GENERATION AND RANDOM GENERATION: FROM SIMPLE GROUPS TO MAXIMAL SUBGROUPS

TIMOTHY C. BURNESS, MARTIN W. LIEBECK, AND ANER SHALEV

Abstract. Let G be a finite group and let $d(G)$ be the minimal number of generators for G. It is well known that $d(G) = 2$ for all (non-abelian) finite simple groups. We prove that $d(H) \leq 4$ for any maximal subgroup H of a finite simple group, and that this bound is best possible.

We also investigate the random generation of maximal subgroups of simple and almost simple groups. By applying a recent theorem of Jaikin-Zapirain and Pyber we show that the expected number of random elements generating such a subgroup is bounded by an absolute constant.

We then apply our results to the study of permutation groups. In particular we show that if G is a finite primitive permutation group with point stabilizer H, then $d(G) - 1 \leq d(H) \leq d(G) + 4$.

1. Introduction

Let G be a finite group and let $d(G)$ be the minimal number of generators for G. We say that G is *d-generator* if $d(G) \leq d$. The investigation of generators for finite simple groups has a rich history, with numerous applications. Perhaps the most well known result in this area is the fact that every finite simple group is 2-generator. For the alternating groups, this was first stated in a 1901 paper of G.A. Miller [47]. In 1962 it was extended by Steinberg [54] to the simple groups of Lie type, and post-Classification, Aschbacher and Guralnick [2] completed the proof by analysing the remaining sporadic groups. More generally, if G is an almost simple group with socle T (so that $T \leqslant G \leqslant \mathrm{Aut}(T)$ with T a non-abelian finite simple group) then $d(G) = \max\{2, d(G/T)\} \leq 3$ (see [14]).

A wide range of related problems on the generation of finite simple groups has been investigated in recent years. For instance, we may consider the abundance of generating pairs: if we pick two elements of a finite simple group G at random, what is the probability that they generate G? In 1969 Dixon [15] proved that if $G = A_n$ then this probability tends to 1 as $n \to \infty$, confirming an 1882 conjecture of Netto [48]. This was extended in [27, 37] to all finite simple groups, as conjectured by Dixon in [15].

Various generalisations have subsequently been studied by imposing restrictions on the orders of the generating pairs. Here there are some interesting special cases. For example, the simple groups that can be generated by a pair of elements of order 2 and 3 coincide with the simple quotients of the modular group $\mathrm{PSL}_2(\mathbb{Z}) \cong Z_2 \star Z_3$, and they have been intensively studied in recent years (see [39, 41], and also [40, 53] for related results). In a different direction, in [21] it is proved that every non-trivial element of a finite simple group belongs to a pair of generating elements, confirming a conjecture of Steinberg [54]. A more general notion of *spread* for 2-generator groups was introduced by Brenner and Wiegold [8], and this has been widely studied in the context of finite almost simple groups (see [10, 9, 22], for example).

Date: October 9, 2013.

Key words and phrases. Finite simple groups; maximal subgroups; minimal generation; primitive permutation groups.

Corresponding author: Dr. T.C. Burness.

Our understanding of the subgroup structure of the finite simple groups has advanced greatly in the last 30 years or so (see [30, 31, 36] for an overview). Indeed, almost all of the above results require detailed information on the maximal subgroups of simple groups. The main purpose of this paper is to investigate various generation properties of the maximal subgroups themselves, establishing some new and rather unexpected results. Our aim is to show that some of the above results for simple groups can be extended, with some suitable small (and necessary) modifications, to all their maximal subgroups. For example, just as every finite simple group is 2-generator, our main result states that any maximal subgroup H can also be generated by very few elements.

Theorem 1. *Every maximal subgroup of a finite simple group is* 4*-generator.*

There are infinitely many examples with G simple and $d(H) = 4$ (see Remarks 4.5 and 5.12, for example), so Theorem 1 is best possible. In fact this theorem follows from a more general result, stated below, dealing also with maximal subgroups of almost simple groups.

Theorem 2. *Let G be a finite almost simple group with socle G_0 and let H be a maximal subgroup of G. Then $d(H \cap G_0) \leq 4$, and also $d(H) \leq 6$.*

It is likely that 4 is also the optimal bound in the more general almost simple situation.

In view of the explicit bounds obtained in Theorem 2, it is natural to investigate the probabilistic generation of maximal subgroups of simple and almost simple groups, in analogy with the aforementioned work on the simple groups themselves.

We introduce some relevant background and notation. For a finite or profinite group G and a positive integer k let $P(G, k)$ denote the probability that k randomly chosen elements of G generate G (topologically, if G is infinite). A profinite group G is said to be *positively finitely generated* (PFG for short) if $P(G, k) > 0$ for some k. Which finitely generated profinite groups are PFG? Various examples have been given in the past two decades; these include prosolvable groups (Mann [45]), groups satisfying the Babai-Cameron-Pálfy condition [4] on their upper composition factors [6], certain iterated wreath products of simple groups, etc.

A characterization of PFG groups in terms of maximal subgroup growth has been obtained in [46]. Let $m_n(G)$ denote the number of maximal subgroups of index n in G. The main result of [46] states that a profinite group G is PFG if and only if $m_n(G)$ grows polynomially with n. Lubotzky [42] provided effective versions of this for finite groups G. Let $\nu(G)$ be the minimal number k such that $P(G, k) \geq 1/e$. Up to a small multiplicative constant, it is known that $\nu(G)$ is the expected number of random elements generating G (see [50] and [42, 1.1]). Define

$$\mathcal{M}(G) = \max_{n \geq 2} \frac{\log m_n(G)}{\log n}.$$

By [42, 1.2] we have $\mathcal{M}(G) < \nu(G) + 4$ for any finite group G.

Remarkable results characterizing PFG profinite groups have been recently obtained by Jaikin-Zapirain and Pyber [26]. Theorem 1 in that paper provides strong bounds on $\nu(G)$ for G finite. Combining this tool with Theorem 2 above we establish random generation of all maximal subgroups of almost simple groups. More precisely we have:

Theorem 3. *There exists an absolute constant c such that $\nu(H) \leq c$ for any maximal subgroup H of a finite almost simple group.*

More generally, by increasing the constant c in Theorem 3, if necessary, we obtain the following corollary.

Corollary 4. *For any given $\epsilon > 0$ there exists an absolute constant $c = c(\epsilon)$ such that $P(H, c) > 1 - \epsilon$ for any maximal subgroup H of a finite almost simple group.*

This is essentially best possible because the strong random generation property in the aforementioned conjecture of Dixon fails to extend to maximal subgroups of simple groups, so there is no universal constant c such that $P(H, c) \to 1$ as $|H| \to \infty$. For example, the symmetric group $H = S_{n-2}$ is a maximal subgroup of A_n, and $P(H, c) \leq 1 - 2^{-c}$ for all c. More generally, many maximal subgroups H have subgroups of bounded index, preventing $P(H, c)$ from tending to 1 as $|H| \to \infty$ if c is fixed.

The maximal subgroup growth of finite simple groups G has been widely studied, see [27], [37], [38], culminating in [32] where it is shown that $m_n(G) \leq n^a$ for any fixed $a > 1$ and sufficiently large n. Combining Theorem 3 with Lubotzky's bound on $\mathcal{M}(G)$ stated above, we obtain a polynomial upper bound on $m_n(H)$ where H is any maximal subgroup of an almost simple group.

Corollary 5. *There is an absolute constant c such that any maximal subgroup of a finite almost simple group has at most n^c maximal subgroups of index n.*

This yields a surprising corollary on *second maximal* subgroups of almost simple groups G, which are defined to be the maximal subgroups of maximal subgroups of G.

Corollary 6. *There is an absolute constant c such that any finite almost simple group has at most n^c second maximal subgroups of index n.*

It is natural to ask whether or not Theorem 2 can be extended to second maximal subgroups of almost simple groups: is there an absolute constant c such that $d(H) \leq c$ for any second maximal subgroup H? The answer to this question appears to depend on a difficult problem in number theory, namely the existence of infinitely many integers of the form $p^k - 1$ (p a fixed prime) with a prime factor r such that $(p^k - 1)/r = o(k)$. This open problem is far beyond the reach of present methods, which only provide prime factors r of the order of magnitude k^c.

To see the connection, let $G = \mathrm{L}_2(p^k)$ and write $p^k - 1 = rb$ with r an odd prime. Set $d = b/2$ if p is odd, otherwise $d = b$. Then $H = Z_p^k.Z_d$ has index r in a Borel subgroup of G, so H is a second maximal subgroup and it is easy to see that $d(H) > k/d$. In particular, if there are infinitely many integers $p^k - 1$ with a prime divisor r as above with $b = o(k)$, then the corresponding second maximal subgroup H of $\mathrm{L}_2(p^k)$ will require arbitrarily many generators. For example, if $p = 2$ then this follows if there are infinitely many Mersenne primes. Similar examples can also be constructed in other small rank groups of Lie type.

We plan to investigate this further in a future paper on the generation properties of second maximal subgroups of simple and almost simple groups. More generally, we will also study the *t-maximal subgroups* of such groups, where a subgroup H of a group G is t-maximal if there exists a chain of subgroups $H = H_t < H_{t-1} < \cdots < H_1 < H_0 = G$ with H_i maximal in H_{i-1} for all i.

Theorems 2 and 3 also have interesting applications to permutation groups. Recall that a transitive permutation group G on a set Ω with point stabilizer H is *primitive* if there is no non-trivial G-invariant partition of Ω, which is equivalent to the condition that H is a maximal subgroup of G. The finite primitive groups can be viewed as the basic building blocks of all finite permutation groups, and they have been studied extensively since the days of Jordan in the 19th century. A key tool here is the O'Nan-Scott theorem (see [16, Theorem 4.1.A]), which partitions these groups into several classes. This often provides a way to reduce a general question about primitive groups to the almost simple case, where one can appeal to the Classification theorem and the wealth of information on the maximal subgroups of almost simple groups.

Let G be a finite primitive permutation group with point stabilizer H. What is the relationship between $d(G)$ and $d(H)$? Clearly, we have $d(G) \leq d(H)+1$, since H is a maximal subgroup of G. For general finite groups G and a maximal subgroup H, $d(H)$ may be much larger than $d(G)$ – indeed the best upper bound on $d(H)$ is $|G:H|(d(G)-1)+1$. It is somewhat surprising that when the core of H in G is trivial, namely when G acts faithfully on the cosets of H, a much better upper bound holds.

Theorem 7. *Let G be a finite primitive permutation group with point stabilizer H. Then*

$$d(G)-1 \leq d(H) \leq d(G)+4.$$

Thus $d(H)$ and $d(G)$ are very close in this case. Note that there are many examples of primitive groups with $d(G)$ arbitrarily large.

Our final result extends Theorem 3 to arbitrary primitive permutation groups, demonstrating that $\nu(H)$ and $\nu(G)$ are also very closely related.

Theorem 8. *There exist absolute constants $0 < c_1 < c_2$ such that*

$$c_1\nu(G) < \nu(H) < c_2\nu(G)$$

for any finite primitive permutation group G with point stabilizer H.

This is the first paper to systematically study the generation of maximal subgroups of finite simple groups. However, explicit generators of some maximal subgroups of simple classical and sporadic groups are described in [24, 25] and [7, 57], respectively, with a view towards practical applications in computational group theory.

In this paper we adopt the notation of [29] for classical groups, so $\mathrm{L}_n(q) = \mathrm{L}_n^+(q)$, $\mathrm{U}_n(q) = \mathrm{L}_n^-(q)$, $\mathrm{PSp}_n(q)$ and $\mathrm{P\Omega}_n^\epsilon(q)$ denote the simple linear, unitary, symplectic and orthogonal groups of dimension n over the finite field $\mathbb{F}_q$, respectively. In addition, if G is a group and n is a positive integer then we write Z_n (or just n) and D_n for the cyclic and dihedral groups of order n, respectively, $[n]$ denotes an arbitrary solvable group of order n, while $Z(G), \Phi(G)$ and G^n represent the centre of G, the Frattini subgroup of G and the direct product of n copies of G, respectively. Further, (a,b) denotes the greatest common divisor of the positive integers a and b.

Let us make some remarks on the layout of the paper. First, in Section 2 we record some preliminary results that we will need in the proof of Theorem 2. Next, in Sections 3 and 4 we prove Theorem 2 for groups with a sporadic and alternating group socle, respectively. This leaves us to deal with groups of Lie type. In Section 5 we consider the non-parabolic subgroups of classical groups, and we do likewise for the exceptional groups in Section 6. We complete the proof of Theorem 2 in Section 7, where we deal with the parabolic subgroups in groups of Lie type. Theorem 3 and Corollary 4 are proved in Section 8, and the short proof of Corollary 6 is given in Section 9. Finally, Theorems 7 and 8 are proved in Section 10.

Acknowledgments

The first author acknowledges the support of EPSRC grant EP/I019545/1. The third author acknowledges the support of an Advanced ERC Grant, an EPSRC Visiting Fellowship, an Israel Science Foundation Grant, and the Miriam and Julius Vinik Chair in Mathematics which he holds.

2. Preliminaries

Here we record a collection of results which we will need in the proof of Theorem 2. Some of these are new, and may be of independent interest.

Proposition 2.1. *The following hold:*

(i) *If G is a finite almost simple group with socle G_0, then*
$$d(G) = \max\{2, d(G/G_0)\} \leq 3,$$
with equality if and only if $G_0 = \mathrm{L}_{2m}(q)$ $(m \geq 2)$, $\mathrm{P}\Omega^{\epsilon}_{2m}(q)$ $(m \geq 5)$ or $\mathrm{P}\Omega^{+}_{8}(q)$, where $q = q_0^2$ is odd and $Z_2 \times Z_2 \times Z_2$ is an epimorphic image of G/G_0.

(ii) *If G is a finite group and N is a minimal normal subgroup of G, then*
$$d(G) \leq d(G/N) + 1.$$

(iii) *If G is a non-cyclic finite group with unique minimal normal subgroup N, then*
$$d(G) = \max\{2, d(G/N)\}.$$

Proof. Parts (i), (ii) and (iii) are the main theorems of [14], [43] and [44], respectively. □

Remark 2.2. In the proof of Theorem 2 we may (and will) assume that $G = HG_0$ (so H has trivial core). Indeed, if $G \neq HG_0$ then H is almost simple and the bound in (i) above implies that $d(H) \leq 3$.

Proposition 2.3. *Let G be an almost simple group with socle G_0, such that G/G_0 is either trivial or has prime order. Then $d(G \times Z_a) = 2$ for any positive integer a. In particular, $d(S_n \times Z_a) \leq 2$ for all n.*

Proof. By Proposition 2.1(i) we have $d(G) = 2$, say $G = \langle x, y\rangle$ and $Z_a = \langle t\rangle$. First suppose G/G_0 has prime order. Without loss, we may assume that G/G_0 is generated by yG_0. Set $H = \langle (x,t), (y,1)\rangle$. We claim that $H = G \times Z_a$. To see this, it suffices to show that the kernel K of the natural projection map $\pi : H \to Z_a$ is isomorphic to G. Clearly, K is isomorphic to a normal subgroup of G, so $K \in \{1, G_0, G\}$ since G/G_0 has prime order. However, $(y, 1) \in K$ and $y \in G \setminus G_0$, so $K = G$ and we are done. An entirely similar argument applies if $G = G_0$. □

Proposition 2.4. *The following hold:*

(i) *Let G be a finite group and suppose N is a normal subgroup of G. Then*
$$d(G/N) \leq d(G) \leq d(G/N) + d(N).$$
If also $N \leqslant \Phi(G)$ then $d(G) = d(G/N)$.

(ii) *Let G_1, G_2 be groups such that there is no non-trivial homomorphism from G_1 into an image of G_2. Then $d(G_1 \times G_2) = \max\{d(G_1), d(G_2)\}$.*

Proof. Part (i) is obvious. For (ii), let $d = \max\{d(G_1), d(G_2)\}$ and note that $d \leq d(G_1 \times G_2)$. Pick generators h_i for G_1 and k_i for G_2 $(i = 1,\dots,d)$. Set $H = \langle (h_1, k_1), \dots, (h_d, k_d)\rangle$. Let π_i $(i = 1, 2)$ be the canonical projection from $G_1 \times G_2$ to G_i, and let $K_i = H \cap \ker \pi_i$. Then $H/K_i \cong G_i$, and there is a canonical homomorphism from H/K_1 to H/K_1K_2, which is an image of G_2. By hypothesis, this homomorphism is trivial, so $H = K_1K_2$ and thus $H = G_1 \times G_2$ and $d(G_1 \times G_2) \leq d$. □

In the next result, we set $\mathcal{L} = \{\mathrm{SL}_2(2), \mathrm{SL}_2(3), \mathrm{SU}_3(2)\}$.

Proposition 2.5. *Let p be a prime and let $G = L \times T$, where $L = \prod_{i=1}^{k} L_i$ is a direct product of groups L_i of Lie type in characteristic p each of which is either quasisimple or in $\mathcal{L}$, and T is an abelian p'-group. Then the following hold:*

(i) $d(G) = \max\{d(L), d(T)\}$;

(ii) *If the groups $L_i/Z(L_i)$ are pairwise non-isomorphic, and at most one of them is in $\mathcal{L}$, then $d(L) = 2$.*

Proof. Part (i) follows from Proposition 2.4(ii), noting that there is no non-trivial homomorphism from $\mathrm{SL}_2(2), \mathrm{SL}_2(3)$ or $\mathrm{SU}_3(2)$ to an abelian p'-group, where $p = 2, 3, 2$ respectively. Now consider (ii). The hypothesis implies that there is no non-trivial homomorphism from L_i to $\prod_{j \neq i} L_j$, so Proposition 2.4(ii) and induction show that $d(L) = \max_i\{d(L_i)\}$. The result follows, using Proposition 2.1(i) and an easy check that the groups in $\mathcal{L}$ are 2-generator. □

Proposition 2.6. *Let G be a finite group with a normal subgroup $L = \prod_{i=1}^{k} L_i$, a central product of groups L_i each of which is either quasisimple or in $\mathcal{L}$, with at most one group in $\mathcal{L}$ occurring (up to isomorphism).*

(i) *Suppose that for any i, j such that $L_i/Z(L_i) \cong L_j/Z(L_j)$, there exists $g \in G$ such that $L_i^g = L_j$. Then $d(G) \leq d(G/L) + 2$.*

(ii) *If the groups $L_i/Z(L_i)$ are pairwise non-isomorphic then $d(G) \leq d(G/L) + 1$.*

Proof. First consider (i). By Proposition 2.5(ii), with two elements we can generate a product $\prod L_{i_j}$, one factor for each isomorphism type among the groups $L_i/Z(L_i)$. Then $d(G/L)$ further elements generate a group covering G/L, and the transitivity hypothesis implies that these $2 + d(G/L)$ elements generate G.

Now let us turn to (ii). Let $r = d(G/L)$ and pick $x, x_2, \ldots, x_r \in G$ such that

$$G = L\langle x, x_2, \ldots, x_r\rangle.$$

We show that $d(L\langle x\rangle) = 2$. The result will then follow by adding $x_2, \ldots, x_r$ to two generators for $L\langle x\rangle$ to generate G.

By the hypothesis of (ii), conjugation by x fixes each factor L_i of L. Consider a factor L_i which is non-solvable (i.e. does not lie in $\mathcal{L}$). By the main theorem of [21], L_i has a conjugacy class C_i such that for any $g \in L_i \setminus Z(L_i)$, there exists an element of C_i which, together with g, generates L_i. Hence we can find $a_i \in C_i$ and $g_i \in L_i$ such that

$$\langle a_i^{x^{-1}}, a_i^{g_i}\rangle = L_i. \tag{1}$$

By inspection, we can also find such $a_i, g_i \in L_i$ when $L_i \in \mathcal{L}$. Set $a = (a_1, \ldots, a_k)$ and $b = (g_1, \ldots, g_k)x$. We claim that $\langle a, b\rangle = L\langle x\rangle$. To see this, observe first that

$$a^b = (a_1^{g_1 x}, \ldots, a_k^{g_k x}),$$

and hence $\langle a, a^b\rangle$ is a subgroup of L whose projection to each factor L_i contains $\langle a_i, a_i^{g_i x}\rangle$, which by (1) is equal to L_i. Since the groups $L_i/Z(L_i)$ are pairwise non-isomorphic by hypothesis, it follows that $\langle a, a^b\rangle = L$. Hence $\langle a, b\rangle = L\langle x\rangle$, and therefore $d(L\langle x\rangle) = 2$, as required. □

Proposition 2.7. *Let G be a finite group with a normal subgroup $L = L_1 \times L_2$, where L_1 is cyclic and L_2 is quasisimple. Then $d(G) \leq d(G/L) + 1$.*

Proof. We need to show that $d(L\langle x\rangle) = 2$ for $x \in G \setminus L$. Since L_1 is cyclic, there exists $a_1 \in L_1$ such that $L_1 = \langle a_1\rangle$, and using the main theorem of [21] we observe that there exist $a_2, g_2 \in L_2$ with $L_2 = \langle a_2^{x^{-1}}, a_2^{g_2}\rangle$. Set $a = (a_1, a_2), b = (1, g_2)x \in L\langle x\rangle$. It suffices to show that $K = \langle a, a^b\rangle = L$. Let π_i be the projection map from K to L_i. Since $a^b = (a_1^x, a_2^{g_2 x})$, it follows that π_i is onto, so $L_1/K \cap L_1 \cong L_2/K \cap L_2$. Then $K \cap L_2 = L_2$ is the only possibility, so $K \cap L_1 = L_1$ and thus $K = L$ as claimed. □

Proposition 2.8. *Let G_1 and G_2 be almost simple groups, with respective socles L_1 and L_2 such that G_1/L_1 and G_2/L_2 are cyclic. Then $d(G_1 \times G_2) = 2$.*

Proof. By Proposition 2.1(i), we have $d(G_i) = 2$, say $G_i = \langle a_i, b_i \rangle$ with $G_1/L_1 = \langle a_1 L_1 \rangle$ and $G_2/L_2 = \langle b_2 L_2 \rangle$. By applying [44, Result 1], we may assume $b_1 \in L_1$ and $a_2 \in L_2$. Let $a = (a_1, a_2)$, $b = (b_1, b_2)$ and set $K = \langle a, b \rangle$.

Let $\pi_i : K \to G_i$ be the i-th projection map and observe that each π_i is onto, so

$$G_1/K \cap G_1 \cong G_2/K \cap G_2. \tag{2}$$

Let $T = K \cap G_1$. Since G_1 is almost simple, one of the following holds:

(i) $T = G_1$; (ii) T contains L_1 but not G_1; (iii) T is trivial.

If (i) holds then $G_1 \leqslant K$ and (2) implies that $K \cap G_2 = G_2$, so $G_2 \leqslant K$ and thus $G_1 \times G_2 = K$ is 2-generator. Next consider (ii). Here G_1/T is cyclic, so (2) implies that $G_2/K \cap G_2$ is cyclic and thus $K \cap G_2$ contains L_2. In particular, K contains $L_1 \times L_2$. By construction, we have $(b_1, b_2) \in K$ and also $(b_1, 1) \in K$ since we chose $b_1 \in L_1$. Therefore $(1, b_2) \in K$, so $K \cap G_2$ contains $\langle L_2, b_2 \rangle = G_2$, which is a contradiction since $T \neq G_1$.

Finally, suppose (iii) holds. By (2), $G_2/K \cap G_2$ is almost simple so $K \cap G_2$ is trivial and thus $G_1 \cong G_2$. If G_1 is simple then $d(G_1 \times G_2) = 2$ (by Proposition 2.10 below, for example), so we may assume $G_1/L_1 = \langle a_1 L_1 \rangle$ is non-trivial and thus $a_1 \notin L_1$. The map $\phi : G_1 \to G_2$ defined by $\phi(x) = y$, where $y \in G_2$ is the unique element of G_2 with $(x, y) \in K$, is an isomorphism. However, $(a_1, a_2) \in K$ by construction, so $\phi(a_1) = a_2$ which is absurd since $a_1 \notin L_1$ but $a_2 \in L_2$. □

Proposition 2.9. *Let G be a 2-generator group and let H be an index-two subgroup of G. Then $d(H) \leq 3$.*

Proof. Let $G = \langle x, y \rangle$, where $x \in H$ and $y \in G \setminus H$. Set $J = \langle x, y^2, y^{-1}xy \rangle$ and note that $x, y \in N_G(J)$, so J is normal in G and $G/J = \langle yJ \rangle$ has order at most 2. However, $J \leqslant H$ and $|G : H| = 2$, whence $J = H$ is 3-generator. □

Proposition 2.10. *Let G be a finite simple group. Then*

$$h_G := \max\{n \mid d(G^n) = 2\} \geq \frac{k(G)}{|\mathrm{Out}(G)|}$$

where $k(G)$ is the number of non-identity conjugacy classes of G. In particular, $h_G \geq 3$ for all G.

Proof. A formula of Philip Hall [23] states that

$$h_G = \frac{\phi_2(G)}{|\mathrm{Aut}(G)|} \tag{3}$$

where $\phi_2(G)$ denotes the number of ordered pairs (a, b) such that $G = \langle a, b \rangle$. By [21, Corollary], for any $1 \neq g \in G$, there exists $h \in G$ such that $G = \langle g, h \rangle$. Also $G = \langle g, h^c \rangle$ for any $c \in C_G(g)$, and the elements h^c are all distinct since $C_G(g) \cap C_G(h) = 1$. Hence

$$\phi_2(G) \geq \sum_{g \in G^\#} |C_G(g)|$$

where $G^\#$ denotes the set of non-identity elements in G. The right hand side is equal to $k(G)\,|G|$, and the conclusion now follows from (3). In particular, if $G \neq A_5, A_6$ then the bound $h_G \geq 3$ follows immediately. For $G = A_5$ we calculate that $h_G = 19$ via (3), and similarly $h_G = 53$ for $G = A_6$. □

Recall that if G is a group of Lie type defined over a field of characteristic p then an element $x \in G$ is *semisimple* (respectively *unipotent*) if the order of x is coprime to p (respectively a power of p).

Proposition 2.11. *Let G be a group of Lie type such that one of the following holds:*

(i) $\mathrm{SL}_n^{\epsilon}(q) \leqslant G \leqslant \mathrm{GL}_n^{\epsilon}(q)$, *where $n \geq 2$ and $G \neq \mathrm{SU}_3(2)$;*

(ii) $G = \mathrm{Sp}_n(q)$*;*

(iii) $G = \Omega_n^{\epsilon}(q)$, *where $n \geq 3$ and $(n,q,\epsilon) \neq (4,2,+)$ or $(4,3,+)$;*

(iv) *G is a simple group of exceptional Lie type.*

Then there exist elements $x,y \in G$ such that $G = \langle x,y \rangle$, where x is semisimple and y is unipotent.

Proof. If G is quasisimple then the main result of [21] provides a semisimple element $s \in G$ with the property that for any non-trivial $y \in G$ there exists $x \in s^G$ with $G = \langle x,y \rangle$. The result follows in this case. Direct calculation deals with the non-quasisimple groups $\mathrm{SL}_2(2)$, $\mathrm{SL}_2(3)$ and $\mathrm{Sp}_4(2)$. (Similarly, it is easy to verify that $\mathrm{SU}_3(2)$ is a genuine exception.)

Next suppose $G = \Omega_4^+(q)$, with $q > 3$. First assume q is even, so $G = \mathrm{SL}_2(q) \times \mathrm{SL}_2(q)$. The cases $q = 4, 8$ can be checked directly, so assume $q \geq 16$. By [21], we have $\mathrm{SL}_2(q) = \langle a_1, b_1 \rangle = \langle a_2, b_2 \rangle$, where $b_1 = b_2$ are involutions and the a_i are regular semisimple elements of order $q+1$. Since $q \geq 16$, there are at least two distinct $\mathrm{Aut}(\mathrm{SL}_2(q))$-classes of regular semisimple elements of order $q+1$, so without loss we may assume $a_2 \neq f(a_1)$ for all $f \in \mathrm{Aut}(\mathrm{SL}_2(q))$. Set $x = (a_1, a_2)$ and $y = (b_1, b_2)$, so x is semisimple and y is unipotent. Our choice of a_1 and a_2 ensures that $\langle x,y \rangle$ is not a diagonal subgroup of G, so $G = \langle x,y \rangle$. If $q > 3$ is odd then it suffices to show that $\mathrm{P}\Omega_4^+(q) = \mathrm{L}_2(q) \times \mathrm{L}_2(q)$ has the desired generation property, and an entirely similar argument applies.

Finally, suppose $\mathrm{SL}_n^{\epsilon}(q) < G \leqslant \mathrm{GL}_n^{\epsilon}(q)$ and $\{\det(x) \mid x \in G\} = \langle \mu \rangle \leqslant \mathbb{F}^*$, where $\mathbb{F} = \mathbb{F}_q$ if $\epsilon = +$, otherwise $\mathbb{F} = \mathbb{F}_{q^2}$. We may as well assume $G/(Z \cap G)$ is almost simple, where $Z = Z(\mathrm{GL}_n^{\epsilon}(q))$, since the handful of exceptional cases can be checked directly. As before, we have $\mathrm{SL}_n^{\epsilon}(q) = \langle x', y' \rangle$, where x' is semisimple and y' is unipotent. The proof of the main theorem of [21] (see [21, Table II]) indicates that there exists a semisimple element $x \in G$ such that $\det(x) = \mu$ and $x^i = x'$ for some i. Therefore $G = \langle x, y' \rangle$. □

Corollary 2.12. *Let G be a non-abelian finite simple group. Then there exist elements $x,y \in G$ of coprime orders such that $G = \langle x,y \rangle$.*

Proof. For groups of Lie type, this follows immediately from Proposition 2.11, while A_n is generated by the permutations $(1,2)(3,4)$ and $(\alpha, \alpha+1, \ldots, n)$ where $\alpha = 1$ if n is odd, otherwise $\alpha = 2$. Finally, if G is a sporadic group then the result follows from [21, 6.2]. □

In our proof of Theorem 2 we require the following extension of Proposition 2.11 to the special orthogonal group $\mathrm{SO}_4^+(q)$.

Proposition 2.13. *Let $G = \mathrm{SO}_4^+(q)$ with $q \geq 4$. Then there exist elements $x,y \in G$ such that $G = \langle x,y \rangle$, where x is semisimple and y is unipotent.*

Proof. First assume q is even, so $G \cong \mathrm{SL}_2(q) \wr S_2 = (\mathrm{SL}_2(q) \times \mathrm{SL}_2(q))\langle \tau \rangle$, where τ interchanges the two $\mathrm{SL}_2(q)$ factors. If $q \leq 8$ then the result is easily checked via MAGMA [5], so let us assume $q \geq 16$ and write $\mathrm{SL}_2(q) = \langle a_1, b \rangle = \langle a_2, b \rangle$ with $|a_1| = |a_2| = q+1$, $|b| = 2$ and $a_2 \neq f(a_1)$ for all $f \in \mathrm{Aut}(\mathrm{SL}_2(q))$. Set $x = (a_1, a_2)$ and $y = (b,1)\tau$. Then $y^2 = (b,b)$ and we deduce that $\langle x, y^2 \rangle = \mathrm{SL}_2(q) \times \mathrm{SL}_2(q)$ as in the proof of Proposition 2.11. Therefore $G = \langle x,y \rangle$.

Now suppose $q \geq 5$ is odd. It is sufficient to show that $\mathrm{PSO}_4^+(q)$ has the desired property. First note that $\mathrm{PSO}_4^+(q) = \mathrm{L}_2(q)^2\langle \delta \rangle = (L_1 \times L_2)\langle \delta \rangle$, where $\delta = (\delta_1, \delta_2)$ induces a diagonal automorphism on each factor. We may assume $|\delta_1| = q-1$ and $|\delta_2| = q+1$. By considering the subgroup structure of $\mathrm{L}_2(q)$ it is easy to see that if $u \in \mathrm{L}_2(q)$ has order $(q-1)/2$ or $(q+1)/2$ then there exists an element $v \in \mathrm{L}_2(q)$ of order p such that $\mathrm{L}_2(q) = \langle u, v\rangle$. In particular, we can choose p-elements $y_i \in L_i$ such that $L_i = \langle \delta_i^2, y_i\rangle$, so $L_1 \times L_2 = \langle x^2, y\rangle$, where $x = (\delta_1, \delta_2)$ is semisimple and $y = (y_1, y_2)$ is unipotent. Therefore $\mathrm{PSO}_4^+(q) = \langle x, y\rangle$ as required. □

Proposition 2.14. *Suppose $G = O_n^\epsilon(q)$ or $\mathrm{SO}_n^\epsilon(q)$, where $n \geq 2$. Then either $d(G) \leq 2$, or $G = \mathrm{SO}_4^+(3)$ and $d(G) = 3$.*

Proof. If $G/Z(G)$ is almost simple then the result follows from Propositions 2.1(i) and 2.4(i) since $d(G/Z(G)) = 2$ and $Z(G)$ is the Frattini subgroup of G. The case $n = 3$ with $q < 4$ can be checked directly, while $O_2^\epsilon(q) \cong D_{2(q-\epsilon)}$ and $\mathrm{SO}_2^\epsilon(q) \cong Z_{q-\epsilon}.(2, q-1)$. It remains to deal with the case $(n, \epsilon) = (4, +)$. For $G = O_4^+(q)$ we refer the reader to [17, 18], while Proposition 2.13 handles $G = \mathrm{SO}_4^+(q)$ (the case $q = 3$ can be checked directly). □

Proposition 2.15. *Let G be a group such that $\mathrm{P\Omega}_4^+(q) \leqslant G \leqslant \mathrm{PGO}_4^+(q)$. Then either $d(G) = 2$, or $G = \mathrm{PSO}_4^+(3)$ and $d(G) = 3$.*

Proof. In view of Proposition 2.13, we may assume q is odd so G is one of the following:

$$\mathrm{PSO}_4^+(q),\ \mathrm{PO}_4^+(q),\ \mathrm{P\Omega}_4^+(q),\ \mathrm{PGL}_2(q)^2,\ \mathrm{L}_2(q)^2.S_2,\ \mathrm{PGL}_2(q)^2.S_2.$$

The case $q = 3$ can be checked directly, so assume $q \geq 5$. In the first two cases we may apply Proposition 2.14, while Proposition 2.11 give the result in the remaining cases. □

3. Sporadic groups

In this section we establish a strong form of Theorem 2 in the case where G_0 is a sporadic simple group.

Proposition 3.1. *Let G be an almost simple sporadic group with socle G_0 and let H be a maximal subgroup of G. Then $\max\{d(H), d(H \cap G_0)\} \leq 3$.*

Proof. If $G_0 \notin \{\mathrm{HN}, \mathrm{Fi}_{23}, \mathrm{Fi}_{24}', \mathrm{Co}_1, \mathbb{B}, \mathbb{M}\}$ then explicit generators for H are given in the Web-Atlas [57] and the result follows. Next suppose $G_0 \in \{\mathrm{HN}, \mathrm{Fi}_{23}, \mathrm{Fi}_{24}', \mathrm{Co}_1\}$. In each of these cases we use a combination of the information in [57] and direct calculation using Magma with a suitable permutation representation of G. For example, consider Conway's group $G = \mathrm{Co}_1$. Now G has 22 conjugacy classes of maximal subgroups, and for 6 of these subgroups an explicit pair of generators is given in [57]. The remaining possibilities are the following:

(1)	$A_9 \times S_3$	(2)	$(D_{10} \times (A_5 \times A_5).2).2$	(3)	$3^6{:}2.\mathrm{M}_{12}$
(4)	$3^{1+4}{:}\mathrm{Sp}_4(3){:}2$	(5)	$3^{3+4}{:}2.(S_4 \times S_4)$	(6)	$5^{1+2}{:}\mathrm{GL}_2(5)$
(7)	$5^3{:}(4 \times A_5).2$	(8)	$7^2{:}(3 \times 2.S_4)$	(9)	$2^{2+12}{:}(A_8 \times S_3)$
(10)	$2^{4+12}.(S_3 \times 3.S_6)$	(11)	$5^2{:}2.A_5$	(12)	$3^2.\mathrm{U}_4(3).D_8$
(13)	$(A_4 \times G_2(4)){:}2$	(14)	$(A_5 \times \mathrm{J}_2){:}2$	(15)	$(A_7 \times \mathrm{L}_2(7)){:}2$
(16)	$(A_6 \times \mathrm{U}_3(3)){:}2$				

In case (1) it is easy to see that $d(H) = 2$, while Proposition 2.6(ii) gives the same conclusion in cases (13)–(16). To deal with the remaining subgroups we first construct G as a permutation group on 98280 points (see [57]). Consider (2). Here $H = N_G(C_G(z))$, where z is a $5B$-element (see [13]), so we can easily construct H using the explicit class representatives given in the Web-Atlas and we quickly obtain two generators for H by random search. In cases (3)–(10), H contains a suitable Sylow subgroup of G and it is easy to construct H and verify $d(H) = 2$ in the same way. Alternatively, we can use Proposition 2.1 to see that $d(H) = 2$. For example, in (4) H has a unique minimal normal subgroup of order 3, so Proposition 2.1(iii) implies that $d(H) = d(3^4{:}\mathrm{Sp}_4(3){:}2)$. Similarly, 3^4 is the unique minimal normal subgroup of $3^4{:}\mathrm{Sp}_4(3){:}2$, so $d(H) = d(\mathrm{Sp}_4(3){:}2) = 2$ by Proposition 2.1(i). Cases (11) and (12) are entirely similar.

Next suppose $G = \mathbb{B}$ is the Baby Monster. The maximal subgroups H of G are listed in the Web-Atlas; either an explicit pair of generators is given, or H is almost simple and Proposition 2.1(i) yields $d(H) = 2$, or H is one of the following:

(1)	$[2^{35}].(S_5 \times \mathrm{L}_3(2))$	(2)	$(3^2{:}D_8 \times \mathrm{U}_4(3).2^2).2$	(3)	$[3^{11}].(S_4 \times 2S_4)$
(4)	$(S_6 \times \mathrm{L}_3(4){:}2).2$	(5)	$5^3.\mathrm{L}_3(5)$	(6)	$(S_6 \times S_6).4$
(7)	$S_5 \times \mathrm{M}_{22}{:}2$	(8)	$5^2{:}4.S_4 \times S_5$		

In each case, it is easy to construct a faithful permutation representation of H (see the proof of [11, 3.3], for example) and we quickly deduce that $d(H) \leq 3$ by random search.

Finally, let us assume $G = \mathbb{M}$ is the Monster. A complete list of the conjugacy classes of maximal subgroups of G is not presently available; to date, some 44 classes have been identified (see [57] for a convenient list, with the addition of $\mathrm{L}_2(41)$ – see [49]), and it is known that any additional maximal subgroup is almost simple with socle $\mathrm{L}_2(13)$, $\mathrm{U}_3(4)$, $\mathrm{U}_3(8)$ or $\mathrm{Sz}(8)$ (see [49]). In particular, Proposition 2.1(i) reveals that each of these additional possibilities is 2-generator, and of course $d(\mathrm{L}_2(41)) = 2$. If H is a representative of one of the remaining 43 known conjugacy classes of maximal subgroups then an explicit pair of generators for H is given in [57], with the exception of the following cases:

(1)	$2.\mathbb{B}$	(2)	$2^{1+24}.\mathrm{Co}_1$	(3)	$2^{10+16}.\Omega_{10}^{+}(2)$
(4)	$2^{5+10+20}.(S_3 \times \mathrm{L}_5(2))$	(5)	$3^{1+12}.2.\mathrm{Suz}{:}2$		

In (1), $H = 2.\mathbb{B}$ is quasisimple and thus $d(H) = 2$ since $d(\mathbb{B}) = 2$. To deal with the cases labelled (2)–(5) we repeatedly apply Proposition 2.1. For example, if $H = 2^{10+16}.\Omega_{10}^{+}(2)$ then Proposition 2.1(iii) yields

$$d(H) = d(2^{16}.\Omega_{10}^{+}(2)) = d(\Omega_{10}^{+}(2))$$

and thus $d(H) = 2$ by Proposition 2.1(i). In the same way, we deduce that $d(H) = 2$ in each of the other cases. In particular, every maximal subgroup of $\mathbb{M}$ is 2-generator. $\square$

4. Alternating groups

Here we establish Theorem 2 in the case where G_0 is an alternating group. We begin by recalling the O'Nan-Scott theorem.

Theorem 4.1 (O'Nan-Scott). *Let $G = A_n$ or S_n, and let H be a maximal subgroup of G. Then one of the following holds:*

(i) *H is intransitive: $H = (S_k \times S_{n-k}) \cap G$, $1 \leq k < n/2$;*
(ii) *H is affine: $H = \mathrm{AGL}_d(p) \cap G$, $n = p^d$, p prime, $d \geq 1$;*
(iii) *H is imprimitive or wreath-type: $H = (S_k \wr S_t) \cap G$, $n = kt$ or k^t;*
(iv) *H is diagonal: $H = (T^k.(\mathrm{Out}(T) \times S_k)) \cap G$, T non-abelian simple, $n = |T|^{k-1}$;*
(v) *H is almost simple.*

The main result of this section is the following:

Proposition 4.2. *Let G be an almost simple group with socle $G_0 = A_n$, and let H be a maximal subgroup of G. Then $\max\{d(H), d(H \cap G_0)\} \leq 4$, with equality only if H is a diagonal-type subgroup.*

Of course, if H is almost simple then Proposition 2.1(i) gives $\max\{d(H), d(H \cap G_0)\} \leq 3$, so we only need to consider the cases labelled (i)–(iv) in Theorem 4.1. The special case $n = 6$ can be checked directly, so we may assume $G = A_n$ or S_n.

Lemma 4.3. *Proposition 4.2 holds in cases (i), (ii) and (iii) of Theorem 4.1.*

Proof. In view of Proposition 2.9 it suffices to show that $d(L) \leq 2$, where $L = S_k \times S_{n-k}$, $\mathrm{AGL}_d(p)$ or $S_k \wr S_t$ in cases (i), (ii) and (iii) of Theorem 4.1.

First consider $L = S_k \times S_{n-k}$. Set $\alpha = 1$ if $n-k$ is odd, otherwise $\alpha = 2$. Similarly, define $\beta = 1$ if k is odd, $\beta = 2$ otherwise. Set $x = ((1,2), x_2)$ and $y = (y_1, (1,2))$, where $x_2 = (\alpha, \alpha+1, \ldots, n-k)$ and $y_1 = (\beta, \beta+1, \ldots, k)$. Then it is easy to see that $L = \langle x, y\rangle$. For example, if $(\alpha, \beta) = (2,1)$ then

$$y^{k+1} = ((1,\ldots,k),1),\ x^{n-k-1} = ((1,2),1),\ x^{n-k} = (1,(2,\ldots,n-k)),\ y^k = (1,(1,2))$$

and $S_k = \langle (1,\ldots,k),(1,2)\rangle$ and $S_{n-k} = \langle (2,\ldots,n-k),(1,2)\rangle$. If $L = \mathrm{AGL}_d(p)$ is affine then $L = V{:}\mathrm{GL}_d(p)$, where V is an elementary abelian normal subgroup of order p^d. Since V is the unique minimal normal subgroup of L, and $d(\mathrm{GL}_d(p)) \leq 2$, Proposition 2.1(iii) yields $d(L) \leq 2$.

Finally, suppose $L = S_k \wr S_t = B.S_t$. Let $(\rho_1, \ldots, \rho_t; \sigma)$ denote a general element of L, where $\rho_i \in S_k$ and $\sigma \in S_t$. Set $\alpha = 1$ if k is odd, otherwise $\alpha = 2$. If $t = 2$ then it is easy to see that $L = \langle x, y\rangle$, where $x = ((1,2),(\alpha,\ldots,k);1)$ and $y = (1,1;(1,2))$. Next suppose $t \geq 4$ is even. Here $L = \langle x,y\rangle$ where

$$x = ((1,2),1,\ldots,1;(2,\ldots,t)),\ y = (1,1,(\alpha,\ldots,k),1,\ldots,1;(1,2)).$$

For example, if k is odd then

$$x^{t-1} = ((1,2),1,\ldots,1;1),\ x^t = (1,\ldots,1;(2,\ldots,t)),\ y^k = (1,\ldots,1;(1,2))$$

and $y^{k+1} = (1,1,(1,\ldots,k),1,\ldots,1;1)$. Similarly, if $t \geq 5$ is odd then $L = \langle x,y\rangle$ with

$$x = ((\alpha,\ldots,k),1,\ldots,1;(2,\ldots,t)),\ y = (1,1,1,(1,2),1,\ldots,1;(1,2,3)).$$

Finally, let us assume $t = 3$. We claim that $L = \langle x,y\rangle$, where $x = ((\alpha,\ldots,k),1,1;(2,3))$ and $y = ((1,2),1,1;(1,3))$. First suppose k is odd, so $x^k = (1,1,1;(2,3))$ and $x^{k+1} = ((1,\ldots,k),1,1;1)$. Now

$$z_1 = (x^k y)^3 = ((1,2),(1,2),(1,2);1),\ y^2 = ((1,2),1,(1,2);1),$$

hence $z_2, z_3 \in \langle x,y\rangle$, where $z_2 = z_1 y^2 = (1,(1,2),1;1)$ and $z_3 = z_2^{x^k} = (1,1,(1,2);1)$. Now $yz_3 = (1,1,1;(1,3))$, so $\langle x^k, yz_3\rangle \cong S_3$ and we are done since $z_1z_2z_3 = ((1,2),1,1;1) \in \langle x,y\rangle$ and $\langle z_1z_2z_3, x^{k+1}\rangle \cong S_k$. A very similar argument applies when k is even. □

We note that there are examples in Lemma 4.3 where $\max\{d(H), d(H \cap G_0)\} = 3$. For instance, $d((S_4 \times S_3) \cap A_7) = 3$.

Lemma 4.4. *Proposition 4.2 holds in case (iv) of Theorem 4.1.*

Proof. First assume $H = T^k.(\mathrm{Out}(T) \times S_k)$. Here $N = T^k$ is the unique minimal normal subgroup of H, so Proposition 2.1(iii) yields $d(H) = \max\{2, d(H/N)\}$. Using Proposition 2.3 it is straightforward to check that $d(\mathrm{Out}(T) \times S_k) \leq 4$ and the result follows.

Now suppose $G = A_n$ and H is an index-two subgroup of $T^k.(\mathrm{Out}(T) \times S_k)$. First assume $k \geq 3$. If we consider the action of $\sigma = (1,2) \in S_k$ on the set Ω of cosets of the diagonal subgroup $D = \{(t,\ldots,t) \mid t \in T\}$ in T^k then σ fixes precisely $|T|^{k-2}$ points, so σ induces an even permutation on Ω and thus $H = T^k.(J \times S_k)$, where J is an index-two subgroup of $\mathrm{Out}(T)$. As before, T^k is the unique minimal normal subgroup of H, so it suffices to show that $d(J \times S_k) \leq 4$. According to Proposition 2.1(i) we have $d(J) \leq 3$, so we may as well assume $d(J) = 3$ since $d(S_k) = 2$ and $d(J \times S_k) \leq d(J) + d(S_k)$. Set $a_1 = (1,2)$ and $a_2 = (\alpha, \alpha+1, \ldots, k)$, where $\alpha = 1$ if k is odd, otherwise $\alpha = 2$. Then $S_k = \langle a_1, a_2\rangle$ and $|a_2|$ is odd. Write $J = \langle b_1, b_2, b_3\rangle$. If $|b_1|$ is odd then $J \times S_k$ is generated

$\mathcal{C}_1$	Stabilizers of subspaces of V
$\mathcal{C}_2$	Stabilizers of decompositions $V = \bigoplus_{i=1}^{t} V_i$, where $\dim V_i = a$
$\mathcal{C}_3$	Stabilizers of prime index extension fields of $\mathbb{F}_q$
$\mathcal{C}_4$	Stabilizers of decompositions $V = V_1 \otimes V_2$
$\mathcal{C}_5$	Stabilizers of prime index subfields of $\mathbb{F}_q$
$\mathcal{C}_6$	Normalizers of symplectic-type r-groups in absolutely irreducible representations
$\mathcal{C}_7$	Stabilizers of decompositions $V = \bigotimes_{i=1}^{t} V_i$, where $\dim V_i = a$
$\mathcal{C}_8$	Stabilizers of non-degenerate forms on V

TABLE 1. The $\mathcal{C}_i$ families

by the elements (b_1, a_1), $(b_2, 1)$, $(b_3, 1)$ and $(1, a_2)$, otherwise (b_1, a_2), $(b_2, 1)$, $(b_3, 1)$ and $(1, a_1)$ do the job. We conclude that $d(J \times S_k) \leq 4$ and thus $d(H) \leq 4$.

Now suppose $k = 2$. Here σ fixes a coset $D(t_1, t_2)$ if and only if $t_2 = t_1 t$ with $t^2 = 1$. Therefore σ has precisely $i_2(T) + 1$ fixed points on Ω, where $i_2(T)$ is the number of involutions in T, whence the number ℓ of 2-cycles of σ on Ω is given by the formula $\ell = \frac{1}{2}(|T| - i_2(T) - 1)$. Consequently, if ℓ is odd then $H \cong T^2.\mathrm{Out}(T)$ and thus $d(H) = \max\{2, d(\mathrm{Out}(T))\} \leq 3$. On the other hand, if ℓ is even then $H = T^2.(J \times S_2)$, where J is an index-two subgroup of $\mathrm{Out}(T)$. As before we get $d(H) = \max\{2, d(J \times S_2)\} \leq 4$. □

Remark 4.5. In case (iv) of Theorem 4.1 there are infinitely many examples with $d(H) = 4$. For example, suppose $T = \mathrm{P}\Omega_{2m}^{+}(p^{2f})$, where $m \geq 6$ is even and p is an odd prime. By [33], $H = (T \times T).(\mathrm{Out}(T) \times Z_2) \leqslant S_{|T|}$ is a maximal subgroup of $A_{|T|}H$, where $\mathrm{Out}(T) \cong D_8 \times Z_{2f}$. Visibly, $Z_2 \times Z_2 \times Z_2$ is an epimorphic image of $\mathrm{Out}(T)$, so the elementary abelian group of order 16 is an image of H and thus $d(H) \geq 4$. We conclude that $d(H) = 4$. In fact, if $m = 6$ then $H \leqslant A_{|T|}$, so in this way we obtain an infinite family of pairs (G, H) where G is simple and H is a maximal subgroup with $d(H) = 4$, demonstrating the sharpness of the bound on $d(H \cap G_0)$ in Theorem 2. To see that $H \leqslant A_{|T|}$ it is sufficient to show that the maps $\iota, \phi_a : T \to T$, defined by $\iota(t) = t^{-1}$ and $\phi_a(t) = t^a$, are even permutations for all involutions $a \in \mathrm{Aut}(T)$. Now $|T|$ is divisible by 4, and the information in [20, Table 4.5.1] reveals that $|\{t \in T \mid t = t^{-1}\}|$ and $|C_T(a)|$ are also divisible by 4 for all involutions $a \in \mathrm{Aut}(T)$, whence ι and ϕ_a are even permutations and thus $H \leqslant A_{|T|}$ as claimed.

5. Classical groups

In this section we prove Theorem 2 for non-parabolic subgroups of classical groups. Let G be an almost simple classical group over $\mathbb{F}_q$ with socle G_0 and natural module V, where $q = p^f$ and p is a prime. The main theorem on the subgroup structure of classical groups is due to Aschbacher. In [1], eight collections of subgroups of G are defined, labelled $\mathcal{C}_i$ for $1 \leq i \leq 8$, and it is shown that if H is a maximal subgroup of G then either H is contained in one of these natural subgroup collections, or it belongs to a family of almost simple subgroups that act irreducibly on V (we use $\mathcal{S}$ to denote this additional subgroup collection). Table 1 provides a rough description of the $\mathcal{C}_i$ families. We refer the reader to [29] for a detailed description of these subgroup collections, and we adopt the notation therein. We also note that a small additional collection of maximal subgroups arises when $G_0 = \mathrm{P}\Omega_8^{+}(q)$ or $\mathrm{Sp}_4(q)'$ (q even), due to the existence of exceptional automorphisms in these cases (see Section 5.4).

It is convenient to postpone the analysis of parabolic subgroups to Section 7, where we also deal with parabolic subgroups of exceptional groups. Throughout this section we set

$$H_0 = H \cap G_0, \quad \tilde{G} = G \cap \mathrm{PGL}(V), \quad \tilde{H} = H \cap \mathrm{PGL}(V).$$

Proposition 5.1. *Theorem 2 holds if $H \in \mathcal{C}_3 \cup \mathcal{C}_5 \cup \mathcal{C}_6 \cup \mathcal{C}_8 \cup \mathcal{S}$.*

Proof. Since $d(G/G_0) \leq 3$ (see Proposition 2.1(i)) it suffices to show that $d(H_0) \leq 3$. This is clear if $H \in \mathcal{S}$, so assume H belongs to one of the relevant $\mathcal{C}_i$ families. Suppose $i \neq 6$. According to [29], in almost all cases H_0 has the form $Z_a.A$, where A is a 2-generator almost simple group, whence $d(H_0) \leq 3$. The few remaining cases are easily dealt with. For example, if $G_0 = \mathrm{U}_4(q)$, q is odd and H is a $\mathcal{C}_5$-subgroup of type $O_4^+(q)$ then [29, 4.5.5] gives $H_0 = \mathrm{PSO}_4^+(q).2 < \mathrm{PGO}_4^+(q)$, so $d(H_0) = 2$ by Proposition 2.15. Finally, if $H \in \mathcal{C}_6$ then [29, §4.6] indicates that either $H_0 = N.A$, where N is a minimal normal subgroup of H_0 and A is 2-generator, or $H_0 = A_4$ or S_4. In the latter situation we have $d(H_0) = 2$, while Proposition 2.1(ii) yields $d(H_0) \leq 3$ in the general case. □

5.1. **Non-parabolic, reducible subgroups.** Here we deal with the non-parabolic subgroups in Aschbacher's $\mathcal{C}_1$ family; the relevant cases are listed in [29, Table 4.1.A].

Lemma 5.2. *Theorem 2 holds if $G_0 = \mathrm{P\Omega}_n^{\epsilon}(q)$ and H is of type $O_m^{\epsilon_1}(q) \perp O_{n-m}^{\epsilon_2}(q)$.*

Proof. Here $1 \leq m \leq n/2$ and $(m, \epsilon_1) \neq (n-m, \epsilon_2)$. According to [29, 4.1.6] we have

$$H_0 \in \{\Omega_{n-1}(q), (\Omega_m^{\epsilon_1}(q) \times \Omega_{n-m}^{\epsilon_2}(q)).[2^i], (\Omega_m^{\epsilon_1}(q) \circ \Omega_{n-m}^{\epsilon_2}(q)).[4]\},$$

where $i = 1$ or 2, and we may assume $(n-m, \epsilon_2) \neq (4, +)$. In particular, if $(m, \epsilon_1) \neq (4, +)$ then Propositions 2.1(i), 2.6(ii) and 2.7 yield $d(H_0) \leq 3$.

Now assume $(m, \epsilon_1) = (4, +)$. If $q = 2$ then $H_0 = (\Omega_4^+(2) \times \Omega_{n-4}^{\epsilon_2}(2)).2$ and Proposition 2.11(iii) implies that $\Omega_{n-4}^{\epsilon_2}(2) = \langle x', y' \rangle$, with x' semisimple and y' unipotent. Now $\Omega_4^+(2) = \langle x, y \rangle$ with $|x| = 2$ and $|y| = 6$, so

$$\Omega_4^+(2) \times \Omega_{n-4}^{\epsilon_2}(2) = \langle (x, x'), (y, 1), (1, y') \rangle$$

and thus $d(H_0) \leq 4$. Similarly, if $q = 3$ then $\Omega_4^+(3) = \langle x, y \rangle$ with $|x| = |y| = 3$, and $\Omega_{n-4}^{\epsilon_2}(3) = \langle x', y' \rangle$, with x', y' semisimple (this follows from the proof of the main theorem of [21]). Therefore $d(H_0) \leq 4$ since $\Omega_4^+(3) \times \Omega_{n-4}^{\epsilon_2}(3)$ is generated by (x, x') and (y, y'). Finally, if $q \geq 4$ then Proposition 2.11(iii) gives $\Omega_4^+(q) = \langle x, y \rangle$ and $\Omega_{n-4}^{\epsilon_2}(q) = \langle x', y' \rangle$ with x, x' semisimple and y, y' unipotent, so $d(\Omega_4^+(q) \circ \Omega_{n-4}^{\epsilon_2}(q)) = 2$ and thus $d(H_0) \leq 4$.

It remains to prove that $d(H) \leq 6$ when $d(G/G_0) = 3$. Here n is even, $\epsilon = +$ and $q = q_0^2$ is odd. Moreover, $\tilde{G}/G_0 = D_8$ or $Z_2 \times Z_2$, and it suffices to show that $d(\tilde{H}) \leq 5$. We quickly reduce to the case $H_0 = (\Omega_4^+(q) \circ \Omega_{n-4}^+(q)).[4]$. If $\tilde{G}/G_0 = D_8$ then $\tilde{H} = (O_4^+(q) \circ O_{n-4}^+(q)).2$ and thus $d(\tilde{H}) \leq 5$ by Proposition 2.14. Now assume $\tilde{G}/G_0 = Z_2 \times Z_2$, so

$$\tilde{H} = (\Omega_4^+(q) \circ \Omega_{n-4}^+(q)).[2^4] = (\mathrm{SO}_4^+(q) \circ \mathrm{SO}_{n-4}^+(q)).[2^2].$$

Using Propositions 2.11(iii) and 2.13 we may write $\mathrm{SO}_4^+(q) = \langle x_1, y_1 \rangle$ and $\mathrm{SO}_{n-4}^+(q) = \langle x_2, y_2, z \rangle$, where the x_i are semisimple and the y_i are unipotent. Then $\mathrm{SO}_4^+(q) \times \mathrm{SO}_{n-4}^+(q)$ is generated by (x_1, y_2), (y_1, x_2) and $(1, z)$, so $d(\tilde{H}) \leq 5$ as required. □

Lemma 5.3. *Theorem 2 holds in the remaining non-parabolic $\mathcal{C}_1$ cases.*

Proof. Suppose $G_0 = \mathrm{L}_n^{\epsilon}(q)$ and H is of type $\mathrm{GL}_m^{\epsilon}(q) \perp \mathrm{GL}_{n-m}^{\epsilon}(q)$. By [29, 4.1.4] we have $H_0 = (\mathrm{SL}_m^{\epsilon}(q) \circ \mathrm{SL}_{n-m}^{\epsilon}(q)).A$ with $A \leqslant Z_{q-\epsilon} \times Z_{q-\epsilon}$, whence $d(H_0) \leq 1 + d(A) \leq 3$ via Propositions 2.6(ii) and 2.7. The other cases are very similar. □

5.2. **Imprimitive subgroups.** The members of Aschbacher's $\mathcal{C}_2$ family are the stabilizers of certain subspace decompositions of the natural G_0-module V,

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_t,$$

where $t \geq 2$, $\dim V_i = a$ for all i, and each V_i is either totally singular, or non-degenerate with V_i orthogonal to V_j for $i \neq j$. The relevant subgroups are listed in [29, Table 4.2.A].

Lemma 5.4. *Theorem 2 holds if $G_0 = \mathrm{L}_n^\epsilon(q)$ and $H \in \mathcal{C}_2$ is of type $\mathrm{GL}_a^\epsilon(q) \wr S_t$.*

Proof. Write $\mathrm{GL}_n^\epsilon(q) = \mathrm{SL}_n^\epsilon(q)\langle\delta\rangle$, and suppose $G \cap \mathrm{PGL}(V)$ lifts to $\mathrm{SL}_n^\epsilon(q)\langle\delta^i\rangle$ for some $i \geq 1$. According to [29, 4.2.9], H lifts to $\hat{H} = \hat{A}.B$, where

$$\hat{A} = \mathrm{SL}_a^\epsilon(q)^t.(q-\epsilon)^{t-1}.Z_{(q-\epsilon)/i}.S_t \leqslant \mathrm{GL}_a^\epsilon(q)^t.S_t$$

and $B = Z_b \times Z_c$ (respectively Z_{bc}) if $\epsilon = +$ (respectively $\epsilon = -$), with $b \in \{1,2\}$ and c a divisor of $\log_p q$. Set $\alpha = 0$ if $G = G_0$, otherwise $\alpha = 1$. Note that B is trivial if $\alpha = 0$. In a slight abuse of notation we also write $\mathrm{GL}_a^\epsilon(q) = \mathrm{SL}_a^\epsilon(q)\langle\delta\rangle$.

If $a = 1$ then $d(H) \leq 4 + \alpha$ since $\hat{H}$ is generated by $(\delta, \delta^{-1}, 1, \ldots, 1)$ and $(\delta^i, 1, \ldots, 1)$, together with at most $2+\alpha$ generators for $S_t \times B$. Now assume $a \geq 2$. If $(a,q,\epsilon) \neq (3,2,-)$ then Proposition 2.11(i) gives $\mathrm{SL}_a^\epsilon(q)\langle\delta^i\rangle = \langle x', y'\rangle$ with x' semisimple and y' unipotent, so $\hat{H}$ is generated by $(x', y', 1, \ldots, 1)$ and $(\delta, \delta^{-1}, 1, \ldots, 1)$, plus at most $2 + \alpha$ generators for $S_t \times B$. Finally suppose $(a,q,\epsilon) = (3,2,-)$. Here $d(G/G_0) \leq 2$ so it suffices to show that $d(H_0) \leq 4$. If $t = 2$ then $G_0 = \mathrm{U}_6(2)$ and direct calculation yields $d(H_0) = 2$ so let us assume $t \geq 3$. Write $\mathrm{SU}_3(2) = \langle x, y\rangle$, where $|x| = 4$ and $|y| = 12$, and note that $|\delta| = 3$. Then $\hat{H}$ is generated by $(x, \delta, \delta^{-1}, 1, \ldots, 1)$, $(y, 1, \ldots, 1)$, plus two more for S_t, hence $d(H_0) \leq 4$ as required. $\square$

Lemma 5.5. *Theorem 2 holds if $G_0 = \mathrm{P\Omega}_n^\epsilon(q)$ and $H \in \mathcal{C}_2$ is of type $O_a(q) \wr S_t$.*

Proof. Here aq is odd. If $a = 1$ then $q = p$ (see [29, Table 4.2.A]) and $H = 2^{n-\alpha}.A$, where $\alpha \in \{1,2\}$ and $A = S_n$ or A_n (see [29, 4.2.15]). Since $2^{n-\alpha}$ is a minimal normal subgroup of H, Proposition 2.1(ii) yields $d(H) \leq d(A) + 1 = 3$.

Now assume $a \geq 3$. Since $d(\tilde{G}/G_0) \leq 2$ it suffices to prove that $d(H) \leq 4$ when $\tilde{G} = G_0$. First suppose t is odd, so n is also odd. Write $\Omega_a(q) = \langle x, y\rangle$, where x is semisimple and y is unipotent (see Proposition 2.11(iii)), and let $\rho \in \mathrm{SO}_a(q)$ be an involution such that $\mathrm{SO}_a(q) = \Omega_a(q)\langle\rho\rangle$. If $\tilde{G} = G_0$ then $d(H) \leq 4$ since H is generated by $(x, y, 1, \ldots, 1)$, $(\rho, -\rho, -1, 1, \ldots, 1)$, together with two generators for $S_t \times Z_b$.

Finally, suppose $a \geq 3$ and t is even. Here H lifts to $\hat{H} = A.(S_t \times Z_b)$, where

$$A \in \{2^{t-1} \times \Omega_a(q)^t.2^{t-1},\ 2^t \times \Omega_a(q)^t.2^{t-1},\ 2^{t-1} \times \mathrm{SO}_a(q)^t,\ 2^t \times \mathrm{SO}_a(q)^t\}$$

and b divides $\log_p q$. If $\tilde{G} = G_0$ then $A = 2^{t-1} \times \Omega_a(q)^t.2^{t-1}$ and for $t \geq 4$ we observe that $\hat{H}$ is generated by $(x, y, 1, \ldots, 1)$ and $(\rho, -\rho, -1, 1, \ldots, 1)$, together with two generators for $S_t \times Z_b$. Similarly, if $t = 2$ then $d(H_0) \leq 4$ since H_0 is generated by (x, y), $(-1, -1)$, (ρ, ρ) and one more for S_2. The general $t = 2$ case is very similar. For example, if $A = 2^2 \times \Omega_a(q)^2.2$ then H is generated by (x, y), $(-1, 1)$ and (ρ, ρ), plus at most two additional generators for $S_2 \times Z_b$. $\square$

Lemma 5.6. *Theorem 2 holds if $G_0 = \mathrm{P\Omega}_n^\epsilon(q)$ and $H \in \mathcal{C}_2$ is of type $O_a^{\epsilon'}(q) \wr S_t$.*

Proof. Here a is even and $\epsilon = (\epsilon')^t$. First assume q is even, so $H_0 = \Omega_a^{\epsilon'}(q)^t.2^{t-1}.S_t$ (see [29, 4.2.11]). Write $O_a^{\epsilon'}(q) = \Omega_a^{\epsilon'}(q)\langle\rho\rangle$. If $a = 2$ then $\Omega_a^{\epsilon'}(q) = \langle z\rangle$ is cyclic and H_0 is generated by $(z, 1, \ldots, 1)$, $(\rho, \rho, 1, \ldots, 1)$ and two more for S_t. On the other hand, if $a \geq 4$ then Proposition 2.11(iii) implies that $\Omega_a^{\epsilon'}(q) = \langle x, y\rangle$ with x semisimple and y unipotent (note that H is non-maximal if $(a,q,\epsilon') = (4,2,+)$ – see [29, Table 3.5.H]), so H_0 is generated by $(x,y,1,\ldots,1)$, $(\rho,\rho,1,\ldots,1)$ and two more for S_t. In general, $d(H) \leq 6$ since $d(G/G_0) \leq 2$.

Now assume q is odd. Let D and D_i denote the discriminants of the defining quadratic forms corresponding to G_0 and $O_a^{\epsilon'}(q)$, respectively (see [29, p.32]). We note that $D_1 = D_i$ for all i, and we write $D = \square$ (respectively $\boxtimes$) if D is a square (respectively non-square) in $\mathbb{F}_q$.

First assume $D = \boxtimes$, so t is odd and $D_i = \boxtimes$ for all i (see [29, 2.5.11(i)]). Write $\mathrm{PO}_a^{\epsilon'}(q) = \Omega_a^{\epsilon'}(q)\langle\rho\rangle$ and observe that $H_0 = (2^{t-1} \times \Omega_a^{\epsilon'}(q)^t.2^{t-1}).S_t$ (see [29, 4.2.11]). If $a \geq 4$ then Proposition 2.11(iii) gives $\Omega_a^{\epsilon'}(q) = \langle x,y\rangle$ with x semisimple and y unipotent (note that $(a,q,\epsilon') \neq (4,3,+)$ since $D_i = \boxtimes$), so H_0 is generated by $(x,y,1,\ldots,1)$, $(\rho,-\rho,-1,1,\ldots,1)$ and two more for S_t. Therefore $d(H_0) \leq 4$ and thus $d(H) \leq 6$ since $d(G/G_0) \leq 2$. Similarly, if $a = 2$ then $\Omega_a^{\epsilon'}(q) = \langle z\rangle$ and we quickly obtain $d(H_0) \leq 4$.

Next suppose $D = \square$ and $D_i = \boxtimes$. Here t is even and H_0 lifts to $(2^{t-1} \times \Omega_a^{\epsilon'}(q)^t.2^{t-1}).S_t$. In particular, if $t \geq 4$ and $\tilde{G} = G_0$ then the analysis of the previous paragraph implies that $d(H) \leq 4$, so for any suitable G we deduce that H is 6-generator since $d(\tilde{G}/G_0) \leq 2$. Similarly, if $t = 2$ and $a \geq 4$ then H_0 is generated by (x,y), (ρ,ρ) and one more for S_2, whence $d(H_0) \leq 3$ and thus $d(H) \leq 6$ since $d(G/G_0) \leq 3$.

Finally suppose $D = D_i = \square$, so H_0 lifts to $\Omega_a^{\epsilon'}(q)^t.2^{2(t-1)}.S_t$. Write $\mathrm{SO}_a^{\epsilon'}(q) = \Omega_a^{\epsilon'}(q)\langle s\rangle$ and $O_a^{\epsilon'}(q) = \mathrm{SO}_a^{\epsilon'}(q)\langle r\rangle$. First assume $t = 2$, so $a \geq 4$ since we may assume $n \geq 8$. If $(a,q,\epsilon') = (4,3,+)$ then $G_0 = \mathrm{P}\Omega_8^+(3)$ and the desired result can be checked directly, otherwise H_0 is generated by (x,y), (r,r), (s,s) and one more for S_2, where x and y are defined as before. To get the general bound in the $t = 2$ case we may assume $d(G/G_0) = 3$, so $\epsilon = +$ and $\mathrm{PSO}_n^+(q) < \tilde{G}$, hence H is generated by (x,y), (r,r), $(s,1)$ and at most three more elements. Now assume $t \geq 3$. If $\tilde{G} = G_0$ and $a = 2$ then H is generated by $(z,1,\ldots,1)$, $(r,rs,s,1,\ldots,1)$ and two more for $S_t \times Z_b$; the case $a \geq 4$ with $(a,q,\epsilon') \neq (4,3,+)$ is very similar. Finally, suppose $t \geq 3$ and $(a,q,\epsilon') = (4,3,+)$. Write $\Omega_4^+(3) = \langle x',y'\rangle$ with $|x'| = |y'| = 3$. Then H_0 is generated by the elements

$$((x,1,\ldots,1);(2,3)),\ ((y,1,\ldots,1);1),\ ((r,sr,s,1,\ldots,1);1),\ ((1,\ldots,1);(1,\ldots,t)),$$

so $d(H_0) \leq 4$ and thus $d(H) \leq 6$ since $d(G/G_0) \leq 2$. $\square$

Lemma 5.7. *Theorem 2 holds in the remaining $\mathcal{C}_2$ cases.*

Proof. Consider the case $G_0 = \mathrm{P}\Omega_n^\epsilon(q)$ with H of type $O_{n/2}(q)^2$, where $qn/2$ is odd. According to [29, 4.2.16], $H = A.Z_b$ where b divides $\log_p q$ and

$$A \in \{\mathrm{SO}_{n/2}(q)^2,\ (\mathrm{SO}_{n/2}(q) \times \mathrm{SO}_{n/2}(q)).2,\ O_{n/2}(q) \circ O_{n/2}(q),\ (O_{n/2}(q) \circ O_{n/2}(q)).2\}.$$

Since $d(\mathrm{SO}_{n/2}(q)) = d(O_{n/2}(q)) = 2$ (see Proposition 2.14) we deduce that $H_0 = \mathrm{SO}_{n/2}(q)^2$ is 4-generator and $d(H) \leq 6$ in general. The remaining cases are similar. For example, if $G_0 = \mathrm{PSp}_n(q)$ and H is of type $\mathrm{GL}_{n/2}(q).2$ (with q odd) then $H_0 = Z_{(q-1)/2}.\mathrm{PGL}_{n/2}(q).2$ is 3-generator and the result follows. Similarly, if $G_0 = \mathrm{U}_n(q)$ and H is of type $\mathrm{GL}_{n/2}(q^2).2$ (with $n \geq 6$) then $d(H) \leq 4$ since $H = Z_a.A$, where a divides $q-1$ and A is an almost simple group with socle $\mathrm{L}_{n/2}(q^2)$. $\square$

5.3. **Tensor product subgroups.** Next we consider the tensor product subgroups which comprise Aschbacher's $\mathcal{C}_4$ and $\mathcal{C}_7$ collections. The members of $\mathcal{C}_4$ are the normalizers of tensor decompositions $V = V_1 \otimes V_2$ of the natural G_0-module, where V_1 and V_2 are not similar (see [29, Table 4.4.A]), while the subgroups in $\mathcal{C}_7$ are the normalizers of tensor decompositions of the form

$$V = V_1 \otimes V_2 \otimes \cdots \otimes V_t,$$

where $t \geq 2$ and the V_i are similar for all i. These subgroups are listed in [29, Table 4.7.A].

Lemma 5.8. *Theorem 2 holds if $G_0 = \mathrm{PSp}_n(q)$ and $H \in \mathcal{C}_4$ is of type $\mathrm{Sp}_{n_1}(q) \otimes O^{\epsilon}_{n_2}(q)$.*

Proof. Here q is odd and $n_2 \geq 3$. Since $d(G/G_0) \leq 2$, it suffices to show that $d(H_0) \leq 4$. If n_2 is odd then $H_0 = \mathrm{PSp}_{n_1}(q) \times \mathrm{PO}_{n_2}(q)$ is clearly 4-generator, so let us assume $n_2 \geq 4$ is even, in which case

$$H_0 = (\mathrm{PSp}_{n_1}(q) \times \mathrm{PO}^{\epsilon}_{n_2}(q)).2 = (\mathrm{PSp}_{n_1}(q) \times \mathrm{P\Omega}^{\epsilon}_{n_2}(q)).[2^i],$$

where $i = 2$ or 3 (see [29, 4.4.11]). If $(n_2, \epsilon) \neq (4, +)$ then Proposition 2.6(ii) implies that $d(H_0) \leq 4$, so assume $(n_2, \epsilon) = (4, +)$. If $(n_1, q) = (2, 3)$ then $G_0 = \mathrm{PSp}_8(3)$ and it is easy to check that $d(\mathrm{PSp}_2(3) \times \mathrm{PO}^{+}_4(3)) = 2$ and thus $d(H_0) \leq 3$. If $n_1 = 2$ and $q \geq 5$ then $H_0 = \mathrm{L}_2(q)^3.D_8$ is 4-generator since $d(\mathrm{L}_2(q)^3) = 2$ by Proposition 2.10.

Finally, suppose $(n_2, \epsilon) = (4, +)$ and $n_1 \geq 4$. First assume $q = 3$. Write $\mathrm{PSp}_{n_1}(3) = \langle x_1, y_1 \rangle$ and $\mathrm{PO}^{+}_4(3) = \langle x_2, y_2 \rangle$, where $|x_1| = 5$, $|x_2| = 2$ and $|y_2| = 6$ (such a generating set for $\mathrm{PSp}_{n_1}(3)$ exists by the main theorem of [21]). Then $\mathrm{PSp}_{n_1}(3) \times \mathrm{PO}^{+}_4(3)$ is generated by (x_1, x_2), $(y_1, 1)$ and $(1, y_2)$, so $d(H_0) \leq 4$. Finally, if $q \geq 5$ then by Propositions 2.11(ii) and 2.13 we may write $\mathrm{PSp}_{n_1}(q) = \langle x_1, y_1 \rangle$ and $\mathrm{PSO}^{+}_4(q) = \langle x_2, y_2 \rangle$, where the x_i are semisimple and the y_i are unipotent. Then $\mathrm{PSp}_{n_1}(q) \times \mathrm{PSO}^{+}_4(q)$ is generated by (x_1, y_2) and (y_1, x_2), whence $d(H_0) \leq 4$. $\square$

Lemma 5.9. *Theorem 2 holds if $G_0 = \mathrm{P\Omega}^{\epsilon}_n(q)$ and $H \in \mathcal{C}_4$ is of type $O^{\epsilon_1}_{n_1}(q) \otimes O^{\epsilon_2}_{n_2}(q)$, where q is odd, $n_i \geq 3$, and $(n_1, \epsilon_1) \neq (n_2, \epsilon_2)$.*

Proof. If n is odd then $3 \leq n_1 < n_2$ and $H_0 = (\Omega_{n_1}(q) \times \Omega_{n_2}(q)).2$ is 2-generator by Proposition 2.6(ii). Similarly, if $n_1 \geq 4$ is even and $n_2 \geq 3$ is odd then $H_0 = \mathrm{P\Omega}^{\epsilon_1}_{n_1}(q) \times \mathrm{SO}_{n_2}(q)$ is 4-generator. In general, if n_1 is even and n_2 is odd then $H = (A \times \mathrm{SO}_{n_2}(q)).Z_a$, where $\mathrm{P\Omega}^{\epsilon_1}_{n_1}(q) \leqslant A \leqslant \mathrm{PGO}^{\epsilon_1}_{n_1}(q)$ and a divides $\log_p q$. If $(n_1, q, \epsilon_1) \neq (4, 3, +)$ then $d(A) = 2$ (see Propositions 2.1(i) and 2.15) and thus $d(H) \leq 5$, otherwise $d(A) \leq 3$ and again we have $d(H) \leq 5$ since $a = 1$.

For the remainder assume n_1 and n_2 are even, so $\epsilon = +$, $n_1, n_2 \geq 4$ and $(n_2, \epsilon_2) \neq (4, +)$. According to [29, 4.4.14–16] we have $H = A.Z_a$, where a divides $\log_p q$ and

$$A = (\mathrm{PSO}^{\epsilon_1}_{n_1}(q) \times \mathrm{PSO}^{\epsilon_2}_{n_2}(q)).[2^i]$$

with $2 \leq i \leq 4$. If $i = 4$ then $d(H) \leq 5$ since $A = \mathrm{PGO}^{\epsilon_1}_{n_1}(q) \times \mathrm{PGO}^{\epsilon_2}_{n_2}(q)$ is 4-generator, therefore we may assume $i \leq 3$ and $d(G/G_0) \leq 2$. Note that $a = 1$ and $i \leq 3$ if $G = G_0$, so it suffices to show that $d(A) \leq 4$. For now, we will assume $(n_1, \epsilon_1) \neq (4, +)$.

If $i = 2$ then $d(A) \leq 4$ since Proposition 2.8 yields $d(\mathrm{PSO}^{\epsilon_1}_{n_1}(q) \times \mathrm{PSO}^{\epsilon_2}_{n_2}(q)) = 2$. Now assume $i = 3$. There are several cases to consider. If both $\mathrm{PSO}^{\epsilon_1}_{n_1}(q)$ and $\mathrm{PSO}^{\epsilon_2}_{n_2}(q)$ are simple then Proposition 2.6(ii) implies that $d(A) \leq 4$, as required. Next suppose neither of these groups are simple, in which case $A = L.[2^5]$ with $L = \mathrm{P\Omega}^{\epsilon_1}_{n_1}(q) \times \mathrm{P\Omega}^{\epsilon_2}_{n_2}(q)$ and $[2^5] < D_8 \times D_8$. Such a subgroup of $D_8 \times D_8$ is either 3-generator, or

$$A \in \{\mathrm{PGO}^{\epsilon_1}_{n_1}(q) \times \mathrm{P\Omega}^{\epsilon_2}_{n_2}(q).2^2, \mathrm{P\Omega}^{\epsilon_1}_{n_1}(q).2^2 \times \mathrm{PGO}^{\epsilon_2}_{n_2}(q)\}.$$

In the former case we get $d(A) \leq 4$ as before, otherwise the same conclusion follows via Proposition 2.1(i). Finally, suppose $\mathrm{PSO}^{\epsilon_1}_{n_1}(q)$ is simple but $\mathrm{PSO}^{\epsilon_2}_{n_2}(q)$ is not. Here $A = L.[2^4]$ with L as before and $[2^4] < D_8 \times 2^2$. The subgroup $[2^4]$ is either 3-generator, or $A = \mathrm{P\Omega}^{\epsilon_1}_{n_1}(q).2^2 \times \mathrm{PGO}^{\epsilon_2}_{n_2}(q)$; in the former case, Proposition 2.6(ii) implies that $d(A) \leq 4$, while in the latter we get $d(A) \leq 4$ by Proposition 2.1(i).

It remains to deal with the case $(n_1, \epsilon_1) = (4, +)$ with n_2 even. Arguing as above, we quickly reduce to the case

$$A = (\mathrm{PSO}^{+}_4(q) \times \mathrm{PSO}^{\epsilon_2}_{n_2}(q)).[2^i] = (\mathrm{L}_2(q) \times \mathrm{L}_2(q) \times \mathrm{P\Omega}^{\epsilon_2}_{n_2}(q)).B$$

with B a 3-generator subgroup of $D_8 \times D_8$. We claim that $d(A) \leq d(B) + 1 \leq 4$.

To see this, set

$$L = \mathrm{L}_2(q) \times \mathrm{L}_2(q) \times \mathrm{P\Omega}_{n_2}^{\epsilon_2}(q) = L_1 \times L_2 \times L_3$$

and write $A = L\langle x, x_2, x_3\rangle$, where conjugation by x fixes the two $\mathrm{L}_2(q)$ factors in L. For now, let us assume $q > 27$. By the main theorem of [21] there exist $a_i, g_i \in L_i$ such that $L_i = \langle a_i^{x^{-1}}, a_i^{g_i}\rangle$ and $a_2 \neq f(a_1)$ for all $f \in \mathrm{Aut}(\mathrm{L}_2(q))$. By arguing as in the proof of Proposition 2.6(ii) we deduce that $d(L\langle x\rangle) = 2$ and thus $d(A) \leq 4$ as claimed.

Next suppose $3 < q \leq 27$. By [21], there exist $a_3, g_3 \in L_3$ such that $L_3 = \langle a_3^{x^{-1}}, a_3^{g_3}\rangle$ and it is easy to check directly that we can find elements $a_1, g_1 \in L_1$ and $a_2, g_2 \in L_2$ such that $L_i = \langle a_i^{x^{-1}}, a_i^{g_i}\rangle$ and $a_2 \neq f(a_1)$ for all $f \in \mathrm{Aut}(\mathrm{L}_2(q))$. For instance, suppose $q = 5$ and $y \in \mathrm{L}_2(q)$ has order r, where $r = 3$ or 5. If C is any conjugacy class of elements of order r in $\mathrm{L}_2(q)$ then there exists $c \in C$ such that $\mathrm{L}_2(q) = \langle y, c\rangle$, so we may take a_1 of order 3 and a_2 of order 5. The other cases are very similar. In particular, the previous argument applies.

Finally, let us assume $q = 3$, so $H = A = L.B$ as above. Suppose there exists an element $x \in B$ acting non-trivially on $L_1 \times L_2$ so that $(L_1 \times L_2)\langle x\rangle \neq \mathrm{PSO}_4^+(3)$. Then Proposition 2.15 implies that $d((L_1 \times L_2)\langle x\rangle) = 2$, say $(L_1 \times L_2)\langle x\rangle = \langle a_1, b_1x\rangle$, while [44, Result 1] gives $L_3\langle x\rangle = \langle a_2, b_2x\rangle$ for some $a_2, b_2 \in L_3$. It follows that $L\langle x\rangle = \langle (a_1, b_1), (a_2, b_2)x\rangle$, and by adding two further generators for B we obtain $d(H) \leq 4$. It remains to justify the existence of such an element $x \in B$.

If $\epsilon_2 = +$ then the proof of [29, 4.4.14] indicates that there exists an element $x = \delta_1 \otimes \delta_2^{-1} \in H_0$, where δ_1 induces a non-trivial diagonal automorphism on the $\mathrm{P\Omega}_4^+(3)$ factor (see [29, (4.4.20)]). Therefore $x \in B$ has the required property. Now assume $\epsilon_2 = -$. Here the proof of [29, 4.4.15] states that the above element $d = \delta_1 \otimes \delta_2^{-1}$ lies in $\mathrm{PSO}_n^+(3)$; if it belongs to G_0 then we are done, so let us assume otherwise. Let D denote the discriminant of the defining quadratic form for L_3 (see [29, p.32]). By [29, 4.4.15(IV)], if $D = \square$ then there exists $x \in B$ swapping the two $\mathrm{L}_2(3)$ factors, so this element has the desired property. Now assume $D = \boxtimes$. Write $V = V_1 \otimes V_2$, where V_1 and V_2 denote the natural modules for $\mathrm{P\Omega}_4^+(3)$ and $\mathrm{P\Omega}_{n_2}^-(3)$, respectively. Let $v \in V_1$ be a non-singular vector and let $r_v : V_1 \to V_1$ be the reflection in v with respect to the underlying non-degenerate symmetric bilinear form on V_1. By [29, 4.4.13(ii)] we have $r = r_v \otimes 1 \in \mathrm{PSO}_n^+(3) \setminus G_0$, so $x = rd \in H_0$ has the desired property since $r_v\delta_1 \in \mathrm{PGO}_4^+(3) \setminus \mathrm{PSO}_4^+(3)$. □

Lemma 5.10. *Theorem 2 holds if $G_0 = \mathrm{L}_n^\epsilon(q)$ and $H \in \mathcal{C}_7$ is of type $\mathrm{GL}_a^\epsilon(q) \wr S_t$.*

Proof. Here $a \geq 3$ and $(a, q, \epsilon) \neq (3, 2, -)$. Write $\mathrm{GL}_a^\epsilon(q) = \mathrm{SL}_a^\epsilon(q)\langle\delta\rangle$ and set $d = (\delta, \delta^{-1}, 1, \ldots, 1) \in \mathrm{GL}_a^\epsilon(q)^t$. For now, let us assume that at least one of the following three conditions do *not* hold:

$$t = 2, \quad a \equiv 2 \,(\mathrm{mod}\ 4), \quad q \equiv -\epsilon \,(\mathrm{mod}\ 4). \tag{4}$$

According to [29, 4.7.3], H is a quotient of $\hat{H} = \langle X^t, d\rangle.(S_t \times A)$, where $\langle X^t, d\rangle \leqslant \mathrm{GL}_a^\epsilon(q)^t$ and $X = \mathrm{SL}_a^\epsilon(q)\langle\delta^i\rangle$ for some $i \geq 0$. In addition, $A = Z_b \times Z_c$ with $c \in \{1, 2\}$ and b a divisor of $\log_p q$ (A is trivial if $G = G_0$). By Proposition 2.11(i) we have $\mathrm{SL}_a^\epsilon(q)\langle\delta^i\rangle = \langle x, y\rangle$ with x semisimple and y unipotent, so $d(H_0) \leq 4$ since $\langle X^t, d\rangle.S_t$ is generated by $(x, y, 1, \ldots, 1)$, d and two more for S_t. In general, $d(H) \leq 5$ since $S_t \times A$ is 3-generator.

Finally, if each of the conditions in (4) hold then H_0 is a quotient of $\hat{H} = \langle X^2, d\rangle$, where X and d are defined as before. Now $X = \mathrm{SL}_a^\epsilon(q)\langle\delta^i\rangle = \langle x, y\rangle$ with x semisimple and y unipotent, so X^2 is 2-generator and thus $d(H_0) \leq 3$. □

Lemma 5.11. *Theorem 2 holds if $G_0 = \mathrm{P\Omega}_n^+(q)$ and $H \in \mathcal{C}_7$ is of type $O_a^\epsilon(q) \wr S_t$.*

Proof. Here $a \geq 4$ is even, q is odd and $(a, \epsilon) \neq (4, +)$. We will assume $\epsilon = +$ since the case $\epsilon = -$ is very similar. Write $\mathrm{PO}_a^+(q) = \langle x, y\rangle$ and $\mathrm{PGO}_a^+(q) = \mathrm{PO}_a^+(q)\langle\delta\rangle$.

First suppose $t = 2$ and $a \equiv 2 \pmod 4$. By [29, 4.7.6] we have $H_0 = \mathrm{PSO}_a^+(q)^2.[2^2]$ and this is 4-generator since $d(\mathrm{PSO}_a^+(q)^2) = 2$ by Proposition 2.8. More generally, [29, 4.7.6] states that

$$H = \mathrm{PSO}_a^+(q)^2.[2^i].(Z_b \times Z_c),$$

where $2 \leq i \leq 4$, $b \in \{1, 2\}$ and c divides $\log_p q$. If $i = 4$ then $H = \mathrm{PGO}_a^+(q)^2.(Z_b \times Z_c)$ is 6-generator by Proposition 2.1(i). Similarly, $d(H) \leq 6$ when $i = 2$ since $d(\mathrm{PSO}_a^+(q)^2) = 2$. Finally, suppose $i = 3$. If $b = 1$ or c is odd then we quickly deduce that $d(H) \leq 6$, so let us assume $b = 2$ and c is even. Here $q \equiv 1 \pmod 4$, so [29, (4.7.20)] implies that $H = \mathrm{PO}_a^+(q)^2.2.(Z_2 \times Z_c)$ and thus $d(H) \leq 5$ since H is generated by $(x, 1)$, $(y, 1)$ and at most 3 more for $2.(Z_2 \times Z_c)$.

Next suppose $t = 3$, $a \equiv 2 \pmod 4$ and $q \equiv 3 \pmod 4$. Here $H = A.Z_b$ where

$$A \in \{\mathrm{PO}_a^+(q)^3.2^2.3,\ \mathrm{PO}_a^+(q)^3.2^2.S_3,\ \mathrm{PGO}_a^+(q)^3.3,\ \mathrm{PGO}_a^+(q)^3.S_3\}$$

and b divides $\log_p q$. Now $H_0 = \mathrm{PO}_a^+(q)^3.2^2.3$ is generated by $(x, 1, 1)$, $(y, 1, 1)$, $(\delta, \delta, 1)$ and one more for Z_3, so $d(H_0) \leq 4$ as required. In general, it is easy to see that $d(H) \leq 5$. For example, if $A = \mathrm{PO}_a^+(q)^3.2^2.S_3$ then H is generated by $(x, 1, 1)$, $(y, 1, 1)$, $(\delta, \delta, 1)$ and two more for $S_3 \times Z_b$.

In the remaining cases we have $H = A.(S_t \times Z_b)$, where $A = \mathrm{PO}_a^+(q)^t.2^{t-1}$ or $\mathrm{PGO}_a^+(q)^t$, and b divides $\log_p q$. Now, if $A = \mathrm{PGO}_a^+(q)^t$ then $d(H) \leq d(\mathrm{PGO}_a^+(q)) + d(S_t \times Z_b) \leq 4$ so let us assume $A = \mathrm{PO}_a^+(q)^t.2^{t-1}$. Here $d(H) \leq 5$ since H is generated by $(x, 1, \ldots, 1)$, $(y, 1, \ldots, 1)$ and $(\delta, \delta, 1, \ldots, 1)$ in A, together with two generators for $S_t \times Z_b$.

We need to work harder to establish $d(H_0) \leq 4$. Here $b = 1$, so the case $t = 2$ is clear. Now assume $t \geq 3$ and let $(y_1, \ldots, y_t; \sigma)$ denote a typical element of $\mathrm{PGO}_a^+(q)^t.S_t$. If $t \geq 5$ then H_0 is generated by the elements

$$(x, 1, \ldots, 1; 1),\ (y, 1, \ldots, 1; 1),\ (\delta, \delta, 1, \ldots, 1; (t-2, t-1, t)),\ (1, \ldots, 1; \sigma),$$

where $\sigma = (1, 2, \ldots, \alpha)$ and $\alpha = t$ if t is even, otherwise $\alpha = t - 1$.

Next suppose $t = 3$. We claim that $H_0 = \langle x_1, x_2, x_3, x_4\rangle$, where

$$x_1 = (x, 1, 1; 1),\ x_2 = (y, 1, 1; (1, 3)),\ x_3 = (\delta, \delta, 1; 1),\ x_4 = (1, 1, 1; (2, 3)).$$

To see this, let $L = \langle x_1, x_2, x_3, x_4\rangle$, $m = |y|$ and first observe that

$$x_2^{2m-2} \cdot (x_4x_2)^3 = (1, y, 1; 1) \in L$$

and thus $(1, y, 1; 1)^{x_4} = (1, 1, y; 1) \in L$. Therefore $x_2 \cdot (1, 1, y^{m-1}; 1) = (1, 1, 1; (1, 3)) \in L$ and the claim follows since $H_0 = \langle x_1, x_3, (1, y, 1; 1), (1, 1, 1; (1, 3))\rangle$. Similar reasoning shows that if $t = 4$ then H_0 is generated by the elements

$$(x, 1, 1, 1; 1),\ (y, 1, 1, 1; (1, 4)),\ (\delta, \delta, 1, 1; 1),\ (1, 1, 1, 1; (1, 2, 3, 4)). \qquad \square$$

Remark 5.12. Suppose $G_0 = \mathrm{P}\Omega_n^+(q)$ and $H \in \mathcal{C}_7$ is of type $O_a^+(q) \wr S_2$, where $a \equiv 2 \pmod 4$ and $q \equiv 1 \pmod 4$. Then [29, 4.7.6] indicates that

$$H_0 = \mathrm{PSO}_a^+(q)^2.[4] = \mathrm{PO}_a^+(q) \times \mathrm{PO}_a^+(q) = (\mathrm{P}\Omega_a^+(q) \times \mathrm{P}\Omega_a^+(q)).2^4$$

(see [29, (4.7.20)]) and thus $d(H_0) = 4$. In this way we obtain an infinite family of examples (G, H), where G is simple and H is a maximal subgroup of G requiring 4 generators, demonstrating the sharpness of the bound on $d(H \cap G_0)$ in Theorem 2.

Lemma 5.13. *Theorem 2 holds in the remaining $\mathcal{C}_4$ and $\mathcal{C}_7$ cases.*

Proof. This is straightforward. For example, suppose $G_0 = \mathrm{P}\Omega_n^+(q)$ and $H \in \mathcal{C}_7$ is of type $\mathrm{Sp}_a(q) \wr S_t$, where tq is even and $(a, q) \neq (2, 2)$. If $t = 2$ and $a \equiv 2 \pmod 4$ then $H_0 = \mathrm{PSp}_a(q)^2$ is 2-generator, otherwise $H = A.(S_t \times Z_b)$, where b divides $\log_p q$ and either $A = \mathrm{PGSp}_a(q)^t$, or q is odd and $A = \mathrm{PSp}_a(q)^t.2^{t-1}$. In the former case we have

$d(H) \leq d(\mathrm{PGSp}_a(q)) + d(S_t \times Z_b) \leq 4$ and as above we observe that the same bound also holds if $A = \mathrm{PSp}_a(q)^t.2^{t-1}$. The other cases are very similar. $\square$

5.4. **Novelty subgroups.** It remains to deal with certain *novelty* subgroups H of G, where $H_0 = H \cap G_0$ is non-maximal in G_0. By [1] and our earlier analysis, we may assume that one of the following holds:

(a) $G_0 = \mathrm{Sp}_4(q)'$, $p = 2$ and G contains a graph automorphism;
(b) $G_0 = \mathrm{P}\Omega_8^+(q)$ and G contains a triality automorphism.

In [1, §14], Aschbacher proves a version of his main theorem which describes the various possibilities in case (a), but his theorem does not apply in case (b); here the possibilities were determined later by Kleidman [28]. We record the relevant non-parabolic subgroups in Table 2. Note that in case (a) we may assume $q > 2$ since $\mathrm{Sp}_4(2)' \cong A_6$.

	G_0	type of H	conditions
(i)	$\mathrm{Sp}_4(q)'$	$O_2^\epsilon(q) \wr S_2$	$q > 2$ even
(ii)		$O_2^-(q^2).2$	$q > 2$ even
(iii)	$\mathrm{P}\Omega_8^+(q)$	$\mathrm{GL}_3^\epsilon(q) \times \mathrm{GL}_1^\epsilon(q)$	
(iv)		$O_2^-(q^2) \times O_2^-(q^2)$	
(v)		$[2^9].\mathrm{SL}_3(2)$	$q = p > 2$

TABLE 2. Some novelty subgroups

In cases (i) and (ii) it is very easy to check that $d(H_0) \leq 3$, so let us consider (iii) – (v).

Lemma 5.14. *Theorem 2 holds in case (iii) of Table 2.*

Proof. It suffices to prove that $d(H_0) \leq 4$ since G/G_0 is a subgroup of $S_4 \times Z_f$ containing a triality (where $q = p^f$), and such a subgroup is 2-generator. If $p = 2$ then $H_0 = (\mathrm{GL}_3^\epsilon(q) \times \mathrm{GL}_1^\epsilon(q)).2$ is clearly 4-generator, so let us assume p is odd. By [28, 3.2.2, 3.2.3], H_0 is a quotient of $\hat{H} \cong (Z_{(q-\epsilon)/2} \times A).2^2$, where A is the index-two subgroup of $\mathrm{GL}_3^\epsilon(q)$ containing $\mathrm{SL}_3^\epsilon(q)$. Write $Z_{(q-\epsilon)/2} = \langle z \rangle$ and $A = \langle x, y \rangle$, where x is semisimple and y is unipotent (see Proposition 2.11(i)). Then $Z_{(q-\epsilon)/2} \times A = \langle (z, y), (1, x) \rangle$, so $d(H_0) \leq 4$ as required. $\square$

Lemma 5.15. *Theorem 2 holds in cases (iv) and (v) of Table 2.*

Proof. Again, it suffices to show that $d(H_0) \leq 4$. According to [28, 3.3.1], in (iv) we have

$$H_0 = N_{G_0}(S) \cong (D_{2l} \times D_{2l}).2^2,$$

where S is a Sylow r-subgroup of G_0 for an odd prime r dividing $q^2 + 1$, and $l = (q^2 + 1)/(2, q-1)$ is odd. Now $D_{2l} = \langle x, y \rangle$ with $|x| = l$ and $|y| = 2$, hence $D_{2l} \times D_{2l}$ is 2-generator and thus $d(H_0) \leq 4$. As explained in [28, §3.4], in (v) we have $H_0 = N_{G_0}(P)$, where $P < G_0$ is a $2A$*-pure* group of order 8 which centralizes an orthogonal decomposition of the natural G_0-module into 1-dimensional non-degenerate subspaces. More precisely, by [28, 3.4.2(ii)] we have $H_0 \cong [2^9].\mathrm{SL}_3(2)$. It is straightforward to explicitly construct H_0 as a subgroup of $\mathrm{P}\Omega_8^+(3)$ and we quickly deduce that $d(H_0) = 2$. $\square$

6. EXCEPTIONAL GROUPS

In this section we complete the proof of Theorem 2 for non-parabolic subgroups of groups of Lie type. Let G be an almost simple group with socle G_0, an exceptional group of Lie type over $\mathbb{F}_q$, and let H be a maximal subgroup of G. Write $\bar{G}$ for the corresponding simple adjoint algebraic group over the algebraic closure $\bar{\mathbb{F}}_q$, and let σ be a Frobenius morphism of $\bar{G}$ such that $G_0 = \bar{G}'_\sigma$. Recall that $\bar{G}_\sigma = \mathrm{Inndiag}(G_0)$, the group generated by all inner and diagonal automorphisms of G_0. As before, we define $H_0 = H \cap G_0$. Since $d(G/G_0) \leq 2$ (see Proposition 2.1(i)), it suffices to prove that $d(H_0) \leq 4$. In this section we assume that H is not a parabolic subgroup; we will deal with these in the next section.

According to [35, Theorem 2], the possibilities for H_0 are as follows. In part (iv) below, $F^*(H_0)$ denotes the generalized Fitting subgroup of H_0.

Proposition 6.1. *One of the following holds:*

(i) H_0 *is almost simple;*

(ii) $H_0 = N_{G_0}(D_\sigma)$, *where D is a connected reductive subgroup of $\bar{G}$ of maximal rank, not a maximal torus; the possibilities are listed in* [34, Table 5.1]*;*

(iii) $H_0 = N_{G_0}(T_\sigma)$, *where T is a maximal torus of $\bar{G}$; the possibilities are listed in* [34, Table 5.2]*;*

(iv) $F^*(H_0)$ *is as in* [35, Table III]*;*

(v) $H_0 = N_{G_0}(E)$, *where E is an elementary abelian group given in* [12, Theorem 1(II)].

In case (i), $d(H_0) \leq 3$ by Proposition 2.1(i), so we need only consider cases (ii)–(v).

Lemma 6.2. *Theorem 2 holds in case (iv) of Proposition 6.1.*

Proof. According to [35, Table III], the possibilities for $N_{\bar{G}_\sigma}(H_0)$ are as follows:

G_0	$N_{\bar{G}_\sigma}(H_0)$
$E_8(q)$	$A \times \mathrm{PGL}_3^\epsilon(q).2$, $G_2(q) \times F_4(q)$, $A \times G_2(q)^2.2$, $A \times G_2(q^2).2$
$E_7(q)$	A^2, $A \times G_2(q)$, $A \times F_4(q)$, $G_2(q) \times \mathrm{PGSp}_6(q)$
$E_6^\epsilon(q)$	$\mathrm{PGL}_3^{\epsilon'}(q).2 \times G_2(q)$
$F_4(q)$	$A \times G_2(q)$

where $A = \mathrm{PGL}_2(q)$ (note that there are also conditions on q for the groups in the table to ensure that all factors are non-solvable). Using Proposition 2.6 we deduce that $d(H_0) \leq 4$ in all cases. $\square$

Lemma 6.3. *Theorem 2 holds in case (v) of Proposition 6.1.*

Proof. By [12, Theorem 1(II)], one of the following holds:

G_0	$N_{\bar{G}_\sigma}(H_0)$
$E_8(q)$	$5^3.\mathrm{SL}_3(5)$, $2^{5+10}.\mathrm{SL}_5(2)$
$E_7(q)$	$(2^2 \times \mathrm{P}\Omega_8^+(q).2^2).S_3$ (q odd)
$E_6^\epsilon(q)$	$3^{3+3}.\mathrm{SL}_3(3)$
$F_4(q)$	$3^3.\mathrm{SL}_3(3)$
$G_2(q)$	$2^3.\mathrm{SL}_3(2)$
${}^2G_2(q)$	$2^3.7$

For $G_0 \neq E_7(q)$ it is immediate that $d(H_0) \leq 3$ in all cases. For $G_0 = E_7(q)$, factoring out the normal 2^2 we obtain the almost simple group $\mathrm{P}\Omega_8^+(q).S_4$, which is 2-generated by Proposition 2.1(i). The S_3 acts faithfully on the normal 2^2, so $d(H_0) \leq 3$. $\square$

Lemma 6.4. *Theorem 2 holds in case (ii) of Proposition 6.1.*

Proof. Here $N_{\bar{G}_\sigma}(H_0)$ is given in [34, Table 5.1]. In Table 3 we summarise enough information to give what we want. In each case H_0 has a normal subgroup K as indicated, and K is a central product $\prod H_i \circ T$, where each H_i is either quasisimple or in $\{\mathrm{SL}_2(2), \mathrm{SL}_2(3), \mathrm{SU}_3(2)\}$, and T is an abelian p'-group. In the table, we use the following

G_0	K	$N_{\bar{G}_\sigma}(H_0)/K$
$E_8(q)$	$D_8(q)$, $A_1(q)E_7(q)$, $A_4^-(q^2)$, ${}^3D_4(q)^2$, ${}^3D_4(q^2)$, $A_2^-(q^2)^2$, $A_2^-(q^4)$	cyclic
	$A_8^\epsilon(q)$, $A_2^\epsilon(q)E_6^\epsilon(q)$, $A_4^\epsilon(q)^2$, $A_4^-(q^2)$	$e.2$, $e.2$, $g.4$, $h.4$
	$D_4(q)^2$, $D_4(q^2)$	$d^2.(S_3\times 2)$, $S_3\times 2$
	$A_2^\epsilon(q)^4$	$e^2.\mathrm{GL}_2(3)$
	$A_1(q)^8$	$d^4.\mathrm{AGL}_3(2)$
$E_7(q)$	$A_1(q)D_6(q)$, $A_1(q^3).{}^3D_4(q)$, $A_1(q^7)$	cyclic
	$A_2^\epsilon(q)A_5^\epsilon(q)$, $E_6^\epsilon(q)\circ(q-\epsilon)$	$de.2$, $e.2$
	$A_7^\epsilon(q)$	$i.(2\times 2/f)$
	$A_1(q)^3D_4(q)$	$d^3.S_3$
	$A_1(q)^7$	$d^4.\mathrm{L}_3(2)$
$E_6^\epsilon(q)$	$A_1(q)A_5^\epsilon(q)$, ${}^3D_4(q)\times(q^2+\epsilon q+1)$, $D_5^\epsilon(q)\circ(q-\epsilon)$	cyclic
	$A_2(q^2)A_2^{-\epsilon}(q)$, $A_2^\epsilon(q^3)$	$j.2$, $e.3$
	$A_2^\epsilon(q)^3$	$e^2.S_3$
	$D_4(q)\circ(q-\epsilon)^2$	$d^2.S_3$
$F_4(q)$	$A_1(q)C_3(q)$, $B_4(q)$, ${}^3D_4(q)$, $B_2(q)^2$ $(p=2)$, $B_2(q^2)$ $(p=2)$	cyclic
	$D_4(q)$, $A_2^\epsilon(q)^2$	S_3, $e.2$
${}^2F_4(q)$	$A_2^-(q)$, ${}^2B_2(q)^2$, $B_2(q)$	cyclic
	$A_2^-(q)$	$k.2$
$G_2(q)$	$A_1(q)^2$, $A_2^\epsilon(q)$	cyclic
${}^3D_4(q)$	$A_1(q)A_1(q^3)$, $A_2^\epsilon(q)\circ(q^2+\epsilon q+1)$	d, $l.2$
${}^2G_2(q)$	$A_1(q)$	2

TABLE 3. Maximal rank subgroups

notation: $d=(2,q-1)$, $e=(3,q-\epsilon)$, $f=(4,q-\epsilon)/d$, $g=(5,q-\epsilon)$, $h=(5,q^2+1)$, $i=(8,q-\epsilon)/d$, $j=(3,q^2-1)$, $k=(3,q+1)$, $l=(3,q^2+\epsilon q+1)$.

Now $d(H_0)\le d(H_0/K)+2$ by Proposition 2.6(i), and H_0/K is either equal to the group $N_{\bar{G}_\sigma}(H_0)/K$ in the right hand column of the table, or has index dividing 2 or 3 in this for $G_0=E_7(q)$ or $E_6^\epsilon(q)$. It is clear that all such groups are 2-generated, except possibly in the following cases:

G_0	H_0/K
$E_8(q)$	$2^2.(S_3\times 2)$, $3^2.\mathrm{GL}_2(3)$, $2^4.\mathrm{AGL}_3(2)$
$E_7(q)$	$2^2.S_3$, $2^3.\mathrm{L}_3(2)$
$E_6^\epsilon(q)$	$3.S_3$, $2.S_3$

However a check using MAGMA verifies that each of these groups, except possibly $3.S_3$ in the last row, is also 2-generated. In the remaining case, $G_0=E_6^\epsilon(q)$ with $e=3$, $K=A_2^\epsilon(q)^3$ and $H_0/K\cong 3.S_3$. If $(q,\epsilon)=(2,-)$ then the Atlas [13] indicates that $H_0/K\cong Z_3\times S_3$ which is 2-generator, so the usual argument applies. Now assume $q>2$. Now H_0 contains a subgroup $K.3=K\langle x\rangle$, where x induces a diagonal automorphism of order 3 on each factor $A_2^\epsilon(q)$ of K. Pick elements a_1,a_2,a_3 of different prime orders in $A_2^\epsilon(q)$. By [21] there exist b_1,b_2,b_3 such that $\langle a_i,b_i\rangle=A_2^\epsilon(q)$ for each i. Then the two elements (a_1,a_2,a_3) and $(b_1,b_2,b_3)x$ generate $K\langle x\rangle$. As $H_0/K\langle x\rangle\cong S_3$, it follows that $d(H_0)\le 4$. □

Lemma 6.5. *Theorem 2 holds in case (iii) of Proposition 6.1.*

Proof. Here $H_0 = N_{G_0}(T_\sigma)$, where T_σ and $W_\sigma := N_{G_0}(T_\sigma)T_\sigma/T_\sigma$ are as in Table 4. In the table we set $\epsilon = \pm 1$, while $W(X)$ denotes the Weyl group of the root system of type X.

G_0	T_σ	W_σ
$E_8(q)$	$(q-\epsilon)^8$	$W(E_8)$
	$(q^4+\epsilon q^3+q^2+\epsilon q+1)^2$	$5 \times \mathrm{SL}_2(5)$
	$(q^2+\epsilon q+1)^4$	$2.(3 \times \mathrm{U}_4(2))$
	$(q^2+1)^4$	$(4 \circ 2^{1+4}).A_6.2$
	$(q^4-q^2+1)^2$	$12 \circ \mathrm{GL}_2(3)$
	$q^8+\epsilon q^7-\epsilon q^5-q^4-\epsilon q^3+\epsilon q+1$	Z_{30}
$E_7(q)$	$(q-\epsilon)^7$	$W(E_7)$
$E_6^\epsilon(q)$	$(q-\epsilon)^6$	$W(E_6)$
	$(q^2+\epsilon q+1)^3$	$3^{1+2}.\mathrm{SL}_2(3)$
$F_4(q)$	$(q-\epsilon)^4$	$W(F_4)$
$(p=2)$	$(q^2+\epsilon q+1)^2$	$3 \times \mathrm{SL}_2(3)$
	$(q^2+1)^2$	$4 \circ \mathrm{GL}_2(3)$
	q^4-q^2+1	Z_{12}
${}^2F_4(q)$	$(q+1)^2$	$\mathrm{GL}_2(3)$
	$(q+\epsilon\sqrt{2q}+1)^2$	$4 \circ \mathrm{GL}_2(3)$
	$q^2+\epsilon\sqrt{2q^3}+q+\epsilon\sqrt{2q}+1$	Z_{12}
$G_2(q)$	$(q-\epsilon)^2$	D_{12}
$(p=3)$	$q^2+\epsilon q+1$	Z_6
${}^3D_4(q)$	$(q^2+\epsilon q+1)^2$	$\mathrm{SL}_2(3)$
	q^4-q^2+1	Z_4
${}^2G_2(q)$	$q+1,\ q+\epsilon\sqrt{3q}+1$	$Z_6,\ Z_6$
${}^2B_2(q)$	$q-1,\ q+\epsilon\sqrt{2q}+1$	$Z_2,\ Z_4$

TABLE 4. Normalizers of maximal tori

First assume $G_0 = E_8(q)$. We claim that $d(H_0) \le 1 + d(W_\sigma)$. To see this, take $t \in T_\sigma$ of maximal order, and $d := d(W_\sigma)$ further elements $h_1, \ldots, h_d$ generating H_0 modulo T_σ. If r is a prime dividing the order of t, then by inspection we see that W_σ acts irreducibly on $\Omega_r := \Omega_1(O_r(T_\sigma))$. Since Ω_r contains a power of t it follows that $\Omega_r \leqslant \langle t, h_1, \ldots, h_d\rangle$. Repeating this argument with H_0/Ω_r, we see that $T_\sigma \leqslant \langle t, h_1, \ldots, h_d\rangle$, and hence $H_0 = \langle t, h_1, \ldots, h_d\rangle$. This proves the claim. Now a check using MAGMA shows that all of the groups W_σ are 2-generated, and so by the claim, $d(H_0) \le 3$, giving the result for $G_0 = E_8(q)$.

The argument is similar for the other types. The only slight difference occurs for $G_0 = E_7(q)$ (with q odd) or $E_6^\epsilon(q)$ (with $q-\epsilon$ divisible by 3), where the irreducibility assertion for W_σ on Ω_r does not necessarily hold for $r = 2$ or 3, respectively. For $E_7(q)$ we have $N_{G_0}(T_\sigma) = ((q-\epsilon)^7/2).W_\sigma$ and $N_{\bar{G}_\sigma}(T_\sigma) = (q-\epsilon)^7.W_\sigma$, and the previous argument still goes through, as we can choose the element t so that $\Omega_r \leqslant \langle t, h_1, \ldots, h_d\rangle$. The same observation also applies in the relevant $E_6^\epsilon(q)$ cases. □

7. PARABOLIC SUBGROUPS

Let G be an almost simple group with socle G_0 of Lie type. In this section we complete the proof of Theorem 2 by handling the case where H is a maximal parabolic subgroup of G. Write $H_0 = H \cap G_0 = QR$, where Q is the unipotent radical and R a Levi subgroup. Denote by $P_{ij\ldots}$ the parabolic subgroup obtained by deleting nodes $i, j, \ldots$ from the Dynkin diagram of G_0. By the maximality of H, one of the following holds:

(a) $H_0 = P_i$ for some i;
(b) G_0 is of type A_n, D_n, E_6, $F_4\,(p = 2)$, $B_2\,(p = 2)$ or $G_2\,(p = 3)$, G contains a graph automorphism τ, and $H_0 = P_{ij}$ where nodes i, j are interchanged by τ;
(c) G_0 is of type D_4, G contains a triality automorphism, and $H_0 = P_{134}$.

Lemma 7.1. *Let $H_0 = QR$ be as above, and exclude case (c), together with the following cases:*

$$p = 2 : \; G_0 = C_n(q),\, F_4(q),\, {}^2F_4(q),\, G_2(q),\, {}^2B_2(q)$$
$$p = 3 : \; G_0 = G_2(q),\, {}^2G_2(q).$$

Then $d(H_0) \leq 1 + d(R)$.

Proof. We refer to [3] for the structure of parabolic subgroups. Note that, owing to the cases excluded in the hypothesis, G_0 is not *special*, in the terminology of [3].

First assume G_0 is untwisted and $H_0 = P_i$ for some i. Then by [3, Theorem 2(a)], Q/Q' is an irreducible $\mathbb{F}_q R$-module. Hence if we generate R with d elements $r_1, \ldots, r_d$, and add one more non-identity element $u \in Q \setminus Q'$, then $r_1, \ldots, r_d, u$ generate P_i modulo Q'. But $Q' \leqslant \Phi(Q)$, so $Q' \leqslant \Phi(P_i)$ and thus $r_1, \ldots, r_d, u$ generate P_i, giving the conclusion.

Now assume that G_0 is twisted, of type ${}^2A_n$, ${}^2D_n$ or ${}^2E_6$. In the first case consider the covering group $\hat{G}_0 = \mathrm{SU}_m(q)$ (where $m = n + 1$). The Levi subgroup

$$\hat{R} \cong \{(A, B) \in \mathrm{GL}_i(q^2) \times \mathrm{GU}_{m-2i}(q) \mid \det(B) = \det(A)^{q-1}\},$$

where $H_0 = P_i$, and [3] (or direct matrix calculation) shows that Q/Q' has the structure of the $\hat{R}$-module $V_i \otimes V_{m-2i} + V_i^{(q)} \otimes V_{m-2i}^*$, where V_i, V_{m-2i} are the natural modules for the factors of $\hat{R}$. As the two composition factors are non-isomorphic $\hat{R}$-modules, we can choose a vector $uQ' \in Q/Q'$ lying in no proper $\hat{R}$-invariant subspace. The conclusion now follows as in the previous paragraph. A similar argument works for the ${}^2D_n$ and ${}^2E_6$ cases: for ${}^2D_n$, the only parabolic for which Q/Q' is reducible is P_{n-1}, in which case R contains a subgroup of index $(2, q-1)$ of $\mathrm{GL}_{n-1}(q)$ and $Q/Q' \cong V_{n-1} + V_{n-1}^*$; and for ${}^2E_6$, Q/Q' is again the sum of at most two non-isomorphic irreducible R-modules. In all cases there is a vector $uQ' \in Q/Q'$ lying in no proper R-invariant subspace, and the conclusion follows.

Next suppose $G_0 = {}^3D_4(q)$. Let R_0 denote the semisimple part of R. If $H_0 = P_2$ then $R_0 = A_1(q^3)$ and Q/Q' is an irreducible R-module $V_2 \otimes V_2^{(q)} \otimes V_2^{(q^2)}$, giving the conclusion in the usual way. And if $H_0 = P_1$ then R contains $A_1(q) \circ (q^3 - 1)$ and again Q/Q' is an irreducible R-module (of dimension 6).

In view of the exclusions in the hypothesis, the only remaining cases to consider are those where G_0 is of type A_n, D_n or E_6, and G contains a graph automorphism. The maximal parabolics in G for which Q/Q' is a reducible R-module are $P_{i,n-i}$ (for A_n), P_{n-1} (for D_n) and P_{16}, P_{35} (for E_6). For these, [3] shows that Q/Q' is a sum of two irreducible R-modules, and the conclusion follows as before. □

Lemma 7.2. *Under the hypotheses of Lemma* 7.1*, we have $d(H_0) \leq 4$.*

Proof. Write $H_0 = QR$ as above. In view of Lemma 7.1, it suffices to show that $d(R) \leq 3$.

First consider classical groups. It is convenient to replace G_0 by the corresponding classical linear group $\mathrm{SL}_n(q)$, $\mathrm{Sp}_n(q)$, etc.

For $G_0 = \mathrm{SL}_n(q)$ we have $H_0 = P_i$ or $P_{i,n-i}$. In the first case $R = (\mathrm{SL}_i(q) \times \mathrm{SL}_{n-i}(q)).(q-1)$, and $d(R) \leq 3$ by Proposition 2.6 (if $i \neq n - i$) and by Proposition 2.10 (if $i = n - i$). In the second case we have

$$R = \{(A, B, C) \in \mathrm{GL}_{n-2i}(q) \times \mathrm{GL}_i(q)^2 \mid \det(ABC) = 1\}.$$

If $i = 1$ then $d(R) \leq 3$ by Proposition 2.6, so assume $i > 1$. By Proposition 2.11, there are semisimple elements x, y and unipotent elements u, v such that

$$\mathrm{GL}_{n-2i}(q) = \langle x, u\rangle,\ \mathrm{GL}_i(q) = \langle y, v\rangle.$$

Furthermore we may take it that $\det(x) = \det(y) = \mu$, a generator of $\mathbb{F}_q^*$. Define the following elements $r, s, t \in R$:

$$r = (x, y^{-1}, v),\ s = (x^{-1}, v, y),\ t = (u, y^{-1}, y).$$

We claim that r, s, t generate R. Indeed, observe first that by taking suitable powers of these elements we see that $\langle r, s, t\rangle$ contains $(1, 1, v)$, $(1, v, 1)$ and $(1, y^{-1}, y)$, hence contains all elements $(1, B, C)$ with $\det(BC) = 1$. It also contains $(u, 1, 1)$ and $(x, y^{-1}, 1)$. Hence it contains $\mathrm{SL}_{n-2i}(q) \times \mathrm{SL}_i(q)^2$ and maps onto Z_{q-1}^2, proving the claim.

Next, if $G_0 = \mathrm{SU}_n(q)$ and $H_0 = P_i$, then $R = (\mathrm{SL}_i(q^2) \times \mathrm{SU}_{n-2i}(q)).(q^2-1)$, and we see that $d(R) \leq 3$ using Proposition 2.6. Similarly, if $G_0 = \mathrm{Sp}_n(q)$ (so q is odd by hypothesis), we have $R = \mathrm{GL}_i(q) \times \mathrm{Sp}_{n-2i}(q)$ and once again we can use Proposition 2.6 (or Proposition 2.10 when $i = n - 2i = 2$).

Now consider $G_0 = \Omega_n^\epsilon(q)$, with $n \geq 7$. By hypothesis, if n is odd then q is odd. If q is even then $R = \mathrm{GL}_i(q) \times \Omega_{n-2i}^\epsilon(q)$, and it is easy to see that $d(R) \leq 3$ using Propositions 2.6 and 2.10, as usual. So assume q is odd. Then

$$R = \{(A, B) \in \mathrm{GL}_i(q) \times \mathrm{SO}_{n-2i}^\epsilon(q) \mid \det(A)\theta(B) \text{ is a square in } \mathbb{F}_q\},$$

where $\theta : \mathrm{SO}_{n-2i}^\epsilon(q) \to \mathbb{F}_q^*/(\mathbb{F}_q^*)^2$ denotes the spinor norm map (see [29, p.29]). If $i = 1$, then R is a cyclic extension of $\Omega_{n-2}^\epsilon(q)$, giving the conclusion by Proposition 2.6. If $i > 1$ and $n - 2i > 4$ or $n - 2i \in \{0, 1, 3\}$, then R is a cyclic extension of $\mathrm{SL}_i(q) \times \Omega_{n-2i}^\epsilon(q)$ and we can again use Proposition 2.6 (or Proposition 2.10 when $(n, i) = (7, 2)$).

It remains to handle the cases where $n = 2m$ is even and $i = m - 2$ or $m - 1$. First let $i = m-2$. Then $R \leqslant \mathrm{GL}_{m-2}(q) \times \mathrm{SO}_4^\epsilon(q)$ and R is a cyclic extension of $\mathrm{SL}_{m-2}(q) \times \Omega_4^\epsilon(q)$. If $m > 4$, or $(m, \epsilon) = (4, -)$, we can use Proposition 2.4(ii) to see that the latter group is 2-generator, giving the result. So suppose $m = 4$ and $\epsilon = +$. If $q \leq 3$ we check the result directly by computation, so take $q > 3$. By Propositions 2.11 and 2.13, there are semisimple elements x, y and unipotent elements u, v such that

$$\mathrm{GL}_2(q) = \langle x, u\rangle,\ \mathrm{SO}_4^+(q) = \langle y, v\rangle.$$

Let $r = (x, y)$, $s = (u, v)$, $t = (x, y^{-1})$, all elements of R. One easily checks that r, s, t generate R, giving the conclusion. Finally, if $i = m - 1$ we have $R \leqslant \mathrm{GL}_{m-1}(q) \times \mathrm{SO}_2^\epsilon(q)$ and we use a similar argument: write $\mathrm{GL}_{m-1}(q) = \langle x, u\rangle$ and $\mathrm{SO}_2^\epsilon(q) = \langle z\rangle$, and see that R is generated by the three elements (x, z), (x^{-1}, z) and $(u, 1)$.

This completes the proof for classical groups. Now consider exceptional groups. Assume $G_0 \neq E_6^\epsilon(q)$ or ${}^3D_4(q)$. Then by hypothesis, G_0 is untwisted and $H_0 = P_i$ for some i. The Levi subgroup $R = R_0J$, where R_0 (the semisimple part of R) is the group generated by all fundamental root subgroups $U_{\pm\alpha_j}$ with $j \neq i$, and J is a Cartan subgroup. Thus R_0 is a central product $\prod L_j$ of total semisimple rank $r - 1$, where r is the rank of G_0 and each L_i is either quasisimple or in $\{\mathrm{SL}_2(2), \mathrm{SL}_2(3)\}$. It follows that R is a cyclic extension of R_0. Moreover, inspection of the Dynkin diagrams of exceptional types shows that the groups $L_j/Z(L_j)$ are pairwise non-isomorphic, and hence R is 2-generator by Proposition 2.6(ii), giving the conclusion.

If $G_0 = {}^3D_4(q)$ then R is a cyclic extension of $A_1(q)$ or $A_1(q^3)$, so $d(R) \leq 2$ by Proposition 2.6. Finally, let $G_0 = E_6^\epsilon(q)$. First suppose $\epsilon = +$ and $H_0 = P_i$. If $i \neq 4$ the argument of the previous paragraph goes through; and if $i = 4$ then $R_0 = A_1(q)A_2(q)^2$. This is easily checked to be 2-generator if $q \leq 3$, and can be seen to be also 2-generator if $q > 3$, using Propositions 2.4(ii) and 2.10. Hence $d(R) \leq 3$.

It remains to consider the cases where $\epsilon = -$, or $\epsilon = +$ and $H_0 = P_{16}, P_{35}$. For $\epsilon = -$ and $H_0 = P_2$ or P_4 we have $R = R_0J$, a cyclic extension of $R_0 = {}^2A_5(q)$ or $A_1(q)A_2(q^2)$; then $d(R_0) = 2$ by Proposition 2.4(ii), so $d(R) \leq 3$, as required. The remaining parabolics are as follows:

(i) P_{16} ($\epsilon = +$), P_1 ($\epsilon = -$): $R_0 = D_4^\epsilon(q)$;
(ii) P_3 ($\epsilon = -$): $R_0 = A_2(q)A_1(q^2)$;
(iii) P_{35} ($\epsilon = +$): $R_0 = A_2(q)A_1(q)^2$.

In all cases, $d(R/R_0) \leq 2$. It follows using Proposition 2.6(ii) that $d(R) \leq 3$ in cases (i) and (ii). As for (iii), we use a slight variation of the argument in the proof of Proposition 2.6(ii) to show that $d(R) \leq 3$. First we check by computation that the conclusion holds for $q \leq 5$, so assume $q > 5$. Let $R_0 = L_1L_2L_3$ with $L_1, L_2 \cong A_1(q)$ and $L_3 \cong A_2(q)$, and let $x \in R \setminus R_0$. As in Proposition 2.6, the aim is to show that $d(R_0\langle x\rangle) = 2$. As x lies in the Levi subgroup R, it fixes all factors of R_0, inducing an inner or diagonal automorphism on each. Using the subgroup structure of $\mathrm{L}_2(q)$, it is easy to see that if $z \in \mathrm{L}_2(q)$ has order $r_1 = (q-1)/d$, where $d = (2, q-1)$, and C is any $\mathrm{L}_2(q)$-class of elements of order r_1 then there exists $c \in C$ such that $\mathrm{L}_2(q) = \langle z, c\rangle$. Similarly for elements of order $r_2 = (q+1)/d$. Therefore, there exist $a_i, g_i \in L_i$ ($i = 1, 2$) such that a_i has order r_i and $L_i = \langle a_i^{x^{-1}}, a_i^{g_i}\rangle$. Pick $a_3, g_3 \in L_3$ as in the proof of Proposition 2.6, and let $a = (a_1, a_2, a_3)$, $b = (g_1, g_2, g_3)x \in R_0\langle x\rangle$. Then $\langle a, a^b\rangle$ projects surjectively onto each factor L_i, and since a_1, a_2 have different orders it follows that $\langle a, a^b\rangle = R_0$. Hence $\langle a, b\rangle = R_0\langle x\rangle$, showing that $d(R_0\langle x\rangle) = 2$, as required. Hence $d(R) \leq 3$. □

Lemma 7.3. *We have $d(H_0) \leq 4$ in the excluded $p = 2, 3$ cases of Lemma 7.1.*

Proof. The cases under consideration are G_0 of type $C_n, F_4, {}^2F_4, G_2, {}^2B_2$ (all with $p = 2$), and $G_2, {}^2G_2$ (with $p = 3$).

Consider $G_0 = C_n(q)$ with q even. If $H_0 = P_i = QR$, then $R = \mathrm{GL}_i(q) \times \mathrm{Sp}_{2n-2i}(q)$ and we can see that $d(R) = 2$ using Proposition 2.11. Also Q/Q' has two R-composition factors, and we deduce that $d(H_0) \leq 4$, as required. The only other case occurs when $G_0 = C_2(q)$, G contains a graph automorphism and H_0 is a Borel subgroup. Here $R = (q-1)^2$ and $Q/Q' \cong (\mathbb{F}_q)^2$, generated by two root groups modulo Q' with R acting as a full group of scalars on each root group, so again $d(H_0) \leq 4$.

Next consider $G_0 = F_4(q)$, q even. If G contains no graph automorphism of G_0, then we may take $H_0 = P_1$ or P_2 (since P_3, P_4 are images of these under a graph automorphism); and if G contains a graph automorphism, $H_0 = P_{14}$ or P_{23}. If $q = 2$ then we can use the explicit permutation representation of degree 69888 for G_0 provided in the Web-Atlas [57] to check that $d(H_0) = 2$ in all cases, so we may assume $q \geq 4$. Write $H_0 = QR$ as before. Since q is even, G_0 is *special* in the terminology of [3], and Q/Q' is no longer necessarily irreducible. Nevertheless, Q/Q' still has a filtration by $\mathbb{F}_qR$-modules, and it is routine to use the commutator relations given in [51, p.404] to calculate its composition factors. In the table below we record these according to their high weights, where R_0 is the semisimple part of R:

H_0	R_0	R_0-composition factors of Q/Q'
P_1	$C_3(q)$	001, 100
P_2	$A_1(q)A_2(q)$	$1 \otimes 20$, $1 \otimes 01$, $0 \otimes 02$
P_{14}	$C_2(q)$	10, 01, 00^2
P_{23}	$A_1(q)^2$	$1 \otimes 0$, $0 \otimes 1$

Hence, we can certainly find two elements $u_1, u_2 \in Q$ such that u_1Q', u_2Q' do not both lie in a proper R-invariant subgroup of Q/Q'. As usual, it follows that $d(H_0) \leq 2 + d(R)$. Finally, we see that $d(R) = 2$ in the usual way, so $d(H_0) \leq 4$ as required.

Next consider $G_0 = {}^2F_4(q)'$. If $q = 2$ we check that $d(H) = 2$ using MAGMA and the Web-Atlas [57], so assume $q > 2$. Write $H_0 = QR$ as usual, so that $R_0 = \mathrm{SL}_2(q)$ or ${}^2B_2(q)$. The structure of H_0 is given by [19, §10]. When $R_0 = \mathrm{SL}_2(q)$ we have $|Q/Q'| = q^2$, and Q/Q' is the natural module for R_0; and when $R_0 = {}^2B_2(q)$, Q/Q' has order q^5 and composition factors of dimensions 1 and 4 as R_0-modules. Hence as before, $d(H_0) \leq 1 + d(R)$, and now the usual argument gives the conclusion.

Next let $G_0 = G_2(q)$. Here we use the commutator relations for G_2 given in [52, p.443]. First assume that $H_0 = QR = P_1$ or P_2. If $p = 2$ then for the short parabolic P_2 (i.e. R_0 a short $A_1(q)$), Q/Q' is an irreducible R-module, while for the long parabolic P_1, Q/Q' is an extension of a trivial module by an irreducible 2-dimensional R-module. And if $p = 3$ then for both P_1 and P_2, Q/Q' is an extension of an irreducible 2-dimensional R-module by a twist of itself. Hence as usual we see that $d(H_0) \leq 2 + d(R)$. Since $d(R) = 2$ the result follows.

Now suppose $G_0 = G_2(q)$, $p = 3$, $H_0 = QR$ is a Borel subgroup and G contains a graph automorphism. From the commutator relations one checks that Q/Q' is generated by 3 root groups modulo Q'. Also $R = (q-1)^2$ acts as a full group of scalars on each of the root groups and it follows in the usual way that $d(H_0) \leq 4$.

Finally, for $G_0 = {}^2G_2(q)$ or ${}^2B_2(q)$, we see from [56], [55] that $|Q/Q'| = q$ and $R = Z_{q-1}$ acts faithfully on Q/Q', so again the usual argument goes through. □

Next we deal with the last excluded case of Lemma 7.1.

Lemma 7.4. *Suppose that $G_0 = D_4(q)$, G contains a triality automorphism, and $H_0 = P_{134}$. Then $d(H_0) \leq 4$.*

Proof. We check this for $q \leq 3$ using MAGMA, so let us assume $q > 3$. Working with $G_0 = \Omega_8^+(q)$ and $H_0 = QR$ as usual, we have

$$R = \{(A, \alpha, \beta) \in \mathrm{GL}_2(q) \times \mathbb{F}_q^* \times \mathbb{F}_q^* \mid \det(A)\alpha\beta \text{ is a square in } \mathbb{F}_q\},$$

whence $d(R) \leq 3$ by Proposition 2.10. As a module for $R_0 = \mathrm{SL}_2(q)$ we have $Q/Q' = V_1 + V_2 + V_3$, a sum of three copies of the natural module, where the V_i are generated by the following root groups:

$$V_1 = \langle U_{1000}, U_{1100}\rangle Q',\ V_2 = \langle U_{0010}, U_{0110}\rangle Q',\ V_3 = \langle U_{0001}, U_{0101}\rangle Q'.$$

One checks that the vector $U_{1000}(1)U_{0010}(1)U_{0001}(1)Q'$ generates Q/Q' under the action of R. Hence $d(H_0) \leq 4$. □

The proof of Theorem 2 for parabolic subgroups is completed by

Lemma 7.5. *If H is a maximal parabolic subgroup of the almost simple group G, then $d(H) \leq 6$.*

Proof. We have already proved that $d(H_0) = d(H \cap G_0) \leq 4$, so the result is automatic if $d(G/G_0) \leq 2$. Hence we may assume that $d(G/G_0) = 3$. The possibilities for G are described in Proposition 2.1(i): $G_0 = \mathrm{L}_{2m}(q)$, $\mathrm{P}\Omega_{2m}^{\epsilon}(q)$ $(m \geq 5)$ or $\mathrm{P}\Omega_8^+(q)$, with q odd and square, and G/G_0 has an image 2^3. As before, write $H = QR$, where Q is the unipotent radical and R a Levi subgroup. As in Lemma 7.1 we have $d(H) \leq 1 + d(R)$, so we need to show that $d(R) \leq 5$.

If $G_0 = \mathrm{L}_{2m}(q)$ then $H = P_{i,2m-i}$ or P_m and we argue in similar fashion to the proof of Lemma 7.2. Writing $I = \mathrm{PGL}_{2m}(q)$, we have $d(G/G \cap I) \leq 2$, so it is enough to show that $d(R \cap I) \leq 3$. For $P_{i,2m-i}$ we have

$$R \cap I = \{(A, B, C) \in \mathrm{GL}_{2m-2i}(q) \times \mathrm{GL}_i(q)^2 \mid \det(ABC) \in \langle \mu^k \rangle\},$$

modulo scalars, for some k (recall that μ is a generator of $\mathbb{F}_q^*$). As in the proof of Lemma 7.2, write $\mathrm{GL}_{2m-2i}(q) = \langle x, u\rangle$ and $\mathrm{GL}_i(q) = \langle y, v\rangle$, where $\det(x) = \det(y) = \mu$. One checks that $R \cap I$ is generated by the three elements (x, y^{k-1}, v), (x^{k-1}, v, y), (u, y, y^{k-1}). This gives the result for $P_{i,2m-i}$, and the P_m case is similar.

Next consider $G_0 = \mathrm{P}\Omega^{\epsilon}_{2m}(q)$ $(m \geq 5)$. Here G/G_0 is a 3-generator subgroup of $D_8 \times Z_f$ where $q = p^f$ (see Proposition 2.1(i)). Let $I = \mathrm{PO}^{\epsilon}_{2m}(q) = G_0.2^2$. Then I is normal in $\mathrm{Aut}(G_0)$ and $\mathrm{Aut}(G_0)/I \cong Z_2 \times Z_f$, so it is enough to show that $d(G \cap I) \leq 3$.

There are five possibilities for the group $G \cap I$: they are G_0, I, $\mathrm{PSO}^{\epsilon}_{2m}(q)$, $G_0\langle r_1\rangle$ and $G_0\langle r_2\rangle$, where r_1, r_2 are reflections in non-singular vectors of square, non-square norm, respectively. We deal with each of these possibilities in similar fashion to the proof of Lemma 7.2. We have $R \leqslant \mathrm{GL}_i(q) \times O^{\epsilon}_{2m-2i}(q)$ (modulo scalars). Write $\mathrm{GL}_i(q) = \langle x, u\rangle$ with x semisimple and u unipotent. Then we can find generators a, b, c for the projection of R to $O^{\epsilon}_{2m-2i}(q)$ such that $(x, a), (u, b), (1, c)$ generate R.

Finally consider $G_0 = \mathrm{P}\Omega^+_8(q)$. If there is no triality automorphism involved in G, then $G/G_0 \leqslant D_8 \times Z_f$ and we argue as above. Otherwise, G/G_0 is a subgroup of $S_4 \times Z_f$ containing a triality, and such a subgroup is 2-generator. This completes the proof. □

This completes the proof of Theorem 2 for parabolic subgroups. Moreover, in view of the results of the previous sections, Theorem 2 is now proved.

8. Random generation

Recall that if G is a finite group then we denote by $\nu(G)$ the minimal number k such that the probability that G is generated by k random elements is at least $1/e$. By an observation of Pak [50], this coincides (up to a small multiplicative constant) with the expected number of random elements generating G. It is known that there exists an absolute constant c such that $\nu(G) \leq c$ for any finite simple group G (indeed, by the main theorem of [37], $\nu(G) = 2$ if $|G|$ is sufficiently large). Here we establish Theorem 3, which provides an extension of this result to maximal subgroups of almost simple groups.

In addition to Theorem 2, the main ingredient in the proof of Theorem 3 is a remarkably explicit bound on $\nu(G)$ due to Jaikin-Zapirain and Pyber, which applies to any finite group G. In order to state this result, we first require some notation. For a non-abelian characteristically simple group A, let $\mathrm{rk}_A(G)$ be the maximal number r such that a normal section of G is the direct product of r chief factors of G isomorphic to A. In addition, let $\ell(A)$ be the minimal degree of a faithful transitive permutation representation of A.

Theorem 8.1 ([26, Theorem 1])**.** *There exist absolute constants $0 < \alpha < \beta$ such that for any finite group G*

$$\alpha \left(d(G) + \max_A \left\{ \frac{\log(\mathrm{rk}_A(G))}{\log(\ell(A))} \right\} \right) < \nu(G) < \beta d(G) + \max_A \left\{ \frac{\log(\mathrm{rk}_A(G))}{\log(\ell(A))} \right\},$$

where A runs through the non-abelian chief factors of G.

Let G be an almost simple group and let H be a maximal subgroup of G. By Theorem 2 we have $d(H) \leq 6$, so in order to prove Theorem 3 it suffices to show that

$$\delta(H) := \max_A \left\{ \frac{\log(\mathrm{rk}_A(H))}{\log(\ell(A))} \right\} \tag{5}$$

is bounded above by an absolute constant, where A runs through the non-abelian chief factors of H.

Lemma 8.2. *Let G be a finite almost simple group and let H be a maximal subgroup of G. Then H has at most three non-abelian chief factors.*

Proof. Let G_0 be the socle of G and let $\gamma(H)$ denote the number of non-abelian chief factors of H. If H is solvable or almost simple then $\gamma(H) \leq 1$, so assume otherwise. If G_0 is a sporadic group then the possibilities for H are conveniently recorded in the Web Atlas [57] and it is easy to check that $\gamma(H) \leq 2$. If G_0 is an alternating group then the maximal subgroups of G are described by the O'Nan-Scott theorem (see Theorem 4.1), and the same conclusion quickly follows. For example, if H is of type $S_k \wr S_t$ then $\gamma(H) \leq 2$, with equality if and only if $k, t \geq 5$.

Now assume G_0 is a classical group. Here H belongs to one of the eight $\mathcal{C}_i$ families that arise in Aschbacher's theorem on the subgroup structure of classical groups (see Table 1 and [1]). If $H \in \mathcal{C}_3 \cup \mathcal{C}_5 \cup \mathcal{C}_6 \cup \mathcal{C}_8$ then the bound $\gamma(H) \leq 2$ is clear. Similarly, if $H \in \mathcal{C}_4$ then $\gamma(H) \leq 2$ unless $G_0 = \mathrm{P}\Omega_n^+(q)$ and H is of type $O_4^+(q) \otimes O_{n/4}^{\epsilon}(q)$ with $q \geq 5$ odd, in which case $\gamma(H) \leq 3$. Next suppose H is a reducible subgroup in the $\mathcal{C}_1$ collection. If H is non-parabolic then either $\gamma(H) \leq 2$, or H is of type $O_4^+(q) \perp O_{n-4}^{\epsilon'}(q)$ with $q \geq 4$ and $\gamma(H) \leq 3$. Similarly, if H is a parabolic subgroup of G then by inspecting the structure of H given in [29, Section 4.1] we deduce that $\gamma(H) \leq 2$ unless $G_0 = \mathrm{P}\Omega_n^+(q)$ and H is of type $P_{n/2-2}$ (with $q \geq 4$), or $G_0 = \mathrm{L}_n(q)$ and H is of type $P_{m,n-m}$ with $2 \leq m < n/2$ and $(m, q) \neq (2, 2), (2, 3)$. In both of these cases it is clear that $\gamma(H) \leq 3$, as required. Finally, suppose $H \in \mathcal{C}_2 \cup \mathcal{C}_7$. If H is a $\mathcal{C}_2$-subgroup of type $O_4^+(q) \wr S_t$ with $t \geq 5$ and $q \geq 4$ then up to isomorphism the collection of non-abelian chief factors of H is either $\{A_t, \mathrm{L}_2(q)^{2t}\}$ or $\{A_t, \mathrm{L}_2(q)^t, \mathrm{L}_2(q)^t\}$, so $\gamma(H) \leq 3$. In each of the remaining cases, it is easy to see that $\gamma(H) \leq 2$. For example, if H is of type $O_4^+(q) \wr S_2$ then H contains an element interchanging the two factors of type $O_4^+(q)$, so either $\mathrm{L}_2(q)^2$ is a minimal normal subgroup of H (and thus $\gamma(H) = 2$), or $\mathrm{L}_2(q)^4$ has this property, in which case $\gamma(H) = 1$.

Finally, let us assume G_0 is an exceptional group of Lie type. The possibilities for H are described in Proposition 6.1 (in addition to the parabolic subgroups), and by inspection we see that $\gamma(H) \leq 3$. □

Remark 8.3. There are examples with $\gamma(H) = 3$. For instance, if $G = \mathrm{P}\Omega_{4m}^+(q)$ and H is a $\mathcal{C}_4$-subgroup of type $O_4^+(q) \otimes O_m(q)$, where qm is odd and $q \geq 5$, then

$$H \cong \mathrm{L}_2(q) \times \mathrm{L}_2(q) \times \mathrm{SO}_m(q)$$

(see [29, 4.4.17]), so the non-abelian chief factors of H are $\mathrm{L}_2(q)$, $\mathrm{L}_2(q)$ and $\Omega_m(q)$.

Corollary 8.4. *Let H be a maximal subgroup of a finite almost simple group. Then $\delta(H) < 1$.*

Proof. By Lemma 8.2 we have $\mathrm{rk}_A(H) \leq 3$ for every non-abelian chief factor A of H. Since $\ell(A) \geq 5$, the result follows. □

By combining Corollary 8.4 with Theorems 2 and 8.1 we obtain the following corollary, which completes the proof of Theorem 3.

Corollary 8.5. *Let G be an almost simple group and let H be a maximal subgroup of G. Then $\nu(H) < 6\beta + 1$, where β is the absolute constant appearing in the statement of Theorem 8.1.*

Finally, let us turn to Corollary 4. For a finite group G and a positive integer k recall that $P(G, k)$ denotes the probability that k randomly chosen elements of G generate G, so $\nu(G)$ is the minimal number k such that $P(G, k) \geq 1/e$. Let $Q(G, k) = 1 - P(G, k)$ be the complementary probability, so

$$Q(G, k) = \frac{|\{(x_1, \ldots, x_k) \in G^k \mid \langle x_1, \ldots, x_k \rangle \neq G\}|}{|G|^k}$$

and we see that $Q(G, kc) \leq Q(G, c)^k$ for all positive integers k and c.

Fix $\epsilon > 0$ and let c be the positive integer in the statement of Theorem 3. Let H be a maximal subgroup of an almost simple group, and let k be the minimal positive integer such that $(1 - 1/e)^k < \epsilon$. Then

$$Q(H, kc) \leq Q(H, c)^k \leq (1 - 1/e)^k < \epsilon$$

and thus $P(H, kc) > 1 - \epsilon$. This completes the proof of Corollary 4.

9. Maximal subgroup growth

Let G be a group and let $m_n(G)$ denote the number of maximal subgroups of index n in G. Recall that G has *polynomial maximal subgroup growth* if $m_n(G) \leq n^c$ for all n, where c is some constant. For example, finite simple groups have this property in the strong sense that there exists an absolute constant c such that $m_n(G) \leq n^c$ for all n and all finite simple groups G. In fact, the main theorem of [32] establishes an even stronger result, namely that if G is simple then $m_n(G) \leq n^a$ for any fixed $a > 1$ and sufficiently large n.

A *second maximal* subgroup of a group G is a maximal subgroup of a maximal subgroup of G. Let $m_n^2(G)$ denote the number of second maximal subgroups of index n in G. Our aim here is to show that $m_n^2(G)$ grows polynomially when G is almost simple, proving Corollary 6. To do this, we combine Corollary 5 with the following lemma, which establishes the analogous property for maximal subgroups.

Lemma 9.1. *There exists an absolute constant c such that any finite almost simple group has at most n^c maximal subgroups of index n.*

Proof. This quickly follows from Theorem 8.1. Let G be an almost simple group and let n be a positive integer. Since $d(G) \leq 3$ and $\delta(G) = 0$ (see Proposition 2.1(i) and (5)), the upper bound in Theorem 8.1 yields $\nu(G) < 3\beta$ and thus $m_n(G) \leq n^{3\beta+4}$ by [42, 1.2].

For completeness we also give an alternative, more elementary argument, which is independent of Theorem 8.1. Write

$$m_n(G) = \alpha_n(G) + \beta_n(G)$$

where $\alpha_n(G)$ (respectively $\beta_n(G)$) denotes the number of maximal subgroups of index n in G with trivial core (respectively, non-trivial core). Note that $\beta_n(G) = m_n(G/G_0)$, where G_0 is the socle of G. By [27, 37, 39] we have $\alpha_n(G) = o(n^2)$ (in fact better bounds hold). We deduce that $\alpha_n(G) \leq n^{c_1}$ for some absolute constant c_1. In addition, by considering the various possibilities for G_0, we see that every subgroup of G/G_0 is a 3-generator solvable group of derived length at most 3. Therefore, the number of subgroups of index n in G/G_0 is at most n^{c_2} for some absolute constant c_2, so $m_n(G/G_0) \leq n^{c_2}$ and the result follows. □

The proof of Corollary 6 is an easy combination of Lemma 9.1 and Corollary 5. Indeed, if G is almost simple and H is a second maximal subgroup of G of index n, then there exists a divisor a of n and a maximal subgroup M of G of index a containing H, such that H is a maximal subgroup of M of index n/a. This yields

$$m_n^2(G) \leq \sum_{a|n} a^{c_1}(n/a)^{c_2} \leq n^{c_1+c_2+1},$$

where c_1 and c_2 are the absolute constants in Lemma 9.1 and Corollary 5, respectively.

10. Primitive permutation groups

In this final section we prove Theorems 7 and 8. Let G be a primitive permutation group on a finite set Ω with point stabilizer $H = G_\alpha$. By the O'Nan-Scott theorem (see [16, Theorem 4.1A]), one of the following holds:

(i) G is almost simple;

(ii) G has a regular minimal normal subgroup N;

(iii) G is of simple diagonal type;

(iv) G is of product type: here $G \leqslant J \wr S_l$ acting with product action on a Cartesian product $\Omega = \Gamma^l$, where J is primitive on Γ of almost simple or simple diagonal type. Moreover, T^l is the socle of G, where T is the socle of J.

Note that if (ii) fails to hold then G has a unique minimal normal subgroup.

10.1. **Proof of Theorem 7.** The lower bound $d(G) - 1 \leq d(H)$ is trivial since H is maximal in G. To establish the upper bound, we consider each of the above four cases in turn. In case (i) we have $d(G_\alpha) \leq 6$ by Theorem 2, and the conclusion of Theorem 7 follows. In case (ii), $G = G_\alpha N$ with $N \cap G_\alpha = 1$, so $G/N \cong G_\alpha$ and thus $d(G_\alpha) = d(G/N) \leq d(G)$.

Now consider case (iii). Let B be the socle of G. Then $B \cong T^k$ and $B_\alpha \cong T$, where T is a non-abelian simple group and $k \geq 2$. Moreover $G = G_\alpha B$, so $G/B \cong G_\alpha/B_\alpha$. Since $B_\alpha \cong T$, it is a minimal normal subgroup of G_α, whence $d(G_\alpha) \leq d(G_\alpha/B_\alpha) + 1$ by Proposition 2.1(ii). Hence

$$d(G_\alpha) \leq d(G_\alpha/B_\alpha) + 1 = d(G/B) + 1 \leq d(G) + 1.$$

Finally, let us consider case (iv). Suppose first that J is almost simple, with socle T, and let $B = T^l$ be the socle of G. As above, $G/B \cong G_\alpha/B_\alpha$, and this group acts transitively on the l factors in B. Let $\gamma \in \Gamma$ and take $\alpha = (\gamma, \ldots, \gamma) \in \Gamma^l = \Omega$. Then $B_\alpha = T_\gamma^l$. Since G_α/B_α acts transitively on the l factors of B_α, it follows that G_α is generated by T_γ together with coset representatives of generators of G_α/B_α, and hence

$$d(G_\alpha) \leq d(T_\gamma) + d(G_\alpha/B_\alpha) = d(T_\gamma) + d(G/B) \leq d(T_\gamma) + d(G).$$

The result follows since $d(T_\gamma) \leq 4$ by Theorem 2.

Now suppose that (J, Γ) is of simple diagonal type. As before, let T and B be the socles of J and G, respectively. Let $\gamma \in \Gamma$ and set $\alpha = (\gamma, \ldots, \gamma) \in \Gamma^l = \Omega$. Then $T = S^k$ with $S \cong T_\gamma$ non-abelian simple, and $B = T^l = S^{kl}$. As above, $G/B \cong G_\alpha/B_\alpha$ acts transitively on the l factors in $B = T^l$, whence

$$d(G_\alpha) \leq d(T_\gamma) + d(G_\alpha/B_\alpha) = d(S) + d(G/B) \leq d(G) + 2$$

and the proof of Theorem 7 is complete.

10.2. **Proof of Theorem 8.** We begin with a couple of preliminary lemmas. Our first result follows immediately from the definition of $\delta(G)$ (see (5)).

Lemma 10.1. *Let G be a finite group and let N be a minimal normal subgroup of G. Then $\delta(G/N) \leq \delta(G) < \delta(G/N) + 1$.*

Lemma 10.2. *Let G be a finite primitive permutation group with point stabilizer H. Then*

$$\delta(G) - 1 < \delta(H) < \delta(G) + 1.$$

Proof. We consider each of the primitive groups of type (i)–(iv) in turn. In case (i), $\delta(G) = 0$ and the result follows from Lemma 8.2. In (ii), G has a minimal normal subgroup N such that $G/N \cong H$, so in this case the result follows from Lemma 10.1. For the remainder we may assume (ii) fails to hold, in which case G has a unique minimal normal subgroup.

If G is of simple diagonal type then the socle of G is of the form $B = T^k$ for a non-abelian simple group T and again the result follows from Lemma 10.1 since $G/B \cong H/T$, where B (respectively T) is a minimal normal subgroup of G (respectively H).

Finally, let us assume G is of product type as in (iv), so $G \leqslant J \wr S_l$ has the product action on $\Omega = \Gamma^l$, and $J \leqslant \mathrm{Sym}(\Gamma)$ is primitive of almost simple or simple diagonal type. Let T denote the socle of J. Then $B = T^l$ (the socle of G) is a minimal normal subgroup of G and we have $G/B \cong H/(H \cap B)$. If J is of simple diagonal type then $H \cap B$ is a minimal normal subgroup of H and the result follows via Lemma 10.1 as before.

Now assume J is almost simple. As in the proof of Theorem 7 we have $H = G_\alpha$ with $\alpha = (\gamma, \ldots, \gamma) \in \Gamma^l = \Omega$, and $H \cap B = B_\alpha = (T_\gamma)^l$. Since $G/B \cong H/(H \cap B)$ acts transitively on the l factors in B, it follows that any non-abelian chief factor of H occurring as a section of $H \cap B$ is of the form $L/K \times \cdots \times L/K$ (l factors), where L/K is a non-abelian chief factor of T_γ. By Lemma 8.2 there are at most 3 possibilities for L/K, so $\delta(H) < \delta(H/(H \cap B)) + 1$ and the desired result quickly follows. □

Corollary 10.3. *Let G be a finite primitive permutation group with point stabilizer H. Then*

$$\nu(H) < \beta\alpha^{-1}\nu(G) + 4\beta + 1 \text{ and } \nu(G) < \beta\alpha^{-1}\nu(H) + \beta + 1,$$

where α and β are the absolute constants in the statement of Theorem 8.1.

Proof. This is an easy application of Theorems 7 and 8.1, together with Lemma 10.2. For the first bound,

$$\nu(H) < \beta d(H) + \delta(H) < \beta(d(G) + 4) + \delta(G) + 1 \leq \beta\alpha^{-1} \cdot \alpha(d(G) + \delta(G)) + 4\beta + 1$$

since we may assume $\beta > 1$, and the result follows since the lower bound in Theorem 8.1 gives $\alpha(d(G) + \delta(G)) < \nu(G)$. To establish the second bound we use the fact that $d(G) \leq d(H) + 1$, so

$$\nu(G) < \beta d(G) + \delta(G) < \beta(d(H) + 1) + \delta(H) + 1 \leq \beta\alpha^{-1} \cdot \alpha(d(H) + \delta(H)) + \beta + 1$$

and once again the result follows by applying the lower bound in Theorem 8.1. □

Theorem 8 follows immediately from Corollary 10.3. Indeed, since $\nu(G), \nu(H) \geq 1$, we deduce that

$$(\beta\alpha^{-1} + \beta + 1)^{-1}\nu(G) < \nu(H) < (\beta\alpha^{-1} + 4\beta + 1)\nu(G).$$

References

[1] M. Aschbacher, *On the maximal subgroups of the finite classical groups*, Invent. Math. **76** (1984), 469–514.

[2] M. Aschbacher and R. Guralnick, *Some applications of the first cohomology group*, J. Algebra **90** (1984), 446–460.

[3] H. Azad, M. Barry, and G.M. Seitz, *On the structure of parabolic subgroups*, Comm. in Alg. **18** (1990), 551–562.

[4] L. Babai, P.J. Cameron and P.P. Pálfy, *On the orders of primitive groups with restricted nonabelian composition factors*, J. Algebra **79** (1982), 161–168.

[5] W. Bosma, J. Cannon and C. Playoust, *The* MAGMA *algebra system I: The user language*, J. Symbolic Comput. **24** (1997), 235–265.

[6] A. Borovik, L. Pyber and A. Shalev, *Maximal subgroups in finite and profinite groups*, Trans. Amer. Math. Soc. **348** (1996), 3745–3761.

[7] J.N. Bray, I.A.I. Suleiman, P.G. Walsh and R.A. Wilson, *Generating maximal subgroups of sporadic simple groups*, Comm. in Alg. **29** (2001), 1325–1337.

[8] J.L. Brenner and J. Wiegold, *Two-generator groups* I, Michigan Math. J. **22** (1975), 53–64.

[9] T. Breuer, R.M. Guralnick and W.M. Kantor, *Probabilistic generation of finite simple groups,* II, J. Algebra **320** (2008), 443–494.

[10] T.C. Burness and S. Guest, *On the uniform spread of almost simple linear groups*, Nagoya Math. J. **109** (2013), 35–109.

[11] T.C. Burness, E.A. O'Brien, and R.A. Wilson, *Base sizes for sporadic simple groups*, Israel J. Math. **177** (2010), 307–334.

[12] A.M. Cohen, M.W. Liebeck, J. Saxl, and G.M. Seitz, *The local maximal subgroups of exceptional groups of Lie type*, Proc. London Math. Soc. **64** (1992), 21–48.

[13] J. Conway, R. Curtis, S. Norton, R. Parker, and R. Wilson, *Atlas of Finite Groups*, Oxford University Press, 1985.

[14] F. Dalla Volta and A. Lucchini, *Generation of almost simple groups*, J. Algebra **178** (1995), 194–223.

[15] J.D. Dixon, *The probability of generating the symmetric group*, Math. Z. **110** (1969), 199–205.

[16] J.D. Dixon and B. Mortimer, *Permutation Groups*, Graduate Texts in Mathematics, vol. 163, Springer-Verlag, 1996.

[17] A.G. Earnest and H. Ishibashi, *Two-element generation of orthogonal groups over finite fields*, J. Algebra **165** (1994), 164–171.

[18] A.G. Earnest and H. Ishibashi, *Addendum: Two-element generation of orthogonal groups over finite fields*, J. Algebra **182** (1996), 805.

[19] P. Fong and G.M. Seitz, *Groups with a (B,N)-pair of rank 2, II*, Invent. Math. **24** (1974), 191–239.

[20] D. Gorenstein, R. Lyons, and R. Solomon, *The Classification of the Finite Simple Groups, Number 3*, Mathematical Surveys and Monographs, vol. 40, Amer. Math. Soc., 1998.

[21] R.M. Guralnick and W.M. Kantor, *Probabilistic generation of finite simple groups*, J. Algebra **234** (2000), 743–792.

[22] R.M. Guralnick and A. Shalev, *On the spread of finite simple groups*, Combinatorica **23** (2003), 73–87.

[23] P. Hall, *The Eulerian functions of a group*, Quarterly J. Math. **7** (1936), 134–151.

[24] D.F. Holt and C.M. Roney-Dougal, *Constructing maximal subgroups of classical groups*, LMS J. Comput. Math. **8** (2005), 46–79.

[25] D.F. Holt and C.M. Roney-Dougal, *Constructing maximal subgroups of orthogonal groups*, LMS J. Comput. Math. **13** (2010), 164–191.

[26] A. Jaikin-Zapirain and L. Pyber, *Random generation of finite and profinite groups and group enumeration*, Annals of Math. **173** (2011), 769–814.

[27] W.M. Kantor and A. Lubotzky, *The probability of generating a finite classical group*, Geom. Dedicata **36** (1990), 67–87.

[28] P.B. Kleidman, *The maximal subgroups of the finite 8-dimensional orthogonal groups* $\mathrm{P}\Omega_8^+(q)$ *and of their automorphism groups*, J. Algebra **110** (1987), 173–242.

[29] P.B. Kleidman and M.W. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, London Math. Soc. Lecture Note Series, vol. 129, Cambridge University Press, 1990.

[30] P.B. Kleidman and M.W. Liebeck, *A survey of the maximal subgroups of the finite simple groups*, Geom. Dedicata **25** (1988), 375–389.

[31] M.W. Liebeck, *Subgroups of simple algebraic groups and of related finite and locally finite groups of Lie type*, Finite and locally finite groups (Istanbul, 1994), 71–96, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 471, Kluwer Acad. Publ., Dordrecht, 1995.

[32] M.W. Liebeck, B.M.S. Martin and A. Shalev, *On conjugacy classes of maximal subgroups of finite simple groups, and a related zeta function*, Duke Math. Journal **128** (2005), 541–557.

[33] M.W. Liebeck, C.E. Praeger, and J. Saxl, *A classification of the maximal subgroups of the finite alternating and symmetric groups*, J. Algebra **111** (1987), 365–383.
[34] M.W. Liebeck, J. Saxl, and G.M. Seitz, *Subgroups of maximal rank in finite exceptional groups of Lie type*, Proc. London Math. Soc. **65** (1992), 297–325.
[35] M.W. Liebeck and G.M. Seitz, *Maximal subgroups of exceptional groups of Lie type, finite and algebraic*, Geom. Dedicata **36** (1990), 353–387.
[36] M.W. Liebeck and G.M. Seitz, *A survey of the maximal subgroups of exceptional groups of Lie type*, Groups, combinatorics & geometry (Durham, 2001), 147–154, World Sci. Publ., Rier Edge, NJ, 2003.
[37] M.W. Liebeck and A. Shalev, *The probability of generating a finite simple group*, Geom. Dedicata **56** (1995), 103–113.
[38] M.W. Liebeck and A. Shalev, *Maximal subgroups of symmetric groups*, J. Comb. Th. Ser. A **75** (1996), 341–352.
[39] M.W. Liebeck and A. Shalev, *Classical groups, probabilistic methods, and the* $(2,3)$*-generation problem*, Annals of Math. **144** (1996), 77–125.
[40] M.W. Liebeck and A. Shalev, *Random* (r,s)*-generation of finite classical groups*, Bull. London Math. Soc. **34** (2002), 185–188.
[41] F. Lübeck and G. Malle, $(2,3)$*-generation of exceptional groups*, J. London Math. Soc. **59** (1999), 109–122.
[42] A. Lubotzky, *The expected number of random elements to generate a finite group*, J. Algebra **257** (2002), 452–459.
[43] A. Lucchini, *Generators and minimal normal subgroups*, Arch. Math. **64** (1995), 173–276.
[44] A. Lucchini and F. Menegazzo, *Generators for finite groups with a unique minimal normal subgroup*, Rend. Sem. Mat. Univ. Padova **98** (1997), 173–191.
[45] A. Mann, *Positively finitely generated groups*, Forum Math. **8** (1996), 429–459.
[46] A. Mann and A. Shalev, *Simple groups, maximal subgroups, and probabilistic aspects of profinite groups*, Israel J. Math. **96** (1996), 449–468.
[47] G.A. Miller, *On the groups generated by two operators*, Bull. Amer. Math. Soc. **7** (1901), 424–426.
[48] E. Netto, *Substitutionentheorie und ihre Anwendungen auf die Algebra*, Teubner, Leipzig, 1882; English transl. 1892, second edition, Chelsea, New York, 1964.
[49] S.P. Norton and R.A. Wilson, *A correction to the* 41*-structure of the Monster, a construction of a new maximal subgroup* $\mathrm{L}_2(41)$*, and a new Moonshine phenomenon*, J. London Math. Soc., to appear.
[50] I. Pak, *On probability of generating a finite group*, preprint (1999).
[51] R. Ree, *A family of simple groups associated with the simple Lie algebra of type* $(F4)$, Amer. J. Math. **83** (1961), 401–420.
[52] R. Ree, *A family of simple groups associated with the simple Lie algebra of type* $(G2)$, Amer. J. Math. **83** (1961), 432–462.
[53] A. Shalev, *Random generation of finite simple groups by p-regular or p-singular elements*, Israel J. Math. **125** (2001), 53–60.
[54] R. Steinberg, *Generators for simple groups*, Canad. J. of Math. **14** (1962), 277–283.
[55] M. Suzuki, *On a class of doubly transitive groups*, Annals of Math. **75** (1962), 105–145.
[56] H.N. Ward, *On Ree's series of simple groups*, Trans. Amer. Math. Soc. **121** (1966), 62–89.
[57] R.A. Wilson et al., *A World-Wide-Web Atlas of finite group representations*, `http://brauer.maths.qmul.ac.uk/Atlas/v3/`.

School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK
E-mail address: `t.burness@soton.ac.uk`

Department of Mathematics, Imperial College, London SW7 2BZ, UK
E-mail address: `m.liebeck@imperial.ac.uk`

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
E-mail address: `shalev@math.huji.ac.il`