, by Equation (1), $\mathrm{Prob}(g \in \mathcal{N}_{good}) = \frac{\rho(G,n,m)}{m}$. Hence, using $n \ge N_0$, and the displayed inequality above, we have

$$\begin{aligned}
&\mathrm{Prob}(g \in \mathcal{N} \mid \textsc{TraceCycle}(g) = \texttt{true}) \\
&\ge\ 1 - \frac{b_4}{n^{1+\frac{1}{6}}} \left(\frac{n}{n-2}\right)^4 \frac{m}{\rho(G,n,m)} \\
&\ge\ 1 - \left(\frac{N_0}{N_0-2}\right)^4 \frac{b_4}{\rho(G,n,m)} \cdot \frac{1}{n^{\frac{1}{6}}} = 1 - \frac{c}{n^{\frac{1}{6}}},
\end{aligned}$$

where $c = \left(\frac{N_0}{N_0-2}\right)^4 \frac{b_4}{\rho(G,n,m)}$. □

5. Preliminaries

It is useful to collect together some of the arithmetic facts we use in the rather delicate estimations in the remaining sections.

Lemma 5.1. *Let n, m, r be as in one of the lines of Table 1, and let d be a divisor of rm with $d \le n$. Then either $d = m$, or $d \le 2m/7$, or r, d are as in Table 5. In particular, either $d \le 2m/7$ or d is one of*

r	d		
1	$\frac{m}{3}$	$\frac{m}{2}$	
2	$\frac{m}{3}$	$\frac{2m}{5}$	$\frac{2m}{3}$
3	$\frac{3m}{5}$	$\frac{3m}{7}$	

Table 5. possibilities for r and d

at most 3 different divisors of rm greater than $2m/7$ and in the latter case $d \le 2m/3 \le 2n/3$.

Proof. We have $d = r_0 \frac{m}{j}$, where r_0 divides r and j divides m. If $j = 1$ then $d = m$ since $2m \ge 2(n-6) > n$. So assume $j \ge 2$. Assume also that $d > 2m/7$, or equivalently $7r_0 > 2j$. If m is even, then (see Table 1) $r = 1$. Hence $r_0 = 1$ and $j \le 2$. Thus $d = m/2$ or $m/3$ as in Table 5. So assume now that m is odd, so $j \ge 3$. If $j = 3$ then we have the examples $(r,d) = (1, \frac{m}{3}), (2, \frac{m}{3}), (1, \frac{2m}{3})$ in Table 5 and no others since if $r = 3$ then (see Table 1) $\gcd(m,6) = 1$. Now assume that $j \ge 5$. Then $r_0 > 1$ and we find $(r,d) = (2, \frac{2m}{5}), (3, \frac{3m}{5}), (3, \frac{3m}{7})$ in Table 5 and no others (since $\gcd(m,6) = 1$ when $r = 3$.). □

The next result follows from the fact that $\log(1-p) > -p$ for $0 < p < 1$.

Lemma 5.2. *Let ε, p be real numbers such that $0 < \varepsilon < 1$ and $0 < p < 1$. Set*

$$N(\varepsilon, p) := \left\lceil \frac{\log(\varepsilon^{-1})}{p} \right\rceil.$$

If $m \geq N(\varepsilon, p)$ then $(1-p)^m \leq \varepsilon$.

Lemma 5.3. *Let s be a real number with $\frac{1}{2} < s < 1$ and n, r, t positive integers such that $12(rn)^s + 6 \leq n$. Then*

(i) $m^s/n < n^s/n < (rn)^s/n < 1/12$.

(ii) $n \geq 156$.

(iii) $2(rn)^s - t > \frac{24-t}{12}(rn)^s$.

(iv) if $s = 2/3$ then $n \geq 1746$.

Proof. (i) This follows directly from $12(rm)^s < 12(rn)^s < 12(rn)^s + 6 \leq n$. (ii) As $s > 1/2$ and $r \geq 1$ we have $12\sqrt{n} + 6 \leq 12\sqrt{rn} + 6 < 12(rn)^s + 6 \leq n$. An easy calculation shows that this implies $n \geq 156$. (iii) Note that $n \geq 156$ implies $n^s > n^{1/2} \geq \sqrt{156} > 12$ and so $2(rn)^s - t = (2r^s - \frac{t}{n^s})n^s > (2r^s - \frac{t}{12})n^s = \frac{24r^s - t}{12} n^s \geq \frac{24-t}{12} r^s n^s$. (iv) By calculator. □

The next inequalities are easily verified.

Lemma 5.4. *Let $x \in \mathbb{R}$ with $x > 12$. Then*

(a) $x\left(\frac{1}{2}\right)^x < \frac{1}{4x}$, *and*

(b) $\left(\frac{11}{12}\right)^x < \frac{5}{x}$.

For the estimates in our last arithmetic result Lemma 5.6, we first restate how to estimate sums via integrals.

Lemma 5.5. *Let $a, b \in \mathbb{Z}$ with $a < b$, and let $f(x)$ be a function defined on the interval $[a-1, b+1]$, satisfying one of the lines of Table 6. Then*

$$\sum_{x=a}^{b} f(x) \leq \int_{a-\delta}^{b+\varepsilon} f(t)\mathrm{d}t.$$

conditions on f	δ	ε
increasing in $[a, b+1]$	0	1
decreasing in $[a-1, b]$	1	0
non-negative in $[a-1, b+1]$ and for some $c \in (a, b)$ decreasing in $[a-1, c]$ and increasing in $[c, b+1]$	1	1

Table 6. Conditions on f

Lemma 5.6. *Let $a, c \in \mathbb{R}^+$ and $n \in \mathbb{Z}^+$ with $n > a > c+2 \geq 3$, and let $t, \ell \in \mathbb{Z}^+$ with $t \geq 2$ and $t \geq \ell$. Then, summing over integers x in the interval $(a, n]$,*

$$\sum_{a<x\leq n} \frac{x^t}{(x-c)^\ell} < \sum_{i=0}^{\ell-2} \binom{t}{i} \frac{c^{t-i}(a-1-c)^{i+1-\ell}}{\ell-i-1} + \binom{t}{\ell-1} c^{t+1-\ell} \log(n) + \sum_{i=\ell}^{t} \binom{t}{i} \frac{c^{t-i}(n+1-c)^{i+1-\ell}}{i+1-\ell}.$$

Proof. Note first that if $t > \ell$ the function $f(x) = \frac{x^t}{(x-c)^\ell}$ is decreasing on $(a, \frac{tc}{t-\ell}]$ and increasing on $[\frac{tc}{t-\ell}, n]$, while if $t = \ell$ then $f(x)$ is decreasing on $(a, n]$. In either case, by Lemma 5.5 we have $\sum_{a<x\leq n} f(x) < \int_{a-1}^{n+1} f(x)dx$. Now

$$\begin{aligned}
\int_{a-1}^{n+1} \frac{x^t}{(x-c)^\ell} dx &= \int_{a-1-c}^{n+1-c} \frac{(y+c)^t}{y^\ell} dy = \int_{a-1-c}^{n+1-c} y^{-\ell} \sum_{i=0}^{t} \binom{t}{i} y^i c^{t-i} dy \\
&= \sum_{i=0}^{t} \binom{t}{i} c^{t-i} \int_{a-1-c}^{n+1-c} y^{i-\ell} dy \\
&= \left[\sum_{0\leq i\leq t, i\neq \ell-1} \binom{t}{i} c^{t-i} \frac{y^{i+1-\ell}}{i+1-\ell} + \binom{t}{\ell-1} c^{t+1-\ell} \log y \right]_{y=a-1-c}^{n+1-c} \\
&= \sum_{0\leq i\leq t, i\neq \ell-1} \binom{t}{i} c^{t-i} \frac{(n+1-c)^{i+1-\ell} - (a-1-c)^{i+1-\ell}}{i+1-\ell} \\
&\quad + \binom{t}{\ell-1} c^{t+1-\ell} (\log(n+1-c) - \log(a-1-c)) \\
&< \sum_{i=0}^{\ell-2} \binom{t}{i} c^{t-i} \frac{(a-1-c)^{i+1-\ell}}{\ell-i-1} + \binom{t}{\ell-1} c^{t+1-\ell} \log(n) \\
&\quad + \sum_{i=\ell}^{t} \binom{t}{i} c^{t-i} \frac{(n+1-c)^{i+1-\ell}}{i+1-\ell}.
\end{aligned}$$

□

6. Binomial inequalities and partitions

In this section we prove a result about partitions that will be needed in Sections 11 and 7. As preparation, we prove an inequality about certain binomial coefficients.

Lemma 6.1. *Let a be an integer such that $a > 1$, and let c, ℓ be integers such that $1 \leq \ell < c$. Then*

$$\binom{ca-1}{a-1}\binom{c}{\ell} \leq \binom{ca}{\ell a}.$$

Proof. The proof is by induction on ℓ, for fixed c, a. Since $\binom{c}{\ell} = \binom{c}{c-\ell}$ and $\binom{ca}{\ell a} = \binom{ca}{(c-\ell)a}$, it is sufficient to prove this for $1 \leq \ell \leq \lfloor c/2 \rfloor$. Suppose first that $\ell = 1$. Here it is straightforward to check that

$$\binom{ca-1}{a-1}\binom{c}{1} = \binom{ca}{a}.$$

Now suppose that $1 \leq \ell < \lfloor c/2 \rfloor$ and that the inequality holds for ℓ. Then, using induction we have

$$\binom{ca-1}{a-1}\binom{c}{\ell+1} = \binom{ca-1}{a-1}\binom{c}{\ell}\frac{c-\ell}{\ell+1} \leq \binom{ca}{\ell a}\frac{c-\ell}{\ell+1}.$$

This latter quantity is at most $\binom{ca}{(\ell+1)a}$ if and only if

$$\frac{c-\ell}{\ell+1} \cdot \frac{1}{(\ell a)!(ca-\ell a)!} \leq \frac{1}{(\ell a+a)!(ca-\ell a-a)!} \tag{16}$$

and this is equivalent to

$$\frac{c-\ell}{\ell+1} \leq \frac{ca-\ell a}{\ell a+a} \cdot \frac{ca-\ell a-1}{\ell a+a-1} \cdots \frac{ca-\ell a-a+1}{\ell a+1}.$$

Now the first factor on the right hand side is equal to $(c-\ell)/(\ell+1)$, and each of the other factors is at least 1 since $c \geq 2\ell+1$. Thus the inequality (16) holds, and so the induction proof is complete. □

Lemma 6.2. *(a) For $2 \leq k \leq d < n$ we have*

$$\binom{d}{k} \leq \left(\frac{d}{n}\right)^k \binom{n}{k}$$

and moreover, if $d \leq \alpha n$ for some $\alpha < 1$ then

$$\frac{\binom{d}{k}}{\binom{n}{k}} \leq \alpha^{k-1}\frac{d-k+1}{n-k+1} \leq \alpha^k.$$

(b) For $2 \leq k \leq 2n/3$ we have

$$\binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} < 2\binom{n}{k}\left(\frac{3k}{4n}\right)^{\lceil k/2 \rceil}.$$

Proof. Every part of the proof depends on the following observation:

Fact 1: For $0 \leq i \leq t \leq n$ with $i < n$ we have $\frac{t-i}{n-i} \leq \frac{t}{n}$ with strict inequality if $t < n$.

For (a) observe that

$$\frac{\binom{v}{k}}{\binom{n}{k}} = \prod_{i=0}^{k-1}\frac{v-i}{n-i} \leq \prod_{i=0}^{k-1}\frac{v}{n} = \left(\frac{v}{n}\right)^k.$$

If $d \leq \alpha n$ for some $\alpha < 1$ then

$$\frac{\binom{d}{k}}{\binom{n}{k}} = \left(\prod_{i=0}^{k-2}\frac{d-i}{n-i}\right)\cdot\frac{d-k+1}{n-k+1} \leq \left(\prod_{i=0}^{k-2}\frac{\alpha n-i}{n-i}\right)\cdot\frac{d-k+1}{n-k+1}.$$

Now, again by Fact 1, $(\alpha n-i)/(n-i) \leq \alpha$ and $\frac{d-k+1}{n-k+1} \leq \frac{d}{n} \leq \alpha$, and therefore

$$\frac{\binom{d}{k}}{\binom{n}{k}} \leq \alpha^{k-1}\frac{d-k+1}{n-k-1} \leq \alpha^k.$$

For (b) let $n_0 = \lfloor n/2 \rfloor$ and $k_0 = \lfloor k/2 \rfloor$. Note then that

$$\begin{aligned}\frac{\binom{n_0}{k_0}}{\binom{n}{k}} &= \frac{n_0(n_0-1)\cdots(n_0-k_0+1)}{n(n-1)\cdots(n-k+1)}k(k-1)\cdots(k_0+1)\\ &= \prod_{i=0}^{k_0-1}\frac{n_0-i}{n-i}\cdot\prod_{j=k_0}^{k-1}\frac{k+k_0-j}{n-j}.\end{aligned}$$

Now $k+k_0 \le 2n/3+n/3 \le n$. Applying Fact 1 with $t=n_0$ to the first product and with $t=k+k_0$ to the second, we obtain

$$\begin{aligned}\frac{\binom{n_0}{k_0}}{\binom{n}{k}} &\le \prod_{i=0}^{k_0-1}\frac{n_0}{n}\cdot\prod_{j=k_0}^{k-1}\frac{k+k_0}{n}\\ &= \left(\frac{n_0}{n}\right)^{k_0}\cdot\left(\frac{k+k_0}{n}\right)^{k-k_0}\\ &\le \left(\frac{1}{2}\right)^{k_0}\cdot\left(\frac{3k}{2n}\right)^{k-k_0}\\ &= \frac{3^{\lceil k/2\rceil}}{2^k}\cdot\left(\frac{k}{n}\right)^{\lceil k/2\rceil}\\ &\le \frac{2\cdot 3^{\lceil k/2\rceil}}{4^{\lceil k/2\rceil}}\cdot\left(\frac{k}{n}\right)^{\lceil k/2\rceil}.\end{aligned}$$

Note that the first inequality is strict if either $k_0 \ge 2$ or $k-1 > k_0$, that is, if $k \ge 3$. If $k=2$ then $\binom{n_0}{k_0} = \lfloor n/2 \rfloor$, while $2\binom{n}{k}\left(\frac{3k}{4n}\right)^{\lceil k/2\rceil} = \frac{3}{2}(n-1) > \lfloor n/2 \rfloor$. Thus (b) is proved for all k.

□

Lemma 6.3. *Let d,k,t be positive integers and $a>0$ such that $k \le d$ and $\frac{t}{d-k+1} \le a$. Then*

$$\begin{aligned}&(d+t)(d+t-1)\ldots(d+t-k+1)\\ &< d(d-1)\ldots(d-k+1)(1+\frac{(1+a)^k t}{a(d-k+1)}).\end{aligned}$$

Proof. Note first that

$$\begin{aligned}&(d+t)(d+t-1)\ldots(d+t-k+1)\\ &= d(1+\frac{t}{d})(d-1)(1+\frac{t}{(d-1)})\ldots(d-k+1)(1+\frac{t}{(d-k+1)})\\ &\le d(d-1)\ldots(d-k+1)(1+\frac{t}{(d-k+1)})^k.\end{aligned}$$

Set $x = \frac{t}{d-k+1}$, so $0 < x \le a$. Then

$$\begin{aligned}(1+x)^k &= \sum_{j=0}^{k}\binom{k}{j}x^j = 1+x\sum_{j=1}^{k}\binom{k}{j}x^{j-1}\\ &\le 1+x\sum_{j=1}^{k}\binom{k}{j}a^{j-1}\end{aligned}$$

$$\begin{aligned} &\leq 1 + \frac{x}{a}\sum_{j=0}^{k}\binom{k}{j}a^j \\ &= 1 + \frac{x}{a}(1+a)^k. \end{aligned}$$

□

Now we state and prove the result on partitions.

Proposition 6.4. *Let $\mathcal{U}$ be a finite set of size $u > 1$, and let $\mathcal{P}$ be a partition of $\mathcal{U}$ in which all parts have size at least 2. For $2 \leq k_0 \leq u$, let $N_{\mathcal{P}}(k_0)$ denote the number of k_0-subsets of $\mathcal{U}$ that are unions of parts of $\mathcal{P}$. Then $N_{\mathcal{P}}(k_0) \leq \binom{\lfloor u/2 \rfloor}{\lfloor k_0/2 \rfloor}$, and moreover, if k_0 is odd and u is even, then $u \geq 4$ and $N_{\mathcal{P}}(k_0) \leq \binom{(u-2)/2}{(k_0-1)/2}$. In particular, $N_{\mathcal{P}}(k_0) = 1$ if $k_0 = u$ and $N_{\mathcal{P}}(k_0) \leq \frac{1}{u-1}\binom{u}{k_0}$ otherwise.*

Proof. First we construct a partition $\mathcal{P}'$ of $\mathcal{U}$ having at most two parts of size 1, and all parts of size at most 2. Start with $\mathcal{P}' = \emptyset$ and run through the parts of $\mathcal{P}$. For each part $P \in \mathcal{P}$ of even size, choose any partition of P with all parts of size 2, and add the parts of this partition to $\mathcal{P}'$. If all parts of $\mathcal{P}$ have even size, then the construction of $\mathcal{P}'$ is completed in this way. So suppose that $\mathcal{P}$ has at least one part of odd size. In this case $\mathcal{P}'$ will have 1 or 2 parts of size 1, and its construction is completed as follows. For each part $P \in \mathcal{P}$ of odd size $p := |P|$, add $(p-1)/2$ parts of size 2 to $\mathcal{P}'$ formed from $p-1$ of the points of P. Let $P_1, \ldots, P_r$ be the odd length parts of $\mathcal{P}$. Pair up the remaining r points into parts of size 2 and add them to $\mathcal{P}'$, leaving exactly 1 or 2 of these points to form singleton parts of $\mathcal{P}'$.

Next we define, for each k_0-subset η of $\mathcal{U}$ that is a union of parts of $\mathcal{P}$, a k_0-subset η' that is a union of parts of $\mathcal{P}'$. Note that if k_0 is odd then η must contain a part of $\mathcal{P}$ of odd size, and in this case $\mathcal{P}'$ has one or two singleton parts. If k_0 is odd and $\mathcal{P}'$ has two singleton parts, then we choose one of them, and we always place this chosen singleton part in η'. To define η' for a given η, we start with $\eta' = \emptyset$ and build it up by considering in turn each of the parts P of $\mathcal{P}$ contained in η. If $|P|$ is even, then P is a union of parts of $\mathcal{P}'$ of size 2, and we add all of these parts to η'. If $|P|$ is odd, then we add to η' all the parts of size 2 of $\mathcal{P}'$ contained in P. At this stage $|\eta'| = k_0 - \ell$, where ℓ is the number of odd sized parts of $\mathcal{P}$ contained in η. Next we add to η' up to $\lfloor \ell/2 \rfloor$ parts of $\mathcal{P}'$ of size 2 that contain points from two different parts of $\mathcal{P}$. If η' cannot be completed in this way then either (i) ℓ is odd, or (ii) ℓ is even and is equal to the number of odd sized parts of $\mathcal{P}$. Case (i) occurs if and only if k_0 is odd, and here we add to η' the designated singleton part of $\mathcal{P}'$. In case (ii) there are two singleton parts of $\mathcal{P}'$, and we add to η' these two singleton parts.

Note that, if $\ell \geq 2$, then we may have had some freedom in choosing the $\lfloor \ell/2 \rfloor$ parts of $\mathcal{P}'$ of size 2 that contain points from two different parts of $\mathcal{P}$, so η' may not be determined uniquely by η. On the other hand, η' always determines η uniquely, since η is the union of the parts of $\mathcal{P}$ that have at least two points in η'. Thus distinct sets η correspond to distinct sets η'.

It follows that $N_{\mathcal{P}}(k_0) \leq N'$ where N' is the number of k_0-subsets $\gamma \subseteq \mathcal{U}$ such that γ is a union of parts of $\mathcal{P}'$ and in addition, if k_0 is odd and $\mathcal{P}'$ has two singleton parts, then γ contains a designated one of these singleton parts.

Suppose that γ is such a k_0-subset. If $\mathcal{P}'$ has at most one part of size 1, then γ contains $\lfloor k_0/2 \rfloor$ of the parts of $\mathcal{P}'$ of size 2 (and also a singleton part if k_0 is odd). Thus $N' \leq \binom{\lfloor u/2 \rfloor}{\lfloor k_0/2 \rfloor}$. Note that in this case, if k_0 were odd, then $\mathcal{P}$ would have at least one odd part, and so $\mathcal{P}'$ would have exactly one odd part, whence u would be odd. Thus the first assertion is proved in this case. So suppose that $\mathcal{P}'$ has two singleton parts, in which case u is even. If k_0 is odd then $k_0 \geq 3$ and γ consists of $\lfloor k_0/2 \rfloor$ of the parts of $\mathcal{P}'$ of size 2 and the designated singleton part, whence $u \geq 4$ and $N' \leq \binom{(u-2)/2}{(k_0-1)/2} < \binom{\lfloor u/2 \rfloor}{\lfloor k_0/2 \rfloor}$. On the other hand, if k_0 is even then γ consists of $k_0/2$ of the two-point parts (or $k_0/2 - 1$ parts of size two and the two singleton parts). Again $N' \leq \binom{\lfloor u/2 \rfloor}{\lfloor k_0/2 \rfloor}$. This proves the first assertion in all cases.

Note that $\lfloor u/2 \rfloor = \lfloor k_0/2 \rfloor$ if and only if either $k_0 = u$, or $k_0 = u - 1$ with u odd. If $k_0 = u$ obviously $N_{\mathcal{P}}(k_0) = N' = 1$. If $k_0 = u - 1$ with u odd then $\mathcal{P}'$ has a unique part of size 1 and its complement is

the unique k_0-subset of U that is a union of parts of $\mathcal{P}'$ - it may or may not be a union of parts of $\mathcal{P}$. Thus $N_{\mathcal{P}}(k_0) \leq N' = 1 \leq \frac{1}{u-1}\binom{u}{k_0}$.

So suppose from now on that $\lfloor k_0/2 \rfloor < \lfloor u/2 \rfloor$, and set $u_1 = \lfloor u/2 \rfloor$ and $k_1 = \lfloor k_0/2 \rfloor$. Then $\binom{\lfloor u/2 \rfloor}{\lfloor k_0/2 \rfloor} = \binom{u_1}{k_1}$, and by Lemma 6.1, this is at most $\frac{1}{2u_1-1}\binom{2u_1}{2k_1}$. If k_0 and u are even, then $k_0 < u$ and this quantity is at most $\frac{1}{u-1}\binom{u}{k_0}$. If k_0 is even and u is odd, then $2 \leq k_0$ and $\frac{1}{2u_1-1}\binom{2u_1}{2k_1} = \frac{1}{u-2}\binom{u-1}{k_0}$. This in turn is at most $\frac{1}{u-1}\binom{u}{k_0}$. Now suppose k_0 and u are odd. Then $\lfloor k_0/2 \rfloor < \lfloor u/2 \rfloor$ implies $k_0 \leq u-2$, and $\frac{1}{2u_1-1}\binom{2u_1}{2k_1} = \frac{1}{u-2}\binom{u-1}{k_0-1}$ which is at most $\frac{1}{u-1}\binom{u}{k_0}$. Finally consider k_0 odd and u is even. As shown above $u \geq 4$ and $N' \leq \binom{(u-2)/2}{(k_0-1)/2}$. By Lemma 6.1, this is at most $\frac{1}{u-3}\binom{u-2}{k_0-1}$, which in turn is at most $\frac{1}{u-1}\binom{u}{k_0}$. □

For a prime p and an integer n, let n_p denote the p-part of n, that is the highest power of p dividing n. Recall that, for a positive integer $k_0 \leq n$, a k_0-subset γ' of Ω, and an element $g \in S_n$, we denote by $c_{k_0}(\gamma', g)$ the length of the g-cycle containing γ' in the action of g on k_0-sets.

Lemma 6.5. *Let $g \in S_n$, let C be a g-cycle of length t, let k_0 be a positive integer such that $k_0 \leq t$ and let p be a prime dividing t.*

(a) Suppose that γ' is a k_0-subset of C such that the p-part t_p does not divide $c_{k_0}(\gamma', g)$. Then γ' is a union of $Z(C,p)$-orbits, where $Z(C,p)$ is the subgroup of order p of the cyclic group $\langle g^C \rangle \cong Z_t$ induced by g on C. In particular p divides $\gcd(k_0, t)$.

(b) The number $\sigma(k_0, C)$ of k_0-subsets γ' of C such that t_p does not divide $c_{k_0}(\gamma', g)$ is at most $\binom{\lfloor t/2 \rfloor}{\lfloor k_0/2 \rfloor}$, and in particular, is 1 if $k_0 = t$, and at most $\frac{1}{t-1}\binom{t}{k_0}$ if $k_0 < t$.

Proof. (a) Since t_p does not divide $c_{k_0}(\gamma', g)$ and $\langle g^C \rangle \cong Z_t$, it follows that the setwise stabiliser H of γ' in $\langle g^C \rangle$ contains the unique subgroup $Z(C,p)$ of $\langle g^C \rangle$ of order p. As γ' is H-invariant, γ' is a union of H-orbits in C, and hence γ' is a union of $Z(C,p)$-orbits in C. In particular, p divides k_0 as well as t.

(b) If $k_0 = t$ then C is its unique k_0-subset and $\sigma(k_0, C) = 1$. If $k_0 < t$ then, by Proposition 6.4, $\sigma(k_0, C) \leq \binom{\lfloor t/2 \rfloor}{\lfloor k_0/2 \rfloor}$ and also $\sigma(k_0, C) \leq \frac{1}{t-1}\binom{t}{k_0}$. □

Corollary 6.6. *Let G, n, m, r be as in one of the lines of Table 1, and let $g \in G$. Let $\Sigma(g)$ be as in Table 2 with $u = |\Sigma(g)|$, and let k_0 be a positive integer such that $k_0 \leq u$. Then the number $\sigma(k_0, \Sigma(g))$ of k_0-subsets γ' of $\Sigma(g)$ such that $c_{k_0}(\gamma', g)$ divides rm satisfies*

$$\sigma(k_0, \Sigma(g)) \leq \begin{cases} 0 & \text{if } k_0 = 1 \\ 1 & \text{if } k_0 = u \\ \frac{1}{u-1}\binom{u}{k_0} & \text{if } 1 < k_0 < u. \end{cases}$$

Proof. For each g-cycle C in $\Sigma(g)$, by the definition of $\Sigma(g)$, $|C|$ does not divide rm, and hence there exists a prime $p(C)$ such that $|C|_{p(C)}$ does not divide rm. Let $Z(C, p(C))$ denote the subgroup of order $p(C)$ of the cyclic group $\langle g^C \rangle$ induced by g on C, let $\mathcal{P}(C)$ denote the set of $Z(C, p(C))$-orbits in C (all of length $p(C)$), and let $\mathcal{P} = \cup_C \mathcal{P}(C)$ denote the corresponding partition of $\Sigma(g)$.

Suppose that γ' is a k_0-subset of $\Sigma(g)$, and for each g-cycle C in $\Sigma(g)$, let $k(C) = |\gamma' \cap C|$. Then $c_{k_0}(\gamma', g)$ is the least common multiple of $c_{k(C)}(\gamma' \cap C, g)$, over all g-cycles C such that $k(C) \neq 0$. Note that $c_{k(C)}(\gamma' \cap C, g)$ divides $|C|$.

Suppose now that $c_{k_0}(\gamma', g)$ divides rm. Then for each C such that $k(C) \neq 0$, also $c_{k(C)}(\gamma' \cap C, g)$ divides rm, and hence $|C|_{p(C)}$ does not divide $c_{k(C)}(\gamma' \cap C, g)$. By Lemma 6.5, $\gamma' \cap C$ is a union of parts of $\mathcal{P}(C)$. Thus γ' is a union of parts of $\mathcal{P}$. Since all parts of $\mathcal{P}$ have size at least 2, this implies that $\sigma(k_0, \Sigma(g)) = 0$ if $k_0 = 1$, and the inequality for $1 < k_0 \leq u$ follows from Proposition 6.4. □

7. Tracing k-subsets

For the remainder of this paper we assume that k is an integer with $2 \leq k \leq n/2$. We use $\Delta(g)$, $\Sigma(g)$ and other notation introduced in Tables 2 and 3. Further, we use without further reference the number M of independent uniformly distributed random k-subsets in Algorithm 2 TraceCycle, where M satisfies (2), in particular $M \geq 4$.

Proposition 7.1. *Let G, n, m, r be as in one of the lines of Table 1, and suppose that $g \in G$ does not contain an m-cycle. Set $v = |\Delta(g)|$ and suppose that $v \leq n-k-1$. Then the proportion of k-subsets γ of Ω such that $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, is at most $\frac{v^k}{n^k} + \frac{1}{n-v-1}$.*

Proof. Set $u = n - v = |\Sigma(g)|$. Suppose that γ is a k-subset of Ω such that $c_k(\gamma, g) = r_0 m$ for some r_0 dividing r, and set $k_0 := |\gamma \cap \Sigma(g)|$. Then $k_0 \leq \min\{k, u\}$. By assumption, $v \leq n-k-1$ and so $u = n - v \geq k+1$ and $k_0 \leq \min\{k, u\} = k$. Also $c_{k_0}(\gamma \cap \Sigma(g), g)$ divides $c_k(\gamma, g)$, and hence divides rm. By Corollary 6.6, the number $\sigma(k_0, \Sigma(g))$ of k_0-subsets γ' of $\Sigma(g)$ such that $c_{k_0}(\gamma', g)$ divides rm is 0 if $k_0 = 1$, 1 if $k_0 = u$, and at most $\frac{1}{u-1}\binom{u}{k_0}$, otherwise. If $k_0 = 0$ then γ is one of the $\binom{v}{k}$ k-subsets of $\Delta(g)$. Thus the number of possibilities for γ is at most

$$\begin{aligned} & \binom{v}{k} + \sum_{k_0=2}^{k} \sigma(k_0, \Sigma(g)) \binom{n-u}{k-k_0} \\ \leq\ & \binom{v}{k} + \frac{1}{u-1} \sum_{k_0=2}^{k} \binom{u}{k_0} \binom{n-u}{k-k_0} \\ <\ & \binom{v}{k} + \frac{1}{u-1} \binom{n}{k}. \end{aligned}$$

Now $u - 1 = n - v - 1$, hence the above is $\binom{v}{k} + \frac{1}{n-v-1}\binom{n}{k}$. By Lemma 6.2(a), $\binom{v}{k}$ is at most $(v/n)^k \binom{n}{k}$, which completes the proof. □

Lemma 7.2. *Let G, n, m, r be as in one of the lines of Table 1. Let g be a uniformly distributed random element of G, and suppose that g does not contain an m-cycle, and that $v = |\Delta(g)| \leq n-k-1$. Then the following both hold.*

(a) $\text{Prob}(\text{TraceCycle}(g) = \texttt{true}) \leq 2^M \left(\left(\frac{v^k}{n^k}\right)^M + \left(\frac{1}{n-v-1}\right)^M \right)$,

(b) $\text{Prob}(\text{TraceCycle}(g) = \texttt{true}) \leq 16 \max\left\{ \left(\frac{v}{n}\right)^4, \left(\frac{1}{n-v-1}\right)^4 \right\}$.

Moreover, if $3 \leq v \leq n-3$ then $\text{Prob}(\text{TraceCycle}(g) = \texttt{true}) \leq 16 \left(\frac{v}{n}\right)^4$.

Proof. Now TraceCycle$(g) = \texttt{true}$ if and only if $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, for each of the M independent uniformly distributed random k-sets γ tested during the algorithm. Thus if g does not contain an m-cycle, the probability that TraceCycle$(g) = \texttt{true}$ is p^M, where p is the proportion of k-subsets γ such that $c_k(\gamma, g) = r_0 m$ for some r_0 dividing r. By Proposition 7.1, $p \leq \frac{v^k}{n^k} + \frac{1}{n-v-1}$. Note that $p^M \leq p^4$ since $p \leq 1$ and $M \geq 4$. Set $x = \frac{v^k}{n^k}$ and $y = \frac{1}{n-v-1}$. If $x \leq y$ then $(x+y)^M \leq (2y)^M = 2^M y^M$, and similarly if $x \geq y$ then $(x+y)^M \leq 2^M x^M$. It follows that $p^M \leq 2^M(x^M + y^M)$, proving part (a).

For (b), we observe that

$$p^M \leq p^4 \leq (x+y)^4 \leq (2\max\{x,y\})^4 = 16 \cdot \max\{x,y\}^4.$$

Part (b) follows on noting that $x \leq v/n$ (since $v \leq n$). Finally suppose that $3 \leq v \leq n-3$. Then $n \geq v+3 \geq v+2+\frac{2}{v-1}$ so $n(v-1) \geq v^2 + v$ and hence $(n-v-1)v \geq n$, that is, $\frac{v}{n} \geq \frac{1}{(n-v-1)}$. The last assertion now follows from part (b). □

Now we analyse the effect of TraceCycle applied to elements of $\mathcal{R}$.

Proposition 7.3. *Let G, n, m, r be as in one of the lines of Table 1 and suppose that $12(rn)^s + 6 \leq n$. Then, for a uniformly distributed random element $g \in G$,*

$$\text{Prob}(\text{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{R}) \leq \left(\frac{33}{8n^{1-s}}\right)^M.$$

Proof. By definition, for $g \in \mathcal{R}$, $v = |\Delta(g)| \leq 4(rn)^s$ and g does not contain an m-cycle. By our assumptions on n and k and the hypothesis, we have $n - k - 1 \geq n/2 - 1 > 4(rn)^s \geq v$.

Thus by Proposition 7.1, the proportion of k-subsets γ such that $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, is at most $\frac{v^k}{n^k} + \frac{1}{n-v-1} \leq \frac{(4r^s)^k}{n^{k(1-s)}} + \frac{1}{n-4(rn)^s-1}$.

Now TraceCycle$(g) = \texttt{true}$ if and only if $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, for each of M independent uniformly distributed random k-sets γ tested during the algorithm. Thus, given $g \in \mathcal{R}$, the probability of this occurring is at most

$$\left(\frac{(4r^s)^k}{n^{k(1-s)}} + \frac{1}{n - 4(rn)^s - 1}\right)^M.$$

Now $12(rn)^s < n$, that is to say, $\frac{4r^s}{n^{1-s}} < \frac{1}{3}$. Also $k \geq 2$, $r \leq 3$ and $s < 1$. Therefore $(\frac{4r^s}{n^{1-s}})^k \leq (\frac{4r^s}{n^{1-s}})^2 < \frac{4r^s}{3n^{1-s}} < \frac{4}{n^{1-s}}$. Also, by assumption, $n - 4(rn)^s - 1 \geq 8(rn)^s + 5 > 8r^s n^s > 8r^s n^{1-s}$. Therefore, the probability that TraceCycle$(g) = \texttt{true}$ is at most

$$\left(\frac{4}{n^{1-s}} + \frac{1}{8r^s n^{1-s}}\right)^M \leq \left(\frac{33}{8n^{1-s}}\right)^M.$$

□

Next we analyse the effect of TraceCycle applied to elements of $\mathcal{N}_{good}$ (defined in Table 3).

Lemma 7.4. *Let G, n, m, r be as in one of the lines of Table 1, and let k_0 be an integer satisfying $0 \leq k_0 \leq k$. Let $g \in \mathcal{N}$ and let C be the m-cycle contained in g. Then the number of k_0-subsets of C that can occur as $\gamma \cap C$, for a k-subset γ of Ω such that $c_k(\gamma, g)$ is not divisible by m, is at most*

$$\sigma_{k_0} = \begin{cases} 1 & \text{if } k_0 = 0 \text{ and } k \leq n - m \\ 0 & \text{if } \gcd(m, k_0) = 1 \text{ or if} \\ & k_0 < k - n + m \\ \omega(\gcd(m, k_0))\binom{\lfloor m/2 \rfloor}{\lfloor k_0/2 \rfloor} & \text{if } \gcd(m, k_0) > 1 \text{ and} \\ & k_0 \geq \max\{1, k - n + m\} \end{cases}$$

where $\omega(d)$ is the number of distinct prime divisors of an integer d.

Proof. Let σ' be the number of k_0-subsets of C that can occur as $\gamma \cap C$, for a k-subset γ of Ω such that $c_k(\gamma, g)$ is not divisible by m. Note that, if γ is such a k-subset, then $\gamma \setminus C$ is contained in the complement $\overline{C}$ of C and hence $k = |\gamma| \leq k_0 + |\overline{C}| = k_0 + n - m$. Thus if $k_0 < k - n + m$ then $\sigma' = 0$. Also if $k_0 = 0 \geq k - n + m$, then $\gamma \cap C = \emptyset$ so $\sigma' \leq 1$. Suppose now that $k_0 > 0$ and $k_0 \geq k - n + m$, that is, $k_0 \geq \max\{1, k - n + m\}$.

Let γ be such that $c_k(\gamma, g)$ is not divisible by m. Then $c_{k_0}(\gamma \cap C, g)$ properly divides m, and hence there exists a prime p dividing m such that the p-part m_p does not divide $c_{k_0}(\gamma \cap C, g)$. By Lemma 6.5(a), p divides $\gcd(m, k_0)$ (and in particular if $\gcd(m, k_0) = 1$ then $\sigma' = 0$). If such a prime p exists then, by Lemma 6.5(b), the number of k_0-subsets $\gamma \cap C$ such that m_p does not divide $c_{k_0}(\gamma \cap C, g)$ is at most $\binom{\lfloor m/2 \rfloor}{\lfloor k_0/2 \rfloor}$. Finally there are at most $\omega(\gcd(m, k_0))$ primes p to consider, and the proof is complete. □

Proposition 7.5. *Let G, n, m, r be as in one of the lines of Table 1 and suppose that $g \in \mathcal{N}_{good}$, and $12(rn)^s + 6 \leq n$. Then the proportion of k-subsets γ of Ω such that $c_k(\gamma, g) \neq mr_0$, for any r_0 dividing r, is at most*

$$\sum_{k_0=\max\{k-(n-m),0\}}^{k} \sigma_{k_0} \binom{n-m}{k-k_0} / \binom{n}{k} \leq \sqrt{8k} \left(\frac{3k}{4m}\right)^{\lceil k/2 \rceil},$$

where σ_{k_0} is as in Lemma 7.4. Moreover, for a uniformly distributed random element $g \in G$,

$$\text{Prob}(\textsc{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{N}_{good}) \geq \left(\frac{n-2}{n}\right)^M.$$

Proof. Let C denote the m-cycle in g and let γ be a k-subset of Ω such that $c_k(\gamma, g) \neq mr_0$ for any r_0 dividing r. By the definition of $\mathcal{N}_{good}$, this implies that $c_{k_0}(\gamma \cap C, g)$ is not divisible by m, where $k_0 = |\gamma \cap C|$. Now $0 \leq k_0 \leq \min\{k, m\} = k$, and moreover $k_0 \geq k - (n-m)$ since $\gamma \subseteq (\gamma \cap C) \cup (\Omega \setminus C)$. Given $\gamma \cap C$, there are at most $\binom{n-m}{k-k_0}$ choices for $\gamma \setminus C$. Hence, by Lemma 7.4, the number of such k-subsets γ is at most

$$X := \sum_{k_0=\max\{k-n+m,0\}}^{k} \sigma_{k_0} \binom{n-m}{k-k_0} \tag{17}$$

where $\sigma_0 = 1$, and $\sigma_{k_0} = \omega((\gcd(m, k_0))\binom{\lfloor m/2 \rfloor}{\lfloor k_0/2 \rfloor}$ for $k_0 > 0$. Now $\omega(\gcd(m, k_0)) \leq \omega(k_0) \leq \sqrt{2k_0} \leq \sqrt{2k}$ (see for example, [11, p. 395]). Hence, $X \leq \sqrt{2}k \sum_{k_0=\max\{k-n+m,0\}}^{k} \binom{\lfloor m/2 \rfloor}{\lfloor k_0/2 \rfloor}\binom{n-m}{k-k_0}$ and by Lemma 6.2(b), we have,

$$\begin{aligned} X &\leq \sqrt{2k} \sum_{k_0=\max\{k-n+m,0\}}^{k} 2\binom{m}{k_0} \left(\frac{3k_0}{4m}\right)^{\lceil k_0/2 \rceil} \binom{n-m}{k-k_0} \\ &\leq \sqrt{8k} \left(\frac{3k}{4m}\right)^{\lceil k/2 \rceil} \sum_{k_0=\max\{k-n+m,0\}}^{k} \binom{m}{k_0}\binom{n-m}{k-k_0} \\ &\leq \sqrt{8k} \left(\frac{3k}{4m}\right)^{\lceil k/2 \rceil} \binom{n}{k} \end{aligned}$$

and hence the proportion $X/\binom{n}{k} \leq p$ where $p := \sqrt{8k}\left(\frac{3k}{4m}\right)^{\lceil k/2 \rceil}$.

Now we consider the final assertion. Note that $\textsc{TraceCycle}(g) = \texttt{true}$ if and only if, for each of the M independent uniformly distributed random k-subsets γ tested, we have $c_k(\gamma, g) = r_0 m$ for some r_0 dividing r. The class $\mathcal{N}_{good}$ is, for some lines of Table 1, a union of several conjugacy classes of elements of S_n, say $\mathcal{N}_{good} = \cup_{\mathcal{C}} \mathcal{N}(\mathcal{C})$. For $g \in \mathcal{N}(\mathcal{C})$, the proportion $p(\mathcal{C})$ of k-subsets γ of Ω, such that $c_k(\gamma, g) \neq r_0 m$ for any r_0 dividing r, may depend on the class $\mathcal{C}$, although, as we have shown above, $p(\mathcal{C}) \leq p$ for all $\mathcal{C}$. Thus, given $g \in \mathcal{N}(\mathcal{C})$, the probability that $\textsc{TraceCycle}(g) = \texttt{true}$ is $(1 - p(\mathcal{C}))^M \geq (1-p)^M$. This implies that

$$\text{Prob}(\textsc{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{N}_{good}) \geq (1-p)^M.$$

Thus to complete the proof it is sufficient to prove that $p \leq \frac{2}{n}$ for some upper bound p of $X/\binom{n}{k}$.

Note that, by Lemma 5.3(ii), $m \geq n - 6 \geq 150$. Suppose first that $4 \leq k \leq \frac{n}{2}$. We consider the function $F(x) = \left(\frac{3x}{4m}\right)^{\frac{x}{2}} = e^{\frac{x}{2}\log\frac{3x}{4m}}$ on the interval $[4, \frac{n}{2}]$. Note that $\frac{3x}{4m} \leq \frac{3n}{8m} < 1$ and $\frac{k}{2} \leq \lceil \frac{k}{2} \rceil$, so

$F(k) \geq \left(\frac{3k}{4m}\right)^{\lceil k/2 \rceil}$, and hence $p \leq \sqrt{8k}F(k)$. Differentiating we have $F'(x) = F(x)\frac{1}{2}\left(\log(\frac{3x}{4m}) + 1\right)$, and since $F(x) > 0$ for $x > 0$, it follows that $F(x)$ has a unique minimum at $\log \frac{3x}{4m} = -1$, that is, when $x = \frac{4m}{3e}$ (which may or may not lie in the interval $[4, \frac{n}{2}]$). Thus the maximum of $F(x)$ on the interval $[4, \frac{n}{2}]$ occurs at one of the endpoints. We claim that $\max\{F(4), F(\frac{n}{2})\} < \frac{1}{n^{3/2}}$. It follows from a proof of this claim that $p \leq \sqrt{8k}F(k) \leq \sqrt{8k}\frac{1}{n^{3/2}} \leq \frac{2}{n}$, since $k \leq \frac{n}{2}$.

Since $m \geq n - 6 \geq 150$, we have $m^2 > 9n^{3/2}$, which implies that $F(4) = (\frac{3}{m})^2 < \frac{1}{n^{3/2}}$. Also $\frac{3n}{8m} \leq \frac{3}{8} + \frac{6}{8m} < \frac{1}{2}$, and $n^{3/2} < 2^{n/4}$. Then, applying Lemma 5.4(a), we find

$$F(\frac{n}{2}) \;=\; \left(\frac{3n}{8m}\right)^{n/4} < \left(\frac{1}{2}\right)^{n/4} < \frac{1}{n^{3/2}}$$

proving the claim for $k \geq 4$. For the remaining cases where $k = 2$ or 3, note that $\omega(\gcd(m, k_0)) \leq 1$ for $1 \leq k_0 \leq 3$, $\sigma_{k_0} = 0$ when $k_0 = 1$, $n \geq 156$, and $n - m \leq 6$. If $k = 2$ then by (17),

$$\frac{X}{\binom{n}{2}} \;\leq\; \frac{\binom{6}{2}}{\binom{n}{2}} + \frac{\binom{\lfloor m/2 \rfloor}{1} \cdot \binom{6}{0}}{\binom{n}{2}} \leq \frac{15 \cdot 2}{155} \cdot \frac{1}{n} + \frac{m}{n-1} \cdot \frac{1}{n} < \frac{2}{n}.$$

If $k = 3$ then, again by (17),

$$\begin{aligned} \frac{X}{\binom{n}{3}} \;&\leq\; \frac{\binom{6}{3}}{\binom{n}{3}} + \frac{\binom{\lfloor m/2 \rfloor}{1}\binom{6}{1}}{\binom{n}{3}} + \frac{\binom{\lfloor m/2 \rfloor}{1}\binom{6}{0}}{\binom{n}{3}} \\ &\leq\; \frac{20 \cdot 6}{154 \cdot 155} \cdot \frac{1}{n} + \frac{3 \cdot m(6+1)}{154(n-1)} \cdot \frac{1}{n} < \frac{2}{n}. \end{aligned}$$

□

8. Bounding $\mathcal{S}_0$

Let G, m, n, r be as in one of the lines of Table 1, so G is A_n or S_n. To estimate the probability of a uniformly distributed random element $g \in G$ being in $\mathcal{S}_0$ or $\mathcal{S}_1^+$, and TraceCycle(g) = `true` we use the following result from [8]. Recall the definitions of an s-small and an s-large cycle and of v from Notation 3.2. Let $i \in \{1, 2, 3\}$. In the next two sections we use the following notation:

Notation 8.1. *1. For $v \geq 1$ let $P(v, rm)$ denote the proportion of elements of S_v of order dividing rm, and let $P(0, rm) = 1$.*

2. For $v \geq 1$ let $P_0(v, rm)$ denote the proportion of elements of S_v of order dividing rm, all of whose cycles are s-small, and let $P_0(0, rm) = 1$.

3. Let $P_1^+(v, rm)$ denote the proportion of elements $g \in S_v$ of order dividing rm, and such that g has exactly one s-large cycle of length d, say, where in addition, d satisfies $(rn)^s \leq d < v - 3(rn)^s$.

4. Let D denote the set of all divisors of rm which are at most n.

5. Let $D_1^+(v)$ denote the set of all divisors d of rm satisfying $(rn)^s \leq d < v - 3(rn)^s$.

Note that $r = 1$ or r is a prime. Hence the number $d(rm)$ of positive divisors of rm is at most $2d(m)$, as $d|rm$ if and only if either $d|m$ or $d = rd_0$ and $d_0|m$.

Remark 8.2. *The following result is essentially [8, Lemma 2.4]. Suppose that s, δ and c_δ are as in Notation 3.2. In particular, $s > \delta$. In Lemma 8.3 we may use as a'_δ any constant such that $a'_\delta \geq$*

a'_δ	m_0
25/4	$c_\delta^{1/(s-\delta)}$
a_δ as in (6)	150

Table 7. Possible values of a'_δ for Lemma 8.3

$\frac{5}{4}(1+3\frac{c_\delta}{(rm)^{s-\delta}}+\left(\frac{c_\delta}{(rm)^{s-\delta}}\right)^2)$ *for all sufficiently large values of* rm, *say* $rm \geq m_0$. *These conditions hold in particular for* a'_δ, m_0 *in one of the lines of Table 7. Note that, for the proof of Theorem 3.1, we have* $n \geq 156$ *by Lemma 5.3(ii), so* $rm \geq 150$, *as in line 2 of Table 7.*

Lemma 8.3. *Let* m, n, r *be as in one of the lines of Table 1. Further, let* $v \geq 16$ *and* s, δ, c_δ *and* a_δ *be as in Notation 3.2. Let* a'_δ *and* m_0 *be as in one of the lines of Table 7 (or more generally as in Remark 8.2) and suppose that* $rm \geq m_0$. *Then*

(a) $P_0(v, rm) < \dfrac{a'_\delta d(rm) r^{2s} n^{2s}}{v^3}$.

(b) If $3(rn)^s < v$ *then* $P_0(v, rm) \leq \dfrac{a'_\delta d(rm)^2 r^{2s} n^{2s}}{v(v-(rn)^s)^3}$.

(c) $P_1^+(v, rm) = \displaystyle\sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm)$.

Proof. This result follows from [8, Lemma 2.4] and its proof. A direct application of [8, Lemma 2.4] would require that $rm \geq v$, which we cannot guarantee to hold. However, the proof of that lemma shows, without the assumption that $rm \geq v$, that

$$P_0(v, rm) \leq \frac{d(rm)(rm)^{2s}(1+3c_\delta(rm)^{\delta-s}+(c_\delta(rm)^{\delta-s})^2)}{v(v-1)(v-2)}$$

whenever $v \geq 3$. Statement (a) follows from this, since $m \leq n$, $\delta < s$ and, for $v \geq 16$, $v(v-1)(v-2) > \frac{4}{5}v^3$. To prove (b) we let D_s denote the set of all divisors d of rm such that $d < \min\{v, (rn)^s\}$. By [8, Lemma 2.3(a)] we have that $P_0(v, rm) = \frac{1}{v}\sum_{d \in D_s} P_0(v-d, rm)$, where $P_0(j, m) = 0$ for $j \leq 0$. Since, using Lemma 5.3(ii), $v - d > 3(rn)^s - (rn)^s > 24$ for $d \in D_s$, we have by (a) that $P_0(v-d, rm) \leq \frac{a'_\delta d(rm) r^{2s} n^{2s}}{(v-d)^3}$. Thus $P_0(v, rm) \leq \frac{1}{v}\sum_{d \in D_s} \frac{a'_\delta d(rm) r^{2s} n^{2s}}{(v-d)^3}$. Since $v - d > v - r^s n^s > 0$ for $d \in D_s$, and $|D_s| \leq d(rm)$, we have $P_0(v, rm) \leq \frac{a'_\delta d(rm)^2 r^{2s} n^{2s}}{v(v-(rn)^s)^3}$.

Finally, we prove (c). The number of permutations in S_v of order dividing rm with exactly one s-large cycle of a given length, d say, where d divides rm and $(rn)^s \leq d < v - 3(rn)^s$ is $\binom{v}{d}(d-1)! P_0(v-d, rm)(v-d)!$. Hence the proportion in S_v of such permutations is $\frac{1}{d} P_0(v-d, rm)$. Summing over all $d \in D_1^+(v)$ yields the desired result. □

Proposition 8.4. *Let* G, m, n, r *be as in one of the lines of Table 1. If* $12(rn)^s + 6 \leq n$ *and* $(rn)^s \log(n) \leq n$ *then, for a uniformly distributed random element* $g \in G$,

$$\text{Prob}(g \in \mathcal{S}_0 \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \leq a_\delta d(rm)^2 r^{2s} \frac{72}{n^{3-2s}}$$

where a_δ *is as in (6).*

Proof. The set $\mathcal{S}_0 = \dot\cup \mathcal{S}_0(v)$, where $\mathcal{S}_0(v)$ is the set of all $g \in \mathcal{S}_0$ with $|\Delta(g)| = v$, where v ranges over all integers satisfying $4(rn)^s < v \le n$.

For $g \in \mathcal{S}_0(v)$, the restriction $g^{\Delta(g)}$ of g to $\Delta(g)$ is a permutation in $\mathrm{Sym}(\Delta(g))$ of order dividing rm with all cycles of length less than $(rn)^s$. Consider a fixed v-set Δ. If $G = S_n$, then all elements of $\mathrm{Sym}(\Delta)$ are induced by permutations in G. On the other hand if $G = A_n$, then one of the lines 4-9 of Table 1 holds and hence rm is odd; thus all elements of $\mathrm{Sym}(\Delta)$ of order dividing rm actually lie in $\mathrm{Alt}(\Delta)$ and are therefore induced by elements of G. Therefore in all cases the number of possibilities for the restriction g^Δ of elements $g \in G$, for a given v-subset $\Delta = \Delta(g)$, is $v!P_0(v, rm)$ and the restriction g^Σ where $\Sigma = \Omega\backslash\Delta$ lies in $\mathrm{Sym}(\Sigma)$ or $\mathrm{Alt}(\Sigma)$ according as $G = S_n$ or A_n, respectively. Hence the number of permutations in $\mathcal{S}_0 \cap G$ corresponding to this value of v satisfies

$$|\mathcal{S}_0(v) \cap G| \le \binom{n}{v} v! P_0(v, rm) \frac{(n-v)!}{|S_n : G|} = n! \frac{P_0(v, rm)}{|S_n : G|} = |G| \cdot P_0(v, rm).$$

As $3(rn)^s < 4(rn)^s < v$, we have $n \ge 156$ by Lemma 5.3(ii) so $rm \ge 150$, and hence we can apply Lemma 8.3(b) with $a'_\delta = a_\delta$. Thus, for a random $g \in G$,

$$\mathrm{Prob}(g \in \mathcal{S}_0(v) \cap G) \le P_0(v, rm) \le \frac{a_\delta d(rm)^2 r^{2s} n^{2s}}{v(v-(rn)^s)^3}.$$

For any $g \in S_n$ with $|\Delta(g)| = v$ and $v \le n-k-1$, we have in particular $3 \le v \le n-3$. Hence by Lemma 7.2(b), given that $g \in \mathcal{S}_0(v) \cap G$ with $v \le n-k-1$,

$$\mathrm{Prob}(\textsc{TraceCycle}(g) = \texttt{true}) \le 16 \left(\frac{v}{n}\right)^4.$$

Hence, if $v \le n-k-1$, then the probability that $g \in \mathcal{S}_0(v)$ and $\textsc{TraceCycle}(g) = \texttt{true}$ is at most $a_\delta d(rm)^2 r^{2s} n^{2s} \frac{16}{v(v-(rn)^s)^3} \left(\frac{v}{n}\right)^4$; and if $n-k-1 < v \le n$, this probability is at most $a_\delta d(rm)^2 r^{2s} n^{2s} \frac{1}{v(v-(rn)^s)^3}$. Summing over the values of v, we find

$$\mathrm{Prob}(g \in \mathcal{S}_0 \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \le \Sigma_1 + \Sigma_2$$

where

$$\begin{aligned}
\Sigma_1 &= 16 a_\delta d(rm)^2 \frac{r^{2s} n^{2s}}{n^4} \sum_{4(rn)^s < v \le n-k-1} \frac{v^3}{(v-(rn)^s)^3}, \\
\Sigma_2 &= a_\delta d(rm)^2 r^{2s} n^{2s} \sum_{n-k \le v \le n} \frac{1}{(v-(rn)^s)^4}.
\end{aligned}$$

We first consider Σ_1 and apply Lemma 5.6 with $a = 4(rn)^s$, $c = (rn)^s$, $t = \ell = 3$, and $n-k-1$ in place of n. We also use $a - 1 - c = 3(rn)^s - 1 > 2(rn)^s$, and find

$$\begin{aligned}
\Sigma_1 &= 16 a_\delta d(rm)^2 \frac{r^{2s} n^{2s}}{n^4} \sum_{4(rn)^s < v \le n-k-1} \frac{v^3}{(v-(rn)^s)^3} \\
&< 16 a_\delta d(rm)^2 \frac{r^{2s} n^{2s}}{n^4} \left(\frac{(rn)^{3s}}{8(rn)^{2s}} + \frac{3(rn)^{2s}}{2(rn)^s} \right. \\
&\quad \left. + \binom{3}{2} (rn)^s \log(n-k-1) + \binom{3}{3} \frac{((rn)^s)^0 (n-k-(rn)^s)^1}{1} \right) \\
&< 16 \frac{a_\delta d(rm)^2 r^{2s}}{n^{3-2s}} \left(\frac{(rn)^s}{8n} + \frac{3(rn)^s}{2n} + \frac{3\log(n)(rn)^s}{n} + \frac{n-(rn)^s}{n} \right).
\end{aligned}$$

The assumption $12(rn)^s + 6 \leq n$ implies by Lemma 5.3(i) that $(rn)^s/n < 1/12$. Also, by our hypothesis, $(rn)^s \log(n) \leq n$ and, therefore, $\Sigma_1 < \frac{16a_\delta d(rm)^2 r^{2s}}{n^{3-2s}}(\frac{1}{96} + \frac{3}{24} + 3 + 1) < \frac{66.2a_\delta d(rm)^2 r^{2s}}{n^{3-2s}}$. Finally, we estimate Σ_2.

$$\Sigma_2 = a_\delta d(rm)^2 r^{2s} n^{2s} \sum_{n-k \leq v \leq n} \frac{1}{(v-(rn)^s)^4}.$$

Since $k \leq n/2$, and since $\frac{1}{v-(rn)^s}$ is decreasing for v in the interval $[n-k-2, n]$, we have by Lemma 5.5 and Lemma 5.3 that

$$\begin{aligned}
\Sigma_2 &< a_\delta d(rm)^2 r^{2s} n^{2s} \int_{n/2-1}^{n} \frac{1}{(v-(rn)^s)^4} \mathrm{d}v \\
&= a_\delta d(rm)^2 r^{2s} n^{2s} \left[\frac{-1}{3(v-(rn)^s)^3} \right]_{n/2-1}^{n} \\
&< a_\delta d(rm)^2 r^{2s} n^{2s} \frac{1}{3(n/2-1-(rn)^s)^3} \\
&= a_\delta d(rm)^2 r^{2s} n^{2s} \frac{8}{3n^3(1-2/n-2(rn)^s/n)^3} \\
&< a_\delta d(rm)^2 r^{2s} n^{2s} \frac{8}{3n^3(1-2/156-2/12)^3} \\
&< 4.83\, a_\delta d(rm)^2 r^{2s} \frac{1}{n^{3-2s}}.
\end{aligned}$$

Adding the upper bounds for Σ_1 and Σ_2 we find that

$$\mathrm{Prob}(g \in \mathcal{S}_0 \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) < a_\delta d(rm)^2 r^{2s} \frac{72}{n^{3-2s}}.$$

□

9. Bounding $\mathcal{S}_1^+$

Let G, m, n, r be as in one of the lines of Table 1, so G is A_n or S_n.

Recall the definitions of an s-small and an s-large cycle and of v from Notation 3.2 and the notation set out in Notation 8.1.

Proposition 9.1. *Let G, n, m, r be as in one of the lines of Table 1. If n is such that $12(rn)^s + 6 \leq n$ and $(rn)^s \log(n) \leq n$, then for a uniformly distributed random element $g \in G$,*

$$\mathrm{Prob}(g \in \mathcal{S}_1^+ \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \leq a_\delta d(rm)^3 \frac{6.24}{n^{1+s}}.$$

Proof. The set $\mathcal{S}_1^+ = \dot{\cup}\mathcal{S}_1^+(v)$, where $\mathcal{S}_1^+(v)$ is the set of all $g \in \mathcal{S}_1^+$ with $|\Delta(g)| = v$ and v ranges over integers satisfying $4(rn)^s < v \leq n$. For a given v, an analogous argument to that given in the second paragraph of the proof of Proposition 8.4 shows that

$$\begin{aligned}
|\mathcal{S}_1^+(v) \cap G| &\leq \binom{n}{v} \cdot v! P_1^+(v, rm) \cdot \frac{(n-v)!}{|S_n : G|} \\
&= n! \frac{P_1^+(v, rm)}{|S_n : G|} = P_1^+(v, rm) \cdot |G|.
\end{aligned}$$

Thus applying Lemma 8.3(c) we have, for a random $g \in G$,

$$\text{Prob}(g \in \mathcal{S}_1^+(v) \cap G) \;\le\; P_1^+(v, rm) = \sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm).$$

If $|\Delta(g)| = v$ and $v \le n-k-1$, then in particular $3 \le v \le n-3$. Hence by Lemma 7.2(b), given that $g \in \mathcal{S}_1^+(v) \cap G$ with $|\Delta(g)| = v$ with $v \le n-k-1$,

$$\text{Prob}(\textsc{TraceCycle}(g) = \texttt{true}) \le 16 \left(\frac{v}{n}\right)^4.$$

Thus, if $v \le n-k-1$, the probability that $g \in \mathcal{S}_1^+(v) \cap G$ and $\textsc{TraceCycle}(g) = \texttt{true}$ is at most

$$\left(\sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm) \right) 16 \left(\frac{v}{n}\right)^4,$$

and if $n-k \le v$ this probability is at most

$$\sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm).$$

Summing over v we find

$$\text{Prob}(g \in \mathcal{S}_1^+ \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \le \Sigma_1 + \Sigma_2$$

where

$$\begin{aligned} \Sigma_1 &= 16 \sum_{4(rn)^s < v \le n-k-1} \left(\sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm) \right) \left(\frac{v}{n}\right)^4, \\ \Sigma_2 &= \sum_{n-k \le v \le n} \left(\sum_{d \in D_1^+(v)} \frac{1}{d} P_0(v-d, rm) \right). \end{aligned}$$

First we consider Σ_1. Interchanging the two summations and taking the sum up to n, we obtain the following upper bound, where D_ℓ denotes the set of all divisors d of rm satisfying $d \ge (rn)^s$. Note that $v > d + 3(rn)^s$ (see Notation 3.2).

$$\Sigma_1 < 16 \sum_{d \in D_\ell} \frac{1}{d} \left(\sum_{3(rn)^s + d < v \le n} P_0(v-d, rm) \cdot \frac{v^4}{n^4} \right).$$

Since $rm \ge 150$ by Lemma 5.3(ii), we may apply Lemma 8.3(b) with $a'_\delta = a_\delta$, and find that this expression is at most

$$\begin{aligned} & 16 \sum_{d \in D_\ell} \frac{1}{d} \left(\sum_{3(rn)^s + d < v \le n} \frac{a_\delta d (rm)^2 r^{2s} n^{2s}}{(v-d)(v-d-(rn)^s)^3} \cdot \frac{v^4}{n^4} \right) \\ < \;& 16 \frac{a_\delta d (rm)^2 r^{2s} n^{2s}}{n^4} \sum_{d \in D_\ell} \frac{1}{d} \left(\sum_{3(rn)^s + d < v \le n} \frac{v^4}{(v-d-(rn)^s)^4} \right). \end{aligned}$$

Now we apply Lemma 5.6 with $t = \ell = 4$, $a = 3(rn)^s + d$ and $c = d + (rn)^s$. Noting that $a - c - 1 = 2(rn)^s - 1$, we obtain that this expression is at most

$$16\frac{a_\delta d(rm)^2 r^{2s} n^{2s}}{n^4} \sum_{d\in D_\ell} \frac{1}{d}\left(\frac{\binom{4}{0}(d+(rn)^s)^4}{3(2(rn)^s-1)^3} + \frac{\binom{4}{1}(d+(rn)^s)^3}{2(2(rn)^s-1)^2} + \frac{\binom{4}{2}(d+(rn)^s)^2}{(2(rn)^s-1)} + \binom{4}{3}(d+(rn)^s)^1\log(n) + \binom{4}{4}(d+(rn)^s)^0(n+1-d-(rn)^s)^1\right).$$

Note that $2(rn)^s - 1 > \frac{23}{12} r^s n^s$ by Lemma 5.3(iii) and, since $d \geq (rn)^s$, also $\frac{d+(rn)^s}{d} \leq 2$. Note also that $d + (rn)^s < n$ and $n + 1 - d - (rn)^s < n$. Hence

$$\begin{aligned}\Sigma_1 &\leq 16\frac{a_\delta d(rm)^3 r^{2s} n^{2s}}{n^4}\left(\frac{2\cdot 12^3\cdot n^3}{3\cdot 23^3\cdot r^{3s}n^{3s}} + \frac{4\cdot 12^2 n^2}{23^2 r^{2s} n^{2s}} + \frac{12\cdot 12 n^1}{23 r^s n^s}\right.\\ &+ \left. 8\cdot\log(n) + \frac{n^1}{(rn)^s}\right)\\ &= \frac{16 a_\delta d(rm)^3}{n^{1+s}}\left(\frac{3456}{36501 r^s} + \frac{576}{529 n^{1-s}} + \frac{144 r^s}{23 n^{2-2s}}\right.\\ &+ \left.\frac{8 r^{2s}\log(n)}{n^{3-3s}} + \frac{r^s}{n^{2-2s}}\right).\end{aligned}$$

Since, by hypothesis $(rn)^s \log(n) \leq n$ and by Lemma 5.3(i) $n^s/n \leq r^s n^s/n \leq 1/12$ and $r \geq 1$, the last expression is at most

$$\begin{aligned}\frac{16 a_\delta d(rm)^3}{n^{1+s}} &\left(\frac{3456}{36501} + \frac{576}{529\cdot 12} + \frac{144}{23\cdot 12^2} + \frac{8}{12^2} + \frac{1}{12^2}\right)\\ &\leq 4.7 a_\delta \frac{d(rm)^3}{n^{1+s}}.\end{aligned}$$

We now consider $\Sigma_2 = \sum_{n-k\leq v\leq n}\left(\sum_{d\in D_1^+(v)} \frac{1}{d} P_0(v-d, rm)\right)$. As $v - d > 3(rn)^s$ and $n - k \geq n/2$ we have by Lemma 8.3(b) (with $a'_\delta = a_\delta$) that

$$\begin{aligned}\frac{\Sigma_2}{a_\delta d(rm)^2 (rn)^{2s}} &\leq \sum_{n/2\leq v\leq n}\left(\sum_{d\in D_1^+(v)} \frac{1}{d(v-d-(rn)^s)^4}\right)\\ &= \sum_{d\in D_1^+(v)} \frac{1}{d}\left(\sum_{v(d)\leq v\leq n} \frac{1}{(v-d-(rn)^s)^4}\right)\end{aligned}$$

where $v(d) = \max\{\frac{n}{2}, d + 3(rn)^s\}$ since, by Notation 8.1, each $d \in D_1^+(v)$ is less than $v - 3(rn)^s$. By

Lemma 5.5, this quantity is at most

$$\begin{aligned}
&\sum_{d\in D_1^+(v)} \frac{1}{d}\left(\int_{v(d)-1}^{n} \frac{1}{(v-d-(rn)^s)^4}\,\mathrm{d}v\right)\\
&= \sum_{d\in D_1^+(v)} \frac{1}{d}\left[-\frac{1}{3}\frac{1}{(v-d-(rn)^s)^3}\right]_{v(d)-1}^{n}\\
&< \sum_{d\in D_1^+(v)} \frac{1}{3d}\,\frac{1}{(v(d)-1-d-(rn)^s)^3}.
\end{aligned}$$

In particular each $d \in D_1^+(v)$ is less than m. By Lemma 5.1, there are at most three divisors of rm which are less than m and greater than $2m/7$, and the sum of the reciprocals $\frac{1}{d}$ of these divisors is at most $\frac{7}{m}$, which is less than $\frac{7.3}{n}$ since $n \geq 156$ (by Lemma 5.3(ii)). Using $v(d) \geq d + 3(rn)^s$ and Lemma 5.3(iii), the contribution from these exceptional divisors is therefore at most

$$\frac{1}{(2(rn)^s-1)^3} \sum_{d\in D_1^+(v), d>2m/7} \frac{1}{3d} < \left(\frac{12}{23(rn)^s}\right)^3 \frac{7.3}{3n} < \frac{0.35}{(rn)^{3s}n}.$$

Finally we estimate the contribution of the remaining elements d of $D_1^+(v)$. We note that each such d is at most $\frac{2n}{7}$ and at least $(rn)^s$, and that $(rn)^s < \frac{n-6}{12}$ by our hypothesis. Thus, using $v(d) \geq \frac{n}{2}$, the remaining contribution is at most

$$\frac{d(rm)}{3(rn)^s}\,\frac{1}{(\frac{n}{2}-1-\frac{2n}{7}-\frac{n-6}{12})^3}.$$

Observe that $\frac{n}{2}-1-\frac{2n}{7}-\frac{n-6}{12} = \frac{11n-42}{84}$ and since $n > 84$ by Lemma 5.3(a) we have $\frac{11n-42}{84} > \frac{n}{8}$. Hence, using also that $\frac{(rn)^s}{n} < \frac{1}{12}$ (by Lemma 5.3(i)), the above expression is less than

$$\frac{d(rm)}{(rn)^s}\,\frac{8^3}{3n^3} < \frac{d(rm)\,8^3}{12^2\cdot 3\,(rn)^{3s}n} < \frac{1.19\,d(rm)}{(rn)^{3s}n}.$$

Thus

$$\frac{\Sigma_2}{a_\delta d(rm)^2(rn)^{2s}} < \frac{0.35}{(rn)^{3s}n} + \frac{1.19\,d(rm)}{(rn)^{3s}n} \leq \frac{1.54\,d(rm)}{(rn)^{2s}n^{1+s}}$$

and hence

$$\mathrm{Prob}(g \in \mathcal{S}_1^+ \cap G \text{ and } \textsc{TraceCycle}(g) = \texttt{true}) \;<\; 6.24\,a_\delta \frac{d(rm)^3}{n^{1+s}}.$$

□

10. Bounding $\mathcal{S}_{\geq 2}$

Proposition 10.1. *Let G, m, n, r be as in one of the lines of Table 1. Then*

$$\frac{|\mathcal{S}_{\geq 2} \cap G|}{|G|} \leq \frac{d(rm)^2}{(rn)^{2s}}.$$

Proof. If g is an element of $\mathcal{S}_{\geq 2} \cap G$ then it has two cycles of lengths d_1, d_2, where $d_i | rm$, and $d_i \geq (rn)^s$. There are at most $d(rm)$ choices for each d_i. Thus, there are at most $d(rm)^2$ choices for the two divisors d_1 and d_2. For a given d_1, d_2, the proportion of elements in G having cycles of lengths d_1 and d_2 is at most

$$(d_1 d_2)^{-1} \leq (rn)^{-2s}.$$

Thus altogether we get a proportion of at most $d(rm)^2 (rn)^{-2s}$. □

11. Bounding $\mathcal{S}_1^-$

Proposition 11.1. *Let G, m, n, r be as in one of the lines of Table 1. Suppose that n is such that $12(rn)^s + 6 \leq n$. Let k be a fixed integer with $2 \leq k \leq n/2$. Then*

(a) the proportion of k-subsets γ such that $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, for $g \in \mathcal{S}_1^- \cap G$, is less than $31/((rn)^{1-s})$.

(b) If TraceCycle *is Algorithm 2 and M is as defined there, then for a uniformly distributed random element $g \in G$,*

$$\text{Prob}(\text{TraceCycle}(g) = \texttt{true} \mid g \in \mathcal{S}_1^- \cap G) < \left(\frac{31}{(rn)^{1-s}} \right)^M,$$

and so

$$\text{Prob}(g \in \mathcal{S}_1^- \cap G \text{ and } \text{TraceCycle}(g) = \texttt{true}) < \left(\frac{31}{(rn)^{1-s}} \right)^M.$$

Proof. We start by recording some important facts used throughout the proof. Let $g \in \mathcal{S}_1^- \cap G$ and put $v = |\Delta(g)|$ and $u = |\Sigma(g)|$, such that $u + v = n$. The definition of $\mathcal{S}_1^-$ implies that g has a unique s-large cycle C in $\Delta(g)$ of length d and we have

(i) $d \leq n$ and $d \neq m$ since $g \in \mathcal{F}$;

(ii) $v > 4(rn)^s$ and $v - d \leq 3(rn)^s$.

By Lemma 5.1 and the hypothesis $n \geq 12(rn)^s + 6$, it follows that $d \leq 2m/3 \leq 2n/3$. Hence $u = n - v \geq n - d - 3(rn)^s \geq \frac{n}{3} - 3(rn)^s \geq 4(rn)^s + 2 - 3(rn)^s = (rn)^s + 2$. Also, $v \leq d + 3(rn)^s \leq \frac{2n}{3} + 3(rn)^s$. This implies that $v = n - u \leq n - 2 - (rn)^s$ and hence in particular

$$v \leq n - 3 \tag{18}$$

and

$$\frac{1}{u-1} < \frac{1}{(rn)^s} < \frac{1}{(rn)^{1-s}}. \tag{19}$$

Set $t = v - d$ so that $t = v - d \leq 3(rn)^s$. Then

$$v = d + t \leq 2n/3 + 3(rn)^s. \tag{20}$$

Suppose that γ is a k-subset for which $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, and set $k_0 := |\gamma \cap \Sigma(g)|$. Then $c_{k_0}(\gamma \cap \Sigma(g), g)$ divides rm, and hence the number of possibilities for the k_0-subset $\gamma \cap \Sigma(g)$ is at

most the number $\sigma(k_0, \Sigma(g))$ of Corollary 6.6. In particular $\sigma(k_0, \Sigma(g)) = 0$ if $k_0 = 1$. Thus $k_0 = 0$ or $2 \le k_0 \le \min\{u, k\}$, and the case $k_0 = 0$ is only possible if $v \ge k$.

First we prove the following upper bound for the number $K_{\neg 0} = K_{\neg 0}(g)$ of k-subsets γ such that $k_0 = |\gamma \cap \Sigma(g)| \ge 2$.

$$\frac{K_{\neg 0}}{\binom{n}{k}} < \frac{97}{96(rn)^{1-s}}. \tag{21}$$

By the remarks above

$$K_{\neg 0} \le \sum_{k_0=2}^{\min\{k,u\}} \sigma(k_0, \Sigma(g)) \binom{n-u}{k-k_0}.$$

If $k_0 \le u-1$ then, by Corollary 6.6 and our considerations above, $\sigma(k_0, \Sigma(g)) \le \frac{1}{u-1}\binom{u}{k_0} \le \frac{1}{(rn)^{(1-s)}}\binom{u}{k_0}$, while if $k_0 = u$ then $\sigma(k_0, \Sigma(g)) = 1$. Thus

$$\frac{K_{\neg 0}}{\binom{n}{k}} \le \frac{1}{\binom{n}{k}(rn)^{1-s}} \sum_{k_0=2}^{\min\{u-1,k\}} \binom{u}{k_0}\binom{n-u}{k-k_0} + R_{\neg 0},$$

where

$$R_{\neg 0} = \begin{cases} 0 & \text{if } k \le u-1, \\ \frac{\binom{n-u}{k-u}}{\binom{n}{k}} & \text{if } k \ge u. \end{cases}$$

Hence

$$\begin{aligned} \frac{K_{\neg 0}}{\binom{n}{k}} &\le \frac{1}{\binom{n}{k}(rn)^{1-s}} \sum_{k_0=0}^{\min\{u,k\}} \binom{u}{k_0}\binom{n-u}{k-k_0} + R_{\neg 0} \\ &= \frac{1}{(rn)^{1-s}} + R_{\neg 0}. \end{aligned} \tag{22}$$

Thus (21) is proved if $k \le u-1$, so suppose that $k \ge u$. Recall that $u > (rn)^{1-s} + 1$ by (19). Hence

$$R_{\neg 0} \le \frac{\binom{n-u}{k-u}}{\binom{n}{k}} = \prod_{i=1}^{u} \frac{(k-u+i)}{(n-u+i)} \le \left(\frac{k}{n}\right)^u \le \left(\frac{1}{2}\right)^u \le \frac{1}{2}\left(\frac{1}{2}\right)^{(rn)^{(1-s)}}.$$

Using $n^{1-s} > 12 > 8$ (see Lemma 5.3(i)) and Lemma 5.4(a), we have $R_{\neg 0} \le \frac{1}{2}\left(\frac{1}{2}\right)^{(rn)^{(1-s)}} < \frac{1}{2}\frac{1}{4(rn)^{2(1-s)}} < \frac{1}{96}\frac{1}{(rn)^{1-s}}$, and now the inequality (21) follows from inequality (22).

To complete the proof of part (a) it remains to estimate the number $K_{=0} = K_{=0}(g)$ of k-subsets $\gamma \subseteq \Delta(g)$ such that $c_k(\gamma, g) = r_0 m$ for some r_0 dividing r. Since this number is zero if $v < k$, we assume that $v \ge k$. Recall that C is the unique s-large cycle of g contained in $\Delta(g)$ and $d = |C|$. By Lemma 5.1, $d \le 2m/3 < 2n/3$. Since m divides $c_k(\gamma, g)$ it follows that $\gamma \not\subseteq C$. We prove

$$\frac{K_{=0}}{\binom{n}{k}} \le \frac{30.6}{(rn)^{1-s}}. \tag{23}$$

The number $K_{=0}$ of such k-subsets is at most $\binom{v}{k} - \binom{d}{k}$.

Set $t = v - d$ so that $t = v - d \leq 3(rn)^s$. Then we have

$$\binom{v}{k} = \frac{1}{k!}(d+t)(d+t-1)\dots(d+t-k+1).$$

We consider separately the cases (i) $(rn)^s < k$, (ii) $k \leq \min\{(rn)^s, d-t+1\}$, and (iii) $d-t+1 < k \leq (rn)^s$. Recall that $(rn)^s \leq d$.

Consider first Case (ii), so $k \leq (rn)^s$ and $d-t+1 \geq k$. If $d \leq m/2$ define $a = 1$ and observe that $\frac{t}{d-k+1} \leq a$. If $d > m/2$ then, by Lemma 5.1, it follows that $d \geq 3m/5$. In this case $\frac{t}{d-k+1} \leq \frac{3(rn)^s}{3m/5-(rn)^s} = \frac{3(rn)^s}{m(3/5-(rn)^s/m)}$. By the hypothesis $(rn)^s \leq (n-6)/12 \leq m/12$ and by Lemma 5.3(i) we have then $\frac{t}{d-k+1} \leq \frac{3}{12}\frac{1}{(3/5-1/12)} = \frac{15}{31}$. In this case define $a = \frac{15}{31}$. Then again $\frac{t}{d-k+1} \leq a$. Setting $d_{(k)} := d(d-1)\dots(d-k+1)$, by Lemma 6.3 we obtain

$$\begin{aligned}\binom{v}{k} &= \frac{1}{k!}(d+t)(d+t-1)\dots(d+t-k+1)\\ &< \frac{1}{k!}\left(d_{(k)}\left(1+\frac{(1+a)^k t}{a(d-k+1)}\right)\right)\\ &= \binom{d}{k}\left(1+\frac{(1+a)^k t}{a(d-k+1)}\right). \qquad (24)\end{aligned}$$

If $d \leq m/2$ we have $a = 1$ and so

$$\binom{v}{k} \leq \binom{d}{k} + \binom{d}{k}\frac{2^k t}{d-k+1}.$$

Applying Lemma 6.2(a) with $\alpha = \frac{1}{2}$,

$$\binom{d}{k} \leq \frac{1}{2^{k-1}}\binom{n}{k}\frac{d-k+1}{n-k+1}.$$

Hence

$$K_{=0} \leq \binom{v}{k} - \binom{d}{k} < \binom{d}{k}\frac{2^k t}{d-k+1} \leq \binom{n}{k}\frac{2t}{n-k+1} < \binom{n}{k}\frac{2t}{n-k}.$$

On the other hand, if $m/2 < d \leq 2n/3$, then $a = \frac{15}{31}$, and (24) becomes

$$\binom{v}{k} \leq \binom{d}{k} + \binom{d}{k}\left(\frac{46}{31}\right)^k \frac{31t}{15(d-k+1)}.$$

By Lemma 6.2(a) with $\alpha = \frac{2}{3}$,

$$\binom{d}{k} \leq \frac{2^{k-1}}{3^{k-1}}\binom{n}{k}\frac{d-k+1}{n-k+1}$$

and hence

$$\begin{aligned}K_{=0} &= \binom{v}{k} - \binom{d}{k} < \binom{d}{k}\frac{(46/31)^k 31t}{15(d-k+1)}\\ &< \binom{n}{k}\frac{2^{k-1}}{3^{k-1}}\frac{(46/31)^k 31t}{15(n-k+1)}\\ &< \binom{n}{k}\frac{92^k}{93^k}\frac{31t}{10(n-k)} < \binom{n}{k}\frac{31t}{10(n-k)}.\end{aligned}$$

Note that by Lemma 5.3, since $k \le (rn)^s$ and by our assumptions, $\frac{t}{n-k} \le \frac{3(rn)^s}{n(1-k/n)} \le \frac{3(rn)^s}{n(1-(rn)^s/n)} \le \frac{3(rn)^s}{n(11/12)} = \frac{36(rn)^s}{11n}$.

Thus for all d we have

$$\frac{K_{=0}}{\binom{n}{k}} < \frac{31 \cdot 36}{10 \cdot 11} \cdot \frac{(rn)^s}{n} < \frac{10.2\, r}{(rn)^{1-s}} \le \frac{30.6}{(rn)^{1-s}}$$

and (23) is proved for Case (ii).

Now consider Cases (i) and (iii). Recall from (20) that $v = d + t \le 2n/3 + 3(rn)^s$. By Lemma 5.3(i), $(rn)^s \le \frac{1}{12}n$. Therefore $v \le 2n/3 + \frac{3}{12}n = \frac{11}{12}n$. This shows, using Lemma 6.2(a), that

$$\begin{aligned} \frac{K_{=0}}{\binom{n}{k}} &\le \frac{\binom{v}{k} - \binom{d}{k}}{\binom{n}{k}} < \frac{\binom{v}{k}}{\binom{n}{k}} \\ &\le \left(\frac{v}{n}\right)^k \le \left(\frac{11}{12}\right)^k. \end{aligned}$$

In Case (i) we have $k > (rn)^s$ and hence, observing that $(rn)^s > n^{1/2} > 12$ by Lemma 5.3(i), and using Lemma 5.4(b), we have $\left(\frac{11}{12}\right)^k < \left(\frac{11}{12}\right)^{(rn)^s} < \frac{5}{(rn)^s} < \frac{5}{(rn)^{1-s}}$. Thus $\frac{K_{=0}}{\binom{n}{k}} < \frac{5}{(rn)^{1-s}}$ and (23) holds for Case (i).

In Case (iii) we have $(rn)^s \ge k > d - t + 1$ and so $d < 4(rn)^s$ as $t \le 3(rn)^s$. Therefore, $v = d + t < 7(rn)^s$, and using Lemmas 5.3(i) and 6.2(a),

$$\begin{aligned} \frac{K_{=0}}{\binom{n}{k}} &\le \left(\frac{v}{n}\right)^k < \left(\frac{7(rn)^s}{n}\right)^k \\ &\le \left(\frac{7(rn)^s}{n}\right)^2 < \frac{49(rn)^s}{12n} \le \frac{49}{4(rn)^{1-s}}. \end{aligned}$$

Thus (23) holds for Case (iii) and hence in all cases.

Combining (23) with (21), we conclude that the proportion of k-subsets γ such that $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, is less than $31/((rn)^{1-s})$ for all values of k and v. This proves (a).

Now TraceCycle(g) = `true` if and only if $c_k(\gamma, g) = r_0 m$, for some r_0 dividing r, for each of the M independent uniformly distributed random k-sets γ tested in the algorithm. Thus, given $g \in \mathcal{S}_1^- \cap G$, the probability that TraceCycle(g) = `true` is at most $\left(31/((rn)^{1-s})\right)^M$.

The last assertion follows on noting that for events A and B we have $\mathrm{Prob}(A \cap B) = \mathrm{Prob}(A)\mathrm{Prob}(B \mid A) \le \mathrm{Prob}(B \mid A)$. □

Acknowledgment: The first author acknowledges the support of EPSRC grant EP/C523229 and the second and third authors acknowledge the support of ARC Discovery grants DP0879134 and DP140100416. We thank Yohei Negi and Sven Reichard for some discussions on an early draft of this paper. We thank an anonymous referee for valuable suggestions.

References

[1] R. M. Beals, C. R. Leedham-Green, A. C. Niemeyer, C. E. Praeger, and Á. Seress, *A black-box algorithm for recognizing finite symmetric and alternating groups, I*, Trans. Amer. Math. Soc., 355, 2097-2113, 2003.

[2] S. Bratus and I. Pak, *Fast constructive recognition of a black box group isomorphic to S_n or A_n using Goldbach's conjecture*, J. Symbolic Comput., 29(1), 33-57, 2000.

[3] The GAP Group, *GAP - Groups, Algorithms, and Programming, Version 4.7.7*, 2015, http://www.gap-system.org.

[4] P. Erdős, and P. Turán, *On some problems of a statistical group-theory. I*, Wahrscheinlichkeitstheorie Verw. Gebiete, 4, 175-186, 1965.

[5] P. Erdős, and P. Turán, *On some problems of a statistical group-theory. III*, Acta Math. Acad. Sci. Hungar., 18 , 309-320, 1967.

[6] S. Linton, A. C. Niemeyer and C. E. Praeger, *Constructive recognition of S_n in its action on k-sets*, in preparation.

[7] Y. Negi, *Recognising large base actions of finite alternating groups*, Honours Thesis, School of Mathematics and Statistifcs, The University of Western Australia, 2006.

[8] A. C. Niemeyer and C. E. Praeger, *On permutations of order dividing a given integer*, J. Algebraic Combinatorics, 26, 125-142, 2007.

[9] A. C. Niemeyer and C. E. Praeger, *On the proportion of permutations of order a multiple of the degree*, J. London Math. Soc., 76, 622-632, 2007.

[10] A. C. Niemeyer, C. E. Praeger and Á. Seress, *Estimation problems and randomised group algorithms*, Probabilistic group theory, combinatorics, and computing, Lecture Notes in Math., 2070, Springer, London, 35-82, 2013.

[11] I. Niven, H. S. Zuckerman and H. L. Montgomery, *An Introduction to the Theory of Numbers*, John Wiley & Sons Inc., New York, fifth edition, 1991.

[12] Á. Seress, *Permutation group algorithms*, Cambridge Tracts in Mathematics, 152, Cambridge University Press, Cambridge, 2003.

[13] R. Warlimont, *Über die Anzahl der Lösungen von $x^n = 1$ in der symmetrischen Gruppe S_n*, Arch. Math., 30, 591-594, 1978.