Forum of Mathematics, Sigma (2017), vol. , e1, 40 pages
doi:

GENERATION OF SECOND MAXIMAL SUBGROUPS AND THE EXISTENCE OF SPECIAL PRIMES

TIMOTHY C. BURNESS[1], MARTIN W. LIEBECK[2] and ANER SHALEV[3]

[1] School of Mathematics, University of Bristol, Bristol BS8 1TW, UK;
t.burness@bristol.ac.uk
[2] Department of Mathematics, Imperial College, London SW7 2BZ, UK;
m.liebeck@imperial.ac.uk
[3] Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel;
shalev@math.huji.ac.il

Received 4 February 2017

Abstract

Let G be a finite almost simple group. It is well known that G can be generated by 3 elements, and in previous work we showed that 6 generators suffice for all maximal subgroups of G. In this paper we consider subgroups at the next level of the subgroup lattice – the so-called second maximal subgroups. We prove that with the possible exception of some families of rank 1 groups of Lie type, the number of generators of every second maximal subgroup of G is bounded by an absolute constant. We also show that such a bound holds without any exceptions if and only if there are only finitely many primes r for which there is a prime power q such that $(q^r - 1)/(q - 1)$ is prime. The latter statement is a formidable open problem in Number Theory. Applications to random generation and polynomial growth are also given.

2010 Mathematics Subject Classification: 20D06 (primary); 20D30, 20P05 (secondary)

1. Introduction

In recent years it has been shown that finite non-abelian simple groups share several fundamental generation properties with their maximal subgroups. For example, both classes can be generated by a small number of elements – the simple groups by 2 elements [3, 32], and their maximal subgroups by 4 elements

[8]. Similarly, both simple groups and their maximal subgroups are randomly generated by boundedly many elements [8, 25]. Analogous results also hold for almost simple groups – that is, groups lying between a non-abelian finite simple group and its automorphism group. These groups are generated by 3 elements [13] and their maximal subgroups by 6 elements [8].

In this paper we investigate analogous questions for subgroups lying deeper in the subgroup lattice of an almost simple group – namely, for second maximal subgroups. We show, somewhat surprisingly, that the question of whether these subgroups are generated by a bounded number of elements is equivalent to a formidable open problem in Number Theory – namely, the existence of primes of the form $\frac{q^r-1}{q-1}$ where r is arbitrarily large and q is a prime power (which may depend on r).

For a finite group G, let $d(G)$ be the minimal number of generators of G. Define the *depth* of a subgroup M of G to be the minimal length of a chain of subgroups

$$G = G_0 > G_1 > \cdots > G_{t-1} > G_t = M,$$

where each G_i is a maximal subgroup of G_{i-1}. A subgroup is *second maximal* if it has depth 2. There has been interest in the study of these subgroups and their overgroups in the context of lattice theory; this includes work of Feit [15], Pálfy [31] and Aschbacher [1]. In addition, the PhD thesis of Basile [5] provides a detailed study of second maximal subgroups of symmetric and alternating groups.

Our first result concerns the number of generators required for second maximal subgroups of almost simple groups.

Theorem 1. *Let G be a finite almost simple group with socle G_0, and let M be a second maximal subgroup of G. Then one of the following holds:*

(i) $d(M) \leqslant 12$*;*

(ii) $d(M) \leqslant 70$*, G_0 is exceptional of Lie type, and M is maximal in a parabolic subgroup of G;*

(iii) $G_0 = \mathrm{L}_2(q)$*, ${}^2B_2(q)$ or ${}^2G_2(q)$, and M is maximal in a Borel subgroup of G.*

The bounds 12 and 70 in parts (i) and (ii) are probably not best possible (see Remark 7.5). In part (iii), $d(M)$ can be enormously large. For example, if $G = \mathrm{L}_2(2^k)$ and $2^k - 1$ is a prime, then the elementary abelian 2-group $M = (Z_2)^k$ is a second maximal subgroup of G requiring k generators. Since the largest currently known prime is a Mersenne prime with $k = 74207281$, we obtain the following.

Proposition 2. *There exists a second maximal subgroup M of a finite simple group such that* $d(M) = 74207281$.

The question of whether $d(M)$ can be arbitrarily large for the groups in part (iii) of Theorem 1 turns out to depend on the open problem in Number Theory mentioned above:

$$\text{\textit{Are there infinitely many primes } } r \text{\textit{ for which there}} \\ \text{\textit{exists a prime power} } q \text{ \textit{such that} } \tfrac{q^r-1}{q-1} \text{ \textit{is prime?}} \tag{1}$$

This would follow, for example, if there exist infinitely many Mersenne primes – but note that in (1), q may be arbitrarily large and may depend on r. It is believed that question (1) has a positive answer. However, existing methods of Number Theory are far from proving this.

We establish the following.

Theorem 3. *The following are equivalent.*

(i) *There exists a constant c such that all second maximal subgroups of finite almost simple groups are generated by at most c elements.*

(ii) *There exists a constant c such that all second maximal subgroups of finite simple groups are generated by at most c elements.*

(iii) *There exists a constant c such that all second maximal subgroups of* $\mathrm{L}_2(q)$ *(q a prime power) are generated by at most c elements.*

(iv) *The question (1) has a negative answer.*

In view of the difficulty of question (1), it seems likely that the validity of part (i) of Theorem 3 will remain open for a long time. However, if we go further down the subgroup lattice and consider *third maximal* subgroups (i.e. subgroups of depth 3), we can show unconditionally that there is no bound on the number of generators:

Proposition 4. *For each real number c there is a third maximal subgroup M of an almost simple group such that $d(M) > c$.*

Next we move on to random generation. For a finite group G and a positive integer k let $P(G,k)$ denote the probability that k randomly chosen elements of G generate G. Let $\nu(G)$ be the minimal number k such that $P(G,k) \geqslant 1/e$. Up to a small multiplicative constant, it is known that $\nu(G)$ is the expected number of random elements generating G (see [30] and [26, Proposition 1.1]). In [8, Theorem 3] it was shown that $\nu(M)$ is bounded by a constant for all maximal subgroups M of almost simple groups. Combining Theorem 1 with results of Jaikin-Zapirain and Pyber [16], we extend this to second maximal subgroups, as follows.

Theorem 5. *There is a constant c such that $\nu(M) \leqslant c$ for all second maximal subgroups M of almost simple groups, with the possible exception of those in part* (iii) *of Theorem* 1.

More precisely, we show that $\nu(M) \leqslant c$ for every second maximal subgroup M of an almost simple group if and only if the question (1) has a negative solution. Indeed, this follows by combining Theorem 3 with Corollary 8.2.

Our final result concerns the growth of third maximal subgroups. Recall that for a group G and a positive integer n, the number of maximal subgroups of index n in G is denoted by $m_n(G)$. The maximal subgroup growth of finite and profinite groups has been widely studied in relation to the notion of positively finitely generated groups – that is, groups G for which, for bounded k, $P(G,k)$ is bounded away from zero (see [28, 29, 26]). For simple groups G, the theory was developed in [17, 25], culminating in [20], where it was proved that $m_n(G) \leqslant n^a$ for any fixed $a > 1$ and sufficiently large n. A polynomial bound for second maximal subgroups was obtained in [8, Corollary 6]. This was based on the random generation of maximal subgroups by a bounded number of elements, together with Lubotzky's inequality $m_n(H) \leqslant n^{\nu(H)+3.5}$ for all finite groups H ([26]). Here we show that, despite the fact that second maximal subgroups may not have such a random generation property, the growth of third maximal subgroups is still polynomial.

Theorem 6. *There is a constant c such that any almost simple group has at most n^c third maximal subgroups of index n.*

Our notation is fairly standard. We adopt the notation of [19] 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, we write Z_n (or just n) and D_n for the cyclic and dihedral groups of order n, respectively, and $[n]$ denotes an arbitrary solvable group of order n.

The paper is organised as follows. In Section 2 we start with some preliminary results that are needed in the proofs of our main theorems. Next, in Sections 3 and 4 we prove Theorem 1 for groups with an alternating group and sporadic socle, respectively. This leaves us to deal with groups of Lie type. In Section 5 we consider second maximal subgroups lying in maximal non-parabolic subgroups of classical groups, and we do likewise for the exceptional groups in Section 6. We complete the proof of Theorem 1 in Section 7, where we deal with the maximal subgroups of parabolic subgroups in groups of Lie type. Here we also present connections with Number Theory and the proof of Theorem 3 is completed at the end of the section. Finally, in Section 8 we discuss random generation and growth, and we prove Proposition 4 and Theorems 5 and 6.

2. Preliminaries

In this section we record several preliminary results that will be needed in the proofs of our main theorems. We start by recalling two of the main results from [8]. The first is [8, Theorem 2]:

Theorem 2.1. ([8]) *Let G be a finite almost simple group with socle G_0 and let H be a maximal subgroup of G. Then $d(H \cap G_0) \leqslant 4$ and $d(H) \leqslant 6$.*

The next result is [8, Theorem 7].

Theorem 2.2. ([8]) *Let G be a finite primitive permutation group with point stabilizer H. Then $d(G) - 1 \leqslant d(H) \leqslant d(G) + 4$.*

Recall that if M is a subgroup of a group H, then

$$\mathrm{core}_H(M) = \bigcap_{h \in H} M^h$$

is the *H-core* of M, which is the largest normal subgroup of H contained in M. The next result, which follows immediately from Theorems 2.1 and 2.2, will play a key role in our analysis of second maximal subgroups.

Lemma 2.3. *Let G be a finite almost simple group and let M be a second maximal subgroup of G, so that $M < H < G$ with each subgroup maximal in the next. If $\mathrm{core}_H(M) = 1$, then $d(M) \leqslant 10$.*

Proof. The condition $\mathrm{core}_H(M) = 1$ implies that H acts faithfully and primitively on the cosets of M. Then $d(M) \leqslant d(H) + 4$ by Theorem 2.2, and $d(H) \leqslant 6$ by Theorem 2.1. □

Remark 2.4. In general, if $N = \mathrm{core}_H(M)$ then Lemma 2.3 implies that $d(M/N) \leqslant 10$, so

$$d(M) \leqslant d_M(N) + d(M/N) \leqslant d_M(N) + 10$$

where $d_M(N)$ is the minimal number of generators of N as a normal subgroup of M (that is, $d_M(N)$ is the minimal d such that $N = \langle x_1^M, \ldots, x_d^M \rangle$ for some $x_i \in N$).

Lemma 2.5. *Let V be a finite dimensional vector space over $\mathbb{F}_q$, and let G be a group such that $\mathrm{SL}(V) \leqslant G \leqslant \Gamma\mathrm{L}(V)$, $\mathrm{Sp}(V) \leqslant G \leqslant \Gamma\mathrm{Sp}(V)$ or $\Omega(V) \leqslant G \leqslant \Gamma\mathrm{O}(V)$. If H is any maximal subgroup of G, then V is a cyclic $\mathbb{F}_q(H \cap \mathrm{GL}(V))$-module.*

Proof. Set $\tilde{G} = G \cap \mathrm{GL}(V)$ and $\tilde{H} = H \cap \mathrm{GL}(V)$. The result is immediate if $\tilde{H}$ acts irreducibly on V, so let us assume $\tilde{H} = \tilde{G}_U$ is the stabilizer of a proper subspace U of V. In the linear case, $\tilde{H}$ stabilizes no other proper non-zero subspace, so any vector $v \in V \setminus U$ generates V as an $\mathbb{F}_q\tilde{H}$-module. Now assume G is symplectic or orthogonal. If U is totally singular (or a non-singular 1-space when G is orthogonal and q is even) then any vector $v \in V \setminus U^{\perp}$ is a generator. Finally, suppose U is non-degenerate. Here U and $U^{\perp}$ are the only proper non-zero $\tilde{H}$-invariant subspaces of V, so any vector $u_1 + u_2 \in U \perp U^{\perp}$ with $u_1, u_2 \neq 0$ is a generator. □

Suppose $G = S_n$ or A_n and let $X = \mathbb{F}_p^n$ be the permutation module for G over $\mathbb{F}_p$, where $n \geqslant 3$ and p is a prime. Set

$$U = \{(a_1, \ldots, a_n) \,:\, \sum a_i = 0\}, \;\; W = \{(a, \ldots, a) \,:\, a \in \mathbb{F}_p\}$$

and note that $W \subseteq U$ if p divides n, otherwise $X = U \oplus W$. It is easy to check that U and W are the only proper non-zero submodules of X, so the quotient $V = U/(U \cap W)$ is irreducible. We call V the *fully deleted permutation module* for G. Note that $\dim V = n - 2$ if p divides n, otherwise $\dim V = n - 1$.

Lemma 2.6. *Let $G = S_n$ or A_n, where $n \geqslant 5$, let p be a prime and let V be the fully deleted permutation module for G over $\mathbb{F}_p$. If H is any maximal subgroup of G, then V is a cyclic $\mathbb{F}_pH$-module.*

Proof. This is an easy exercise if H is an intransitive or imprimitive maximal subgroup $(S_k \times S_{n-k}) \cap G$ or $(S_t \wr S_{n/t}) \cap G$. So assume now that H is primitive on $I := \{1, \ldots, n\}$. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{F}_p^n$ and let $v = e_1 - e_2 = (1, -1, 0, \ldots, 0) \in U$. We show that the orbit v^H spans U. For a subset $J \subseteq I$, let $V(J) = \langle e_i - e_j \,:\, i, j \in J \rangle$. Note that $\langle v \rangle = V(\{1, 2\})$.

Define S to be the span of v^H. We claim that if $V(J) \subseteq S$ (where $1 < |J| < n$) then there is a larger set J' containing J such that $V(J') \subseteq S$. To see this, note that as H is primitive, J is not a block for H, so there exists $h \in H$ such that $J \cap J^h$ is neither empty nor J. Say h sends $i \mapsto x$, $j \mapsto y$, where $i, j \in J$, $x \in J$ and $y \notin J$. Then h sends $e_i - e_j \mapsto e_x - e_y$, and so $\langle V(J), (e_i - e_j)^h \rangle$ contains $V(J')$, where $J' = J \cup \{y\}$. Hence the claim, and the lemma follows. □

The next result concerns the minimal generation of maximal subgroups of certain wreath products. In the statement of the lemma, we use the notation $\frac{1}{e}H$ for a normal subgroup of index e in H, and we write V_4 for the Klein four-group $Z_2 \times Z_2$.

Lemma 2.7. *Let G be one of the following groups, where $n \geqslant 2$ and $A = S_n$ or A_n.*

(i) *$G = \frac{1}{e}(Z_d \wr A)$, where $d \geqslant 3$, e divides d, and furthermore the natural projection map from G to A is surjective.*

(ii) *$G = \frac{1}{e}(Z_2 \wr A)$ with $e = 1$ or 2.*

(iii) *$G = \frac{1}{e}(V_4 \wr A)$ with $e = 1, 2$ or 4.*

(iv) *$G = \frac{1}{e}(D_8 \wr A)$ with $e = 1$ or 2.*

(v) *$G = \frac{1}{e}(Q_8 \wr A)$ with $e = 1$ or 2.*

Then $d(H) \leqslant 6$ for every maximal subgroup H of G.

Proof. The result is trivial for $n = 2$ and for $n = 3$, $A = A_3$, so assume that $n \geqslant 3$ and $A \neq A_3$. First consider (i) and (ii). Without loss of generality, we may assume that $G = BA$, where the base group B is the kernel of an A-invariant homomorphism from $(Z_d)^n$ to Z_e (here $d = 2$ in case (ii)). Then using the action of A we see that $B = B(e)$, where

$$B(e) = \{(\lambda_1, \dots, \lambda_n) \in (Z_d)^n \ : \ \sum \lambda_i \equiv 0 \ (\mathrm{mod}\ e)\}$$

(writing Z_d as the additive group of integers modulo d). Let H be a maximal subgroup of G.

Suppose first that $B \leqslant H$. Then $H = BM$ where M is a maximal subgroup of A. As in the previous proof we see that there is a vector $v \in B$ such that $\langle v^M \rangle$ contains $B(0)$. Since $B(e)/B(0)$ is cyclic, it follows that $d_H(B) \leqslant 2$ and thus $d(H) \leqslant 2 + d(M) \leqslant 6$ since $d(M) \leqslant 4$ by [8, Proposition 4.2].

Now suppose that $B \nleqslant H$. Then $H/(H \cap B) \cong A$ and $H \cap B$ is a maximal A-invariant subgroup of B. Let $d = \prod p_i^{a_i}$ where the p_i are distinct primes, and let P_i be a Sylow p_i-subgroup of B. Order the p_i so that $P_1 \nleqslant H$. As each P_i is A-invariant, we have

$$H \cap B = (H \cap P_1) \prod_{i \geqslant 2} P_i.$$

Write $p = p_1$, $a = a_1$ and $p^b = e_p$ for the p-part of e, so that

$$P_1 = \{(\lambda_1, \dots, \lambda_n) \in (Z_{p^a})^n \ : \ \sum \lambda_i \equiv 0 \ (\mathrm{mod}\ p^b)\}.$$

Let $\phi : P_1 \to (Z_p)^n$ be the map sending $(\lambda_1, \dots, \lambda_n) \mapsto (p^{a-1}\lambda_1, \dots, p^{a-1}\lambda_n)$. Then $\phi(H \cap P_1)$ is a non-zero A-invariant subspace of $(Z_p)^n$, so is one of U, W or $(Z_p)^n$ (where U, W are as defined above). If $\phi(H \cap P_1)$ contains U, then $H \cap P_1$ has an element h of the form

$$h = (1 + p\lambda_1', -1 + p\lambda_2', p\lambda_3', \dots, p\lambda_n'),$$

and $\langle h^A \rangle$ is a subgroup $(Z_{p^a})^{n-1}$ of P_1. Thus $d_H(H \cap P_1) \leqslant 2$, and similarly $d_H(\prod_{i \geqslant 2} P_i) \leqslant 2$, so $d(H) \leqslant d_H(H \cap B) + d(A) \leqslant 6$. Finally, if $\phi(H \cap P_1) = W$ then $H \cap P_1 = \phi^{-1}(W)$ by maximality, and again we see that $d_H(H \cap P_1) \leqslant 2$, giving the result as above.

The remaining cases are similar to, but easier than, (i) and (ii). Consider part (iv), for example. Let $B = G \cap (D_8)^n$ be the base group of G, and let $C = G \cap (Z_4)^n < B$. The result follows in the usual way if $B \leqslant H$, so assume this is not the case. As in the proof of (i) we see that $d_H(H \cap C) \leqslant 2$. Also $B/C = \frac{1}{e'}(Z_2)^n$, and we see in the usual way that $d_{H/H \cap C}(H \cap B/H \cap C) \leqslant 2$. Hence $d(H) \leqslant 4 + d(A) \leqslant 6$. □

We will also need some results on the generation of maximal subgroups of certain non-simple classical groups.

Lemma 2.8. *Let G be a group such that $G_0 \leqslant G \leqslant \mathrm{Aut}(G_0)$, where $G_0 = \mathrm{P}\Omega_4^+(q)$, and let H be a maximal subgroup of G. Then $d(G) \leqslant 6$, $d(H) \leqslant 8$ and $d(H \cap G_0) \leqslant 4$.*

Proof. Here $G_0 = S \times S$ with $S = \mathrm{L}_2(q)$ and it is easy to check that the result holds when $q \in \{2, 3\}$. Now assume $q \geqslant 4$, so S is simple. Write $q = p^f$ with p prime and set $H_0 = H \cap G_0$. Since $d(G_0) = 2$ and every subgroup of

$$\mathrm{Out}(G_0) = (Z_{(2,q-1)} \times Z_f) \wr S_2$$

is 4-generator, it suffices to show that $d(H_0) \leqslant 4$. Write $G = G_0.A$.

If H contains G_0 then $H_0 = G_0$ and thus $d(H_0) = 2$. Otherwise $H = H_0.A$ and H_0 is a maximal A-invariant subgroup of G_0. It follows that H_0 is either a diagonal subgroup isomorphic to S, or it is of the form $S \times B$, $B \times S$, $B \times B$, where $B = C \cap S$ and C is a maximal subgroup of an almost simple group with socle S. By inspecting [7, Table 8.1], we observe that $d(B) \leqslant 2$ and thus $d(H_0) \leqslant 4$ as required. □

Lemma 2.9. *Let $G_0 \in \{\mathrm{L}_2(2), \mathrm{L}_2(3), \mathrm{U}_3(2)\}$ and let H be a maximal subgroup of G, where $G_0 \leqslant G \leqslant \mathrm{Aut}(G_0)$. Then $d(H \cap G_0) \leqslant 3$.*

Proof. This is a straightforward calculation. □

3. Symmetric and alternating groups

In this section we begin the proof of Theorem 1 by handling the case where G_0 is an alternating group. Our main result is the following.

Proposition 3.1. *Let G be an almost simple group with socle A_n. Then $d(M) \leqslant 10$ for every second maximal subgroup M of G.*

Proof. If $n \leqslant 8$ then it is easy to check that $d(M) \leqslant 3$, so for the remainder we may assume that $G = A_n$ or S_n, with $n \geqslant 9$. Write $M < H < G$, where M is maximal in H, and H is maximal in G. The possibilities for H are given by the O'Nan-Scott theorem and we deduce that one of the following holds:

1. H is intransitive: $H = (S_k \times S_{n-k}) \cap G$, $1 \leqslant k < n/2$;

2. H is affine: $H = \mathrm{AGL}_d(p) \cap G$, $n = p^d$, p prime, $d \geqslant 1$;

3. H is imprimitive or wreath-type: $H = (S_k \wr S_t) \cap G$, $n = kt$ or k^t;

4. H is diagonal: $H = (T^k.(\mathrm{Out}(T) \times S_k)) \cap G$, T non-abelian simple, $n = |T|^{k-1}$;

5. H is almost simple.

If H is almost simple, then $d(M) \leqslant 6$ by Theorem 2.1, so we need to deal with the first four cases. Set $C = \mathrm{core}_H(M)$. If $C = 1$ then $d(M) \leqslant 10$ by Lemma 2.3, so we may assume otherwise.

Case 1: H is intransitive.

First assume $k \geqslant 5$. If C contains $A_k \times A_{n-k}$ then [8, Proposition 2.8] implies that $d(M) \leqslant 3$. Otherwise, C and H/C are 2-generator almost simple groups and thus Theorem 2.2 implies that

$$d(M) \leqslant d(M/C) + d(C) \leqslant d(H/C) + d(C) + 4 \leqslant 8.$$

Next suppose $k = 4$. The result quickly follows if C contains $V_4 \times A_{n-4}$, so assume otherwise. Then either C is a subgroup of S_4 and H/C has socle A_{n-4}, or vice versa, whence $d(M) \leqslant 8$ as before. A very similar argument applies if $k \leqslant 3$.

Case 2: H is affine.

Here $H = \mathrm{AGL}(V) \cap G = V.L$, where $V = \mathbb{F}_p^d$ is the unique minimal normal subgroup of H and $\mathrm{SL}(V) \leqslant L = \mathrm{GL}(V) \cap G$. Note that $n = p^d$. Since we may assume $C \neq 1$ it follows that $M = V.J$ and $J < L$ is maximal. If $d = 1$ or

$(d,p) = (2,3)$ then it is easy to see that $d(M) \leqslant 2$, so we may assume that $\mathrm{SL}(V)$ is quasisimple. Let $Z = Z(L)$ and note that Z is cyclic. Then L/Z is almost simple and thus $d(JZ/Z) \leqslant 6$ by Theorem 2.1. Therefore $d(J) \leqslant 7$, and by applying Lemma 2.5 we deduce that $d(M) \leqslant 8$.

Case 3: H is imprimitive or wreath-type.

First assume $G = S_n$. Write $H = S_k \wr S_t = N.S_t$, where $N = (S_k)^t$ and $k, t \geqslant 2$. If $k = 2$ then Lemma 2.7 implies that $d(M) \leqslant 6$, so we may assume that $k \geqslant 3$. Suppose M contains N, so $M = N.J$ and $J < S_t$ is maximal. Now J has $s \leqslant 2$ orbits on $\{1,\ldots,t\}$, and $d(J) \leqslant 4$ by [8, Proposition 4.2] (the cases with $t \leqslant 4$ can be checked directly), so

$$d(M) \leqslant d((S_k)^s) + d(J) \leqslant 6$$

since $d(S_k \times S_k) = 2$ (see [8, Proposition 2.8]). Now assume M does not contain N, so $M = (M \cap N).S_t$ and $M \cap N$ is a maximal S_t-invariant subgroup of N. If $k \neq 4$ then $A = (A_k)^t$ is the unique minimal normal subgroup of H, so we may assume H contains A and we can consider $\bar{M} = M/A < \bar{H} = H/A = S_2 \wr S_t$. By Lemma 2.7 we have $d(\bar{M}) \leqslant 6$ and thus $d(M) \leqslant d(A_k) + d(\bar{M}) \leqslant 8$. Similarly, if $k = 4$ then $A = (V_4)^t$ is the unique minimal normal subgroup of H, so we may assume M contains A. Note that $B = (Z_3)^t$ is the unique minimal normal subgroup of $H/A = S_3 \wr S_t$. If M/A does not contain B, then Theorem 2.2 implies that $d(M/A) \leqslant d(H/A) + 4 = 6$ and thus $d(M) \leqslant 8$. Therefore, we may assume that M/A contains B, so M contains $(A_4)^t$ and the above argument goes through (via Lemma 2.7).

Now assume $G = A_n$ and $H = (S_k \wr S_t) \cap G$. If $H = S_k \wr S_t$ (which can happen if $n = k^t$) then the previous argument applies. Therefore, we may assume that H is an index-two subgroup of $S_k \wr S_t$, so $H = ((A_k)^t.2^{t-1}).S_t$ or $S_k \wr A_t$. The latter case is handled as above, so let us assume $H = ((A_k)^t.2^{t-1}).S_t = N.S_t$. If $k = 2$ then $H = \frac{1}{2}(S_2 \wr S_t)$ and thus $d(M) \leqslant 6$ by Lemma 2.7. Now assume $k \geqslant 3$. If M contains N then $M = N.J$ with $J < S_t$ maximal and it is easy to see that $d(M) \leqslant (2+1)s + d(J) \leqslant 10$, where $s \leqslant 2$ is the number of orbits of J on $\{1,\ldots,t\}$. If $N \nleqslant M$ then we can reduce to the case where M contains $(A_k)^t$ and by applying Lemma 2.7 we deduce that $d(M) \leqslant 8$.

Case 4: H is diagonal.

Write $H = (T^k.(\mathrm{Out}(T) \times S_k)) \cap G$. First assume $G = S_n$. Here $H = T^k.(\mathrm{Out}(T) \times S_k)$ and T^k is the unique minimal normal subgroup of H, so $M = T^k.J$ for some maximal subgroup $J < \mathrm{Out}(T) \times S_k$. The projection of J to the S_k factor has $s \leqslant 2$ orbits on $\{1,\ldots,k\}$ and thus $d(M) \leqslant d(T^s) + d(J) = 2 + d(J)$.

If J is a standard maximal subgroup of $\mathrm{Out}(T) \times S_k$ (i.e. J is of the form $A \times S_k$ or $\mathrm{Out}(T) \times B$, where A, B are maximal in the respective factors), then $d(J) \leqslant 7$ since every subgroup of $\mathrm{Out}(T)$ is 3-generator, $d(S_k) \leqslant 2$ and every maximal subgroup of S_k is 4-generator. The only other possibility is $J = (L \times A_k).2$, where $|\mathrm{Out}(T) : L| = 2$ (see [34, Lemma 1.3], for example). Clearly, $d(J) \leqslant 6$ in this case.

Now suppose $G = A_n$. We may as well assume that H is an index-two subgroup of $T^k.(\mathrm{Out}(T) \times S_k)$, otherwise the previous argument applies. If $k \geqslant 3$, then $H = T^k.(L \times S_k)$, where $|\mathrm{Out}(T) : L| = 2$ (see the proof of [8, Lemma 4.4]), and T^k is the unique minimal normal subgroup of H. In this situation, the above argument goes through unchanged. Finally, assume $k = 2$. Set $\ell = \frac{1}{2}(|T| - i_2(T) - 1)$, where $i_2(T)$ denotes the number of involutions in T. As explained in the proof of [8, Lemma 4.4], if ℓ is even then $H = T^2.(L \times S_2)$ as above, and the usual argument applies. If ℓ is odd then $H = T^2.\mathrm{Out}(T)$, so H has two minimal normal subgroups N_1 and N_2 (both isomorphic to T). If M contains T^2 then $M = T^2.J$ (with $J < \mathrm{Out}(T)$ maximal) and thus $d(M) \leqslant d(T^2) + d(J) \leqslant 2 + 3 = 5$. Otherwise we may assume that M contains N_1, but not N_2, in which case M/N_1 is a maximal subgroup of $H/N_1 \cong \mathrm{Aut}(T)$. By Theorem 2.1 we have $d(M/N_1) \leqslant 6$ and we conclude that $d(M) \leqslant 8$. □

4. Sporadic groups

Our main result on second maximal subgroups of sporadic groups is the following.

Proposition 4.1. *Let G be an almost simple group with sporadic socle G_0. Then $d(M) \leqslant 10$ for every second maximal subgroup M of G.*

As before, write $M < H < G$ where H is a maximal subgroup of G. Set

$$\mathcal{A} = \{\mathrm{M}_{11}, \mathrm{M}_{12}, \mathrm{M}_{22}, \mathrm{M}_{23}, \mathrm{M}_{24}, \mathrm{HS}, \mathrm{J}_1, \mathrm{J}_2, \mathrm{J}_3, \mathrm{Co}_2, \mathrm{Co}_3, \mathrm{McL}, \mathrm{Suz}, \mathrm{He}, \mathrm{Fi}_{22}, \mathrm{Ru}\}$$

and

$$\mathcal{B} = \{\mathrm{O'N}, \mathrm{J}_4, \mathrm{Th}, \mathrm{Ly}, \mathrm{HN}\}.$$

Lemma 4.2. *If $G_0 \in \mathcal{A} \cup \mathcal{B}$, then $d(M) \leqslant 5$.*

Proof. It is convenient to use Magma [6], together with the detailed information on sporadic groups and their maximal subgroups provided in the Web-Atlas [36]. First assume $G_0 \in \mathcal{A}$. Here we use the Web-Atlas to construct G as a permutation group of degree $n \leqslant 6156$ (with equality if $G_0 = \mathrm{J}_3$), and we use the Magma command MaximalSubgroups to construct H and M as permutation groups of degree n. In each case it is straightforward to find five generators for M by random search.

A similar approach is effective if $G_0 \in \mathcal{B}$. For example, suppose $G = \mathrm{O'N}.2$. First we use the Web-Atlas to construct G as a permutation group on 245520 points, and then we construct the maximal subgroups H of G using the generators given in the Web-Atlas. As before, we can use Magma to find the maximal subgroups of H, and the desired result quickly follows. The remaining cases are similar, working with a suitable matrix representation when $G = \mathrm{J}_4$, Th or Ly. □

Remark 4.3. The bound $d(M) \leqslant 5$ in Lemma 4.2 is sharp. For example, take

$$G = \mathrm{Fi}_{22}.2, \quad H = \mathrm{U}_4(3).2^2 \times S_3, \quad M = J.2^2 \times S_3,$$

where $J < \mathrm{U}_4(3)$ is a maximal subgroup of type $\mathrm{GU}_2(3) \wr S_2$. Then M has a normal subgroup N such that M/N is elementary abelian of order 2^5, so $d(M) = 5$. (More precisely, $M = 2.\mathrm{L}_2(3)^2.2^4 \times S_3$ and $N = 2.\mathrm{L}_2(3)^2 \times 3$.)

Lemma 4.4. *If* $G_0 \in \{\mathrm{Fi}_{23}, \mathrm{Fi}_{24}', \mathrm{Co}_1\}$, *then* $d(M) \leqslant 8$.

Proof. First observe that Theorem 2.1 implies that $d(M) \leqslant 6$ if H is almost simple, so we may assume otherwise.

Suppose $G = \mathrm{Fi}_{23}$. Using the Web-Atlas, we construct G as a permutation group of degree 31671. In all but four cases, generators for H are given in the Web-Atlas, and we can proceed as in the proof of the previous lemma. The exceptions are the following:

$$H \in \{3^{1+8}.2^{1+6}.3^{1+2}.2S_4,\ [3^{10}].(\mathrm{L}_3(3) \times 2),\ 2^{6+8}{:}(A_7 \times S_3),\ \mathrm{Sp}_6(2) \times S_4\}.$$

It is easy to construct $H = \mathrm{Sp}_6(2) \times S_4$ as a permutation group of degree 67, and the bound $d(M) \leqslant 3$ quickly follows. We can obtain $H = 2^{6+8}{:}(A_7 \times S_3)$ as the normalizer in G of a normal subgroup of order 2^{14} in a Sylow 2-subgroup of G. The 3-local subgroups $3^{1+8}.2^{1+6}.3^{1+2}.2S_4$ and $[3^{10}].(\mathrm{L}_3(3)\times 2)$ can be constructed in a similar fashion, using a Sylow 3-subgroup. In all three cases, it is easy to check that $d(M) \leqslant 3$.

Next suppose $G = \mathrm{Fi}_{24}$. Here we start with a permutation representation of degree 306936, and we construct the maximal subgroups H of G using the generators given in the Web-Atlas. In all but two cases, we can use Magma to find the maximal subgroups M of H, and verify the bound $d(M) \leqslant 4$. The exceptions are the cases

$$H \in \{S_3 \times \mathrm{P}\Omega_8^+(3){:}S_3,\ \mathrm{Fi}_{23} \times 2\}.$$

If $H = \mathrm{Fi}_{23} \times 2$ then $M = \mathrm{Fi}_{23}$ or $J \times 2$, where $J < \mathrm{Fi}_{23}$ is maximal, so $d(M) \leqslant 4$. Suppose $H = S_3 \times \mathrm{P\Omega}_8^+(3){:}S_3$. Then [34, Lemma 1.3] implies that

$$M \in \{J \times \mathrm{P\Omega}_8^+(3){:}S_3,\ S_3 \times L,\ (3 \times \mathrm{P\Omega}_8^+(3){:}3).2\},$$

where $J < S_3$ and $L < \mathrm{P\Omega}_8^+(3){:}S_3$ are maximal. Since J is cyclic and $d(L) \leqslant 6$, it follows that $d(M) \leqslant 8$ in the first two cases. In the final case, it is clear that $d(M) \leqslant 4$ (note that $d(\mathrm{P\Omega}_8^+(3){:}3) = 2$).

Now assume $G = \mathrm{Fi}_{24}'$. Every maximal subgroup H of G is the intersection with G of a maximal subgroup of $G.2 = \mathrm{Fi}_{24}$ (with the exception of the almost simple maximal subgroups He:2, $\mathrm{U}_3(3){:}2$ and $\mathrm{L}_2(13){:}2$), so we can construct H as above, use Magma to obtain the maximal subgroups M of H, and then finally verify the desired bound on $d(M)$. This approach is effective unless $H = (3 \times \mathrm{P\Omega}_8^+(3){:}3).2$. Let M be a maximal subgroup of H. If $M = 3 \times \mathrm{P\Omega}_8^+(3){:}3$ then clearly $d(M) \leqslant 3$, so assume otherwise. Then $M = J.2$, and either $M \cong \mathrm{P\Omega}_8^+(3){:}S_3$ is almost simple, or $J = 3 \times L$ with $L = K \cap \mathrm{P\Omega}_8^+(3){:}3$ for some maximal subgroup $K < \mathrm{P\Omega}_8^+(3){:}S_3$. Since $d(L) \leqslant 5$ by Theorem 2.1, we conclude that $d(M) \leqslant 7$.

Finally, suppose $G = \mathrm{Co}_1$. Here we work with a permutation representation of degree 98280. Explicit generators for the six largest maximal subgroups are given in the Web-Atlas, and in the usual way we deduce that $d(M) \leqslant 3$. Representatives of the remaining sixteen conjugacy classes of maximal subgroups H of G can be constructed using the information provided in the Atlas [12] and Web-Atlas, and once again we find that $d(M) \leqslant 3$. As before, the p-local maximal subgroups can be constructed by taking normalizers of appropriate normal subgroups of a Sylow p-subgroup of G. We leave the reader to check the details. □

Lemma 4.5. *If $G = \mathbb{B}$ or $\mathbb{M}$, then $d(M) \leqslant 10$.*

Proof. First assume $G = \mathbb{B}$. If H is almost simple then $d(M) \leqslant 6$, so we may assume otherwise. Also recall that $d(M) \leqslant 10$ if $\mathrm{core}_H(M) = 1$, so we may also assume that M contains a nontrivial normal subgroup of H. The maximal subgroups of G are listed in the Web-Atlas.

Suppose $H = 2.{}^2E_6(2){:}2$. Here $Z(H) \cong Z_2$ is the unique minimal normal subgroup of H, so we may assume that $M = 2.J$, where $J < {}^2E_6(2){:}2$ is maximal. Since ${}^2E_6(2){:}2$ is almost simple, Theorem 2.1 implies that $d(M) \leqslant 1 + 6 = 7$.

If $H = 2^{1+22}.\mathrm{Co}_2$ then Z_2 is the unique minimal normal subgroup of H, so we can assume that $M = Z_2 \times \mathrm{Co}_2$ or $2^{1+22}.J$, where $J < \mathrm{Co}_2$ is maximal. If $M = Z_2 \times \mathrm{Co}_2$ then $d(M) = 2$, so let us assume $M = 2^{1+22}.J$. Here Z_2 is the unique minimal normal subgroup of M, so $d(M) = d(2^{22}.J)$ by the main theorem of [27]. Using Magma, we calculate that J has at most 7 composition factors on the irreducible $\mathbb{F}_2\mathrm{Co}_2$-module 2^{22}, and by applying [8, Proposition 3.1] we conclude that $d(M) \leqslant 7 + d(J) \leqslant 10$.

Next assume $H = (2^2 \times F_4(2)){:}2$. If $M = 2^2 \times F_4(2)$ then $d(M) \leqslant 4$, otherwise $M = (2 \times F_4(2)).2$ or $(2^2 \times J).2$, where $J = L \cap F_4(2)$ for some maximal subgroup $L < F_4(2).2$. In the first case it is clear that $d(M) \leqslant 4$. In the latter, Theorem 2.1 gives $d(J) \leqslant 4$, so $d(M) \leqslant 7$.

If $H = 2^{2+10+20}.(\mathrm{M}_{22}{:}2 \times S_3)$, $[2^{35}].(S_5 \times \mathrm{L}_3(2))$ or $5^3.\mathrm{L}_3(5)$, then a permutation representation of H of degree 6144 is given in the Web-Atlas, and it is straightforward to show that M is 3-generator. Similarly, we can use a matrix representation of $H = 2^{9+16}.\mathrm{Sp}_8(2)$ of dimension 180 over $\mathbb{F}_2$ to check that $d(M) \leqslant 4$. The Web-Atlas also provides a matrix representation of $H = [2^{30}].\mathrm{L}_5(2)$ of dimension 144 over $\mathbb{F}_2$ and one can check that $d(M) = 2$ (we thank Eamonn O'Brien for his assistance with this computation).

In each of the remaining cases, we can take a suitable permutation representation of H (see the proof of [9, Proposition 3.3], for example), and it is straightforward to check that $d(M) \leqslant 4$.

The case $G = \mathbb{M}$ is similar. Again we may assume that H is not almost simple and $\mathrm{core}_H(M) \neq 1$, so H belongs to one of the conjugacy classes of maximal subgroups of G listed in the Web-Atlas. If $|H| < 5 \times 10^9$ then a permutation representation of H is given in the Web-Atlas, and it is straightforward to check that $d(M) \leqslant 4$. The remaining cases can be handled by arguing as above. For example, suppose $H = 2^{5+10+20}.(S_3 \times \mathrm{L}_5(2))$. Here 2^5 is the unique minimal normal subgroup of H, so we may assume that $M = 2^{5+10+20}.J$ or $2^{5+10}.(S_3 \times \mathrm{L}_5(2))$, where $J < S_3 \times \mathrm{L}_5(2)$ is maximal. In the latter case, the main theorem of [27] implies that $d(M) = d(S_3 \times \mathrm{L}_5(2)) = 2$. Now assume $M = 2^{5+10+20}.J$. Using Magma, we calculate that J has at most 8 composition factors on 2^5, 2^{10} and 2^{20}, in total. Since every maximal subgroup of $S_3 \times \mathrm{L}_5(2)$ is 2-generator, it follows that $d(M) \leqslant 8 + d(J) \leqslant 10$.

Another possibility is $H = 3^8.\mathrm{P}\Omega_8^-(3).2_3$. In this case 3^8 is the unique minimal normal subgroup of H, so we may assume that $M = 3^8.J$, where $J < \mathrm{P}\Omega_8^-(3).2_3$ is maximal. Here 3^8 is the natural module for $\mathrm{P}\Omega_8^-(3)$ and Lemma 2.5 implies that $d(M) \leqslant 1 + d(J) \leqslant 6$. The other cases are similar and we omit the details. □

5. Classical groups

Let G be an almost simple classical group over $\mathbb{F}_q$ with socle G_0, where $q = p^f$ for a prime p. Let V be the natural G_0-module. Write $M < H < G$, where M is maximal in H, and H is maximal in G.

Let n denote the dimension of V. Due to the existence of exceptional isomorphisms between certain low-dimensional classical groups (see [19, Proposition

2.9.1], for example), we may (and will) assume that $n \geqslant 3$ if $G_0 = \mathrm{U}_n(q)$, $n \geqslant 4$ if $G_0 = \mathrm{PSp}_n(q)'$, and $n \geqslant 7$ if $G_0 = \mathrm{P}\Omega_n^{\epsilon}(q)$. We also assume that $(n,q) \neq (4,2)$ if $G_0 = \mathrm{PSp}_n(q)'$, since $\mathrm{PSp}_4(2)' \cong A_6$.

The purpose of this section is to prove Theorem 1 in the case where G_0 is classical and M is contained in a maximal non-parabolic subgroup H of G. It is convenient to postpone the analysis of maximal subgroups of parabolic subgroups to Section 7, where we also deal with parabolic subgroups of exceptional groups.

By Aschbacher's subgroup structure theorem for finite classical groups (see [2]), with some exceptional cases for $G_0 = \mathrm{P}\Omega_8^+(q)$ or $\mathrm{PSp}_4(q)$ (with q even), the maximal subgroup H of G is either almost simple, or it belongs to one of eight subgroup collections, denoted $C_1, \ldots, C_8$, which are roughly described in Table 1. In order to prove Theorem 1 for classical groups, we will consider each of these subgroup collections in turn.

C_1	Stabilizers of subspaces of V
C_2	Stabilizers of decompositions $V = \bigoplus_i V_i$, where $\dim V_i = a$
C_3	Stabilizers of prime index extension fields of $\mathbb{F}_q$
C_4	Stabilizers of decompositions $V = V_1 \otimes V_2$
C_5	Stabilizers of prime index subfields of $\mathbb{F}_q$
C_6	Normalizers of symplectic-type r-groups in absolutely irreducible representations
C_7	Stabilizers of decompositions $V = \bigotimes_i V_i$, where $\dim V_i = a$
C_8	Stabilizers of non-degenerate forms on V

Table 1. The C_i subgroup collections

The main result of this section is the following.

PROPOSITION 5.1. *Let M be a second maximal subgroup of an almost simple classical group G with socle G_0, where $M < H < G$ and H is a maximal non-parabolic subgroup of G. Then $d(M) \leqslant 12$.*

Set $M_0 = M \cap G_0$ and $H_0 = H \cap G_0$, and note that $G/G_0 \cong H/H_0$. If M contains H_0 then $d(M) \leqslant d(M_0) + d(M/H_0) \leqslant d(M_0) + 3$ since every subgroup of G/G_0 is 3-generator. Otherwise $H/H_0 \cong M/M_0$ and again we deduce that $d(M) \leqslant d(M_0) + 3$. Therefore, it suffices to show that $d(M_0) \leqslant 9$. This follows from Theorem 2.1 if H is almost simple, so we may assume that H belongs to one of the collections C_i, $i = 1, \ldots, 8$ (or a small additional collection of maximal subgroups that arises when $G_0 = \mathrm{P}\Omega_8^+(q)$ or $\mathrm{PSp}_4(q)$ (with q even)).

We begin with a useful preliminary result. Recall that the *solvable residual* of a finite group is the smallest normal subgroup such that the respective quotient is solvable (equivalently, it is the last term in the derived series).

Lemma 5.2. *Let $E = H^{\infty}$ be the solvable residual of H, and assume that E is quasisimple and acts irreducibly on V. Then $d(M) \leqslant 9$.*

Proof. Set $\tilde{G} = G \cap \mathrm{PGL}(V)$ and $\tilde{M} = M \cap \mathrm{PGL}(V)$. If H contains G_0 then H is almost simple and thus $d(M) \leqslant 6$ by Theorem 2.1, so assume otherwise. Set $C = C_G(E)$ and note that C is a normal subgroup of $N_G(E) = H$. The irreducibility of E on V implies that $C_{\tilde{G}}(E)$ is cyclic, so $C_{\tilde{M}}(E)$ is also cyclic, and thus $d(M \cap C) \leqslant 3$ since every subgroup of $G/\tilde{G}$ is 2-generator. Therefore, in order to prove the lemma it suffices to show that $d(MC/C) \leqslant 6$.

To see this, first note that H/C is almost simple (with socle $EC/C \cong E/Z(E)$). If M contains C then $MC/C = M/C$ is a maximal subgroup of H/C and thus $d(MC/C) \leqslant 6$ as required. On the other hand, if M does not contain C then $MC = H$, so $MC/C = H/C$ is almost simple and thus $d(MC/C) \leqslant 3$. □

Lemma 5.3. *Proposition* 5.1 *holds if $H \in \mathcal{C}_1$.*

Proof. The possibilities for G and H are listed in [19, Table 4.1.A]. Recall that H is non-parabolic.

First assume $G_0 = \mathrm{L}_n(q)$ and H is of type $\mathrm{GL}_m(q) \oplus \mathrm{GL}_{n-m}(q)$, where $1 \leqslant m < n/2$. It is convenient to work in the quasisimple group $\mathrm{SL}_n(q)$, so [19, Proposition 4.1.4] implies that $H = N.A$ where $N = \mathrm{SL}_m(q) \times \mathrm{SL}_{n-m}(q)$ and $A \leqslant (Z_{q-1} \times Z_{q-1}).(Z_f \times Z_2)$ with $q = p^f$. Note that $d(N) = 2$ (see [8, Proposition 2.5(ii)]) and every subgroup of $(Z_{q-1} \times Z_{q-1}).(Z_f \times Z_2)$ is 4-generator. In particular, if M contains N then $d(M) \leqslant 6$, so assume otherwise. Then $M = (M \cap N).A$ and it suffices to show that $d(M \cap N) \leqslant 8$. Since $M \cap N$ is a maximal A-invariant subgroup of N, it is of the form $C \times \mathrm{SL}_{n-m}(q)$ or $\mathrm{SL}_m(q) \times D$, where $C = E \cap \mathrm{SL}_m(q)$ and E is maximal in a group F such that $\mathrm{SL}_m(q) \leqslant F \leqslant \Gamma\mathrm{L}_m(q).\langle\gamma\rangle$ (where γ is a graph automorphism if $m \geqslant 3$, otherwise $\gamma = 1$), and similarly for D. Since C and D are 5-generator by Theorem 2.1 (the cases $m = 1$ and $(m, q) = (2, 2)$ or $(2, 3)$ can be checked directly), we conclude that $d(M \cap N) \leqslant 7$ and the result follows.

A very similar argument applies if $G_0 = \mathrm{U}_n(q)$ and H is of type $\mathrm{GU}_m(q) \perp \mathrm{GU}_{n-m}(q)$, and also if $G_0 = \mathrm{PSp}_n(q)$ and H is of type $\mathrm{Sp}_m(q) \perp \mathrm{Sp}_{n-m}(q)$. We omit the details. To complete the proof, we may assume that $G_0 = \mathrm{P}\Omega_n^{\epsilon}(q)$ and $n \geqslant 7$. If n, q are even and H is of type $\mathrm{Sp}_{n-2}(q)$, then H is almost simple and thus $d(M) \leqslant 6$. Now assume that H is of type $O_m^{\epsilon_1}(q) \perp O_{n-m}^{\epsilon_2}(q)$, where $(m, \epsilon_1) \neq (n-m, \epsilon_2)$. Note that q is odd if m or $n-m$ is odd. Again, it will be convenient to work in the quasisimple group $\Omega_n^{\epsilon}(q)$.

If $m = 1$ then $H_0 = \Omega_{n-1}(q).2$ and it is easy to see that $d(M_0) \leqslant 5$. Now assume $m \geqslant 2$, so [19, Proposition 4.1.6] implies that $H_0 = N.A$, where $N = \Omega_m^{\epsilon_1}(q) \times \Omega_{n-m}^{\epsilon_2}(q)$ and $A = [2^i]$ with $i = 1$ or 2. Note that N is 4-generator. If M contains N then $d(M_0) \leqslant 6$, so let us assume otherwise. Then $M_0 = (M \cap N).A$ and it suffices to show that $M \cap N$ is 7-generator. If $\Omega_m^{\epsilon_1}(q)$ and $\Omega_{n-m}^{\epsilon_2}(q)$ are both quasisimple then we can repeat the argument in the second paragraph of the proof, using Theorem 2.1, to deduce that $d(M \cap N) \leqslant 7$. Therefore, we may assume that

$$\Omega_m^{\epsilon_1}(q) \in \{\Omega_2^{\pm}(q), \Omega_3(3), \Omega_4^{+}(q)\}$$

and $\Omega_{n-m}^{\epsilon_2}(q)$ is quasisimple (if $G_0 = \Omega_7(3)$ and H is of type $O_4^+(3) \perp O_3(3)$, then it is easy to check that $d(M) \leqslant 4$). The first two cases are straightforward since $\Omega_2^{\pm}(q)$ is cyclic, and every subgroup of $\Omega_3(3)$ is 2-generated. Finally, if $\Omega_m^{\epsilon_1}(q) = \Omega_4^{+}(q)$ then the usual argument goes through, using Lemma 2.8 in place of Theorem 2.1. □

Lemma 5.4. *Proposition* 5.1 *holds if* $H \in \mathcal{C}_2$.

Proof. The various possibilities for G and H are recorded in [19, Table 4.2.A]. First assume $G_0 = \mathrm{U}_n(q)$ and H is of type $\mathrm{GL}_{n/2}(q^2)$. Here $n \geqslant 4$ is even and it is convenient to work in the quasisimple group $\mathrm{SU}_n(q)$, so $H_0 = N.A$ where $N = \mathrm{SL}_{n/2}(q^2)$ and $A = Z_{q-1}.Z_2$. If M contains N then $M_0 = N.B$ for some subgroup $B \leqslant A$, whence $d(M_0) \leqslant 4$ and the result follows. Otherwise, $M_0 = (M \cap N).A$ and $M \cap N = C \cap \mathrm{SL}_{n/2}(q^2)$, where C is a maximal subgroup of a group D such that $\mathrm{SL}_{n/2}(q^2) \leqslant D \leqslant \Gamma\mathrm{L}_{n/2}(q^2)\langle\gamma\rangle$ (here γ is a graph automorphism if $n \geqslant 6$, otherwise $\gamma = 1$). Therefore Theorem 2.1 implies that $d(M \cap N) \leqslant 4$ and thus $d(M_0) \leqslant 6$.

Next suppose $G_0 = \mathrm{P}\Omega_n^{\epsilon}(q)$ and H is of type $O_{n/2}(q)^2$. Here $qn/2$ is odd, $n \geqslant 10$ and $H = N.A$, where $N = \Omega_{n/2}(q) \times \Omega_{n/2}(q)$ and $A \leqslant [2^4].Z_f$. Note that $d(N) = 2$. If M contains N then $M = N.B$ with $B \leqslant A$ and thus $d(M) \leqslant 7$, so assume otherwise. Then $M = (M \cap N).A$ and $M \cap N$ is a maximal A-invariant subgroup of N. If M does not contain an element that interchanges the two $\Omega_{n/2}(q)$ factors of N then $M \cap N = C \times \Omega_{n/2}(q)$ or $\Omega_{n/2}(q) \times C$, where $C = D \cap \Omega_{n/2}(q)$ and D is maximal in an almost simple group with socle $\Omega_{n/2}(q)$. By Theorem 2.1 we have $d(C) \leqslant 4$, so $d(M \cap N) \leqslant 6$ and thus $d(M) \leqslant 11$. Now assume M has an element that interchanges the two $\Omega_{n/2}(q)$ factors. Then either $M \cap N$ is a diagonal subgroup isomorphic to $\Omega_{n/2}(q)$, or $M \cap N = C \times C$ with C as above. In the former case, $d(M \cap N) = 2$ and therefore $d(M) \leqslant 7$. Finally, suppose $M \cap N = C \times C$. Write $M_0 = (M \cap N).B$ with $B \leqslant Z_2 \times Z_2$. To obtain a generating set for M, take 4 generators for one of the factors C, take an element in M that swaps the two $\Omega_{n/2}(q)$ factors, and take two generators for B. These 7 elements generate a subgroup $M_0.2 \leqslant M$ and we can obtain a set of generators for M by choosing at most two further elements. We conclude that $d(M) \leqslant 9$.

Similar arguments apply in each of the remaining cases. For brevity, we only provide details in the two most difficult cases:

(a) $G_0 = \mathrm{L}_n^{\epsilon}(q)$ and H is of type $\mathrm{GL}_a^{\epsilon}(q) \wr S_t$;

(b) $G_0 = \mathrm{P}\Omega_n^{\epsilon}(q)$ and H is of type $O_a^{\epsilon'}(q) \wr S_t$, where $a \geqslant 2$ is even and q is odd.

Consider case (a). To begin with, let us assume $\epsilon = +$ and $a \geqslant 2$ (the special case $a = 1$ will be handled later). Note that $(a, q) \neq (2, 2)$ (see [7, 19]). Set $d = (a, q-1)$. By [19, Proposition 4.2.9] we have $H_0 = N_0.S_t$ and $H = N.S_t$, where $N_0 = A_0.B_0$, $N = A.B$ such that A_0 and A are sections of $(Z_{q-1})^t$, $B_0 = \mathrm{L}_a(q)^t.\frac{1}{d}(Z_d)^t$ and $B = \mathrm{L}_a(q)^t.C.2^b.Z_k$ where $C = \frac{1}{e}(Z_d)^t$ for some divisor e of d, $b \in \{0, 1\}$ and k is a divisor of $\log_p q$. Write $\mathrm{L}_a(q) = \langle x, y\rangle$, where x and y have coprime orders (see [8, Proposition 2.11]), and fix δ such that $\mathrm{PGL}_a(q) = \mathrm{L}_a(q).\langle\delta\rangle$. Also write $\mathbb{F}_q^{\times} = \langle\lambda\rangle$ and fix an element $\mu \in \mathbb{F}_q^{\times}$ of order d.

Suppose M contains N. Then $M = N.J$ with $J < S_t$ maximal, hence $M_0 = N_0.J$. If J is transitive on $\{1, \ldots, t\}$ then M_0 is a quotient of the subgroup of $\mathrm{GL}_a(q) \wr J$ generated by the elements $(\lambda, \lambda^{-1}, 1, \ldots, 1)$, $(\mu, 1, \ldots, 1)$, $(x, y, 1, \ldots, 1)$ and $(\delta, \delta^{-1}, 1, \ldots, 1)$ in $\mathrm{GL}_a(q)^t$, plus at most four generators for J, whence $d(M_0) \leqslant 8$ and the result follows. Similarly, if J is intransitive then $d(J) \leqslant 2$ and once again we deduce that $d(M_0) \leqslant 8$.

Now assume $N \nleqslant M$, so $M = (M \cap N).S_t$ and $M \cap N$ is a maximal S_t-invariant subgroup of N. Suppose A is not contained in M. Then $M = (M \cap A).B.S_t$ and $M \cap A$ is a maximal S_t-invariant subgroup of A. In other words, $(M \cap A).S_t$ is a maximal subgroup of $A.S_t$. Since $A.S_t$ is a quotient of a group of the form $\frac{1}{s}(Z_{q-1} \wr S_t)$ for some divisor s of $q-1$, Lemma 2.7 implies that $d((M \cap A).S_t) \leqslant 6$ and we deduce that $d(M) \leqslant 11$. Now assume M contains A. Set $\bar{M} = M/A$, $\bar{H} = H/A = B.S_t$ and let us assume that $(a, q) \neq (2, 3)$. Here $S = \mathrm{L}_a(q)^t$ is the unique minimal normal subgroup of $\bar{H}$, so we may assume that $\bar{M}$ contains S (if not, then Theorem 2.2 implies that $d(\bar{M}) \leqslant 10$ and thus $d(M) \leqslant 12$ since A is 2-generator as a normal subgroup of M). We now consider the quotient groups $\tilde{M} = \bar{M}/S$ and $\tilde{H} = \bar{H}/S = C.2^b.Z_k.S_t$. If $\tilde{M}$ does not contain $C = \frac{1}{e}(Z_d)^t$ then $\tilde{M} = (\tilde{M} \cap C).2^b.Z_k.S_t$ and $(\tilde{M} \cap C).S_t < C.S_t$ is maximal. Now Lemma 2.7 implies that $(\tilde{M} \cap C).S_t$ is 6-generator, hence $d(\tilde{M}) \leqslant 8$ so $d(\bar{M}) \leqslant 9$ and thus $d(M) \leqslant 11$. We have now reduced to the case where $C \leqslant \tilde{M}$, hence $M_0 = H_0$ and Theorem 2.1 implies that $d(M_0) \leqslant 4$.

Now assume $\epsilon = +$ and $(a, q) = (2, 3)$. As above, we may assume that $M = (M \cap N).S_t$ contains A, but the rest of the argument needs to be slightly modified since $\mathrm{L}_2(3) = A_4 = V_4{:}3$ is not simple. Set $\bar{M} = M/A$ and $\bar{H} = H/A = B.S_t$, where $B = (A_4)^t.C.2^b$. Now $D = (V_4)^t$ is the unique minimal normal subgroup of $\bar{H}$, so we may as well assume it is contained in $\bar{M}$. Set $\tilde{M} = \bar{M}/D$ and $\tilde{H} = \bar{H}/D = E.C.2^b.S_t$ with $E = 3^t$. If $\tilde{M}$ does not contain E then $\tilde{M} = (\tilde{M} \cap E).C.2^b.S_t$ and $(\tilde{M} \cap E).S_t < E.S_t$ is maximal, so $\tilde{M} \cap E = 3$ or 3^{t-1} and $(\tilde{M} \cap E).S_t$ is 3-generator. It follows that $d(\tilde{M}) \leqslant 6$, so $d(\bar{M}) \leqslant 8$ and $d(M) \leqslant 10$. We have now reduced to the case where $E \leqslant \tilde{M}$, so $S = (A_4)^t \leqslant M$ and the remainder of the previous argument now goes through.

To complete the analysis of the case $\epsilon = +$, we may assume that $a = 1$. Here $q \geqslant 5$ and $H = N.S_t$, where $N = A.2^b.Z_k.S_t$ with A, b and k as above. It is easy to reduce to the case where $M = (M \cap N).S_t$. If M contains A then $M_0 = H_0$ is 4-generator, so assume otherwise. Then $M = (M \cap A).2^b.Z_k.S_t$ and $(M \cap A).S_t < A.S_t$ is maximal. This is a situation we considered above, and by applying Lemma 2.7 we deduce that $d(M) \leqslant 8$.

A similar argument applies when $\epsilon = -$, so we will only give details in the special case $(a, q) = (3, 2)$. Here $\mathrm{U}_3(2) = 3^2{:}Q_8$, $H_0 = N_0.S_t$ and $H = N.S_t$, where $N_0 = A_0.B_0$, $N = A.B$ such that A_0 and A are sections of $(Z_3)^t$, $B_0 = \mathrm{U}_3(2)^t.\frac{1}{3}(Z_3)^t$ and $B = \mathrm{U}_3(2)^t.C.2^b$ where $C = \frac{1}{3}(Z_3)^t$ or $(Z_3)^t$ and $b \in \{0, 1\}$. It is straightforward to reduce to the case where $M = (M \cap N).S_t$, and by arguing as above we may assume that M contains A. Set $\bar{M} = M/A$ and $\bar{H} = H/A = (3^2{:}Q_8)^t.C.2^b.S_t$. Now $D = (3^2)^t$ is the unique minimal normal subgroup of $\bar{H}$, so we may assume that $\bar{M}$ contains D. Now set $\tilde{M} = \bar{M}/D$ and $\tilde{H} = \bar{H}/D = (Q_8)^t.C.2^b.S_t$. Let $E = (Q_8)^t$. If E is not contained in $\tilde{M}$ then $\tilde{M} = (\tilde{M} \cap E).C.2^b.S_t$ and $(\tilde{M} \cap E).S_t < E.S_t$ is maximal, so Lemma 2.7 implies that $(\tilde{M} \cap E).S_t$ is 6-generator and we deduce that $d(\bar{M}) \leqslant 9$ and $d(M) \leqslant 11$. On the other hand, if $E \leqslant \tilde{M}$ then M contains $\mathrm{U}_3(2)^t$ and we can complete the proof as above.

Finally, let us turn to case (b). Let D and D' denote the discriminants of the quadratic forms corresponding to $O_n^{\epsilon}(q)$ and $O_a^{\epsilon'}(q)$ (see [19, p.32], for example). To begin with, we will assume that $a \geqslant 4$ and $(a, \epsilon') \neq (4, +)$.

First assume $D' = \square$. By [8, Proposition 2.11] we have $\mathrm{P}\Omega_a^{\epsilon'}(q) = \langle x, y\rangle$, where $|x|$ and $|y|$ are coprime. Fix involutions r and s such that $\mathrm{PSO}_a^{\epsilon'}(q) = \mathrm{P}\Omega_a^{\epsilon'}(q).\langle s\rangle$, $\mathrm{PO}_a^{\epsilon'}(q) = \mathrm{PSO}_a^{\epsilon'}(q).\langle r\rangle$ and $[r, s] = 1$, so $\langle r, s\rangle = V_4$. By [19, Proposition 4.2.11] we have $H_0 = N_0.S_t$ and $H = N.S_t$ where

$$N_0 = 2^{t-1}.\mathrm{P}\Omega_a^{\epsilon'}(q)^t.2^{2(t-1)}, \quad N = 2^{t-1}.\mathrm{P}\Omega_a^{\epsilon'}(q)^t.2^{2(t-1)}.[2^b].Z_k \qquad (2)$$

with $0 \leqslant b \leqslant 3$ and k a divisor of $\log_p q$. Note that $[2^b] \leqslant D_8$ is 2-generator.

Suppose M contains N, so $M_0 = N_0.J$ for some maximal subgroup $J < S_t$. If J is transitive then M_0 is generated by $(-1, 1, \ldots, 1)$, $(x, y, 1, \ldots, 1)$, $(r, r, 1, \ldots, 1)$

and $(s,s,1,\ldots,1)$, together with at most 4 more for J. This gives $d(M_0) \leqslant 8$. Similarly, if J is intransitive then $d(J) \leqslant 2$ and we need at most 8 generators for M_0.

Now assume $N \not\leqslant M$, so $M = (M\cap N).S_t$ and $M \cap N$ is a maximal S_t-invariant subgroup of N. Write $N = A.B$, where $A = 2^{t-1}$ and $B = \mathrm{P}\Omega_a^{\epsilon'}(q)^t.2^{2(t-1)}.[2^b].Z_k$. If $A \not\leqslant M$ then $M = (M\cap A).B.S_t$ and $M\cap A$ is a maximal S_t-invariant subgroup of A. Therefore $M \cap A = 2^v$ with $v \in \{0, t-2\}$, so $d((M\cap A).S_t) \leqslant 3$ and we deduce that $d(M) \leqslant 9$. Now assume M contains A and set $\bar{M} = M/A$ and $\bar{H} = H/A$. Here $S = \mathrm{P}\Omega_a^{\epsilon'}(q)^t$ is the unique minimal normal subgroup of $\bar{H}$, so we may assume that $S \leqslant \bar{M}$ (if not, Theorem 2.2 implies that $d(\bar{M}) \leqslant d(\bar{H}) + 4 = 10$ and thus $d(M) \leqslant 11$). Set $\tilde{M} = \bar{M}/S$ and $\tilde{H} = \bar{H}/S = C.[2^b].Z_k.S_t$, where $C = 2^{2(t-1)}$. If $\tilde{M}$ contains C then $M_0 = H_0$ and thus $d(M_0) \leqslant 4$, so assume otherwise. Then $\tilde{M} = (\tilde{M} \cap C).[2^b].Z_k.S_t$ and $(\tilde{M} \cap C).S_t < \frac{1}{4}(V_4 \wr S_t)$ is maximal. Since $(\tilde{M}\cap C).S_t$ is 6-generator by Lemma 2.7, we conclude that $d(\tilde{M}) \leqslant 6+2+1 = 9$, so $d(\bar{M}) \leqslant 10$ and thus $d(M) \leqslant 11$ as required.

Next suppose that $D' = \boxtimes$, so $\mathrm{P}\Omega_a^{\epsilon'}(q) = \Omega_a^{\epsilon'}(q) = \mathrm{PSO}_a^{\epsilon'}(q)$. We continue to assume that $a \geqslant 4$ and $(a,\epsilon') \neq (4,+)$. By [19, Proposition 4.2.11] we have $H_0 = N_0.S_t$ and $H = N.S_t$ where

$$N_0 = 2^d \times \Omega_a^{\epsilon'}(q)^t.2^{t-1}, \quad N = 2^e \times \Omega_a^{\epsilon'}(q)^t.2^b.2^c.Z_k$$

with $b \in \{t-1, t\}$, $c \in \{0,1\}$ and k a divisor of $\log_p q$. Also, $d = e = t-1$ if t is odd, otherwise $d = t-2$ and $e \in \{t-2, t-1\}$. Define the elements x, y and r as above. It is straightforward to reduce to the case where $M = (M \cap N).S_t$.

Write $N = A \times B$, where $A = 2^e$ and $B = \Omega_a^{\epsilon'}(q)^t.2^b.2^c.Z_k$. If M contains A then we may assume that $\bar{M} = M/A$ contains $\Omega_a^{\epsilon'}(q)^t$, which is the unique minimal normal subgroup of $\bar{H} = H/A$. Therefore, $M_0 = (2^d \times \Omega_a^{\epsilon'}(q)^t.2^v).S_t$ and the S_t-invariance of $M_0 \cap N$ implies that $v \in \{0, 1, t-1\}$, so $d(M_0) \leqslant 5$. Now assume $A \not\leqslant M$, so $M = (M \cap A).B.S_t$ and $M \cap A$ is a maximal S_t-invariant subgroup of A. Therefore $M \cap A = 2^v$ with $v \in \{0, t-2\}$, so $d((M \cap A).S_t) \leqslant 3$ and $d(M) \leqslant 7$.

To complete the proof, we may assume that $(a,\epsilon') = (4,+)$ or $a = 2$. Suppose $(a,\epsilon') = (4,+)$. Define the involutions r and s as above and note that $D' = \square$, $\mathrm{P}\Omega_4^+(q) = \mathrm{L}_2(q) \times \mathrm{L}_2(q)$ and $\mathrm{P}\Omega_4^+(q).\langle r\rangle = \mathrm{L}_2(q) \wr S_2$. If $q \geqslant 5$ then we can still write $\mathrm{P}\Omega_4^+(q) = \langle x, y\rangle$, where $|x|$ and $|y|$ are coprime, but this is not possible when $q = 3$ (note that $\mathrm{P}\Omega_4^+(3) = A_4 \times A_4$ can be generated by x and y, where $|x| = 6$ and $|y| = 3$). As above, we have $H_0 = N_0.S_t$ and $H = N.S_t$, where N_0 and N are given in (2). One now checks that the argument above goes through essentially unchanged. Indeed, the only difference is for $q = 3$, where we require two generators for $\mathrm{P}\Omega_4^+(q)^t$ as a normal subgroup of $\mathrm{P}\Omega_4^+(q)^t.S_t$, rather than one. However, it is clear that the desired bound $d(M) \leqslant 12$ still holds in this case.

For example, if $q = 3$ and M contains N then we get $d(M_0) \leqslant 9$ and the result follows.

Finally, suppose $a = 2$. We will assume $q \equiv \epsilon' \pmod 4$ (the other case is very similar), so $D' = \square$ and $\mathrm{P\Omega}_2^{\epsilon'}(q) = Z_m$ is cyclic, where $m = (q-\epsilon')/4$. Write $H = N.S_t$, where $N = A.B$, $A = 2^{t-1}$ and $B = (Z_m)^t.2^{2(t-1)}.[2^b].Z_k$ with $0 \leqslant b \leqslant 3$ and k a divisor of $\log_p q$. In the usual way, we reduce to the case where $M = (M \cap N).S_t$. If M does not contain A then $M = (M \cap A).B.S_t$, where $M \cap A$ is a maximal S_t-invariant subgroup of A, so $M \cap A = 2^v$ with $v \in \{0, t-2\}$. Therefore, $d((M \cap A).S_t) \leqslant 3$ and we deduce that $d(M) \leqslant 9$. Now assume M contains A. Set $\bar{M} = M/A$ and $\bar{H} = H/A = C.2^{2(t-1)}.[2^b].Z_k.S_t$, where $C = (Z_m)^t$. If $C \nleqslant \bar{M}$ then $\bar{M} = (\bar{M} \cap C).2^{2(t-1)}.[2^b].Z_k.S_t$ and $(\bar{M} \cap C).S_t < Z_m \wr S_t$ is maximal. Therefore Lemma 2.7 implies that $d((\bar{M} \cap C).S_t) \leqslant 6$, so $d(\bar{M}) \leqslant 11$ and thus $d(M) \leqslant 12$. Now assume $C \leqslant \bar{M}$ and set $\tilde{M} = \bar{M}/C$ and $\tilde{H} = \bar{H}/C = D.[2^b].Z_k.S_t$, where $D = 2^{2(t-1)}$. If $D \nleqslant \tilde{M}$ then $\tilde{M} = (\tilde{M} \cap D).[2^b].Z_k.S_t$ and $(\tilde{M} \cap D).S_t$ is a maximal subgroup of $D.S_t = \frac{1}{4}(V_4 \wr S_t)$. By Lemma 2.7 we have $d((\tilde{M} \cap D).S_t) \leqslant 6$, so $d(\tilde{M}) \leqslant 9$, $d(\bar{M}) \leqslant 10$ and thus $d(M) \leqslant 11$. Finally, if $\tilde{M}$ contains D then $M_0 = H_0$ and $d(M_0) \leqslant 4$. □

Lemma 5.5. *Proposition* 5.1 *holds if $H \in \mathcal{C}_3$.*

Proof. First assume $G_0 = \mathrm{L}_n^\epsilon(q)$, so H is of type $\mathrm{GL}_{n/k}^\epsilon(q^k)$ for some prime k (note that k is odd if $\epsilon = -$). If $n = k$ then $H_0 = Z_a.Z_k$ for some $a \geqslant 1$ (see [19, Proposition 4.3.6]) and thus $d(M_0) \leqslant 2$. On the other hand, if $n > k$ then H^∞ is quasisimple and irreducible, so Lemma 5.2 implies that $d(M) \leqslant 9$.

Next suppose $G_0 = \mathrm{PSp}_n(q)$. If H is of type $\mathrm{Sp}_{n/k}(q^k)$, or if $n \geqslant 6$ and H is of type $\mathrm{GU}_{n/2}(q)$, then the result follows from Lemma 5.2. Now assume $n = 4$ and H is of type $\mathrm{GU}_2(q)$, so $q \geqslant 5$ is odd (see [7, Table 8.12]). Here H^∞ is reducible (see [19, Lemma 4.3.2]) so we need to argue differently. According to [19, Proposition 4.3.7] we have

$$H_0 = Z_{(q+1)/2}.(\mathrm{PGU}_2(q) \times Z_2).$$

In general, $H = N.A$ where $N = Z_{(q+1)/2}$ or Z_{q+1}, $A/Z(A)$ has socle $\mathrm{L}_2(q)$ and $Z(A) \leqslant Z_2$. If M contains N then M/N is a maximal subgroup of $H/N \cong A$ and we deduce that $d(M) \leqslant 8$ since every maximal subgroup of $A/Z(A)$ is 6-generator by Theorem 2.1. On the other hand, if $N \nleqslant M$ then $M = (M \cap N).A$ and $d(M) \leqslant d(M \cap N) + d(A) \leqslant 5$.

Finally, suppose $G_0 = \mathrm{P\Omega}_n^\epsilon(q)$. If $n \equiv 2 \pmod 4$ and H is of type $O_{n/2}(q^2)$ (with q odd) then H^∞ is quasisimple and irreducible, so the result follows from Lemma 5.2. The same argument applies if n is even and H is of type $\mathrm{GU}_{n/2}(q)$. Finally, let us assume that H is of type $O_{n/k}^\epsilon(q^k)$, where k is a prime and $n/k \geqslant 3$.

By applying Lemma 5.2 we reduce to the case where H is of type $O_4^+(q^k)$, so $\epsilon = +$ and

$$H_0 = \mathrm{P}\Omega_4^+(q^k).[\ell] = (\mathrm{L}_2(q^k) \times \mathrm{L}_2(q^k)).[\ell] = N.[\ell],$$

where $\ell = (1+\delta_{2,k})k$ (see [19, Proposition 4.3.14]). Then $N < H/Z(H) \leqslant \mathrm{Aut}(N)$ and $Z(H) \leqslant Z_2$. If M contains $Z(H)$ then Lemma 2.8 implies that $d(M/Z(H)) \leqslant 8$ and thus $d(M) \leqslant 9$. Otherwise $M \cong H/Z(H)$ and $d(M) \leqslant 6$ by Lemma 2.8. □

Lemma 5.6. *Proposition* 5.1 *holds if* $H \in \mathcal{C}_4$.

Proof. First assume $G_0 = \mathrm{L}_n^\epsilon(q)$ and H is of type $\mathrm{GL}_a^\epsilon(q) \otimes \mathrm{GL}_b^\epsilon(q)$, where $n = ab$ and $2 \leqslant a < b$. By [19, Proposition 4.4.10] we have $H = N.A$ where $N = \mathrm{L}_a^\epsilon(q) \times \mathrm{L}_b^\epsilon(q)$ and $A \leqslant (Z_{(a,q-\epsilon)} \times Z_{(b,q-\epsilon)}).(Z_f \times Z_2)$. Since $d(N) = 2$ we deduce that $d(M) \leqslant 6$ if M contains N, so assume otherwise. Then $M = (M \cap N).A$ and $d(A) \leqslant 4$, so it suffices to show that $d(M \cap N) \leqslant 8$. Now $M \cap N$ is a maximal A-invariant subgroup of N, so $M \cap N = C \times \mathrm{L}_b^\epsilon(q)$ or $\mathrm{L}_a^\epsilon(q) \times D$, where $C = E \cap \mathrm{L}_a^\epsilon(q)$ for some maximal subgroup E of a group F with $\mathrm{L}_a^\epsilon(q) \leqslant F \leqslant \mathrm{Aut}(\mathrm{L}_a^\epsilon(q))$, and similarly for D. By applying Theorem 2.1 and Lemma 2.9 we deduce that C and D are 4-generator, so $d(M \cap N) \leqslant 6$.

Next suppose $G_0 = \mathrm{PSp}_n(q)$ and H is of type $\mathrm{Sp}_a(q) \otimes O_b^\epsilon(q)$, where $n = ab$, $b \geqslant 3$ and q is odd. Here $H = N.A$, where $N = \mathrm{PSp}_a(q) \times \mathrm{P}\Omega_b^\epsilon(q)$ and $A \leqslant [2^3].(Z_f \times Z_2)$. In particular, $d(M) \leqslant 9$ if M contains N. Otherwise $M = (M \cap N).A$, where $M \cap N$ is a maximal A-invariant subgroup of N, and it suffices to show that $d(M \cap N) \leqslant 7$. If both factors of N are simple then $M \cap N = C \times \mathrm{P}\Omega_b^\epsilon(q)$ or $\mathrm{PSp}_a(q) \times D$, where $C = E \cap \mathrm{PSp}_a(q)$ and E is maximal in an almost simple group with socle $\mathrm{PSp}_a(q)$, and similarly for D. By applying Theorem 2.1 we deduce that $d(M \cap N) \leqslant 6$. A very similar argument applies if $(a,q) = (2,3)$, $(b,q) = (3,3)$ or $(b,\epsilon) = (4,+)$, using Lemmas 2.8 and 2.9.

Finally, let us assume $G_0 = \mathrm{P}\Omega_n^\epsilon(q)$. The usual argument applies if H is of type $\mathrm{Sp}_a(q) \otimes \mathrm{Sp}_b(q)$, so let us take H to be of type $O_a^{\epsilon_1}(q) \otimes O_b^{\epsilon_2}(q)$. Here q is odd, $a,b \geqslant 3$ and $(a,\epsilon_1) \neq (b,\epsilon_2)$. For brevity, we will assume that $\epsilon_1 = \epsilon_2 = +$, so $\epsilon = +$ and $4 \leqslant a < b$ (the other cases are very similar). By [19, Proposition 4.4.14] we have $H = N.A$, where $N = \mathrm{P}\Omega_a^+(q) \times \mathrm{P}\Omega_b^+(q)$ and $A \leqslant (D_8 \times D_8).Z_f$. Note that $d(N) \leqslant 4$ and every subgroup of $(D_8 \times D_8).Z_f$ is 5-generator. In particular, if M contains N then $d(M) \leqslant 9$, so assume otherwise. Then $M = (M \cap N).A$ and the usual argument (using Theorem 2.1 and Lemma 2.8) shows that $d(M \cap N) \leqslant 6$, whence $d(M) \leqslant 11$. □

Lemma 5.7. *Proposition* 5.1 *holds if* $H \in \mathcal{C}_5$.

Proof. First assume $G_0 = \mathrm{L}_n^{\epsilon}(q)$ and H is of type $\mathrm{GL}_n^{\epsilon}(q_0)$, where $q = q_0^k$ for a prime k (with k odd if $\epsilon = -$). Note that $(n, q_0) \neq (2,2)$ (see [7, Table 8.1]). If $(n, q_0) = (2,3)$ then $H \cong A \times B$, where $A \in \{A_4, S_4\}$ and $B \leqslant Z_k$, and we deduce that $d(M) \leqslant 3$. The same conclusion holds if $\epsilon = -$ and $(n, q_0) = (3,2)$. In every other case, Lemma 5.2 implies that $d(M) \leqslant 9$. Similarly, we can apply Lemma 5.2 if G_0 is symplectic or orthogonal, and also if $G_0 = \mathrm{U}_n(q)$ and H is of type $\mathrm{Sp}_n(q)$.

Finally, let us assume $G_0 = \mathrm{U}_n(q)$ and H is of type $O_n^{\epsilon}(q)$ (so q is odd and $n \geqslant 3$). In view of Lemma 5.2 we may assume that $(n, \epsilon) = (4, +)$ (note that $(n,q) \neq (3,3)$; see [7, Table 8.5]). Here $q \geqslant 5$ (see [7, Table 8.10]) and

$$H_0 = \mathrm{PSO}_4^+(q).2 = (\mathrm{L}_2(q) \times \mathrm{L}_2(q)).[2^2] = N.[2^2]. \tag{3}$$

More precisely, $N < H/Z(H) \leqslant \mathrm{Aut}(N)$ with $Z(H) \leqslant Z_2$, and the result follows by applying Lemma 2.8. □

Lemma 5.8. *Proposition* 5.1 *holds if* $H \in \mathcal{C}_6$.

Proof. First assume $G_0 = \mathrm{L}_n(q)$ and H is of type $r^{1+2m}.\mathrm{Sp}_{2m}(r)$, where $n = r^m$ and r is an odd prime. If $n = 3$ then $q = p \equiv 1 \pmod 3$ (see [7, Table 8.3]), $3^2{:}Q_8 \leqslant H \leqslant \mathrm{AGL}_2(3)$ and it is easy to check that $d(M) \leqslant 3$. Now assume $n \geqslant 5$, in which case

$$H = W{:}(\mathrm{Sp}_{2m}(r).A) \leqslant W{:}\mathrm{GSp}_{2m}(r)$$

and $A \leqslant Z_{2f}$, where $W = r^{2m}$ and $q = p^f$, with f an odd divisor of $r-1$ (see [19, Proposition 4.6.5]). Since W is the unique minimal normal subgroup of H we may assume that $M = W.J$ for some maximal subgroup $J < \mathrm{Sp}_{2m}(r).A$, so Lemma 2.5 implies that $d(M) \leqslant 8$. An entirely similar argument applies if $G_0 = \mathrm{U}_n(q)$. If $G_0 = \mathrm{L}_2(q)$ and H is of type $2_-^{1+2}.O_2^-(2)$ then $H = A_4$ or S_4 and the result follows.

Next assume $G_0 = \mathrm{P}\Omega_n^+(q)$ and H is of type $2_+^{1+2m}.O_{2m}^+(2)$, so $q = p \geqslant 3$ and $n = 2^m$ with $m \geqslant 3$. By [19, Proposition 4.6.8] we have $H = W.A$ with $W = 2^{2m}$ and $A = \Omega_{2m}^+(2)$ or $O_{2m}^+(2)$. In particular, W is the unique minimal normal subgroup of H so we may assume that $M = W.J$ with J maximal in A. By applying Lemma 2.5 we deduce that $d(M) \leqslant 1 + d(J) \leqslant 7$. The case where $G_0 = \mathrm{PSp}_n(q)$ and H is of type $2_-^{1+2m}.O_{2m}^-(2)$ is entirely similar. □

Lemma 5.9. *Proposition* 5.1 *holds if* $H \in \mathcal{C}_7$.

Proof. We refer the reader to [19, Table 4.7.A] for the list of cases that we need to consider. First assume $G_0 = \mathrm{L}_n^{\epsilon}(q)$ and H is of type $\mathrm{GL}_a^{\epsilon}(q) \wr S_t$ with $a \geqslant 3$. Here $n = a^t$ and $(a, q, \epsilon) \neq (3, 2, -)$. We will assume $\epsilon = +$ since the case $\epsilon = -$ is very similar. To begin with, let us assume that at least one of the following three conditions does *not* hold:

$$t = 2, \quad a \equiv 2 \,(\mathrm{mod}\ 4), \quad q \equiv -1 \,(\mathrm{mod}\ 4). \tag{4}$$

Write $\mathrm{PGL}_a(q) = \mathrm{L}_a(q).\langle\delta\rangle$ and $\mathrm{L}_a(q) = \langle x, y\rangle$ where $|x|$ and $|y|$ are coprime. Set $d = (a, q-1)$. According to [19, Proposition 4.7.3] we have $H_0 = N_0.S_t$ and $H = N.S_t$, where

$$N_0 = \mathrm{L}_a(q)^t.A_0, \quad N = \mathrm{L}_a(q)^t.A.2^b.Z_k$$

where $A_0 = \frac{1}{c}(Z_d)^t \leqslant \frac{1}{e}(Z_d)^t = A$, $b \in \{0, 1\}$ and k divides $\log_p q$, for some divisors c, e of d. If M contains N then $M_0 = N_0.J$ for some maximal subgroup $J < S_t$ and the result quickly follows. For example, if J is a transitive subgroup then M_0 is generated by $(x, y, 1, \ldots, 1)$, $(\delta, \delta^{-1}, 1, \ldots, 1)$ and $(\delta^{\ell}, 1, \ldots, 1)$ for some $\ell \geqslant 0$, together with at most 4 generators for J.

Now assume $N \nleqslant M$, so $M = (M \cap N).S_t$ and $M \cap N$ is a maximal S_t-invariant subgroup of N. Since $S = \mathrm{L}_a(q)^t$ is the unique minimal normal subgroup of H, we may assume that M contains S. Set $\bar{M} = M/S$ and $\bar{H} = H/S = A.2^b.Z_k.S_t$. If $\bar{M}$ contains A then $M_0 = H_0$ and thus $d(M_0) \leqslant 4$, so assume otherwise. Then $\bar{M} = (\bar{M} \cap A).2^b.Z_k.S_t$ and $(\bar{M} \cap A).S_t < A.S_t$ is maximal, so Lemma 2.7 implies that $d((\bar{M} \cap A).S_t) \leqslant 6$ and we deduce that $d(\bar{M}) \leqslant 8$ and $d(M) \leqslant 10$.

To complete the analysis of this case, we may assume that all of the conditions in (4) are satisfied. The above argument goes through unchanged if H contains an element that interchanges the two copies of $\mathrm{L}_a(q)$ in the socle of H, so we may assume that

$$H = (N_1 \times N_2).A.2^b.Z_k,$$

where $N_i = \mathrm{L}_a(q)$ and A, b and k are as above. Note that N_1 and N_2 are the minimal normal subgroups of H. If M contains both of these subgroups then the previous argument goes through, so we may assume that M contains N_1 but not N_2. Set $\bar{M} = M/N_1$ and $\bar{H} = H/N_1 = N_2.A.2^b.Z_k$. Since $N_2 \nleqslant \bar{M}$ we have $\bar{M} = (\bar{M} \cap N_2).A.2^b.Z_k$ and $\bar{M} \cap N_2 = \mathrm{L}_a(q) \cap B$ where B is a maximal subgroup of an almost simple group with socle $\mathrm{L}_a(q)$. By Theorem 2.1 we have $d(\bar{M} \cap N_2) \leqslant 4$, so $d(\bar{M}) \leqslant 8$ and thus $d(M) \leqslant 10$.

The remaining $\mathcal{C}_7$ cases are similar, so we only give details in the situation where $G_0 = \mathrm{P\Omega}_n^{+}(q)$ and H is of type $O_a^{+}(q) \wr S_t$, with $a \geqslant 6$ and q odd. Let D and D' be the discriminants of the quadratic forms corresponding to $O_n^{+}(q)$ and $O_a^{+}(q)$, respectively. Note that $n(q-1)/4$ is always even, so $D = \square$ (see [19, Proposition 2.5.10]). Write $\mathrm{PO}_a^{+}(q) = \mathrm{PSO}_a^{+}(q).\langle r\rangle$ and $\mathrm{PGO}_a^{+}(q) = \mathrm{PO}_a^{+}(q).\langle\delta\rangle$ for involutions r and δ. Also fix $x, y \in \mathrm{PSO}_a^{+}(q)$ such that $\mathrm{PSO}_a^{+}(q) = \langle x, y\rangle$. Two cases require special attention:

(a) $t = 2$ and $a \equiv 2 \pmod 4$;

(b) $t = 3$, $a \equiv 2 \pmod 4$ and $q \equiv 3 \pmod 4$.

For now, we will assume that we are not in one of these cases. By [19, Proposition 4.7.6] we have $H_0 = N_0.S_t$ and $H = N.S_t$, where

$$N_0 = \mathrm{PSO}_a^+(q)^t.[2^{2t-1}], \quad N = \mathrm{PSO}_a^+(q)^t.[2^i].Z_k$$

with $i \in \{2t-1, 2t\}$ and k a divisor of $\log_p q$. If M contains N then $M_0 = N_0.J$ with $J < S_t$ maximal and the result quickly follows. For example, if J is transitive then M_0 is generated by $(x, 1, \ldots, 1)$, $(y, 1, \ldots, 1)$, $(r, 1, \ldots, 1)$ and $(\delta, \delta^{-1}, 1, \ldots, 1)$, together with at most 4 generators for J.

Now assume $M = (M \cap N).S_t$. First consider the case where $D' = \square$, so $\mathrm{PSO}_a^+(q) = \mathrm{P\Omega}_a^+(q).2$ and $S = \mathrm{P\Omega}_a^+(q)^t$ is the unique minimal normal subgroup of H. As usual, we may assume that M contains S, so set $\bar{M} = M/S$ and $\bar{H} = H/S = A.Z_k.S_t$, where $A = [2^{t+i}]$. If $\bar{M}$ contains A then $M_0 = H_0$ and thus $d(M_0) \leqslant 4$. Otherwise, $\bar{M} = (\bar{M} \cap A).Z_k.S_t$ and $(\bar{M} \cap A).S_t$ is a maximal subgroup of $A.S_t = \frac{1}{b}(D_8 \wr S_t)$, where $b = 1$ or 2. By Lemma 2.7 we have $d((\bar{M} \cap A).S_t) \leqslant 6$, so $d(\bar{M}) \leqslant 7$ and thus $d(M) \leqslant 9$. A similar argument applies if $D' = \boxtimes$. Here $\mathrm{PSO}_a^+(q) = \mathrm{P\Omega}_a^+(q)$ and once again we may assume that M contains $S = \mathrm{P\Omega}_a^+(q)^t$. The rest of the argument goes through, replacing D_8 by V_4.

It remains to handle the cases described in (a) and (b) above. First consider (a). We will assume $D' = \square$ (the other case is very similar), so $H = \mathrm{P\Omega}_a^+(q)^2.[2^{b+2}].Z_k.Z_c$, where $b \in \{2, 3, 4\}$, $c \in \{1, 2\}$ and k divides $\log_p q$. Note that $[2^{b+2}] \leqslant D_8 \times D_8$ is 4-generator. If $c = 2$ then the previous argument goes through, so let us assume $c = 1$. Here H has two minimal normal subgroups N_1 and N_2, both isomorphic to $\mathrm{P\Omega}_a^+(q)$. If M contains $S = N_1 \times N_2$ then $M/S < H/S = [2^{b+2}].Z_k$, so $d(M/S) \leqslant 5$ and thus $d(M) \leqslant 7$. Therefore, we may assume that H contains N_1 but not N_2. Set $\bar{M} = M/N_1$ and $\bar{H} = H/N_1 = N_2.[2^{b+2}].Z_k$. Then $\bar{M} = (\bar{M} \cap N_2).[2^{b+2}].Z_k$ and Theorem 2.1 implies that $d(\bar{M} \cap N_2) \leqslant 4$, so $d(\bar{M}) \leqslant 9$ and thus $d(M) \leqslant 11$.

Finally, let us assume that the conditions in (b) hold, so $D' = \boxtimes$ and $H = S.A.Z_k.B$, where $S = \mathrm{P\Omega}_a^+(q)^3$, $A = [2^b]$ with $b \in \{5, 6\}$, k divides $\log_p q$ and $B \in \{Z_3, S_3\}$. In the usual way, it is easy to reduce to the case where $M = (M \cap N).B$. Now B acts transitively on the factors of S, so S is the unique minimal normal subgroup of H and we may assume that M contains S. Set $\bar{M} = M/S$ and $\bar{H} = H/S = A.Z_k.B$. If $\bar{M}$ contains A then $M_0 = H_0$ and $d(M_0) \leqslant 4$, so assume otherwise. Then $\bar{M} = (\bar{M} \cap A).Z_k.B$ and $(\bar{M} \cap A).B$ is a maximal subgroup of $A.B = \frac{1}{c}(V_4 \wr B)$, where $c \in \{1, 2\}$. Therefore $d((\bar{M} \cap A).B) \leqslant 6$ by Lemma 2.7, so $d(\bar{M}) \leqslant 7$ and thus $d(M) \leqslant 9$ since S is 2-generator. □

Lemma 5.10. *Proposition 5.1 holds if $H \in \mathcal{C}_8$.*

Proof. First assume $G_0 = \mathrm{L}_n(q)$. If H is of type $\mathrm{Sp}_n(q)$, then $n \geqslant 4$ and Lemma 5.2 applies. Next suppose H is of type $O_n^{\epsilon}(q)$. If $(n,q) \neq (3,3)$ and $(n,\epsilon) \neq (4,+)$, then we can use Lemma 5.2 once again. It is easy to check that $d(M) \leqslant 3$ if $(n,q) = (3,3)$. If $(n,\epsilon) = (4,+)$ then (3) holds and we can repeat the argument in the proof of Lemma 5.7. Finally, suppose that H is of type $\mathrm{U}_n(q_0)$, where $n \geqslant 3$ and $q = q_0^2$. If $(n,q) = (3,4)$ then $d(M) \leqslant 3$, otherwise the result follows from Lemma 5.2.

Finally let us assume that $G_0 = \mathrm{PSp}_n(q)$ and H is of type $O_n^{\epsilon}(q)$, where q is even, $n \geqslant 4$ and $(n,q) \neq (4,2)$. If $(n,\epsilon) \neq (4,+)$ then H is almost simple and thus $d(M) \leqslant 6$. On the other hand, if $(n,\epsilon) = (4,+)$ then

$$H_0 = O_4^+(q) = \mathrm{L}_2(q) \wr S_2 = (\mathrm{L}_2(q) \times \mathrm{L}_2(q)).2 = N.2$$

and $N < H \leqslant \mathrm{Aut}(N)$, so Lemma 2.8 implies that $d(M) \leqslant 8$. □

To complete the proof of Proposition 5.1, it remains to deal with certain *novelty* subgroups H of G, where $H_0 = H \cap G_0$ is non-maximal in G_0. In view of [2] and our earlier work, we may assume that one of the following holds:

(a) $G_0 = \mathrm{PSp}_4(q)$, q even and G contains a graph-field automorphism;

(b) $G_0 = \mathrm{P\Omega}_8^+(q)$ and G contains a triality automorphism.

In [2, Section 14], Aschbacher proves a version of his main theorem which describes the various possibilities for H in case (a), but his theorem does not apply in case (b); here the possibilities were determined later by Kleidman [18]. We record the relevant non-parabolic subgroups in Table 2. Note that in case (a) we may assume $q > 2$ since $\mathrm{PSp}_4(2)' \cong A_6$.

G_0	Type of H	Conditions
$\mathrm{PSp}_4(q)$	$O_2^{\epsilon}(q) \wr S_2$	$q > 2$ even
	$O_2^-(q^2).2$	$q > 2$ even
$\mathrm{P\Omega}_8^+(q)$	$\mathrm{GL}_3^{\epsilon}(q) \times \mathrm{GL}_1^{\epsilon}(q)$	
	$O_2^-(q^2) \times O_2^-(q^2)$	
	$[2^9].\mathrm{SL}_3(2)$	$q = p > 2$

Table 2. Some novelty subgroups

Lemma 5.11. *Proposition 5.1 holds if $G_0 = \mathrm{PSp}_4(q)$ and H is in Table 2.*

Proof. Here $H_0 = D_{2(q\pm 1)} \wr S_2$ or $Z_{q^2+1}.4$, so $d(M_0) \leqslant 4$ and the result follows. □

Lemma 5.12. *Proposition 5.1 holds if $G_0 = \mathrm{P}\Omega_8^+(q)$ and H is in Table 2.*

Proof. As before, it suffices to show that $d(M_0) \leqslant 9$. First assume H is of type $\mathrm{GL}_3^\epsilon(q) \times \mathrm{GL}_1^\epsilon(q)$. Set $d = (2, q-1)$. Working in $\Omega_8^+(q)$ we have $H_0 = N_0.Z_{(q-\epsilon)/d}.[2^2]$ and $H = N_0.A$, where $N_0 = \frac{1}{d}\mathrm{GL}_3^\epsilon(q)$ and $A = Z_{(q-\epsilon)/d}.[2^a].B.Z_k$ with $a \in \{2,3,4\}$, $B \in \{Z_3, S_3\}$ and k a divisor of $\log_p q$. If M contains N_0 then $M_0 = N_0.C$ and $C \leqslant Z_{(q-\epsilon)/d}.[2^2]$ is 3-generator, so $d(M_0) \leqslant 5$. Now assume $N_0 \not\leqslant M$, so $M = (M \cap N_0).A$ and $M_0 = (M \cap N_0).Z_{(q-\epsilon)/d}.[2^2]$, where $M \cap N_0$ is a maximal A-invariant subgroup of N_0. Now $M \cap \mathrm{SL}_3^\epsilon(q) = D \cap \mathrm{SL}_3^\epsilon(q)$, where D is maximal in a group E of the form

$$\mathrm{SL}_3^\epsilon(q) \leqslant E \leqslant \Gamma\mathrm{L}_3^\epsilon(q).\langle\gamma\rangle,$$

where γ is a graph automorphism. By applying Theorem 2.1 we deduce that $d(M \cap \mathrm{SL}_3^\epsilon(q)) \leqslant 5$, so $d(M \cap N_0) \leqslant 6$ and thus $d(M_0) \leqslant 9$ as required.

If H is of type $O_2^-(q^2) \times O_2^-(q^2)$ then $H_0 = (D_{2l} \times D_{2l}).2^2$, where $l = (q^2+1)/(2, q-1)$ is odd, and we deduce that $d(M_0) \leqslant 5$ since every subgroup of $D_{2l} \times D_{2l}$ is 3-generator. In the final case we have $H_0 = [2^9].\mathrm{SL}_3(2)$ and using Magma one can check that every subgroup of H_0 is 8-generator. In particular, $d(M_0) \leqslant 8$ and the result follows. □

This completes the proof of Proposition 5.1.

6. Exceptional groups

In this section we turn to the exceptional groups of Lie type, establishing Theorem 1 for the second maximal subgroups lying in a maximal non-parabolic subgroup.

Proposition 6.1. *Suppose $M < H < G$ with each subgroup maximal in the next, where G is an almost simple exceptional group of Lie type and H is non-parabolic. Then $d(M) \leqslant 12$.*

Proof. Let G_0 be the socle of G, and $H_0 = H \cap G_0$, $M_0 = M \cap G_0$. Write $G_0 = G(q)$, an exceptional simple group of Lie type over $\mathbb{F}_q$, where $q = p^e$, p prime. With the aid of Magma, it is easy to check that $d(M) \leqslant 4$ if $G_0 = {}^2F_4(2)'$, $G_2(3)$ or ${}^3D_4(2)$, so we may assume otherwise. As $d(G/G_0) \leqslant 2$ it is sufficient to show that $d(M_0) \leqslant 10$.

According to [23, Theorem 2], the possibilities for H_0 are as follows:

(i) H_0 is almost simple;

(ii) $H_0 = N_{G_0}(K)$, where K is a reductive subgroup of G_0 of maximal rank, not a maximal torus; the possibilities are listed in [22, Table 5.1];

(iii) $H_0 = N_{G_0}(T)$, where T is a maximal torus of G_0; the possibilities are listed in [22, Table 5.2];

(iv) The generalized Fitting subgroup $F^*(H_0)$ is as in [23, Table III];

(v) $H_0 = N_{G_0}(E)$, where E is an elementary abelian group given in [11, Theorem 1(II)].

In case (i), $d(M_0) \leqslant 4$ by Theorem 2.1.

In case (iv), with two exceptions H_0 has a subgroup H_1 of index at most 6 that is a direct product $L_1 \times L_2$ of non-isomorphic simple groups L_1, L_2; in the exceptional cases, H_0 has a subgroup $H_1 \cong \mathrm{L}_2(q)^2$ or $\mathrm{L}_2(q) \times G_2(q)^2$ of index dividing 4. Excluding the exceptional cases, we must have $M_0 \cap H_1 = L_1 \times M_2$ where either $M_2 = L_2$ or M_2 is a maximal H-invariant subgroup of L_2. Using Theorem 2.1 we see that $d(M_2) \leqslant 4$, so $d(M_0) \leqslant 8$. The first exceptional case $H_1 = L_1 \times L_2 \cong \mathrm{L}_2(q)^2$ is entirely similar: either $M_0 \cap H_1 = L_1 \times M_2$ as above, or it is a diagonal subgroup isomorphic to $\mathrm{L}_2(q)$. In the second exceptional case, the two $G_2(q)$ factors are interchanged by an element of H_0, so either $M_0 \cap H_1$ is $M_1 \times G_2(q)^2$ with M_1 maximal H-invariant in $\mathrm{L}_2(q)$, or it is $\mathrm{L}_2(q) \times D$ where D is a diagonal subgroup of $G_2(q)^2$ isomorphic to $G_2(q)$. In every case we easily see that $d(M_0) \leqslant 8$ using Theorem 2.1.

Next consider case (v). In this case, either H_0 is one of the groups

$$5^3.\mathrm{SL}_3(5),\ 2^{5+10}.\mathrm{SL}_5(2),\ 3^{3+3}.\mathrm{SL}_3(3),\ 3^3.\mathrm{SL}_3(3),\ 2^3.7,\ 2^3.\mathrm{SL}_3(2), \qquad (5)$$

or $G_0 = E_7(q)$ and $H_0 = (2^2 \times \mathrm{P}\Omega_8^+(q).2).S_3$ with q odd. In the latter case, either M_0 contains $\mathrm{P}\Omega_8^+(q)$ in which case $d(M_0) \leqslant 4$, or $M_0 \cap \mathrm{P}\Omega_8^+(q)$ is a maximal S_3-invariant subgroup, in which case $d(M_0) \leqslant 7$ by Theorem 2.1. The only problematic possibility in the list (5) is $H_0 = 2^{5+10}.\mathrm{SL}_5(2)$. Let $P = O_2(H_0)$. Then $\Phi(P) = 2^5$ and $H_0/P \cong \mathrm{SL}_5(2)$ acts on $P/\Phi(P)$ as the wedge-square of the natural module. If M_0 contains P, then $M_0 = P.X$ where X is maximal in $\mathrm{SL}_5(2)$; by inspecting [7, Tables 8.18, 8.19] we see that X is either a parabolic subgroup or 31:5, and so has at most 3 composition factors on $P/\Phi(P)$. In particular, we deduce that $d(M_0) \leqslant 3 + d(X) \leqslant 7$ in this case. And if $P \nleqslant M_0$ then $M_0 = \Phi(P).\mathrm{SL}_5(2)$ and hence $d(M_0) \leqslant 3$.

Next we handle case (iii). Here $H_0 = N_{G_0}(T)$, where T is a maximal torus of G_0, as listed in [22, Table 5.2]. The groups $W = N_{G_0}(T)/T$ are also listed in Table 5.2 of [22]; these are subgroups of the Weyl group of G_0.

Suppose first that $T \leqslant M_0$, so that $M = T.X$ with X maximal in $N_G(T)/T$ (which is $W \times \langle \phi \rangle$, possibly extended by a graph automorphism, where ϕ is a field automorphism). If $T \neq (q \pm 1)^r$ with $r \in \{7, 8\}$, it is clear from the list that $d(T) \leqslant 6$, and one checks that $d(X) \leqslant 6$ also, giving $d(M) \leqslant 12$. And if $T = (q \pm 1)^r$ then $W = W(E_r)$ and one checks that $d(X) \leqslant 4$ for a maximal subgroup in this case, giving $d(M) \leqslant r + 4 \leqslant 12$.

Now suppose $T \nleqslant M_0$. Then $M_0 = (M \cap T).W$. A check gives $d(W) \leqslant 2$, hence $d(M_0) \leqslant d(M \cap T) + 2 \leqslant 10$.

It remains to handle case (ii), in which $H_0 = N_{G_0}(K)$, where K is a reductive subgroup of G_0 of maximal rank, not a maximal torus. The possibilities for K and H_0/K are listed in [22, Table 5.1]. In all cases K is a central product $\prod L_i \circ R$, where each L_i is either quasisimple or in $\{\mathrm{SL}_2(2), \mathrm{SL}_2(3), \mathrm{SU}_3(2)\}$, and R is an abelian p'-group of rank at most 2 (also $R = 1$ unless G_0 is of type E_7, E_6^ϵ or ${}^3D_4$).

The cases where K is solvable are those in Table 3. We exclude these cases from consideration until the end of the proof.

G_0	K	q
$E_8(q)$	$A_1(q)^8$	$2, 3$
$E_8(q)$	$A_2^-(q)^4$	2
$E_7(q)$	$A_1(q)^7$	$2, 3$
${}^2E_6(q)$	$A_2^-(q)^3$	2
$F_4(q)$	$A_2^-(q)^2$	2

Table 3. Cases with K solvable

Let $N = \mathrm{core}_H(M)$. By Lemma 2.3 we may assume that $N \neq 1$. Assume first that $K \leqslant N$. Then $M = K.X$ where X is maximal in H/K. Inspecting the list of possibilities for K and H/K, it is easy to check that $d(K) \leqslant 4$ and $d(X) \leqslant 8$, giving the conclusion.

Next assume that $N \leqslant Z(K)$. Then $H = MK$ so $d_M(N) = d_H(N)$. Inspection of the list shows that $d(N) \leqslant 2$ except for the cases $K = A_1(q)^r$ $(r = 7, 8)$, and in these cases $Z(K) = 2^{r-4}$ and $d_M(N) \leqslant 2$. Hence by Remark 2.4 we have $d(M) \leqslant d_M(N) + 10 \leqslant 12$, as required.

Now assume $K \nleqslant N$ and $N \nleqslant Z(K)$. Then N contains a product N_0 of factors L_i of K. In all but two cases in the list where K has at least two isomorphic factors L_i, H_0/K acts transitively on these factors; the two exceptional cases are $K = A_2^\epsilon(q)^2$ in $F_4(q)$ and $K = A_1(q)^2$ in $G_2(q)$. Hence inspecting the list, we see that K is in Table 4, with N_0 equal to one of the factors (or $A_1(q)^3$).

G_0	K
$E_8(q)$	$A_1(q)E_7(q)$, $A_2^{\epsilon}(q)E_6^{\epsilon}(q)$
$E_7(q)$	$A_1(q)D_6(q)$, $A_2^{\epsilon}(q)A_5^{\epsilon}(q)$, ${}^3D_4(q)A_1(q^3)$, $D_4(q)A_1(q)^3$, $E_6^{\epsilon}(q) \circ (q-\epsilon)$
$E_6^{\epsilon}(q)$	$A_1(q)A_5^{\epsilon}(q)$, $A_2(q^2)A_2^{-\epsilon}(q)$, ${}^3D_4(q) \times (q^2+\epsilon q+1)$, $D_5^{\epsilon}(q) \circ (q-\epsilon)$, $D_4(q) \circ (q-\epsilon)^2$
$F_4(q)$	$A_1(q)C_3(q)$, $A_2^{\epsilon}(q)^2$
$G_2(q)$	$A_1(q)^2$
${}^3D_4(q)$	$A_1(q)A_1(q^3)$, $A_2^{\epsilon}(q) \circ (q^2+\epsilon q+1)$

Table 4. Cases with $K \not\leqslant N$ and $N \not\leqslant Z(K)$

Write $K = N_0K_0$, where K_0 is the product of the factors L_i (or R) not in N_0. Then $M \cap K = N_0M_0$, where M_0 is a maximal H-invariant subgroup of K_0. From the above table, K_0 is either a single factor L_i or R of K, or it is $A_1(q)^3$. In the former case, using Theorem 2.1 we see that $d(M_0) \leqslant 4$, whence $d(M) \leqslant d(N_0)+d(M_0)+d(H/K) \leqslant 12$. The other possibility is that $K_0 = A_1(q)^3$, $N_0 = D_4(q)$. If $q > 3$ then M_0 must be a diagonal subgroup of K_0, so $d(M_0) \leqslant 2$; and if $q \leqslant 3$ then $H/N_0 \cong A_1(q)^3.d^3.S_3$ where $d = (2, q-1)$, and we easily check that $d(M/N_0) \leqslant 10$, so that $d(M) \leqslant d(N_0) + 10 \leqslant 12$.

It remains to handle the cases where K is solvable, given in Table 3. The most complicated example is $K = A_1(q)^8$ in $E_8(q)$ with $q = 3$. We deal with this case and leave the others to the reader. In this case $Z(K) = 2^4$, $H/K \cong 2^4.\mathrm{AGL}_3(2)$, so

$$H = 2^4.2^{16}.3^8.2^4.2^3.\mathrm{L}_3(2).$$

Let R denote the solvable radical of H. If $R \leqslant M$ then $M = R.X$ where X is maximal in $\mathrm{L}_3(2)$; since $d(X) = 2$ and $d_M(A_1(3)^8) = 2$, it follows that $d(M) \leqslant 2 + d(2^4.2^3) + d(X) < 12$. And if $R \not\leqslant M$ then $M/M \cap R \cong \mathrm{L}_3(2)$ and it follows that $d_M(M \cap R) \leqslant 10$, whence $d(M) \leqslant 10 + d(\mathrm{L}_3(2)) = 12$. □

7. Parabolic subgroups and Number Theory

In this section we complete the proof of Theorems 1 and 3 by handling second maximal subgroups M lying in parabolic subgroups. In particular we relate the boundedness of $d(M)$ to the number-theoretic question (1) stated in the Introduction.

LEMMA 7.1. *Let $q = p^k$, where p is a prime and $k \geqslant 1$, let e be a divisor of $q-1$ and let E be the subgroup of order e of the multiplicative group $\mathbb{F}_q^{\times}$. Let $M = \mathbb{F}_q.E$*

be the corresponding subgroup of the semidirect product $\mathbb{F}_q.\mathbb{F}_q^\times \cong \mathrm{AGL}_1(q)$. *Then*

$$k/\ell \leqslant d(M) \leqslant k/\ell + 1,$$

where $\ell = \min\{i : e \text{ divides } p^i - 1\}$ *is the multiplicative order of p modulo e.*

Proof. Let K be the minimal subfield of $\mathbb{F}_q$ containing E. Then K has order p^ℓ where ℓ divides k. Therefore $\mathbb{F}_q$ has dimension k/ℓ as a vector space over K. Thus M is generated by a basis of that vector space together with a generator of the cyclic group E, so $d(M) \leqslant k/\ell + 1$.

To prove the other inequality, suppose (a_i, b_i) are generators for M, where $a_i \in \mathbb{F}_q$, $b_i \in E$ and $i = 1, \dots, d$. Then $a_1, \dots, a_d$ generate $\mathbb{F}_q$ as a vector space over K, so $d \geqslant k/\ell$, as required. □

The next result helps in establishing a connection between bounding the number of generators of second maximal subgroups and the answer to the number-theoretic question (1) stated in Section 1.

Lemma 7.2. *Let* $G = \mathrm{L}_2(q)$, ${}^2B_2(q)$ *or* ${}^2G_2(q)$ *where* $q = p^k$ *(p prime), let* $d = (2, q-1)$, 1 *or* 1 *respectively, and let* $B = UT$ *be a Borel subgroup of G with unipotent normal subgroup U and Cartan subgroup T of index d in* $\mathbb{F}_q^\times$. *Let s be a prime divisor of* $q - 1$ *and let* $e = \frac{q-1}{ds}$, *so that B has a maximal subgroup* $M = U.e$ *of index s. Let* ℓ *be the multiplicative order of p modulo e. Then*

(i) *we have* $k/\ell \leqslant d(M) \leqslant k/\ell + 1$*;*

(ii) $d(M)$ *is unbounded if and only if* $\ell = o(k)$*;*

(iii) *either* $k \in \{\ell, 2\ell\}$ *(in which case* $d(M) \leqslant 3$*), or* $\frac{p^k-1}{p^\ell-1} = s$ *is prime.*

Proof. We first prove part (i). If $G = \mathrm{L}_2(q)$, then $U \cong \mathbb{F}_q$ and so $k/\ell \leqslant d(M) \leqslant k/\ell + 1$ by the previous lemma. The other families ${}^2B_2(q)$ and ${}^2G_2(q)$ are handled by the same argument, noting that $U/\Phi(U) \cong \mathbb{F}_q$ with T acting by scalar multiplication (see [33, 35]).

Part (ii) follows immediately from part (i). To prove (iii), note that $ds = \frac{p^k-1}{p^\ell-1} \cdot \frac{p^\ell-1}{e}$ and s is a prime. If $\frac{p^k-1}{p^\ell-1} \neq 1, s$, then $d = 2$ and $2s = \frac{p^k-1}{p^\ell-1}$. This implies that $\frac{k}{\ell}$ is even, say $\frac{k}{\ell} = 2m$. Then writing $q_0 = p^\ell$, we have $2s = \frac{(q_0^m-1)(q_0^m+1)}{q_0-1}$, which forces $m = 1$, hence $k = 2\ell$. This proves (iii). □

Lemma 7.3. *Let G be an almost simple group of Lie type with socle* G_0. *Suppose G has a maximal subgroup which is a Borel subgroup B, and suppose B has a maximal subgroup M with* $d(M) > 12$. *Then* $G_0 = \mathrm{L}_2(q)$, ${}^2B_2(q)$ *or* ${}^2G_2(q)$, *and* $M \cap G_0$ *is as in Lemma 7.2.*

Proof. These are the cases where G_0 has BN-rank 1, or is $\mathrm{L}_3(q)$, $C_2(2^e)$ or $G_2(3^e)$ and G contains a graph or graph-field automorphism. We need to rule out the latter three cases, and also the case where $G_0 = \mathrm{U}_3(q)$. As before, set $M_0 = M \cap G_0$. Note that if $G_0 = \mathrm{L}_2(q)$, ${}^2B_2(q)$ or ${}^2G_2(q)$, then Lemma 2.3 shows that $M \cap G_0$ is as in Lemma 7.2.

Consider first $G_0 = \mathrm{U}_3(q)$, so that $B \cap G_0 = QT$ where $Q = q^{1+2}$ is a special group with $Q' = \Phi(Q) \cong \mathbb{F}_q$, $Q/Q' \cong \mathbb{F}_q^2$ and $T \cong Z_{(q^2-1)/d}$ with $d = (3, q+1)$. If $Q \leqslant M$ then $M = QS$ where S contains either Z_{q-1} or $Z_{(q+1)/d}$ (note that $d(S) \leqslant 2$). Using Lemma 7.1 we see that $d_{M/Q'}(Q/Q') \leqslant 4$ and it follows that $d(M) \leqslant 4 + d(S) + 1 \leqslant 7$, a contradiction. And if $Q \nleqslant M$ then $M_0 = (M \cap Q).T$ and $M \cap Q$ is a maximal T-invariant subgroup of Q; it follows that $d_{M_0}(M \cap Q) \leqslant 2$, so $d(M_0) \leqslant 2 + d(T) = 3$ and thus $d(M) \leqslant 5$, a contradiction.

Next consider $G_0 = C_2(q)$ where $q = 2^e$ and G contains an element inducing a graph-field automorphism on G_0. Adopting the notation of [10], let $B \cap G_0 = QT$ where Q is generated by the positive root groups relative to a fixed root system (so $|Q| = q^4$), and $T = \langle h_\alpha(t), h_\beta(u) \,:\, t, u \in \mathbb{F}_q^\times \rangle$, where α, β are fundamental roots with α long and β short. By assumption, $G = G_0\langle \tau \rangle$, where τ is a graph-field automorphism of G_0 normalizing Q and T, sending

$$h_\alpha(t) \mapsto h_\beta(t^r), \;\; h_\beta(u) \mapsto h_\alpha(u^{2r}),$$

where $r = 2^f$ for some $f \leqslant e$. Let $\pi_1, \pi_2 : T \to \mathbb{F}_q^\times$ be the maps sending $h_\alpha(t), h_\beta(u)$ to t, u respectively.

Assume first that $Q \leqslant M$, so $M_0 = QT_0$ and T_0 is a maximal τ-invariant subgroup of T.

If $\pi_1(T_0) = \mathbb{F}_q^\times$ then $\pi_2(T_0) = \mathbb{F}_q^\times$ also (as T_0 is τ-invariant), and so T_0 acts as the full group of scalars on each factor of a series $1 = Q_0 < Q_1 < \cdots < Q_4 = Q$ with $Q_i/Q_{i-1} \cong \mathbb{F}_q$ for all i; hence $d_{M_0}(Q) \leqslant 4$ and it follows that $d(M) \leqslant 4 + d(T_0) + 1 \leqslant 7$, a contradiction.

Now assume $\pi_1(T_0) = A < \mathbb{F}_q^\times$. As T_0 is τ-invariant, $\pi_2(T_0) = A$ as well, and so by maximality $T_0 = \{h_\alpha(t)h_\beta(u) \,:\, t, u \in A\}$. If e is even (recall that $q = 2^e$) then (again by maximality) $|A|$ is divisible by $q^{1/2} - \epsilon$ for some $\epsilon = \pm 1$, and now the result follows as in the previous paragraph, using Lemma 7.1. On the other hand, if e is odd, then the automorphism $t \mapsto t^2$ of $\mathbb{F}_q$ has odd order, so there is an automorphism ϕ of $\mathbb{F}_q$ such that $\phi^2(t) = t^2$ for all $t \in \mathbb{F}_q$. But then

$$\langle \{h_\alpha(t)h_\beta(u),\; h_\alpha(\phi(v))h_\beta(v) \,:\, t, u \in A,\; v \in \mathbb{F}_q^\times\} \rangle$$

is a proper τ-invariant subgroup of T, contradicting the maximality of T_0.

Finally for this case ($G_0 = C_2(q)$), if $Q \nleqslant M$ then $M_0 = (M \cap Q).T$ and $M \cap Q$ is a maximal T-invariant subgroup of Q; it follows that $d_{M_0}(M \cap Q) \leqslant 3$ and so $d(M) \leqslant 3 + d(T) + 1 \leqslant 6$, a contradiction.

The case where $G_0 = G_2(3^e)$ and G contains a graph or graph-field automorphism is handled in very similar fashion. The case $G_0 = \mathrm{L}_3(q)$ is also similar, but this time τ sends $h_\alpha(t) \mapsto h_\beta(t^r)$, $h_\beta(u) \mapsto h_\alpha(u^r)$ for all $t, u \in \mathbb{F}_q^\times$, and in the case of the above argument where $M_0 = QT_0$, we must have $\pi_i(T_0) = \mathbb{F}_q^\times$ for $i = 1, 2$, giving $d_{M_0}(Q) \leqslant 3$. □

Proposition 7.4. *Theorem* 1 *holds in the case where $M < H < G$ with G an almost simple group of Lie type and H a maximal parabolic subgroup of G.*

Proof. Let G_0 denote the socle of G, which is a simple group of Lie type over $\mathbb{F}_q$, a field of characteristic p.

Let $M_0 = M \cap G_0$, and write $H_0 = H \cap G_0 = P = QR$, a parabolic subgroup with unipotent radical Q and Levi subgroup R. We use the notation $P = P_{ij\ldots}$ to mean a parabolic with excluded nodes $i, j, \ldots$ from the Dynkin diagram.

By Lemma 7.3, we may assume that H_0 is not a Borel subgroup. In particular, G_0 is not of type ${}^2G_2$ or ${}^2B_2$. We also exclude for now the cases where (G_0, p) is *special* in the sense of [4] – that is to say, $p = 2$ and G_0 is of type C_n, F_4, ${}^2F_4$, G_2, or $p = 3$ and G_0 is of type G_2. We shall deal with these excluded cases at the end of the proof.

Suppose first that G_0 is untwisted and $H_0 = P_i$ for some i. Then by [4, Theorem 2(a)], Q/Q' has the structure of an irreducible $\mathbb{F}_qR$-module, and $Q' \leqslant \Phi(Q)$, so $Q' \leqslant M_0$. It follows that either $M_0 = QK$ with K a maximal H/Q-invariant subgroup of R, or $M_0 = Q'.R$.

Consider the case where $M_0 = QK$. Now $R = R_0Z$, where Z is a central torus of rank 1 inducing scalars on the module Q/Q'. Hence either $K = R_0Z_0$ with $Z_0 < Z$, or $K = K_0Z$ with $K_0 < R_0$. For G_0 classical, R_0 is of type $\mathrm{SL}_i(q) \times \mathrm{SL}_{n-i}(q)$ or $\mathrm{SL}_i(q) \times Cl_{n-2i}(q)$ and Q/Q' is the corresponding tensor product space $U \otimes W$ with $\dim U = i$, $\dim W = n - i$ or $n - 2i$ (here $Cl_{n-2i}(q)$ denotes an appropriate classical group of dimension $n - 2i$ over $\mathbb{F}_q$). Then using Lemma 2.5 we see that Q/Q' is a cyclic K-module. Using Theorem 2.1 we deduce that $d(K) \leqslant 6$. Hence $d(M_0) \leqslant 1 + d(K) \leqslant 7$. For G_0 of exceptional type, the irreducible module Q/Q' has dimension at most 64 with equality for $(G_0, R_0) = (E_8(q), D_7(q))$, so we get $d(M_0) \leqslant \dim(Q/Q') + d(K) \leqslant 70$.

Now suppose $M_0 = Q'.R$. Here we bound $d(M_0)$ by $d_{M_0}(Q') + d(R)$. Now R_0 is a commuting product of at most 3 factors which are either quasisimple or groups in $\{\mathrm{SL}_2(q), \Omega_3(q), \Omega_4^+(q) : q \leqslant 3\}$; hence it is straightforward to check that $d(R) \leqslant 4$. Also $d_{M_0}(Q')$ is at most the number of R-composition factors in Q'. By [4, Theorem 2], this is 1 less than the i-th coefficient of the highest root in the root system of G_0, hence is at most 1 for G_0 classical, and at most 5 for G_0 exceptional. We conclude that $d(M_0) \leqslant 9$ in this case.

Next assume that G_0 is twisted (and not special) – hence of type ${}^2A_n$, ${}^2D_n$, ${}^2E_6$ or ${}^3D_4$. In the first case consider the covering group $\hat{G}_0 = \mathrm{SU}_m(q)$ (where $m = n+1$), where $H_0 = P_i = QR$ with R of type $\mathrm{SL}_i(q^2) \times \mathrm{SU}_{m-2i}(q)$. Here Q/Q' has the structure of the R-module $V_1 + V_2$ with $V_1 = U \otimes W$ and $V_2 = U^{(q)} \otimes W^*$, where U, W are the natural modules for the factors of R. Hence as above, the possibilities for M_0 are QK, $Q_1.R$ and $Q_2.R$, where $Q_i = Q'.V_i < Q$. We deal with the possibilities just as before. The ${}^2D_n$ or ${}^2E_6$ cases are very similar – again, Q/Q' is a sum of at most two irreducible R-submodules, leading to three possibilities for M_0 as above. Finally, if $G_0 = {}^3D_4(q)$ then $H_0 = P_i$ with $i = 1$ or 2. If $i = 2$ then $R_0 = A_1(q^3)$ and Q/Q' is the irreducible $\mathbb{F}_qR$-module $V_2 \otimes V_2^{(q)} \otimes V_2^{(q^2)}$; and if $i = 1$ then R contains $A_1(q) \circ (q^3 - 1)$ and again Q/Q' is an irreducible $\mathbb{F}_qR$-module (of dimension 6). In either case the result follows in the usual way.

The case where G_0 is of type A_n, D_n, D_4 or E_6 and G contains a graph automorphism is very similar. In these cases, the maximal parabolics of G for which Q/Q' is a reducible R-module are $P_{i,n-i}$ (for A_n), P_{n-1} (for D_n), P_{134} (for D_4 when G contains a triality automorphism) and P_{16}, P_{35} (for E_6). For these, [4] shows that Q/Q' is a sum of two irreducible R-modules (three for the D_4 case), and we argue as in the previous paragraph.

It remains to handle the cases where G_0 is special. These are dealt with by the same method as above. By the proof of [8, Lemma 7.3], Q/Q' has at most 4 $\mathbb{F}_qR$-module composition factors, so we can compute the possibilities for M_0 and bound $d(M_0)$ just as before. □

By combining this result with Propositions 3.1, 4.1, 5.1 and 6.1, we conclude that the proof of Theorem 1 is complete.

Remark 7.5. The upper bound of 70 in part (ii) of Theorem 1 is not sharp, and we make some remarks here about how one could go about improving it. As observed in the proof of Proposition 7.4, we have this upper bound of 70 because of second maximal subgroups $M < QR = P_1$, a D_7-parabolic subgroup of $E_8(q)$, of the form $M = QK_0Z$ where K_0 is a maximal subgroup of $D_7(q)$. To improve the bound significantly, one would have to study the actions of such subgroups K_0 on Q/Q', which is a 64-dimensional spin module for $D_7(q)$. Likewise, the E_7-parabolic P_8 of $E_8(q)$ has maximal subgroups $M = QK_0Z$ with K_0 a maximal subgroup in $E_7(q)$ (not all of which are known); consequently, in order to improve the obvious upper bound $d(M) \leqslant \dim(Q/Q') + d(K_0) \leqslant 60$ in this case, one would have to study the actions of such K_0 on the 56-dimensional $E_7(q)$-module Q/Q'.

We are also in a position to give a proof of Theorem 3.

Proof of Theorem 3. Clearly, part (i) of Theorem 3 implies (ii), and (ii) implies (iii). For the next implication, note that the question (1) stated in Section 1 has a negative answer if and only if there exists a constant c such that if p is a prime and $(p^k - 1)/(p^\ell - 1)$ is prime for some natural numbers k, ℓ, then $k \leqslant c\ell$. Hence the fact that (iii) implies (iv) follows from Lemma 7.2.

Finally, we show that (iv) implies (i). Assume (iv) holds, and let G be an almost simple group with socle G_0. Let M be second maximal in G. By Theorem 1, we have $d(M) \leqslant 70$ except possibly if $G_0 = \mathrm{L}_2(q)$, ${}^2B_2(q)$ or ${}^2G_2(q)$, and M is maximal in a Borel subgroup B of G. In the latter cases, $B \cap G_0 = UT$ as in Lemma 7.2. If $U \nleqslant M$ then $d(M) \leqslant 10$ by Lemma 2.3; and if $U \leqslant M$, then $d(M)$ is bounded by Lemma 7.2 together with the assumption (iv). Hence (iv) implies (i) and the proof of Theorem 3 is complete. □

8. Random generation and third maximal subgroups

In this final section we prove Proposition 4 and Theorems 5 and 6.

Proof of Proposition 4. Let $p \geqslant 5$ be a prime such that $p \equiv \pm 3 \pmod 8$. The group $\mathrm{PGL}_2(p)$ has a maximal subgroup S_4 (cf. [14]), and S_{p+1} has a maximal subgroup $\mathrm{PGL}_2(p)$ (by [21]). Moreover, for $n = 2(p+1)$, the imprimitive subgroup $S_2 \wr S_{p+1}$ is maximal in S_n (again by [21]). Hence we have the following chain of subgroups of S_n, each maximal in the previous one:

$$S_n > S_2 \wr S_{p+1} > S_2 \wr \mathrm{PGL}_2(p) > (S_2)^{p+1}.S_4.$$

Write $M = (S_2)^{p+1}.S_4$, and let B be the base group $(S_2)^{p+1}$. By the Schreier index formula, $d(B) - 1 \leqslant |M : B|\,(d(M) - 1)$, and hence

$$d(M) > \frac{d(B) - 1}{24} = \frac{p}{24}.$$

Since M is third maximal in S_n and p can be arbitrarily large, this completes the proof of the proposition. □

For the proof of Theorem 5, we need the following result on chief factors of second maximal subgroups. For a finite group G, we define $\gamma(G)$ to be the number of non-abelian chief factors of G.

Proposition 8.1. *If M is a second maximal subgroup of an almost simple group, then $\gamma(M) \leqslant 5$.*

Proof. Let G be an almost simple group with socle G_0 and write $M < H < G$ with M maximal in H and H maximal in G. Note that if N is a normal subgroup of M, then $\gamma(M) \leqslant \gamma(N) + \gamma(M/N)$. In particular, if N is solvable then $\gamma(M) = \gamma(M/N)$. By [8, Lemma 8.2], we have $\gamma(M) \leqslant 3$ if H is almost simple, so we may assume otherwise. More generally, if H is of the form $H = N.A$, where N is solvable and A is almost simple, then either $M = (M \cap N).A$ and $\gamma(M) = 1$, or $M = N.J$ and $J < A$ is maximal, so $\gamma(M) \leqslant 3$. Similarly, if $H = N.(A \times B)$ with N solvable and A and B almost simple, then either $\gamma(M) = 2$ or $M = N.J$ with $J < A \times B$ maximal and it is easy to check that $\gamma(M) = \gamma(J) \leqslant 4$.

If G_0 is sporadic then all the maximal subgroups of G are known (apart from a handful of small almost simple candidates in the Monster) and it is straightforward to verify the bound $\gamma(M) \leqslant 4$ by direct inspection. Next suppose $G_0 = A_n$ is an alternating group. As noted in the proof of Proposition 3.1, the possibilities for H are determined by the O'Nan-Scott theorem and once again it is easy to check that $\gamma(M) \leqslant 4$. This bound is sharp. For example, if $G = S_n$ and $H = S_k \wr S_t$, where $k \geqslant 5$ and $t \geqslant 11$, then $M = (S_k)^t.(S_5 \times S_{t-5})$ is a maximal subgroup of H with $\gamma(M) = 4$.

Next assume G_0 is a classical group. Here we use [19] to inspect the possibilities for H (recall that we may assume H is not almost simple) and one checks that $\gamma(M) \leqslant 4$ if H is non-parabolic. In fact, the same bound holds in all cases, with the possible exception of the case where $G_0 = \mathrm{L}_n(q)$ and H is a parabolic subgroup of type $P_{m,n-m}$ as described in [19, Proposition 4.1.22]. In the latter case, we could have $\gamma(M) = \gamma(J) + 2$ where $J = K \cap \mathrm{L}_a(q)$ for some maximal subgroup K of an almost simple group with socle $\mathrm{L}_a(q)$ (here $a = m$ or $n - m$). Therefore, $\gamma(M) \leqslant 5$. Similar reasoning applies when G_0 is an exceptional group. A convenient description of the maximal subgroups of G is given in [24, Theorem 8] and it is straightforward to show that $\gamma(M) \leqslant 5$. □

We now derive consequences concerning the invariant $\nu(M)$ defined in Section 1. Our main tool is Theorem 1 of Jaikin-Zapirain and Pyber [16].

Corollary 8.2. *There exists an absolute constant c such that if M is a second maximal subgroup of an almost simple group, then $\nu(M) \leqslant c\,d(M)$. Consequently $\nu(M)$ is bounded if and only if $d(M)$ is bounded.*

Proof. Let β be the constant in [16, Theorem 1]. By combining Proposition 8.1 with this theorem, we obtain

$$\nu(M) < \beta d(M) + \frac{\log(\gamma(M))}{\log 5} \leqslant \beta d(M) + 1.$$

The result follows. □

Proof of Theorem 5. Let G be an almost simple group with socle G_0 and let M be a second maximal subgroup of G which is not as in part (iii) of Theorem 1. Then $d(M) \leqslant 70$ by Theorem 1, and the result follows from Corollary 8.2. □

For the proof of Theorem 6 we need the following result, which may be of some independent interest.

Lemma 8.3. *Let R be a finite-dimensional algebra over a finite field $\mathbb{F}$. Let M be an R-module of finite dimension over $\mathbb{F}$. Then M has at most $|M/JM|-1$ maximal submodules, where J is the Jacobson radical of R. Moreover, this upper bound is best possible.*

Proof. It is well known that every maximal submodule of M contains JM. Therefore the number of maximal submodules of M equals the number of maximal submodules of M/JM (as an R/J-module). This enables us to reduce to the case where $J = 0$, so that R is a semisimple algebra and M is a semisimple R-module.

Hence we may write

$$M = \bigoplus_{i=1}^{m} n_i S_i,$$

where the S_i $(1 \leqslant i \leqslant m)$ are pairwise non-isomorphic simple R-modules, and $n_i \geqslant 1$ is the multiplicity of S_i.

Let $M_0 < M$ be a maximal submodule. Then $M/M_0 \cong S_i$ for some unique i with $1 \leqslant i \leqslant m$. It follows that $M_0 \supseteq M_i$ where $M_i = \bigoplus_{j \neq i} n_j S_j$. Hence M_0/M_i may be regarded as a maximal submodule of $n_i S_i$.

The number of such maximal submodules is less than $|\mathrm{Hom}(n_i S_i, S_i)| = |\mathrm{End}(S_i)|^{n_i}$. Since S_i (being simple) is a cyclic module we have $|\mathrm{End}(S_i)| \leqslant |S_i|$. It follows that M has less than $|S_i|^{n_i}$ maximal submodules M_0 satisfying $M/M_0 \cong S_i$. Summing over i we see that the number of maximal submodules of M is less than

$$\sum_{i=1}^{m} |S_i|^{n_i} \leqslant \prod_{i=1}^{m} |S_i|^{n_i} = |M|.$$

This completes the proof of the upper bound.

To show that this upper bound is best possible, let $R = \mathbb{F} = \mathbb{F}_2$ and let M be a d-dimensional vector space over $\mathbb{F}$. Then $|M/JM| = 2^d$ and M has $|M/JM| - 1$ maximal submodules. □

Proof of Theorem 6. Let G be an almost simple group with socle G_0. By [8, Corollary 6], G has at most n^a second maximal subgroups of index n for some absolute constant a and for all $n \geqslant 1$. It therefore suffices to show the following.

Claim. *There is an absolute constant b such that, for every $n \geqslant 1$, every second maximal subgroup M of G has at most n^b maximal subgroups of index n in G.*

Indeed, assuming the claim, a third maximal subgroup N of index n in G is contained in some second maximal subgroup M of G, which – being of index at most n – can be chosen in at most n^{a+1} ways. Given M, the third maximal subgroup N can be chosen in at most n^b ways. Thus G has at most n^{a+b+1} third maximal subgroups of index n.

To prove the claim, let M be a second maximal subgroup of G. Recall that $m_n(M)$ denotes the number of maximal subgroups of M of index n in M. If $m_n(M) \leqslant n^b$ for an absolute constant b and for all n then the claim follows immediately.

We show that this is the case assuming G_0 is not $\mathrm{L}_2(q)$, ${}^2B_2(q)$ or ${}^2G_2(q)$. Indeed, in this case we have $\nu(M) \leqslant c$ by Theorem 5, so by [26, Proposition 1.2] we have $m_n(M) \leqslant n^b$ where $b = c + 3.5$.

Now assume that $G_0 = \mathrm{L}_2(q)$, ${}^2B_2(q)$ or ${}^2G_2(q)$. We apply Lemma 7.3 which describes the second maximal subgroups M of G for which $d(M)$ is possibly unbounded. By Corollary 8.2 these are the ones for which $\nu(M)$ is possibly unbounded.

Suppose $G_0 = \mathrm{L}_2(q)$ with $q = p^k$, and let $G = G_1.A$ where $G_1 = G \cap \mathrm{PGL}_2(q)$ and A is a group of field automorphisms of order dividing k. Set $f = |A|$. The relevant second maximal subgroups M are of the form $U.T_1.A$, where $U \cong \mathbb{F}_q$ and $T_1 \leqslant \mathbb{F}_q^\times$ has order e. Let ℓ be the multiplicative order of p modulo e as in Lemma 7.2. Note that $|G : M| > q$. We shall show that such subgroups M have less than n^4 maximal subgroups of index n in G.

The maximal subgroups of such a group M split naturally into two types. The first type is $U.X$ where X is maximal in $T_1.A$. Now, $T_1.A$ is metacyclic, and so are its subgroups. Since all subgroups of $T_1.A$ are 2-generated, there are at most $|T_1.A|^2 = e^2f^2 < q^2k^2 < q^4$ such subgroups (including non-maximal ones). This proves the claim with $b = 4$ for subgroups of M of the first type.

The second type of subgroups of M is $U_0.T_1.A$, where U_0 is a proper $\mathbb{F}_{p^\ell}$-subspace of $U \cong \mathbb{F}_q$ that is maximal A-invariant. Let $q_0 = p^\ell$ and consider the group algebra $R = \mathbb{F}_{q_0}[A]$. Then U, U_0 are R-modules and U_0 is a maximal submodule of U.

Applying Lemma 8.3, there are fewer than $|U| = q$ possibilities for U_0. We now claim that, given U_0, there are less than q^3 subgroups of M of type $U_0.T_1.A$. Indeed, T_1 is split in $U_0.T_1$, so there are less than q possibilities for $U_0.T_1$; and the cyclic group A is generated by an element of the form $u\phi$ where $u \in U_0.T_1$ and ϕ is a fixed field automorphism, so there are less than q^2 possibilities for such a generator, hence less than q^3 possibilities in all for $U_0.T_1.A$.

We conclude that the number of maximal subgroups of M of the second type is also less than q^4. Since $|G : M| > q$ this completes the proof of the claim for $G_0 = \mathrm{L}_2(q)$, with $b = 4$.

The proofs for Suzuki and Ree groups are similar, and this completes the proof of the claim, and hence of the theorem. □

Acknowledgements

The authors are grateful for the hospitality of the Centre Interfacultaire Bernoulli at EPFL, where this work was completed. The second and third authors acknowledge the support of EPSRC Mathematics Platform grant EP/I019111/1. The third author also acknowledges the support of Advanced ERC Grant 247034, an ISF grant 1117/13, and the Miriam and Julius Vinik Chair in Mathematics which he holds.

References

[1] M. Aschbacher, *On intervals in subgroup lattices of finite groups*, J. Amer. Math. Soc. **21** (2008), 809–830.

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

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

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

[5] A. Basile, *Second maximal subgroups of the finite alternating and symmetric groups*, PhD thesis (Australian National University, 2001), arxiv:0810.3721.

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

[7] J.N. Bray, D.F. Holt and C.M. Roney-Dougal, *The Maximal Subgroups of the Low-dimensional Finite Classical Groups*, London Math. Soc. Lecture Note Series, vol. 407, Cambridge University Press, 2013.

[8] T.C. Burness, M.W. Liebeck and A. Shalev, *Generation and random generation: From simple groups to maximal subgroups*, Advances in Math. **248** (2013), 59–95.

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

[10] R.W. Carter, *Simple groups of Lie type*, John Wiley and Sons, 1972.

[11] 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.

[12] J.H. Conway, R.T. Curtis, S.P. Norton, R.A. Parker, and R.A. Wilson, *Atlas of Finite Groups*, Oxford University Press, 1985.

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

[14] L.E. Dickson, *Linear groups with an exposition of the Galois field theory*, Teubner, Leipzig 1901 (Dover reprint 1958).

[15] W. Feit, *An interval in the subgroup lattice of a finite group which is isomorphic to M_7*, Algebra Universalis **17** (1983), 220–221.

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

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

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] M.W. Liebeck and G.M. Seitz, *Maximal subgroups of exceptional groups of Lie type, finite and algebraic*, Geom. Dedicata **36** (1990), 353–387.

[24] M.W. Liebeck and G.M. Seitz, *A survey of of maximal subgroups of exceptional groups of Lie type*, in Groups, combinatorics & geometry (Durham, 2001), 139–146, World Sci. Publ., 2003.

[25] M.W. Liebeck and A. Shalev, *The probability of generating a finite simple group*, Geom. Dedicata **56** (1995), 103–113.

[26] A. Lubotzky, *The expected number of random elements to generate a finite group*, J. Algebra **257** (2002), 452–459.

[27] A. Lucchini and F. Menegazzo, *Generators for finite groups with a unique minimal normal subgroup*, Rend. Sem. Mat. Univ. Padova **98** (1997), 173–191.

[28] A. Mann, *Positively finitely generated groups*, Forum Math. **8** (1996), 429–459.

[29] A. Mann and A. Shalev, *Simple groups, maximal subgroups, and probabilistic aspects of profinite groups*, Israel J. Math. **96** (1996), 449–468.

[30] I. Pak, *On probability of generating a finite group*, preprint (1999).

[31] P.P. Pálfy, *On Feit's examples of intervals in subgroup lattices*, J. Algebra **116** (1988), 471–479.

[32] R. Steinberg, *Generators for simple groups*, Canad. J. Math. **14** (1962), 277–283.

[33] M. Suzuki, *On a class of doubly transitive groups*, Annals of Math. **75** (1962), 105–145.

[34] J. Thévenaz, *Maximal subgroups of direct products*, J. Algebra **198** (1997), 352–361.

[35] H.N. Ward, *On Ree's series of simple groups*, Trans. Amer. Math. Soc. **121** (1966), 62–89.

[36] R.A. Wilson et al., *A World-Wide-Web Atlas of finite group representations*, `http://brauer.maths.qmul.ac.uk/Atlas/v3/`.