

# THE DIAMETER OF RANDOM CAYLEY DIGRAPHS OF GIVEN DEGREE

MANUEL E. LLADSER, PRIMOŽ POTOČNIK, JANA ŠIAGIOVÁ, JOZEF ŠIRÁŇ, AND MARK C. WILSON

ABSTRACT. We consider random Cayley digraphs of order  $n$  with uniformly distributed generating sets of size  $k$ . Specifically, we are interested in the asymptotics of the probability that such a Cayley digraph has diameter two as  $n \rightarrow \infty$  and  $k = f(n)$ , focusing on the functions  $f(n) = \lfloor cn \rfloor$  and  $f(n) = \lfloor n^\delta \rfloor$ . In both instances we show that the probability converges to 1 as  $n \rightarrow \infty$ , for any fixed  $c \in (0, 1/2)$  and any fixed  $\delta \in (1/2, 1)$ , respectively. We obtain sharper results for Abelian groups. The proofs use detailed asymptotic analysis in several regimes of the combinatorial function  $a(n, k, t)$ , equal to the number of ways of choosing a subset of size  $k$  from a set of size  $n$  while not choosing any of  $t$  preassigned disjoint pairs.

## 1. INTRODUCTION

It is well known that almost all graphs and digraphs have diameter two [1]. This result has been generalized and strengthened in various directions, of which we shall be interested in restrictions to Cayley graphs and digraphs.

In [5] it was proved that almost all Cayley digraphs have diameter two, and in [4] this was extended to Cayley graphs. The random model used in [5, 4] is the most straightforward one: in terms of Cayley digraphs for a given group  $G$ , one chooses a random generating set by choosing its elements among the non-identity elements of  $G$  independently and uniformly, each with probability  $2^{-n+1}$  where  $n$  is the order of  $G$ . Observe that such generating sets have size at least  $n/2$  with probability at least  $1/2$ , in which case a simple counting argument shows that the corresponding Cayley digraphs have diameter at most two. The less trivial part of [5] therefore concerns random Cayley digraphs in which the number of generators is at most half of the order of the group.

This motivates a study of random Cayley digraphs in which the number of generators is restricted. In this case one cannot use the model of [5]. Instead, we let every generating set of the Cayley digraph of a fixed degree appear with equal probability. The fundamental question here is: For which functions  $f$  is it true that the diameter of a random Cayley digraph of an arbitrary group of order  $n$  and of degree  $f(n)$  is asymptotically almost surely equal to 2 as  $n$  tends to infinity? By the well known Moore bound for graphs or digraphs of diameter two we know that  $f$  has to increase at least as fast as  $\sqrt{n}$ . A study of the behaviour of the problem for functions of the form  $f(n) = \lfloor n^\delta \rfloor$  for the powers  $\delta$  satisfying  $1/2 \leq \delta < 1$  is therefore natural in this context. However, even the case when  $f(n) = \lfloor cn \rfloor$  for a constant  $c$  seems not to have been investigated before and, as we shall see, leads to an interesting asymptotic analysis.

The probability that a random Cayley digraph of (in- and out-) degree  $k$  on a group of order  $n$  has diameter 2 will be estimated in Section 3 in terms of a certain combinatorial function that depends on  $n$ ,  $f(n)$ , and a third parameter reflecting the class of groups considered. It turns out that in the case of Abelian groups the combinatorial function can be analyzed by elementary methods and yields the following results, proved in Section 4:

- For any  $c$  such that  $0 < c < 1/2$ , the probability of a random Cayley digraph on an Abelian group of order  $n$  and degree  $\lfloor cn \rfloor$  having diameter 2 is at least  $1 - O(\exp(-c^2n/2))$ .
- For any  $\delta$  such that  $1/2 < \delta < 1$ , the probability of a random Cayley digraph on an Abelian group of order  $n$  and degree  $\lfloor n^\delta \rfloor$  having diameter 2 is at least  $1 - O(\exp(-n^{2\delta-1}/2))$ .
- For each  $\mu(n) \in (0, 1]$  and  $\varepsilon > 0$ , the probability of a random Cayley digraph of degree  $k = \lfloor \sqrt{2n \ln(n/\mu(n))} \rfloor$  on an Abelian 2-group of order  $n$  having diameter 2 is at least  $1 - \mu(n) - \varepsilon$  for all sufficiently large  $n$ .
- If  $f(n)$  is such that  $f(n)/\sqrt{n \ln n} \rightarrow \infty$  as  $n \rightarrow \infty$ , then the probability of a random Cayley digraph of degree  $f(n)$  on an Abelian group of order  $n$  having diameter 2 converges to 1.

The elementary methods of our analysis of random Cayley digraphs on Abelian groups fail when Cayley digraphs on arbitrary groups are considered. However, as we show in Section 5, more sophisticated methods can be used to extend several of the results mentioned above to arbitrary groups. More precisely, we prove the following:

- The probability of a random Cayley digraph of order  $n$  and degree  $k$  having diameter 2 tends exponentially fast to 1 as  $n$  goes to infinity, provided that  $k/n$  remains in a compact subset of the interval  $(0, 1/2)$ .
- For any constant  $\delta$  such that  $1/2 < \delta < 1$ , the probability of a random Cayley digraph of order  $n$  and degree  $\lfloor n^\delta \rfloor$  having diameter 2 tends to 1 as  $n$  goes to infinity.
- If  $f(n)$  is such that  $f(n)/\sqrt{n \ln n} \rightarrow \infty$  as  $n \rightarrow \infty$ , then the probability of a random Cayley digraph of order  $n$  and degree  $f(n)$  having diameter 2 converges to 1.

## 2. THE MODEL

Throughout, let  $G$  be a finite group of order  $n$  and  $k$  a positive integer not exceeding  $n-2$ . The set of non-trivial elements of  $G$  will be denoted by  $G^*$ . For a set  $A$  and an integer  $r$ , the symbol  $\binom{A}{r}$  stands for the set of all subsets of  $A$  of size  $r$ .

For  $S \in \binom{G^*}{k}$ , the *Cayley digraph on  $G$  relative to  $S$* , denoted by  $\text{Cay}(G, S)$ , is the  $k$ -valent digraph with vertex set  $G$  and arc set  $\{(g, gs) : g \in G, s \in S\}$ . The *distance*  $\partial_S(g, h)$  from the vertex  $g$  to the vertex  $h$  in  $\text{Cay}(G, S)$  is the length of the shortest directed path from  $g$  to  $h$  in  $\text{Cay}(G, S)$ . The *diameter*  $\text{diam}(\text{Cay}(G, S))$  is the smallest integer  $d$  such that for every ordered pair  $(g, h)$  the distance from  $g$  to  $h$  is at most  $d$ .

We are now ready to introduce our model for random Cayley digraphs of a given valence. Let  $\mathcal{P}(G, k)$  be the probability space  $(\mathcal{B}, 2^{\mathcal{B}}, \text{Pr})$  where  $\mathcal{B} = \binom{G^*}{k}$ ,  $2^{\mathcal{B}}$  is the power set of  $\mathcal{B}$ , and  $\text{Pr}$  is the uniformly distributed probability measure on  $\mathcal{B}$ . Since  $|\mathcal{B}| = \binom{n-1}{k}$ ,  $\text{Pr}(\{S\}) = \binom{n-1}{k}^{-1}$  for all  $S \in \mathcal{B}$ . More generally, for every subset  $L \subseteq G^*$  of size  $\ell$ , the probability that a random set  $S \in \mathcal{B}$  contains  $L$  as a subset is given by

$$(1) \quad \Pr(S \supseteq L) = \Pr\left(\left\{S \in \binom{G^*}{k} : L \subseteq S\right\}\right) = \binom{n-1-\ell}{k-\ell} \binom{n-1}{k}^{-1}.$$

We can now define a random variable  $\text{Diam} : \mathcal{B} \rightarrow \mathbb{R}$  on the probability space  $\mathcal{P}(G, k)$  by letting, for every  $S \in \binom{G^*}{k}$ ,

$$(2) \quad \text{Diam}(S) = \text{diam}(\text{Cay}(G, S)).$$

The main goal of this article is to derive bounds on the probability of the event  $\text{Diam}(S) > 2$  and study the asymptotic behavior of the bounds.

Since Cayley digraphs are vertex-transitive, the diameter of  $\text{Cay}(G, S)$  coincides with the maximum value of  $\partial_S(1, y)$  over all  $y \in G^*$ . Clearly,  $\partial_S(1, y) \leq 2$  if and only if  $y \in S$ , or there exists  $x \in S$  such that  $(1, x, y)$  is a directed path from 1 to  $y$  of length 2. The latter is equivalent to requiring that  $\{x, x^{-1}y\} \subseteq S$ ; in particular, the events in the following definition play an important role in the analysis.

**Definition 2.1.** For  $x, y \in G^*$ , let

$$T(x, y) = \left\{S : S \in \binom{G^*}{k}, \{x, x^{-1}y\} \subseteq S\right\} \quad \text{and} \quad X(y) = \bigcup_{x \in G^* \setminus \{y\}} T(x, y).$$

If  $S$  is an arbitrary element of  $\binom{G^*}{k}$  and  $\text{Diam}(S) \leq 2$  then, for each  $y \in G^*$ ,  $y \in S$  or  $S \in X(y)$ . Hence  $\Pr(\text{Diam} \leq 2) \leq \min_{y \in G^*} \Pr(y \in S \text{ or } S \in X(y))$ . In particular, if  $\Pr(\overline{X(y)} \mid y \notin S)$  denotes the conditional probability of the complement  $\overline{X(y)}$  of  $X(y)$  given that  $y \notin S$ , we have the following inequality:

$$\Pr(\text{Diam} > 2) \geq \max_{y \in G^*} \Pr(y \notin S \text{ and } S \notin X(y)) = \left(1 - \frac{k}{n-1}\right) \cdot \max_{y \in G^*} \Pr(\overline{X(y)} \mid y \notin S),$$

where the last identity is a direct consequence of (1). On the other hand, if  $\text{Diam}(S) > 2$  then there exists  $y \in G^*$  such that  $y \notin S$  and  $S \not\subseteq X(y)$ . As a result:

$$\Pr(\text{Diam} > 2) \leq \sum_{y \in G^*} \Pr(\overline{X(y)} \mid y \notin S) \cdot \Pr(y \notin S) \leq (n - k - 1) \cdot \max_{y \in G^*} \Pr(\overline{X(y)} \mid y \notin S).$$

Due to these inequalities, if we define

$$(3) \quad M = \max_{y \in G^*} \Pr(\overline{X(y)} \mid y \notin S)$$

then we obtain

$$(4) \quad \left(1 - \frac{k}{n-1}\right) \cdot M \leq \Pr(\text{Diam} > 2) \leq (n - k - 1) \cdot M.$$

Notice that the upper- and lower-bounds in equation (4) differ by a linear factor of order  $n$ . These inequalities provide the basis of our investigation. In what follows, we obtain asymptotic estimates on the probability of the event  $[\text{Diam} > 2]$  from estimates on  $M$ .

### 3. PRELIMINARY ESTIMATES

In this section we derive bounds on the quantity  $M$  appearing in (3) and (4). Our analysis is based on the following coefficients.

**Definition 3.1.** Let  $A$  be a fixed set of size  $n$  and  $J$  be a set of  $t$  pairwise disjoint unordered pairs from  $A$ , where  $0 \leq 2t \leq n$ . Define  $a(n, k, t)$  to be the number of  $k$ -subsets of  $A$  that contain, as a subset, none of the pairs from  $J$ .

Notice that if  $n = k + t$  then  $a(n, k, t) = 2^t$  by the pigeonhole principle. Furthermore, if  $n < k + t$  then  $a(n, k, t) = 0$  for the same reason.

**Lemma 3.2.** For any  $n \geq k + t$  and  $n \geq 2t$  we have

$$(5) \quad a(n, k, t) = \sum_{\ell=0}^k \binom{n-2t}{\ell} \binom{t}{k-\ell} 2^{k-\ell}.$$

*Proof.* Let  $J$  be a set of  $t$  pairwise disjoint unordered pairs from  $A$ , and let  $X$  be the set of  $k$ -subsets of  $A$  containing none of the pairs from  $J$ . Clearly,  $a(n, k, t) = |X|$ . To determine the cardinality of  $X$ , let  $U = A \setminus \cup J$  be the set of all elements of  $A$  appearing in none of the pairs from  $J$ . Fix  $\ell$ ,  $0 \leq \ell \leq \min\{k, n - 2t\}$ , and count the number of  $k$ -subsets  $S \in X$  that intersect  $U$  in a set of size  $\ell$ . Note that  $S \cap U$  can be chosen in  $\binom{n-2t}{\ell}$  ways. On the other hand,  $S \setminus U$  can be selected by first choosing the  $k - \ell$  pairs from  $J$  that  $S$  intersects, and then choosing which of the two elements of each intersecting pair it contains. Hence, there are precisely  $\binom{n-2t}{\ell} \binom{t}{k-\ell} 2^{k-\ell}$  subsets  $S \in X$  with  $|S \cap U| = \ell$ . The result follows by summing over all integers  $0 \leq \ell \leq \min\{k, n - 2t\}$ , and noticing that the possibly remaining terms in the summation vanish when  $\ell > (n - 2t)$ .  $\square$

A straightforward consequence of the definition is that the function  $a(n, k, t)$  is decreasing in  $t$  and increasing in  $n$ . Later on we need the following direct consequence of Lemma 3.2.

**Lemma 3.3.** If  $\alpha \geq 2$  and  $\beta \geq 0$  are integers then  $f(k, t) = a(\alpha t + \beta, k, t)$  is increasing in  $t$ .

*Proof.* Note that  $f(k, t) = \sum_{j=0}^k \binom{(\alpha-2)t+\beta}{j} \binom{t}{k-j} 2^{k-j}$ . Since  $\alpha \geq 2$ , each of the factors in this sum are increasing functions of  $t$ , and hence so is the function  $f(k, t)$ .  $\square$

We are now ready to bound the quantity  $M$  appearing in (3) and (4). In what follows, the ‘‘square roots’’ of elements in  $G$  will play an important role. We will therefore first introduce and discuss the set

$$(6) \quad \sigma(y) = \{x \in G : x^2 = y\},$$

defined for an arbitrary element  $y \in G$ .

**Lemma 3.4.** Let  $G$  be a finite group. If the order of  $G$  is odd, then  $\sigma(y)$  is a singleton for every  $y \in G$ . If the order of  $G$  is even, then there exists at least one element  $y \in G$  with  $\sigma(y) = \emptyset$ .

*Proof.* Consider the mapping  $s : G \rightarrow G$ ,  $x \mapsto x^2$ . Note that  $|\sigma(y)| = 1$ , for all  $y \in G$ , if and only if  $s$  is a bijection. Suppose first that  $|G| = 2m + 1$  is odd; in particular,  $x^{2m+1} = 1$ , for all  $x \in G$ , due to Lagrange's theorem. Hence, for every  $x, y \in G$ , we see that  $x^2 = y^2$  implies that  $x = x^{2m+2} = y^{2m+2} = y$ . So  $s$  is a bijection and the claim of the lemma follows. On the other hand, if  $|G|$  is even,  $G$  has a non-trivial involution  $x$ . Since  $s(x) = 1 = s(1)$ ,  $s$  is not a bijection. Since  $G$  is finite, this implies that  $s$  is not surjection, which in turn implies that there is an element  $y \in G$  such that  $\sigma(y) = \emptyset$ .  $\square$

We can now determine the exact value of  $M$  when  $G$  is Abelian.

**Lemma 3.5.** *If  $G$  is an Abelian group of order  $n$  and  $k$  an integer such that  $1 \leq k \leq n - 2$  then*

$$(7) \quad M = \binom{n-2}{k}^{-1} \cdot a(2t, k, t),$$

where

$$t = \left\lceil \frac{n-2}{2} \right\rceil = \begin{cases} (n-2)/2 & \text{if } n \text{ is even;} \\ (n-3)/2 & \text{if } n \text{ is odd.} \end{cases}$$

*Proof.* Notice that, for each  $y \in G^*$ , the distribution of  $S$  when we condition of the event  $[y \notin S]$  is uniformly distributed over  $\mathcal{B}_y := \binom{G^* \setminus \{y\}}{k}$ ; in particular, if we define  $\Pr_y(A) = |A| / \binom{n-2}{k}$ , for all  $A \subset \mathcal{B}_y$ , then  $\Pr(\overline{X(y)} \mid y \notin S) = \Pr_y(X'(y))$ , where we have defined

$$(8) \quad X'(y) = \{S \in \mathcal{B}_y : \{x, x^{-1}y\} \not\subset S, \text{ for all } x \in G^* \setminus \{y\}\}.$$

As a result,  $M = \binom{n-2}{k}^{-1} \cdot m$ , where  $m = \max\{|X'(y)| : y \in G^*\}$ .

Fix  $y \in G^*$  and, for each  $x \in G^* \setminus \{y\}$ , define  $T_x = \{x, x^{-1}y\} \subset G^* \setminus \{y\}$ . Because  $G$  is Abelian,  $T_x \cap T_z \neq \emptyset$  if and only if  $T_x = T_z$ ; in particular, the set  $\{T_x : x \in G^* \setminus \{y\}\}$  partitions  $G^* \setminus \{y\}$ . Clearly,  $T_x$  is a singleton if and only if  $x \in \sigma(y)$ , in which case  $T_x = \{x\}$ . As a result,  $\{T_x : x \in G^* \setminus (\{y\} \cup \sigma(y))\}$  partitions  $G^* \setminus (\{y\} \cup \sigma(y))$  into  $\tau = (n-2 - |\sigma(y)|)/2$  unordered pairs. From the definition in (8), we see that a  $k$ -subset  $S \subset G^* \setminus \{y\}$  belongs to  $X'(y)$  if and only if it is a subset of  $G^* \setminus (\{y\} \cup \sigma(y))$  which does not contain, as a subset, any of the aforementioned pairs. Due to Definition 3.1, there are precisely  $a(2\tau, k, \tau)$  such subsets  $S$ . But notice that  $a(2\tau, k, \tau)$  is an increasing function of  $\tau$  because of Lemma 3.4; in particular,  $|X'(y)|$  is maximal when  $|\sigma(y)|$  is minimal for  $y \in G^*$ . By Lemma 3.3, if  $G$  has odd order then  $|\sigma(y)| = 1$  for all  $y \in G^*$ , and hence  $m = a(2t, k, t)$  with  $t = (n-3)/2$ . On the other hand, if  $G$  has even order, then there exists an element  $y \in G^*$  with  $\sigma(y) = \emptyset$ , implying that  $m = a(2t, k, t)$  with  $t = (n-2)/2$ .  $\square$

If  $G$  is not Abelian then we may bound  $M$  from above, as stated in the following theorem.

**Lemma 3.6.** *If  $G$  is a finite group of order  $n$  and  $k$  an integer such that  $1 \leq k \leq n - 2$  then*

$$(9) \quad M \leq \binom{n-2}{k}^{-1} \cdot a(3t+1, k, t),$$

where

$$t = \begin{cases} \lfloor (n-1)/3 \rfloor & \text{if } |G| \text{ is even;} \\ \lfloor (n-2)/3 \rfloor & \text{if } |G| \text{ is odd.} \end{cases}$$

*Proof.* Let  $y \in G^*$  and  $s = |\sigma(y)|$ . Define  $C = G^* \setminus (\{y\} \cup \sigma(y))$ . We first show that there exists a set  $J \subseteq C$  of size  $\tau = \lfloor \frac{n-1-s}{3} \rfloor$  such that the sets  $\{x, x^{-1}y\}$ , with  $x \in J$ , are distinct, pairwise disjoint, and of cardinality 2. (Notice that for each  $x \in C$ ,  $\{x, x^{-1}y\} \subset C$ .) We shall define such a set  $J$  recursively.

If  $s \geq n-3$  then  $\tau = 0$ , and  $J = \emptyset$  has the desired properties. So we may assume that  $s \leq n-4$ . Then the set  $C$  is non-empty, and we can choose  $x_1 \in C$  and set  $J_1 = \{x_1\}$ . Now suppose that  $J_\ell$ , with  $1 \leq \ell < \tau$ , has been defined so that  $J_\ell \subseteq C$ ,  $|J_\ell| = \ell$ , and  $|\cup_{x \in J_\ell} \{x, x^{-1}y\}| = 2\ell$ . Let  $K_\ell = \cup_{x \in J_\ell} \{x, yx^{-1}, x^{-1}y\}$ . Since  $|K_\ell| \leq 3\ell \leq 3\tau - 3 \leq n - s - 4 < n - s - 2 = |C|$ , we can choose an element  $x_{\ell+1} \in C \setminus K_\ell$  and define  $J_{\ell+1} = J_\ell \cup \{x_{\ell+1}\}$ . Clearly,  $|J_{\ell+1}| = \ell+1$  and  $J_{\ell+1} \subseteq C$ . Now suppose that  $|\cup_{x \in J_{\ell+1}} \{x, x^{-1}y\}| < 2\ell+2$ . Then one of the elements  $x_{\ell+1}$  or  $x_{\ell+1}^{-1}y$  must belong to  $\cup_{x \in J_\ell} \{x, x^{-1}y\}$ . However, both of these cases imply that  $x_{\ell+1} \in K_\ell$ , which is not possible. This construction yields therefore a set  $J = J_\tau$  with the desired properties.

Notice that

$$M = \binom{n-2}{k}^{-1} \cdot \max_{y \in G^*} |X'(y)|,$$

where  $X'(y)$  is as defined in (8). We first show that  $|X'(y)| \leq a(3\tau+1, k, \tau)$ , where  $\tau = \lfloor \frac{n-1-|\sigma(y)|}{3} \rfloor$  is the cardinality of the aforementioned set  $J$ . Note that  $\{x, x^{-1}y\}$  is a singleton whenever  $x \in \sigma(y)$ . In particular, for  $k$ -subset  $S \subset G^* \setminus \{y\}$ ,  $S \in X'(y)$  if and only if, for each  $x \in G^* \setminus (\{y\} \cup \sigma(y))$ ,  $\{x, x^{-1}y\} \not\subset S$ . Since  $J \subset G^* \setminus (\{y\} \cup \sigma(y))$ , for each  $x \in J$ ,  $\{x, x^{-1}y\} \subset G^* \setminus (\{y\} \cup \sigma(y))$ , and the definition of the coefficients  $a(n, k, t)$  implies that  $|X'(y)| \leq a(n-1-|\sigma(y)|, k, \tau)$ . But observe that  $n-2-|\sigma(y)| \leq 3\tau+1$ . Since  $a(n, k, t)$  is an increasing function of  $n$ ,  $|X'(y)| \leq a(3\tau+1, k, \tau)$ . Finally, by Lemma 3.3, the function  $a(3\tau+1, k, \tau)$  is increasing in  $\tau$  and since the maximum value of  $\tau$  is achieved for those  $y$  which have  $|\sigma(y)|$  minimal, it follows from Lemma 3.4 that  $|X'(y)| \leq a(3t+1, k, t)$ , where  $t$  is now as in the statement of the lemma. Since this holds for every  $y \in G^*$ , the result follows.  $\square$

Knowledge of the asymptotic behaviour of the function  $a(n, k, t)$  when  $n$  is asymptotically linear in  $t$  would allow us to make conclusions about the asymptotic behaviour of the random variable  $\text{Diam}$ . For example, Theorem 5.1 shows that if  $\lim_{n \rightarrow \infty} (n-k-1) \binom{n-2}{k}^{-1} a(3t+1, \lfloor cn \rfloor, t) = 0$  for a constant  $c$  such that  $0 < c < 1/2$ , then the diameter of a random Cayley digraph of order  $n$  and degree  $\lfloor cn \rfloor$  is asymptotically almost surely equal to two. Similar statements, with  $t$  equal to different linear functions of  $n$ , could be made for random Abelian Cayley digraphs on the basis of Theorem 4.1. In the next sections we will show that limits such as the above are indeed equal to zero and therefore the corresponding random Cayley digraphs almost surely have diameter two. We also consider the lower bound in the Abelian case, and describe more precisely the threshold at which  $\Pr(\text{Diam} \leq 2)$  jumps asymptotically away from 0.

#### 4. ANALYSIS IN ABELIAN GROUPS

In this section we restrict ourselves to Abelian groups, for which we can prove the following result. The proof is omitted as it is a direct consequence of equations (3) and (4), Lemma 3.5 and the first equality in Lemma 3.2.

**Theorem 4.1.** *If  $G$  is Abelian of finite order  $n$  and  $k$  an integer such that  $1 \leq k \leq n-2$  then*

$$(10) \quad \left(1 - \frac{k}{n-1}\right) \cdot \binom{n-2}{k}^{-1} \cdot \binom{t}{k} \cdot 2^k \leq \Pr(\text{Diam} > 2) \leq (n-k-1) \cdot \binom{n-2}{k}^{-1} \cdot \binom{t}{k} \cdot 2^k,$$

where  $t$  is as in Lemma 3.5.

We start the analysis by observing that in the context of Lemma 3.5 we have that  $n-2 = 2t+d$  where  $d = 0$  if  $n$  is even and  $d = 1$  if  $n$  is odd. Therefore,

$$(11) \quad M = \binom{n-2}{k}^{-1} a(2t, k, t) = \binom{2t+d}{k}^{-1} \left[ 2^k \binom{t}{k} \right] = b(t, k) c(t, k),$$

where

$$b(t, k) = \binom{2t}{k}^{-1} 2^k \binom{t}{k} \quad \text{and} \quad c(t, k) = \left( \frac{2t-k+1}{2t+1} \right)^d.$$

Observe that the function  $c(t, k)$  converges to 1 as  $t \rightarrow \infty$  uniformly on the set  $\{(t, k) \in \mathbb{R}_+^2 : 0 \leq k \leq f(t)\}$  for any function  $f(t)$  such that  $f(t)/t \rightarrow 0$  as  $t \rightarrow \infty$ . Furthermore, for any constants  $c_1, c_2$  satisfying  $0 \leq c_1 < c_2 < 1$  there exist constants  $d_1, d_2$  such that  $0 < d_1 \leq c(t, k) \leq d_2$  whenever  $c_1 t \leq k \leq c_2 t$ . In particular, if the ratio  $\lambda = k/t$  remains within prescribed bounds as  $t \rightarrow \infty$ , away from zero, then the value of  $c(t, k)$  remains within a bounded interval, away from zero.

For the asymptotic analysis of the behaviour of binomial coefficients appearing in  $b(t, k)$  we use Stirling's approximation. To state the result of the corresponding routine calculations in a concise

form, for  $0 < \lambda < 1$  let

$$\begin{aligned} R(\lambda) &= (2 - \lambda) \ln(1 - \lambda/2) - (1 - \lambda) \ln(1 - \lambda) , \\ P(\lambda) &= \left( \frac{2 - \lambda}{2 - 2\lambda} \right)^{1/2} , \text{ and} \\ C(t, \lambda) &= 1 + O(t^{-1}) + O((t\lambda)^{-1}) + O((t(1 - \lambda))^{-1}) \quad \text{as } t \rightarrow \infty. \end{aligned}$$

Then, writing  $k = \lambda t$ , routine calculation with the help of Stirling's approximation yields

$$(12) \quad b(t, k) = \exp(tR(\lambda)) P(\lambda) C(t, \lambda) ;$$

the terms  $R$ ,  $P$  and  $C$  represent the exponential rate, the leading coefficient, and the correction term. The exponential rate  $R(\lambda)$  is easily seen to be negative for  $0 < \lambda < 1$ . Furthermore  $C(t, \lambda)$  tends to 1 for any fixed  $\lambda \in (0, 1)$  as  $t \rightarrow \infty$ . Note also that

$$(13) \quad R(\lambda) = (2 - \lambda) \ln(1 - \lambda/2) - (1 - \lambda) \ln(1 - \lambda) = -\lambda^2/4 + O(\lambda^3) \text{ for } 0 < \lambda < 1.$$

Our first result about the behaviour of  $\Pr(\text{Diam} > 2)$  in the case of Abelian groups deals with the situation where  $k \approx cn$  for some  $c < 1/2$ . In our asymptotic calculations for  $n \rightarrow \infty$  we may then replace  $\lambda = k/t$  for  $t = \lfloor (n - 2)/2 \rfloor$  with the value  $2c$ . Combining now (11), (12) and (13) with Theorem 4.1 we arrive at the following result:

**Theorem 4.2.** *For any constant  $c$  such that  $0 < c < 1/2$ , the probability of a random Cayley digraph on an Abelian group of order  $n$  and degree  $\lfloor cn \rfloor$  having diameter 2 is at least  $1 - O(\exp(-c^2n/2))$ .  $\square$*

We now turn to the case where  $k \approx n^\delta$  with  $1/2 < \delta < 1$ . For  $k = \lambda t$  with  $\lambda = o(1)$  as  $t \rightarrow \infty$ , the approximation (13) is still valid, and we also have  $C(t, \lambda) = 1 + O(\lambda)$ . Thus, if  $k$  grows at least as fast as  $n^\delta$  with  $\delta > 1/2$ , for asymptotic computation we may replace  $k$  with  $2n^{\delta-1}t$  and set  $\lambda = 2n^{\delta-1}$ . Then, the exponent  $tR(\lambda)$  in (12) may be replaced with  $-n^{2\delta-1}/2$ , which implies exponential decay if  $\delta > 1/2$ . Using Theorem 4.1 again, we have the following conclusion.

**Theorem 4.3.** *For any constant  $\delta$  such that  $1/2 < \delta < 1$ , the probability of a random Cayley digraph on an elementary Abelian 2-group of order  $n$  and degree  $\lfloor n^\delta \rfloor$  having diameter 2 is at least  $1 - O(\exp(-n^{2\delta-1}/2))$ .  $\square$*

It is natural to ask what happens when  $k \approx \sqrt{n}$ . By the Moore bound for diameter two, the probability of a random Cayley digraph of degree  $k$  and order  $n$  (for a general group) having diameter 2 is zero if  $k < \sqrt{n}$ . It is interesting to note that if the right-hand side is increased by a factor of 2, then for any  $n$  of the form  $n = 2^{2d}$  there exists a Cayley (di)graph of order  $n$  and degree  $k = 2\sqrt{n} = 2 \cdot 2^d$  on an elementary Abelian group of order  $n$ , which has diameter 2. Indeed, representing vertices of the graph as  $2d$ -dimensional 0-1 vectors, it is sufficient to consider a generating set of the form  $S_1 \cup S_2$  where  $S_1$  and  $S_2$  consists of all non-zero vectors having the first  $d$  and the last  $d$  coordinates equal to zero, respectively. It follows that the probability that a random Cayley (di)graph on an elementary Abelian 2-group of order  $n$  and degree  $k = 2\sqrt{n}$  has diameter 2 is positive. This, of course, does not allow to make any conclusion as to how large this probability might be.

However, our approximation above shows that if  $k = \lfloor c\sqrt{n} \rfloor$ , then the upper bound from part (10) of Theorem 4.1 tends to infinity as  $n \rightarrow \infty$ , while the lower bound converges to  $\exp(-c^2/2)$ . In particular, this shows that, when restricted to Cayley graphs on Abelian groups of order  $n$  and valence  $c\sqrt{n}$ , the probability  $\Pr(\text{Diam} = 2)$  does *not* tend to 1 as  $n \rightarrow \infty$ , for any value of  $c$ . Note that when  $c \leq 1$ , our lower bound is inferior to the Moore bound.

This brings us to the question in the Introduction concerning the threshold for  $k = f(n)$  at which the asymptotic value of the upper bound on  $\Pr(\text{Diam} > 2)$  undergoes a phase transition, switching abruptly from 1 to 0 as  $k$  increases. The facts above give a lot of information on this point in the case of Abelian groups for which we have both an upper and a lower estimate (10) of Theorem 4.1 on the probability of having diameter greater than 2. Note that we have already seen that the Moore bound does not give the correct asymptotic order for the threshold.

**Theorem 4.4.** *Let  $G$  be a finite Abelian group of order  $n$  and let  $P(n, k)$  denote the probability that a random Cayley digraph of degree  $k = f(n)$  on  $G$  has diameter at most 2. Suppose that  $0 < \mu(n) \leq \mu'(n) \leq 1$ .*

- If  $f(n)/\sqrt{n \ln n} \rightarrow \infty$ , then  $P(n, k) \rightarrow 1$  as  $n \rightarrow \infty$ .
- If  $f(n)/\sqrt{2n \ln(n/\mu(n))} \rightarrow 1$  then  $P(n, k) \geq 1 - \mu'(n)$  for all sufficiently large  $n$ .
- If  $f(n)/\sqrt{2n(-\ln \mu'(n))} \rightarrow 1$  then  $P(n, k) \leq 1 - \mu(n)$  for all sufficiently large  $n$ .

*Proof.* To simplify the calculations, in the asymptotics we may replace  $t$  with  $n/2$  and  $\lambda = k/t$  with  $2k/n$ ; moreover, by (13) and other facts above we may replace the upper bound in question with  $n \exp[-k^2/(2n) + O(k^3/n^2)]$ . Note that this approximation works well if  $k = o(n^{2/3})$ , which includes the region of interest to us here. Now  $n \exp[-k^2/(2n) + O(k^3/n^2)]/\mu(n) \rightarrow 1$  if and only if  $k/\sqrt{2n \ln(n/\mu(n))} \rightarrow 1$ . Part (ii) follows directly from the definition of limit. By letting  $\mu(n) \rightarrow 0$ , we obtain (i). By the same method, we can deal with the lower bound and obtain (iii).  $\square$

We move on in the next section to general finite groups. Results similar to those we have already obtained for Abelian groups could perhaps in principle be developed in a way similar to the elementary approach above, but with much more effort and facing difficulties which we now outline.

The upper bound on  $M$  for general groups now involves the factor  $a(3t+1, k, t)$ . The expression for this factor involving binomial coefficients in (5) will have a number of summands of order  $k$ , which is not constant. It seems unlikely therefore that an elementary asymptotic analysis will work. We cannot hope for better — as a consequence of [6, 10] this combinatorial sum cannot be converted to a ‘closed formula’ of length independent on the parameters  $n$  and  $k$  for  $k = cn$  or  $k = n^\delta$ . We will deal with this difficulty in the next section by using a more sophisticated approach.

## 5. ASYMPTOTIC ANALYSIS FOR GENERAL GROUPS

In this section we address the case of general finite groups. For this case, Lemma 3.6 provides an upper-bound for  $M$  in terms of the coefficients  $a(n, k, t)$ . The following result is now a direct consequence of equation (4).

**Theorem 5.1.** *If  $G$  is a finite group of order  $n$  and  $k$  an integer such that  $1 \leq k \leq n - 2$  then*

$$(14) \quad \Pr(\text{Diam} > 2) \leq (n - k - 1) \cdot \binom{n-2}{k}^{-1} \cdot a(3t+1, k, t),$$

where  $t$  is as in Lemma 3.6.

In what follows, we aim to analyze the asymptotic behavior of the upper-bound in equation (14) for the case where  $k = \lfloor cn \rfloor$  and  $k = \lfloor n^\delta \rfloor$ , with  $0 < c < 1$  and  $1/2 < \delta < 1$ . We do so in general by considering the asymptotic behavior of  $a(n, k, t)$  when  $n \geq 2t$ . The most technical aspect of the asymptotic analysis is when  $n > k + t$ , which we address reducing it to a uniform asymptotic expansion of a one-dimensional parameter-varying integral. When  $k = \Theta(n)$ , the asymptotic behavior of the resulting integral relies on the stationary phase method, as found for example in [11]. Instead, when  $k = o(n)$ , the analysis follows the lines of [2, 3]. We note in passing that the first of these regimes could be easily addressed using the methods recently developed in [7, 8] (see [9, Section 4.9] for more details). However, the methods of [7, 8] do not work directly in the sublinear case, unlike the methods of the present paper.

**5.1. Integral representation.** We begin our study of the asymptotic behaviour of  $a(n, k, t)$ , when  $n \geq 2t$ . Define  $d_1 = (n - 2t)/n$ ,  $d_2 = t/n$  and  $d_3 = k/n$ . In what follows, unless otherwise stated,  $d_1$ ,  $d_2$  and  $d_3$  always stand as short forms of these functions of  $(n, k, t)$ . There are three possible asymptotic regimes to consider in terms of the relationship between  $d_1 + d_2$  and  $d_3$ .

If  $d_1 + d_2 < d_3$  i.e.  $k + t > n$  then

$$(15) \quad a(n, k, t) = 0.$$

On the other hand, if  $d_1 + d_2 = d_3$  i.e.  $k + t = n$  then

$$(16) \quad a(n, k, t) = 2^t.$$

In what remains of this section, we assume that  $d_1 + d_2 > d_3$  i.e.  $n > k + t$ .

Recalling formula (5), for fixed  $n$  and  $t$ , consider the generating function

$$\begin{aligned} \sum_{k \geq 0} a(n, k, t) x^k &= \sum_{k \geq 0} x^k \sum_{\ell=0}^k \binom{n-2t}{\ell} \binom{t}{k-\ell} 2^{k-\ell} \\ &= \sum_{i \geq 0} \binom{n-2t}{i} x^i \cdot \sum_{j \geq 0} \binom{t}{j} 2^j x^j = (1+x)^{n-2t} (1+2x)^t. \end{aligned}$$

Due to Cauchy's formula, the above implies that

$$a(n, k, t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} r^{-k} (1 + r e^{i\theta})^{n-2t} (1 + 2r e^{i\theta})^t e^{-ik\theta} d\theta,$$

for all  $r > 0$ . Since the modulus of the integrand is maximized at  $\theta = 0$ , we normalize the integral by the factor  $(1+r)^{n-2t} (1+2r)^t$ , and rewrite the integrand in an exponential-logarithmic form to obtain

$$(17) \quad a(n, k, t) = E(r; n, k, t) \cdot I(r; n, k, t),$$

where

$$\begin{aligned} E(r; n, k, t) &= (2\pi)^{-1} \cdot r^{-k} (1+r)^{n-2t} (1+2r)^t, \\ I(r; n, k, t) &= \int_{-\pi}^{\pi} \exp(-n \cdot F(\theta; r, d_1, d_2, d_3)) d\theta, \\ F(\theta; r, d_1, d_2, d_3) &= d_3 \cdot i\theta - d_1 \cdot \ln \left\{ \frac{1 + r e^{i\theta}}{1+r} \right\} - d_2 \cdot \ln \left\{ \frac{1 + 2r e^{i\theta}}{1+2r} \right\}, \end{aligned}$$

where logarithms are always to be interpreted in the principal sense.

To determine the asymptotic behavior of the coefficients  $a(n, k, t)$ , the goal is to tune  $r$  with  $(n, k, t)$  so that  $I(r; n, k, t)$  decays polynomially with  $n$ , in which case  $E(r; n, k, t)$  captures the precise exponential growth rate of the coefficients. We note for later that

$$(18) \quad \frac{\ln E(r; n, k, t)}{n} = d_1 \ln(1+r) + d_2 \ln(1+2r) - d_3 \ln r.$$

To accomplish the above notice that

$$\frac{\partial F}{\partial \theta}(0; r, d_1, d_2, d_3) = i \left( d_3 - \frac{d_1 r}{1+r} - \frac{2d_2 r}{1+2r} \right).$$

Thus, in order for  $\theta = 0$  to be a stationary point of  $F(\theta; r, d_1, d_2, d_3)$ ,  $r$  and  $(n, k, t)$  must satisfy the relation  $d_1 r / (1+r) + 2d_2 r / (1+2r) = d_3$ . A routine calculation shows that the only non-negative real solution to this equation is given by the formula

$$(19) \quad r = \frac{2d_3}{(1-3d_3) + \sqrt{(1-3d_3)^2 + 8d_3(d_1 + d_2 - d_3)}} > 0.$$

In what remains of the manuscript, and unless otherwise stated,  $r$  always stands for the above function of  $(d_1, d_2, d_3)$  when  $d_1 + d_2 > d_3$ . Furthermore, it is understood that  $r$  is given by (19) when it is omitted from the notation. Thus, for example,  $F(\theta; d_1, d_2, d_3)$  denotes  $F(\theta; r; d_1, d_2, d_3)$ ; in particular,  $\theta = 0$  is a stationary point of  $F(\theta; d_1, d_2, d_3)$ .

On the other hand, notice that

$$\frac{\partial^2 F}{\partial \theta^2}(0; d_1, d_2, d_3) = \frac{d_1 r}{2(1+r)^2} + \frac{d_2 r}{(1+2r)^2}.$$

Using that  $d_1 + 2d_2 = 1$ , it follows almost immediately that

$$(20) \quad \frac{\partial^2 F}{\partial \theta^2}(0; d_1, d_2, d_3) \geq \frac{r}{2(1+2r)^2}.$$

Similarly, but after using that  $\ln(1-w) \leq -w/2$ , for all  $0 \leq w \leq 1$ , it also follows that

$$(21) \quad \operatorname{Re}\{F(\theta; d_1, d_2, d_3)\} \geq \frac{(1 - \cos \theta)r}{2(1+2r)^2}.$$

Three distinct regimes are possible when  $d_1 + d_2 > d_3$ , depending on the behavior of  $r$  as  $n \rightarrow \infty$ .



If  $r \rightarrow \infty$  then (17), (18) and (21) imply that  $a(n, k, t) = O(2^t \cdot e^{O(n/r + (d_1 + d_2 - d_3) \ln(r))})$ . In the special case that  $d_3$  stays bounded away from  $1/3$ , it follows from (19) that  $r(d_1 + d_2 - d_3) = \Theta(1)$ ; in particular,  $a(n, k, t) = O(2^t \cdot e^{O(n(d_1 + d_2 - d_3))})$ . Hence, if  $n(d_1 + d_2 - d_3)$  remains bounded then

$$(22) \quad a(n, k, t) = O(2^t).$$

This bound is sharp because  $a(n, k, t) = 2^t$  when  $d_1 + d_2 = d_3$ .

On the other hand, if  $r$  stays in a compact subset  $C$  of  $(0, +\infty)$  then

$$(23) \quad a(n, k, t) = \Theta(n^{-1/2} e^{n(d_1 \ln(1+r) + d_2 \ln(1+2r) - d_3 \ln r)}).$$

Due to the identity in (18), the above is equivalent to having  $I(n, k, t) = \Theta(n^{-1/2})$ , uniformly for all  $(n, k, t)$  such that  $r \in C$ . To see this, first think of  $(\theta; r, d_1, d_2, d_3)$  as a vector of unrelated variables, and notice that there exists a sufficiently small  $0 < \delta < \pi$  such that  $F(\theta; r, d_1, d_2, d_3)$  is an analytic function of  $\theta$  for  $|\theta| < \delta$ , for all  $(r, d_1, d_2, d_3)$  such that  $r \in C$ . Due to (21),  $I(n, k, t)$  is localizable around  $\theta = 0$  and, because of the analyticity, the end-points of the integral may be shifted by incurring exponentially small errors that stay uniform for  $r \in C$ . Since  $\theta = 0$  is a stationary point of  $F(\theta; d_1, d_2, d_3)$ , the inequality in (20) together with Laplace's method imply the claimed asymptotics for  $I(n, k, t)$ .

Finally, if  $r \rightarrow 0^+$  then

$$(24) \quad a(n, k, t) = \Theta(k^{-1/2} e^{n(d_1 \ln(1+r) + d_2 \ln(1+2r) - d_3 \ln r)}).$$

Similarly as before, all reduces to show now that  $I(n, k, t) = \Theta(k^{-1/2})$ . The difficulty here is that  $F(\theta; d_1, d_2, d_3)$  converges uniformly to 0 for all  $-\pi \leq \theta \leq \pi$  when  $r \rightarrow 0^+$ . We resolve this issue exploiting that  $F(\theta; r, d_1, d_2, d_3)$  is also analytic with respect to  $r$  near  $r = 0$ . Indeed, thinking again of  $(\theta; r, d_1, d_2, d_3)$  as a vector of unrelated variables, notice that there is  $\eta > 0$  such that  $F(\theta; r, d_1, d_2, d_3)$  is an analytic function of  $(\theta; r)$  for  $|\theta| < 2\pi$  and  $|r| < \eta$ , for all  $(d_1, d_2, d_3)$ . In particular, since  $F(\theta; 0, d_1, d_2, d_3) - \frac{\partial F}{\partial \theta}(0; 0, d_1, d_2, d_3)\theta = 0$ , for all  $|\theta| < 2\pi$ , it follows from the Hartogs' series of this function that

$$F(\theta; r, d_1, d_2, d_3) - \frac{\partial F}{\partial \theta}(0; r, d_1, d_2, d_3)\theta = r \cdot G(\theta; r, d_1, d_2, d_3),$$

for certain function  $G(\theta; r, d_1, d_2, d_3)$ , analytic in  $(\theta; r)$  for  $|\theta| < 2\pi$  and  $|r| < \eta$ . Recall that  $G(\theta; d_1, d_2, d_3)$  stands for  $G(\theta; r, d_1, d_2, d_3)$  when  $r$  is given by formula (19). Since  $\theta = 0$  is a stationary point of  $F(\theta; d_1, d_2, d_3)$ , the above identity implies that  $F(\theta; d_1, d_2, d_3) = r \cdot G(\theta; d_1, d_2, d_3)$ . Hence

$$(25) \quad I(n, k, t) = \int_{-\pi}^{\pi} e^{-nr \cdot G(\theta; d_1, d_2, d_3)} d\theta.$$

But notice that  $nr \rightarrow \infty$  because  $nr \sim k$  for  $d_1 r / (1+r) + 2d_2 r / (1+2r) = d_3$  and  $d_1 + 2d_2 = 1$ . On the other hand, due to (19), (20) and (21), it is immediate that  $\theta = 0$  is a stationary point of  $G(\theta; d_1, d_2, d_3)$ , and

$$\begin{aligned} \frac{\partial^2 G}{\partial \theta^2}(0; d_1, d_2, d_3) &\geq \frac{1}{2(1+2r)^2}, \\ \operatorname{Re}\{G(\theta; d_1, d_2, d_3)\} &\geq \frac{1 - \cos \theta}{2(1+2r)^2}. \end{aligned}$$

The arguments used for the previous regime now imply that  $I(n, k, t) = \Theta((nr)^{-1/2}) = \Theta(k^{-1/2})$ , as claimed.

**5.2. The linear case.** We first specialize the results of the previous section to determine the asymptotic order of the coefficients upper-bound for the probability of the event  $[\text{Diam} > 2]$  in Theorem 5.1, when  $k = \lfloor cn \rfloor$ , with  $0 < c < 1$ , and  $t \sim n/3$  is as in Lemma 3.6; in particular,  $d_1 = (t+1)/(3t+1) \sim 1/3$ ,  $d_2 = t/(3t+1) \sim 1/3$  and  $d_3 = k/(3t+1) \sim c$ . We distinguish between three possible regimes, according to the value of  $c$  relative to  $2/3$ .

If  $0 < c < 2/3$  then, for large enough  $n$ ,  $d_1 + d_2 > d_3$  and  $r \rightarrow r_c$ , where

$$(26) \quad r_c = \frac{2c}{(1-3c) + \sqrt{(1-3c)^2 + 8c(2/3-c)}}.$$

In particular, if  $c$  is bounded away from 0 and  $2/3$  then, for sufficiently large  $n$  independent of  $c$ ,  $r$  is bounded to a compact subset of  $(0, +\infty)$  and the asymptotic regime of the coefficients  $a(3t+1, k, t)$  is described by equation (23). The exponential rate of these coefficients w.r.t.  $n$  is therefore given by  $-c \ln(r_c) + \ln(1+r_c)/3 + \ln(1+2r_c)/3$ .

On the other hand, using that  $n! = \Gamma(n+1)$  and the asymptotic expansion of the Gamma function at infinity, we find for  $0 \leq k \leq n-2$  but  $k/(n-2)$  bounded away from 1 that

$$(27) \quad \frac{1}{n} \ln \left\{ \binom{n-2}{k}^{-1} \right\} = \frac{k}{n} \log \left( \frac{k}{n} \right) + \left( 1 - \frac{k}{n} \right) \log \left( 1 - \frac{k}{n} \right) + O \left( \frac{\log(n)}{n} \right).$$

Recall that when  $c \geq 1/2$ , every Cayley graph of degree  $\lfloor cn \rfloor$  and order  $n$  has diameter at most 2. Deferring the study of the value  $c = 0$  to the next subsection dealing with the sub-linear case, the discussion above yields the following result.

**Theorem 5.2.** *If  $C \subset (0, 1/2]$  is a compact set and  $0 < \gamma < \inf_{c \in C} r_c$ , then the probability of a random Cayley digraph of order  $n$  and degree  $k$  having diameter 2 is at least  $1 - O(\sqrt{n} \cdot e^{-n^\gamma})$ , uniformly for all  $n$  and  $k$  such that  $k/n \in C$ , as  $n$  goes to infinity.*

**5.3. The sublinear case.** Finally, we analyze the asymptotic order of the upper-bound for the probability of the event  $[\text{Diam} > 2]$  in Theorem 5.1 when  $k = \Theta(n^\delta)$ , with  $1/2 < \delta < 1$ , and  $t \sim n/3$  is as in Lemma 3.6. Note that the case with  $\delta = 1$  falls into the context of the linear case in the previous section.

As before,  $d_1 \rightarrow 1/3$  and  $d_2 \rightarrow 1/3$ , however,  $d_3 \rightarrow 0^+$ ; in particular,  $r \rightarrow 0^+$ . The asymptotic regime of the coefficients  $a(3t+1, k, t)$  is therefore described by equation (24). Using equation (27), we find that the exponential growth rate of the coefficients  $\binom{n-2}{k}^{-1} a(3t+1, k, t)$  is

$$d_3 \log d_3 + (1 - d_3) \log(1 - d_3) - d_3 \log r + (1 - 2d_2) \log(1 + r) + d_2 \log(1 + 2r) + O \left( \frac{\log(n)}{n} \right).$$

Due to equation (19), this rate is asymptotic to  $-r^2/3$ . When  $k = \Theta(n^\delta)$ , with  $\delta > 1/2$ , the upper-bound of  $\Pr[\text{Diam} > 2]$  is of order  $n^{1-\delta/2} \exp(-n^{(2\delta-1)}(1+o(1))/3)$ , which leads to the following result.

**Theorem 5.3.** *For any constant  $\delta$  such that  $1/2 < \delta < 1$ , the probability of a random Cayley digraph of order  $n$  and degree  $k = \Theta(n^\delta)$  having diameter 2 is at least  $1 - O(n^{1-\delta/2} \exp(-n^{(2\delta-1)}/3))$ .*

## 6. CONCLUSION

We have derived rather detailed information on the event that a random Cayley digraph of a group has diameter 2 in the case of elementary Abelian 2-groups, and slightly less precise information in the general case. Many natural questions have been answered by our asymptotic analysis of upper and lower bounds on probability, but some remain. For example, the case  $k \sim c\sqrt{n}$  is not settled by our analysis, nor is the exact asymptotic order of the phase transition where  $\Pr(\text{Diam} > 2)$  switches from almost inevitable to almost impossible. We know that this order is at most  $\sqrt{n \ln n}$  and definitely exceeds  $\sqrt{n}$ , and we conjecture that for both Abelian and general groups,  $\sqrt{n \ln n}$  is the correct order.

Our upper and lower bounds for the probability differ by a factor of order  $n$ , and tighter bounds will be needed in order to make better progress on these problems.

Finally, recall that the function  $a(n, k, t)$  has a natural combinatorial interpretation which may be more widely applicable. Our asymptotic analysis yields immediate consequences for this function, in regimes including more cases than were needed for the results in this paper, and even more could be deduced by similar methods.

**Acknowledgement.** Work on this paper begun at the University of Auckland, New Zealand, when the second and the third authors were visiting the fourth author thanks to the support of the local Department of Mathematics and NZIMA. Their enquiries about asymptotic analysis led from Auckland to Slovenia (M. Petkovšek) to Pennsylvania (H. Wilf) and then via Robin Pemantle to the first and the fifth author, the latter being blissfully unaware in Auckland of the existence of the work going on in the same building!

Research of the third and the fourth author was also supported by the VEGA Research Grants No. 1/0489/08 and 1/0280/10, the APVV Research Grants No. 0040-06 and 0104-07, and the APVV LPP Research Grants No. 0145-06 and 0203-06. The second author was partially supported by the ARRS grant no. J1-0540.

## REFERENCES

- [1] B. Bollobás. *Graph Theory, An Introductory Course*. Springer, 1979.
- [2] Manuel Lladser. Mixed powers of generating functions. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 171–182. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006.
- [3] Manuel Lladser. Uniform formulae for coefficients of meromorphic functions in two variables. Part I. *SIAM J. Discret. Math.*, 20:811–828, 2006.
- [4] J. Meng and Q. Huang. Almost all Cayley graphs have diameter 2. *Discrete Math.*, 178:267–269, 1998.
- [5] J. Meng and X. Liu. The diameters of almost all Cayley digraphs. *Acta Math. Appl. Sinica (English Ser.)*, 13:400–413, 1997.
- [6] P. Paule and M. Schorn. A Mathematica version of Zeilberger’s algorithm for proving binomial coefficient identities. *J. Symbolic Comput.*, 20:673–698, 1995.
- [7] R. Pemantle and M. C. Wilson. Asymptotics of multivariate sequences. I. Smooth points of the singular variety. *J. Combin. Theory Ser. A*, 97(1):129–161, 2002.
- [8] R. Pemantle and M. C. Wilson. Asymptotics of multivariate sequences. II. Multiple points of the singular variety. *Combin. Probab. Comput.*, 13(4-5):735–761, 2004.
- [9] R. Pemantle and M. C. Wilson. Twenty combinatorial examples of asymptotics derived from multivariate generating functions. *SIAM Rev.*, 50:199–272, 2008.
- [10] M. Petkovšek, H. S. Wilf, and D. Zeilberger. *A = B*. AK Peters, Ltd., 1996.
- [11] R. Wong. *Asymptotic approximations of integrals*. Academic Press Inc., Boston, MA, 1989.

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, CO 80309-0526, USA  
*E-mail address:* `manuel.lladser@colorado.edu`

FACULTY OF MATHEMATICS AND PHYSICS, UNIVERSITY OF LJUBLJANA, LJUBLJANA, SLOVENIA  
*E-mail address:* `primoz.potocnik@fmf.uni-lj.si`

DEPARTMENT OF MATHEMATICS, SvF, SLOVAK UNIVERSITY OF TECHNOLOGY, BRATISLAVA, SLOVAKIA  
*E-mail address:* `siagiova@math.sk`

DEPARTMENT OF MATHEMATICS, OPEN UNIVERSITY, U.K. AND DEPARTMENT OF MATHEMATICS, SvF, SLOVAK UNIVERSITY OF TECHNOLOGY, BRATISLAVA, SLOVAKIA  
*E-mail address:* `j.siran@open.ac.uk`

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF AUCKLAND, PRIVATE BAG 92019 AUCKLAND, NEW ZEALAND  
*E-mail address:* `mcw@cs.auckland.ac.nz`