*Submitted exclusively to the London Mathematical Society*
doi:10.1112/0000/000000

# An improved diameter bound for finite simple groups of Lie type

Zoltán Halasi, Attila Maróti, László Pyber, Youming Qiao

Abstract

For a finite group $G$, let $\mathrm{diam}(G)$ denote the maximum diameter of a connected Cayley graph of $G$. A well-known conjecture of Babai states that $\mathrm{diam}(G)$ is bounded by $(\log_2 |G|)^{O(1)}$ in case $G$ is a non-abelian finite simple group. Let $G$ be a finite simple group of Lie type of Lie rank $n$ over the field $\mathbb{F}_q$. Babai's conjecture has been verified in case $n$ is bounded, but it is wide open in case $n$ is unbounded. Recently, Biswas and Yang proved that $\mathrm{diam}(G)$ is bounded by $q^{O(n(\log_2 n+\log_2 q)^3)}$. We show that in fact $\mathrm{diam}(G) < q^{O(n(\log_2 n)^2)}$ holds. Note that our bound is significantly smaller than the order of $G$ for $n$ large, even if $q$ is large. As an application, we show that more generally $\mathrm{diam}(H) < q^{O(n(\log_2 n)^2)}$ holds for any subgroup $H$ of $\mathrm{GL}(V)$, where $V$ is a vector space of dimension $n$ defined over the field $\mathbb{F}_q$.

## 1. *Introduction*

Given a finite group $G$ and a set $S$ of generators of $G$, the associated Cayley graph $\Gamma$ is defined to have vertex set the elements of $G$ and edge set $\{\{g, gs\} : g \in G,\ s \in S\}$. The diameter $\mathrm{diam}_S(G)$ of $\Gamma$ is the maximum over $g \in G$ of the length of a shortest expression of $g$ as a product of generators in $S$ and their inverses. The maximum of $\mathrm{diam}_S(G)$, as $S$ runs over all possible generating sets of $G$, is denoted by $\mathrm{diam}(G)$.

In 1992, Babai [**4**] conjectured that $\mathrm{diam}(G) < (\log |G|)^{O(1)}$ holds for any non-abelian finite simple group $G$. (Here and throughout the paper the base of the logarithms will always be 2, unless otherwise stated.) The first class of simple groups for which Babai's conjecture was proved [**12**] were the groups $\mathrm{PSL}(2, p)$ where $p$ is prime. Following Helfgott's paper [**12**], the conjecture was verified for finite simple groups of Lie type of bounded rank by Pyber and Szabó [**22**] and Breuillard, Green, Tao [**7**]. In particular, Babai's conjecture holds for exceptional simple groups of Lie type. However, the conjecture remains wide open for finite simple groups of Lie type of large rank, that is, for simple classical groups of large rank.

Babai's conjecture is open even in the case of alternating groups. Babai and Seress [**3**] proved that $\mathrm{diam}(\mathrm{A}_n) < \exp(\sqrt{n \ln n}(1 + o(1)))$ and in [**4**] they showed that the same bound holds for arbitrary permutation groups of degree $n$.

The strongest bound to date is $\mathrm{diam}(\mathrm{A}_n) < \exp(O(\log n)^4 \log\log n)$ $(n > 2)$, due to Helfgott and Seress [**14**]. The same estimate is shown to hold in [**14**] for arbitrary transitive groups of

---

2000 *Mathematics Subject Classification* 20F69, 20G40, 20C30, 20C99 (primary), 05C25, 20D05, 51N30, 11N05 (secondary).

This work on the project leading to this application has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 741420). The first, second and third authors were partly supported by the National Research, Development and Innovation Office (NKFIH) Grant No. K115799. The first and second authors were also supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The fourth author was also supported by the Australian Research Council DE150100720.

degree $n$. The inductive proof of Helfgott and Seress relies heavily on the fact that their result extends to transitive groups. For a greatly simplified argument see [**13**].

In connection with Babai's conjecture, we remark that Breuillard and Tointon [**8**] showed, without the use of the classification theorem of finite simple groups, that for any $\epsilon > 0$ there is a constant $C_\epsilon$ depending only on $\epsilon$ such that every non-abelian finite simple group $G$ with a symmetric generating set $S$ satisfies

$$\mathrm{diam}_S(G) \leq \max\left\{\left(\frac{|G|}{|S|}\right)^{\epsilon}, C_\epsilon\right\}.$$

Breuillard remarks in his ICM survey [**6**], that it would be interesting to get non-trivial bounds for all finite simple groups of Lie type also when the rank grows and see if one can improve the above "crude bound".

Let $G$ be a finite simple group of Lie type of Lie rank $n$ defined over $\mathbb{F}_q$. Biswas and Yang [**5**] proved that $\mathrm{diam}(G) < q^{O(n(\log n+\log q)^3)}$. The first result of the present paper provides an improvement of this bound showing that the exponent need not depend on $q$.

THEOREM 1.1. *If $G$ is a finite simple group of Lie type of Lie rank $n$ defined over the field of size $q$, then* $\mathrm{diam}(G) < q^{O(n(\log n)^2)}$.

Let $\Gamma$ be a finitely generated group and $S$ a finite set of generators of $\Gamma$. For a positive integer $n$, let $\gamma_S(n)$ denote the number of elements in $\Gamma$ which may be expressed as a product of $n$ elements of $S \cup S^{-1} \cup \{1\}$. The celebrated theorem of Gromov [**11**] asserts that $\Gamma$ is virtually nilpotent if and only if the function $\gamma_S$ is bounded from above by a polynomial in $n$. Recently Shalom and Tao [**24**] obtained a strengthening of this theorem, namely that if $\gamma_S(n) \leq n^{c(\log\log n)^c}$ for some $n > 1/c$ with $c > 0$ a sufficiently small absolute constant, then $\Gamma$ is virtually nilpotent.

The Gap Conjecture asserts that if a finitely generated group $\Gamma$ has growth type strictly less than $e^{\sqrt{n}}$ then it is virtually nilpotent (see [**10**] for a precise formulation of the conjecture). As the famous Grigorchuk groups show this would be best possible even within the class of residually finite $p$-groups.

The above conjecture was shown to hold for residually nilpotent groups [**10**], [**17**]. For $\Gamma$ a residually solvable group the Gap Conjecture, with $\sqrt{n}$ replaced by $n^{1/7}$, has been proved by Wilson [**27**] (see also [**25**] and the slides [**26**] of his talk at the 2010 Ischia Group Theory Conference). One of the main ingredients of the proof was to establish upper bounds for $\mathrm{diam}(G)$ in case $G$ is a solvable subgroup of $\mathrm{GL}(V)$ acting completely reducibly on the finite vector space $V$. Wilson shows that in general $\mathrm{diam}(G) \leq O(1)|V|$. He also points out that this bound is sharp since $G$ may be taken to be a Singer cycle in $\mathrm{GL}(V)$.

Motivated by the above results, we consider the diameters of arbitrary linear groups over finite vector spaces.

THEOREM 1.2. *Let $G$ be a subgroup of $\mathrm{GL}(V)$ where $V$ is a vector space of dimension $n$ defined over the field of size $q$ and characteristic $p$. Let $h = \max_S\{\mathrm{diam}(S)\}$ where $S$ runs over the (non-abelian) classical composition factors of $G$ defined over fields of characteristic $p$, if such exist, otherwise put $h = 1$. Then*

$$\mathrm{diam}(G) < |V|^{O(1)}h^2 < q^{O(n(\log n)^2)}.$$

Note that Theorem 1.2 may be viewed as an extension of Theorem 1.1. Actually, both results also extend to directed Cayley graphs by a result of Babai [**1**].

Theorem 1.2 is deduced from a structure theorem for a finite group acting completely reducibly on a vector space (see Theorem 3.3).

For $G = \mathrm{GL}(V)$ we must have $\mathrm{diam}(G) \geq \mathrm{diam}_S(G) \geq (q-1)/2$ where $S$ is a generating set of $G$ where all but one element in $S$ has determinant 1. This shows that the diameter of absolutely irreducible (almost simple) subgroups of $\mathrm{GL}(n,q)$ may be much larger than the bound predicted by Babai's conjecture for $\mathrm{PSL}(n,q)$.

Kornhauser, Miller and Spirakis [**16**] asked in 1984 whether or not the diameter of a transitive group is always polynomially bounded in terms of the degree. A positive answer (which is supported by the results in [**14**]) would show that the best possible bound for $\mathrm{S}_n$ and for its transitive subgroups is the same. (As the example of Singer cycles in $\mathrm{SL}(V)$ shows, the analogue of this is unlikely to be true for $\mathrm{SL}(V)$ where $V$ is a finite vector space.) Since the minimal degree of a permutation representation of a simple group of Lie type of rank $n$ over the field $\mathbb{F}_q$ is roughly $q^n$, our Theorem 1.1 also supports a positive answer to the above question.

## 2. *Proof of Theorem 1.1*

### 2.1. *A new degree reduction lemma*

In this section we prove Theorem 1.1. To achieve this, we prove a new degree reduction lemma for matrices over finite fields (Lemma 2.2). This is a linear algebraic analogue of the degree reduction lemma for permutations by Babai and Seress [**2**, Lemma 3]. It improves the corresponding one by Biswas and Yang [**5**, Lemma 4.4 (ii)]. Theorem 1.1 then follows by combining Lemma 2.2 with the rest of the Biswas-Yang machinery.

We first state our degree reduction lemma, and indicate how Theorem 1.1 follows from this together with [**5**]. We then prove this lemma in Section 2.3.

Let us set up some notation. For $n \in \mathbb{N}$, let $[n] := \{1, 2, \ldots, n\}$. Fix a finite field $\mathbb{F}_q$ of characteristic $p$. Let $\overline{\mathbb{F}_q}$ be the algebraic closure of $\mathbb{F}_q$. We use $I$ to denote identity matrices. Let $\mathrm{M}(n,q)$ denote the linear space of $n \times n$ matrices over $\mathbb{F}_q$, and $\mathrm{GL}(n,q)$ the group of $n \times n$ invertible matrices over $\mathbb{F}_q$. For $A \in \mathrm{M}(n,q)$, we use $\mathrm{charpoly}(A,x)$ and $\mathrm{minpoly}(A,x)$ to denote the characteristic polynomial and the minimal polynomial of $A$ in the variable $x$, respectively.

Definition 2.1. *The degree* $\deg(A)$ *of a matrix* $A \in \mathrm{M}(n,q)$ *is the rank of* $A - I$.

We now state the degree reduction lemma, whose proof is postponed to Section 2.3.

Lemma 2.2. *Suppose we are given* $A \in \mathrm{GL}(n,q)$, *such that* $\mathrm{charpoly}(A,x)$ *has irreducible factors* $f_1, \ldots, f_r$ *of degrees* $p_1, \ldots, p_r$ *respectively, where the* $p_i$ *are distinct primes larger than* 2 *for which the inequality* $\prod_{i \in [r]} p_i > n^4$ *holds. Then there exists* $m \in \mathbb{N}$, *such that* $A^m$ *is a non-identity matrix of degree at most* $\deg(A)/4$. *Furthermore, if each* $f_i$ *has a root of order*

*$q^{p_i}-1$ over $\overline{\mathbb{F}_q}$, then there exists $m' \in \mathbb{N}$, such that $A^{mm'}$ has the additional property that 1 is its only eigenvalue lying in $\mathbb{F}_q$.*

Note that we do not require $p_1, \ldots, p_r$ to be the first $r$ odd primes.

Note that an irreducible polynomial $f_i$ of degree $p_i$ over $\mathbb{F}_q$ has a root of order $q^{p_i}-1$ over $\overline{\mathbb{F}_q}$ if and only if $f_i$ is the minimal polynomial of some Singer cycle element in $\mathrm{GL}(p_i, q)$. Such polynomials $f_i$ exist for every $p_i$ and $q$.

Compare Lemma 2.2 with [**5**, Lemma 4.4 (ii)]. The key difference is that Biswas and Yang required the primes to be coprime with $p(q-1)$, while we do not have such a restriction. This leads to the desired improvement, because of the following easy number-theoretic bounds, as already used in [**2**, Sec. 3].

By a classical result of Erdős [**9**], there exist constants $\gamma_1$ and $\gamma_2$ larger than 1 such that for every number $x \geq 1$ we have

$$\gamma_1^x < \prod_{\substack{x < p' \leq 2x \\ p' \text{prime}}} p' < \gamma_2^x.$$

For $y \geq 2$ let $f(y)$ be the product of all primes no greater than $y$. For $y \geq 4$ we have $\gamma_1^{y/2} \cdot f(y/2) < f(y) < \gamma_2^{y/2} \cdot f(y/2)$, and by induction this gives $\gamma_1^y < f(y) < \gamma_2^y$. Let $\bar{p}$ be a prime. From $f(\bar{p}) < {\gamma_2}^{\bar{p}}$ we get

$$\sum_{\substack{p' \leq \bar{p} \\ p' \text{ prime}}} p' = \sum_{\substack{p' \leq \bar{p} \\ p' \text{ prime}}} \Big( \frac{p'}{\log p'} \cdot \log p' \Big) < \frac{2\bar{p}}{\log \bar{p}} \cdot \Big( \sum_{\substack{p' \leq \bar{p} \\ p' \text{ prime}}} \log p' \Big) < \frac{2\bar{p}^2 \log \gamma_2}{\log \bar{p}}.$$

For a fixed integer $n$ we may take $\bar{p}$ to be the smallest prime such that ${\gamma_1}^{\bar{p}} \geq n^4$. This assures that the product of all primes no greater than $\bar{p}$ is larger than $n^4$ and also that the sum of all primes no greater than $\bar{p}$ is bounded by

$$\frac{2\bar{p}^2 \log \gamma_2}{\log \bar{p}} < \gamma_3 \frac{(\log n)^2}{\log \log n}$$

for some constant $\gamma_3$ and all $n \geq 3$. To see the latter claim note that $\bar{p} = O(\log n)$ by the Bertrand-Chebyshev theorem.

Compare the above estimates with [**5**, Lemma 4.4 (i)]. There, because of the coprime with $p(q-1)$ condition, the sum over the orders of $q$ in $\mathbb{Z}/p_i\mathbb{Z}$ can only be bounded from above by $O((\log n + \log q)^3)$, provided that the least common multiple of these orders is larger than $n^4$.

### 2.2. *The Biswas-Yang machinery*

A proof of Theorem 1.1 follows by plugging in Lemma 2.2 to the rest of the Biswas-Yang machinery [**5**]. We briefly outline the procedure.

In order to prove Theorem 1.1, it is sufficient to assume that $G$ is a finite simple classical group (of unbounded dimension $n$), by the fact that Babai's conjecture is known to hold in the bounded rank case (see [**22**] and [**7**]). Moreover, it is sufficient to establish the estimate $\mathrm{diam}(G) < q^{O(n(\log n)^2)}$ for every group $G$ isomorphic to $\mathrm{SL}(n, q)$, $\mathrm{Sp}(n, q)$, $\mathrm{SU}(n, q)$, or $\Omega(n, q)$, with $n$ sufficiently large.

Let $V$ be a vector space of dimension $n$ over the field $\mathbb{F}_q$. If $G$ is different from $\mathrm{SL}(V)$, we view $V$ as a non-degenerate formed space with a non-degenerate alternating bilinear form

in the symplectic case, with a non-degenerate Hermitian form in the unitary case, or with a non-degenerate quadratic form in the orthogonal case.

Let $t$ be a positive integer. Following [**5**, Definition 2.1], we say that a subset $H$ of $\mathrm{GL}(V)$ is a *$t$-transversal set* if, given any embedding $X$ of a subspace $W$ of dimension $t$ into $V$, there is a linear transformation in $H$ whose restriction to $W$ is $X$. If $V$ is equipped with a non-degenerate form, we say, following [**5**, Definition 6.4], that a subset $H$ of $G$ is a *singularly $t$-transversal set* if, for any isometric embedding $X$ of a totally singular subspace $W$ of dimension $t$ into $V$, there is an element of $H$ whose restriction to $W$ is $X$. Given any symmetric generating set $S$ for $G$, the set $S^{(t)} = \cup_{i=1}^{q^{nt}} S^i$ is $t$-transversal if $G = \mathrm{SL}(V)$ and $t < n$ (see [**5**, Corollary 2.4]) and is singularly $t$-transversal if $G \neq \mathrm{SL}(V)$ and $t \leq (n-2)/5$ (see [**5**, Corollary 6.8]).

The proof of Theorem 1.1 consists of two steps. The first step is Proposition 2.3, which is [**5**, Proposition 5.5] and [**5**, Proposition 7.7] with different bounds.

PROPOSITION 2.3. *There are universal positive constants $\gamma_4$ and $\gamma_5$ such that for any symmetric generating set $S$ in $G$ where $G$ is any of the groups $\mathrm{SL}(n,q)$, $\mathrm{Sp}(n,q)$, $\mathrm{SU}(n,q)$, $\Omega(n,q)$, with $n > 2$, there is a non-scalar matrix $A$ in $G$ such that $\deg(A) < \gamma_4((\log n)^2/\log\log n)$ and $A$ may be expressed as the product of less than $q^{\gamma_5 \cdot n \cdot ((\log n)^2/\log\log n)}$ elements from $S$.*

*Proof.* We apply Lemma 2.2 to the argument of Biswas and Yang [**5**].

Let $G = \mathrm{SL}(V) = \mathrm{SL}(n,q)$. We may assume that $n$ is sufficiently large. Put $\gamma_4 = 2\gamma_3$ where $\gamma_3$ is as in the previous subsection and assume that $d$, defined to be the integer part of $\gamma_3((\log n)^2/\log\log n)$, is less than $n$. Since $S^{(d)}$ is a $d$-transversal set for $1 \leq d < n$, there is $A_0 \in S^{(d)}$ that maps some $d$-dimensional subspace $W$ to itself, and the restriction of $A_0$ to $W$ is a diagonal block matrix $C$, where the blocks are companion matrices of irreducible polynomials $f_i$ of degrees $p_i$, and possibly an identity matrix of an appropriate size, such that the $p_i$ range over all primes from 3 to $\bar{p}$ as in Section 2.1 and each $f_i$ has a root of order $q^{p_i}-1$ over $\overline{\mathbb{F}_q}$. Then $A_0$ satisfies the condition of Lemma 2.2, and the length of $A_0$ is bounded by $q^{nd}$. By Lemma 2.2, raise $A_0$ to an appropriate power to obtain a non-identity matrix $A_1$ of degree at most $\deg(A_0)/4$ with eigenvalues being either 1 or outside $\mathbb{F}_q$. The length of $A_1$ is bounded by $q^{nd+n}$ since the order of $A_0$ is bounded by $q^n$. If $\deg(A_1) < 2d$, then we are done. Otherwise, we enter the inductive step. The key in the inductive step is to locate a subspace $W_1$ of dimension $d$ such that $A_1W_1 \cap W_1 = 0$, whose existence is guaranteed by [**5**, Lemma 5.3]. Then use the $2d$-transversal set $S^{(2d)}$ to obtain a matrix $M_1$ of length at most $q^{2nd}$ that fixes $A_1W_1$ pointwise, $W_1$ setwise, and when restricting to $W_1$, realises the diagonal block $C$ as before. The commutator $A_1' = M_1A_1^{-1}M_1^{-1}A_1$ then realises $C$ when restricting on $W_1$, so it satisfies the condition of Lemma 2.2. Furthermore, $\deg(A_1') \leq 2\deg(A_1)$ by [**5**, Proposition 5.2]. By Lemma 2.2, raise $A_1'$ to an appropriate power to get a non-identity matrix $A_2$ such that

$$\deg(A_2) \leq \deg(A_1')/4 \leq 2\deg(A_1)/4 = \deg(A_1)/2.$$

It can be checked that the length of $A_2$ is bounded by

$$2(q^{2nd} + q^{nd+n})q^n \leq q^{2nd+2(n+2)}.$$

Suppose we have obtained a non-scalar matrix $A_j$ with eigenvalues either 1 or outside $\mathbb{F}_q$ with $\deg(A_j) \leq n/2^{j+1}$ and length at most $q^{2nd+j(n+2)}$. If $\deg(A_j)$ is not small enough, then we construct a matrix $A_{j+1}$ of length at most

$$2(q^{2nd} + q^{2nd+j(n+2)})q^n \leq q^{2nd+(j+1)(n+2)}.$$

Repeat this by at most $\log n$ times to reach the desired matrix $A$.

For $G \neq \mathrm{SL}(V)$ the argument is very similar as for $\mathrm{SL}(V)$ above. Here Witt's decomposition theorem (see [**5**, Theorem 6.2]) and Witt's extension lemma (see [**5**, Lemma 6.5]) are used. The latter is that $G$ is a singularly $t$-transversal set for any $t$. Moreover, we mention [**5**, Lemma 7.6]. If $A$ is a matrix in $G$ of degree $d$ such that the eigenvalues of $A$ are either 1 or outside $\mathbb{F}_q$, then there is a totally singular subspace $W$ of $V$ such that $W \cap AW = \{0\}$, $W \perp AW$, and $\dim W \geq (d/32) - (7/4)$. □

Given a non-scalar matrix $A$ of degree $d$ and length $\ell$, the second step is to show that the diameter of $G$ with respect to $S$ is bounded by $O((q^{2nd} + \ell) \cdot \frac{n}{d})$ (cf. [**5**, Proposition 8.3]). This is due to the following. Firstly, any conjugate of $A$ can be obtained by conjugating by a matrix of length less than $q^{2nd}$, as the number of conjugates of $A$ is bounded by such (see [**5**, Lemma 8.1]). Secondly, by Liebeck and Shalev [**18**], every element in $G$ is a product of at most $O(n/d)$ conjugates of $A$.

We may take $d$ to be less than $\gamma_4((\log n)^2/\log\log n)$ and $\ell$ to be less than $q^{\gamma_5 \cdot n \cdot ((\log n)^2/\log\log n)}$ by Proposition 2.3. Then

$$\mathrm{diam}_S(G) \leq O((q^{2nd} + \ell) \cdot \frac{n}{d}) \leq q^{O(n(\log n)^2)}.$$

This completes the proof of Theorem 1.1 (modulo Lemma 2.2).

We remind the reader that in the above procedure, the exponent with respect to the base $q$ in the length bound of $A$ is always bounded by $O(nd)$. It follows that the $\log q$ term does not appear in the exponent if $d = O((\log n)^2)$.

### 2.3. *Proof of Lemma 2.2*

We first need the following preparations.

FACT 2.4. *Let $f = f(x) \in \mathbb{F}_q[x]$ be an irreducible monic polynomial of degree $d$. Let $C_f \in \mathrm{GL}(d, q)$ be its companion matrix.*

(i) *For any $a \in \mathbb{N}$, $C_f^{p^a}$ is similar to the companion matrix of an irreducible polynomial in $\mathbb{F}_q[x]$ of degree $d$.*

(ii) *For $m \in \mathbb{N}$, $C_f^{q^m-1} = I$ if and only if $d \mid m$.*

*Proof.* (1) First observe that $(f(x))^{p^a} = \tilde{f}(x^{p^a})$ where $\tilde{f}$ is the polynomial obtained by raising every coefficient of $f$ to the $p^a$th power. Then we can verify that $\mathrm{charpoly}(C_f^{p^a}, x) = \tilde{f}(x)$.

(2) Recall that $\mathrm{minpoly}(C_f, x) = \mathrm{charpoly}(C_f, x) = f(x)$, and $f(x) \mid x^{q^m} - x$ if and only if $d \mid m$. The claim then follows. □

THEOREM 2.5 ([**20**], generalized Jordan normal form). *Let $\mathbb{F}$ be a perfect field, and $A \in M(n, \mathbb{F})$. Suppose* $\mathrm{charpoly}(A, x)$ *decomposes into a product of irreducible monic polynomials as $f_1^{e_1} \cdot \ldots \cdot f_k^{e_k}$, where $f_i \in \mathbb{F}[x]$ is of degree $d_i$. Then $A$ is similar to a block diagonal matrix*

$\mathrm{Diag}(J_1, \ldots, J_\ell)$, *where each* $J_i$, *called a (generalized) Jordan block, is of the form*

$$\begin{bmatrix} C_{f_{b_i}} & I & 0 & \ldots & 0 & 0 \\ 0 & C_{f_{b_i}} & I & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & I & 0 \\ 0 & 0 & 0 & \ldots & C_{f_{b_i}} & I \\ 0 & 0 & 0 & \ldots & 0 & C_{f_{b_i}} \end{bmatrix}, \tag{2.1}$$

*where* $b_i \in [k]$, $I$ *is the identity matrix of size* $d_{b_i}$, *and* $0$ *is the all-zero matrix of size* $d_{b_i} \times d_{b_i}$.

We are ready to prove Lemma 2.2.

*Proof of Lemma 2.2.* Suppose $\mathrm{charpoly}(A, x)$ decomposes into a product of irreducible monic polynomials as

$$f_1^{e_1} \cdot \ldots \cdot f_k^{e_k} \cdot (x-1)^t,$$

where $t \in \mathbb{N}$, $e_i \in \mathbb{Z}^+$, $f_i \in \mathbb{F}[x]$ is irreducible, monic, and of degree $d_i$, and $f_i \neq x-1$ for $i \in [k]$. Observe that $n = t + \sum_{i=1}^{k} d_i e_i$. Let $f_0 = x-1$, and $s = n - t$. Clearly, $\deg(A) \geq s$. By the assumption of this lemma, we have that $k \geq r$, and $\deg(f_i) = p_i$ for $i \in [r]$. It is possible that $\deg(f_{i'}) = p_i$ for some $r < i' \leq k$, but this does not concern us and we can order $f_j$, $r < j \leq k$, arbitrarily.

For our purpose, we can replace $A$ with any of its conjugates. Therefore by Theorem 2.5, we assume $A = \mathrm{Diag}(J_1, \ldots, J_\ell)$ where each $J_i$ is a Jordan block of the form (2.1).

We first raise $A$ to the $p^a$th power, where $a$ is an integer larger than $\log n$. This operation helps to eliminate the identity blocks as in Eq. 2.1. That is, for any $i \in [\ell]$, $J_i^{p^a} \cong \mathrm{Diag}(C_{\tilde{f}_{b_i}}, \ldots, C_{\tilde{f}_{b_i}})$ for some $b_i \in \{0, 1, \ldots, k\}$, where $\tilde{f}_{b_i}$ is an irreducible polynomial of degree $d_{b_i}$ by Fact 2.4 (1). Let $\tilde{A} = A^{p^a}$. By arranging the diagonal blocks via conjugation transformations, we can assume that

$$\tilde{A} = \mathrm{Diag}(C_{\tilde{f}_1}, \ldots, C_{\tilde{f}_1}, \ldots, C_{\tilde{f}_k}, \ldots, C_{\tilde{f}_k}, 1, \ldots, 1),$$

where the number of $C_{\tilde{f}_i}$ is $e_i$, and the number of 1 is $t$. In particular, $\deg(\tilde{A}) = s \leq \deg(A)$.

For $j \in [s]$, let $c_j \in [k]$ be such that the $j$th diagonal position of $\tilde{A}$ is in the diagonal block $C_{\tilde{f}_{c_j}}$. For example, for $0 < j \leq d_1 e_1$, $c_j = 1$. For $d_1 e_1 < j \leq d_1 e_1 + d_2 e_2$, $c_j = 2$. That is, $c_j = \ell$ if and only if $\sum_{i=1}^{\ell-1} d_i e_i < j \leq \sum_{i=1}^{\ell} d_i e_i$. We then build a zero-one matrix $D$ of size $r \times s$ as follows. For $i \in [r]$ and $j \in [s]$, $D(i,j) = 1$ if $p_i | d_{c_j}$, and 0 otherwise. We then deduce the following.

(a) For any $j \in [s]$, $\prod_{i \in [r]} p_i^{D(i,j)} \leq d_{c_j} \leq s$.

(b) For $i \in [r]$, let $n_i = \sum_{j \in [s]} D(i,j)$. We claim that there exists $i' \in [r]$, such that $n_{i'} \leq s/4$. For this, consider the weighted average $W$ of $n_i$ with weights $\log p_i$. We have

$$\begin{aligned} W &= \frac{\sum_{i \in [r]} n_i \log p_i}{\sum_{i \in [r]} \log p_i} = \frac{\sum_{j \in [s]} \sum_{i \in [r]} D(i,j) \log p_i}{\sum_{i \in [r]} \log p_i} \\ &\leq \frac{\sum_{j \in [s]} \sum_{i \in [r]} D(i,j) \log p_i}{4 \log n} \\ &= \frac{\sum_{j \in [s]} \log(\prod_{i \in [r]} p_i^{D(i,j)})}{4 \log n} \\ &\leq \frac{s \cdot \log s}{4 \log n} \leq \frac{s}{4}. \end{aligned}$$

In the above, the first $\leq$ is due to the choice of the $p_i$, namely we have chosen those $p_i$ to satisfy $\prod_{i\in[r]} p_i > n^4$. The second $\leq$ is due to item (a) we just described. The existence of such $i' \in [r]$ satisfying $n_{i'} \leq s/4$ then follows. In other words, the number of diagonal positions in those $C_{\tilde{f}_i}$ with $p_{i'} \mid d_i$ is bounded from above by $s/4$.

Let $i' \in [r]$ be an index satisfying (b) and let $p' = p_{i'}$. Let $s'$ be the lowest common multiple of these $\deg(f_i)$ which are coprime to $p'$. Then $\hat{A} = \tilde{A}^{q^{s'}-1}$ satisfies the following: Firstly, $\hat{A}$ is not identity. This is because the existence of $C_{\tilde{f}_{i'}}$ where $\deg(\tilde{f}_{i'}) = p_{i'}$ and Fact 2.4 (2). Secondly, $\hat{A}$ is of degree at most $s/4 \leq \deg(A)/4$. This is because for any $C_{\tilde{f}_i}$ with $p_{i'} \nmid \deg(\tilde{f}_i)$, $C_{\tilde{f}_i}^{q^{s'}-1}$ becomes identity by Fact 2.4 (2), and from (b) we know the sum of the sizes of such blocks is at least $(3s)/4$. This shows the existence of $m \in \mathbb{N}$ as required.

We now prove the statement of the furthermore part in Lemma 2.2. For this, it is sufficient to show that a $(q-1)^k$-power of $\hat{A}$, for some integer $k$, is not the identity matrix. Suppose otherwise. If the $(q-1)^k$-power of $\hat{A}$ is the identity, then, in particular, the $(q^{s'}-1)(q-1)^k$-power of the companion matrix $C_{\tilde{f}_i}$ is the identity. Since $C_{\tilde{f}_i}$ has order $q^{p'}-1$ by the assumption on the root orders of $f_i$ (taking the $p^a$-power of $A$ does not do harm), we must have $q^{p'}-1 \mid (q^{s'}-1)(q-1)^k$. Since $p' \nmid s'$, the greatest common divisor of $q^{p'}-1$ and $q^{s'}-1$ is $q-1$, from which it follows that $q^{p'}-1 \mid (q-1)^{k+1}$. By applying the following Claim 2.6, with $p'$ in place of $t$ and noting that $p' > 2$, we arrive to a contradiction.

CLAIM 2.6. *Let $t$ be a prime and $q$ an integer larger than 1. If $q^t - 1$ divides some power of $q-1$, then $t = 2$.*

*Proof.* Notice that the condition $q^t - 1$ divides some power of $q-1$ is equivalent to the condition that every prime divisor of $q^t-1$ divides $q-1$.

We claim that $\frac{q^t-1}{q-1} = t^s$ for some integer $s \geq 2$. Let $r$ be a prime divisor of $(q^t-1)/(q-1)$. Then $r \mid q-1$ by the condition $(q^t-1) \mid (q-1)^z$ for some $z \in \mathbb{N}$ and, since $q^{t-1} + \cdots + q + 1$ is congruent to $t$ modulo $q-1$, the primes $r$ and $t$ must be equal. This proves that $\frac{q^t-1}{q-1} = t^s$ for some integer $s \geq 1$. We also have $s \geq 2$ by $q > 1$.

On the other hand,

$$\frac{q^t-1}{q-1} = \frac{((q-1)+1)^t - 1}{q-1} = \sum_{k=1}^{t} \binom{t}{k} (q-1)^{k-1}$$

is congruent to $(q-1)^{t-1} + t$ modulo $t^2$, as the intermediate terms $\binom{t}{k}(q-1)^{k-1}$, $1 < k < t$, are divisible by $t^2$ if they do appear. Since $t^2$ does not divide $t$, it cannot divide $(q-1)^{t-1}$ either. This forces $t = 2$ as $t$ divides $q-1$ by our condition. □

This concludes the proof of Lemma 2.2. □

## 3. *A structure theorem for completely reducible groups*

In this section we will prove Theorem 3.3 which, in the next section, will be used to deduce Theorem 1.2 (in case the group acts completely reducibly on its module).

Let us fix some notation. Let $V$ be the vector space of dimension $n$ over $\mathbb{F}_q$. Let $G$ be a subgroup of $\mathrm{GL}(V)$ acting completely reducibly on $V$. The $G$-module $V$ is the direct sum $V_1 \oplus \cdots \oplus V_m$ of irreducible $G$-modules $V_i$ with $1 \leq i \leq m$. It is natural to write each vector space $V_i$ as a direct sum $W_{i1} \oplus \cdots \oplus W_{ik_i}$ of isomorphic vector spaces $W_{ij}$ with $1 \leq j \leq k_i$

such that $\{W_{i1}, \ldots, W_{ik_i}\}$ is preserved by the action of $G$ and with $k_i$ as large as possible. It follows that for each pair $(i, j)$ the stabilizer of $W_{ij}$ in $G$ acts irreducibly and primitively (but not necessarily faithfully) on $W_{ij}$.

To simplify notation, write the vector space $V$ as a direct sum $W_1 \oplus \cdots \oplus W_k$ such that $G$ preserves $\Omega = \{W_1, \ldots, W_k\}$, the stabilizer $G_i$ of $W_i$ in $G$ acts irreducibly and primitively on $W_i$ for each $i$ with $1 \leq i \leq k$ and $k = \sum_{i=1}^{m} k_i$ in the above notation. For each $i$ let the action of $G_i$ on $W_i$ be $P_i$. The group $G$ is a subgroup of $(P_1 \times \cdots \times P_k) : \mathrm{S}_k$. Let $N$ denote the intersection of $G$ with $P_1 \times \cdots \times P_k$, that is, the kernel of the action of $G$ on $\Omega$. The factor group $G/N$ may be viewed as a subgroup of $\mathrm{S}_k \leq \mathrm{S}_n$.

We continue with a slightly simplified version of [**15**, Proposition 5.7]. Here a *quasisimple* group is a finite perfect group $H$ such that $H/Z(H)$ is simple. The *generalized Fitting subgroup* $F^*(X)$ of a finite group $X$ is defined to be the product of the Fitting subgroup of $X$ and all subnormal quasisimple subgroups of $X$.

THEOREM 3.1 (Jaikin-Zapirain, Pyber; 2011). *Let $Q$ be a subgroup of $\mathrm{GL}(W)$ with $Q$ acting irreducibly and primitively on the finite vector space $W$ defined over the prime field $\mathbb{F}_p$. For the generalized Fitting subgroup $F^*(Q)$ of $Q$ let $F$ be the field $Z(\mathrm{End}_{F^*(Q)}(W))$. There exists a universal constant $\gamma_6$ such that whenever $|Q| > |W|^{\gamma_6}$, then*

*(i) there is a tensor product decomposition $U' \otimes_F U$ of $W$ such that $\dim(U) \geq \dim(U')$;*
*(ii) there is a quasisimple normal subgroup $R$ in $Q$ isomorphic to $\mathrm{A}_\ell$ or to a classical group $\mathrm{Cl}(d, K)$ for some $K \leq F$ and $d \geq 2$;*
*(iii) if $R = \mathrm{A}_\ell$, then $U$ is the natural $\mathrm{A}_\ell$-module, while if $R = \mathrm{Cl}(d, K)$, then $U$ is $F \otimes_K U''$ where $U''$ is the natural $\mathrm{Cl}(d, K)$-module;*
*(iv) $|Q/R| \leq |W|^5$.*

Let $P$ be a subgroup of $\mathrm{GL}(W)$ acting irreducibly and primitively on the finite vector space $W$ defined over the field $\mathbb{F}_q$ (possibly different from its prime field $\mathbb{F}_p$). It centralizes a cyclic subgroup $Z$ of $\mathrm{GL}(W)$ isomorphic to $\mathbb{F}_q^*$. According to a claim of Liebeck and Shalev (see [**19**, p. 112]) $PZ$ acts irreducibly and primitively on $W$ viewed over the field $\mathbb{F}_p$. For the sake of completeness, we present a proof for this fact. If $U$ is a $PZ$-invariant subspace of $W$, then $U$ must be an $\mathbb{F}_q$-space. Thus $PZ$ acts irreducibly on $W$. For a contradiction, assume that $W = W_1 + \cdots + W_t$ is an imprimitivity decomposition of the $PZ$-module $W$ over $\mathbb{F}_p$ where $t > 1$. Let $Z_0$ be the stabilizer of $W_1$ in $Z$. Clearly $Z_0 < Z$ since otherwise the $W_i$ are $\mathbb{F}_q$-spaces contradicting the fact that $P$ acts primitively on $W$ viewed over $\mathbb{F}_q$. Let $z$ be an element of $Z$ outside $Z_0$. Without loss of generality, assume that $z$ maps $W_1$ to $W_2$. Since $W_1$ is an $\mathbb{F}_p$-space, it follows that $z \notin \mathbb{F}_p$. In particular, $1 + z \neq 0$ and so $1 + z \in Z$. For a non-zero vector $w_1$ in $W_1$ we have $w_1(1 + z) = w_1 + w_1 z \in W_1 + W_2$. Since $w_1 \neq 0$, the element $w_1(1 + z)$ is neither in $W_1$ nor in $W_2$. This is a contradiction.

As a Corollary to Theorem 3.1 we obtain the surprising fact that primitive linear groups are not far from being simple groups. To simplify notation, let $\mathrm{Cl}(d, r)$ be $\mathrm{Cl}(d, \mathbb{F}_r)$ for any prime power $r$.

THEOREM 3.2. *If $P$ is a subgroup of $\mathrm{GL}(W)$ with $P$ acting irreducibly and primitively on the finite vector space $W$ defined over the field $\mathbb{F}_q$ with $|P| > |W|^{\gamma_6}$ where $\gamma_6$ is as in Theorem 3.1, then there is a quasisimple normal subgroup $R$ in $P$ isomorphic to $\mathrm{A}_\ell$ such that*

*$\ell \leq \dim_{\mathbb{F}_q}(W)$ or to a classical group $\mathrm{Cl}(d,r)$ such that $d \leq \dim_{\mathbb{F}_q}(W)$ with $\mathbb{F}_r$ and $\mathbb{F}_q$ of the same characteristic, and $|P/R| \leq |W|^5$. Moreover, if $R$ is isomorphic to $\mathrm{Cl}(d,r)$, then $r^d \leq |W|$.*

*Proof.* By Theorem 3.1 and the claim of Liebeck and Shalev (see the paragraph after Theorem 3.1), there is a quasisimple normal subgroup $R$ in $PZ$ isomorphic to $\mathrm{A}_\ell$ such that $\ell \leq \dim_{\mathbb{F}_q}(W)$ or to a classical group $\mathrm{Cl}(d,r)$ such that $d \leq \dim_{\mathbb{F}_q}(W)$ (the bounds for $\ell$ and $d$ follow from the fact that the field $F$ in Theorem 3.1 contains $\mathbb{F}_q$) and $\mathbb{F}_r$ and $\mathbb{F}_q$ have the same characteristic. In the latter case we have $r^d \leq |W|$ by Theorem 3.1. It also follows that $|PZ/R| \leq |W|^5$. Since $R$ is quasisimple, $R = [R,R] \leq [PZ,PZ] \leq P$. This completes the proof of the theorem. □

We are now in position to prove our structure theorem.

THEOREM 3.3. *Let $V$ be a vector space of dimension $n$ over the field $\mathbb{F}_q$. Let $G \leq \mathrm{GL}(V)$ be a group acting completely reducibly on $V$. Write $V$ as a direct sum $W_1 \oplus \cdots \oplus W_k$ of (non-trivial) subspaces of $V$ in such a way that $G$ permutes the set $\Omega = \{W_1, \ldots, W_k\}$ and the stabilizer of each $W_i$ in $G$ acts primitively on $W_i$ for every $i$ with $1 \leq i \leq k$. Let $N$ be the kernel of the action of $G$ on $\Omega$. In particular, $G/N$ may be viewed as a subgroup of $\mathrm{S}_n$. There exists a constant $\gamma_7$ such that whenever $|N| > |V|^{\gamma_7}$,*

*(i) there is a normal subgroup $C$ of $G$ contained in $N$ such that $C = Q_1 \circ \cdots \circ Q_w$ is a central product of quasisimple groups $Q_i$ with $w \leq k$;*

*(ii) each $Q_i$ has a factor group $T_i$ such that for some $j \in \{1, \ldots, k\}$, $T_i$ is an alternating group $\mathrm{A}_{\ell_j}$ with $\ell_j \leq \dim_{\mathbb{F}_q}(W_j)$, or $T_i$ is a classical simple group $\mathrm{Cl}(d_j, r_j)$ such that $\mathbb{F}_{r_j}$ and $\mathbb{F}_q$ have the same characteristic, $d_j \leq \dim_{\mathbb{F}_q}(W_j)$ and $r_j^{d_j} \leq |W_j|$;*

*(iii) $|N/C| \leq |V|^{\gamma_7}$.*

*Proof.* For each $i$ let the stabilizer of $W_i$ in $G$ be $G_i$ and let the action of $G_i$ on $W_i$ be $P_i$. Let $\gamma_7$ be the maximum of 6 and $\gamma_6$. Without loss of generality, we may assume that there is an integer $t \geq 0$ such that $|P_i| > |W_i|^{\gamma_7}$ for every $i$ with $i \leq t$ and $|P_i| \leq |W_i|^{\gamma_7}$ for every $i$ with $t < i \leq k$. For every $i$ with $i \leq t$, let $R_i$ be the quasisimple normal subgroup of $P_i$ whose existence is assured by Theorem 3.2 (and is $R$ in that notation).

If $N$ denotes the intersection of $G$ with $P_1 \times \cdots \times P_k$, that is, the kernel of the action of $G$ on $\Omega$, then the factor group $G/N$ may be viewed as a subgroup of $\mathrm{S}_k \leq \mathrm{S}_n$. In order to prove the theorem, we may assume that $|N| > |V|^{\gamma_7}$. In particular, $t > 0$.

Let $M$ be the normal subgroup of $G$ defined to be the intersection of $N$ and $R_1 \times \cdots \times R_t$. Since the image $M_i$ of the natural projection from $M$ to $P_i$ is normal in $P_i$, the group $M_i$ must also be normal in $R_i$. Since $R_i$ is quasisimple, $M_i = R_i$ or $M_i$ is central in $R_i$. In the latter case $|M_i| < |W_i|$. Without loss of generality, we may assume that there is a $u \geq 0$ such that $M_i = R_i$ for every index $i$ at most $u$ and $M_i$ is abelian for $i > u$. Thus the commutator subgroup $M'$ may be viewed as a subgroup of $R_1 \times \cdots \times R_u$ (where $u \geq 0$) which projects onto $R_i$ for every $i$ with $i \leq u$. Clearly, $|N/M'| \leq |V|^{\gamma_7}$ by Theorem 3.2.

We may thus assume that $M' \neq 1$, that is, $u \geq 1$. Now $M'/Z(M')$ may be viewed as a subgroup of $F_1 \times \cdots \times F_u$ where $F_i = R_i/Z(R_i)$ is a non-abelian simple group for every $i$ with $1 \leq i \leq u$. Moreover, $M'/Z(M')$ projects onto every $F_i$. It follows, by [**23**, p. 328, Lemma], that $M'/Z(M')$ is a direct product $\prod_{j=1}^{w} D_j$ of full diagonal subgroups $D_j$ of subproducts $\prod_{i \in I_j} F_i$ where the $I_j$ form a partition of $\{1, \ldots, u\}$. The preimage in $M'$ of any simple factor $D_j$ of $M'/Z(M')$ contains a normal quasisimple subgroup of $M'$ which is subnormal in $G$. Let $C$ be the product of all components, that is, all subnormal quasisimple subgroups, of $G$ contained

in the group $M'$. Since any two distinct components in a finite group commute, $C$ may be expressed in the form $Q_1 \circ \cdots \circ Q_w$ where the $Q_j$ are components of $G$ contained in $M'$.

The group $C$ is normal in $G$ and so (i) is established. Since $C \cdot Z(M') = M'$, it is easy to see that there is a refinement of our previous bound for $|N/M'|$ in the form $|N/C| \leq |V|^{\gamma_7}$. This is (iii).

Fix an index $i$ at most $w$. The component $Q_i$ projects onto $F_j$ for some $j$ at most $u$. The group $F_j$ is isomorphic to $\mathrm{A}_{\ell_j}$ such that $\ell_j \leq \dim_{\mathbb{F}_q}(W_j)$ or to a classical simple group $\mathrm{Cl}(d_j, r_j)$ such that $d_j \leq \dim_{\mathbb{F}_q}(W_j)$, $r_j^{d_j} \leq |W_j|$, and $\mathbb{F}_{r_j}$ and $\mathbb{F}_q$ have the same characteristic, by Theorem 3.2. Thus $Q_i$ has a factor group $T_i$ such that $T_i$ is $\mathrm{A}_{\ell_j}$ or $T_i$ is the classical simple group $\mathrm{Cl}(d_j, r_j)$. This gives (ii).

This completes the proof of the theorem. □

## 4. *A bound for* $\mathrm{diam}(G)$ *for* $G$ *a linear group*

In this section we prove Theorem 1.2.

A main tool in our argument is Lemma 5.1 of Babai and Seress [**4**].

LEMMA 4.1 (Babai, Seress; 1992). *If $N$ is a non-trivial, proper normal subgroup in a finite group $G$, then* $\mathrm{diam}(G) \leq 4 \cdot \mathrm{diam}(N) \cdot \mathrm{diam}(G/N)$.

We also need Theorem 1.3 of [**4**] mentioned in the third paragraph of the Introduction.

THEOREM 4.2 (Babai, Seress; 1992). *If $G$ is a permutation group of degree $n$, then* $\mathrm{diam}(G) < \exp((n \ln n)^{1/2}(1 + o(1)))$.

Now let $G$ be a subgroup of $\mathrm{GL}(V)$ acting on the finite vector space $V$ of dimension $n$ over the field of size $q$ and characteristic $p$. In case $h \neq 1$ let $S$ be a classical (non-abelian) composition factor of $G$ defined over a field of characteristic $p$ such that $h = \mathrm{diam}(S)$.

First assume that $G$ acts completely reducibly on $V$. In this case we rely on Theorem 3.3 to prove Theorem 1.2.

We use the notation of Theorem 3.3. Write $V$ as a direct sum $W_1 \oplus \cdots \oplus W_k$ of (non-trivial) subspaces of $V$ in such a way that $G$ permutes the set $\Omega = \{W_1, \ldots, W_k\}$ and the stabilizer of each $W_i$ in $G$ acts primitively on $W_i$ for every $i$ with $1 \leq i \leq k$. Let $N$ be the kernel of the action of $G$ on $\Omega$. In particular, $G/N$ may be viewed as a subgroup of $\mathrm{S}_n$. First, $\mathrm{diam}(G/N)$ is less than exponential in $n$ by Theorem 4.2. Thus, in order to establish our bound for $\mathrm{diam}(G)$, it is sufficient to show that $\mathrm{diam}(N) < |V|^{O(1)}h^2 < q^{O(n(\log n)^2)}$ by Lemma 4.1. This is certainly true in case $|N| \leq |V|^{\gamma_7}$ where $\gamma_7$ is as in Theorem 3.3. Thus assume that $|N| > |V|^{\gamma_7}$. Let $C$ be the normal subgroup of $G$, as in Theorem 3.3, such that $|N/C| \leq |V|^{\gamma_7}$. It follows by Lemma 4.1 that it is sufficient to show that

$$\mathrm{diam}(C) < |V|^{O(1)}h^2 < q^{O(n(\log n)^2)}. \tag{4.1}$$

We claim that $h < |V|^{O(1)}$ or $S$ is a composition factor of $C$. For a proof we may assume that $S$ is a composition factor of $G/N$. A standard argument using the classical result of Praeger and Saxl [**21**] shows that $h < |S| < |V|^{O(1)}$.

Since $C$ is normal in $G$ and the center $Z(C)$ of $C$ is characteristic in $C$, the group $Z(C)$ is normal in $G$. Since $G$ acts completely reducibly on $V$, so does the abelian group $Z(C)$, by Clifford's theorem. By Schur's lemma and the fact that a finite division ring is a field, an abelian group $A \leq \mathrm{GL}(W)$ acting irreducibly on a finite vector space $W$ is cyclic and has order at most $|W| - 1$. From these it follows that $|Z(C)| < |V|$.

The factor group $C/Z(C)$ is the direct product of non-abelian simple groups each isomorphic to an alternating group or to a classical group in characteristic $p$. Let $A$ be the product of all factors of $C/Z(C)$ which are isomorphic to alternating groups, if such exist, otherwise let $A = 1$. Let $B$ be the product of all other simple factors of $C/Z(C)$, that is, $C/Z(C) = A \times B$. In view of (4.1), the bound $Z(C) < |V|$ and Lemma 4.1, it is sufficient to establish the bound

$$\mathrm{diam}(A \times B) < |V|^{O(1)} h^2 < q^{O(n(\log n)^2)}. \tag{4.2}$$

We first handle $A$ and $B$ separately as follows. The sum of degrees of all simple factors in $A$, if such exist, is at most $n$ by Theorem 3.3. Hence $A$ may be considered as a permutation group of degree at most $n$ and hence $\mathrm{diam}(A) < O(1)|V|$ by Theorem 4.2. We have $\mathrm{diam}(B) \leq 20\ n^3\ h^2 < |V|^{O(1)} h^2$ by [**4**, Lemma 5.4].

We claim that $h = q^{O(n(\log n)^2)}$. We may assume by the above that $S$ is a composition factor of $C$ (and a direct factor of $B$). In this case $S$ is isomorphic to the non-abelian composition factor $S_i$ of some component $Q_i$ of $G$ (normal in $C$). The group $S_i$ is a simple classical group of dimension $d_j$ defined over the field $\mathbb{F}_{r_j}$, for some $j$. Thus $h = r_j^{O(d_j(\log d_j)^2)}$ by Theorem 1.1. Since $d_j \leq n$ and $r_j^{d_j} \leq q^n$, we conclude that $r_j^{O(d_j(\log d_j)^2)} = q^{O(n(\log n)^2)}$.

Now (4.2) follows by Lemma 4.1 and the previous two paragraphs. This completes the proof of Theorem 1.2 when $G$ acts completely reducibly on $V$.

Now let $G$ be an arbitrary subgroup of $\mathrm{GL}(V)$. Let $O_p(G)$ denote the largest normal $p$-subgroup of $G$. The factor group $G/O_p(G)$ may be viewed as a completely reducible linear group acting on the direct sum of the composition factors of the $G$-module $V$. Thus

$$\mathrm{diam}(G/O_p(G)) < |V|^{O(1)} h^2 < q^{O(n(\log n)^2)} \tag{4.3}$$

by the above.

In order to complete the proof of Theorem 1.2, it is sufficient, by Lemma 4.1 and (4.3), to show that $\mathrm{diam}(P) < |V|^{O(1)}$ for every $p$-subgroup $P$ of $\mathrm{GL}(V)$.

Let $Q$ be a $p$-group and $\mathcal{C}$ a normal chain in $Q$ such that every associated factor in the chain $\mathcal{C}$ is elementary abelian. Let $\ell(Q, \mathcal{C})$ be the length of the chain $\mathcal{C}$ and let $r(Q, \mathcal{C})$ be the maximum rank of the associated factors in $\mathcal{C}$. It is easy to see that

$$\mathrm{diam}(Q) \leq 4^{\ell(Q,\mathcal{C})-1} \cdot (p \cdot r(Q, \mathcal{C}))^{\ell(Q,\mathcal{C})} \tag{4.4}$$

using Lemma 4.1 and Lemma 5.2 of [**4**].

Now let $P$ be an arbitrary $p$-subgroup of $\mathrm{GL}(V)$. This is a subgroup of a Sylow $p$-subgroup $S$ of $\mathrm{GL}(m, q)$ where $m$ is the smallest power of 2 which is larger than $n$. We have

$$\mathrm{diam}(P) \leq 4^{\ell(S,\mathcal{C})-1} \cdot (p \cdot r(S, \mathcal{C}))^{\ell(S,\mathcal{C})} \tag{4.5}$$

by (4.4), for any chain $\mathcal{C}$ of normal subgroups in $S$ such that the associated factor groups are elementary abelian.

There exists an elementary abelian normal subgroup $A$ in $S$ such that $|A| = q^{m^2/4}$ and $S/A$ is the direct product of two copies of a Sylow $p$-subgroup in $\mathrm{GL}(m/2, q)$. It follows, by induction on $m$, that there is a chain $\mathcal{C}$ of normal subgroups in $S$ such that

(i) the associated factor groups are elementary abelian;
(ii) the first group is $A$;
(iii) $r(S, \mathcal{C}) = (m^2/4) \cdot \log_p q \leq n^2 \cdot \log_p q$; and
(iv) $\ell(S, \mathcal{C}) = 1 + \log_2 m \leq 2 + \log_2 n$.

From this and (4.5) it follows that $\mathrm{diam}(P) < |V|^{O(1)}$.

This completes the proof of Theorem 1.2.

## *References*


**1.** Babai, L. On the diameter of Eulerian orientations of graphs. Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, 822–831, ACM, New York, 2006.
**2.** Babai, L. and Seress, Á. On the degree of transitivity of permutation groups: a short proof. *J. Combin. Theory Ser. A* **45** (1987), no. 2, 310–315.
**3.** Babai, L. and Seress, Á. On the diameter of Cayley graphs of the symmetric group. *J. Combin. Theory Ser. A* **49** (1988), no. 1, 175–179.
**4.** Babai, L. and Seress, Á. On the diameter of permutation groups. *European J. Combin.* **13** (1992), no. 4, 231–243.
**5.** Biswas, A. and Yang, Y. A diameter bound for finite simple groups of large rank. *J. Lond. Math. Soc.* (2) **95** (2017), no. 2, 455–474.
**6.** `https://www.math.u-psud.fr/~breuilla/BreuillardICMtalk.pdf`.
**7.** Breuillard, E.; Green, B.; Tao, T. Approximate subgroups of linear groups. *Geom. Funct. Anal.* **21** (2011), no. 4, 774–819.
**8.** Breuillard, E. and Tointon, M. C. H. Nilprogressions and groups with moderate growth. *Adv. Math.* **289** (2016), 1008–1055.
**9.** Erdős, P. Beweis eines Satzes von Tschebyschef. *Acta Litt. Sci. Szeged* **5** (1932), 194–198.
**10.** Grigorchuk, R. I. On growth in group theory. *Proceedings of the International Congress of Mathematicians*, Vol. I, II (Kyoto, 1990), 325–338, Math. Soc. Japan, Tokyo, 1991.
**11.** Gromov, M. Groups of polynomial growth and expanding maps. *Inst. Hautes tudes Sci. Publ. Math.* No. **53** (1981), 53–73.
**12.** Helfgott, H. A. Growth and generation in $SL_2(Z/pZ)$. *Ann. of Math.* (2) **167** (2008), no. 2, 601–623.
**13.** Helfgott, H. A. Growth in linear algebraic groups and permutation groups: towards a unified perspective. ArXiv:1804.03049.
**14.** Helfgott, H. A. and Seress, Á. On the diameter of permutation groups. *Ann. of Math.* (2) **179** (2014), no. 2, 611–658.
**15.** Jaikin-Zapirain, A. and Pyber, L. Random generation of finite and profinite groups and group enumeration. *Ann. of Math.* (2) **173** (2011), no. 2, 769–814.
**16.** Kornhauser, D.; Miller, G.; Spirakis, P. Coordinating pebble motion on graphs, the diameter of permutation groups, and applications. Proceedings of the 25th IEEE Symposium on Foundations of Computer Science, Singer Island, FL, IEEE Computer Society Press, New York (1984), pp. 241–250.
**17.** Lubotzky, A. and Mann, A. On groups of polynomial subgroup growth. *Invent. Math.* **104** (1991), no. 3, 521–533.
**18.** Liebeck, M. W. and Shalev, A. Diameters of finite simple groups: sharp bounds and applications. *Ann. of Math.* (2) **154** (2001), no. 2, 383–406.
**19.** Liebeck, M. W. and Shalev, A. Bases of primitive linear groups. *J. Algebra* **252** (2002), no. 1, 95–113.
**20.** Mal'cev, A. I. Foundations of linear algebra. W. H. Freeman, San Francisco, Calif.-London 1963.
**21.** Praeger, C. E. and Saxl, J. On the orders of primitive permutation groups. *Bull. London Math. Soc.* **12** (1980), no. 4, 303–307.
**22.** Pyber, L. and Szabó, E. Growth in finite simple groups of Lie type. *J. Amer. Math. Soc.* **29** (2016), no. 1, 95–146.
**23.** Scott, L. L. Representations in characteristic $p$, The Santa Cruz Conference on Finite Groups (Univ. California, Santa Cruz, Calif., 1979) *Proc. Sympos. Pure Math.*, Vol. **37**, Amer. Math. Soc., Providence, R.I., 1980, pp. 319–331.
**24.** Shalom, Y. and Tao, T. A finitary version of Gromov's polynomial growth theorem. *Geom. Funct. Anal.* **20** (2010), no. 6, 1502–1547.
**25.** Wilson, J. S. On the growth of residually soluble groups. *J. London Math. Soc.* (2) **71** (2005), no. 1, 121–132.

**26.** http://www.dipmat2.unisa.it/ischiagrouptheory/IGT2010/talks/Wilson.pdf.
**27.** Wilson, J. S. The gap in the growth of residually soluble groups. *Bull. Lond. Math. Soc.* **43** (2011), no. 3, 576–582.

*Zoltán Halasi*
*Department of Algebra and Number Theory, Eötvös University, Pázmány Péter Sétány 1/c, H-1117, Budapest and Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Reáltanoda utca 13-15, H-1053, Budapest*
*Hungary*

zhalasi@cs.elte.hu and
halasi.zoltan@renyi.mta.hu

*Attila Maróti*
*Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Reáltanoda utca 13-15, H-1053, Budapest*
*Hungary*

maroti.attila@renyi.mta.hu

*László Pyber*
*Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Reáltanoda utca 13-15, H-1053, Budapest*
*Hungary*

pyber.laszlo@renyi.mta.hu

*Youming Qiao*
*Centre for Quantum Software and Information, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007*
*Australia*

Youming.Qiao@uts.edu.au