Recognising simplicity of black-box groups by constructing involutions and their centralisers

Christopher W. Parker
School of Mathematics,
University of Birmingham,
Edgbaston, Birmingham B15 2TT, U.K.

and

Robert A. Wilson
School of Mathematical Sciences,
Queen Mary, University of London,
Mile End Road, London E1 4NS, U.K.

June 1, 2009

Abstract

We investigate the complexity of constructing involutions and their centralisers in groups of Lie type over finite fields of odd order, and discuss applications to the problem of deciding whether a matrix group, or a black-box group of known characteristic, is simple. We show that if the characteristic is odd, then simplicity can be recognised in Monte Carlo polynomial time.

Mathematics Subject Classification: 20D06, 20-04

1 Introduction

It has been a fundamental tenet of abstract group theory for half a century that the key to studying finite simple groups is to study their involution centralisers. Also at the computational level many practical problems can be reduced to corresponding problems in involution centralisers. However, most of these practical methods have the flavour of ad hoc tricks, and have not always been developed into general algorithms. One reason for this is that these methods do not perform

well in the worst case (they are frequently exponential-time algorithms), and even in good cases their complexity is hard to analyse.

The purpose of the present paper is to promote the use of involution-centraliser methods in computational group theory, both by presenting some practical algorithms for solving particular problems which appear to be computationally hard, and by analysing the complexity of the methods proposed for finding involutions and their centralisers in certain cases.

The main results of this paper are as follows. The first three are to do with computing centralisers of involutions.

Theorem 1. *If G is a simple exceptional group of Lie type defined over a field of odd order, and z is any involution in G, then the proportion of ordered pairs (z, z^g) with zz^g of odd order is bounded below by a positive constant.*

Theorem 2. *If G is a simple classical group defined over a field of odd order, with natural module of dimension n, and z is any involution in G, then the proportion of ordered pairs (z, z^g) with zz^g of odd order is bounded below by n^{-1} times a positive constant.*

We shall also show in Section 2.4 that this bound is best possible.

Corollary 3. *If G is a black-box group which is isomorphic to a (known) simple group S of Lie type defined over a field of odd order, or to an extension of an odd-order normal subgroup by S, and z is an involution in G, then there exists a Monte Carlo polynomial time algorithm to compute $C_G(z)$.*

The next three results are to do with computing involutions.

Theorem 4. *If G is a simple exceptional group of Lie type defined over a field of odd order, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power up to an element of $\mathcal{C}$ is at least a positive constant.*

Theorem 5. *If G is a simple classical group defined over a field of odd order, with natural module of dimension n, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power up to an element of $\mathcal{C}$ is at least n^{-3} times a positive constant.*

Corollary 6. *If G is a black-box group which is isomorphic to a (known) simple group S of Lie type defined over a field of odd order, or to an extension of an odd-order normal subgroup by S, and $\mathcal{C}$ is a fixed conjugacy class of involutions in G, then there exists a Monte Carlo polynomial time algorithm to compute an element of $\mathcal{C}$.*

Finally we have the applications to recognising simple groups.

Theorem 7. *If G is a black-box group such that $G/O_p(G)$ is a known simple group of Lie type in characteristic p, then there is a Monte Carlo polynomial time algorithm to decide whether $O_p(G) = 1$.*

Corollary 8. *A finite simple group of Lie type over a field of (known) odd characteristic can be recognised in the class of all black-box groups in Monte Carlo polynomial time.*

Corollary 9. *A finite simple group of Lie type over a field of odd order can be recognised in the class of all matrix groups in Monte Carlo polynomial time.*

Theorem 1 is proved in Section 2.2 and Theorem 2 is proved in Section 2.4. Theorem 4 is proved in Section 3.2 and Theorem 5 in Section 3.4. The classical group cases (Theorems 2 and 5) rely on some technical counting arguments, which are collected in Sections 2.3 and 3.3. Corollaries 3 and 6 are proved in Section 4, as are Theorem 7 and Corollaries 8 and 9.

The catalyst for this paper was provided by the appearance of [16], which contains a paper by Altseimer and Borovik [1] and another by Babai and Shalev [3]. The former uses involution centralisers to distinguish $\mathrm{PSp}_{2n}(q)$ from $\Omega_{2n+1}(q)$ in a computational setting, while in the latter the main obstacle to determining whether a black box group is simple is exemplified by the 'Challenge Problem' of distinguishing a simple group such as $\mathrm{PSL}_2(p^2)$ from an extension of shape $p^4{:}\mathrm{PSL}_2(p^2)$, when p is a very large prime. We immediately realised that the involution centraliser method provides a neat and easy solution to this latter problem.

Assume for the moment that we are in a computational setting in which it is possible to find involutions, and their centralisers. If z is any involution in $\mathrm{PSL}_2(p^2)$ or $p^4{:}\mathrm{PSL}_2(p^2)$ then $C(z)/\langle z\rangle$ is $D_{(p^2-1)/2}$ or $p^2{:}D_{(p^2-1)/2}$ respectively. But these two groups can be distinguished (with arbitrarily high probability) because in the first case two random commutators commute, while in the second case they fail to commute with probability approximately 15/16. The solves the 'Challenge Problem' in odd characteristic. This paper is essentially just a generalisation of this result.

We refer to Babai and Shalev [3] for background to the problem, and any definitions and preliminary results which are not found here can be found either there, or in Babai and Beals [2]. The fundamental notion of black-box group was introduced by Babai and Szemeredi in [4]. We have stated our results in such a way that an order oracle is not required. This is because in a known simple group, suitable pseudo-orders can be computed in Monte Carlo polynomial time, and these suffice for all our computations.

Information about, and notation for, involutions and their centralisers is taken from Table 4.5.1 of [13] unless otherwise stated. Structures of maximal tori are taken from [15]. A great deal of information about (maximal or other) subgroups of classical groups is used implicitly: this can mostly be found in Kleidman and Liebeck's book [17].

2 Finding the centraliser of a given involution

2.1 General strategy

The standard method for constructing the centraliser of an involution in a black-box group with a pseudo-order oracle (see [2]) is Bray's algorithm [9]. The generators of the centraliser are constructed from the generators of the group and the involution itself, making use of the following two results, which have been formulated so that they can be applied to groups in which only pseudo-orders are available.

Theorem 10. *If z is an involution in a group G, and g is any element of G, then, for some integer k, either*

(i) $[z,g]^{2k+1} = 1$ *and* $g[z,g]^k \in C_G(z)$, *or*

(ii) $[z,g]^{2k}$ *and both* $[z,g]^k$ *and* $[z,g^{-1}]^k$ *lie in* $C_G(z)$.

Proof. If $[z,g]$ has odd (pseudo-)order, then say $[z,g]^{2k+1} = 1$, and therefore $zg[z,g]^k = gz[z,g]^{k+1} = gz[z,g]^{-k} = g[z,g]^k z$ since z is an involution; otherwise, say $[z,g]^{2k} = 1$, and so $z[z,g^{\pm 1}]^k = z[z,g^{\pm 1}]^{-k} = [z,g^{\pm 1}]^{-k}z$. □

Thus if we have a method of producing independent nearly uniformly distributed random elements of G, this theorem gives us a method of producing some elements of $C_G(z)$. Moreover, in case (i) the resulting elements are nearly uniformly distributed, as the following result of Richard Parker shows [9].

Theorem 11. *With the notation of Theorem 10, if g is (nearly) uniformly distributed among the elements of G for which $[z,g]^{2k+1} = 1$ for some integer k, then $g[z,g]^k$ is (nearly) uniformly distributed among the elements of $C_G(z)$.*

Proof. If $h = yg$, where $y \in C_G(z)$, then $[z,h] = [z,g]$ so that $h[z,h]^k = yg[z,g]^k$. Therefore each element of $C_G(z)$ occurs exactly once as g runs over any coset of $C_G(z)$. □

Unfortunately there is no such result in case (ii) of Theorem 10: this is obvious since $[x,g^{\pm 1}]^k$ is an involution. However, if the odd order case occurs sufficiently often, then we can construct nearly-uniformly distributed random elements of the involution centraliser efficiently. Of course in practice, we use the even order case of Bray's algorithm as well: although the resulting elements of the centraliser are not nearly uniformly distributed, they do significantly speed up the process of constructing the centraliser.

If G is a finite simple group of Lie type defined over a field of odd order, then the involution centralisers are well-understood, and they are generated with arbitrarily high probability by a constant number of nearly uniformly distributed random elements [19]. Therefore in order to prove Corollary 3 it suffices to prove that case (i) of Theorem 10 occurs with probability at least a positive

rational function of the input size. More details of this reduction can be found in Theorem 7 of [14].

Thus we take G to be a simple group of Lie type, of Lie rank r, defined over a field of odd order. For each class of involutions we find suitable dihedral groups (of twice odd order), and show that a proportion cr^{-1} of pairs of involutions in this class generate such a dihedral group. In order to avoid double counting, we only count the cases where the cyclic part of the dihedral group is generated by a regular semisimple element in a suitable subgroup of G, so that the centraliser of the dihedral group is easy to calculate. (A semisimple element is called *regular* in H if it lies in a unique maximal torus of H or equivalently of the ambient algebraic group.)

In the proof of asymptotic results such as Theorems 1 and 2 we may neglect any finite number of simple groups, so we may assume that either the Lie rank or the field is 'large'. This enables us to show that 'most' elements in our chosen cyclic groups are regular semisimple elements. The easiest cases are the exceptional groups, and, perhaps surprisingly, the hardest are $\mathrm{PSL}_n(q)$ and $\mathrm{PSU}_n(q)$. Therefore we treat the exceptional groups first.

The proof breaks into two parts. The first is finding a suitable class of dihedral subgroups of twice odd order, and using a dimension-counting argument to show that the proportion of involution pairs which lie in these groups is, asymptotically, independent of q. (Here we define the *dimension* of a group defined over $\mathrm{GF}(q)$ to be the same as the dimension of the corresponding algebraic group as an algebraic variety.)

To be more specific, the number of involutions conjugate to z is, up to a constant factor, roughly q^k, where $k = \dim G - \dim C_G(z)$, so the number of involution pairs is asymptotically a constant times q^{2k}. If T is a torus all of whose elements are inverted by z, then it contains roughly $q^{2\dim T}$ pairs of such involutions. Moreover the number of conjugates of T is roughly $q^{\dim G - \dim C_G(T)}$, so, up to a constant factor, the number of pairs of involutions accounted for in this way is q^l, where $l = \dim G + 2\dim T - \dim C_G(T)$. We want this to be a constant proportion of all the pairs of involutions, independent of q, so we need to show that the dimensions are equal, that is

$$2\dim G - 2\dim C_G(z) = \dim G + 2\dim T - \dim C_G(T)$$

or, simplifying,

$$2\dim T + \mathrm{codim}\, C_G(T) = 2\mathrm{codim}\, C_G(z).$$

The second part of the proof is estimating the constants. There are several sources of constants we need to control:

(i) $N_G(T)/C_G(T)$. This is a subgroup of the Weyl group, so in the exceptional groups has bounded (but possibly large) order. In the classical groups we need to take care over our choice of T to make sure this group is not too large.

(ii) The difference between the simple group and the adjoint group. This is the group of diagonal automorphisms, so in the exceptional groups has order 1, 2 or 3. In the symplectic and orthogonal groups we get a factor of at most 4, but in the linear and unitary groups we have $(n, q \pm 1)$ so more care is required.

(iii) The 2-part of the order of T. We choose T carefully so that this is never greater than 4.

(iv) The proportion of odd-order elements of T which have larger centraliser than T. In the exceptional groups we can choose q large enough so that this proportion is small and can be neglected. In the classical groups when q is small this is the source of some difficulty.

(v) The difference between $|H|$ and $q^{\dim H}$. Again, if q is large enough this can be neglected. In the case of the classical groups we must explicitly estimate this error factor (see Corollary 15).

2.2 Products of involutions in exceptional groups of Lie type in odd characteristic

There are eight families of exceptional groups of Lie type in odd characteristic, and each seems to need individual treatment. In each case the rank is a constant, as is the order of the Weyl group. Moreover, by neglecting a finite number of groups (in which the results hold trivially), we may assume the order of the field is as large as we like. Thus of the five sources of constants listed in Section 2.1, only (iii) remains to be taken care of.

We consider first the cases where the Weyl group has a central involution, and the involution z fuses into this conjugacy class in the algebraic group. Notice that there may be more than one such conjugacy class in the finite group of Lie type.

Theorem 12. *If G is an exceptional group of Lie type over a field of odd order, and z is conjugate in the ambient algebraic group to the central involution in the Weyl group of G, then the proportion of ordered pairs (z, z^g) with zz^g of odd order is bounded below by a positive constant.*

Proof. The involution z inverts every type of maximal torus T in G. Since T is maximal, $T = C_G(T)$, so the dimension formula which we need to prove simplifies to

$$\dim T + 2 \dim C_G(z) = \dim G.$$

We choose the following maximal tori of odd order:

G	$\lvert T\rvert$	$\lvert N_G(T)/T\rvert$	$\dim G$	type of z	$\dim C_G(z)$
${}^2G_2(q)$	$q+\sqrt{3q}+1$	6	7	t_1	3
$G_2(q)$	q^2+q+1	6	14	t_1	6
${}^3D_4(q)$	q^4-q^2+1	4	28	t_2	12
$F_4(q)$	q^4-q^2+1	12	52	t_1	24
$E_7(q)$	$(q^7\pm 1)/2$	14	133	t_4/t_4'	63
$E_8(q)$	q^8-q^4+1	24	248	t_1	120

(In the case of $E_7(q)$, there are two classes t_4, t_4' of involutions in the adjoint group $E_7(q).2$ fusing to the central involution of the Weyl group. Class t_4 lies in the simple group if $q \equiv 1 \bmod 4$, and class t_4' does if $q \equiv 3 \bmod 4$. We choose T of order $(q^7+\varepsilon)/2$ where $q \equiv \varepsilon \bmod 4$.)

In particular, we observe that in every case there are some nontrivial odd-order products of two involutions in the given class, and therefore the theorem holds for an arbitrary finite number of cases. Thus we can ignore finitely many values of q, at the expense of possibly having to change the constant in the theorem. Now for large q, the proportion of pairs of inverting involutions whose product is a regular semisimple element tends to 1. Therefore the number of pairs of involutions accounted for in this way is $\sim q^k/c$, where $k = 2\dim T + (\dim G - \dim T) = \dim G + \dim T$ and c is a constant, equal to $|N_G(T)/T|$ in all cases except E_7, where it is $4|N_G(T)/T|$. On the other hand, the dimension of the set of pairs of involutions in this class is $2(\dim G - \dim C_G(z))$. Using the table above, we readily check that $2(\dim G - \dim C_G(z)) = k$. Hence the proportion of pairs of involutions whose product is a regular semisimple element in a torus of this type tends to $1/c$ as q tends to infinity. □

Theorem 13. *If G is an exceptional group of Lie type over a field of odd order, and z is an involution which is not conjugate to the central involution (if any) in the Weyl group of G, then the proportion of ordered pairs (z, z^g) with zz^g of odd order is bounded below by a positive constant.*

Proof. The classes of involutions which we need to consider are listed below. The first four columns contain information from [13], and the last two columns summarise our choice of torus, and information about the centraliser of the torus which we shall prove as we go along.

G	$\dim G$	type of z	$\dim C_G(z)$	$\lvert T\rvert$	$\dim C_G(T)$
$F_4(q)$	52	t_4	36	$q\pm 1$	22
$E_6^{\pm}(q)$	78	t_1	46	$(q\pm 1)\times(q\pm 1)$	18
		t_2	38	q^4-q^2+1	6
$E_7(q)$	133	t_1	69	q^4+1	13
		t_7/t_7'	79	$(q\pm 1)\times(q^2+1)$	31
$E_8(q)$	248	t_8	136	q^4+1	32

In $E_8(q)$ consider the subgroup $2{\cdot}(\mathrm{P}\Omega_8^-(q) \times \mathrm{P}\Omega_8^-(q))$ inside $2{\cdot}\mathrm{P}\Omega_{16}^+(q)$. The involutions of type $-1^4 1^4$ in $\Omega_8^-(q)$ lift to involutions in the spin group $2{\cdot}\Omega_8^-(q)$, and it is straightforward to calculate the eigenvalues of these involutions acting on the Lie algebra: these are $-1^{112}1^{136}$, and therefore they are involutions of type t_8 in the notation of [13]. In particular $\operatorname{codim} C_G(z) = 112$. Also these involutions fuse to the central involution of the Weyl group of type D_4, so invert every maximal torus of $\mathrm{O}_8^-(q)$. In particular they invert the cyclic torus of order q^4+1, which is twice an odd number. Finally, the centraliser of T is, up to a constant factor, $T \circ 2{\cdot}\Omega_8^-(q)$, so has dimension $4+28=32$ and codimension 216. So $2\dim T + \operatorname{codim} C_G(T) = 8 + 216 = 224 = 2 \times 112$ as required.

In $E_7(q)$ we look inside $(\mathrm{SL}_2(q) \circ 2\Omega_{12}^+(q).2).2$ at the involutions of type $(-1,1)\otimes(-1^2 1^{10})$ and calculate their eigenvalues on the Lie algebra to be $-1^{54}1^{79}$. Therefore they are involutions of type t_7 or t_7' according as $q \equiv 1$ or $3 \bmod 4$, and thus $\operatorname{codim} C_G(z) = 54$. Such involutions can simultaneously invert cyclic groups of order $q \pm 1$ in $\mathrm{SL}_2(q)$ and q^2+1 in $2{\cdot}\Omega_4^-(q)$ so we obtain a torus T of dimension 3 and at most 4 times odd order, with centraliser $T \circ 2{\cdot}\Omega_8^-(q)$. Thus $2\dim T + \operatorname{codim} C_G(T) = 6 + 133 - 3 - 28 = 108 = 2 \times 54$ as required.

The other class of involutions in $E_7(q)$ can be dealt with again in the group $(\mathrm{SL}_2(q) \circ 2\Omega_{12}^+(q).2).2$, this time looking at the involutions of type $-1^4 1^8$ in $\mathrm{O}_{12}^+(q)$. These have eigenvalues $-1^{64}1^{69}$ on the Lie algebra, and so are of type t_1, and have centraliser of codimension $133 - 3 - 66 = 64$ in G. They invert a cyclic torus T of order q^4+1 (and so of twice odd order) inside $2{\cdot}\Omega_8^-(q)$, and therefore T has centraliser $T \circ \mathrm{SL}_2(q) \circ \mathrm{SL}_2(q^2)$ of dimension $4+3+6=13$. Finally we calculate $2\dim T + \operatorname{codim} C_G(T) = 8 + 133 - 13 = 128 = 2 \times 64$ as required.

In $E_6(q)$ or ${}^2E_6(q)$ we look inside the subgroup $({}^3D_4(q) \times (q^2 \pm q + 1)).3$, and find involutions inverting a cyclic torus of order $q^4 - q^2 + 1$ (and hence odd order) inside ${}^3D_4(q)$. This torus has centraliser $T \times C_{q^2 \pm q + 1}$ of dimension 6 only. The involutions centralise $\mathrm{SL}_2(q) \circ \mathrm{SL}_2(q^3)$ inside ${}^3D_4(q)$, so are of type t_2 in $E_6^{\pm}(q)$, and therefore have centraliser of codimension $78 - 3 - 35 = 40$. Finally, $2\dim T + \operatorname{codim} C_G(T) = 8 + 78 - 6 = 80 = 2 \times 40$ as required.

The other class of involutions in $E_6(q)$ or ${}^2E_6(q)$ can be dealt with by looking inside the subgroup $2{\cdot}(\mathrm{PSL}_2(q) \times \mathrm{PSL}_6(q)).2$, respectively $2{\cdot}(\mathrm{PSL}_2(q) \times \mathrm{PSU}_6(q)).2$, at an involution of type $(-1,1) \otimes (-1,1^5)$. This involution has eigenvalues $-1^{32}1^{46}$ in its action on the Lie algebra, so is of type t_1. It inverts various tori T of rank 2 with centralisers $T \times GL_4(q)$ of dimension 18, and we have $\operatorname{codim} C_G(z) = 78 - 1 - 45 = 32$, so $2\dim T + \operatorname{codim} C_G(T) = 4 + 78 - 18 = 64 = 2\operatorname{codim} C_G(z)$ as required. More precisely, in $E_6(q)$ these tori can be seen inside $2^2.(\mathrm{PSL}_2(q) \times (\mathrm{PSL}_2(q) \times \mathrm{PSL}_4(q)).C_{q-1}).2$ in $2.(\mathrm{PSL}_2(q) \times \mathrm{PSL}_6(q)).2$ and we can choose either $C_{q-1} \times C_{q-1}$ or $C_{q+1} \times C_{q+1}$, one of which has 4 times odd order. Similarly in ${}^2E_6(q)$ we have $2^2.(\mathrm{PSL}_2(q) \times (\mathrm{PSL}_2(q) \times \mathrm{PSU}_4(q)).C_{q+1}).2$ in $2.(\mathrm{PSL}_2(q) \times \mathrm{PSU}_6(q)).2$, and the same argument applies.

Finally consider the involutions of type t_4 in $F_4(q)$. Note that the negatives of reflections in $\mathrm{O}_9(q)$ lift to involutions of this type. Therefore they invert tori T of

dimension 1, centralising $T \circ 2 \cdot \Omega_7(q)$. In particular $\operatorname{codim} C_G(T) = 52-1-21 = 30$ and $\operatorname{codim} C_G(z) = 52 - 36 = 16$, so $2 \dim T + \operatorname{codim} C_G(T) = 2 + 30 = 2 \times 16$ as required. We can choose T to have order $q \pm 1$, so of twice odd order. □

Putting together Theorems 12 and 13 we have Theorem 1.

2.3 Some counting arguments

It is more difficult to prove analogous results for classical groups, as dimension counting arguments alone are not sufficient. We need explicit bounds on the numbers of regular semisimple elements of odd order in various subgroups in order to deal with small fields. Indeed, occasionally our generic proofs do not work for the fields of orders 3 or 5, and separate arguments are required. We collect together in this section the various technical counting arguments we shall need.

Lemma 14. *If q is any real number with $q \geq 3$, and k, m are positive integers with $k \leq m$, then*

$$\prod_{j=k}^{m} \left(1 - \frac{1}{q^j}\right) \geq 1 - \frac{1}{(q-1)q^{k-1}} \geq \frac{1}{2}.$$

Proof. Given any $n \geq 1$ and any $0 < x_l < 1$ for $1 \leq l \leq n$, one proves immediately by induction on n that $\prod_{l=1}^{n}(1 - x_l) \geq 1 - \sum_{l=1}^{n} x_l$. In particular,

$$\prod_{j=k}^{m} \left(1 - \frac{1}{q^j}\right) \;\geq\; 1 - \sum_{j=k}^{\infty} q^{-j} = 1 - \frac{1}{q^{k-1}(q-1)} \geq 1 - \frac{1}{2q^{k-1}} \geq \frac{1}{2}$$

as required. □

We obtain as an immediate corollary the following useful bounds on the orders of certain classical groups (somewhat better bounds can obviously be obtained with a more careful analysis):

Corollary 15. *If $q \geq 3$ then*

(i) $\frac{1}{2}q^{n^2} \leq |\mathrm{GL}_n(q)| \leq q^{n^2}$;

(ii) $\frac{1}{2}q^{n^2} \leq |\mathrm{GU}_n(q)| \leq 2q^{n^2}$;

(iii) $\frac{1}{2}q^{\frac{1}{2}n(n+1)} \leq |\mathrm{Sp}_n(q)| \leq q^{\frac{1}{2}n(n+1)}$; *and*

(iv) $\frac{1}{2}q^{\frac{1}{2}n(n-1)} \leq |\mathrm{SO}_n(q)| \leq 2q^{\frac{1}{2}n(n-1)}$.

□

These bounds will be used for estimating the orders of centralisers of involutions, as well as cyclic and dihedral groups, and thereby estimating the numbers of involutions, cyclic and dihedral groups of various types. This deals with part (v) of the list of sources of constants in Section 2.1.

For small fields, especially the field of order 3, we also need to deal with part (iv). The next lemma will be used to estimate the numbers of elements in certain cyclic groups which have the same centraliser as the cyclic group itself. Since the cyclic group is a torus, it is a well-understood subgroup of the multiplicative group of a well-defined field, and the crux is to eliminate the elements ± 1 and all elements which lie in any subfield.

Lemma 16. *(i) Suppose that C is a subgroup of the multiplicative group of the field $F = \mathrm{GF}(p^k)$, where p is an odd prime. Assume that C has order $(p^k-1)/d$ where d divides $(p^k-1, 4)$. Then either the proportion of elements of C which are not ± 1 and lie in no proper subfield of F is at least $1/2$ or $(p, k, d) = (3, 1, 1)$, $(3, 1, 2)$, $(5, 1, 2)$, $(5, 1, 4)$ or $(3, 2, 4)$.*

(ii) Suppose C is a subgroup of the subgroup of order $(p^k + 1)$ in the field $F = \mathrm{GF}(p^{2k})$, where p is an odd prime. Assume that C has order $(p^k + 1)/d$ where d divides $(p^k + 1, 4)$. Then either the proportion of elements of C which lie in no proper subfield of F is at least $1/2$ or $(p, k, d) = (3, 1, 2)$ or $(7, 1, 4)$.

Parts (i) and (ii) also hold in the quotient of C by $C \cap \{\pm 1\}$.

Proof. (i) Assume that C has order $(p^k - 1)/d$ with d dividing $(p^k - 1, 4)$. We prove that the proportion of elements of C which are also in a proper subfield of F or are ± 1 is at most one half. If $k = 1$, then F has no proper subfield and the required proportion is at least $2d/(p - 1)$, which is at most $1/2$ provided $p \geq 17$. If $d = 1$, we just require $p \neq 3$. For $d = 2$, if $p = 7$, we note that the proportion of elements not equal to $\pm 1 = 1$ in the cyclic group of order 3 is $\frac{2}{3}$, so we only need to exclude $p = 3$ and 5. For $d = 4$, we have that 4 divides $p - 1$. Plainly we must exclude $p = 5$. For $p = 13$, we have that C has order 3 and so $\frac{2}{3}$ of its elements are not ± 1.

For $k = 2$, there are $(p - 1)$ elements in the proper subfield. Provided $p \geq 7$ we have $\frac{d}{p+1} \leq \frac{1}{2}$. For $p = 5$, this can fail only if $d = 4$, when C has order 6 and 4 of its elements are not in proper subfields. So suppose that $p = 3$, then again $d = 4$ and this case is excluded from our consideration.

For $k = 3$, it is sufficient to show $2d \leq p^2 + p + 1$ and for $k = 4$ it is sufficient to show $2d \leq p^2 + 1$. Both these hold as $p \geq 3$. So the result is true for $k < 5$. Suppose that $k \geq 5$. Then

$$p + p^2 + \cdots + p^{\lfloor k/2 \rfloor} < \frac{p^{k-2} - 1}{p - 1} < \frac{p^k - 1}{(p - 1)^3}$$

Since

$$\frac{d}{p^k-1}.\frac{p^k-1}{(p-1)^3} = \frac{d}{(p-1)^3} \leq \frac{1}{2},$$

this proves that at least half of the elements of C are contained in no proper subfield of F. So (i) holds.

(ii) We follow the same method of proof as for part (i). For $k = 1$, the required inequality is $4d \leq p+1$. For $d = 1$, the inequality holds for all odd p. For $d = 2$, we note that the inequality holds for $p \geq 7$. For $p = 5$, we have that C has order 3 and $\frac{2}{3}$ of its elements are not ± 1. For $d = 4$, we have to consider $p = 11$ specially. But then C has order 3 and again $2/3$ of its elements are not ± 1.

Suppose that $k = 2$. Then the result holds if $4d \leq p^2 + 1$. Otherwise, $d = 4$ and $p = 3$. But then 4 doesn't divide $p^2 + 1$. So the result holds for $k = 2$.

Assume that $k = 3$. Then the result holds if $2d \leq p^2 - p + 1$. Otherwise, $p = 3$ and $d = 4$. In this case $|C| = 7$ and so in fact $6/7$ of its elements are not in proper subfields. For $k = 4$, the result holds if $4d \leq p^2 - 1$, which is always true as 4 divides $p^2 + 1$. The $k = 5$ and 6 cases are trivial to check. So suppose that $k \geq 7$. Then the number of elements in proper subfields of F is at most

$$p^2 + p^4 + \cdots + p^{\lfloor 2k/3 \rfloor} \leq \frac{p^{2k/3+1}-1}{p-1} \leq \frac{p^k-1}{(p-1)^3}.$$

This is less than one half of the elements of C. This concludes the proof of (ii).

Since every subfield contains -1, the same arguments works in the quotient by $C \cap \{\pm 1\}$. □

2.4 Products of involutions in classical groups

We consider first the symplectic groups $\mathrm{PSp}_{2n}(q)$, since they are the easiest groups to deal with. The extra complications in the other cases will be less confusing if we have first seen the basic ideas in action. We may assume $n \geq 2$.

Theorem 17. *If $n \geq 2$ and $G \cong \mathrm{PSp}_{2n}(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that $z.z^g$ has odd order, as g ranges over the elements of G, is bounded below by n^{-1} times a positive constant, independent of q and n.*

Proof. First consider the involutions in $\mathrm{PSp}_{2n}(q)$ which lift to elements of order 4 in $\mathrm{Sp}_{2n}(q)$. The centraliser C of such an involution lifts to $\mathrm{GL}_n(q).2$ or $\mathrm{GU}_n(q).2$, according as $q^n \equiv 1$ or $3 \bmod 4$. Certainly $\mathrm{Sp}_{2n}(q) > \mathrm{Sp}_2(q^n) \cong \mathrm{SL}_2(q^n)$, which contains maximal tori of orders $q^n \pm 1$ each inverted by such elements of order 4. Moreover one of these two tori has twice odd order, so maps to a torus T of odd order $(q^n \pm 1)/2$ in G. We need to estimate the proportion of elements of T which have centraliser T, that is, the proportion of elements of T which are regular semisimple elements. In the case $T \cong C_{(q^n-1)/2}$, such an element lifts to

an element of the form $\mathrm{diag}(\lambda, \lambda^{-1})$ with $\lambda \in \mathrm{GF}(q^n)$, and it is sufficient that λ lies in no proper subfield. Lemma 16 shows that the proportion of such λ is at least $1/2$, for any value of q. In the case $T \cong C_{(q^n+1)/2}$, we may lift T to the subgroup $\mathrm{GU}_1(q^n)$ of unitary elements of $\mathrm{GF}(q^{2n})$. In this case Lemma 16 yields the same conclusion. Hence the number of pairs of involutions in this class with product of odd order is at least

$$\frac{1}{4}.\frac{|G|}{|N_G(T)|}.|T|^2 = \frac{1}{8n}.|G|.|T|,$$

since $N_G(T)/T \cong C_{2n}$, so the proportion of such pairs is at least

$$\frac{1}{8n}.\frac{|T|}{|G|}.|C|^2 \geq \frac{1}{8n}.\frac{\frac{1}{4}q^n.\frac{1}{2}q^{n^2}.\frac{1}{2}q^{n^2}}{q^{n(2n+1)}} \geq \frac{1}{128n}$$

by Corollary 15.

Next consider the involutions with centraliser $C \cong \mathrm{Sp}_{2k}(q) \times \mathrm{Sp}_{2n-2k}(q)$, and $2k < n$. (The same argument applies in the case $2k = n$, except that there is an extra factor of 2 in the order of the centraliser, which affects the constants.) Provided $q^k \neq 3$, we can choose a torus T of twice odd order, $q^k \pm 1$, with centraliser lifting to $\mathrm{GL}_2(q^k) \times \mathrm{Sp}_{2n-4k}(q)$ or $\mathrm{GU}_2(q^k) \times \mathrm{Sp}_{2n-4k}(q)$, which, since $(q^k \pm 1)/2$ is odd, can also be expressed as $C_{(q^k\pm1)/2} \times \mathrm{SL}_2(q^k).2 \times \mathrm{Sp}_{2n-4k}(q)$. We need to estimate the numbers of elements of T which have the same centraliser as T. It is sufficient that these elements of T, regarded as a subgroup of the multiplicative group of the field $\mathrm{GF}(q^k)$ or $\mathrm{GF}(q^{2k})$, should not be ±1 and not lie in any proper subfield, so by Lemma 16 at least half of the elements of T have the same centraliser as T. Now inside $\mathrm{Sp}_4(q^k) \leq \mathrm{Sp}_{4k}(q)$ there is a subgroup $D_{q^k\pm1} \times \mathrm{SL}_2(q^k).2$, containing involutions inverting T. These involutions negate a $2k$-space over $\mathrm{GF}(q)$, so belong to our chosen conjugacy class. The number of pairs of such involutions in $N_G(T)$ with odd order product whose centraliser is contained in $C_G(T)$ is at least $\frac{1}{4}|T|^2$, so by the same argument as before the proportion of such pairs in G is at least $|T|^2|C|^2/4|N_G(T)|.|G|$. Using the fact that $N_G(T) = D_{q^k\pm1}.k \times \mathrm{SL}_2(q^k).2 \times \mathrm{Sp}_{2n-4k}(q)$, which, by Corollary 15, has order less than $4kq^{k+3k+(n-2k)(2n-4k+1)}$, we find that the proportion is at least $1/128k$, which is at least $1/128n$.

When $q = 3$ and $k = 1$, there are no non-trivial odd-order elements in T and we need to modify the argument slightly. But we only need an asymptotic result as $n \to \infty$. Each of the involutions is defined by a non-singular 2-space in the ambient $2n$-dimensional symplectic space, and almost all pairs of 2-spaces span a 4-space, which in this case may be either non-singular or singular (with 2-dimensional radical). A straightforward counting argument shows that the latter case occurs with probability $8/27$ in the limit as $n \to \infty$, and it is easy to see that the corresponding involutions have product of order 3 in this case. This concludes the proof. □

Remark 18. The almost simple group $\mathrm{PSp}_{2n}(q).2$ obtained by adjoining diagonal automorphisms has just one additional class of involutions, for n even only, with centraliser $2 \times \mathrm{PSp}_n(q^2)$.

Next we consider the orthogonal groups. In this case we shall not calculate the constants explicitly. We acquire a finite number of factors of 2 by employing the estimates in Corollary 15, and by taking a torus in $\mathrm{P}\Omega_n(q)$ rather than $\mathrm{O}_n(q)$. We also acquire factors of 2 when we use Lemma 16 to estimate the proportion of elements in our chosen torus which have the stated centraliser. We shall show that the powers of q cancel out by using a dimension-counting argument just as in Section 2.2. The only other contribution to our estimates is a factor of $|N_G(T)/C_G(T)|$ in the denominator, which will always be bounded by $1/n$ where n is the dimension of the natural module.

Theorem 19. *If $n \geq 7$, and $G \cong \mathrm{P}\Omega_n^\varepsilon(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that $z.z^g$ has odd order, as g ranges over the elements of G, is bounded below by n^{-1} times a positive constant.*

Proof. The main part of the proof is the dimension-counting argument. The orthogonal groups $\mathrm{O}_n(q)$ all have dimension $n(n-1)/2$ and both conjugacy classes of involutions of type $\pm(-1^k 1^{n-k})$ (with $k \leq n/2$) have centralisers of shape $\mathrm{O}_k(q) \times \mathrm{O}_{n-k}(q)$ and dimension $k(k-1)/2 + (n-k)(n-k-1)/2$. Hence $\operatorname{codim} C_G(z) = k(n-k)$. Now consider the subgroups of shape $\mathrm{O}_2^\pm(q^k) \times \mathrm{O}_{n-2k}(q)$. Both possible factors $\mathrm{O}_2^\pm(q^k) \cong D_{2(q^k \mp 1)}$ contain involutions of both classes, inverting the cyclic subgroup T of order $q^k \mp 1$. We can choose the sign here, so that T has twice odd order, except when $q^k = 3$, which case is treated separately below. Now the centraliser of T in $\mathrm{O}_n(q)$ is $T \times \mathrm{O}_{n-2k}(q)$ and has dimension $k + (n-2k)(n-2k)/2$ and codimension $2k(n-k-1)$. Again we calculate $2 \dim T + \operatorname{codim} C_G(T) = 2k(n-k) = 2\operatorname{codim} C_G(z)$ as required.

The remaining involutions in the simple groups of orthogonal type lift to elements of order 4 in $\Omega_{2m}(q)$, squaring to -1. If m is even, there are two classes of such elements, fused in the full orthogonal group, which has + type. Thus we may consider just one of these two classes. On the other hand, if m is odd, such elements exist in $\Omega_{2m}^\varepsilon(q)$ just when $q \equiv \varepsilon \bmod 8$. In all cases, the involution centraliser lifts to $\mathrm{GL}_m(q).2$ if $q \equiv 1 \bmod 4$ and to $\mathrm{GU}_m(q).2$ if $q \equiv 3 \bmod 4$.

We consider first the cases when m is even, say $m = 2k$. Then the group $\Omega_{4k}^+(q)$ contains $\Omega_4^+(q^k)$, which is of shape $\mathrm{SL}_2(q^k) \circ \mathrm{SL}_2(q^k)$. The elements of order 4 in one of the two factors $\mathrm{SL}_2(q^k)$ square to -1, so are of the correct type. Indeed, the two factors contain elements from different conjugacy classes, and we may take whichever one we like. We take a torus T of order $\frac{1}{2}(q^k \pm 1)$, whichever has odd order, and find its centraliser is $T \times \mathrm{SL}_2(q^k)$. The dimension-counting argument now gives $\dim T + \operatorname{codim} C_G(T) = 2k + 2k(4k-1) - 4k = 2k(4k-2)$ while $\operatorname{codim} C_G(z) = 2k(4k-1) - 4k^2 = 2k(2k-1)$ and therefore $\dim T + \operatorname{codim} C_G(T) = 2\operatorname{codim} C_G(z)$ as required.

In the case when m is odd, say $m = 2k+1$, we apply almost the same argument inside $\mathrm{O}_2^\varepsilon(q) \times \mathrm{O}_{4k}^+(q)$. The torus still has dimension k, but this time its centraliser has dimension $4k+1$. Similarly G has dimension $(2k+1)(4k+1)$ and $C_G(z)$ has dimension $(2k+1)^2$, and again it is easy to check the required condition.

Secondly we need to check the five sources of constants listed in Section 2.1. First, $N_G(T)/C_G(T)$ has order $2k \leq n$. Cases (ii), (iii) and (v) have already been dealt with. So it remains to check that the proportion of elements of T whose centraliser is the same as that of T is at least a positive constant. The arguments are similar to those used in the proof of Theorem 17. Lemma 16 implies that there are enough such elements in T, except in the case $k = 1$ and $q = 3$ or 5.

If $q = 3$ or 5, and $k = 1$, we need a separate argument. In this case the involutions are (negatives of) reflections, in vectors of specified non-zero norm, and two distinct such reflections have product of order q if and only if the corresponding vectors fail to be orthogonal. Asymptotically, as $n \to \infty$, there are roughly q^{n-1} vectors of each norm, and roughly q^{n-2} of these are orthogonal to a given one. Therefore fewer than half the vectors of any given norm are orthogonal to a fixed non-isotropic vector. This concludes the proof. □

Remark 20. This argument actually proves the corresponding result for the simple orthogonal group extended by all diagonal and graph automorphisms, except for the involutions with centraliser containing $2 \times \mathrm{P}\Omega_{n/2}(q^2)$.

Next we consider the linear groups. For simplicity we work first in $\mathrm{PGL}_n(q)$, and then deduce the required result for $\mathrm{PSL}_n(q)$.

Theorem 21. *If $n \geq 2$ and $G \cong \mathrm{PGL}_n(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that $z.z^g$ has odd order, as g ranges over the elements of G, is bounded below by n^{-1} times a positive constant.*

Proof. Consider first the case when n is even and z is an involution of type $t_{n/2}$ or $t'_{n/2}$ in the notation of [13]. Write $n = 2m$ for convenience, and let ζ be a pre-image of z in $\mathrm{GL}_{2m}(q)$. If z is of type t_m then ζ can be chosen of order 2, and its centraliser in $\mathrm{GL}_{2m}(q)$ is $\mathrm{GL}_m(q) \times \mathrm{GL}_m(q)$. But ζ is conjugate to its negative in $\mathrm{GL}_{2m}(q)$, so the centraliser of z in $\mathrm{PGL}_{2m}(q)$ has shape $C_{q-1}.(\mathrm{PGL}_m(q) \wr 2)$. On the other hand, if z is of type t'_m, then ζ squares to a scalar of order containing the full 2-part of the centre of $\mathrm{SL}_{2m}(q)$, and the centraliser of z in $\mathrm{PGL}_n(q)$ has shape $C_{q+1}.\mathrm{PGL}_m(q^2).2$. Now $\mathrm{GL}_{2m}(q)$ contains $\mathrm{GL}_2(q^m)$, which contains elements mapping modulo the scalars of $\mathrm{GF}(q)$ to involutions of both types. In particular, the element $\mathrm{diag}(1,-1)$ in $\mathrm{GL}_2(q^m)$ maps to an involution of type t_m. Moreover, the subgroup $\mathrm{GL}_2(q^m)/C_{q-1}$ of $\mathrm{PGL}_{2m}(q)$ has shape $(C_{(q^m-1)/(q-1)} \times \mathrm{PSL}_2(q^m)).2$ and contains two classes of involutions, one of which squares to the full 2-part of the central C_{q-1} of scalars, so is of type t'_m. In some cases, depending

on m and q, one is inner and one is outer in the quotient $\mathrm{PGL}_2(q^m)$, while in the other cases, both map to involutions in $\mathrm{PSL}_2(q^m)$.

In any case, there are dihedral groups of order $q^m \pm 1$ in $\mathrm{PGL}_2(q^m)$ generated by involutions in the chosen class t_m or t'_m. Lifting to $\mathrm{GL}_2(q^m)$, the corresponding group has centraliser of order $2(q^m - 1)$. Adjoining the field automorphisms of order m we obtain the full normaliser, which has order $2(q^m - 1).2(q^m \pm 1).m$. (This contributes a factor $1/m$ to the calculation, and gives rise to the factor $1/n$ in the statement of the theorem.) We may choose the sign so that $(q^m \pm 1)/2$ is odd and greater than 1 except when $q^m = 3$, and use Lemma 16 to show that there are always enough elements in this cyclic group whose centraliser is no bigger than the centraliser of the whole group. Thus the dimension-counting argument using Corollary 15 finishes the proof: $\dim T = m$ and $\dim C_G(T) = 2m - 1$, so $\operatorname{codim} C_G(T) = 4m^2 - 2m$ and $2\dim T + \operatorname{codim} C_G(T) = 4m^2$; on the other hand, $C_G(z)$ has dimension $2m^2 - 1$, so has codimension $2m^2$.

The proof for the remaining classes t_m in the general case of $\mathrm{PGL}_n(q)$ where $n > 2m$ is just a modification of the above argument. We use the dihedral groups of order $2(q^m \pm 1)$ in $\mathrm{GL}_2(q^m)$, and just recalculate the dimensions of the centralisers. Working in $\mathrm{GL}_n(q)$ for simplicity, the centraliser of z is $\mathrm{GL}_m(q) \times \mathrm{GL}_{n-m}(q)$, which has codimension $n^2 - m^2 - (n-m)^2 = 2m(n-m)$, while the centraliser of T is $T.C_{q^m-1} \times \mathrm{GL}_{n-2m}(q)$ which has codimension $n^2 - 2m - (n-2m)^2 = 4m(n-m) - 2m$. Hence $2\dim T + \operatorname{codim} C_G(T) = 4m(n-m) = 2\operatorname{codim} C_G(z)$ as required. In all cases we note that $|N_G(T)/C_G(T)| = 2m < n$.

Finally we consider the case when $m = 1$ and $q = 3$. The involution centraliser in $\mathrm{GL}_n(3)$ is $C_2 \times \mathrm{GL}_{n-1}(3)$, and two such involutions can have product of order 3. Since the sum of the 1-dimensional eigenspaces of the involutions is a 2-space fixed by the D_6, and the intersection of the $(n-1)$-dimensional eigenspaces is an $(n-2)$-space fixed by the D_6, it follows easily that the centraliser of this D_6 is $2 \times \mathrm{GL}_{n-2}(3)$, which has index bounded above and below by constants times 3^{4n}. Also, the centraliser of the involution has index bounded by constants times 3^{2n}, so the usual counting argument works. Calculating suitable constants is left as an exercise for the interested reader. □

Corollary 22. *If $\mathrm{PSL}_n(q) \leq G \leq \mathrm{PGL}_n(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that zz^g has odd order, as g ranges over the elements of G, is at least n^{-1} times a positive constant.*

Proof. The G-class of z is the same as the $\mathrm{PGL}_n(q)$-class of z. □

We note that the bound in Theorem 21 and Corollary 22 is best possible, essentially because there are groups $\mathrm{PSL}_n(q)$ in which the Singer cycles contain almost all the odd-order elements of the group. To see this, let p be any odd prime, and a any (large) positive integer (so that 2^a is large compared to n, say), and let $q = p^{2^{a-1}}$ so that $q - 1$ is divisible by 2^a. Now all the maximal tori

except the Singer cycle have order divisible by $q-1$, so by a large power of 2. In particular, the proportion of odd-order elements which lie outside the Singer cycles is at most $1/2^a$.

Now pick an involution z in the class $t_{n/2}$ or $t'_{n/2}$. The dimension-counting argument shows that almost all the elements inverted by z are regular semisimple elements. Moreover, the numbers of such elements are essentially determined by the normaliser of the maximal torus they are contained in. The above remarks show that in the given cases almost all the regular semisimple elements of odd order lie in the Singer cycle. But these occur (asymptotically, as $q \to \infty$) at most a proportion $1/n$ of the time. Thus our bound is best possible, as claimed.

The proof of our main theorem for the unitary groups is an easy modification of the proof for the linear groups.

Theorem 23. *If $n \geq 3$ and $G \cong \mathrm{PGU}_n(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that $z.z^g$ has odd order, as g ranges over the elements of G, is bounded below by n^{-1} times a positive constant.*

Proof. We begin with the case $n = 2m$, and an involution of type t_m or t'_m. An involution of type t_m has centraliser which lifts to $\mathrm{GU}_m(q) \wr 2$ in $\mathrm{GU}_{2m}(q)$, while the centraliser of an involution of type t'_m lifts to $\mathrm{GL}_m(q^2).2$. In both cases the codimension in $\mathrm{GU}_{2m}(q)$ is $(2m)^2 - 2m^2 = 2m^2$. Now to find a suitable torus, look inside the subgroup $\mathrm{GU}_2(q^m)$ if m is odd, or inside $\mathrm{GL}_2(q^m) \leq \mathrm{GL}_m(q^2)$ if m is even. The same argument as in $\mathrm{GL}_{2m}(q)$ goes through with a few sign-changes, which only affect the constant.

For the involutions of type t_m in $\mathrm{PGU}_n(q)$ for $n > 2m$, the centralisers differ slightly from the $\mathrm{PGL}_n(q)$ case, but the dimensions are always the same. The only case where the argument breaks down is the case $m = 1$, $q = 3$, and a similar fix works as before. Indeed, in $\mathrm{PGU}_n(3)$ there are asymptotically 3^{2n-1} vectors of each norm (in GF(3)), and 3^{2n-2} of these have given inner product with a fixed one. Since the order of the product of the corresponding reflections depends only on the norms and the inner product of the reflecting vectors, at least a constant proportion of the products have order 3. □

Corollary 24. *If $\mathrm{PSU}_n(q) \leq G \leq \mathrm{PGU}_n(q)$, and $z \in G$ is an involution, then the proportion of pairs (z, z^g) such that zz^g has odd order, as g ranges over the elements of G, is at least n^{-1} times a positive constant.*

Proof. The G-class of z is the same as the $\mathrm{PGU}_n(q)$-class of z. □

Finally we put together Theorems 17, 19, and Corollaries 22 and 24 to obtain our main Theorem 2.

3 Finding involutions

3.1 General strategy

The basic method for finding an involution is as described in [14]: take a supply of (pseudorandom) elements of even order in the group, and power them up to involutions. Two problems may arise, however. The first is that there may not be enough elements of even order in the group, so that one does not find an element of even order after a polynomial number of attempts. The second is that the resulting involutions are not necessarily (nearly) uniformly distributed.

The first problem appears intractable in the general 'black box' context. For in a group of Lie type in characteristic 2, the proportion of elements of even order is bounded above (and below) by a constant times q^{-1}, where q is the order of the field of definition. Thus the time taken to find an element of even order by random search is proportional to q, whereas the input size may be only $O(\log q)$. If we are to have any hope of obtaining a polynomial time algorithm, therefore, we must assume that G is not a Lie type group defined in characteristic 2, and we do this from now on.

We choose to solve the second problem by making it harder: we seek an algorithm which returns a (pseudorandom) involution in a specified conjugacy class. Our aim is to show that if G is a simple group of Lie type, with Lie rank r, defined over a field of odd order, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements in G which power up to an element of $\mathcal{C}$ is at least $c'r^{-c}$, where c and c' are positive constants. Our method is to choose a suitable maximal torus in G and estimate the proportion of its elements that are regular semisimple elements and power up to an element of $\mathcal{C}$. The following easy counting lemma implies that at least half of the regular semisimple elements in any cyclic torus of even order power up to the involution, and that this remains true in any subgroup or quotient.

Lemma 25. *Suppose that G is a Lie type group, $T \leq G$ is maximal torus and $T_0 \leq T$. If T_0 is cyclic and $|T_0| = 2^a.b$ where b is odd, then at least one half of the regular semisimple elements of T_0 have order divisible by 2^a.*

Proof. Suppose that r is the number of regular semisimple elements in T_0 and that r_i is the number of these elements which have order divisible by exactly 2^i where $0 \leq i \leq a$. If $a = 0$ then there is nothing to prove. So assume that $a > 0$. Let z be the unique involution in T_0. Now, if x is an odd order regular semisimple element, then zx is also regular semisimple as $T = C_G(x) = C_G(zx)$. Thus $r_0 \leq r_1$. Now for each regular semisimple element x of even order $2^c.d$ with $c < a$, there are precisely two elements of T which square to x, and these elements are also regular semisimple. Therefore we have $2r_1 \leq r_2, \ldots, 2r_{a-1} \leq r_a$. Since $r = r_0 + r_1 + \cdots + r_a$, we infer that $r_a \geq r/2$. □

For the exceptional groups, r is bounded, so $c'r^{-c}$ is effectively a constant. For the classical groups, we show that we can take $c = 3$, which appears to be best possible for the linear and unitary groups. This can be improved to $c = 2$ for the symplectic and orthogonal groups. On the other hand, $c \geq 1$, since the number of conjugacy classes of involutions is linear in r. It is therefore not clear to us whether our lower bounds are best possible. [Certainly better bounds are attainable for restricted classes of involutions, as the proofs below will make clear. The question is whether better bounds can be found for all classes of involutions simultaneously.]

The following lemma explains our strategy: we choose a maximal torus T and estimate (i) the proportion of elements of T which power up to the desired involution, and (ii) the index of T in its normaliser.

Lemma 26. *Let G be a Lie type group, T a maximal torus in G and $\mathcal{C}$ be a conjugacy class of G. Assume that at least a proportion k of the regular semisimple elements of T power to a member of $\mathcal{C}$. Then at least a proportion $k/[N_G(T) : T]$ of the elements of G power to an element of $\mathcal{C}$.*

Proof. Since the regular semisimple elements of T lie in a unique conjugate of T, $\bigcup_{g \in G} T^g$ contains at least $k|T||G|/|N_G(T)|$ elements which power to an element of $\mathcal{C}$. Hence at least a proportion $k/[N_G(T) : T]$ of the elements of G power to an element of $\mathcal{C}$. □

We also silently use the following easy observation which allows us to calculate in universal groups rather than their simple quotients.

Lemma 27. *Suppose that G is a group, N is a normal subgroup of G and Y is a subset of G. Let X be the set of elements of G which power to elements of Y. Write $\overline{G} = G/N$ and, for subsets Z of G, write $\overline{Z} = \{zN \mid z \in Z\}$. Then $\overline{X}$ consists of elements of $\overline{G}$ which power to elements of $\overline{Y}$ and $|\overline{X}|/|\overline{G}| \geq |X|/|G|$.*

3.2 Elements of even order in odd characteristic exceptional groups of Lie type

In Table 1 we list the shapes of the centralisers of involutions z in the finite simple exceptional groups of Lie type defined over fields of odd order, as well as our choice of maximal torus containing z. The information about centralisers comes from [13], and the information about shapes of tori comes from [15].

For most classes of involutions in the exceptional groups we can choose a *cyclic* maximal torus T containing an involution of the chosen class. Therefore by Lemma 25 at least half of the regular semisimple elements in this torus power up to the involution. Also, the number of regular semisimple elements in the torus is (at least) a monic polynomial in q of degree r. (More precisely, it is given by one of a finite number of such polynomials, depending on certain congruences:

Table 1: Involution centralisers in simple exceptional groups, q odd

Group	Involution centraliser	Conditions	Maximal torus
${}^3D_4(q)$	$(\mathrm{SL}_2(q) \circ \mathrm{SL}_2(q^3)).2$		$C_{(q-1)(q^3+1)}$
$G_2(q)$	$(\mathrm{SL}_2(q) \circ \mathrm{SL}_2(q)).2$		C_{q^2-1}
${}^2G_2(q)$	$C_2 \times \mathrm{PSL}_2(q)$		C_{q-1}
$F_4(q)$	$(\mathrm{SL}_2(q) \circ \mathrm{Sp}_6(q)).2$		$C_{(q-1)(q^3+1)}$
	$2.\mathrm{P}\Omega_9(q)$		C_{q^4+1}
$E_6(q)$	$(C_{(q-1)/3} \circ 4.\mathrm{P}\Omega^+_{10}(q)).4$	$q \equiv 1 \pmod{12}$	$C_{q-1} \circ_3 C_{q^5-1}$
	$(C_{q-1} \circ 4.\mathrm{P}\Omega^+_{10}(q)).4$	$q \equiv 5 \pmod{12}$	$C_{q-1} \times C_{q^5-1}$
	$(C_{(q-1)/3} \circ 2.\mathrm{P}\Omega^+_{10}(q)).2$	$q \equiv 7 \pmod{12}$	$C_{q-1} \circ_3 C_{q^5-1}$
	$(C_{q-1} \circ 2.\mathrm{P}\Omega^+_{10}(q)).2$	$q \equiv 11 \pmod{12}$	$C_{q-1} \times C_{q^5-1}$
	$2.(\mathrm{PSL}_2(q) \times \mathrm{PSL}_6(q)).2$	$d = (3, q-1)$	$C_{(q+1)} \circ_d C_{(q^2+q+1)(q^3+1)}$
${}^2E_6(q)$	$(C_{q+1} \circ 2.\mathrm{P}\Omega^-_{10}(q)).2$	$q \equiv 1 \pmod{12}$	$C_{q+1} \times C_{q^5+1}$
	$(C_{(q+1)/3} \circ 2.\mathrm{P}\Omega^-_{10}(q)).2$	$q \equiv 5 \pmod{12}$	$C_{q+1} \circ_3 C_{q^5+1}$
	$(C_{q+1} \circ 4.\mathrm{P}\Omega^-_{10}(q)).4$	$q \equiv 7 \pmod{12}$	$C_{q+1} \times C_{q^5+1}$
	$(C_{(q+1)/3} \circ 4.\mathrm{P}\Omega^-_{10}(q)).4$	$q \equiv 11 \pmod{12}$	$C_{q+1} \circ_3 C_{q^5+1}$
	$2.(\mathrm{PSL}_2(q) \times \mathrm{PSU}_6(q)).2$	$d = (3, q+1)$	$C_{q+1} \circ_d C_{(q^2+q+1)(q^3+1)}$
$E_7(q)$	$(\mathrm{SL}_2(q) \circ 2.\mathrm{P}\Omega^+_{12}(q).2).2$	$q \equiv 1 \pmod{4}$	$C_{q+1} \circ_2 C_{q^6-1}$
	$(\mathrm{SL}_2(q) \circ 2.\mathrm{P}\Omega^+_{12}(q).2).2$	$q \equiv 3 \pmod{4}$	$C_{q-1} \circ_2 C_{q^6-1}$
	$2.\mathrm{PSL}_8(q).4.2$	$q \equiv 1 \pmod{8}$	$C_{(q^7-1)/2}$
	$2 \times \mathrm{PSU}_8(q).2.2$	$q \equiv 3 \pmod{8}$	$C_{(q^7+1)/2}$
	$2 \times \mathrm{PSL}_8(q).2.2$	$q \equiv 5 \pmod{8}$	$C_{(q^7-1)/2}$
	$2.\mathrm{PSU}_8(q).4.2$	$q \equiv 7 \pmod{8}$	$C_{(q^7+1)/2}$
	$(3.E_6(q) \circ C_{(q-1)/2}).S_3$	$q \equiv 1 \pmod{12}$	$C_{(q^6+q^3+1)(q-1)/2}$
	$(E_6(q) \circ C_{(q-1)/2}).2$	$q \equiv 5 \pmod{12}$	$C_{(q^6+q^3+1)(q-1)/2}$
	$({}^2E_6(q) \circ C_{(q+1)/2}).2$	$q \equiv 7 \pmod{12}$	$C_{(q^6-q^3+1)(q+1)/2}$
	$(3.{}^2E_6(q) \circ C_{(q+1)/2}).S_3$	$q \equiv 11 \pmod{12}$	$C_{(q^6-q^3+1)(q+1)/2}$
$E_8(q)$	$2.\mathrm{P}\Omega^+_{16}(q).2$		C_{q^8-1}
	$(\mathrm{SL}_2(q) \circ 2.E_7(q)).2$		$C_{(q+1)(q^7-1)}$

Note: the notation $\circ_d$ means that the central product in which the subgroups C_d of the two factors are identified. Of course, as abstract groups these central products are isomorphic to direct products of smaller groups, but we use the central product notation to make it clear which elements of the torus are central in the universal group of Lie type.

see [12].) Thus, provided q is large enough, at least half of the elements in the torus are regular semisimple. It follows that the proportion of elements in G which power to an involution in the chosen class is at least $1/(4|N_G(T)/T|)$. But $N_G(T)/T$ is a subgroup of the Weyl group, so has bounded order. (A better bound can be obtained with more work, as we can calculate the precise subgroup $N_G(T)/T$ of the Weyl group in each case (see Carter [11, 3.6.5]).) This general method enables us to prove:

Theorem 28. *There is an absolute constant c such that if G is an exceptional simple group of Lie type, defined over the field $\mathrm{GF}(q)$ of odd order, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power into $\mathcal{C}$ is at least c.*

Proof. We put flesh on the bones of the above argument, starting with the case $G \cong {}^2G_2(q)$, where $q = 3^{2n+1}$. The involution centraliser in G has shape $2 \times \mathrm{PSL}_2(q)$, which contains a maximal torus $T \cong C_{q-1}$ which is cyclic of twice odd order. Precisely two elements of this torus are not regular semisimple. The normaliser of the torus is $D_{2(q-1)}$, so there are $|G|/2(q-1)$ such tori, each containing exactly $(q-3)/2$ regular semisimple elements of even order. Therefore the total number of such elements in G is $|G|.(q-3)/4(q-1)$, so the proportion of them in the group is $(q-3)/4(q-1) \geq 3/13$ since $q \geq 27$.

Essentially the same argument deals with the groups ${}^3D_4(q)$, $G_2(q)$, $F_4(q)$ and $E_8(q)$, as well as involutions of type t_4/t_4' and t_7/t_7' in $E_7(q)$. Since T is a maximal torus, $N_G(T)/T$ is a subgroup of the Weyl group, so has bounded order. Ignoring a finite number of groups if necessary, we may assume that at least half the elements of T are regular semisimple, so at least a quarter of the elements of T are regular semisimple of even order, so power to the desired involution.

This argument does not work directly in $E_7(q)$ with the involutions of type t_1, in the torus of shape $C_{q-1} \times C_{q^6-1}$ or $C_{q+1} \times C_{q^6-1}$. However, in this case the diagonal involution is central in $2{\cdot}E_7(q)$, and the chosen factor $C_{q\pm1}$ has twice odd order, so that the 2-part of the torus in the simple group $E_7(q)$ is cyclic. Therefore the proportion of regular semisimple elements in T which power to z is again at least one half, and the argument goes through.

We are left with involutions of type t_1 and t_2 in $E_6(q)$ and ${}^2E_6(q)$. In these cases the best maximal tori we could find are direct or central products of two cyclic groups, whose orders have a common factor of $q \pm 1$. There are now three involutions in the torus, only one of which is necessarily in our chosen conjugacy class. By inspection of Table 1 the 2-parts of the orders of the cyclic factors of T are equal. Hence, for each involution z in T, at least one quarter of the elements of T power up to z. Now we may choose q large enough so that at least $7/8$ of the elements in the torus are regular semisimple, whence at least $1/8$ of the regular semisimple elements in T power to our chosen involution. □

3.3 More counting arguments

In this section we prove results about the proportion of regular semisimple elements in certain specified cyclic maximal tori, which will be used in the proof of Theorem 5 for linear and unitary groups. Recall that a Singer cycle in $\mathrm{GL}_m(q)$ is just the subgroup $\mathrm{GL}_1(q^m)$ and this subgroup is cyclic of order q^m-1. We observe the convention that every element of $\mathrm{GL}_1(q)$ is regular semisimple, for our fixed value of q. Note that a subgroup of a Singer cycle in G contains a regular semisimple element if and only if it acts irreducibly on the natural module for $\mathrm{GL}_n(q)$.

Note also that scalars do not affect whether an element is regular semisimple, so all the results in this section hold also for the corresponding projective groups.

Lemma 29. *Suppose q is odd, $m \geq 1$ and $T \leq \mathrm{GL}_m(q)$ is a Singer cycle. Let $T^0 = T \cap \mathrm{SL}_m(q)$. Then in each coset of T^0 in T the proportion of regular semisimple elements is at least $1-2/q$. In particular, if $T \geq S \geq T^0$, then the proportion of the elements of S which are regular semisimple is at least $1-2/q$.*

Proof. The result trivially holds for $m=1$. If $m=2$, then there are exactly $q-1$ elements of T which are not regular semisimple and in each coset of T^0 there are at most two such elements. Thus the proportion of regular semisimple elements in each coset of T^0 is at least $(q+1-2)/(q+1) > 1-2/q$ and so the result holds for $m=2$. Suppose that $m \geq 3$. We have that T^0 has order $q^{m-1}+\cdots+1$. The elements of T which are not regular semisimple are contained in one of the cyclic subgroups of order dividing q^d-1 for some divisor d of m. The maximal such subgroups are those for which m/d is prime, so the number of elements of T which are not regular semisimple is at most $\sum_{m/d \text{ prime}}(q^d-1)$. Then, in a worst case situation, all these elements lie in the same coset of T^0 and so the number of regular semisimple elements in a coset of T^0 is at least

$$\sum_{j=0}^{m-1} q^j - \sum_{m/d \text{ prime}} (q^d-1) \geq q^{m-1}.$$

Therefore the proportion of regular semisimple elements in this coset is at least

$$\frac{q^{m-1}(q-1)}{(q^m-1)} = 1 - \frac{q^{m-1}-1}{q^m-1} > 1 - \frac{1}{q}$$

as required. □

We now answer the same question for Singer cycles of $\mathrm{GU}_m(q)$. These subgroups come from the embedding of $\mathrm{GU}_1(q^m)$ into $\mathrm{GU}_m(q)$ when m is odd and of $\mathrm{GL}_1(q^m)$ into $\mathrm{GU}_m(q)$ when m is even. Thus a Singer cycle of $\mathrm{GU}_m(q)$ is a cyclic group of order $q^m-(-1^m)$. As in the linear case we consider every element of $\mathrm{GU}_1(q)$ to be regular semisimple.

Lemma 30. *Suppose that q is odd, $m \geq 1$ and $T \leq \mathrm{GU}_m(q)$ is a Singer cycle. Let $T^0 = T \cap \mathrm{SU}_m(q)$. Then the proportion of regular semisimple elements in each coset of T^0 in T is at least $1 - 2/q$ if $m \neq 2$ and at least $1 - 3/q$ if $m = 2$. In particular, if $T \geq S \geq T^0$, then the proportion of the elements of S which are regular semisimple is at least $1 - 2/q$ if $m \neq 2$ and at least $1 - 3/q$ if $m = 2$.*

Proof. For $m = 1$, all the elements of T are regular semisimple. So we may assume that $m \geq 2$. For $m = 2$, each coset of T^0 contains at most two elements which are not regular semisimple. Thus each coset of T^0 contains at least $(q-1)-2$ regular semisimple elements. Hence the result holds when $m = 2$.

Assume that $m \geq 3$. Note that

$$|T^0| = \frac{q^m - (-1^m)}{q+1} = q^{m-1} - q^{m-2} + \cdots - (-1)^m.$$

The elements of T which are not regular semisimple are contained in cyclic subgroups of order $q^d - (-1)^{m/d}$, where d divides m. Hence the number of elements of each coset of T^0 in T which are not regular semisimple is at most

$$\begin{aligned} 1 + \sum_{m/d \text{ prime}} q^d &\leq 1 + q + q^2 + \cdots + q^{m-2} \\ &= \frac{q^{m-1} - 1}{q-1} \\ &= \frac{q+1}{q(q-1)} . \frac{q^m - q}{q+1} \\ &\leq \frac{q+1}{q(q-1)} . \left(\frac{q^m - (-1)^m}{q+1} \right). \end{aligned}$$

Hence the proportion of elements of each coset of T^0 in T which are not regular semisimple is at most $(q+1)/(q(q-1)) \leq 2/q$ since $q \geq 3$. □

In fact in $\mathrm{GU}_2(3)$, T^0 is central and so contains no regular semisimple elements.

The next two results describe the proportions of elements in a Singer cycle whose *square* is regular semisimple.

Lemma 31. *Suppose that q is odd, $m \geq 2$, and $T \leq \mathrm{GL}_m(q)$ is a Singer cycle. Assume that if $m = 2$, then $q > 3$. Set $T^0 = T \cap \mathrm{SL}_m(q)$.*

(i) The proportion of elements in each coset of T^0 in T whose square is regular semisimple is at least $1 - \frac{2}{q}$ if $m > 2$ and at least $1 - \frac{4}{q}$ if $m = 2$.

(ii) Assume that $|T^0| = 2^a b$ where b is odd and xT^0 is an odd-order element of T/T^0. Then the proportion of regular semisimple elements in xT^0 which have order not divisible by 2^a is at least $\frac{1}{2}(1 - \frac{2}{q})$ if $m \neq 2$ and at least $\frac{1}{2}(1 - \frac{4}{q})$ if $m = 2$.

Proof. Since $m \geq 2$, the number of elements whose square is not regular semisimple is at most $2\sum_{m/d \text{ prime}}(q^d - 1)$. Therefore the number of elements in each coset of T_0 in T whose square is regular semisimple is at least

$$\sum_{j=0}^{m-1} q^j - 2 \sum_{m/d \text{ prime}} (q^d - 1).$$

If integers e and $e+1$ both divide m, and m/e and $m/(e+1)$ are both primes, then $e = 2$ and $m = 6$. So for $m > 6$, we may remove the terms $q^{d+1} + q^d$ from $\sum_{j=0}^{m-1} q^j$ for every d with m/d prime. Hence

$$\sum_{j=0}^{m-1} q^j - 2 \sum_{m/d \text{ prime}} (q^d - 1) \geq q^{m-1}.$$

For $m = 6$, we have $(q^5 + \cdots + 1) - 2(q^3 - 1) - 2(q^2 - 1) \geq q^5$, since $2(q^3 + q^2) \leq q^4 + q^3 + q^2$. The same inequality holds when $m = 5$. For $m = 4$, the greatest common divisor of $q^3 + q^2 + q + 1$ and $2(q^2 - 1)$ is $2(q + 1)$ and so we obtain

$$(q^3 + q^2 + q + 1) - 2(q + 1) \geq q^3.$$

Since $|T^0|$ is odd when $m = 3$, each coset of T^0 contains at least q^2 elements whose square is regular semisimple. Finally, for $m = 2$ we have at least $q - 3$ elements whose square is regular semisimple. Thus for $m \geq 3$, we deduce that a proportion of $1 - 2/q$ of the elements of T^0 have regular semisimple square. For $m = 2$, as we have omitted $q = 3$, we have a proportion of $1 - \frac{4}{q+1} > 1 - \frac{4}{q}$. This proves (i).

To see that part (ii) holds, we simply square all the regular semisimple elements in $\sqrt{x}T^0$ whose square is regular semisimple and obtain the proportions described. □

We note that $\mathrm{SL}_2(3)$ has no regular semisimple elements in T^0 whose square is regular semisimple.

We need the analogous result to Lemma 31 for the unitary groups.

Lemma 32. *Suppose that q is odd, $m \geq 1$ and $T \leq \mathrm{GU}_m(q)$ is a Singer cycle and set $T^0 = T \cap \mathrm{SU}_m(q)$. Assume that $(m, q) \neq (2, 3)$ or $(2, 5)$. Then the following hold.*

(i) The proportion of elements in every coset of T^0 in T whose square is regular semisimple is at least $1 - \frac{2}{q}$ if $m \geq 3$ and at least $1 - \frac{5}{q}$ if $m = 2$.

(ii) Assume that $|T^0| = 2^a b$ where b is odd and xT^0 is an odd-order element of T/T^0. Then the proportion of regular semisimple elements in xT^0 which have order not divisible by 2^a is at least $\frac{1}{2}(1 - \frac{2}{q})$ if $m \neq 2$ and $\frac{1}{2}(1 - \frac{5}{q})$ if $m = 2$.

Proof. We follow the proof of Lemma 30. The generic case is when $m \geq 5$. In this case, the number of elements of T whose square is not regular semisimple is at most $2(q^{m-3} + \cdots + 1) \leq q^{m-2}$ as $q \geq 3$. It follows that in each coset of T^0 in T, a proportion of at most $\frac{(q+1)q^{m-2}}{q^{m-1}-(-1)^m} < \frac{2}{q}$ of the elements have a square which is not regular semisimple. Hence (i) holds when $m \geq 5$. For $m = 4$, the subgroup of T which contains all the elements which are not regular semisimple is cyclic of order $q^2 - 1$ and this subgroup intersects T^0 in a subgroup of order $2(q-1)$. Hence the proportion of elements of any coset of T^0 in T whose square is regular semisimple is at least

$$\frac{(q^3 - q^2 + q - 1) - 4(q-1)}{q^3 - q^2 + q - 1} = \frac{(q-1)(q^2-3)}{(q-1)(q^2+1)} > 1 - \frac{2}{q}.$$

For $m = 3$, we have that T^0 has at most 3 elements which are not regular semisimple. Since $|T^0|$ is odd, at most 3 elements of each coset of T^0 have square which is not regular semisimple. Hence the proportion of elements whose square is not regular semisimple in each coset is at greater than $1 - \frac{2}{q}$. Suppose finally that $m = 2$. Then T^0 has order $(q-1)$ and the elements which square to an element which is not regular semisimple are contained in the subgroup X of order $(q+1)(4, q-1)/2$. Thus $|X \cap T^0| \leq 4$. So each coset contains at most four elements whose square is not regular semisimple. Hence the proportion of elements of each coset of T^0 in T which square to regular semisimple elements is at least

$$\frac{(q-1)-4}{q-1} = 1 - \frac{4}{q-1} > 1 - \frac{5}{q}.$$

Thus (i) holds and part (ii) easily follows. □

Again we remark that there are no regular semisimple elements in T^0 whose square is regular semisimple in the case that $G \cong \mathrm{SU}_2(3)$ or $\mathrm{SU}_2(5)$.

The last two results in this section estimate the numbers of regular semisimple elements in the direct product of two Singer cycles in suitable subgroups.

Lemma 33. *Assume that $T_k \leq \mathrm{GL}_k(q)$ and $T_l \leq \mathrm{GL}_l(q)$ are Singer cycles and set $T = T_k \times T_l \leq \mathrm{GL}_k(q) \times \mathrm{GL}_l(q) \leq \mathrm{GL}_{k+l}(q)$. Suppose that $x \in T_k$ is regular semisimple in $\mathrm{GL}_k(q)$ and $y \in T_l$ is regular semisimple in $\mathrm{GL}_l(q)$. Then (x, y) is regular semisimple in $\mathrm{GL}_n(q)$ unless $k = l$ and the order x is equal to the order of y. Furthermore, in any case, if x is a given regular semisimple element in T_k, then the number of elements of $y \in T_l$ such that (x, y) is regular semisimple and $\det(x, y) = 1$ is at least $|T_l|(1 - \frac{2}{q})/(q-1) - n/2$.*

Proof. Set $n = k + l$. We consider $G = \mathrm{GL}_n(q)$ as a subgroup of $\widetilde{G} = \mathrm{GL}_n(F)$ where F is algebraically closed of characteristic p. Then the subgroup T of G can be diagonalised in $\widetilde{G}$. The elements of T which are a product of a regular

semisimple element x in $\mathrm{GL}_k(q)$ and a regular semisimple element y in $\mathrm{GL}_l(q)$ then have the form

$$(x,y) = d = \mathrm{diag}(\lambda, \lambda^q, \ldots, \lambda^{q^{k-1}}, \mu, \mu^q, \ldots, \mu^{q^{l-1}})$$

where λ has order dividing $q^k - 1$ and μ has order dividing $q^l - 1$. Furthermore, the fact that x and y are regular semisimple means that λ does not have order dividing $q^m - 1$ for any proper divisor m of k and that μ does not have order dividing by $q^m - 1$ for any proper divisor m of l. Now the condition for an element to be regular semisimple in G is that the diagonalisation in $\widetilde{G}$ has no repeated entry (no repeated eigenvalue). Now if $k \neq l$, then d certainly has this property. Thus, in this case, d is regular semisimple whenever x and y are. So assume that $k = l$. Once x is specified there are precisely k possible ways to choose y so that d is not regular. Using Lemma 29, this means that there are at least $|T_l|(1 - \frac{2}{q})/(q-1) - n/2$. $y \in T_l$ such that (x,y) is regular semisimple and has determinant 1. □

For the other classical groups G acting on natural module V we frequently use the fact that if an element x of G is regular semisimple in the supergroup $\mathrm{GL}(V)$, then it is certainly regular semisimple in G. Thus our check for being regular semisimple is the same as in Lemma 33: having distinct eigenvalues. Note that if n is even, then the eigenvalues of elements of the Singer cycle in $\mathrm{GU}_n(q)$ are $\lambda, \lambda^q, \ldots, \lambda^{q^{n-1}}, \lambda^{-1}, \lambda^{-q}, \ldots, \lambda^{-q^{n-1}}$. So we see that these elements are regular semisimple so long as $\lambda^{q^e} \neq \lambda$ for all $1 \leq e \leq n$ where $\lambda \in \mathrm{GF}(q^n)$. Arguing as in Lemma 33 we obtain:

Lemma 34. *Suppose that* $\mathrm{GU}_k(q) \times \mathrm{GU}_l(q) \leq \mathrm{GU}_n(q)$ *with* $k + l = n$. *Let* T_k *be a Singer cycle in* $\mathrm{GU}_k(q)$ *and* T_l *be a Singer cycle in* $\mathrm{GU}_l(q)$. *Assume that* x *regular semisimple in* T_k *with respect to* $\mathrm{GU}_k(q)$ *and* $y \in T_l$ *is regular semisimple with respect to* $\mathrm{GU}_l(q)$.

(i) If either $k \neq l$ *or* x *and* y *have different orders, then* (x,y) *is regular semisimple in* $\mathrm{GU}_n(q)$.

(ii) If $k = n/2$, *and* $x \in T_k$ *is fixed, then the number of* $y \in T_l$ *such that* $\det(x,y) = 1$ *is at least* $|T_l|(1 - \frac{2}{q})/(q-1) - n/2$ *if* $n \neq 4$ *and at least* $(q-1)(1 - \frac{3}{q}) - 2 = \frac{(q-1)(q-5)-2}{q}$ *if* $n = 4$.

□

Similar results hold for the orthogonal and the symplectic groups.

3.4 Elements of even order in classical groups of odd characteristic

For classical groups of bounded dimension we can use analogous arguments to those used for exceptional groups. However, if the dimension (or equivalently the

Lie rank) is unbounded, we must deal with all field orders $q \geq 3$. The problem is that, although the number of regular semisimple elements is given by a monic polynomial of degree r in q, for small values of q this polynomial might, a priori, evaluate to zero. This means we need much tighter control over the proportions of regular semisimple elements in our chosen tori.

Just as with the exceptional groups, we aim for cyclic maximal tori whenever possible. This has three advantages. First, they contain a unique involution. Second, most of their elements are regular semisimple. Third, $|N_G(T)/T|$ is small in these cases. Unfortunately this is not usually possible, and so we need a product of two cyclic tori instead.

For an integer m, we let m_2 be the largest power of 2 dividing m.

We begin by estimating the proportion of elements in $\mathrm{PSL}_n(q)$ which power to an element of a given conjugacy class of involutions. To do this we start by presenting a collection of elements of $\mathrm{SL}_n(q)$ which map to involutions of $\mathrm{PSL}_n(q)$. If $(q-1)_2 > n_2$, we let $\lambda \in \mathrm{GF}(q)$ be such that $\lambda^n = -1$ and we fix this element. Then, for $1 \leq m \leq n-1$, we define the following elements of $\mathrm{SL}_n(q)$,

$$\zeta_m = \begin{cases} \mathrm{diag}(-1^m, 1^{n-m}) & m \text{ even} \\ \mathrm{diag}(-\lambda^m, \lambda^{n-m}) & m \text{ odd}, (q-1)_2 > n_2 \end{cases}.$$

Also let

$$\zeta_m^* = \begin{cases} \mathrm{diag}(1^{n-m}, -1^m) & m \text{ even} \\ \mathrm{diag}(\lambda^{n-m}, -\lambda^m) & m \text{ odd}, (q-1)_2 > n_2 \end{cases}.$$

Then ζ_m and ζ_m^* are conjugate in $\mathrm{SL}_n(q)$ and $\zeta_{n-m}\zeta_m^*$ is central in $\mathrm{SL}_n(q)$. Hence the image of ζ_{n-m} and ζ_m in $\mathrm{PSL}_n(q)$ are conjugate and we may assume that $m \leq n/2$. We denote the conjugacy class of the image of ζ_m in $\mathrm{PSL}_n(q)$ by t_m as in [13, Table 4.5.1]. There is just one further class of involutions in $\mathrm{PSL}_n(q)$, and they occur only if n is even. The involutions in this class are images in $\mathrm{PSL}_n(q)$ of central elements of $\mathrm{GL}_{n/2}(q^2)$ embedded in $\mathrm{GL}_n(q)$. We denote their conjugacy class by $t'_{n/2}$.

Lemma 35. *If $G \cong \mathrm{PSL}_2(q)$, q odd, and $\mathcal{C}$ is the conjugacy class of involutions in G, then the proportion of elements of G which power up to an element of $\mathcal{C}$ is at least $\frac{1}{4}$.*

Proof. We choose $\varepsilon = \pm 1$ such that $q - \varepsilon \equiv 0 \pmod 4$. Let T be the torus of G of order $(q-\varepsilon)/2$. Then at least one half of the elements of T have even order. Now any two conjugates of T intersect trivially and there are $q(q+\varepsilon)/2$ conjugates of T. It follows that the proportion of elements of G which have even order is at least one quarter as claimed. □

Theorem 36. *If $G \cong \mathrm{PSL}_n(q)$, q odd, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power up to an element of $\mathcal{C}$ is at least c/n^3 where c is a positive constant.*

Proof. By Lemma 35 we may assume that $n > 2$. We may also suppose that if n is small then q is large. Set $G = \mathrm{SL}_n(q)$ and regard G as a subgroup of $\mathrm{GL}_n(q)$.

We first examine the involutions of $\mathrm{PSL}_n(q)$ that are images in $\mathrm{PSL}_n(q)$ of ζ_k with $1 \leq k \leq n-1$. Set $l = n-k$. Since the images of ζ_k and ζ_l are conjugate in $\mathrm{PSL}_n(q)$, it suffices to find a lower estimate of the number of elements of G which power to either one of them. Therefore, if either one of k or l is even we may as well assume that it is k. So either k and l are both odd or k is even. Note that we may well have $k > n/2$ with this choice. As indicated in its definition ζ_k is contained in the centre of the subgroup $X_{k,l} = \mathrm{GL}_k(q) \times \mathrm{GL}_l(q)$ of $\mathrm{GL}_n(q)$. Let T_k be a Singer cycle in the $\mathrm{GL}_k(q)$ factor of $X_{k,l}$ and T_l be a Singer cycle in the $\mathrm{GL}_l(q)$ factor of $X_{k,l}$. So T_k is a cyclic group of order $q^k - 1$ and T_l is a cyclic group of order $q^l - 1$. Furthermore, $\zeta_k \in T_k$. Set $T = T_k T_l \cong T_k \times T_l$. For a subgroup H of $\mathrm{GL}_n(q)$, define $H^0 = H \cap G$. Thus

$$T^0 = \{xy \in T_k T_l \mid \det x \det y = 1\}$$

and has order $(q^k - 1)(q^l - 1)/(q - 1)$.

We consider three different cases:

(i) $(q^k - 1)_2 \neq (q^l - 1)_2$,

(ii) $(q^k - 1)_2 = (q^l - 1)_2$ and $(q - 1)_2 > n_2$, and

(iii) $(q^k - 1)_2 = (q^l - 1)_2$ and $(q - 1)_2 \leq n_2$.

We first address case (i). Assume that $(q^k - 1)_2 \neq (q^l - 1)_2$ and note that in this case, k must be even for otherwise $(q^k - 1)_2 = (q^l - 1)_2 = (q - 1)_2$. In particular, ζ_k is an involution. Without loss of generality, we may assume that $(q^k - 1)_2 > (q^l - 1)_2$. We consider the subset $\mathcal{S}$ of T^0 consisting of elements xy such that x is regular semisimple in T_k with respect to $\mathrm{GL}_k(q)$, $|x|_2 = (q^k - 1)_2$ and y is regular semisimple in T_l with respect to $\mathrm{GL}_l(q)$. As x and y have different orders, their product is regular semisimple in $\mathrm{GL}_n(q)$ by Lemma 33 and of course such elements power to ζ_k. By Lemmas 25 and 29, there are at least $|T_k|\frac{1}{2}(1 - \frac{2}{q})$ choices for x. The element y must be chosen so that $\det y = \det x^{-1}$. This specifies a particular coset of T_l^0 in T_l from which we must choose y. Hence Lemma 29 implies there are at least $(1 - \frac{2}{q})|T_l^0| \geq 1$ choices for y. Thus in case (i), at least a proportion $\frac{1}{2}(1 - \frac{2}{q})^2 \geq \frac{1}{18}$ elements of T^0 are regular semisimple and power up to z_k. This completes the analysis of case (i).

Note now that if n is odd, then k is even and l is odd which means we are in case (i). Hence from now on we know n is even.

In case (ii), $(q^k - 1)_2 = (q^l - 1)_2$ and $(q - 1)_2 > n_2$. We consider the same set of elements $\mathcal{S}$ as above. There are two potential problems: the first is that xy is not regular semisimple in $\mathrm{GL}_n(q)$ and the second is that it does not power to an element of $z_k Z(G)$. For the first problem, we simply use Lemma 33 to initially

see that the problem arises only when $k = n/2$ and then to get that in this case the number of elements of $\mathcal{S}$ which are regular semisimple is at least

$$|T_k|\frac{1}{2}(1-\frac{2}{q})(|T_k^0|(1-\frac{2}{q})-\frac{n}{2}) \geq |T^0|\frac{1}{6}(\frac{1}{3}-\frac{n}{2|T_k^0|}) \geq \frac{|T^0|}{36}$$

as $\frac{n}{2|T_k^0|} \leq \frac{1}{6}$ for all $n \geq 8$ and q odd and for $n = 4$ and 6 so long as $q > 11$ (say). So, given that q is large whenever n is small, whatever k is, we have plenty of regular semisimple elements.

As n is even and $(q-1)_2 > n_2 \geq 2$ by assumption, $q \equiv 1 \pmod 4$. Consequently $(q^k-1)_2 = (q-1)_2 k_2$ and $(q^l-1)_2 = (q-1)_2 l_2$. In particular, as $(q^k-1)_2 = (q^l-1)_2$, we have $k_2 = l_2$ and $(q-1)_2 > n_2 > k_2$. Therefore, in both T_k and T_l, any element w of 2-power order has the property that $w^{k_2} = w^{l_2} \in Z(X_{k,l})$. This means that, for $xy \in \mathcal{S}$, the 2-part of $(xy)^{k_2} = x^{k_2}y^{k_2}$ is contained in the centre of $X_{k,l}$ and has 2-part of its order $(q-1)_2$. Since $(q-1)_2 > n_2$, the 2-part of $(xy)^{k_2}$ is not central in G. Thus we may power xy to an element ζ which squares to an element of $Z(G)$ and commutes with $X_{k,l}$. It follows that $\zeta Z(G)$ is an involution in $G/Z(G)$ which centralises $X_{k,l}/Z(G)$. So from [13, Table 4.5.1] we infer that $\zeta Z(G)$ is in class t_k. Thus the proportion of elements of T^0 which are regular semisimple and power to an element which projects to a conjugate of $\zeta_k Z(G)$ is at least $\frac{1}{36}$ in case (ii).

Now we consider case (iii). In this case ζ_k is by definition an involution and k must be even. Since n is even, l is even and so we may suppose that $k \leq l$. Set $S = T_k^0 T_l^0$ and let T^* be the subgroup of T^0 of index $(q-1)_2$ containing S. As $|T^0/S| = q-1$, T^*/S has odd order. Define T_k^* and T_l^* to be subgroups of T_k and T_l respectively such that $|T_k : T_k^*| = |T_l : T_l^*| = (q-1)_2$.

As k is even, $|T_k^0|$ is even and so $T^*Z(G)/Z(G)$ has even order. Assume that $xy \in T^*$ with x regular semisimple in T_k^* with respect to $\mathrm{GL}_k(q)$, y regular semisimple in T_l^* with respect to $\mathrm{GL}_l(q)$, $|x|_2 = |T_k^0|_2$ and $|y|_2 < |T_l^0|_2 = |T_k^0|_2$. Then, as x and y have different orders, xy is regular semisimple in $\mathrm{GL}_n(q)$ and by powering xy we obtain ζ_k. Thus we determine the proportion of these elements in T^*. By Lemmas 25 and 29, we have at least $\frac{1}{2}(1-\frac{2}{q})|T_k^*| \geq \frac{1}{6}|T_k^*|$ choices for x and then y has to be chosen from the correct coset of T_l^0 in T_l^* so that $xy \in T^*$ and such that $|y|_2 < |T_l^0|_2$. If $l \leq 2$, then, as $k \leq l$, in this case, we have $3 \leq n \leq 4$. Since for small n we may assume that q is large, we may apply Lemma 31 to get that there are $\frac{1}{2}(1-\frac{2}{q})|T_l^0| \geq \frac{1}{6}|T_l^0|$ choices for y if $l > 2$ and $\frac{1}{2}(1-\frac{4}{q})|T_l^0| \geq \frac{1}{10}|T_l^0|$ if $l = 2$. Therefore the proportion of regular semisimple elements of T^0 which power up to ζ_k is at least

$$\frac{\frac{1}{6}|T_k^*|\frac{1}{10}|T_l^0|}{|T^0|} = \frac{1}{60(q-1)_2} \geq \frac{1}{60n}.$$

Bringing the results of the investigations of (i), (ii) and (iii) together we see that at least a proportion of $\frac{1}{60n}$ of the elements of T^0 are regular semisimple and

power to an element which projects to an involution in class t_k, $1 \leq k \leq n/2$. Since $|N_G(T^0)/T^0| = lk$ if $k \neq l$ and $|N_G(T^0)/T^0| = n^2/2$ if $k = l$, we apply Lemma 26 to obtain at least a proportion of at least

$$\frac{|T^0|}{|N_G(T^0)|}\frac{1}{60n} \geq \frac{1}{30n^3}$$

elements of G power to elements of G which project into our given class $\mathcal{C}$.

The elements of G which project into class $t'_{n/2}$ of $\mathrm{PSL}_n(q)$ are dealt with in a far easier way as in this case the appropriate torus is cyclic. Assume that class $t'_{n/2}$ is contained in $\mathrm{PSL}_n(q)$. We let T be the Singer cycle of $\mathrm{GL}_n(q)$ and $T^0 = T \cap G$. Then the unique element of order 2 in $T^0/Z(G)$ is a representative for the class $t'_{n/2}$. We apply Lemmas 25 and 29 to see that at least a proportion of $\frac{1}{2}(1 - \frac{2}{q}) \geq \frac{1}{6}$ of the elements of T^0 are regular semisimple and power to such an element. As $|N_G(T^0)/T^0| = n$, we have at least $\frac{1}{6n}$ of the elements of G power to an element which projects to an involution in $t'_{n/2}$.

This completes the proof of Theorem 36. □

We next consider the unitary groups. The considerations are similar to those for the linear groups and so we will abbreviate the arguments slightly though there are sufficient differences to merit presenting a full proof. The involutions in $\mathrm{PSU}_n(q)$ are images of the following elements of $\mathrm{SU}_n(q)$:

$$\zeta_m = \begin{cases} \mathrm{diag}(-1^m, 1^{n-m}) & m \text{ even} \\ \mathrm{diag}(-\lambda^m, \lambda^{n-m}) & m \text{ odd}, (q-1)_2 > n_2 \end{cases}.$$

where, if $(q+1)_2 > n_2$, we let $\lambda \in \mathrm{GF}(q)$ be such that $\lambda^n = -1$ and $1 \leq m \leq n-1$. Just as in the linear case the image of ζ_{n-m} and ζ_m in $\mathrm{PSU}_n(q)$ are conjugate. The conjugacy class of the image of ζ_m in $\mathrm{PSU}_n(q)$ is called t_m as in [13, Table 4.5.1]. Again, precisely as in the linear groups, there is just one further class of involutions in $\mathrm{PSU}_n(q)$, and they occur only if n is even. The involutions in this class are images in $\mathrm{PSU}_n(q)$ of central elements of $\mathrm{GU}_{n/2}(q^2)$ embedded in $\mathrm{GU}_n(q)$. We denote their conjugacy class by $t'_{n/2}$.

Theorem 37. *If $G \cong \mathrm{PSU}_n(q)$, q odd, $n \geq 3$, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power up to an element of $\mathcal{C}$ is at least c/n^3 where c is a positive constant.*

Proof. Set $G = \mathrm{SU}_n(q)$. We may suppose that if n is small then q is large.

We begin with the images of ζ_k, $1 \leq k \leq n-1$. Set $l = n-k$. Then as before, it suffices to find a lower estimate of the number of elements of G which power to either of z_k or z_l. Thus whenever possible, we assume that k is even. Therefore either k and l are both odd or k is even. Set $X_{k,l} = \mathrm{GU}_k(q) \times \mathrm{GU}_l(q)$ of $\mathrm{GU}_n(q)$. Let T_k and T_l be Singer cycles in $\mathrm{GU}_k(q)$ and $\mathrm{GL}_l(q)$ respectively. So T_k is a cyclic

group of order $q^k-(-1)^k$ and T_l is a cyclic group of order $q^l-(-1)^l$. Furthermore, $\zeta_k \in T_k$ and ζ_k is centralized by $X_{k,l}$. Set $T = T_kT_l$ and, for subgroups H of $\mathrm{GU}_n(q)$, define $H^0 = H \cap G$. Thus $T^0 = \{xy \in T_kT_l \mid \det x \det y = 1\}$ and has order $(q^k-(-1)^k)(q^l-(-1)^l)/(q+1)$. This time the three cases are:

(i) $(q^k-(-1)^k)_2 \neq (q^l-(-1)^l)_2$,

(ii) $(q^k-(-1)^k)_2 = (q^l-(-1)^l)_2$ and $(q+1)_2 > n_2$, and

(iii) $(q^k-(-1)^k)_2 = (q^l-(-1)^l)_2$ and $(q+1)_2 \leq n_2$.

We first address case (i). Then k is even for otherwise $(q^k+1)_2 = (q^l+1)_2 = (q+1)_2$. In particular, $(q^k-1)_2 = (q^2-1)_2(k/2)_2$ and so we may assume $(q^k-1)_2 > (q^l-(-1)^l)_2$. Let $\mathcal{S}$ be the subset of T^0 with elements xy such that x is regular semisimple in T_k with respect to $\mathrm{GU}_k(q)$, $|x|_2 = (q^k-1)_2$ and y is regular semisimple in T_l with respect to $\mathrm{GU}_l(q)$. As x and y have different orders, their product is regular semisimple in $\mathrm{GU}_n(q)$ and it powers to ζ_k. By Lemmas 25 and 30, there are at least $|T_k|\frac{1}{2}(1-\frac{2}{q})$ choices for x when $k \neq 2$ and when $k = 2$ at least $|T_k|\frac{1}{2}(1-\frac{3}{q})$ choices as long as $q \neq 3$. In the extreme case when $q = 3$ and $k = 2$, exactly half the elements of T_k have order 8 and are regular semisimple. The element y is picked so that $\det y = \det x^{-1}$. If $(l,q) \neq (2,3)$, Lemma 30 implies there are at least $(1-\frac{3}{q})|T_l^0|$ choices for y if $l = 2$ and otherwise $(1-\frac{2}{q})|T_l^0|$ choices. In the exceptional case when $l = 2$ and $q = 3$, we note that as $|x|_2 = (q^k-1)_2$, $\det x \neq 1$. But then y can be chosen arbitrarily in the coset T_l^0 with the correct determinant as both elements are regular semisimple. Thus in case (i), at least a proportion $\frac{1}{18}$ of the elements of T^0 are regular semisimple and power to ζ_k.

Case (i) deals with the case when n is odd as in this case k is even and then $(q^k-1)_2 > (q^l+1)_2 = (q+1)_2$. Hence we now have n is even.

In case (ii), $(q^k-(-1)^k)_2 = (q^l-(-1)^l)_2$ and $(q+1)_2 > n_2$. We consider $\mathcal{S}$ again. Using Lemma 33 we have that the elements of $\mathcal{S}$ are regular semisimple unless $k = n/2$. If $k \neq n/2$, we obtain a proportion of at least $1/18$ of the elements of T^0 are regular semisimple and power to ζ_k. If $k = n/2$ and $k > 2$, then the number of elements of $\mathcal{S}$ which power to ζ_k is at least

$$|T_k|\frac{1}{2}(1-\frac{2}{q})(|T_k^0|(1-\frac{2}{q})-\frac{n}{2}) \geq \frac{|T^0|}{36}$$

as $\frac{n}{2|T_k^0|} \leq \frac{1}{6}$ unless both n and q are small. For $n/2 = 2$ using $q \neq 3,5$ this lower bound becomes

$$\begin{aligned} |T^0|\frac{1}{2}(1-\frac{3}{q})\frac{((q-5)(q-1)-2)}{q(q-1)} &\geq |T^0|\frac{1}{2}(1-\frac{3}{q})\frac{((q-5)(q-1)-(q-1))}{q(q-1)} \\ &\geq |T^0|\frac{1}{2}(1-\frac{3}{q})(1-\frac{6}{q}) \geq \frac{4}{49}|T^0|. \end{aligned}$$

As n is even, $(q+1)_2 > n_2 \geq 2$ and so $q \equiv 3 \pmod 4$. Consequently $(q^k-1)_2 = (q+1)_2 k_2$ and $(q^l-1)_2 = (q+1)_2 l_2$. In particular, $k_2 = l_2$ and $(q+1)_2 > n_2 > k_2$. We now argue just as in the case for the linear groups to deduce that the proportion of elements of T^0 which are regular semisimple and power to an element of $\zeta_k Z(G)$ is at least $\frac{1}{36}$ in case (ii).

We now deal with (iii). Then ζ_k is an involution so k is even and as n is even, l is also even and so we may suppose that $k \leq l$. As in the linear case, set $S = T_k^0 T_l^0$ and let T^* be the subgroup of T^0 of index $(q+1)_2$ containing S. So T^*/S has odd order. Define T_k^* and T_l^* to be subgroups of T_k and T_l such that $|T_k : T_k^*| = |T_l : T_l^*| = (q+1)_2$. Then as k is even, $|T_k^0|$ is even and so $T^*/Z(G)$ has even order.

Consider the elements $xy \in T^*$ with x regular semisimple in T_k^* with respect to $\mathrm{GU}_k(q)$, y regular semisimple in T_l^* with respect to $\mathrm{GU}_l(q)$, $|x|_2 = |T_k^0|_2$ and $|y|_2 < |T_l^0|_2 = |T_k^0|_2$. Then, as x and y have different orders, xy is regular semisimple in $\mathrm{GU}_n(q)$ and by powering xy we obtain ζ_k. Thus we determine the proportion of these elements in T^*. Suppose now that $(k,q) \neq (2,3)$. By Lemmas 25 and 30, we have at least $\frac{1}{6}|T_k^*|$ choices for x and then y has to be chosen from the correct coset of T_l^0 in T_l so that $xy \in T^*$ and such that $|y|_2 < |T_l^0|_2$. If $l \leq 2$, then, as $k \leq l$, in this case, we have $3 \leq n \leq 4$. Since we have eliminated the possibilities that $(n,q) = (4,3)$ or $(4,5)$, we may apply Lemma 32 to get that there are $\frac{1}{2}(1-\frac{2}{q})|T_l^0| \geq \frac{1}{7}|T_l^0|$ choices for y. Therefore the proportion of regular semisimple elements of T^0 which power up to ζ_k is at least

$$\frac{\frac{1}{6}|T_k^*|\frac{1}{7}|T_l^0|}{|T^0|} = \frac{1}{42(q+1)_2} \geq \frac{1}{42n}.$$

In total then, we have shown that at least a proportion of $\frac{1}{42n}$ of the elements of T^0 are regular semisimple and power to ζ_k.

Thus, so long as $(k,q) \neq 3$ in case (iii), using Lemma 26 and $|N_G(T^0)/T^0| \leq n^2/2$, we obtain at the very least a proportion of

$$\frac{|T^0|}{|N_G(T^0)|}\frac{1}{42n} \geq \frac{1}{21n^3}$$

elements of G power to elements of G which project into our given class $\mathcal{C}$.

We now return to the case that $(k,q) = (2,3)$. We may suppose that $n \geq 5$. Let $R_k \leq \mathrm{GU}_2(3)$ with R_k a product of two cyclic groups of order 4. Set $R = R_k \times T_l$ and let R^* be the unique subgroup of R^0 of index 4. Now choose x in R_k^* of order 4 and note that x is a regular semisimple element in $\mathrm{GU}_2(q)$. Since $q = 3$, $l \geq 3$ and so there are $\frac{1}{2}(1-\frac{2}{q})|T_l^0|$ choices for y such that xy is regular semisimple in R^0. It follows that at least $\frac{1}{48}$ of the elements of R^0 power to z_k in this case. As $|N_G(R^0)/R^0| = 2(n-2)$, we use Lemma 26 to get our result.

The conjugacy class $t'_{n/2}$ is dealt with exactly as in the linear case. We obtain 1/6 of the elements of T^0 are regular semisimple and power to the correct involution. So Lemma 26 again delivers the result.

This concludes the proof of Theorem 37. □

We continue with the symplectic groups $\mathrm{PSp}_{2n}(q)$, in which the involution centralisers lift to $\mathrm{Sp}_{2k}(q) \times \mathrm{Sp}_{2n-2k}(q)$ or $\mathrm{Sp}_n(q) \wr 2$ (if n is even) or $\mathrm{GL}_n(q).2$ or $\mathrm{GU}_n(q).2$. We pick tori which are products of two cyclic groups one of order $q^k \pm 1$ and the other of order $q^{n-k} \pm 1$ in the first two cases, and in the last two cases we take a cyclic torus of order $q^n \pm 1$.

Theorem 38. *If $G \cong \mathrm{PSp}_{2n}(q)$, q odd, $n \geq 2$, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power to an element of $\mathcal{C}$ is at least c/n^2 where c is a positive constant.*

Proof. As usual we may suppose that if q is small then n is large. Consider first the involutions which lift to elements of order 4 in $\mathrm{Sp}_{2n}(q)$. Their centralisers lift either to $\mathrm{GL}_n(q).2$ or $\mathrm{GU}_n(q).2$ (according as $q \equiv 1$ or $3 \bmod 4$). These groups contain cyclic maximal tori of $\mathrm{Sp}_{2n}(q)$, of order $q^n - 1$ or $q^n + 1$, which contain the given elements of order 4. In either case $|N_G(T)/T| = 2n$. Lemma 16 shows that at least half of the elements of this torus are regular semisimple, and by Lemma 25 at least half of the regular semisimple elements have order divisible by the full 2-power in $|T|$. Hence at least a fraction $\frac{1}{4}.\frac{1}{2n} = \frac{1}{8n}$ of the elements of $\mathrm{PSp}_{2n}(q)$ power up to an involution in this conjugacy class.

The remaining involutions lift to involutions in $\mathrm{Sp}_{2n}(q)$, and their centralisers lift to $\mathrm{Sp}_{2k}(q) \times \mathrm{Sp}_{2n-2k}(q)$ if $2k < n$, and $\mathrm{Sp}_{2k}(q) \wr 2$ if $2k = n$. In these cases we choose maximal tori of shape $C_{q^k \pm 1} \times C_{q^{n-k} \pm 1}$. The signs may be chosen so that $q^k \pm 1$ is divisible by 4, and $q^{n-k} \pm 1$ is not. Therefore at least half the elements in the second factor are regular semisimple, and at least a quarter of the elements in the first factor are regular semisimple and have the full 2-part of the order. Therefore, as the product of these regular semisimple elements is regular semisimple at least one eighth of the elements of this torus are regular semisimple elements which power up to the required involution. In all cases $|N_G(T)/T| = 4k(n-k)$ (or $8k(n-k) = 2n^2$ if $n = 2k$), so the proportion of elements of G which power into this conjugacy class of involutions is at least $1/16n^2$. □

In odd dimensions, the orthogonal groups behave very much like the symplectic groups.

Theorem 39. *There exists a positive constant c such that, if $G \cong \mathrm{P}\Omega_n(q)$, q odd, n odd, $n \geq 7$, and $\mathcal{C}$ is a conjugacy class of involutions in G, then the proportion of elements of G which power to an element of $\mathcal{C}$ is at least c/n^2.*

Proof. Again assume that if q is small, then n is large. Since n is odd, the simple group $G = \Omega_n(q) \cong \mathrm{P}\Omega_n(q)$ has index 2 in $\mathrm{SO}_n(q)$, so we can work in $\mathrm{GO}_n(q)$ or $\mathrm{SO}_n(q)$. We use the information provided in [13, Table 4.5.1] to describe the involutions in G. Thus we have involutions z_k^ε, for all $1 \leq k \leq (n-1)/2$ and

$\varepsilon = \pm 1$, with centralisers $\mathrm{O}^\varepsilon_{2k}(q) \times \mathrm{O}_{n-2k}(q)$. In fact $z_k^\varepsilon \in G$ if and only if $q^k \equiv \varepsilon \pmod 4$. For $z_k^\varepsilon \in G$, we see that z_k^ε is contained in a maximal torus T_{2k} of $\mathrm{O}^\varepsilon_{2k}(q)$ which is cyclic of order $q^k - \varepsilon$ (contained in $\mathrm{O}^\varepsilon_2(q^k)$). Note that $T_{2k} \leq \mathrm{SO}_n(q)$ but intersects G in a subgroup of index 2. We select a maximal cyclic torus T_{n-2k-1} of $\mathrm{O}_{n-2k}(q)$ contained in $\mathrm{O}^\theta_{n-2k-1}(q)$ such that $q^{(n-2k-1)/2} - \theta$ is twice an odd number. Provided $(k,q) \neq (1,3)$ or $(1,5)$ and $(n-2k-1,q) \neq (2,3)$, Lemmas 25 and 16 imply at least one quarter of the elements of $T_k \cap G$ are regular semisimple and have maximal 2-part in their order and one half of the elements of $T_{n-2k-1} \cap G$ are regular semisimple (in fact they have odd order). Since these elements have different orders, their product is regular semisimple. Hence at least one sixteenth of the elements of $T \cap G$ are regular semisimple and power to z_k^ε in the typical cases. Returning to the special cases with $(k,q) = (1,3)$ or $(1,5)$ and $(n-2k-1,q) \neq (2,3)$, in the first case let $x = z_k^\varepsilon$ and choose y to be a regular semisimple element of $T_{n-3} \cap G$ (here we use that as q is small we may assume that $n \geq 6$ and we also note that y has odd order). In these cases the product has centraliser contained in $N_G(T)$ and so such elements uniquely determine T and so are just as good as regular semisimple elements. Since $N_G(T)/T$ has order at most n^2, the result follows from Lemma 26 in all cases except $(n-2k-1,q) \neq (2,3)$.

Suppose that $(n-2k-1,q) = (2,3)$. We may suppose that $n > 8$. By Lemmas 16 and 25 at least one quarter of the elements of T_{2k} are regular semisimple with respect to $\mathrm{O}^\varepsilon_{2k}(3)$ and power to z_k^ε. Let x be such an element, let $H = C_G(z_k^\varepsilon)$ and let K be the component of H. Then $K \cong \mathrm{O}^\varepsilon_{2k}(3)$. Let $R = C_G(H)$, so that $R \cong \Omega_3(3)$. Then $U = C_G(x) = T_{2k}R$ has order $12(3^k - \varepsilon)$. It follows that at least a proportion $\frac{1}{48}$ of the elements of U have centraliser U and power to z_k^ε. Since $Z(U) = T_{2k}$, the proportion of elements of G which power to z_k^ε is at least $\frac{1}{|G|}\frac{|U|}{48}\frac{|G|}{|N_G(U)|} = \frac{1}{96}$. Now we have that at least a proportion d/n power to a conjugate of z_k^ε, for some constant d. This completes the analysis of the case when G is orthogonal in odd dimension. □

If n is even, this free choice of tori which we exploited when n was odd is not available, and the argument needs to be more subtle.

Theorem 40. *There exists a positive constant c such that, if $G \cong \mathrm{P}\Omega^\varepsilon_{2n}(q)$, q odd, $n \geq 4$, and C is a conjugacy class of involutions in G, then the proportion of elements of G which power to an element of C is at least c/n^2.*

Proof. Again we work in $\mathrm{O}^\varepsilon_{2n}(q)$, let $G = \Omega^\varepsilon_{2n}(q)$ and note that if q is small we may assume that n is large. The involutions of $\mathrm{P}\Omega^\varepsilon_{2n}(q)$ which lift to elements of order 4 in $\mathrm{O}_{2n}(q)$ have centralisers which lift to $\mathrm{GL}_n(q).2$ or $\mathrm{GU}_n(q).2$. So, for such involutions in the projective group, the same argument as for the symplectic groups shows that the proportion of elements of G which power into this class is at least a constant times n^{-1}.

We are left therefore with the involutions of $\mathrm{P}\Omega^\varepsilon_{2n}(q)$ which lift to involutions $z_k^{\varepsilon_1}$ in G and have centralisers of shape $\mathrm{O}^{\varepsilon_1}_{2k}(q) \times \mathrm{O}^{\varepsilon_2}_{2m}(q)$ inside $\mathrm{O}^\varepsilon_{2n}(q)$, where

$k+m=n$ and $\varepsilon=\varepsilon_1.\varepsilon_2$. These lie in $G=\Omega^{\varepsilon}_{2n}(q)$ exactly when at least one of $q^k-\varepsilon_1$ and $q^m-\varepsilon_2$ is divisible by 4. Let T_{2k} be a torus of order $q^k-\varepsilon_1$ contained in the first factor, and T_{2m} be a torus of order $q^m-\varepsilon_1$ in the second factor of $\mathrm{O}^{\varepsilon_1}_{2k}(q)\times\mathrm{O}^{\varepsilon_2}_{2m}(q)$. Without loss of generality we may assume that $q^k\equiv\varepsilon_1 \pmod 4$ and that if $q^m\equiv\varepsilon_2 \pmod 4$ then $m\leq k$. Set $T=(T_{2k}\times T_{2m})\cap G$. Then T has index 4 in $T_{2k}\times T_{2m}$

If $|T_{2k}|_2\neq|T_{2m}|_2$, then Lemmas 25 and 16 ensures that, so long as $(k,q,\varepsilon_1)\neq(1,3,-)$ or $(1,5,+)$ and $(m,q,\varepsilon_2)\neq(1,3,\pm)$ or $(1,5,+)$, at least one quarter of the regular semisimple elements in $T_{2k}\cap G$ and have the full power of 2 in their order and one half of the elements of $T_{2m}\cap G$ are regular semisimple. So, as the orders of such elements are different, their product powers up to $z_k^{\varepsilon_1}$ and are regular semisimple. Since $|N_G(T)/T|<n^2$, we deduce that the proportion of elements of the group which power up to such an involution is at least $1/16n^2$. Consider the exceptional cases $(k,q,\varepsilon_1)=(1,3,-)$ or $(1,5,+)$ and $(m,q,\varepsilon_2)=(1,3,\pm)$ or $(1,5,+)$. In the first two cases, we have $q-\varepsilon_1=4$. Hence, as $n\geq 8$, our choice of m and k implies $(q^m-\varepsilon_2)_2=2$. It follows that $T_{2m}\cap G$ has odd order. We let $w=(t_k^{\varepsilon_1},y)\in(T_{2k}\cap G)\times(T_{2m}\cap G)\leq T$ where $y\in T_{2l}\cap G$ is regular semisimple of odd order. By Lemma 16, at least one half of the elements in $T_{2m}\cap G$ are regular semisimple and so at least one eighth of the elements of T are described in this way. Now for such elements we have $C_G(w)\leq N_G(T)$ and so, in this case, there is a constant c such that at least a proportion of c/n elements of G power to $z_k^{\varepsilon_1}$. If $(m,q,\varepsilon_2)=(1,3,+)$, we have $C_G(x)=T_{2k}$ for regular semisimple elements of T_{2k} and so as one quarter of the elements of T_k are regular semisimple and power to $z_k^{\varepsilon_1}$, at least a proportion $1/4n$ of the elements of G power to $z_k^{\varepsilon_1}$. Finally for $(m,q,\varepsilon_2)=(1,3,-)$ or $(1,5,+)$ we have $(q-\varepsilon_2)=4$. Thus $(q^k-\varepsilon_1)_2\geq 8$ and this time we consider elements $w=(x,y)\in T$ where x is a regular semisimple element of T_{2k} and has the full power of 2 in its order and $y\in T_{2m}$ is an involution. Then $C_G(w)\leq N_G(T)$ and so at least a proportion c/n of the elements of G power to $z_k^{\varepsilon_1}$, where again c is an appropriate constant.

Now consider the case $|q^k-\varepsilon_1|_2=|q^m-\varepsilon_2|_2$. Note that the central involutions of T_{2k} and T_{2m} project to the same involution in $\mathrm{P}\Omega_{2m}(q)$. We may assume that $k\leq m\leq n/2$. We estimate the number of regular semisimple elements in $(T_{2k}\cap G)\times(T_{2m}\cap G)$ which have the full 2-power of the order in the first factor but not in the second. We use Lemmas 16 and 25 to see that so long as $(k,q,\varepsilon_1)\neq(1,3,-)$, $(1,5,+)$, at least one quarter of the elements of T_{2k} have the required property. Since $m\geq k$, we may suppose that m is large if q is small. Thus Lemma 16 implies that at least $\frac{1}{4}$ of the elements of T_{2m} are regular semisimple and do not have full 2-power in their order. Since $|N_G(T)/T|\leq n^2$, Lemma 26 now gives the result so long as $(k,q,\varepsilon_1)\neq(1,3,-)$, $(1,5,+)$. The result for $(k,q,\varepsilon_1)=(1,3,-)$ and $(1,5,+)$ follows in exactly the same way as in the previous case using Lemma 16 and the fact that we may assume m is large as q is small. This completes the investigation of orthogonal groups in even dimensions. $\square$

Taken together Lemma 35 and Theorems 36, 37, 38, 39 and 40 prove Theorem 5.

4 Applications

The first purpose of proving the rather technical main theorems is to deduce Corollaries 3 and 6. Corollary 3 follows easily from Theorems 1 and 2. A proof in the case when $G = S$ is given in Theorem 7 of [14]. (Note that the assumption of an order oracle there is not necessary if we assume that the isomorphism type of S is known, since we can factorise $|S|$ into 'pretend primes' and thence calculate pseudo-orders of elements in Monte Carlo polynomial time.) The general case follows by the same argument, since an element of G has odd order if and only its image in $S = G/O_p(G)$ has odd order.

Corollary 6 follows easily from Theorems 4 and 5. The only issue which has not been addressed so far is that having found an involution, we need to identify which conjugacy class it lies in. It suffices to prove the result for S. We clearly cannot distinguish abstractly conjugacy classes in S which are fused in $\mathrm{Aut}(S)$, but we claim that all other pairs of classes can be distinguished. If there is only one $\mathrm{Aut}(S)$-conjugacy class of involutions in S, there is nothing to prove. Otherwise, we use Corollary 3 to construct the centraliser of the involution. Then Corollary 4.4 of [3] allows us to compute the names of the non-abelian composition factors of this centraliser. Inspection of the list of involution centralisers in Table 4.5.1 of [13] shows that this suffices to determine the $\mathrm{Aut}(S)$-class of the involution. (Note however that for arbitrary fields, D_1 is soluble, so disappears from the information we compute, and for $q = 3$, also A_1, B_1, C_1 and D_2 are soluble, so for a few small-rank groups over GF(3) we need some other, if necessary brute-force, computation.)

Our original motivation for proving these results was to obtain an effective algorithm for testing whether $O_p(G) = 1$, in black box groups of characteristic p. This so-called 'p-core problem' (otherwise known as the problem of 'O_p or not O_p') is discussed in [2] and [3] as being one of the important open problems in the development of polynomial-time algorithms for black-box groups.

In the eight years since the first draft of this paper was written, however, the subject has moved on, and many other problems have been solved by methods similar to the ones we proposed. Borovik himself proposed such methods in [8], and his former student Yalçinkaya has carried out some of them [23]. Indeed, the whole idea of using involution centralisers in computational group theory has really taken off, and many applications of our results are already in the literature. For example, our results are used in [14] to underpin an effective algorithm for constructive membership testing in black-box groups of Lie type in odd characteristic. The paper by Liebeck and O'Brien [18] on recognising the characteristic of a black-box group also makes fundamental use of involution centraliser methods.

The long preprint of Leedham-Green and O'Brien on constructive recognition of classical groups of odd characteristic [20] is also based on these methods.

The heart of the general p-core problem is the much more specific problem to distinguish between a simple group and a non-simple group. In other words, given a black-box group G we wish to certify, in polynomial time, with arbitrarily small probability of error, either that G is simple, or that G is not simple. In the latter case, moreover, we wish to provide a witness in the form of an element whose normal closure is a proper non-trivial normal subgroup. We claim that if G is a black-box group which is in fact a simple group of Lie type over a finite field of odd characteristic, then we can certify this in Monte Carlo polynomial time, subject to the existence of an order oracle.

This claim rests on various reductions described in [2, 3, 18]. First, the characteristic p can be found in Monte Carlo polynomial time by [18], using an order oracle. Once the characteristic is known, Theorem 4.17 of [3] reduces the problem to the problem of distinguishing between a simple group of known characteristic p, and a group which is an extension of a p-group by the same simple group.

We shall show that this last problem can be solved in Monte Carlo polynomial time, without the necessity for an order oracle. The ingredients we shall need are as follows:

(i) $O_p(G)$ can be recognised in Monte Carlo polynomial time. Therefore $G/O_p(G)$ is a black-box group of characteristic p. This is proved in Section 3.4 of [2].

(ii) If subgroups $H < K$ of G have been constructed, then the normal closure of H in K can be constructed in Monte Carlo polynomial time. This is Theorem 1.5 of [5].

(iii) If H is a black-box group, then the derived group H' can be constructed in Monte Carlo polynomial time. This is in Corollary 1.6 of [5].

(iv) If H is a black-box group which is a central product of quasisimple groups (i.e. its components), then the components of H can be constructed in Monte Carlo polynomial time. This is proved in [2, Theorem 5.1] and [7, Remark 1.9].

(v) If H is a black-box group which is quasisimple, then $Z(H)$ can be constructed in Monte Carlo polynomial time. This is Theorem 4.15 of [3].

(vi) If G is a black-box group which has a non-trivial p-quotient, then more than half (or better: a proportion $1 - 1/p$ of) its elements have order divisible by p. This is obvious.

(vii) If $S = G/O_p(G)$ is a simple group of Lie type in odd characteristic, and $z \in G$ is an involution, then $O_p(C_G(z)) \neq 1$ if and only if $O_p(G) \neq 1$, and moreover, $O_p(C_G(z)) \leq O_p(G)$. This follows easily from the structure of the involution centralisers in S given in Table 4.5.1 of [13].

As an example, suppose that G has a normal (possibly trivial) elementary abelian p-subgroup A, such that $G/A \cong \mathrm{PSp}_{2n}(q)$, with $q > 3$. We test whether $A = 1$ by finding an involution $z \in G$ and its centraliser $C_G(z)$. Clearly $A \neq 1$ if and only if $C_A(z) \neq 1$. Thus $H = C_G(z)/\langle z \rangle$ has a normal (possibly trivial) elementary abelian p-subgroup B, such that $H/B \cong \mathrm{PSp}_{2k}(q) \times \mathrm{PSp}_{2n-2k}(q)$ (provided $n \neq 2k$). Moreover, for any $n > 2$, we may choose $k = 1$ if we like. Now we look for elements of order $(q^{n-1}+1).(q-1)/4$ and power them up to get elements of order $(q-1)/2$, which map to the factor $\mathrm{PSp}_2(q)$ in H/B. Similarly, we can power them up to get elements of order $(q^{n-1}+1)/2$, which map to the factor $\mathrm{PSp}_{2n-2}(q)$ in H/B. Now if $B \neq 1$ then at least one of these two groups has a non-trivial normal p-subgroup (and in any case, if this is not true, then the involution centraliser acts trivially on its O_p-subgroup, so p-singular elements are easy to find, and we immediately obtain a non-trivial element of B), so we proceed by induction. In the case $n = 2$, we have $H/B \cong \mathrm{PSp}_2(q) \wr 2$, so we first pass to the subgroup of index 2, and then proceed as before. After $n-1$ steps we have reduced to n groups H_i such that $H_i/O_p(H_i) \cong \mathrm{PSp}_2(q) \cong \mathrm{PSL}_2(q)$, with the property that $O_p(G) = 1$ if and only if $O_p(H_i) = 1$ for all i. The latter criterion can be checked by the method given in the introduction.

More generally, suppose that G is a group with a possibly trivial normal p-subgroup A, such that $S = G/A$ is a simple group of Lie type in characteristic p. The following recursive algorithm determines, in Monte Carlo polynomial time, whether or not $A = 1$.

Input: a black-box group G of characteristic p, such that $G/O_p(G)$ is a known simple group of Lie type in characteristic p.

Output: either a non-trivial element of $O_p(G)$, or an assertion that $O_p(G) = 1$.

1. If S is defined over GF(3), of Lie rank at most 6, use brute force. Return $O_p(G) \neq 1$ with a witness, or $O_p(G) = 1$, as appropriate.

2. Working in $S = G/O_p(G)$, find an involution z, and determine its class (in $\mathrm{Aut}(S)$). More specifically, we can target a particular conjugacy class of involution: we ensure in this way that its centraliser $C_S(z)$ has the property that $C_S(z)''$ is perfect, except in the case $S \cong \mathrm{PSL}_2(q)$.

3. Lift z to G: if this is not an involution, we have found a non-trivial element z^2 of $O_p(G)$, so return it and the statement $O_p(G) \neq 1$.

4. Construct $C_G(z)$, using Corollary 3. We know that $O_p(C_G(z)) \neq 1$ if and only if $O_p(G) \neq 1$, and moreover, $O_p(C_G(z)) \leq O_p(G)$.

5. Test for a non-trivial p-quotient of $C_G(z)$ by looking at the orders of a few random elements. If successful, power up to an element of order p which necessarily lies in $O_p(G)$. Return the element of order p and the statement $O_p(G) \neq 1$.

6. Calculate $C_G(z)'$ and test for a non-trivial p-quotient. If successful, return the element of order p and the statement $O_p(G) \neq 1$.

7. Calculate $C_G(z)''$ and test for a non-trivial p-quotient. If successful, return the element of order p and the statement $O_p(G) \neq 1$.

8. If $C_G(z)'' = 1$ (and therefore $S = G = \mathrm{PSL}_2(q)$), return $O_p(G) = 1$.

9. Work in $C_G(z)''/O_p(C_G(z)'')$, which is a central product of quasisimple groups, to construct the components K_i.

10. Lift random elements of each K_i to G, and construct the normal closure G_i in $C_G(z)''$. Now we know that $O_p(G) \neq 1$ if and only if there exists i such that $O_p(G_i) \neq 1$. (This is because if all the $O_p(G_i) = 1$, then elements of order p in $O_p(G)$ must have been found in one of the previous steps, unless $O_p(G) = 1$.)

11. For each i, construct the centre Z_i of the quasisimple group $G_i/O_p(G_i)$.

12. Lift random elements of Z_i to G_i: if any two fail to commute, their commutator is an element of $O_p(G)$, so return this commutator and the statement $O_p(G) \neq 1$.

13. Otherwise, each Z_i is central in G_i, and we work in the black box groups G_i/Z_i, each of which now satisfies the conditions required of the input group.

14. For each i, call the algorithm with input G_i, from step 1. If any call returns $O_p(G_i) \neq 1$, then return $O_p(G) \neq 1$ with the witness. Otherwise, return $O_p(G) = 1$.

We have shown that each step can be accomplished in Monte Carlo polynomial time. It remains to show that the number of calls to the algorithm is not too large. But at each reduction, the sum of the Lie ranks of the components of the involution centraliser is at most the Lie rank of the input simple group S, so the total number of nodes in the recursion tree is less than twice the Lie rank of S.

We have proved Theorem 7, and Corollary 8 follows. Corollary 9 follows, since matrix groups have a polynomial time order oracle.

In fact, Yalçinkaya [23] proves more or less the same result, independently, using similar methods. The main difference between his work and ours, is that he restricts himself to one class of involutions, namely the class of so-called classical involutions, whose centralisers have a component which is a long root $\mathrm{SL}_2(q)$. While this seems to make little difference to the complexity of the algorithm for this particular problem, the flexibility we have to choose the class of involutions freely has important benefits for the solution of other problems, such as those proposed in [14] and [20].

The preprint of Lübeck, Niemeyer and Praeger [22] is devoted to proving stronger bounds than can be deduced directly ours, by allowing a range of conjugacy classes of involutions rather than a single one. These stronger bounds will be of greater use in designing practical implementations of our algorithms and others, as they will more effectively limit the degree of the polynomial describing the time complexity.

The only finite simple groups which cannot at present be recognised in Monte Carlo polynomial time are the groups of Lie type over a (large) field of even order. The first obstacle here is that we have no method of finding an element of even order in Monte Carlo polynomial time. If we could find an element of even order, and power it up to an involution, then it would be easy to find its centraliser, since (presumably, although we have not proved it) almost all products of two conjugate involutions have odd order. However, even with the involution centraliser it is not obvious how to proceed, as already in the simple group the involution centraliser has a large normal 2-subgroup. This means that we no longer have a simple criterion for $O_p(G) \neq 1$ in terms of the involution centraliser.

As far as our more general aim is concerned, namely to determine in Monte Carlo polynomial time whether $O_p(G) = 1$ for an *arbitrary* black-box group G, we believe this should now be possible for any odd p. Indeed, Seress announced just such a result, relying heavily on our work, at the conference Group Theory, Combinatorics and Computation, University of Western Australia, in January 2009.

References

[1] C. Altseimer and A. Borovik, Probabilistic recognition of orthogonal and symplectic groups, pp. 1–20 in [16]. With corrections in `http://www.ma.umist.ac.uk/avb/pdf/alt-avb4.pdf`.

[2] L. Babai and R. Beals, A polynomial-time theory of black-box groups I, pp. 30–64 in [10].

[3] L. Babai and A. Shalev, Recognizing simplicity of black-box groups and the frequency of p-singular elements in affine groups, pp. 39–62 in [16].

[4] L. Babai and Szemeredi, On the complexity of matrix group problems I, in: Proc. 25th IEEE Symp. Found. Comp. Sci., Palm Beach, FL (1984) pp. 229–240

[5] L. Babai, G. Cooperman, L. Finkelstein, E. Luks, and Á. Seress, Fast Monte Carlo algorithms for permutation groups, *J. Comput. System Sci.* **50** (1995), 296–308.

[6] L. Babai, W. Kantor, P. Palfy and Á. Seress, Black box recognition of finite simple groups of Lie type by statistics of element orders,

[7] L. Babai, P. Palfy and J. Saxl, On the number of p-regular elements in finite simple groups, to appear in *LMS JCM*.

[8] A. V. Borovik, Centralisers of involutions in black box groups, *Contemp. Math.* **298** (2002), 7–20.

[9] J. N. Bray, An improved method for generating the centralizer of an involution, *Arch. Math. (Basel)* **74** (2000), 241–245.

[10] C. M. Campbell, E. F. Robertson, N. Ruskuc and G. C. Smith (eds.), *Groups St Andrews 1997 in Bath, I*, Cambridge Univ. Press, 1999.

[11] R. W. Carter, *Finite groups of Lie type. Conjugacy classes and complex characters.* Reprint of the 1985 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1993.

[12] D. I. Deriziotis, On the number of conjugacy classes in finite groups of Lie type, *Comm. Algebra* **13** (1985), 1019–1045.

[13] D. Gorenstein, R. Lyons, and R. Solomon, *The classification of the finite simple groups, 3*, Amer. Math. Soc., 1998.

[14] P. E. Holmes, S. A. Linton, E. A. O'Brien, A. J. E. Ryba and R. A. Wilson, Constructive membership in black-box groups, *J. Group Theory* **11** (2008), 747-763.

[15] W. M. Kantor and Á. Seress, Prime power graphs for groups of Lie type, *J. Algebra* **247** (2002), 370–434.

[16] W. M. Kantor and Á. Seress (eds.), *Groups and Computation III*, de Gruyter, Berlin/New York, 2001.

[17] P. B. Kleidman and M. W. Liebeck, *The subgroup structure of the finite classical groups*, London Mathematical Society Lecture Note Series, 129. Cambridge University Press, Cambridge, 1990.

[18] M. W. Liebeck and E. A. O'Brien, Finding the characteristic of a group of Lie type, *J. London Math. Soc.* **75** (2007), 741–754.

[19] M. W. Liebeck and A. Shalev, The probability of generating a finite simple group, *Geom. Dedicata* **56** (1995), no. 1, 103–113.

[20] C. R. Leedham-Green and E. A. O'Brien, Constructive recognition of classical groups in odd characteristic, preprint, 2008.

[21] F. Lübeck, Finding p'-elements in finite groups of Lie type, pp. 249–255 in [16].

[22] F. Lübeck, A. Niemeyer and C. Praeger, Finding involutions in finite Lie type groups of odd characteristic, to appear in *J. Algebra*.

[23] Ş. Yalçinkaya, Black-box groups, *Suppl. Turkish J. Math.* **31** (2007), 171–210.