J. Eur. Math. Soc. 28, 4063–4103 (2026) DOI 10.4171/JEMS/1547

Henry Bradford · Andreas Thom

Short laws for finite groups of Lie type

Received October 4, 2022

Abstract. We produce new short laws in two variables valid in finite groups of Lie type. Our result improves upon results of Kozma and the second named author, and is sharp up to logarithmic factors, for all families except possibly the Suzuki groups. We also produce short laws valid for generating pairs and random pairs in finite groups of Lie type, and, conditional on Babai's diameter conjecture, make effective the dependence of our bounds on the rank. Our proof uses, among other tools, the classification of finite simple groups, Aschbacher's structure theorem for maximal subgroups for classical groups, and upper bounds on the diameters of finite simple groups due to Breuillard, Green, Guralnick, Pyber, Szabó and Tao.

Keywords: groups of Lie type, laws, word maps.

1. Introduction

A *law* for a group G is an equation holding identically in G. Every finite group satisfies a law, and the length of the shortest law satisfied by the finite group G is a very natural measure of the complexity of G. In this paper we study the lengths of shortest laws in finite groups of Lie type.

1.1. Statement of results

Our main result is as follows. Let p be prime and let q be a power of p.

Theorem 1.1. *Let $G = X(q)$ be a finite simple group of Lie type over a finite field of order q, where X is the type of G. Then there exists a word $w_G \in F_2$ of length*

$$O_X(q^a \log(q)^{O_X(1)})$$

which is a law for G, where $a = a(X, p) \in \mathbb{N}$ is as in Table 1 below. Moreover, every

Henry Bradford: Christ's College, University of Cambridge, St Andrew's Street, Cambridge CB2 3BU, UK; hb470@cam.ac.uk

Andreas Thom: Institut für Geometrie, TU Dresden, 01062 Dresden, Germany; andreas.thom@tu-dresden.de

Mathematics Subject Classification 2020: 20F10 (primary); 20D05 (secondary).

law for G is of length $\Omega(q^a)$ unless $G = {}^2B_2(q)$, in which case every law for G is of length $\Omega(q^{1/2})$.

X	A_l	${}^2A_l$	B_l	C_l
a	$\lfloor (l+1)/2 \rfloor$	$\lfloor (l+1)/2 \rfloor$	$2\lfloor l/2 \rfloor$ ($l \geq 3$, q odd) l (otherwise)	l

X	D_l	${}^2D_l$	E_6	${}^2E_6$	E_7	E_8	F_4	G_2
a	$l-2$ ($l \geq 4$ even, q odd) $l-1$ (otherwise)	$2\lfloor l/2 \rfloor$	4	4	7	7	4	1

X	${}^3D_4$	${}^2B_2$	${}^2F_4$	${}^2G_2$
a	3	1	2	1

Tab. 1. Degree of polynomial part in lengths of shortest laws.

Observe that for a given X, a depends only on the parity of p, and that only for groups of type B or D. Hence for each fixed type X (except for ${}^2B_2$), and restricting q to be either odd or even, the asymptotic behaviour of the length of the shortest law holding in $X(q)$, as $q \to \infty$, is known up to a polylogarithmic factor. The fact of Suzuki groups being a difficult case of general statements about groups of Lie type, especially those with a quantitative or asymptotic flavour, has been observed with respect to several other problems. We discuss briefly some of these below.

The integer exponent a in Theorem 1.1 arises naturally in our proof in the following way: it is maximal such that $\mathrm{PSL}_2(q^a)$ occurs as a section of G (apart from for the Suzuki groups). To contextualise this bound, we may compare it with the minimal dimension n of a projective representation of G. It turns out that $a \leq \lfloor n/2 \rfloor$ in all cases (see Section A.1.2 in the appendix). Hence we have the following easy-to-state consequence of Theorem 1.1.

Corollary 1.2. *Let $G = X(q)$ be as in Theorem 1.1. Let $n = n(X, p)$ be the minimal dimension of a faithful projective module for G over $\overline{\mathbb{F}}_q$. Then there exists a word $w_G \in F_2$ of length*

$$O_n(q^{\lfloor n/2 \rfloor} \log(q)^{O_n(1)})$$

which is a law for G.

The exponent $\lfloor n/2 \rfloor$ in Corollary 1.2 agrees with the exponent a from Theorem 1.1 in some cases, for instance if $X = A_l$ or C_l. In general, however, a may be much smaller than $\lfloor n/2 \rfloor$. For comparison, the values of n for each pair (X, q) are listed in Section A.1.2.

The conclusion of Theorem 1.1 may easily be extended to other groups closely related to simple groups of Lie type.

Corollary 1.3. *Let $G = X(q)$, $a = a(X, p)$ be as in Theorem 1.1. Let $G \leq H \leq \mathrm{Aut}(G)$ and let $\hat{H}$ be a central extension of H. Then there exists a word $w_{\hat{H}} \in F_2$ of length $O_X(q^a \log(q)^{O_X(1)})$ which is a law for $\hat{H}$.*

Since a law for a group is also a law for any of its sections, the lower bounds for the length of the shortest law in F_2 satisfied by G also apply to $\hat{H}$. Groups $\hat{H}$ satisfying the hypothesis of Corollary 1.3 include the isometry groups $\mathrm{GL}_n(q)$, $\mathrm{GU}_n(q)$, $\mathrm{Sp}_n(q)$ and $\mathrm{GO}^{\varepsilon}_n(q)$ of the classical forms, corresponding respectively to $G = A_l(q)$, $G = {}^2A_l(q)$, $G = C_l(q)$ and $G = B_l(q)$, $D_l(q)$ or ${}^2D_l(q)$.

If we consider word maps which need only vanish on generating pairs for groups of Lie type then we may use much shorter words. For a group G and a word $w \in F_2$, define the *vanishing set* of w in G to be

$$Z(G, w) = \{(g, h) \in G \times G : w(g, h) = 1\}. \tag{1.1}$$

Define $c = c(X) \in \mathbb{N}$ by $c({}^2D_l) = c({}^3D_4) = 2$ and $c(X) = 1$ in all other cases.

Theorem 1.4. *Let $G = X(q)$ be as in Theorem 1.1. Then there exists a non-trivial reduced word $w_G \in F_2$ of length*

$$O_X(q^c \log(q)^{O_X(1)})$$

such that

$$\{(g, h) \in G : \langle g, h\rangle = G\} \subseteq Z(G, w).$$

Theorem 1.4 will be a tool in the proof of Theorem 1.1, as well as being a result of interest in its own right.

We make extensive use in the proofs of Theorems 1.1 and 1.4 of upper bounds on the diameters of finite simple groups. The key conjecture in this area is due to Babai.

Conjecture 1.5 ([2]). Let G be a non-abelian finite simple group. Then

$$\mathrm{diam}(G) = \log(|G|)^{O(1)}.$$

The definition of the diameter $\mathrm{diam}(G)$ of a finite group G is deferred until Section 2.3. Babai's Conjecture is still some way from being proven, in spite of remarkable progress in recent years. For the groups of Lie type, the state of the art is the following result.

Theorem 1.6 ([9, 37]). *Let $G = X(q)$ be as in Theorem 1.1 and let n be as in Corollary 1.2. Then*

$$\mathrm{diam}(G) \leq \log(|G|)^{O_n(1)}.$$

The dependence of the implied constant in Theorem 1.6 upon n is not given explicitly, and it is to be expected that the constants arising from existing proofs would be quite large. The implied constants in Theorem 1.1, Corollary 1.2 and Theorem 1.4 are similarly inexplicit. Nevertheless, conditional on Babai's Conjecture, and using the result of [26] we can say a little more.

Remark 1.7. Assume that Conjecture 1.5 is true. Let $G = X(q)$ and n be as in Corollary 1.2, let a be as in Theorem 1.1 and let $c(X)$ be as in Theorem 1.4. Then there exists a word $w_G \in F_2$ of length

$$O(a_n q^{a(X)} \log(q)^{O(b_n)})$$

which is a law for G, and a non-trivial reduced word $w'_G \in F_2$ of length

$$O(c_n q^{c(X)} \log(q)^{O(d_n)})$$

such that

$$\{(g,h) \in G : \langle g,h \rangle = G\} \subseteq Z(G,w'),$$

where a_n, b_n, c_n and d_n are functions of n which are explicitly computable.

For the sake of avoiding a large amount of tedious book-keeping, we will not give explicit bounds on the growth of a_n, b_n, c_n or d_n, and will content ourselves with remarking, at the appropriate points in the proofs of Theorems 1.1 and 1.4, where computing the dependence of the laws on n is a non-trivial matter, according to the best currently known dependences.

It is by now well-known that a generic pair of elements in a finite simple group of Lie type generates the group [19, 29, 32]. The words w_G arising in Theorem 1.4 are *almost laws* for G, in the sense that the probability that a random pair (g,h) of elements of G lies in $Z(G,w_G)$ tends to 1 as $|G| \to \infty$. In fact, for groups of bounded rank we can do even better: Breuillard, Green, Guralnick and Tao [8] showed that Cayley graphs of such groups with respect to random pairs of generators form *expanders*. In particular, these Cayley graphs have logarithmic diameter and lazy random walks on them have logarithmic mixing time. From this we conclude the following bound for the length of almost laws.

Theorem 1.8. *Let $G = X(q)$ be as in Theorem 1.1 and let c be as in Theorem 1.4. Then there exist non-trivial reduced words $w_G \in F_2$ of length*

$$O_X(q^c \log(q))$$

such that if g,h are independent uniform random variables on G then

$$\mathbb{P}[(g,h) \in Z(G,w_G)] \to 1 \quad \text{as } q \to \infty.$$

Theorem 1.8 complements the work of Zyrus [44], who produced short almost laws for $\mathrm{PSL}_n(q)$, and also provided lower bounds for the length of almost laws in this case.

1.2. Background

The study of the structure of laws in groups is a classical subject, growing out of the work of Birkhoff [3] in universal algebra, and further developed by many authors (see [36] and the references therein). At this point the behaviour of laws for finite simple groups was already a matter of considerable interest. For instance it was noted in [36] that non-isomorphic finite simple groups generate distinct varieties, and the question was posed whether there exists an infinite family of non-abelian finite simple groups satisfying a common law. Jones [18] answered this question in the negative for the alternating groups and groups of Lie type and the later completion of the classification of finite simple groups established that this was sufficient to give a negative answer in general. Jones' result tells

us that for any sequence $(G_i)_{i\in\mathbb{N}}$ of distinct finite simple groups of Lie type, the length of the shortest laws satisfied by G_i tends to infinity. The results of the present paper address the rate of this divergence, emphasizing the case of groups of bounded rank.

Prior to our work, the best upper bound on the length of laws for finite simple groups of Lie type was given by Kozma and the second named author.

Theorem 1.9. *Let G be a finite simple group of Lie type of Lie rank r, over a field of order q. Then there exists a word $w_G \in F_2$ of length at most $q^{O(r)}$ which is a law for G. Moreover, if $G = \mathrm{PSL}_n(q)$, then w_G is of length at most*

$$\exp(O(n^{1/2}\log(n)))q^{n-1}.$$

Theorem 1.9 builds upon the work of Hadad [17], and uses the Jordan decomposition to restrict the possibilities for the order of an element in $\mathrm{PGL}_n(q)$. The behaviour of element orders in groups of Lie type is a theme to which we shall return several times in what follows. The main result of [17] claimed a stronger upper bound, but the proof was found to contain a gap. Nevertheless, we have from [17] the following observation concerning lower bounds on the length of laws for finite simple groups of Lie type, which in particular shows that the exponent $\lfloor n/2\rfloor$ in Corollary 1.2 is best possible.

Theorem 1.10. *Let $k \geq 1$ and let $w \in F_k$ be a law for $\mathrm{PSL}_2(q)$. Then w has length at least $(q-1)/3$. We have $\mathrm{PSL}_2(q^{\lfloor n/2\rfloor})$ as a section of $\mathrm{PSL}_n(q)$ by restriction of scalars, so $\mathrm{PSL}_n(q)$ has no law of length less than $\Omega(q^{\lfloor n/2\rfloor})$.*

Complementary to these results for groups of Lie type, one may ask for short laws for the alternating and symmetric groups. The strongest available result in this direction is that found in [26].

Theorem 1.11. *There exists a law for $\mathrm{Sym}(n)$ of length at most*

$$\exp(O(\log(n)^4\log\log(n))).$$

Further, assuming Conjecture 1.5 holds, there exists a law for $\mathrm{Sym}(n)$ of length at most

$$\exp(O(\log(n)\log\log(n))).$$

Theorem 1.11 will also be useful for bounding the functions a_n, b_n, c_n and d_n in Remark 1.7, since groups of Lie type of large rank will contain large symmetric or alternating subgroups.

Meanwhile, short almost laws for the symmetric groups and for $\mathrm{PSL}_n(q)$ were produced by Zyrus [44]. In the latter case, lower bounds on the length of almost laws are also given.

Theorem 1.12 ([44]). *There are non-trivial reduced words $w_n \in F_2$ of length*

$$O(n^8\log(n)^{O(1)})$$

such that if g, h are independent uniform random variables on $\mathrm{Sym}(n)$ then

$$\mathbb{P}[(g,h) \in Z(\mathrm{Sym}(n), w_n)] \to 1 \quad \text{as } n \to \infty.$$

Theorem 1.13 ([44]). *There are non-trivial reduced words $w_{q,n} \in F_2$ of length*

$$O_n(q \log(q)^{O_n(1)})$$

such that if g, h are independent uniform random variables on $\mathrm{PSL}_n(q)$ then

$$\mathbb{P}[(g,h) \in Z(\mathrm{PSL}_n(q), w_{q,n})] \to 1 \quad \text{as } q \to \infty.$$

Further, any such words $w_{q,n} \in F_2$ are of length $\Omega_n(q)$.

It is expected that the methods used to prove the lower bound in this last result will extend to other finite simple groups of Lie type. This will be explored elsewhere.

As well as being of interest in their own right, the existence of short laws for finite simple groups may be applied to the provision of short laws for other groups, and indeed of laws holding simultaneously in all sufficiently small finite groups. This latter problem is also of interest in geometric group theory, where it is relevant to the residual finiteness growth of free groups, originally studied by Bou-Rabee [4] and later in [22]. The best known result in this direction is contained in a previous paper of the authors [6].

Theorem 1.14. *Let $\delta > 0$. For all $n \in \mathbb{N}$ there exists a word $w_n \in F_2$ of length*

$$O_\delta(n^{2/3} \log(n)^{3+\delta})$$

such that for every finite group G satisfying $|G| \leq n$, w_n is a law for G.

Theorem 1.1 in the specific cases of the groups $\mathrm{PSL}_3(q)$ and $\mathrm{PSU}_3(q)$ was applied in the proof of Theorem 1.14. We will also apply Theorem 1.14 when bounding the functions a_n, b_n, c_n and d_n occurring in Remark 1.7, since many small groups of undetermined structure arise as subquotients of G, and we shall require explicit bounds on lengths of laws satisfied by these. It is important to stress that there is no circularity in our reasoning here, since the application of Theorem 1.1 in the proof of Theorem 1.14 makes no use of the dependence of the implied constants in Theorem 1.1 on n.

1.3. Outline of the proof

Our approach to the proof of Theorem 1.1 was inspired by that of Theorem 1.11 in [26]. Indeed, the fact that the strategy of the proof of Theorem 1.11 could potentially provide a blueprint for producing short laws for groups of Lie type was already remarked upon in [26]. In both cases the problem is first divided into a search for words vanishing, respectively, on *generating* and *non-generating* pairs of elements in our group G.

In the present setting, the generating case is precisely the content of Theorem 1.4. Producing the desired word has two stages: first, we identify a large subset E of G on which some short word vanishes. In most cases, E will be the set of elements of G lying in some maximally split maximal torus T of G, so that all elements of E satisfy a power law of length equal to the exponent $e = \exp(T)$ of T, the latter being some small-degree polynomial in q. Second, we prove the existence of a small set of sufficiently short non-trivial words u_i with the following property: for any generating pair g, h of G, there exists

i such that the evaluation $u_i(g,h)$ of u_i at (g,h) lies in E. From this it will follow that the vanishing sets of the u_i^e's cover all generating pairs, and combining these words by a standard commutator trick, we will have the required conclusion. It is at this stage in the argument that bounds on the diameter of G become relevant: Theorem 1.6 guarantees that the evaluation of a random word u of length $\log(|G|)^{O_X(1)}$ at a fixed generating pair g,h is almost uniformly distributed on G. In particular, since E contains a positive proportion of the elements of G, $u(g,h)$ lies in E with probability bounded from below. It follows that if we pick our set of u_i sufficiently large and independent at random, the desired property will hold with positive probability, so at least one such set must exist. In the setting of Theorem 1.8 we have available the results of [8], and need only take random words of length about $\log(|G|)$.

In the non-generating case, we seek a word vanishing on all pairs in G which generate a proper subgroup. It therefore suffices to find a law holding simultaneously for all maximal subgroups of G. When G is a group of Lie type, there is a vast literature devoted to determining the structure of maximal subgroups of G, of which we shall avail ourselves.

For classical groups the seminal result on the structure of maximal subgroups is Aschbacher's Theorem [1], which draws a dichotomy between "geometric" and "non-geometric" subgroups. The geometric subgroups are those that preserve some extra geometric structure on the natural projective module for G: a direct-sum or tensor-product decomposition, for instance. They all have well-understood structure, being an extension built out of nilpotent groups, smaller groups of Lie type, and permutation groups of small degree. All the levels of the extension satisfy short laws (those for the groups of Lie type being obtained by induction) and from these we may easily produce a law valid on the whole extension.

The non-geometric subgroups are also very restricted: for instance they are all central extensions of almost simple groups. We may therefore invoke the CFSG, and examine each family separately. An alternating group or a group of Lie type in characteristic different from that of G cannot embed into G unless it is of very small order compared to G: this follows from the work of Wagner [40–42] for the alternating groups, and from that of Landazuri and Seitz [27] for the groups of Lie type in cross-characteristic. Sufficiently short laws for these groups are therefore easy to provide using Theorems 1.11 and 1.14. This leaves us with the case of groups of Lie type in defining characteristic (the sporadic simple groups are trivial for the purposes of asymptotic statements such as ours). Here the possibilities for the embedded group are restricted thanks to the representation-theoretic work of Donkin [14] and Liebeck [28], and we have sufficiently short laws for all subgroups that arise by induction.

For the exceptional groups of Lie type, the Aschbacher classes are not strictly defined, but the overall shape of the classification of maximal subgroups is very similar to that for the classical groups, as was elucidated in a series of papers (see [31] and the references therein, [24,25,34] and the discussion of the Suzuki groups in [7]) so that we may pursue a similar strategy to the classical case.

The *lower bounds* for the length of the shortest law for G appearing in Theorem 1.1 all follow from Theorem 1.10: the largest PSL_2 occurring as a section of $X(q)$ is $\mathrm{PSL}_2(q^a)$.

The reason that our upper and lower bounds do not match up to a polylogarithmic factor in the case of the Suzuki groups ${}^2B_2(q)$, and in this case alone, is that ${}^2B_2(q)$ does not have PSL_2-sections of unbounded size as q varies. The best available lower bound of $\Omega(q^{1/2})$ comes from [10], and is based on [18]. Roughly, the algebraic geometry of ${}^2B_2(q)$ is sufficiently well-controlled by that of the $\mathrm{Sp}_4(q)$ in which it sits, that any law for ${}^2B_2(q)$ of length much less than $q^{1/2}$ would also be a law for $\mathrm{Sp}_4(q)$. Since $\mathrm{Sp}_4(q)$ *does* contain $\mathrm{SL}_2(q)$, this is impossible. It is amusing to note that the Suzuki groups are outliers with respect to several other statements about groups of Lie type. For instance, as is well-known, they are the only non-abelian finite simple groups of order not divisible by 3. To give a deeper example, Kassabov, Lubotzky and Nikolov [21] sought to construct generating sets with respect to which the entire family of finite simple groups would form an expander family. Alas the Suzuki groups fell outside the scope of their methods (also owing to the absence of large SL_2 subgroups) and it was not until the later work of [10] that this gap was filled, by different methods.

The paper is structured as followed. In Section 2 we specify our notation; introduce some preliminaries on laws in groups, diameters and mixing times for random walks on finite groups, and prove Theorems 1.4 and 1.8. We also show how Corollary 1.3 follows from Theorem 1.1. In Section 3 we gather results on the structure of maximal subgroups in finite simple groups of Lie type and implement our inductive argument to show that they satisfy short laws. In Section 4 we put everything together and prove Theorem 1.1. This includes identifying subgroups which witness the lower bound in Theorem 1.1. In an appendix we gather together background material on algebraic groups, groups of Lie type, and automorphisms of finite groups, and prove a technical result (Proposition 2.14) about the orders of elements in groups of Lie type.

2. Preliminaries and laws for generating pairs

2.1. Notation

We make use of some notations which are standard in the theory of finite groups: for A and B groups, $A \times B$ refers to the direct product of A and B, while $A.B$ refers to an extension of undetermined structure with kernel A and quotient B. We denote by $A \circ B$ a *central product* of A and B, that is, a group of the form $(A \times B)/N$, where $N \lhd A \times B$ is the graph of an isomorphism between subgroups of $Z(A)$ and $Z(B)$.

For $n \in \mathbb{N}$, n will also denote the cyclic group of order n. In many of the sources to which we refer, $[n]$ will denote a group of order n of undetermined structure. For G a group and $g \in G$, $\mathrm{ccl}_G(g)$ will denote the *conjugacy class* of g in G. For $H < G$, $C_G(H)$ denotes the centraliser of H in G.

We use the Dynkin notation $X(q)$ for a finite simple group of Lie type over a field of order q, where

$$X \in \{A_l, {}^2A_l, B_l, C_l, D_l, {}^2D_l, {}^3D_4, E_6, {}^2E_6, E_7, E_8, F_4, G_2, {}^2B_2, {}^2G_2, {}^2F_4\}.$$

If X is a twisted type, we write $X(q)$ for the group whose defining Frobenius automorphism has fixed field of order q. By contrast, many authors use $X(q)$ to denote the group whose natural representation is defined over a field of order q. For instance, the group that we would denote ${}^2A_l(q)$, they would denote ${}^2A_l(q^2)$. The (generally) simple groups of linear, symplectic, unitary and orthogonal type will also be denoted $\mathrm{PSL}_n(q)$, $\mathrm{PSp}_n(q)$, $\mathrm{PSU}_n(q)$ and $\mathrm{P}\Omega_n^\varepsilon(q)$ (for $\varepsilon \in \{+,-,\circ\}$), with similar notation used to denote other groups in the same families in the standard fashion. Some sources, including [7, 23], also use the notation **L**, **S**, **U** and **O** to refer to these families of simple classical groups, respectively. We avoid this convention.

We use the *Landau notation* for functions: for $U \subseteq \mathbb{R}$ and $f, g : U \to \mathbb{R}$ we write $f = O(g)$ if there exists a positive constant C such that for all $x \in U$, $|f(x)| \leq C|g(x)|$. More generally, for $\{f_a : U \to \mathbb{R}\}_{a\in A}$ a family of functions, we write $f_a = O_a(g)$ if there exists a function $C\colon A \to (0,\infty)$ such that for all $a \in A$ and $x \in U$, $|f_a(x)| \leq C(a)|g(x)|$. Conversely, we write $f = \Omega(g)$ (respectively $f_a = \Omega_a(g)$) if $g = O(f)$ (respectively $g = O_a(f_a)$). We are therefore following the (stronger) definition of the symbol Ω due to Knuth, as opposed to that of Hardy–Littlewood. There should be no confusion in our use of the symbols O, Ω for both the Landau notation and for groups of orthogonal type, since the latter always appear with a superscript $+$, $-$ or $\circ$.

2.2. *Laws in finite groups*

Definition 2.1. Fix x, y an ordered basis for the free group F_2 and let $w \in F_2 \setminus \{1\}$. For any group G define the *evaluation map* $w : G \times G \to G$ by $w(g,h) = \pi_{(g,h)}(w)$, where $\pi_{(g,h)}$ is the unique homomorphism $F_2 \to G$ extending $x \mapsto g$, $y \mapsto h$. We call w a *law for* G if $w(G \times G) = \{1_G\}$.

We stress that the identity element of F_2 is by definition *not* a law for any group G. We could of course more generally have defined laws in the free group F_k of any finite rank $k \geq 1$, but taking an embedding $F_k \leq F_2$ allows us to transform any law in F_k into a law in F_2, changing the length by at most a constant factor.

Example 2.2. (i) If w is a law for G, then it is also a law for every subgroup and every quotient of G.

(ii) G is abelian iff $x^{-1}y^{-1}xy$ is a law for G.

(iii) If G is a finite group, then $x^{|G|}$ is a law for G. In particular, G satisfies some law.

We note two basic facts about the structure of laws in finite groups, which will enable us to construct new laws from old. The first allows us to combine words vanishing on subsets of a group to a new word vanishing on the union of those subsets, and is proved in [26, Lemma 2.2]. To this end, recall for G a group and $w \in F_2$ a word, the definition of the *vanishing set* $Z(G,w)$ of w on G from [39]:

$$Z(G,w) = \{(g,h) \in G \times G : w(g,h) = 1_G\}.$$

Lemma 2.3. *Let $w_1, \dots, w_m \in F_2$ be non-trivial words. Then there exists a non-trivial word $w \in F_2$ of length at most $16m^2 \max_i |w_i|$ such that for all groups G,*

$$Z(G, w) \supseteq Z(G, w_1) \cup \cdots \cup Z(G, w_m).$$

Example 2.4. From Lemma 2.3 we may quickly prove the upper bound in Theorem 1.1 for $X = A_1$. It is well-known that for any $g \in \mathrm{SL}_2(q)$, the order $o(g)$ of g divides $q-1$, q or $q+1$. Applying Lemma 2.3 to the words $w_1 = x^{q-1}$, $w_2 = x^q$ and $w_3 = x^{q+1}$ we obtain a word of length $O(q)$ which is a law for $\mathrm{SL}_2(q)$, and hence also for $A_1(q) = \mathrm{PSL}_2(q)$. Note however that this approach to Theorem 1.1 already fails for $X = A_2$, since $A_2(q) = \mathrm{PSL}_3(q)$ has elements of order $\Omega(q^2)$.

Note that, as well as allowing us to increase the vanishing set of words within a single group, Lemma 2.3 allows us to take a family of groups and, given a law for each group in the family, produce a new law which holds in every group in the family simultaneously. For instance we have the following observation, which previously appeared in [5].

Example 2.5. For $1 \leq i \leq m$ let $w_i = x^i$. Applying Lemma 2.3 to the words w_i, we obtain a non-trivial word $w \in F_2$ of length at most $16m^3$ such that for every group G satisfying

$$\max\{o(g) : g \in G\} \leq m,$$

w is a law for G.

Relatively short laws for finite simple groups of Lie type were already constructed in [39] using Example 2.5 (since these groups do not contain elements of very large order, relative to their size). Although these laws are too weak for the conclusion of Theorem 1.1, they will be useful in the proof of Theorem 1.1 nonetheless, when in the course of our induction argument for dealing with maximal subgroups of $G = X(q)$, we encounter a large number of subgroups defined over proper subfields of the field over which G is defined.

Proposition 2.6. *Let $a(X, p)$ be as in Table 1. For every $N \in \mathbb{N}$, there exists a word $w_{X,N} \in F_2$ of length $O(N^{6a(X,p)})$ such that for every prime power $q \leq N$, $w_{X,N}$ is a law for $X(q)$.*

Proof. Bounds on the maximal element orders of the $X(q)$ are given in Proposition A.4. Comparing Tables 1 and 3, we have $d(X) \leq 2a(X, p)$ in all cases, except for $X = D_1$ or ${}^2D_1$. By Theorem A.1 (ii), $D_1(q)$ and ${}^2D_1(q)$ are abelian for all q, so satisfy laws of bounded length. In all other cases the result is now immediate from Example 2.5. ∎

Remark 2.7. Although Proposition 2.6 produces laws which are simultaneously valid in all sufficiently small groups of a fixed type X, they are longer than the analogous simultaneous laws arising from Theorem 1.1 in almost all cases. For we may combine by Lemma 2.3 the laws produced in Theorem 1.1 for $X(q)$ as q ranges over prime powers less than N (or only over powers of 2 or 3 in the cases $X = {}^2B_2, {}^2F_4$ or ${}^2G_2$). The laws

obtained this way are shorter than those constructed in Proposition 2.6 in all cases except $X = A_1$.

We also obtain from Lemma 2.3 a construction of laws for direct products of groups.

Corollary 2.8. *Let $G_1, \dots, G_m$ be groups, and suppose that for $1 \leq i \leq m$, $w_i \in F_2$ is a law for G_i. Then $G = G_1 \times \cdots \times G_m$ has a law of length at most $16m^2 \max_i |w_i|$.*

Proof. Let w be as in Lemma 2.3. Then for each i,

$$Z(G_i, w) \supseteq Z(G_i, w_i) = G_i \times G_i.$$

Thus, for any $g = (g_i), h = (h_i) \in G$, $w(g,h) = (w(g_i, h_i)) = (1_{G_i}) = 1_G$. ∎

Our second fact is that the length of shortest laws behaves well for group extensions. It appears (in slightly weaker form) in [39, Lemma 2.1].

Lemma 2.9. *Let $1 \to N \to G \to Q \to 1$ be an extension of groups. Suppose N, Q satisfy non-trivial laws in F_2 of length n, m, respectively. Then G satisfies a non-trivial law in F_2 of length at most nm.*

Proof. Let w_N, w_Q be laws of minimal length for N, Q, respectively. We may assume both w_N and w_Q are cyclically reduced. Note that for any $g, h \in G$, $w_Q(g,h) \in N$. Suppose first that $w_Q(x,y)$ is a power of one of our basis elements x or y. Then $w_Q(x,x)$, $w_Q(y,y)$ are both laws for Q, and freely generate a non-abelian free subgroup of $F(x,y)$, so $w_N(w_Q(x,x), w_Q(y,y))$ is a law for G of the required length.

If w_Q is not a power of a basis element, then there exists a cyclic permutation w_Q' of w_Q such that w_Q, w_Q' freely generate a non-abelian free subgroup of $F(x,y)$. Moreover, w_Q' is also a law for Q, so $w_N(w_Q, w_Q')$ is a law for G. ∎

Example 2.10. Combining Lemma 2.9 with Example 2.2 (ii), we obtain for every $d \geq 1$ a non-trivial word $w_d \in F_2$ of length at most 4^d which is a law for every soluble group of derived length at most d, and therefore for every nilpotent group of step at most 2^d. These conclusions have been improved upon by Elkasapy and the second author [15, 16].

Proof of Corollary 1.3. Let G be a finite simple group of Lie type. Let l be the length of the shortest word in F_2 which is a law for G. Let H and $\hat{H}$ be as in the statement of Corollary 1.3. Then $H/G \leq \mathrm{Out}(G)$, hence by Theorem A.5, H/G is soluble of derived length at most 3, so that by Lemma 2.9 and Example 2.10, H satisfies a law of length at most $64l$. There is an abelian normal subgroup $Z \lhd \hat{H}$ such that $\hat{H}/Z \cong H$, so that by Lemma 2.9 again, $\hat{H}$ has a law of length at most $256l$. The result now follows from Theorem 1.1. ∎

2.3. Diameter, expansion and random walks

Let G be an arbitrary finite group, and let $S \subseteq G$ be a generating set. Recall that S determines a left-invariant *word metric* on G; the ball about 1 of radius $n \in \mathbb{N}$ in this

metric is

$$B_S(n) = \{s_1 \cdots s_n : s_1, \ldots, s_n \in S \cup S^{-1} \cup \{1\}\}.$$

The *diameter of G with respect to S* is the quantity

$$\operatorname{diam}(G, S) = \min\{n \in \mathbb{N} : B_S(n) = G\}$$

and the diameter of G itself (often referred to as the *worst-case diameter* of G) is

$$\operatorname{diam}(G) = \max\{\operatorname{diam}(G, S) : S \subseteq G, \langle S\rangle = G\}.$$

We shall make use of known bounds on diam for finite simple groups, or more specifically the consequences of such bounds for random walks on such groups. Suppose S is symmetric. Let $x_1, \ldots, x_L$ be independent random variables, each with distribution function

$$\frac{1}{2|S|}\chi_S + \frac{1}{2}\delta_{1_G}, \tag{2.1}$$

where χ_S is the indicator function of S and δ_{1_G} is the Dirac mass at the identity, and let ω_L be the random variable on G given by $\omega_L = x_1 \cdots x_L$. It is well-known that the number L of steps taken for ω_L to approach the uniform distribution is controlled by the *spectral gap* of the distribution function (2.1) [33, Theorem 5.1], which in turn is controlled by the diameter of G with respect to S [13, Corollary 3.1]. Hence we have the following result.

Theorem 2.11. *Let $S \subseteq G$ be a symmetric generating set and let $E \subseteq G$. Then*

$$\mathbb{P}[\omega_L \in E] \geq |E|/(2|G|) \quad \textit{for all } L \geq 2|S| \operatorname{diam}(G, S)^2 \log(2|G|).$$

Given this theorem, the following is an immediate consequence of Theorem 1.6 and, where relevant, the conclusion of Conjecture 1.5.

Corollary 2.12. *Let $G = X(q)$ be as in Theorem 1.1; let $S \subseteq G$ be a generating set and let $E \subseteq G$. Then*

$$\mathbb{P}[\omega_L \in E] \geq |E|/(2|G|) \quad \textit{for all } L \geq \Omega_X(|S| \log(q)^{\Omega_X(1)}).$$

Moreover, assuming Conjecture 1.5, the same conclusion holds for all $L \geq \Omega(|S| \log(q)^{\Omega(1)})$.

For the conclusion of Theorem 1.8, we will need the stronger conclusion of logarithmic mixing time satisfied by generic generating pairs, which follows from the results of [8] and Theorem 2.11.

Theorem 2.13. *Let $G = X(q)$ be as in Theorem 1.1. Let $g, h \in G$ be elements chosen independently uniformly at random. Then with probability tending to 1 as $q \to \infty$, $S = \{g^{\pm 1}, h^{\pm 1}\}$ satisfies the following conclusion. For any $E \subseteq G$,*

$$\mathbb{P}[\omega_L \in E] \geq |E|/(2|G|) \quad \textit{for all } L \geq \Omega_X(\log(q)).$$

2.4. Laws for generating pairs

We now prove Theorems 1.4 and 1.8. We thereby also reduce Theorems 1.1 to known structural results on maximal subgroups of finite simple groups of Lie type, to be described in the following section. The proof of Theorem 1.1 will be completed in Section 4.

The strategy in this subsection closely mimics that employed in [26]. To wit, we identify a large subset $E \subseteq G$ which may be seen to lie within the vanishing set of a short word. Then we run simultaneously a large number of short random walks on G. Using Theorem 1.6 (or Conjecture 1.5), we see that with high probability, at least one of our random walks lands in E. It follows that *as a deterministic fact*, there exists a set W of short words (of controlled size) such that under any evaluation in G, some member of W lies in E. We can then easily substitute the elements of W into a word vanishing on E and combine the words arising to obtain the required result.

First, let us specify the set E. In [26], $E \subseteq \mathrm{Sym}(n)$ was the set of n-cycles, so that the vanishing set of the word x^n contained E. Our set E will similarly satisfy a short power-law $x^{q^c \pm 1}$, where q is the order of the underlying field of G. In all cases the exponent $c = c(X)$ will be as in Theorem 1.4.

Proposition 2.14. *For G a finite group and $m \in \mathbb{N}$, let*

$$E_G(m) = \{g \in G : o(g) \text{ divides } m\}.$$

Let $G = X(q)$ be as in Theorem 1.1. Define $b(X,q) \in \mathbb{N}$ to be $q+1$ for $X = {}^2A_l$ or ${}^2E_6$; q^2-1 for $X = {}^2D_l$; q^2-q+1 for $X = {}^3D_4$; and $q-1$ otherwise. Then

$$|E_G(b(X,q))| = \Omega_X(|G|).$$

We imagine that the conclusion of Proposition 2.14 is well-known to the experts, but we have been unable to locate a unified proof in the literature, hence we present one in Appendix A below. For now, let us deduce Theorems 1.4 and 1.8 by combining Proposition 2.14 with, respectively, Corollary 2.12 and Theorem 2.13.

Proof of Theorem 1.4. Let $E_G = E_G(b(X,q))$ be as in Proposition 2.14. Let $u_1, \ldots, u_m$ be the results of m independent lazy random walks of length $L = C_1 \log(|G|)^{C_1}$ on a free generating set for F_2, where $C_1 = C_1(X)$ is sufficiently large (to be determined).

Fix (temporarily) a generating pair $g, h \in G$. For each $1 \leq i \leq m$, the probability that $u_i(g,h) \in E_G$ is at least $C_2 = C_2(X) > 0$, by Proposition 2.14 and Corollary 2.12 (since C_1 is assumed to be sufficiently large, depending on X, Corollary 2.12 does indeed apply here).

By independence of the u_i, the probability that $u_i(g,h) \notin E_G$ for every $1 \leq i \leq m$ is at most $(1-C_2)^m$. Setting $m = C_3 \log(|G|)$ for $C_3 = C_3(X) > 0$ sufficiently large, we may take $(1-C_2)^m < |G|^{-2}$.

Now, the number of possible generating pairs (g,h) for G is at most $|G|^2$. Taking a union bound over all such pairs (g,h), the probability of the event "for every

$1 \leq i \leq m$ there exists a generating pair $g, h \in G$ such that $u_i(g,h) \notin E_G$" is at most $(1-C_2)^m |G|^2 < 1$.

Therefore there exist *deterministically* words $u_1, \ldots, u_m \in F_2$ with $m = C_3 \log(|G|) = O_X(\log(q))$ of length at most $L = \log|G|^{O_X(1)} = O_X(\log(q)^{O_X(1)})$ such that for every generating pair $g, h \in G$, there exists $1 \leq i \leq m$ such that $u_i(g,h) \in E_G$. Note also that, since we may just as well remove 1_G from E_G without changing the argument, we may assume without loss of generality that all u_i are non-trivial in F_2. Since F_2 is torsion-free,

$$u_1^{b(X,q)}, \ldots, u_m^{b(X,q)}$$

are non-trivial words of length at most $O_X(q^{c(X)} \log(q)^{O_X(1)})$. By the definition of E_G, every generating pair $g, h \in G$ lies in $Z(G, u_i^{b(X,q)})$ for some $1 \leq i \leq m$. Combining the $u_i^{b(X,q)}$ by applying Lemma 2.3, we obtain a word satisfying the conditions of Theorem 1.4, of length $O_X(q^{c(X)} \log(q)^{O_X(1)})$. ∎

For the improvements required in Remark 1.7, we argue as before, but by assuming Conjecture 1.5 we may take walks of length $L = C_1 \log(|G|)^{C_1}$ with C_1 an absolute constant.

Proof of Theorem 1.8. The same proof applies, but we restrict to generating pairs g, h for which the conclusion of Theorem 2.13 applies, so that we may take walks of length $L = C_1 \log(|G|)$ with $C_1 = C_1(X) > 0$. ∎

3. Laws for non-generating pairs

Having established laws valid for generating pairs in groups of Lie type in Section 2.4, we turn our attention to laws valid for *non-generating* pairs. Trivially, if $g, h \in G$ satisfy $\langle g, h \rangle \neq G$, then there exists a maximal subgroup $M \lneq G$ such that $g, h \in M$. Therefore, our goal will be to describe the maximal subgroups of finite quasisimple groups of Lie type, and produce short laws which they satisfy.

3.1. Structure of maximal subgroups

For a given Lie type, we will identify finitely many families of subgroups, such that every maximal subgroup lies in at least one family, and produce a law valid in each family in turn. We will then combine the laws for the various families, using Lemma 2.3. Crucially, the number of laws we produce in this way will depend only on the Lie type, and not on the field order q.

Fortunately, there is an extensive literature on the maximal subgroups of finite simple groups, much of it developed in the decade following the completion of the CFSG. As discussed in the Introduction, it transpires that all maximal subgroups are an extension of groups for which sufficiently short laws are already available: they are of small order, nipotent of small class, permutation groups of small degree, or smaller groups of Lie type.

For the groups of Lie type which occur, we obtain sufficiently short laws by invoking Theorem 1.1 for those groups and applying induction. To implement our induction, we introduce the following strict partial ordering on finite simple groups of Lie type:

Definition 3.1. Let $H = Y(p^\mu), G = X(p^\lambda)$ be groups of Lie type in characteristic p. We declare that $H \prec G$ if one of the following holds:

(i) $n(Y, p) < n(X, p)$;

(ii) $n(Y, p) = n(X, p)$, X is classical and Y is exceptional;

(iii) $n(Y, p) = n(X, p)$, X and Y are exceptional and either

(a) $X = Y$ and a' is a proper divisor of a with a/a' prime, or

(b) (G, H) is one of $(E_6(p^{2a'}), {}^2E_6(p^{a'}))$, $(F_4(p^a), {}^2F_4(p^a))$ (for $p = 2$) and $(G_2(p^a), {}^2G_2(p^a))$ (for $p = 3$);

(iv) $n(Y, p) = n(X, p)$, X and Y are classical and either

(a) $X \neq {}^2A_l$ and a' is a proper divisor of a, or

(b) $X = {}^2A_l$ and a' is a proper divisor of $2a$;

(v) $n(Y, p) = n(X, p)$ and either

(a) $G = A_l(p^a)$ and H is one of $B_{l'}(p^a)$, $C_{l'}(p^a)$, $D_{l'}(p^a)$ or ${}^2D_{l'}(p^a)$, or

(b) $p = 2$, $G = C_l(p^a)$ and H is one of $B_{l'}(p^a)$, $D_{l'}(p^a)$ or ${}^2D_{l'}(p^a)$

(we refer to Theorem A.2 for the pairs of values (l, l') which yield the same value of $n(X, p)$ in (v)).

We then extend $\prec$ to be a transitive relation.

Remark 3.2. (i) It is straightforward to verify that "$\prec$" is a well-defined strict partial ordering.

(ii) It is especially important to note that the following relations hold, as consequences of the above:

$$\begin{aligned} \mathrm{PSL}_n(q) &\prec \mathrm{PSU}_n(q) && \text{(by (iv)(b));} \\ \mathrm{PSp}_n(q) &\prec \mathrm{PSU}_n(q) && \text{(by (v)(a) and transitivity);} \\ \mathrm{PSU}_n(q) &\prec \mathrm{PSL}_n(q^2) && \text{(by (iv)(a)).} \end{aligned}$$

(iii) As we shall see, if H is a simple group of Lie type in the same characteristic as G, which arises as a proper section of G, then $H \prec G$. The converse does not hold, but it will be much easier in practice to work with "$\prec$" than to attempt to perform induction on the family of sections directly.

3.2. Geometric subgroups of classical groups

For the classical groups, the key result on the structure of maximal subgroups is Aschbacher's Theorem [1]. Aschbacher's paper defines eight classes of subgroups, denoted $\mathcal{C}_1$–$\mathcal{C}_8$, and known collectively as the "geometric subgroups". We shall not define

these classes precisely, so suffice it to say that each type is associated with some extra geometric structure on the natural module associated with our group. Aschbacher's Theorem asserts that every maximal subgroup of a quasisimple classical group either belongs to one of the classes $\mathcal{C}_i$, or belongs to the class of "non-geometric subgroups", denoted $\mathcal{S}$. Moreover, every subgroup in $\mathcal{S}$ is an almost simple group satisfying certain additional irreducibility conditions.

We make no claims as to the disjointness of the families of subgroups described below, or that every subgroup we consider is indeed maximal in the corresponding classical group. All that we require is that every maximal subgroup appears at least once in one of the nine classes.

Before stating the structure theorem for maximal subgroups of classical groups we introduce some additional terminology. Let $n = n(X, p)$ be as in Corollary 1.2.

Definition 3.3. Let $G = X(q)$ be a finite simple group of Lie type with $q = p^\lambda$, p prime. Let S be a section of G.

(i) S is a *Lie-like level for* G if there exist $m, n_i, \lambda_i \in \mathbb{N}_{>0}$ with

$$\sum_{i=1}^{m} n_i \lambda_i \leq n(X, p); \tag{3.1}$$

groups of Lie type $G_1 = Y_1(q^{\lambda_1}), \ldots, G_m = Y_m(q^{\lambda_m})$ such that $n_i = n(Y_i, p)$, $\lambda_i a(Y_i, p) \leq a(X, p)$ and $G_i \prec G$; and a finite abelian group A such that S is a quotient of

$$L = A \times G_1 \times \cdots \times G_m.$$

(ii) S is a *subfield level for* G if either (a) there is a group $H = Y(p^\mu)$ of Lie type with $S \cong H$ such that μ is a proper divisor of λ, with λ/μ prime and $X = Y$ or $(X, Y) = (D_l, {}^2D_l)$, or (b) $G = \mathrm{PSU}_n(q)$ and $S = \mathrm{Sp}_n(q)$ (for n even) or $\mathrm{SO}^\epsilon_n(q)$.

(iii) S is a *p-level for* G if it is a p-group.

Remark 3.4. The number of possibilities for L/A is bounded by a function of n alone where L, A are as in Definition 3.3 (i).

Theorem 3.5. *Let M be a maximal subgroup of G. Then either* (geometric type) *there exists a subnormal series*

$$M = K_1 \rhd K_2 \rhd K_3 \rhd K_4 \rhd K_5 = 1$$

such that one of the following holds:

$\mathcal{C}_1$: *$K_4 = 1$, K_3/K_4 is a p-level for G, K_2/K_3 is a Lie-like level for G, and K_1/K_2 is abelian;*

$\mathcal{C}_2$: *$K_4 = 1$, K_3/K_4 is a Lie-like level for G, K_2/K_3 is abelian, and K_1/K_2 is a subgroup of* $\mathrm{Sym}(n)$*;*

$\mathcal{C}_3$: *$K_4 = 1$, K_3/K_4 is a Lie-like level for G, K_2/K_3 is abelian, and $|K_1/K_2| \leq 2n$;*

$\mathcal{C}_4$: $K_3 = 1$, K_2/K_3 *is a Lie-like level for* G, *and* $|K_1/K_2| \leq n$;

$\mathcal{C}_5$: $K_3 = 1$, K_2/K_3 *is a subfield level for* G, *and* $|K_1/K_2| \leq n$;

$\mathcal{C}_6$: $K_3 = 1$, K_2/K_3 *is a* 2*-step nilpotent group, and* K_1/K_2 *is a subgroup of* $\mathrm{Sym}(n^2)$;

$\mathcal{C}_7$: K_4 *is abelian,* K_3/K_4 *is a Lie-like level for* G, $|K_2/K_3| \leq n^2$, *and* K_1/K_2 *is a subgroup of* $\mathrm{Sym}(n)$;

$\mathcal{C}_8$: $K_4 = 1$, K_3/K_4 *is abelian,* K_2/K_3 *is a Lie-like level for* G, *and* K_1/K_2 *is abelian;*

or (non-geometric type *or* type $\mathcal{S}$) *there is a non-abelian finite simple group* S *such that* $S \leq M \leq \mathrm{Aut}(S)$ *and the preimage* $\hat{S}$ *of* S *in the full covering group* $\hat{G}$ *of* G *is absolutely irreducible on the natural module for* $\hat{G}$.

Proof. This is immediate from [7, Theorem 2.2.19], which in turn is based on [23, Main Theorem]. The class $\mathcal{S}$ is described in [7, Definition 2.1.3], and the structure of the subnormal series for the groups in cases $\mathcal{C}_1$–$\mathcal{C}_8$ follows from [7, Tables 2.3, 2.5–2.11], noting that in case $\mathcal{C}_6$, $n = r^m$ for r prime, so that the natural module for $K_1/K_2 = \mathrm{Sp}_{2m}(r)$, $\mathrm{SO}^{\pm}_{2m}(r)$ or $\Omega^{\pm}_{2m}(r)$ is a set of order $r^{2m} = n^2$ on which K_1/K_2 acts faithfully. ∎

Recall the strict partial order $\prec$ introduced in Definition 3.1.

Proposition 3.6. *Assume that all finite simple groups of Lie type* H *with* $H \prec G$ *satisfy a law as in Theorem 1.1. Then there is a word* $w_{\mathrm{geom}} \in F_2$ *of length* $O_X(q^{a(X,p)} \log(q)^{O_X(1)})$ *such that if* $M \leq G$ *is a geometric maximal subgroup as in cases* $\mathcal{C}_1$–$\mathcal{C}_8$ *of Theorem 3.5, then* w_{geom} *is a law for* M.

The following will be used to deal with the p-levels.

Lemma 3.7. *Suppose* q *is a power of the prime* p, *and let* P *be a Sylow* p*-subgroup of* $\mathrm{PGL}_n(q)$. *Then* P *is nilpotent of class at most* $n - 1$.

Corollary 3.8. *There is a non-trivial word in* F_2, *of length depending only on* n, *which is a law for every* p*-subgroup of* $\mathrm{PGL}_n(q)$.

Proof. This follows from Lemma 3.7 and Example 2.10. ∎

Proof of Proposition 3.6. By Lemmas 2.3 and 2.9, it will suffice to provide a law of appropriate length for each of the factors K_i/K_{i+1} in the subnormal series for M in each of the cases $\mathcal{C}_1$–$\mathcal{C}_8$.

Abelian factors and the 2-step nilpotent factor from case $\mathcal{C}_6$ satisfy laws of bounded length. The p-level from case $\mathcal{C}_1$ is dealt with by Corollary 3.8 (with the length of the law obtained as in Example 2.10). Those factors of order bounded by a polynomial function of n (occurring in cases $\mathcal{C}_3$, $\mathcal{C}_4$, $\mathcal{C}_5$ and $\mathcal{C}_7$) are handled by Theorem 1.14. Those factors embedding into $\mathrm{Sym}(n)$ (cases $\mathcal{C}_2$ and $\mathcal{C}_7$) or $\mathrm{Sym}(n^2)$ (case $\mathcal{C}_6$) are handled by Theorem 1.11 (with the second, stronger bound being available in the setting of Remark 1.7).

We are left with the groups occurring as Lie-like or subfield levels for G (these being the only cases in which the induction hypothesis is actually used).

First suppose S is a Lie-like level for G. Let L, A, m, G_i, n_i and λ_i be as in Definition 3.3 (i). By Lemma 2.9 it suffices to produce a short law valid simultaneously in all possible L/A. Indeed, by Remark 3.4 and Lemma 2.3, it suffices to produce a short law for each of the possible groups L/A individually. By induction, each G_i satisfies a law of length

$$O_{X_i}(q^{\lambda_i a(X_i,p)}(\lambda_i \log(q))^{O_{X_i}(1)}) = O_X(q^{a(X,p)} \log(q)^{O_X(1)})$$

(since by (3.1), $\lambda_i \leq n$ for all i, and either $n_i < n$ for all i, or $m = 1$ and $\lambda_1 = 1$). By Lemma 2.9, L satisfies a law of length at most

$$O_X(q^{a(X,p)} \log(q)^{O_X(1)}).$$

Therefore suppose S is a subfield level for G. In case (b) of Definition 3.3 (ii), S satisfies a sufficiently short law by induction (using Definition 3.1 (iv)(b)). Suppose therefore that we are in case (a) of Definition 3.3 (ii). Let $d = \lambda/\mu$ be the degree of the extension of the field over which G is defined, over the field over which S is defined. By Proposition 2.6, there exists a word w_{small} of acceptable length which is a law for all S such that $d \geq 7$. Meanwhile for $d = 2, 3$ or 5 there exists by induction a word w_d of acceptable length which is a law for S. Combining these laws by applying Lemma 2.3 we have a word of acceptable length which is a law for all subfield levels. ∎

3.3. Non-geometric subgroups of classical groups

Recall that, by Aschbacher's Theorem, if M is a maximal subgroup of G not lying in any of the classes $\mathcal{C}_1$–$\mathcal{C}_8$ above (that is, M is a *non-geometric* subgroup), there is a non-abelian finite simple group S such that $S \leq M \leq \mathrm{Aut}(S)$ and the full covering group $\tilde{S}$ of S is absolutely irreducible on the natural module V for G. In this subsection we show that there is a short law satisfied by all such M.

Proposition 3.9. *Assume that all finite simple groups H of Lie type with $H \prec G$ satisfy a law as in Theorem 1.1. Then there is $w_{\text{nongeom}} \in F_2$ of length $O_X(q^{a(X,p)} \log(q)^{O_X(1)})$ such that if $M \leq G$ is a non-geometric maximal subgroup as in cases 8 of Theorem 3.5, then w_{nongeom} is a law for M.*

As always, by Lemma 2.3 it suffices to produce boundedly many sufficiently short laws such that each M satisfies one of them. First, by the CFSG we may exclude the case of S sporadic: there are finitely many possibilities for the corresponding M, so M satisfies a law of *bounded* length. For the same reason, we may omit from consideration an arbitrarily large bounded number of other possibilities for S.

Second, $S \lhd M$, so we may naturally identify M with a subgroup of $\mathrm{Aut}(S)$, and M/S with a subgroup of $\mathrm{Out}(S)$. Thus M/S is soluble of derived length at most 3, so applying Lemma 2.9 and Example 2.10, it suffices to find a short law for S (note that we are not using the full power of the Schreier hypothesis here, since we already excluded sporadic groups). Let $\mathrm{Lie}(p)$ be the class of all finite simple groups of Lie type in characteristic p. We deal separately with the three cases: of S an alternating group; S a group of Lie type but $S \notin \mathrm{Lie}(p)$; and $S \in \mathrm{Lie}(p)$.

The following standard observation (following from Schur's Lemma) will allow us to move between linear representations of $\tilde{S}$ and projective representations of S.

Lemma 3.10. *Let G be a group, and let ρ be an absolutely irreducible linear representation of G. Then $\rho(Z(G))$ consists of scalar matrices.*

Consider first the case where $S \cong \mathrm{Alt}(m)$ is alternating. The work of Wagner allows us to bound m in terms of the dimension n of V alone. In particular, there is a law of length depending only on n which is satisfied simultaneously by all such S.

Theorem 3.11. *Let $m \geq 9$ and let ρ be a non-trivial absolutely irreducible representation of the full covering group $\widetilde{\mathrm{Alt}}(m)$ of $\mathrm{Alt}(m)$ over an arbitrary field $\mathbb{F}$.*

(i) *If $\rho(Z(\widetilde{\mathrm{Alt}}(m))) \neq 1$, then $\mathrm{char}(\mathbb{F}) \neq 2$, and $\dim(\rho) \geq 2^{\lfloor (m-s-1)/2 \rfloor}$, where there are non-negative integers $w_1 > \cdots > w_s$ such that*

$$m = 2^{w_1} + \cdots + 2^{w_s}.$$

In particular, for any $\varepsilon > 0$, $\dim(\rho) \geq 2^{(1-\varepsilon)m/2}$ for m sufficiently large.

(ii) *If $\rho(Z(\widetilde{\mathrm{Alt}}(m))) = 1$ and $\mathrm{char}(\mathbb{F}) \neq 0$, then $\dim(\rho) \geq m - 2$.*

Proof. If the hypothesis of (i) holds, then Theorem 1.3 of [42] applies, as the induced projective representation of $\mathrm{Alt}(m)$ is *proper*, in the sense of [42].

In the setting of (ii), ρ descends to a linear representation of $\mathrm{Alt}(m)$; then we are done by [40, Theorem 1.1] or [41, Theorem 1.1]. ∎

The bound for the length of the law satisfied concurrently by the alternating S arising in this case, required for Remark 1.7, follows straightforwardly from the dimension bounds in Theorem 3.11, Theorem 1.11 (recalling that Conjecture 1.5 is assumed in Remark 1.7) and the fact that $(\mathrm{Alt}(m))_m$ is an ascending nested sequence.

Second, suppose S is a simple group of Lie type with $S \notin \mathrm{Lie}(p)$. Lower bounds on the dimensions of cross-characteristic representations of finite simple groups of Lie type are provided by the work of Landazuri and Seitz [27]. From their very detailed results, we require only the following, which may be read off directly from the table in the main result of [27].

Theorem 3.12. *There exist explicit absolute constants $c_1, c_2 > 0$ such that the following holds. Let $S = Y(r)$ be a finite simple group of Lie type over a finite field of order r, and let m be the minimal dimension of a faithful projective representation of S over $\overline{\mathbb{F}}_r$. Let ρ be a non-trivial projective representation of S over $\mathbb{F}_q$. If $(r, q) = 1$ then*

$$\dim(\rho) \geq c_1 r^{c_2 m}.$$

We use Theorem 3.12 to bound the order of S by an explicit function of $n = n(X, p)$ alone. All such S will therefore satisfy a sufficiently short law by Theorem 1.14. Indeed, $\log_r(n/c_1) \geq c_2 m$ by Theorem 3.12, so that

$$\log_2(n/c_1) \log_r(n/c_1)/c_2^2 \geq \log_r(n/c_1)^2/c_2^2 \geq m^2$$

(since $2 \leq r$) and

$$(n/c_1)^{\log_2(n/c_1)/c_2^2} \geq r^{m^2} \geq |S|$$

as required.

Finally, suppose $S = Y(p^\mu) \in \mathrm{Lie}(p)$. The representations of almost simple groups of Lie type in defining characteristic were studied by Liebeck [28], building on work of Donkin [14], as a key step in bounding the orders of non-geometric maximal subgroups in classical groups. Recall that $n(Y, p)$ is the minimal dimension of a faithful irreducible projective $\overline{\mathbb{F}}_p S$-module (consult Section A.1.2 for the value of $n(Y, p)$ for each possible S).

Theorem 3.13 ([14], [28, Theorems 2.1–2.3]). *Let $S = Y(p^\mu)$ be a finite simple group of Lie type Y. Let M be a faithful absolutely irreducible projective S-module over $\mathbb{F}_{p^\lambda}$. Then*

$$\dim(M) \geq n(Y, p)^{\mu/(\lambda,\mu)}.$$

Recall that $G = X(q)$ is a classical group in characteristic p. Let $n(X, p)$ be as in Corollary 1.2. Note that complete lists of maximal subgroups of G are known for $n(X, p) \leq 12$ and are recorded in [7, Tables 8.1–8.5]. The possibilities for S being read off from these tables, it may be verified that there is a law of the required length satisfied by all of them, using our induction hypothesis and Lemma 2.3. We therefore henceforth assume that $n(X, p) \geq 13$. Write $p^\lambda = q^2$ if $X = {}^2A_l$ and $p^\lambda = q$ otherwise.

First suppose that $\mu/(\lambda, \mu) = t \geq 2$. Then $\mu \leq \lambda t$ and (by Theorem 3.13) $n(X, p) \geq n(Y, p)^t$. By induction, S satisfies a law of length

$$O_Y(p^{\mu\lfloor n(Y,p)/2\rfloor} \log(q)^{O_Y(1)})$$

(since $a(Y, p) \leq \lfloor n(Y, p)/2\rfloor$) which is acceptable since

$$\mu\lfloor n(Y, p)/2\rfloor \leq \lambda t \lfloor n(X, p)^{1/t}/2\rfloor \leq a(X, p)\lambda$$

for $X \neq {}^2A_l$ and

$$\lambda t \lfloor n(X, p)^{1/t}/2\rfloor \leq a(X, p)\lambda/2$$

for $X = {}^2A_l$, since $n(X, p) \geq 13$. Moreover, $n(Y, p) \geq 2$, so $t \leq \log_2(n(X, p))$ and $\lambda \leq 2\log_p(q)$, and the total number of possibilities for S is at most

$$O\Big(\log_p(q) \sum_{t=2}^{\log_2(n(X,p))} t n(X, p)^{1/t}\Big),$$

so by Lemma 2.3 we have a law of the required length satisfied by all such S (the factor of $\log_p(q)^2$ introduced when we apply Lemma 2.3 is acceptable, since $n(Y, p) \leq \sqrt{n(X, p)}$, so that $Y \neq X$).

Therefore we may suppose $\mu/(\lambda, \mu) = 1$, so $\mu \mid \lambda$. Write $\lambda = \mu r$ for $r \in \mathbb{N}$. First, for $r \geq 12$ we may argue as in the case of subfield levels in geometric subgroups: by Proposition 2.6 there is a word of length $O(q^{a(X,p)})$ which is a law for all possible S.

Next suppose $2 \leq r \leq 11$. Then by induction S satisfies a law of length

$$O_Y(p^{\mu\lfloor n(Y,p)/2\rfloor}\log(q)^{O_Y(1)}),$$

which is acceptable since

$$\mu\lfloor n(Y,p)/2\rfloor \leq \lambda\lfloor n(X,p)/2\rfloor/2 \leq a(X,p)\lambda$$

for $X \neq {}^2A_l$, while $\lambda\lfloor n(X,p)/2\rfloor/2 = a(X,p)\lambda/2$ for $X = {}^2A_l$. Moreover, at most $O(n(X,p))$ groups S occur in this case, so by Lemma 2.3 we have an acceptable law for all of them.

Finally, suppose $r = 1$, so that $\lambda = \mu$. Once again, recall that $\tilde{S}$ is a central extension of $Y(q)$ (respectively $Y(q^2)$ if $X = {}^2A_l$). Since $n(Y,p) \leq n(X,p)$, at most boundedly many Y arise for each X. It remains to prove that $Y(q) \preceq X(q)$ (respectively $Y(q^2) \preceq X(q)$) and $a(Y,p) \leq a(X,p)$ (respectively $2a(Y,p) \leq a(X,p)$), from which it follows by Lemma 2.3 there is a sufficiently short law satisfied by all S which occur.

We may easily exclude the case of Y exceptional. Indeed, X is classical and we have $n(X,p) \geq 13$, so if Y is exceptional with $n(Y,p) \leq n(X,p)$, then $Y(q) \preceq X(q)$ (respectively $Y(q^2) \preceq X(q)$) by Definition 3.1 (ii), and it may be seen by inspection of Table 1 above that $2a(Y,p) \leq a(X,p)$.

Therefore suppose Y is classical. We recall some more information from [28] on the possible dimensions of irreducible modular representations of $\tilde{S}$. Let K be the splitting field of $\tilde{S}$ (so that $K = \mathbb{F}_{p^{2\mu}}$ in the case $Y = {}^2A_l$ or ${}^2D_l$ and $K = \mathbb{F}_{p^\mu}$ otherwise).

Theorem 3.14 ([28, Theorem 1.1]). *Let M be an irreducible $K\tilde{S}$-module. Then either* (i) $\dim_K(M) = n(Y,p)$, (ii) $\dim_K(M) \geq n(Y,p)^2/2$ *or* (iii) *one of the following holds:*

(a) $\dim_K(M)$ *is as in the following table:*

X	$\dim_K(M)$	
A_l *or* ${}^2A_l$	$l(l+1)/2$	
B_l	$l(2l+1)$	
C_l	$l(2l-1)-2$	*if* $p \mid l$
	$l(2l-1)-1$	*if* $p \nmid l$
D_l *or* ${}^2D_l$	$l(2l-1)$	*for* p *odd*
	$l(2l-1)-1$	*for* $p=2$, l *odd*
	$l(2l-1)-2$	*for* $p=2$, l *even*

(b) $Y = B_l$ *or* C_l, $2 \leq l \leq 6$ *and* $\dim_K(M) = 2^l$;

(c) $Y = D_l$ *or* ${}^2D_l$, $4 \leq l \leq 7$ *and* $\dim_K(M) = 2^{l-1}$;

(d) $Y = C_3$ *and* $\dim_K(M) = 14$.

Suppose first that $n(Y,p) < n(X,p)$ (so that Definition 3.1 (i) applies). Since the action of $\tilde{S}$ on the natural module V for G is absolutely irreducible, $M = V \otimes_{\mathbb{F}_{p^\lambda}} K$ is an irreducible $K\tilde{S}$-module satisfying $\dim_K(M) > n(Y,p)$. Thus $n(X,p) = \dim_K(M)$ is

as in Theorem 3.14 (ii) or (iii). We can now verify by brute computation (and using Theorem A.2) that $2a(Y, p) \leq a(X, p)$ for X of type ${}^2A_l$ and $a(Y, p) \leq a(X, p)$ otherwise. These computations are greatly expedited by the observation that $a(Y, p) \leq \lfloor n(Y, p)/2 \rfloor$ and $a(X, p) \geq \lfloor n(X, p)/2 \rfloor - 2$ (since X is classical) and $a({}^2A_l, p) = \lfloor (l + 1)/2 \rfloor = \lfloor n({}^2A_l, p)/2 \rfloor$.

We may therefore assume $n(X, p) = n(Y, p)$. There are elementary methods which could exclude most possibilities for Y which would be potential obstructions to the conclusion of Theorem 1.1. However, we have found that the most efficient way of excluding all such Y simultaneously is to use the conclusion of Liebeck's investigations as a black box.

Theorem 3.15 ([28, Theorem 4.1]). *Let $G = X(q)$ for X of classical type with natural module of dimension n. Let H be a maximal subgroup of G. Then one of the following holds:*

(i) *H lies in $\mathcal{C}_1$–$\mathcal{C}_8$ (as described in Theorem 3.5) and $X \neq D_4$;*

(ii) *$H \in \{\mathrm{Alt}(n + 1), \mathrm{Sym}(n + 1), \mathrm{Alt}(n + 2), \mathrm{Sym}(n + 2)\}$;*

(iii) *$|H| < q^{3n}$ (respectively $|H| < q^{6n}$ for $X = {}^2A_l$).*

Cases (i) and (ii) of Theorem 3.15 having been dealt with above, we may assume that the order of S satisfies the upper bound from Theorem 3.15 (iii). This will contradict the following lower bound on the orders of classical groups.

Theorem 3.16. *Let Y be of classical type. Then*

$$|Y(q)| = \Omega_Y(q^{(n(Y,p)^2 - n(Y,p))/2}).$$

Proof. The orders of the finite simple groups of Lie type are computed in many places, for instance [11, 43]. ∎

Since $n(X, p) = n(Y, p) \geq 13$, we have $3n(X, p) < (n(Y, p)^2 - n(Y, p))/2$. Combining Theorem 3.15 (iii) with Theorem 3.16 yields $\Omega_Y(q^{(n(Y,p)^2 - n(Y,p))/2}) = |Y(q)| < q^{3n(X,p)}$ (respectively $\Omega_Y(q^{n(Y,p)^2 - n(Y,p)}) = |Y(q^2)| < q^{6n(X,p)}$ for $X = {}^2A_l$). This is a contradiction for q sufficiently large depending on $n(X, p)$. This concludes the last case of the proof of Proposition 3.9.

Remark 3.17. For the sake of Remark 1.7, the dependence of the implied constant in Theorem 3.16 may be made explicit.

It is important to note that the work of Section 2.4 and Section 3 up to this point do not by themselves constitute a proof of the upper bound in Theorem 1.1 for the classical groups. This is because a classical group may contain a large exceptional group lying prior to it in our induction, so that we must assume Theorem 1.1 for the exceptional subgroup in order to proceed. This means that our inductive argument in the proof of the upper bound in Theorem 1.1 must handle exceptional and classical groups simultaneously.

3.4. *Exceptional groups*

The broad shape of maximal subgroups of exceptional groups of Lie type is very similar to that for the classical groups, though the Aschbacher classes $\mathcal{C}_1$–$\mathcal{C}_8$ are not formally defined for the exceptional groups. Once again, every maximal subgroup has a short subnormal series, all the factors of which are either of small order; abelian, or a group of Lie type for which we may assume a sufficiently short law exists by induction.

Proposition 3.18. *Let $G = X(q)$ and $a(X, p)$ be as in Theorem 1.1, with*

$$X \in \{E_6, E_7, E_8, F_4, G_2, {}^3D_4, {}^2E_6, {}^2B_2, {}^2F_4, {}^2G_2\}.$$

Suppose that the conclusion of Theorem 1.1 holds for all groups H satisfying $H \prec G$ as in Definition 3.1. Then there is a word in F_2 of length $O(q^{a(X,p)} \log(q)^{O(1)})$ which is a law for all maximal subgroups M of G.

Proof. As for the classical groups, it will suffice by Lemma 2.3 to divide the maximal subgroups of G into a number of classes independent of q and to show that a law of the required length holds in each family.

First, we deal with the class of *subfield subgroups*, that is, subgroups of the form $M = X(q^{1/r})$, where r is a prime divisor of $\log_p(q)$. A law valid in all such subgroups is produced by the same argument that was used for the subfield levels in Proposition 3.6: there is a law of acceptable length valid in all such M with $r \geq 7$ by Proposition 2.6. For each of $r = 2, 3$ or 5, a law of acceptable length holds in M by induction. Thus by Lemma 2.3 there is a law of acceptable length valid in all subfield subgroups.

The case of the Suzuki groups ${}^2B_2(q)$ ($q = 2^{2m+1}$) follows from [7, Table 8.16]: every maximal subgroup of G is either a *subfield subgroup* or is soluble of derived length at most 3.

The case of the small Ree groups ${}^2G_2(q)$ ($q = 3^{2m+1}$) follows from [7, Table 8.43] (which in turn is based on [25]): every maximal subgroup is either a *subfield subgroup*, or is soluble of derived length at most 4, or is isomorphic to $2 \times \mathrm{PSL}_2(q)$.

The case of the Steinberg triality groups ${}^3D_4(q)$ follows from [7, Table 8.51] (which in turn is based on [24]). Other than the *subfield subgroups*, all maximal subgroups M of G have the following structure: there is a subnormal series

$$M = K_1 \rhd K_2 \rhd K_3 \rhd K_4 \rhd K_5 = 1$$

such that K_1/K_2 has order at most 24; K_2/K_3 is abelian; K_4 has order a power of q; and K_3/K_4 is isomorphic to one of the following:

$$\mathrm{SL}_2(q^3) \circ (q-1); \quad \mathrm{SL}_2(q) \circ (q^3-1); \quad G_2(q); \quad \mathrm{SL}_2(q^3) \circ \mathrm{SL}_2(q);$$
$$\mathrm{SL}_3(q) \circ (q^2+q+1); \quad \mathrm{SU}_3(q) \circ (q^2-q+1); \quad \mathrm{PGL}_3(q); \quad \mathrm{PGU}_3(q).$$

There is a sufficiently short law for K_4 by Corollary 3.8, for K_3/K_4 by hypothesis and Corollary 2.8, and thus also for M by Lemma 2.9.

The case of the large Ree groups ${}^2F_4(q)$ ($q = 2^{2m+1}$) follows from [34, Main Theorem] (note that the order q of the underlying field is, for historical reasons, denoted q^2 in [34], in spite of being an odd power of 2). Several isomorphism types of groups which are extensions of an abelian group by a group of bounded order occur, as do *subfield subgroups*. Other than these, every maximal subgroup M has a subnormal series:

$$M = K_1 \rhd K_2 \rhd K_3 \rhd K_4 = 1$$

such that K_1/K_2 is abelian, K_3 is a 2-group, and K_2/K_3 is isomorphic to one of the following:

$$\mathrm{PSL}_2(q) \times (q-1); \quad {}^2B_2(q) \times (q-1);$$
$$\mathrm{SU}_3(q); \quad \mathrm{PGU}_3(q); \quad {}^2B_2(q) \times {}^2B_2(q); \quad B_2(q).$$

We have a sufficiently short law for K_3 by Corollary 3.8, for K_2/K_3 by hypothesis and Corollary 2.8, and thus also for M by Lemma 2.9.

We may therefore assume that G is one of $G_2(q)$, $F_4(q)$, $E_6(q)$, ${}^2E_6(q)$, $E_7(q)$ or $E_8(q)$. Theorem 8 of [31] describes the possibilities for M:

(i-a) M is a maximal parabolic subgroup;

(i-b) M is a subgroup of *maximal rank*: the possibilities for M are given in [30, Tables 5.1, 5.2];

(i-c) $G = E_7(q)$ and $M = (2^2 \times P\Omega_8^+(q).2^2).\mathrm{Sym}(3)$ or ${}^3D_4(q).3$;

(i-d) $G = E_8(q)$ and $M = \mathrm{PGL}_2(q) \times \mathrm{Sym}(5)$;

(i-e) $F^*(M)$ is one of the groups appearing in [31, Table 3];

(ii) M is either a *subfield subgroup* or a *twisted type*, that is, one of the following: ${}^2E_6(q^{1/2}) \leq E_6(q)$ (for q a square), ${}^2F_4(q) \leq F_4(q)$ (for $q = 2^{2m+1}$) or ${}^2G_2(q) \leq G_2(q)$ (for $q = 3^{2m+1}$);

(iii) M is an *exotic local subgroup*, isomorphic to one of $2^3.\mathrm{SL}_3(2)$, $3^3.\mathrm{SL}_3(3)$, $3^{3+3}.\mathrm{SL}_3(3)$, $5^3.\mathrm{SL}_3(5)$ or $2^{5+10}.\mathrm{SL}_5(2)$;

(iv) $G = E_8(q)$ and $M = (\mathrm{Alt}(5) \times \mathrm{Alt}(6)).2^2$;

(v) $F^*(M)$ is one of the simple groups appearing in [31, Table 2];

(vi) $F^*(M)$ is a finite simple group of Lie type over a field $\mathbb{F}_{q_0}$ of characteristic p; $\mathrm{rk}(F^*(M)) \leq \mathrm{rk}(G)/2$, and there exists a constant $t(G)$, depending only on the root system of G, such that one of the following holds:

  (a) $q_0 \leq 9$;

  (b) $F^*(M) = \mathrm{PSL}_3(16)$ or $\mathrm{PSU}_3(16)$;

  (c) $q_0 \leq t(G)$ and $F^*(M) = \mathrm{PSU}_2(q_0)$, ${}^2B_2(q_0)$ or ${}^2G_2(q_0)$.

Recall that $F^*(M)$ is the generalized Fitting subgroup of M.

In case (ia), there exists a subnormal series

$$M = K_1 \rhd K_2 \rhd K_3 \rhd K_4 \rhd K_5 = 1$$

such that K_1/K_2 is of bounded order, K_2/K_3 is abelian, K_4 is a p-group, and $K_3/K_4 =$

$L \circ A$, where A is abelian and L is a central extension of a group $T \leq \bar{L} \leq \mathrm{Aut}(T)$, for T one of the following direct products of non-abelian simple groups.

G	T
$G_2(q)$	$A_1(q)$
$F_4(q)$	$B_3(q)$; $C_3(q)$; $A_1(q) \times A_2(q)$
$E_6(q)$	$D_5(q)$; $A_1(q) \times A_4(q)$; $A_1(q) \times A_2(q) \times A_2(q)$; $A_5(q)$
${}^2E_6(q)$	${}^2A_5(q)$; $A_1(q) \times A_2(q^2)$; $A_2(q) \times A_1(q^2)$; ${}^2D_4(q)$
$E_7(q)$	$E_6(q)$; $A_1(q) \times D_5(q)$; $A_2(q) \times A_4(q)$; $D_6(q)$; $A_1(q) \times A_5(q)$; $A_1(q) \times A_2(q) \times A_3(q)$; $A_6(q)$
$E_8(q)$	$E_7(q)$; $A_1(q) \times E_6(q)$; $D_5(q) \times A_2(q)$; $A_3(q) \times A_4(q)$; $A_1(q) \times A_2(q) \times A_4(q)$; $D_7(q)$; $A_1(q) \times A_6(q)$; $A_7(q)$

Each simple factor of each T arising satisfies a sufficiently short law by induction, and by the Schreier hypothesis $\mathrm{Out}(T)$ is the extension of a soluble group of derived length at most 3, by a permutation group of the isomorphic direct factors. By Lemmas 2.8 and 2.9, therefore, L satisfies a sufficiently short law.

In case (ib), for every group M appearing in [30, Table 5.1], there exists a subnormal series

$$M = K_1 \rhd K_2 \rhd K_3 \rhd K_4 \rhd K_5 = 1$$

such that K_1/K_2 is of bounded order, K_2/K_3 and K_4 are abelian, and K_3/K_4 is a direct product of a bounded number of finite simple groups $H_i \in \mathrm{Lie}(p)$, each satisfying $H_i \prec G$. Meanwhile every group appearing in [30, Table 5.2] is the extension of an abelian group by a group of bounded order.

In cases (ic) and (id), we have sufficiently short laws for $P\Omega_8^+(q)$, ${}^3D_4(q)$ and $\mathrm{PGL}_2(q)$ by hypothesis, so we may produce sufficiently short laws for M by Lemma 2.9.

For case (ie), $F^*(M)$ is the direct product of at most three non-abelian finite simple groups of Lie type in characteristic p, each of which satisfies a sufficiently short law by induction (seen by inspection of [31, Table 3]). By Proposition A.7, $M \leq \mathrm{Aut}(F^*(M))$. By Theorem A.5 and Proposition A.6, $\mathrm{Out}(F^*(M))$ is the extension of a soluble group of derived length at most 3 by a subgroup of Sym(3). By Lemma 2.9, it suffices that $F^*(M)$ satisfies a sufficiently short law. This is so, since each simple factor does, and by Corollary 2.8.

Among the subgroups arising in case (ii), the untwisted subfield subgroups (that is, those of the form $M = X(q^{1/r})$ for r a prime divisor of $\log_p(q)$) are dealt with as above. Those of twisted type satisfy a sufficiently short law by induction.

Finally, the subgroups arising in cases (iii)–(vi) are all of bounded order (in cases (v) and (vi) this follows from Proposition A.7, since $F^*(M)$ is simple of bounded order). ∎

4. Completing the proof of Theorem 1.1

At last we are ready to put everything together and prove our main result.

Proof of Theorem 1.1. First we show that a law of the required length does indeed exist. This is the upshot of the last two sections. To be explicit: Let G be as in Theorem 1.1. Let $w_{\text{gen}} \in F_2$ be the word produced in Theorem 1.4. Suppose by induction that a law of the required length exists for all $H \prec G$ as in Definition 3.1. If G is classical, let w_{geom} and w_{nongeom} be as in Propositions 3.6 and 3.9, respectively. If G is exceptional, let w_{exc} be as in Proposition 3.18. For $(g, h) \in G \times G$, either $\{g, h\}$ generates G or $\langle g, h\rangle$ is contained in a maximal subgroup of G. If G is classical then this subgroup is either geometric or non-geometric. Thus applying Lemma 2.3 to $w_{\text{gen}}, w_{\text{geom}}, w_{\text{nongeom}}$ (for G classical) or $w_{\text{gen}}, w_{\text{exc}}$ (for G exceptional) we have the required law.

The lower bound on the length of the shortest law in G from Theorem 1.1 is witnessed by the following subgroups, which exist for q larger than an absolute constant. There is no other way than doing this case by case.

- By Theorem 1.10, $A_l(q)$ has shortest law of length $\Omega(q^{\lfloor (l+1)/2 \rfloor})$.
- ${}^2A_1(q)$ has shortest law of length $\Omega(q)$ by Theorem A.1 (i). Suppose $l \geq 2$ and let $n = l + 1$. We claim that the shortest law for ${}^2A_l(q)$ has length $\Omega(q^{\lfloor n/2 \rfloor})$. If n is odd, then ${}^2A_{l-1}(q)$ is a subquotient of ${}^2A_l(q)$ [7, Table 2.3] and the result follows by induction. If n is divisible by 4, then $A_{n/2-1}(q^2)$ is a subquotient of ${}^2A_l(q)$ ([7, Table 2.3] again) and the conclusion follows. Otherwise write $n = ms$ for s an odd prime. Then ${}^2A_{m-1}(q^s)$ is a subquotient of ${}^2A_l(q)$ [7, Table 2.6]. Since m is even, we are done by induction.
- $C_1(q)$ has shortest law of length $\Omega(q)$ by Theorem A.1 (i). For $l \geq 2$, claim the shortest law for $C_l(q)$ has length $\Omega(q^l)$. Write $l = ms$, for s a prime. Then $C_m(q^s)$ is a subquotient of $C_l(q)$ [7, Table 2.6], so we are done by induction.
- By Theorem A.1 (iv) and (vii) and the previous paragraphs, ${}^2D_2(q)$ and ${}^2D_3(q)$ each have shortest law of length $\Omega(q^2)$. Assume $l \geq 4$; we claim the shortest law for ${}^2D_l(q)$ has length $\Omega(q^{2\lfloor l/2 \rfloor})$. If l is even, and $l/2 = ms$, for s a prime, then ${}^2D_{l/s}(q^s)$ is a subquotient of ${}^2D_l(q)$ [7, Table 2.6]. If l is odd, then ${}^2D_{l-1}(q)$ is a subquotient of ${}^2D_l(q)$ [7, Table 2.3]. In both cases we are done by induction.
- By Theorem A.1 (v) and the above, $B_2(q)$ has shortest law of length $\Omega(q^2)$. Suppose $l \geq 3$. By Theorem A.1 (viii) we need only consider $B_l(q)$ for q odd. If l is even, then ${}^2D_l(q)$ is a subquotient of $B_l(q)$, whereas if l is odd, then ${}^2D_{l-1}(q)$ is a subquotient of $B_l(q)$ [7, Table 2.3]. In both cases it follows from the previous paragraph that the shortest law for $B_l(q)$ has length $\Omega(q^{2\lfloor l/2 \rfloor})$.
- For $D_l(q)$, with $l \geq 4$, then ${}^2D_{l-1}(q)$ is a subquotient of $D_l(q)$. If l is odd, it follows that the shortest law for $D_l(q)$ has length $\Omega(q^{l-1})$, whereas if l is even, it has length $\Omega(q^{l-2})$. Moreover, if l and q are both even, then $C_{l-1}(q)$ is a subquotient of $D_l(q)$, whose shortest law therefore has length $\Omega(q^{l-1})$ [7, Table 2.3].
- $G_2(q)$ has a parabolic subgroup with Levi factor $\mathrm{GL}_2(q)$: this has shortest law of length $\Omega(q)$.

- ${}^2G_2(q)$ has a subgroup $\mathrm{PSL}_2(q)$: this has shortest law of length $\Omega(q)$; as is noted in [25, Theorem C], this subgroup is contained in an involution-centraliser, and exists for $q \geq 27$.
- ${}^3D_4(q)$ has a maximal subgroup with a subquotient $\mathrm{PSL}_2(q^3)$: this has shortest law of length $\Omega(q^3)$; as noted in [24], this is a maximal parabolic subgroup.
- $F_4(q)$ has a maximal subgroup with $\Omega_9(q)$ as a quotient: this has shortest law of length $\Omega(q^4)$; this is a subgroup of maximal rank [30, Table 5.1]. The subgroup $\mathrm{Sp}_4(q^2)$, occurring inside a maximal subgroup of maximal rank as noted above, also has shortest law of length $\Omega(q^4)$, but as is noted in [30], this only arises for q even.
- ${}^2F_4(q)$ has a subquotient $B_2(q)$, whose shortest law has length $\Omega(q^2)$ [34, Main Theorem]; note however the difference in notational convention: our q is denoted q^2 in [34].
- $E_6(q)$ has a parabolic subgroup, the Levi factor of which contains $\Omega_{10}^+(q)$, whose shortest law has length $\Omega(q^4)$.
- ${}^2E_6(q)$ has a parabolic subgroup, the Levi factor of which contains $\Omega_8^-(q)$, whose shortest law has length $\Omega(q^4)$.
- $E_7(q)$ has a maximal subgroup of maximal rank containing $\mathrm{PSL}_2(q^7)$ [30, Table 5.1], whose shortest law has length $\Omega(q^7)$.
- $E_8(q)$ has a maximal parabolic subgroup, the Levi factor of which contains $E_7(q)$, whose shortest law has length $\Omega(q^7)$.
- Finally, the shortest law in ${}^2B_2(q)$ has length $\Omega(q^{1/2})$ [10, Lemma 3.4].

This finishes the proof of Theorem 1.1. ∎

Appendix A. Background on groups of Lie type

The goal of this appendix is to provide background on finite simple groups and prove Proposition 2.14. We also collect various tables and case studies in order to make the main text more transparent. We start with a quick overview that recalls the classical theory.

A.1. Groups of Lie type

A.1.1. Isomorphisms in small rank or characteristic. We note some exceptional isomorphisms of low-dimensional classical groups. Items (i)–(vii) of Theorem A.1 below are recorded, for example, in [7, Proposition 1.10.1]. Item (viii) is discussed for instance in [7, Section 1.5.5].

Theorem A.1. *Let q be a prime power.*

(i) $\mathrm{PSL}_2(q) \cong \mathrm{PSp}_2(q) \cong \mathrm{PSU}_2(q) \cong P\Omega_3^\circ(q)$.

(ii) $P\Omega_2^\pm(q)$ *is abelian.*

(iii) $P\Omega_4^+(q) \cong \mathrm{PSL}_2(q) \times \mathrm{PSL}_2(q)$.

(iv) $P\Omega_4^-(q) \cong \mathrm{PSL}_2(q^2)$.

(v) $P\Omega_5^\circ(q) \cong \mathrm{PSp}_4(q)$.

(vi) $P\Omega_6^+(q) \cong \mathrm{PSL}_4(q)$.

(vii) $P\Omega_6^-(q) \cong \mathrm{PSU}_4(q)$.

(viii) *Suppose q is even. Then for all $n \geq 1$, $P\Omega_{2n+1}^\circ(q) \cong \mathrm{PSp}_{2n}(q)$.*

It follows that, in proving Theorem 1.1, we may assume the following restrictions on X and q hold:

(i) If $X = {}^2A_l$ or C_l then $l \geq 2$.

(ii) If $X = B_l$ then $l \geq 3$.

(iii) If $X = D_l$ or ${}^2D_l$ then $l \geq 4$.

(iv) If $X = B_l$ then q is odd.

A.1.2. Representations. We record the minimal dimension $n = n(X, p)$ of faithful projective representations of the simple group $G = X(q)$ over the algebraic closure of $\mathbb{F}_q$ (recalling that q is a power of the prime p: implicit in writing $n = n(X, p)$ is the claim that the dimension does not depend on the degree of the field extension $(\mathbb{F}_q : \mathbb{F}_p)$; we see below that this is indeed the case). In addition to providing context to the statement of Corollary 1.2, knowing $n(X, p)$ will be important in the proof of Theorem 1.1, in that it will provide a restriction on the possible embeddings of one quasisimple group of Lie type into another of matched characteristic as a non-geometric subgroup.

The various possible values for $n(X, p)$ are recorded in [23, Proposition 5.4.13]. We reproduce their conclusions in the next theorem. Note that every group in our list is isomorphic to a group in their list by Theorem A.1.

Theorem A.2. *Let $G = X(q)$ be as in Theorem 1.1. If $X = C_2$ then suppose $q \geq 3$. Let $n = n(X, p) \in \mathbb{N}$ be the minimal dimension of a faithful projective module for G over $\overline{\mathbb{F}}_q$ (in other words, let n be minimal such that G is isomorphic to a subgroup of $\mathrm{PGL}_n(\overline{\mathbb{F}}_q)$). Then n is as in Table 2 below.*

Remark A.3. Corollary 1.2 follows from Theorem 1.1, by inspection of Tables 1 and 2, since $a(X, p) \leq \lfloor n(X, p)/2 \rfloor$ in all cases.

A.1.3. Maximal element orders. We give upper bounds on the orders of elements in finite simple groups of Lie type. For this we refer to the tables from [39], which in turn are based on [20].

Proposition A.4. *Let $d(X)$ be as in Table 3. Then*

$$\max\{o(g) : g \in X(q)\} = O(q^{d(X)}).$$

X	A_l or ${}^2A_l$	B_l (p odd)	C_l	D_l	${}^2D_l$
n	$l+1$	2 ($l=1$)	$2l$	4 ($l=2$)	2 ($l=2$)
		4 ($l=2$)		4 ($l=3$)	4 ($l=3$)
		$2l+1$ ($l\geq 3$)		$2l$ ($l\geq 4$)	$2l$ ($l\geq 4$)

X	E_6 or ${}^2E_6$	E_7	E_8	F_4	G_2
n	27	56	248	25 ($p=3$)	7 (p odd)
				26 ($p\neq 3$)	6 ($p=2$)

X	${}^3D_4$	${}^2B_2$	${}^2F_4$	${}^2G_2$
n	8	4	26	7

Tab. 2. Minimal dimensions of projective representations of simple groups of Lie type.

X	A_l	${}^2A_l$	B_l	C_l	D_l	${}^2D_l$
$d(X)$	l	l	l	l	l	l

X	E_6	${}^2E_6$	E_7	E_8	F_4	G_2	${}^3D_4$	${}^2B_2$	${}^2F_4$	${}^2G_2$
$d(X)$	6	6	7	8	4	2	4	1	2	1

Tab. 3. Degree of length of laws coming from maximal element orders.

A.2. Automorphisms of finite groups

If G is a non-abelian finite simple group, then $Z(G)=1$, so we naturally have a short exact sequence $1\to G\to \mathrm{Aut}(G)\to \mathrm{Out}(G)\to 1$. The following is widely known and classical.

Theorem A.5. *Let G be either a finite simple group of Lie type or* $\mathrm{Alt}(n)$ *for $n\geq 5$. Then* $\mathrm{Out}(G)$ *is soluble of derived length at most* 3.

The famous Schreier conjecture extends the conclusion of Theorem A.5 to all non-abelian finite simple groups, but this is known only as a consequence of CFSG, and we shall not need it.

The structure of automorphism groups of direct products of non-abelian finite simple groups is folklore.

Proposition A.6. *Let $G_1,\dots,G_m$ be non-abelian finite simple groups. Then there exists $H\leq \mathrm{Sym}(m)$ such that*

$$\mathrm{Out}(G_1\times\cdots\times G_m)\leq (\mathrm{Out}(G_1)\times\cdots\times\mathrm{Out}(G_m)).H.$$

For a general finite group G, let $F^*(G)$ be the *generalized Fitting subgroup*.

Proposition A.7. $C_G(F^*(G))\leq F^*(G)$. *In particular, if $F^*(G)$ is a product of non-abelian finite simple groups, then G embeds as a subgroup of* $\mathrm{Aut}(F^*(G))$.

A.3. Algebraic groups and Chevalley groups

We recall some concepts from the theory of algebraic groups, well-covered in any standard text on the subject such as [35].

Let $\mathbb{G}$ be a connected semisimple linear algebraic group of rank l over an algebraically closed field K (so that $\mathbb{G}$ is isomorphic to a closed subgroup of $\mathrm{GL}_d(K)$). An element of $\mathbb{G}$ is *semisimple* if it is diagonalizable. A *torus* in $\mathbb{G}$ is a connected subgroup consisting of simultaneously diagonalizable elements; a subgroup of $\mathbb{G}$ is a torus iff it is isomorphic to $(K^*)^m$ for some $m \in \mathbb{N}$. A *maximal torus* in $\mathbb{G}$ is a torus of maximal dimension; $\mathbb{G}$ contains at least one maximal torus, all maximal tori are conjugate in $\mathbb{G}$, and every torus is contained in at least one maximal torus. The rank l of $\mathbb{G}$ is by definition the common dimension of the maximal tori in $\mathbb{G}$; hence a torus in $\mathbb{G}$ is maximal iff it is isomorphic to $(K^*)^l$.

Let $\mathfrak{g}$ be the Lie algebra of $\mathbb{G}$ and let T be a maximal torus of $\mathbb{G}$. Let $X(T) = \mathrm{Hom}(T, K^*)$ be the *character group* of T. Then $X(T)$ is a group under pointwise multiplication; since $T \cong (K^*)^l$, $X(T) \cong \mathbb{Z}^l$. For $\alpha \in X(T)$, let

$$\mathfrak{u}_\alpha = \{x \in \mathfrak{g} : \forall t \in T,\ (\mathrm{Ad}(t))(x) = \alpha(t)x\}.$$

Let $\Phi(T, \mathbb{G}) = \{\alpha \in X(T) : \mathfrak{u}_\alpha \neq 0\}$ be the set of *roots of* $\mathbb{G}$ *relative to* T. Embedding $X(T)$ into $X(T) \otimes_{\mathbb{Z}} \mathbb{R}$, identified with Euclidean l-space, $\Phi(T, \mathbb{G})$ is an abstract root system (see [35, Sections 8.1, 9.1]).

For each $\alpha \in \Phi(T, \mathbb{G})$, there is a unique closed connected one-dimensional subgroup U_α of $\mathbb{G}$ with Lie algebra $\mathfrak{u}_\alpha$, normalized by T (the *root subgroup* corresponding to α). There is an isomorphism $\lambda \mapsto x_\alpha(\lambda)$ from $(K, +)$ to U_α satisfying

$$t x_\alpha(\lambda) t^{-1} = x_\alpha(\alpha(t)\lambda) \tag{A.1}$$

for all $t \in T$ [35, Theorem 8.17].

It is clear that every semisimple s element of $\mathbb{G}$ is contained in at least one maximal torus. The element s is called *regular* if it lies in a unique maximal torus.

Proposition A.8 ([35, Corollary 14.10]). *For $s \in \mathbb{G}$ semisimple, and T a maximal torus in $\mathbb{G}$ containing s, the following conditions are equivalent:*

(i) *s is regular;*

(ii) *for every root α of $\mathbb{G}$ relative to T, $\alpha(s) \neq 1$.*

Proposition A.9. *Let $s \in \mathbb{G}$ be regular semisimple, let T be the unique maximal torus of $\mathbb{G}$ containing s and let $g \in \mathbb{G}$ be such that $s^g \in T$. Then $g \in N_{\mathbb{G}}(T)$. In particular, $C_{\mathbb{G}}(s) \subseteq N_{\mathbb{G}}(T)$.*

Proof. T^g is also a maximal torus of $\mathbb{G}$ and s^g is regular semisimple. By uniqueness of T^g, $T = T^g$. ∎

Theorem A.10. *For every maximal torus T of $\mathbb{G}$, $N_{\mathbb{G}}(T)/T$ is isomorphic to the Weyl group $W_{\mathbb{G}}$ of $\mathbb{G}$.*

Proof. By [35, Corollary 6.5], all maximal tori are conjugate in $\mathbb{G}$, so the isomorphism type of $N_{\mathbb{G}}(T)/T$ is independent of T. ∎

We recall the construction of the Chevalley groups given in [11] and use this to give an explicit linear representation of the untwisted finite simple groups of Lie type. Let $K = \overline{\mathbb{F}}_p$ be the algebraic closure of the finite field with p elements, and let Φ be an irreducible root system of rank l, with fundamental system of roots Π. Let $\mathcal{L}_K$ be the K-Lie algebra defined in [11, Section 4.4]; $\mathcal{L}_K$ has a *Chevalley basis* $\{h_\alpha : \alpha \in \Pi\} \cup \{e_\beta : \beta \in \Phi\}$, with respect to which the structure constants are as in [11, Theorem 4.2.1].

Let $\mathcal{L}(K) \leq \mathrm{GL}_{|\Phi|+|\Pi|}(K)$ be the group generated by the automorphisms $\{x_\beta(\lambda) : \beta \in \Phi, \lambda \in K\}$ of the Lie algebra $\mathcal{L}_K$ (the $x_\beta(\lambda)$ are defined in [11, Section 4.4]). Then $\mathcal{L}$ is a simple adjoint algebraic group of rank l and $\mathcal{L}_K$ is its Lie algebra.

Definition A.11. For q a power of p, the (untwisted) *finite simple group of type* Φ over the field $\mathbb{F}_q$ is the subgroup $\mathcal{L}(\mathbb{F}_q)$ of $\mathcal{L}(K)$ generated by $\{x_\beta(\lambda) : \beta \in \Phi, \lambda \in \mathbb{F}_q\}$.

We describe an explicit maximal torus in $\mathcal{L}(K)$. By [11, Proposition 6.4.1], for $\alpha \in \Pi$ there is a homomorphism $\phi_\alpha : K^* \to \mathcal{L}(K)$ given by

$$\phi_\alpha(z)h_\gamma = h_\gamma \text{ for } \gamma \in \Pi, \quad \phi_\alpha(z)e_\beta = z^{A_{\alpha,\beta}}e_\beta \text{ for } \beta \in \Phi, \tag{A.2}$$

where, for $\alpha, \beta \in \Phi$, $A_{\alpha,\beta} = 2(\alpha,\beta)/(\alpha,\alpha) \in \mathbb{Z}$ is the *Cartan integer*. By [11, Lemma 6.4.4], $\phi_\alpha(z) \in \mathcal{L}(K)$ may be explicitly realized as

$$\phi_\alpha(z) = x_\alpha(z)x_{-\alpha}(-z^{-1})x_\alpha(z)x_\alpha(-1)x_{-\alpha}(1)x_\alpha(-1). \tag{A.3}$$

By (A.2), for any $z \in K^*$, $\phi_\alpha(z)$ is a diagonal matrix and we have a homomorphism $\phi : (K^*)^\Pi \to \mathcal{L}(K)$ defined, for $\underline{z} = (z_\alpha)_{\alpha\in\Pi} \in (K^*)^\Pi$, by

$$\phi(\underline{z}) = \prod_{\alpha\in\Pi} \phi_\alpha(z_\alpha)$$

so that

$$\phi(\underline{z})h_\gamma = h_\gamma \text{ for } \gamma \in \Pi, \quad \phi(\underline{z})e_\beta = \phi(\underline{z})_\beta e_\beta \text{ for } \beta \in \Phi,$$

where we write

$$\phi(\underline{z})_\beta = \prod_{\alpha\in\Pi} z_\alpha^{A_{\alpha,\beta}} \tag{A.4}$$

for $\beta \in \Phi$. From this description, it is clear that ϕ is a morphism of algebraic groups.

Lemma A.12. *The morphism ϕ has finite kernel.*

Proof. The Cartan matrix $A = (A_{\alpha,\beta})_{\alpha,\beta\in\Pi}$ is non-singular (this is clear from inspection of the tables in [11, Section 3.6]; alternatively, an inverse for A in $\mathrm{GL}_\Pi(\mathbb{Q})$ is calculated explicitly in [38, Section 1.6, (1.157), (1.158)]). Putting A into Smith normal form, there exist $P, Q \in \mathrm{GL}_\Pi(\mathbb{Z})$ and $0 \neq D_\alpha \in \mathbb{Z}$ such that

$$(PAQ)_{\alpha,\beta} = \begin{cases} D_\alpha, & \alpha = \beta, \\ 0, & \alpha \neq \beta. \end{cases}$$

For $B \in \mathrm{GL}_\Pi(\mathbb{Z})$, define the morphism $\psi^{(B)} : (K^*)^\Pi \to (K^*)^\Pi$ by

$$\psi^{(B)}(\underline{z})_\alpha = \prod_{\beta \in \Pi} z_\beta^{B_{\beta,\alpha}},$$

so that for $B, C \in \mathrm{GL}_\Pi(\mathbb{Z})$, $\psi^{(BC)} = \psi^{(C)} \circ \psi^{(B)}$ and $\psi^{(I)} = \mathrm{Id}$. Then

$$(\psi^{(Q)} \circ \psi^{(A)} \circ \psi^{(P)})(\underline{z})_\alpha = z_\alpha^{D_\alpha} \quad \text{and} \quad \psi^{(A)}(\underline{z})_\alpha = \phi(\underline{z})_\alpha, \tag{A.5}$$

so if $\psi^{(P)}(\underline{z}) \in \ker(\phi)$, then z_α is a D_αth root of unity, for all $\alpha \in \Pi$. However $\psi^{(P)}$ is invertible, with inverse $\psi^{(P^{-1})}$, so $|\ker(\phi)| \leq \prod_{\alpha \in \Pi} D_\alpha$. ∎

Proposition A.13. *The image $T_0 = \mathrm{im}(\phi)$ of ϕ is a maximal torus in $\mathcal{L}(K)$.*

Proof. $\mathrm{im}(\phi)$ is a closed connected abelian subgroup of $\mathcal{L}(K)$ (being the image of such under a morphism of algebraic groups). It consists of diagonal matrices, hence is a torus, so is contained in a maximal torus T (of dimension l, the rank of $\mathcal{L}(K)$). Since $\ker(\phi)$ is finite, $\mathrm{im}(\phi) \subseteq T$ also has dimension l, hence they are equal. ∎

Consider $T_0(q) = T_0 \cap \mathcal{L}(\mathbb{F}_q)$. By (A.3), $T_0(q)$ contains $\phi(\underline{z})$ for every $\underline{z} \in (\mathbb{F}_q^*)^\Pi$, so $|T_0(q)| \geq (q-1)^l/|\ker(\phi)|$. In the other direction, using the notation of Lemma A.12, suppose $\underline{z} \in (K^*)^\Pi$ with $\phi(\underline{z}) \in T_0(q)$. Writing $\underline{z} = \psi^{(P)}(\underline{y})$, by (A.5) we have $y_\alpha^{D_\alpha} \in \mathbb{F}_q^*$ for all $\alpha \in \Pi$, hence

$$(q-1)^l/|\ker(\phi)| \leq |T_0(q)| \leq (q-1)^l \prod_{\alpha \in \Pi} D_\alpha. \tag{A.6}$$

Now, the roots of $\mathcal{L}(K)$ relative to $T_0 = \mathrm{im}(\phi)$ are indexed by the elements of Φ: for each $\beta \in \Phi$ there is a morphism $T_0 \to K^*$ (which we also denote by β) given by

$$\beta(\phi(\underline{z})) = \phi(\underline{z})_\beta$$

(see (A.4)). Thus $\langle e_\beta \rangle \subseteq \mathfrak{u}_\beta$, and since the maps $\beta : T_0 \to K^*$ are all distinct, non-trivial, and T_0 acts trivially on $\langle h_\alpha : \alpha \in \Pi \rangle$, we have equality. The root subgroup corresponding to β is precisely $\{x_\beta(\lambda) : \lambda \in K\}$; in accordance with (A.1),

$$\phi(\underline{z}) x_\beta(\lambda) \phi(\underline{z})^{-1} = x_\beta(\beta(\phi(\underline{z}))\lambda)$$

(see [11, Section 7.1]).

Proposition A.14. *Let $R(q) = \{s \in T_0(q) : s \text{ is regular in } \mathcal{L}(K)\}$. Then*

$$|T_0(q) \setminus R(q)| = O_\Phi(|T_0(q)|/q). \tag{A.7}$$

Proof. Since Φ lies in the $\mathbb{R}$-span of Π, for every $\beta \in \Phi$ there exists $\alpha \in \Pi$ such that $A_{\alpha,\beta} \neq 0$. Then for $z \in \mathbb{F}_q^*$, $\beta(\phi((z^{\delta_{\alpha,\gamma}})_{\gamma \in \Pi})) = z^{A_{\alpha,\beta}}$. Letting $N(\beta) = \mathrm{hcf}(\{A_{\alpha,\beta} : \alpha \in \Pi\} \setminus \{0\})$, we have $|\mathrm{im}(\beta|_{T_0(q)})| \geq (q-1)/N(\beta)$. By (A.6),

$$|T_0(q) \cap \ker(\beta)| \leq (q-1)^{l-1} D N(\beta),$$

where $D = \prod_{\alpha\in\Pi} D_\alpha$. Using the criterion for regularity from Proposition A.8 (ii), and (A.6), we have

$$|T_0(q) \setminus R(q)| \le (q-1)^{l-1} D \sum_{\beta\in\Phi} N(\beta) \quad \text{while} \quad |T_0(q)| \ge (q-1)^l/|\ker(\phi)|$$

as desired. ∎

A.4. Proof of Proposition 2.14

Let us recall the result, which it is our objective to prove here.

Proposition A.15. *For G a finite group and $m \in \mathbb{N}$, let*

$$E_G(m) = \{g \in G : o(g) \text{ divides } m\}.$$

Let $G = X(q)$ be as in Theorem 1.1. Define $b(X,q) \in \mathbb{N}$ to be $q+1$ for $X = {}^2A_l$ or ${}^2E_6$; q^2-1 for $X = {}^2D_l$; q^2-q+1 for $X = {}^3D_4$; and $q-1$ otherwise. Then

$$|E_G(b(X,q))| = \Omega_X(|G|).$$

Lemma A.16. *Let $\Delta_n(q) \le \mathrm{GL}_n(q)$ be the subgroup of diagonal matrices. Let $g \in \Delta_n(q)$ and suppose g has no repeated eigenvalues. Then $C_{\mathrm{GL}_n(q)}(g) = \Delta_n(q)$. More generally, let $g \in \mathrm{GL}_n(q)$ have the form*

$$g = \begin{pmatrix} I_m & 0 \\ 0 & h \end{pmatrix}$$

for some $h \in \Delta_{n-m}(q)$ with no 1-eigenvectors and no repeated eigenvalues. Then any $c \in C_{\mathrm{GL}_n(q)}(g)$ has the form

$$c = \begin{pmatrix} k & 0 \\ 0 & d \end{pmatrix}$$

for some $k \in \mathrm{GL}_m(q)$ and $d \in \Delta_{n-m}(q)$.

Lemma A.17. *Let G be a finite group and let $N \lhd G$. Then for any $m \in \mathbb{N}$,*

$$|E_{G/N}(m)|/|G/N| \ge |E_G(m)|/|G|.$$

Proof. If $g \in E_G(m)$, then $gN \in E_{G/N}(m)$, so $E_{G/N}(m)$ is the union of all cosets of N intersecting $E_G(m)$. Thus $|E_{G/N}(m)| \ge |E_G(m)|/|N|$. ∎

Lemma A.18. *Let $l, m \in \mathbb{N}$; for $1 \le i \le l$ and $1 \le j \le m$ let $a_i^{(j)} \in \mathbb{Z}$, $\lambda^{(j)} \in \mathbb{F}_q^*$ and*

$$f^{(j)}(X_1,\dots,X_l) = \lambda^{(j)} X_1^{a_1^{(j)}} \cdots X_l^{a_l^{(j)}} \in \mathbb{F}_q[X_1^{\pm1},\dots,X_l^{\pm1}].$$

Suppose that the $f^{(j)}$ are all distinct, and none is identically 1. Let $S \subseteq \mathbb{F}_q$ and $D > 0$, and suppose that for all i and j, $|a_i^{(j)}| \le D$. Then there is a subset $A \subseteq S^l$ such that

(i) $|A| \ge (|S| - Dm(m+1))^l$;

(ii) *for all $\mathbf{a} \in A$, the $f^{(j)}(\mathbf{a}) \in \mathbb{F}_q$ are distinct and $\ne 1$.*

Proof. We proceed by induction on l. If $l = 1$ then for $s \in S$, if $f^{(j)}(s) = f^{(k)}(s)$ or $f^{(j)}(s) = 1$ for some $j \neq k$, then s satisfies one of a set of $m(m+1)/2$ non-trivial polynomials of degree at most $2D$, so there are at most $Dm(m+1)$ possibilities for s.

For $l \geq 2$, for at least $|S| - Dm(m+1)$ values $s_l \in S$, for any $1 \leq j,k \leq m$, $\lambda^{(j)} s_l^{a_l^{(j)}} = \lambda^{(k)} s_l^{a_l^{(k)}}$ implies $\lambda^{(j)} = \lambda^{(k)}$ and $a_l^{(j)} = a_l^{(k)}$ (by the preceding paragraph). Given such a value s_l, the Laurent polynomials

$$f^{(j)}(X_1, \ldots, X_{l-1}, s_l) \in \mathbb{F}_q[X_1^{\pm 1}, \ldots, X_{l-1}^{\pm 1}]$$

are all distinct. Likewise for such values s_l, $\lambda^{(j)} s_l^{a_l^{(j)}} = 1$ implies $\lambda^{(j)} = 1$ and $a_l^{(j)} = 0$, so for such values s_l, no $f^{(j)}(X_1, \ldots, X_{l-1}, s_l)$ is identically 1. ∎

Remark A.19. In all cases in which we apply Lemma A.18, either (i) $|S| = \Omega(q)$ and $D = O(\sqrt{q})$, or (ii) $|S| = \Omega(\sqrt{q})$ and D is bounded. Under either of these conditions, it follows that for all $\varepsilon > 0$, if $q = \Omega_{\varepsilon,l,m}(1)$ then

$$|A| \geq (1-\varepsilon)|S|^l. \tag{A.8}$$

In particular, $|A| = \Omega_{l,m}(|S|^l)$.

A.4.1. Untwisted groups. Here $G = \mathcal{L}(\mathbb{F}_q)$ is as in Definition A.11. Let T_0 be as in Proposition A.13; let $T_0(q) = T_0 \cap G$, and let $R(q) \subseteq T_0(q)$ be the set of regular elements. Note that any conjugate in G of an element of $T_0(q)$ has order dividing $q-1$, so the conclusion of Proposition 2.14 reduces to the following.

Proposition A.20. *Let G, $R(q)$ be as above. Then*

$$\Big| \bigcup_{r \in R(q)} \mathrm{ccl}_G(r) \Big| = \Omega_\Phi(|G|). \tag{A.9}$$

Proof. If $r_1, r_2 \in R(q)$ and $g_1, g_2 \in G$ are such that $r_1^{g_1} = r_2^{g_2}$, then by Proposition A.9 we have $g_1 g_2^{-1} \in N_G(T_0)$ so the fibres of the map sending $(r,g) \in R(q) \times G$ to r^g are of size bounded by $|N_G(T_0)| \leq |W| \cdot |T_0(q)|$ (by Theorem A.10). Thus

$$\Big| \bigcup_{r \in R(q)} \mathrm{ccl}_G(r) \Big| \geq \frac{|R(q)|}{|W| \cdot |T_0(q)|} |G|$$

and the result follows, since by Proposition A.14, $|R(q)|/|T_0(q)| = \Omega_\Phi(1)$. ∎

Remark A.21. The bound (A.9) establishes Proposition 2.14 for the groups $X(q)$ with $X = A_l, B_l, C_l, D_l, E_6, E_7, E_8, F_4$ or G_2. The basic strategy for the other families will be the same: in each case we produce $R(q) \subseteq G = X(q)$ such that (a) for all $r \in R(q)$, $o(r)$ divides $b(X,q)$; (b) the fibres of the map sending $(r,g) \in R(q) \times G$ to r^g are of size $O_X(|R(q)|)$. Note in particular that (b) holds whenever, for all $r \in R(q)$, we have (b)(i) $|\mathrm{ccl}_G(r) \cap R(q)| = O_X(1)$, and (b)(ii) $|C_G(r)| = O_X(|R(q)|)$. As above we have a bound of the form (A.9), from which the result follows.

A.4.2. Unitary groups. For ${}^2A_l(q) = \mathrm{PSU}_{l+1}(q)$, using Lemma A.17 it suffices to show that $G = \mathrm{SU}_{l+1}(q)$ satisfies $|E_G(q+1)| = \Omega_l(|G|)$. Recall that $\mathrm{GU}_{l+1}(q)$ is the subgroup of $\mathrm{GL}_{l+1}(q^2)$ preserving a non-degenerate Hermitian form on $\mathbb{F}_{q^2}^{l+1}$. Up to similarity, there is only one such form, given by I_{l+1}. Given $\underline{\lambda} \in \mathbb{F}_{q^2}^{l+1}$, $\mathrm{diag}(\underline{\lambda}) \in \mathrm{GU}_{l+1}(q)$ iff $\lambda_1, \ldots, \lambda_{l+1} \in S$, the set of $(q+1)$th roots of unity in $\mathbb{F}_{q^2}$. Indeed, $\mathrm{diag}(\underline{\lambda}) \in \mathrm{GU}_{l+1}(q)$ iff

$$I_{l+1} = \phi(\mathrm{diag}(\underline{\lambda}))^t \,\mathrm{diag}(\underline{\lambda}) = \mathrm{diag}(\underline{\lambda})^{q+1},$$

where $\phi : \mathrm{GL}_{l+1}(q^2) \to \mathrm{GL}_{l+1}(q^2)$ is the Frobenius automorphism given by $\phi(g)_{i,j} = g_{i,j}^q$.

Thus $\mathrm{diag}(\underline{\lambda}) \in \mathrm{SU}_{l+1}(q)$ iff $\lambda_1, \ldots, \lambda_l \in S$ and $\lambda_{l+1} = (\lambda_1 \cdots \lambda_l)^{-1}$. We apply Lemma A.18 with $m = l+1$, S as above, $D = 1$, and

$$f^{(1)} = X_1, \ \ldots, \ f^{(l)} = X_l, \quad f^{(l+1)} = (X_1 \cdots X_l)^{-1}.$$

Using Remark A.19, we see that the set $R(q) \subseteq \mathrm{SU}_{l+1}(q)$ of diagonal matrices without repeated eigenvalues has size $\Omega_l((q+1)^l)$. We check the conditions of Remark A.21. Certainly any element of $R(q)$ has order dividing $q+1$. Since any two conjugate matrices have the same eigenvalues, we have (b)(i), and by Lemma A.16 the centraliser of any $r \in R(q)$ consists of diagonal elements of $\mathrm{SU}_{l+1}(q)$ (of which there are $(q+1)^l$), whence (b)(ii) holds. Finally, we pass from $\mathrm{SU}_{l+1}(q)$ to $\mathrm{PSU}_{l+1}(q)$ using Lemma A.17.

A.4.3. Orthogonal groups of type ${}^2D_l$. Note that ${}^2D_l(q) = P\Omega_n^-(q)$, where $n = 2l$. By Lemma A.17 it suffices to show that $G = \Omega_n^-(q)$ satisfies $|E_G(q^2-1)| = \Omega_n(|G|)$. We start by recalling some background information on the forms preserved by these groups, taken from [11]. If q is odd then the orthogonal group $\mathrm{O}_n^-(q)$ is the group of isometries of the symmetric bilinear form B^- on $\mathbb{F}_q^n$, represented by

$$\begin{pmatrix} 0 & I_{l-1} & 0 & 0 \\ I_{l-1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\gamma \end{pmatrix} \in \mathbb{M}_n(\mathbb{F}_q)$$

with respect to the standard ordered basis $\mathcal{B} = \{b_i\}_{i=1}^n$ of $\mathbb{F}_q^n$, where $\gamma \in \mathbb{F}_q$ is a non-square. If on the other hand q is even then $\mathrm{O}_n^-(q)$ is the group of isometries of the quadratic form on $\mathbb{F}_q^n$, represented by

$$f^-(\underline{x}) = \sum_{i=1}^{l-1} x_i x_{l-1+i} + (x_{2l-1} - \alpha x_{2l})(x_{2l-1} - \bar{\alpha} x_{2l}),$$

where α generates $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$ (note that f^- is nonetheless defined over $\mathbb{F}_q$). In this case, let B^- be the *polar form* of a quadratic form f^-, given by

$$B^-(\underline{x}, \underline{y}) = f^-(\underline{x} + \underline{y}) - f^-(\underline{x}) - f^-(\underline{y}).$$

Then B^- is a symmetric bilinear form preserved by $\mathrm{O}_n^-(q)$; its matrix with respect to $\mathcal{B}$ is

$$\begin{pmatrix} 0 & I_{l-1} & 0 & 0 \\ I_{l-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta \\ 0 & 0 & \beta & 0 \end{pmatrix} \in \mathbb{M}_n(\mathbb{F}_q),$$

where $\beta = -(\alpha + \overline{\alpha}) \in \mathbb{F}_q$. In all cases, $\Omega_n^-(q)$ is a subgroup of index 2 in $\mathrm{SO}_n^-(q)$; see for instance [7, Definition 1.6.13] for a precise description of $\Omega_n^-(q)$ as a subgroup of $\mathrm{SO}_n^-(q)$.

The group $\mathrm{SO}_2^-(q)$ is cyclic of order $q+1$; let $\zeta \in \mathrm{SO}_2^-(q)$ be a generator. Since $\mathrm{hcf}(q+1, q-1) \leq 2$, ζ^m is not diagonalizable over $\mathbb{F}_q$ for any $1 \leq m \leq q$ with $m \neq (q+1)/2$.

Now let $D(q) \leq \mathrm{SO}_{2l}^-(q)$ be the subgroup of elements of the form

$$\begin{pmatrix} d(\underline{\lambda}) & 0 \\ 0 & \zeta^m \end{pmatrix}, \quad \text{where} \quad d(\underline{\lambda}) = \mathrm{diag}(\lambda_1, \ldots, \lambda_{l-1}, \lambda_1^{-1}, \ldots, \lambda_{l-1}^{-1}) \in \Delta_{n-2}(q)$$

for some $\underline{\lambda} = (\lambda_1, \ldots, \lambda_{l-1}) \in (\mathbb{F}_q^*)^{l-1}$, and let $R(q) \subseteq D(q)$ be the set of elements such that $1 \neq \lambda_1, \ldots, \lambda_{l-1}, \lambda_1^{-1}, \ldots, \lambda_{l-1}^{-1}$ are all distinct and $1 \leq m \leq q$ with $m \neq (q+1)/2$. Applying Lemma A.18 to $X_1^{\pm 1}, \ldots, X_{l-1}^{\pm 1}$, we see from (A.8) that $|R(q)| \geq 3|D(q)|/4$ for q larger than an (explicit) absolute constant, so $|R(q) \cap \Omega_n^-(q)| \geq |D(q) \cap \Omega_n^-(q)|/2$. Let

$$r_1 = \begin{pmatrix} d(\underline{\lambda}) & 0 \\ 0 & \zeta^m \end{pmatrix}, r_2 = \begin{pmatrix} d(\underline{\mu}) & 0 \\ 0 & \zeta^k \end{pmatrix} \in R(q) \cap \Omega_n^-(q)$$

and suppose r_1 and r_2 are conjugate in $\Omega_n^-(q)$. The eigenvalues of ζ in $\mathbb{F}_{q^2}^*$ are $\omega^{\pm 1}$ for some $\omega \in \mathbb{F}_{q^2}^*$ of order $q+1$. Since $k, m \neq (q+1)/2$, $\omega^{\pm m}$ and $\omega^{\pm k}$ do not lie in $\mathbb{F}_q$, hence $\{\omega^{\pm m}\} = \{\omega^{\pm k}\}$, so given m there are only two possibilities for k. The (common) eigenvalues of r_1 and r_2 in $\mathbb{F}_q$ are the (common) entries of $\underline{\lambda}$ and $\underline{\mu}$. Thus for fixed $\underline{\lambda}$, there are at most $2^{l-1}(l-1)!$ possibilities for $\underline{\mu}$. Thus $|\mathrm{ccl}_{\Omega_n^-(q)}(r_1) \cap R(q)| \leq 2^l(l-1)!$.

Now suppose $r \in R(q) \cap \Omega_n^-(q)$ and let $g \in C_{\mathrm{SO}_n^-(q)}(r)$. Then g preserves each eigenspace $\langle b_i \rangle$ for $1 \leq i \leq n-2$, hence preserves the orthogonal complement of $W = \langle b_1, \ldots, b_{n-2} \rangle$ under B^-, namely $U = \langle b_{n-1}, b_n \rangle$. The fact that g preserves B^- (and preserves $f^-|_U$ in the case of q even) implies that g has the form

$$\begin{pmatrix} d(\underline{\mu}) & 0 \\ 0 & h \end{pmatrix} \quad \text{for some } \underline{\mu} \in (\mathbb{F}_q^*)^{l-1} \text{ and } h \in \mathrm{SO}_2^-(q),$$

so that $|C_{\Omega_n^-(q)}(r)| \leq |C_{\mathrm{SO}_n^-(q)}(r)| \leq 2|D(q)| \leq 8|R(q) \cap \Omega_n^-(q)|$. We have satisfied the conditions of Remark A.21.

A.4.4. Suzuki groups. For the groups ${}^2B_2(q)$ we use the 4-dimensional linear representation over $\mathbb{F}_q$ described in [43, Section 4.2]. Here $q = 2^{2n+1}$. The form of the elements of ${}^2B_2(q) \cap \Delta_4(q)$ is described in [43, (4.10), p. 115]; they are

$$\{\mathrm{diag}(\alpha, \alpha^{2^{n+1}-1}, \alpha^{-2^{n+1}+1}, \alpha^{-1}) : \alpha \in \mathbb{F}_q^*\}$$

(so that $|{}^2B_2(q) \cap \Delta_4(q)| = q - 1$). Applying Lemma A.18 with $l = 1, m = 4, S = \mathbb{F}_q^*$, $D = 4\sqrt{q}$, and

$$f^{(1)}(X_1) = X_1,\ f^{(2)}(X_1) = X_1^{2^{n+1}-1},\ f^{(3)}(X_1) = X_1^{-2^{n+1}+1},\ f^{(4)}(X_1) = X_1^{-1},$$

we conclude that there exists $R(q) \subseteq {}^2B_2(q) \cap \Delta_4(q)$ consisting of elements without repeated eigenvalues and $|R(q)| \geq q/2 \geq |{}^2B_2(q) \cap \Delta_4(q)|/2$. As before, any element of $R(q)$ has order dividing $q - 1$. Given $r \in R(q)$, we have

$$|\mathrm{ccl}_{{}^2B_2(q)}(r) \cap R(q)| \leq |\mathrm{ccl}_{{}^2B_2(q)}(r) \cap \Delta_4(q)| \leq 24,$$

since any two conjugate elements have the same eigenvalues, and by Lemma A.16, $C_{{}^2B_2(q)}(r) \leq {}^2B_2(q) \cap \Delta_4(q)$. Thus $R(q)$ satisfies Remark A.21.

A.4.5. Small Ree groups. The argument for ${}^2G_2(q)$ is essentially identical (with $q = 3^{2n+1}$, using the 7-dimensional linear representation explained in [43, Section 4.5]; see in particular the description of the diagonal matrices lying in ${}^2G_2(q)$, given in [43, (4.53), p. 137]).

A.4.6. Large Ree groups. For ${}^2F_4(q)$ we use the model described in [43, Section 4.9]. First consider the group $F_4(q)$. As described in [43, Section 4.8], this group has a faithful 27-dimensional (reducible) linear representation on a non-associative $\mathbb{F}_q$-algebra $\mathcal{A}$. There is a preferred $\mathbb{F}_q$-basis $\mathcal{W} = \{w_i, w_i', w_i''\}_{i=0}^8$ for $\mathcal{A}$ defined in [43, Section 4.8.4]. Let $\circ$ denote the product on $\mathcal{A}$. Then $\circ$ is commutative and satisfies the following relations (among others):

$$\begin{array}{lll} w_0 \circ w_1 = 0; & w_0 \circ w_1' = w_1'; & w_0 \circ w_1'' = w_1''; \\ w_0' \circ w_1 = w_1; & w_0' \circ w_1' = 0; & w_0' \circ w_1'' = w_1''; \\ w_0'' \circ w_1 = w_1; & w_0'' \circ w_1' = w_1''; & w_0'' \circ w_1'' = 0; \\ w_1 \circ w_8 = w_0' + w_0''; & w_1' \circ w_8' = w_0 + w_0''; & w_1'' \circ w_8'' = w_0 + w_0'. \end{array} \tag{A.10}$$

Moreover, $\circ$ is preserved by the action of $F_4(q)$.

Let $q = 2^{2n+1}$; identify $F_4(q)$ with its image in $\mathrm{GL}(\mathcal{A})$ and embed ${}^2F_4(q)$ as a subgroup of $F_4(q)$ as in [43, Section 4.9.1]. The elements of ${}^2F_4(q)$ which are diagonal with respect to the basis $\mathcal{W}$ may be parametrised as $\{g_{(\alpha,\beta)}\}_{\alpha,\beta\in\mathbb{F}_q^*}$, where the eigenvalues of $g_{(\alpha,\beta)}$ on the basis $\mathcal{W}$ are as given in [43, Section 4.9.2]. The vectors w_0, w_0', w_0'' (called w_9, w_9', w_9'' in [43, Section 4.9.2], in contrast to the notation in [43, Section 4.8]) are 1-eigenvectors of $g_{(\alpha,\beta)}$, and there are Laurent polynomials $f^{(1)}, \ldots, f^{(24)} \in \mathbb{F}_q[X_1^{\pm 1}, X_2^{\pm 1}]$ such that the remaining elements of $\mathcal{W}$ are eigenvectors of $g_{(\alpha,\beta)}$ with eigenvalues $f^{(1)}(\alpha,\beta), \ldots, f^{(24)}(\alpha,\beta)$ [43, table in Section 4.9.2]. Moreover, the family $1, f^{(1)}, \ldots, f^{(24)}$ satisfies Lemma A.18 (with $l = 2$, $m = 24$, $S = \mathbb{F}_q^*$ and $D = 4\sqrt{q}$). By Lemma A.18 and Remark A.19, there exists a set $R(q) \subseteq {}^2F_4(q)$, with $|R(q)| \geq q^2/2$ for q larger than an explicit absolute constant, such that $R(q)$ consists of diagonal matrices (with respect to the basis $\mathcal{W}$) of the form

$$g = \begin{pmatrix} I_3 & 0 \\ 0 & h \end{pmatrix}$$

for some $h \in \Delta_{24}(q)$ with no 1-eigenvectors and no repeated eigenvalues. The set $R(q)$ satisfies conditions (a) and (b)(i) of Remark A.21 as in the previous cases. Moreover, by Lemma A.16 any $c \in C_{(^2F_4(q))}(g)$ has the form

$$c = \begin{pmatrix} k & 0 \\ 0 & d \end{pmatrix}$$

for some $k \in \mathrm{GL}_3(q)$ and $d \in \Delta_{24}(q)$. The fibres of the projection of $C_{(^2F_4(q))}(g)$ onto the top-left 3×3 block have order bounded by $|(\mathbb{F}_q^*)^2| = O(q^2)$, so to verify condition (b)(ii) of Remark A.21, it suffices to bound the image of this projection map, that is, to bound the number of possibilities for k.

Using the fact that the action of ${}^2F_4(q)$ preserves "$\circ$", the first three rows of relations from (A.11) yield

$$\begin{array}{lll} k_{1,1} + k_{2,1} = 1; & k_{2,1} + k_{3,1} = 0; & k_{3,1} + k_{1,1} = 1; \\ k_{1,2} + k_{2,2} = 1; & k_{2,2} + k_{3,2} = 1; & k_{3,2} + k_{1,2} = 0; \\ k_{1,3} + k_{2,3} = 0; & k_{2,3} + k_{3,3} = 1; & k_{3,3} + k_{1,3} = 1, \end{array} \tag{A.11}$$

so that k has the form

$$k = I_3 + \begin{pmatrix} x & y & z \\ x & y & z \\ x & y & z \end{pmatrix}.$$

However, the final row of relations from (A.11) then shows that $x + y = y + z = z + x = 0$, so $x = y = z$. Finally, since the eigenvalues of d occur in inverse pairs (see [43, Section 4.9.2]), we have $\det(k) = 1$, so there are only three possibilities for x, and hence for k.

A.4.7. Type ${}^2E_6$. The group $E_6(q)$ admits a central extension $\mathrm{SE}_6(q)$ of degree dividing 3. As discussed in [43, Section 4.10], $\mathrm{SE}_6(q)$ admits a faithful 27-dimensional linear representation over $\mathbb{F}_q$. A preferred basis $\mathcal{W} = \{w_0, \ldots, w_8''\}$ is given, and we can define a Hermitian form, with respect to which $\mathcal{W}$ is orthonormal. The group ${}^2E_6(q)$ is the image in $E_6(q^2)$ of the subgroup $H(q)$ of $\mathrm{SE}_6(q^2)$ consisting of those elements preserving this Hermitian form (see [43, Section 4.11]). By Lemma A.17 it suffices to show that $|E_{H(q)}(q+1)| = \Omega(|H(q)|)$.

There is described in [43, Section 4.10.3] a subgroup $T = T(q)$ of $\mathrm{SE}_6(q)$, which is the intersection of $\mathrm{SE}_6(q)$ with a maximal torus of the simple linear algebraic group SE_6 (defined over $K = \overline{\mathbb{F}}_q$). The elements of $T(q)$ are diagonal with respect to the basis $\mathcal{W}$, and are parametrised by six elements $\alpha, \beta, \gamma, \delta, \lambda, \mu \in \mathbb{F}_q^*$. Arguing as in the case of unitary groups, an element of $\mathrm{SE}_6(q^2)$ lying in $T(q^2)$ is in $H(q)$ iff it is supported on the set S of $(q+1)$th roots of unity. We apply Lemma A.18 to this set S (with $l = 6$, $m = 27$ and $D = 2$) and the 27 Laurent polynomials described in [43, table in Section 4.10.3]. We conclude that there exists a set $R(q) \subseteq T(q^2) \cap H(q)$ such that $|R(q)| \geq q^6/2$ for q larger than an explicit absolute constant, and $R(q)$ consists of elements of $T(q^2) \cap H(q)$, all of whose eigenvalues are distinct from each other and distinct from 1. Conditions (a) and

(b)(i) of Remark A.21 are then clear, as in the previous cases. The centraliser $C_{H(q)}(g)$ of an element $g \in R(q)$ is diagonal by Lemma A.16. It is not self-evident that $C_{H(q)}(g)$ is contained in $T(q^2) \cap H(q)$, but modulo scalars $C_{H(q)}(g)$ lies in the normaliser N of $T(q^2)$ in $E_6(q^2)$. The quotient map $C_{H(q)}(g) \to N/T(q^2)$ has fibres of size bounded by $|T(q^2) \cap H(q)|(q+1)^6$, and image a subgroup of $W(E_6) \cong \mathrm{SO}_5(3)$ (see [43, Section 4.10.3] again), and we have condition (b)(ii) of Remark A.21.

A.4.8. Type ${}^3D_4$. The conclusion of Proposition 2.14 for the groups ${}^3D_4(q)$ is essentially contained in [12]. Therein, G is a simple, simply-connected group of Dynkin type D_4 over $K = \overline{\mathbb{F}}_p$; q is a power of p; T is a maximal torus of G and σ is an automorphism of G such that T is σ-stable and ${}^3D_4(q) = G_\sigma$, the set of σ-fixed points of G.

There is another σ-stable maximal torus T' of G, and $g \in G$ such that $(T'_\sigma)^g = T_4 \subseteq T$, where $T_4 \cong C^2_{q^2-q+1}$ is as described in [12, Table 1.1]. The regular semisimple elements of G lying in T_4 are described in [12, Propositions 2.1–2.2]: they are the elements of the form s_{13} appearing in Table 2.1 from that paper.

By [12, Table 4.4] and the discussion preceding it, the number of conjugacy classes of elements of type s_{13} in G_σ (which is the same as the number of irreducible characters of type χ_{13}) is $(q^4 - 2q^3 - q^2 + 2q)/24$. It therefore suffices to show that each such element x has centraliser in G_σ of order $O(q^4)$. This is so, because x is regular: by Proposition A.9, $C_{G_\sigma}(x) \leq N_{G_\sigma}(T')$, so

$$\begin{aligned} |C_{G_\sigma}(x)| &\leq |N_{G_\sigma}(T') : T'_\sigma| \cdot |T'_\sigma| \\ &\leq |N_G(T') : T'| \cdot |T_4| \\ &\leq |W|(q^2 - q + 1)^2 \quad \text{(by Theorem A.10)} \end{aligned}$$

as required.

This concludes the proof of Proposition 2.14.

Acknowledgements. HB would like to thank Tim Burness, Martin Liebeck, Nikolay Nikolov and Emilio Pierro for enlightening conversations.

Funding. This research was supported by ERC CoG 681207 and ERC grant "GRANT" no. 648329.

References

[1] Aschbacher, M.: On the maximal subgroups of the finite classical groups. Invent. Math. **76**, 469–514 (1984) Zbl 0537.20023 MR 0746539

[2] Babai, L., Seress, Á.: On the diameter of permutation groups. Eur. J. Combin. **13**, 231–243 (1992) Zbl 0783.20001 MR 1179520

[3] Birkhoff, G.: On the structure of abstract algebras. Proc. Cambridge Philos. Soc. **31**, 433–454 (1935), Zbl 0013.00105 Zbl 61.1026.07

[4] Bou-Rabee, K.: Quantifying residual finiteness. J. Algebra **323**, 729–737 (2010) Zbl 1222.20020 MR 2574859

[5] Bou-Rabee, K., McReynolds, D. B.: Asymptotic growth and least common multiples in groups. Bull. London Math. Soc. **43**, 1059–1068 (2011) Zbl 1253.20029 MR 2861528

[6] Bradford, H., Thom, A.: Short laws for finite groups and residual finiteness growth. Trans. Amer. Math. Soc. **371**, 6447–6462 (2019) Zbl 1515.20117 MR 3937332

[7] Bray, J. N., Holt, D. F., Roney-Dougal, C. M.: The maximal subgroups of the low-dimensional finite classical groups. London Math. Soc. Lecture Note Ser. 407, Cambridge University Press, Cambridge (2013) Zbl 1303.20053 MR 3098485

[8] Breuillard, E., Green, B., Guralnick, R., Tao, T.: Expansion in finite simple groups of Lie type. J. Eur. Math. Soc. **17**, 1367–1434 (2015) Zbl 1331.20060 MR 3353804

[9] Breuillard, E., Green, B., Tao, T.: Approximate subgroups of linear groups. Geom. Funct. Anal. **21**, 774–819 (2011) Zbl 1229.20045 MR 2827010

[10] Breuillard, E., Green, B., Tao, T.: Suzuki groups as expanders. Groups Geom. Dynam. **5**, 281–299 (2011) Zbl 1247.20017 MR 2782174

[11] Carter, R. W.: Simple groups of Lie type. Pure Appl. Math. 28, John Wiley & Sons, London (1972) Zbl 0248.20015 MR 0407163

[12] Deriziotis, D. I., Michler, G. O.: Character table and blocks of finite simple triality groups ${}^3D_4(q)$. Trans. Amer. Math. Soc. **303**, 39–70 (1987) Zbl 0628.20014 MR 0896007

[13] Diaconis, P., Saloff-Coste, L.: Comparison techniques for random walk on finite groups. Ann. Probab. **21**, 2131–2156 (1993) Zbl 0790.60011 MR 1245303

[14] Donkin, S.: A note on decomposition numbers of reductive algebraic groups. J. Algebra **80**, 226–234 (1983) Zbl 0505.20028 MR 0690715

[15] Elkasapy, A.: A new construction for the shortest non-trivial element in the lower central series. arXiv:1610.09725 (2016)

[16] Elkasapy, A., Thom, A.: On the length of the shortest non-trivial element in the derived and the lower central series. J. Group Theory **18**, 793–804 (2015) Zbl 1346.20045 MR 3393415

[17] Hadad, U.: On the shortest identity in finite simple groups of Lie type. J. Group Theory **14**, 37–47 (2011) Zbl 1234.20019 MR 2764921

[18] Jones, G. A.: Varieties and simple groups. J. Austral. Math. Soc. **17**, 163–173 (1974) Zbl 0286.20028 MR 0344342

[19] Kantor, W. M., Lubotzky, A.: The probability of generating a finite classical group. Geom. Dedicata **36**, 67–87 (1990) Zbl 0718.20011 MR 1065213

[20] Kantor, W. M., Seress, Á.: Large element orders and the characteristic of Lie-type simple groups. J. Algebra **322**, 802–832 (2009) Zbl 1180.20009 MR 2531224

[21] Kassabov, M., Lubotzky, A., Nikolov, N.: Finite simple groups as expanders. Proc. Nat. Acad. Sci. USA **103**, 6116–6119 (2006) Zbl 1161.20010 MR 2221038

[22] Kassabov, M., Matucci, F.: Bounding the residual finiteness of free groups. Proc. Amer. Math. Soc. **139**, 2281–2286 (2011) Zbl 1230.20045 MR 2784792

[23] Kleidman, P., Liebeck, M.: The subgroup structure of the finite classical groups. London Math. Soc. Lecture Note Ser. 129, Cambridge University Press, Cambridge (1990) Zbl 0697.20004 MR 1057341

[24] Kleidman, P. B.: The maximal subgroups of the Chevalley groups $G_2(q)$ with q odd, the Ree groups ${}^2G_2(q)$, and their automorphism groups. J. Algebra **117**, 30–71 (1988) Zbl 0651.20020 MR 0955589

[25] Kleidman, P. B.: The maximal subgroups of the Steinberg triality groups ${}^3D_4(q)$ and of their automorphism groups. J. Algebra **115**, 182–199 (1988) Zbl 0642.20013 MR 0937609

[26] Kozma, G., Thom, A.: Divisibility and laws in finite simple groups. Math. Ann. **364**, 79–95 (2016) Zbl 1344.20024 MR 3451381

[27] Landazuri, V., Seitz, G. M.: On the minimal degrees of projective representations of the finite Chevalley groups. J. Algebra **32**, 418–443 (1974) Zbl 0325.20008 MR 0360852

[28] Liebeck, M. W.: On the orders of maximal subgroups of the finite classical groups. Proc. London Math. Soc. (3) **50**, 426–446 (1985) Zbl 0591.20021 MR 0779398

[29] Liebeck, M. W.: Probabilistic and asymptotic aspects of finite simple groups. In: Probabilistic group theory, combinatorics, and computing, Lecture Notes in Math. 2070, Springer, London, 1–34 (2013) Zbl 1288.20103 MR 3026185

[30] Liebeck, M. W., Saxl, J., Seitz, G. M.: Subgroups of maximal rank in finite exceptional groups of Lie type. Proc. London Math. Soc. (3) **65**, 297–325 (1992) Zbl 0776.20012 MR 1168190

[31] Liebeck, M. W., Seitz, G. M.: A survey of maximal subgroups of exceptional groups of Lie type. In: Groups, combinatorics & geometry (Durham, 2001), World Scientific Publishing, River Edge, NJ, 139–146 (2003) Zbl 1032.20010 MR 1994964

[32] Liebeck, M. W., Shalev, A.: The probability of generating a finite simple group. Geom. Dedicata **56**, 103–113 (1995) Zbl 0836.20068 MR 1338320

[33] Lovász, L.: Random walks on graphs: a survey. In: Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud. 2, János Bolyai Math. Soc., Budapest, 353–397 (1996) Zbl 0854.60071 MR 1395866

[34] Malle, G.: The maximal subgroups of ${}^2F_4(q^2)$. J. Algebra **139**, 52–69 (1991) Zbl 0725.20014 MR 1106340

[35] Malle, G., Testerman, D.: Linear algebraic groups and finite groups of Lie type. Cambridge Stud. Adv. Math. 133, Cambridge University Press, Cambridge (2011) Zbl 1256.20045 MR 2850737

[36] Neumann, H.: Varieties of groups. Ergeb. Math. Grenzgeb. 37, Springer, Berlin (1967) Zbl 0251.20001 MR 0215899

[37] Pyber, L., Szabó, E.: Growth in finite simple groups of Lie type. J. Amer. Math. Soc. **29**, 95–146 (2016) Zbl 1371.20010 MR 3402696

[38] Rosenfeld, B., Wiebe, B.: Geometry of Lie groups. Springer (2013) Zbl 0867.53002 MR 1443207

[39] Thom, A.: About the length of laws for finite groups. Israel J. Math. **219**, 469–478 (2017) Zbl 1475.20038 MR 3642030

[40] Wagner, A.: The faithful linear representation of least degree of S_n and A_n over a field of characteristic 2. Math. Z. **151**, 127–137 (1976) Zbl 0321.20008 MR 0419581

[41] Wagner, A.: The faithful linear representations of least degree of S_n and A_n over a field of odd characteristic. Math. Z. **154**, 103–114 (1977) Zbl 0336.20008 MR 0437631

[42] Wagner, A.: An observation on the degrees of projective representations of the symmetric and alternating group over an arbitrary field. Arch. Math. (Basel) **29**, 583–589 (1977) Zbl 0383.20009 MR 0460451

[43] Wilson, R. A.: The finite simple groups. Grad. Texts in Math. 251, Springer London, London (2009) Zbl 1203.20012 MR 2562037

[44] Zyrus, C.: Almost laws for finite simple groups. Ph.D. thesis, Technische Universität Dresden (2020)