Mathematische Annalen (2025) 393:1863–1880
https://doi.org/10.1007/s00208-025-03285-5

Mathematische Annalen

Pairs of commuting integer matrices

Tim Browning[1] · Will Sawin[2] · Victor Y. Wang[3]

Received: 1 October 2024 / Revised: 29 July 2025 / Accepted: 1 September 2025 /
Published online: 10 September 2025

Abstract

We prove upper and lower bounds on the number of pairs of commuting $n \times n$ matrices with integer entries in $[-T, T]$, as $T \to \infty$. Our work uses Fourier analysis and leads to an analysis of exponential sums involving matrices over finite fields. These are bounded by combining a stratification result of Fouvry and Katz with a new result about the flatness of the commutator Lie bracket.

Mathematics Subject Classification 11D45 · 11T23 · 15B33 · 15B36

1 Introduction

Let M_n denote the scheme of $n \times n$ matrices, for an integer $n \geqslant 2$. Let $\mathsf{V}_n \subset \mathsf{M}_n$ denote the closed subscheme $\mathrm{tr} = 0$. Concretely, if R is a ring, then $\mathsf{M}_n(R)$ denotes the set of $n \times n$ matrices A with entries in R, and $\mathsf{V}_n(R)$ denotes the set of such matrices for which $\mathrm{tr}(A) = 0$. The aim of this note is to count integral points of height $\leqslant T$, as $T \to \infty$, on the *commuting variety*

$$\mathfrak{C}_n = \{(X, Y) \in \mathsf{M}_n(\mathbb{C})^2 : XY = YX\},$$

viewed as an affine variety in $\mathsf{M}_n(\mathbb{C})^2 \cong \mathbb{C}^{2n^2}$. Thus we are interested in the behaviour of the counting function

$$N(T) := \#\{(X, Y) \in \mathsf{M}_n(\mathbb{Z})^2 : |X|, |Y| \leqslant T, \ XY = YX\},$$

✉ Tim Browning
tdb@ist.ac.at

Will Sawin
wsawin@math.princeton.edu

Victor Y. Wang
vywang@as.edu.tw

1 ISTA, Am Campus 1, 3400 Klosterneuburg, Austria

2 Princeton University, Fine Hall, 304 Washington Rd, Princeton, NJ 08540, USA

3 Institute of Mathematics, Academia Sinica, Taipei 106319, Taiwan

as $T \to \infty$, where $|X|$ denotes the maximum modulus of any of the entries in a matrix $X \in \mathsf{M}_n(\mathbb{Z})$. It was shown by Motzkin and Taussky in the paragraph after [16, Thm. 6] that $\mathfrak{C}_n$ is irreducible and it follows from work of Basili [1, Thm. 1.2] that $\dim \mathfrak{C}_n = n^2 + n$. Furthermore, $\mathfrak{C}_n$ cannot be linear, since it contains $(X, 0)$ and $(0, Y)$, for any $X, Y \in \mathsf{M}_n(\mathbb{C})$, but it does not include the point (X, Y) if X and Y are non-commuting. In fact, a recursive formula for the degree of $\mathfrak{C}_n$ has been provided by Knutson and Zinn-Justin [13].

Since $\mathfrak{C}_n$ is defined by a system of homogeneous quadratic equations with coefficients in $\mathbb{Q}$, we can view $\mathfrak{C}_n$ as an affine cone over a non-linear irreducible variety in $\mathbb{P}^{2n^2-1}$ of dimension n^2+n-1, which is definable over $\mathbb{Q}$. Appealing to the resolution of the *dimension growth conjecture* by Salberger [19], it therefore follows that

$$N(T) \ll_{n,\epsilon} T^{n^2+n-1+\epsilon} \tag{1.1}$$

for any $\epsilon > 0$, where the implied constant depends at most on n and ε. Moreover, the ϵ can be removed if the variety has degree $\geqslant 5$ by [5], which is likely to be the case for $n \geqslant 3$ by the work in [13]. On the other hand, we have the lower bound

$$N(T) \geqslant \#\{(X, Y) \in \mathsf{M}_n(\mathbb{Z}) \times \mathbb{Z}I_n : |X|, |Y| \leqslant T\} \gg T^{n^2+1} \tag{1.2}$$

coming from the special subvariety of scalar matrices Y, where I_n denotes the $n \times n$ identity matrix. The following is our main result.

Theorem 1.1 *Let $T \geqslant 1$. Then*

$$N(T) \ll_n T^{n^2+2-\frac{2}{n+1}}.$$

Note that Theorem 1.1 improves on (1.1) for $n \geqslant 3$. In fact our exponent is always within 1 of that occurring in the lower bound (1.2), which we expect to be the truth. Our approach uses matrix identities and ideas from harmonic analysis, which will lead us to analyse the fibres of the commutator map

$$[\cdot,\cdot]\colon \mathsf{M}_n^2 \to \mathsf{V}_n, \quad (X, Y) \mapsto XY - YX, \tag{1.3}$$

where we recall that $\mathsf{V}_n \subset \mathsf{M}_n$ denotes the closed subscheme $\mathrm{tr} = 0$. In this way we have been led to prove the following result, the analogue of which was proved by Larsen and Lu [14] in the multiplicative setting.

Theorem 1.2 *The map*

$$[\cdot,\cdot](\mathbb{C})\colon \mathsf{M}_n(\mathbb{C})^2 \to \mathsf{V}_n(\mathbb{C}), \quad (X, Y) \mapsto XY - YX \tag{1.4}$$

is flat over the open set $\mathsf{V}_n(\mathbb{C}) \setminus \{0\}$.

As explained below in Proposition 2.1, Theorem 1.2 is equivalent to the *pointwise* bound

$$\dim \left\{(U, V) \in \mathsf{M}_n(\mathbb{C})^2 : UV - VU = M\right\} \leqslant n^2 + 1 \tag{1.5}$$

on the dimensions of the fibres of the commutator map over matrices $M \in \mathsf{V}_n(\mathbb{C})\setminus\{0\}$.[1] For comparison, earlier work has examined the dimensions of the fibres *on average* over certain small families of matrices in $\mathsf{V}_n(\mathbb{C}) \setminus \{0\}$, such as diagonal matrices (Knutson [12]) or rank-one matrices (Neubauer [17]). Moreover, Young [25] proved a finer conjecture of Knutson [12] concerning the irreducible components of the diagonal commutator scheme. Our new methods may be capable of re-proving, and generalising to other families, some of these earlier results.

2 Background on flatness

Before proceeding, we recall some classical criteria for flatness, which are useful both for interpreting Theorem 1.2 and for proving it. Throughout this paper, a *variety* will be a locally closed subset of a projective space $\mathbb{P}^N$ over a field k (which is often called a *quasi-projective variety*). If k is algebraically closed, then varieties V over k may be completely understood in terms of their k-points $V(k)$ by [10, Chapter II, Proposition 2.6], and we will take this classical viewpoint whenever possible.

Proposition 2.1 *Let $f: E_1 \to E_2$ be a morphism of smooth, irreducible varieties over a field k. Let k' be a field containing k. Then the following are equivalent:*

(1) The morphism f is flat.
(2) The morphism $f \otimes k': E_1 \otimes k' \to E_2 \otimes k'$ is flat.
(3) For every point $y \in E_2(\overline{k})$, the fibre $f^{-1}(y)$ is either empty or of dimension exactly $\dim E_1 - \dim E_2$.
(4) For every point $y \in E_2(\overline{k})$, we have $\dim(f^{-1}(y)) \leqslant \dim E_1 - \dim E_2$.

Proof (1)⇔(2): Flatness is preserved by base change from k to k'. (See [23, Tag 047C] or [23, Tag 02JZ].) The key point is that the inclusion homomorphism $k \to k'$ is flat.

(1)⇒(3): This follows from [15, Theorem 15.1 (or paragraph 1 of § 23)].

(3)⇒(1): This follows from miracle flatness [15, Corollary to Theorem 23.1], which applies since E_1 and E_2 are smooth, irreducible varieties over k.

(3)⇒(4): This is trivial.

(4)⇒(3): Any non-empty fibre $f^{-1}(y)$ has dimension at least $\dim E_1 - \dim E_2$, and thus exactly $\dim E_1 - \dim E_2$ if (4) holds. □

While a direct proof of Theorem 1.2 is possible by induction on n (via algebraic geometry, Grassmannians, symmetry, and dimension counting), we shall find it more convenient to prove instead a characteristic p version. We will prove the following result in § 5 by induction on n for $p \geqslant n$, using symmetry and point counting.

Theorem 2.2 *For all primes $p \gg_n 1$, and for all points $M \in \mathsf{V}_n(\overline{\mathbb{F}}_p) \setminus \{0\}$, we have*

$$\dim \left\{ (U, V) \in \mathsf{M}_n(\overline{\mathbb{F}}_p)^2 : UV - VU = M \right\} \leqslant n^2 + 1.$$

[1] In fact, equality must hold in (1.5), since the map (1.4) is surjective by work of Shoda [20].

We shall prove that this result is equivalent to Theorem 1.2. In fact Theorem 1.2 and this equivalence are not needed for the Diophantine results in this paper, but are included purely for the relevance of Theorem 1.2 to algebraic geometry. The key point is that miracle flatness, such as the equivalence (1) $\Leftrightarrow$ (3) in Proposition 2.1, converts flatness statements like Theorem 1.2 into dimension statements like Theorem 2.2. To deal with the fact that one statement occurs in characteristic zero and one statement occurs in characteristic p, we use standard "spreading out" arguments,[2] to convert between characteristics 0 and p. This consists of setting everything up over $\operatorname{Spec}\mathbb{Z}$ and applying general results that natural properties (in this case, flatness) define open subsets of varieties over $\operatorname{Spec}\mathbb{Z}$. At the heart of this lies the observation that an open subset of $\operatorname{Spec}\mathbb{Z}$ contains the characteristic 0 point $\operatorname{Spec}\mathbb{Q}$ if and only if it contains the characteristic p point $\operatorname{Spec}\mathbb{F}_p$, for all but finitely many primes p.

Proof of the equivalence of Theorems 1.2 and 2.2 We broadly follow the argument of [14, § 5]. Let $\mathsf{V}_n - 0$ be the open subscheme of V_n with zero section removed, over $\mathbb{Z}$. Thus $(\mathsf{V}_n - 0)(k) = \mathsf{V}_n(k) \setminus \{0\}$, for every field k. Let $\mathsf{C}_n \subseteq \mathsf{M}_n^2$ be the commuting scheme $XY = YX$, over $\mathbb{Z}$. Let $f\colon \mathsf{M}_n^2 - \mathsf{C}_n \to \mathsf{V}_n - 0$ be the commutator map in (1.3), restricted to $\mathsf{M}_n^2 - \mathsf{C}_n$. Let $O \subseteq \mathsf{M}_n^2 - \mathsf{C}_n$ be the *flat locus* of f; i.e. the set of points $x \in \mathsf{M}_n^2 - \mathsf{C}_n$ at which the morphism f is flat. Then O is open by [23, Tag 0399]. Theorem 1.2 is the statement that the base change $f \otimes \mathbb{C}\colon (\mathsf{M}_n^2 - \mathsf{C}_n)_{\mathbb{C}} \to (\mathsf{V}_n - 0)_{\mathbb{C}}$ is flat. By the equivalence (1)$\Leftrightarrow$(2) in Proposition 2.1, this is equivalent to the flatness of $f \otimes \mathbb{Q}\colon (\mathsf{M}_n^2 - \mathsf{C}_n)_{\mathbb{Q}} \to (\mathsf{V}_n - 0)_{\mathbb{Q}}$, which is equivalent to O containing the entire characteristic 0 fibre, $(\mathsf{M}_n^2 - \mathsf{C}_n)_{\mathbb{Q}}$. An open set of a scheme of finite type over $\mathbb{Z}$ contains the entire characteristic 0 fibre if and only if it contains all but finitely many characteristic p fibres,[3] so Theorem 1.2 is equivalent to O containing $(\mathsf{M}_n^2 - \mathsf{C}_n)_{\mathbb{F}_p}$ for all but finitely many p. We call this equivalence ($\star$).

If O contains $(\mathsf{M}_n^2 - \mathsf{C}_n)_{\mathbb{F}_p}$, then since the base change of a flat morphism is flat (by part (1) of [23, Tag 047C]), the base change of f along $(\mathsf{V}_n - 0)_{\overline{\mathbb{F}}_p} \to (\mathsf{V}_n - 0)_{\mathbb{Z}}$ is flat. This base change is $f \otimes \overline{\mathbb{F}}_p\colon (\mathsf{M}_n^2 - \mathsf{C}_n)_{\overline{\mathbb{F}}_p} \to (\mathsf{V}_n - 0)_{\overline{\mathbb{F}}_p}$. From this (and the implication (1)$\Rightarrow$(4) in Proposition 2.1) the statement of Theorem 2.2 follows, since

$$\dim \mathsf{M}_n(\overline{\mathbb{F}}_p)^2 - \dim \mathsf{V}_n(\overline{\mathbb{F}}_p) = 2n^2 - (n^2 - 1) = n^2 + 1.$$

Thus Theorem 1.2 implies Theorem 2.2.

For the reverse implication, we appeal to miracle flatness, which for the generality we need is established in [23, Tag 00R4] or [15, Theorem 23.1]. This states that a morphism between regular Noetherian schemes is flat in a neighbourhood of any point where the dimension of the fibre plus the dimension of the target equals the dimension of the source. But then Theorem 2.2 implies that for all $p \gg_n 1$, the set O contains the entire characteristic p fibre of $\mathsf{M}_n^2 - \mathsf{C}_n$, which by the equivalence ($\star$) described above implies Theorem 1.2. $\square$

[2] A general reference for such arguments is [18, Theorem 3.2.1(iv)] but we don't really need this particular form of the theorem.

[3] The analytically minded reader may think of this as a consequence of the Lang–Weil estimate for the complement of the open set in question.

3 Matrix exponential sums

Over any field k, we will often make the identification $\mathsf{M}_n(k) \cong k^{n^2}$. The usual dot product on $X, Y \in \mathsf{M}_n(k)$ is then given by $\mathrm{tr}(X^tY) = \sum_i \sum_j X_{ji}Y_{ji}$. For the purposes of harmonic analysis to come, however, it will be more convenient to work with the pairing $\mathsf{M}_n(k) \times \mathsf{M}_n(k) \to k$ given by $(X, Y) \mapsto \mathrm{tr}(XY) = \sum_i \sum_j X_{ij}Y_{ji}$, which can be viewed as the usual dot product after an invertible linear change of variables $X \mapsto X^t$. (Abstractly, the key point is that this pairing is *non-degenerate*, in the sense that for any non-zero $X \in \mathsf{M}_n(k)$, the linear form $Y \mapsto \mathrm{tr}(XY)$ is non-zero.)

Given $A, B \in \mathsf{M}_n(\mathbb{Z})$, our work will lead us to analyse the complete exponential sum

$$S(A, B; p) := \sum_{\substack{(U,V)\in \mathsf{M}_n(\mathbb{F}_p)^2\\ UV-VU=0}} e_p(\mathrm{tr}(AU + BV)), \tag{3.1}$$

for a sufficiently large prime p. It follows from the point count of Feit–Fine [6] that

$$S(0, 0; p) \ll_n p^{n^2+n}. \tag{3.2}$$

We seek to find conditions on A, B under which we can show that cancellation occurs in $S(A, B; p)$.

We begin by recording the observation that

$$S(A, B; p) = \frac{1}{\#\mathsf{M}_n(\mathbb{F}_p)} \sum_{U,V,Z\in \mathsf{M}_n(\mathbb{F}_p)} e_p(\mathrm{tr}(Z(UV - VU) + AU + BV)). \tag{3.3}$$

The cyclic property of the trace ensures that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ and $\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$, for any $A, B, C \in \mathsf{M}_n(\mathbb{F}_p)$. Hence $\mathrm{tr}(Z(UV - VU) + AU) = \mathrm{tr}(U(VZ - ZV + A))$, so that we can average over U to conclude that

$$S(A, B; p) = \sum_{\substack{V,Z\in \mathsf{M}_n(\mathbb{F}_p)\\ VZ-ZV+A=0}} e_p(\mathrm{tr}(BV)).$$

In particular, it is clear that

$$|S(A, B; p)| \le S(A, 0; p) = \#\left\{(V, Z) \in \mathsf{M}_n(\mathbb{F}_p)^2 : VZ - ZV + A = 0\right\}, \tag{3.4}$$

for any $B \in \mathsf{M}_n(\mathbb{Z})$. We are now ready to prove the following result.

Lemma 3.1 *If $p \nmid A$ or $p \nmid B$, then*

$$S(A, B; p) \ll_n p^{n^2+1}\mathbf{1}_{p|\mathrm{tr}(A)}\mathbf{1}_{p|\mathrm{tr}(B)}.$$

Proof We may assume that p is sufficiently large in terms of n, else the result is trivial. If $p \nmid A$, then it follows from applying Theorem 2.2 and the Lang–Weil bound in (3.4) that

$$S(A, B; p) \ll_n p^{n^2+1} \mathbf{1}_{p|\mathrm{tr}(A)}.$$

Similarly, if $p \nmid B$, then averaging over V in (3.3) gives

$$S(A, B; p) \ll \#\left\{(U, Z) \in \mathsf{M}_n(\mathbb{F}_p)^2 : ZU - UZ + B = 0\right\} \ll_n p^{n^2+1} \mathbf{1}_{p|\mathrm{tr}(B)}.$$

Combining these bounds, we arrive at the statement of the lemma. □

4 Fourier analysis and proof of the main result

In this section we prove Theorem 1.1 via harmonic analysis, by combining work of Fouvry and Katz [7] with our discussion in the previous section on exponential sums over finite fields. We shall assume throughout this section that $n \geqslant 2$.

Let p be an auxiliary prime in the interval $[T, 2T^2]$ whose size will be optimised later. Then

$$N(T) \le \#\{(X, Y) \in \mathsf{M}_n(\mathbb{Z})^2 : |X|, |Y| \leqslant T,\ \ p \mid XY - YX\}.$$

Consider the function

$$w(X) := \prod_{1 \le i, j \le n} \left(\frac{\sin \pi X_{ij}}{\pi X_{ij}}\right)^2,$$

for any $X \in \mathsf{M}_n(\mathbb{R})$. Identifying $\mathsf{M}_n(\mathbb{R})$ with $\mathbb{R}^{n^2}$, this has Fourier transform

$$\widehat{w}(A) = \int_{\mathsf{M}_n(\mathbb{R})} w(X) e(\mathrm{tr}(AX)) \mathrm{d}X = \prod_{1 \le i, j \le n} \max\{1 - |A_{ij}|, 0\}, \tag{4.1}$$

where $\mathrm{d}X = \prod_{i,j} \mathrm{d}X_{ij}$. Moreover it is clear that $(\frac{4}{\pi^2})^{n^2} \le w(\frac{1}{2}X) \le 1$ if $|X| \le 1$. It now follows that

$$\begin{aligned} N(T) &\ll_n \sum_{\substack{(X,Y) \in \mathsf{M}_n(\mathbb{Z})^2 \\ p | XY - YX}} w\Big(\frac{1}{2T} X\Big) w\Big(\frac{1}{2T} Y\Big) \\ &= \sum_{\substack{(U,V) \in \mathsf{M}_n(\mathbb{F}_p)^2 \\ UV - VU = 0}} \sum_{X' \in \mathsf{M}_n(\mathbb{Z})} w\Big(\frac{U + pX'}{2T}\Big) \sum_{Y' \in \mathsf{M}_n(\mathbb{Z})} w\Big(\frac{V + pY'}{2T}\Big). \end{aligned}$$

An application of Poisson summation reveals that

$$\sum_{X'\in \mathsf{M}_n(\mathbb{Z})} w\Big(\frac{U+pX'}{2T}\Big) = \sum_{A\in \mathsf{M}_n(\mathbb{Z})} \int_{\mathsf{M}_n(\mathbb{R})} w\Big(\frac{U+pX'}{2T}\Big) e(\mathrm{tr}(AX'))\mathrm{d}X'$$
$$= \left(\frac{2T}{p}\right)^{n^2} \sum_{A\in \mathsf{M}_n(\mathbb{Z})} e_p(-\operatorname{tr}(AU))\widehat{w}\left(\frac{2T}{p}A\right),$$

on making an obvious change of variables. We have a similar identity for the sum over Y', all of which leads to the expression

$$N(T) \ll_n \left(\frac{T}{p}\right)^{2n^2} \sum_{(A,B)\in \mathsf{M}_n(\mathbb{Z})^2} \widehat{w}\left(\frac{2T}{p}A\right) \widehat{w}\left(\frac{2T}{p}B\right) S(A, B; p),$$

in the notation of (3.1).

Appealing to (4.1), we finally deduce that

$$N(T) \ll_n \left(\frac{T}{p}\right)^{2n^2} \sum_{\substack{(A,B)\in \mathsf{M}_n(\mathbb{Z})^2\\ |A|,|B|\leqslant \frac{1}{2}p/T}} |S(A, B; p)|.$$

If $p \mid (A, B)$ then only $A = B = 0$ contribute to the sum. Thus it follows from (3.2) that

$$N(T) \ll_n \left(\frac{T}{p}\right)^{2n^2} \left(p^{n^2+n} + U(p, T)\right), \tag{4.2}$$

where

$$U(p, T) = \sum_{\substack{(A,B)\in \mathsf{M}_n(\mathbb{Z})^2\\ |A|,|B|\leqslant \frac{1}{2}p/T\\ p\nmid(A,B)}} |S(A, B; p)|.$$

We now appeal to a *stratification* result by Fouvry and Katz [7, Thm. 1.1] to estimate $U(p, T)$. This produces subschemes $V_{2n^2} \subset V_{2n^2-1} \subset \cdots \subset V_2 \subset V_1 \subset \mathbb{A}_{\mathbb{Z}}^{2n^2}$, with $\dim(V_j \otimes \mathbb{C}) \le 2n^2 - j$ for $1 \le j \le 2n^2$, such that

$$S(A, B; p) \ll_n p^{\frac{n^2+n}{2}+\frac{j-1}{2}},$$

whenever the reduction modulo p of (A, B) corresponds to vector in $\mathbb{A}^{2n^2}(\mathbb{F}_p) \setminus V_j(\mathbb{F}_p)$, under the isomorphism $\mathsf{M}_n^2 \cong \mathbb{A}^{2n^2}$. It is convenient to put $V_0 = \mathbb{A}_{\mathbb{Z}}^{2n^2}$ and $V_{2n^2+1} = \emptyset$. Combining this with Lemma 3.1, we therefore deduce that

$$U(p, T) \ll_n \sum_{j=1}^{2n^2+1} \sum_{(A,B)\in R_j} \min\left\{p^{\frac{n^2+n}{2}+\frac{j-1}{2}}, \ p^{n^2+1}\right\},$$

where R_j denotes the set of $(A, B) \in \mathsf{M}_n(\mathbb{Z})^2$ with $|A|, |B| \leqslant \frac{1}{2}p/T$ for which the reduction of (A, B) mod p is non-zero and belongs to the set $V_{j-1}(\mathbb{F}_p) \setminus V_j(\mathbb{F}_p)$. We have $R_j \subseteq R'_j$, where R'_j is the set of $(A, B) \in \mathsf{M}_n(\mathbb{Z})^2$ with $|A|, |B| \leqslant \frac{1}{2}p/T$ for which the reduction of (A, B) mod p belongs to $V_{j-1}(\mathbb{F}_p)$. We now recall the statement of [4, Lemma 4] with $r = 1$. Given a subscheme $W \subset \mathbb{A}^N_{\mathbb{Z}}$ and an affine variety $V \subset \mathbb{A}^N_{\mathbb{F}_p}$, with $\dim(W \otimes \mathbb{C}) \le \ell$ and $\dim V \le k$, this states that

$$\#\left\{\mathbf{t} \in W \cap \mathbb{Z}^N : |\mathbf{t}| \le H,\ \mathbf{t} \bmod p \in V(\mathbb{F}_p)\right\} \ll_{D,N} H^{\ell} p^{k-\ell} + H^k,$$

for any $H \geqslant 1$, where $D = \max\{\deg(W \otimes \mathbb{C}), \deg V\}$. Note that $H^{\ell} p^{k-\ell} \le H^k$ if $k \le \ell$ and $H \le p$. We apply this with $W = \mathsf{M}_n^2$, $V = V_{j-1}$, $\ell = 2n^2$, $k = 2n^2 - j + 1$, and $H = p/T$. Noting that $p \geqslant T$, it now follows that

$$\#R_j \le \#R'_j \ll_n \left(\frac{p}{T}\right)^{2n^2-j+1}.$$

But then we may conclude that

$$U(p, T) \ll_n \left(\frac{p}{T}\right)^{2n^2+1} \left(\sum_{j=1}^{n^2-n+3} \left(\frac{T}{\sqrt{p}}\right)^j p^{\frac{n^2+n-1}{2}} + \sum_{j=n^2-n+4}^{2n^2+1} \left(\frac{T}{p}\right)^j p^{n^2+1} \right).$$

Since $p \in [T, 2T^2]$ we see that the first term is maximised at the largest value of j, while the second term is maximised when j is least. This leads to the expression

$$\begin{aligned} U(p, T) &\ll_n \left(\frac{p}{T}\right)^{2n^2+1} \left(T^{n^2-n+3} p^{n-2} + T^{n^2-n+4} p^{n-3}\right) \\ &\ll_n \left(\frac{p}{T}\right)^{2n^2} \cdot T^{n^2-n+2} p^{n-1}, \end{aligned}$$

since $p \geqslant T$.

Returning to (4.2), we have now established that

$$N(T) \ll_n \frac{T^{2n^2}}{p^{n^2-n}} + T^{n^2-n+2} p^{n-1}.$$

This is optimised at $p^{n^2-1} \asymp T^{n^2+n-2}$, which leads to the final bound for $N(T)$ recorded in Theorem 1.1.

The idea of choosing an auxiliary modulus p and using harmonic analysis has been applied to other varieties in the past, such as by Fujiwara [8, 9], and by Shparlinski and Skorobogatov [21, 22]. Moreover, using finer geometric information, Heath-Brown [11] has given stronger results in many cases by working with a composite modulus pq. Once combined with Fouvry–Katz stratification [7], the simplest version of Fujiwara's method gives the following general result.

Theorem 4.1 *Let $V \subset \mathbb{A}^N_{\mathbb{Z}}$ be a subscheme with* $\dim(V \otimes \mathbb{C}) = D$*. Assume that there is a positive density set $\mathcal{P}$ of primes such that for all $p \in \mathcal{P}$, and for all non-zero vectors $c \in \mathbb{F}_p^N$, we have*

$$\sum_{x \in V(\mathbb{F}_p)} e_p(c_1 x_1 + \cdots + c_N x_N) \ll_V p^{D-L}, \tag{4.3}$$

for some L such that $2L \in \mathbb{Z}$. Then for $T \geqslant 1$, we have

$$\#\{x \in V(\mathbb{Z}) \cap [-T, T]^N\} \ll_V T^{D-L+\frac{L^2}{N-D+L}}. \tag{4.4}$$

This bound improves on the *dimension growth bound* $O_{\varepsilon,V}(T^{D-1+\epsilon})$ of Salberger [19] and Vermeulen [24] (when it applies), provided only that $N-D \geqslant 4$ and $L \geqslant \frac{3}{2}$ in (4.3). Note that (4.4) recovers the bound in Theorem 1.1 when $N = 2n^2$, $D = n^2 + n$ and $L = n - 1$. Moreover, if V lies in an algebraic family[4] F in the sense of Bonolis–Kowalski–Woo [2], then one could hope for (4.3) and (4.4) to be made fairly uniform over V, perhaps with a polylogarithmic dependence on the coefficients as in work of Bonolis–Pierce–Woo [3].

5 Flatness of the commutator map

In this section we shall prove Theorem 2.2. We shall often find it notationally convenient to adopt the expectation notation $\mathbb{E}_{g \in S} = \frac{1}{\#S}\sum_{g \in S}$, for any non-empty subset $S \subseteq \mathrm{GL}_n(\mathbb{F}_q)$. Let p be any prime. (Eventually, we will assume $p \geqslant n$.) Fix a non-zero $M \in \mathsf{V}_n(\overline{\mathbb{F}}_p)$ and note that

$$\dim\{(U, V) \in \mathsf{M}_n(\overline{\mathbb{F}}_p)^2 : UV - VU = M\} = -1 + \dim \mathcal{X},$$

where

$$\mathcal{X} := \{(U, V, \lambda) \in \mathsf{M}_n(\overline{\mathbb{F}}_p)^2 \times \mathbb{G}_m(\overline{\mathbb{F}}_p) : UV - VU = \lambda M\}.$$

We shall bound $\dim \mathcal{X}$ via point counting on $\mathcal{X}$. We take $q = p^m \to \infty$, with $m \in \mathbb{N}$ sufficiently large so that $M \in \mathsf{V}_n(\mathbb{F}_q)$ and the variety $\mathcal{X}$ has an irreducible component that is invariant under $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_q)$. It will suffice to estimate the quantity

$$\Sigma(M) = \frac{1}{q} \sum_{U,V \in \mathsf{M}_n(\mathbb{F}_q)} \sum_{\lambda \in \mathbb{G}_m(\mathbb{F}_q)} \mathbf{1}_{UV - VU = \lambda M},$$

since then

$$\Sigma(M) \geqslant (1 + O_n(q^{-1/2}))\, q^{\dim\{(U,V) \in \mathsf{M}_n(\overline{\mathbb{F}}_p)^2 : UV - VU = M\}} \tag{5.1}$$

[4] That is, V is the fibre over a $\mathbb{Z}$-point of a dominant morphism $\Delta \colon W \to M$, where W is a subscheme of an affine space and M is a closed subscheme of an affine space. For example, if V is a complete intersection, then one could simply ask for the sum of the degrees and number of variables of the defining equations to be bounded.

by the Lang–Weil estimate for the variety $\mathcal{X}$.

Let $\psi(\cdot):=e_p(\mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p}(\cdot))$ on $\mathbb{F}_q$. We proceed by noting that

$$\begin{aligned}\Sigma(M) &= \frac{1}{q\cdot\#\mathsf{M}_n(\mathbb{F}_q)}\sum_{U,V,Z\in\mathsf{M}_n(\mathbb{F}_q)}\sum_{\lambda\in\mathbb{G}_m(\mathbb{F}_q)}\psi(\mathrm{tr}(Z(UV-VU-\lambda M)))\\ &= \frac{1}{\#\mathsf{M}_n(\mathbb{F}_q)}\sum_{U,V,Z\in\mathsf{M}_n(\mathbb{F}_q)}\psi(\mathrm{tr}(Z(UV-VU)))\,(\mathbf{1}_{\mathrm{tr}(ZM)=0}-q^{-1})\\ &= \sum_{\substack{V,Z\in\mathsf{M}_n(\mathbb{F}_q)\\ VZ-ZV=0}}(\mathbf{1}_{\mathrm{tr}(ZM)=0}-q^{-1}),\end{aligned}$$

where in the last step we average over U using the cyclic property of the trace function. We now exploit the linearity of the equations $VZ-ZV=0$ and $\mathrm{tr}(ZM)=0$ in the coordinates of Z. Since the hyperplane $\mathrm{tr}(ZM)=0$ has codimension either 0 or 1 in the linear space $C(V):=\{Z\in\mathsf{M}_n(\overline{\mathbb{F}}_q):VZ-ZV=0\}$, we find that

$$\begin{aligned}\Sigma(M) &= \sum_{V\in\mathsf{M}_n(\mathbb{F}_q)} q^{\dim C(V)}(q^{-1}\mathbf{1}_{C(V)\not\subseteq M^\perp}+\mathbf{1}_{C(V)\subseteq M^\perp}-q^{-1})\\ &= \sum_{V\in\mathsf{M}_n(\mathbb{F}_q)} q^{\dim C(V)}(1-q^{-1})\mathbf{1}_{C(V)\subseteq M^\perp},\end{aligned}$$

where $M^\perp:=\{A\in\mathsf{M}_n(\overline{\mathbb{F}}_q):\mathrm{tr}(AM)=0\}$. Then, since $\dim C(V)$ depends only on the conjugacy class of V, we find on averaging over conjugates of V that

$$\Sigma(M)=\sum_{V\in\mathsf{M}_n(\mathbb{F}_q)} q^{\dim C(V)}(1-q^{-1})L(V,M), \tag{5.2}$$

where

$$L(V,M):=\mathbb{E}_{g\in\mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{C(gVg^{-1})\subseteq M^\perp})\leqslant 1. \tag{5.3}$$

The orbit of V under the conjugation action of $\mathrm{GL}_n(\mathbb{F}_q)$ has cardinality

$$\mathscr{O}(V)=\frac{\#\,\mathrm{GL}_n(\mathbb{F}_q)}{\#(\mathrm{GL}_n(\mathbb{F}_q)\cap C(V))},$$

by the orbit-stabiliser theorem. Observe that $\#\,\mathrm{GL}_n(\mathbb{F}_q)\le\#M_n(\mathbb{F}_q)=q^{n^2}$. Moreover, the variety $\mathrm{GL}_n(\overline{\mathbb{F}}_q)\cap C(V)$, which contains the identity matrix I_n, is a non-empty open subset of the linear space $C(V)$. Thus $\#(\mathrm{GL}_n(\mathbb{F}_q)\cap C(V))\gg_n q^{\dim C(V)}$ by the Lang–Weil estimate, whence

$$\mathscr{O}(V)\ll_n q^{n^2-\dim C(V)}.$$

Therefore, if $K_n(\mathbb{F}_q) \subseteq \mathsf{M}_n(\mathbb{F}_q)$ denotes a complete set of representatives for the conjugation action, then breaking (5.2) into orbits gives

$$\begin{aligned}\Sigma(M) &= \sum_{V \in K_n(\mathbb{F}_q)} \mathscr{O}(V) q^{\dim C(V)} (1-q^{-1}) L(V, M) \\ &\ll_n q^{n^2} \sum_{V \in K_n(\mathbb{F}_q)} L(V, M).\end{aligned} \tag{5.4}$$

Lemma 5.4 will non-trivially bound $L(V, M)$. The proof inducts on n and uses an auxiliary averaging result, Lemma 5.3. First, we define some useful projection maps. Given integers $1 \leqslant k \leqslant n-1$ and a matrix $M \in \mathsf{M}_n(\mathbb{F}_q)$, define block matrices $p_{ij}(M)$ so that

$$M = \begin{bmatrix} p_{11}(M) & p_{12}(M) \\ p_{21}(M) & p_{22}(M) \end{bmatrix},$$

where $p_{ij}(M)$ has $k\mathbf{1}_{i=1} + (n-k)\mathbf{1}_{i=2}$ rows and $k\mathbf{1}_{j=1} + (n-k)\mathbf{1}_{j=2}$ columns.

Lemma 5.3 concerns some averages over $\mathrm{GL}_n(\mathbb{F}_q)$. First we study averages over a certain subgroup. Let

$$E := \{N \in \mathsf{M}_n(\mathbb{F}_q) : p_{11}(N) = 0,\ \ p_{21}(N) = 0,\ \ p_{22}(N) = 0\} \cong \mathbb{F}_q^{k \times (n-k)}.$$

Then E is a vector space such that $N_1 N_2 = 0$ for all $N_1, N_2 \in E$. In particular,

$$1 + E := \{1 + N : N \in E\}$$

is an abelian subgroup of $\mathrm{GL}_n(\mathbb{F}_q)$, where $(1+N)^{-1} = 1 - N$ for all $N \in E$.

Lemma 5.1 *Let $M \in \mathsf{M}_n(\mathbb{F}_q)$, with $M \neq 0$. Let $1 \leqslant k \leqslant n-1$. Then the following hold:*

(1) We have

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{p_{11}(h^{-1}Mh)=0}) \leqslant q^{-k} + \min_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{21})(h^{-1}Mh)=0}).$$

(2) We have

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{22})(h^{-1}Mh)=0}) \leqslant q^{-(n-1)} + \min_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{21},p_{22})(h^{-1}Mh)=0}).$$

(3) We have

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{12},p_{22})(h^{-1}Mh)=0}) \leqslant q^{-(n-1)}.$$

Proof First we record a general calculation: for $N \in E$ we have

$$\begin{aligned}(1-N)&M(1+N)\\ &= (1-N)(M+MN)\\ &= M+MN-NM-NMN\\ &= M+\left[\begin{smallmatrix}0 & p_{11}(M)p_{12}(N)\\ 0 & p_{21}(M)p_{12}(N)\end{smallmatrix}\right]-\left[\begin{smallmatrix}p_{12}(N)p_{21}(M) & p_{12}(N)p_{22}(M)\\ 0 & 0\end{smallmatrix}\right]-\left[\begin{smallmatrix}0 & p_{12}(N)p_{21}(M)p_{12}(N)\\ 0 & 0\end{smallmatrix}\right].\end{aligned} \tag{5.5}$$

(1): Both sides of (1) are invariant under conjugation of M by any element of $1+E$. Moreover, if $p_{11}(h^{-1}Mh)$ is never zero, then (1) is trivial. Thus, after conjugating M if necessary, we may assume that $p_{11}(M)=0$. Then, writing $h=1+N$, we get

$$p_{11}(h^{-1}Mh) = -p_{12}(N)p_{21}(M), \tag{5.6}$$

$$p_{21}(h^{-1}Mh) = p_{21}(M), \tag{5.7}$$

by (5.5). In particular, by (5.6),

$$\mathbb{E}_{h\in 1+E}(\mathbf{1}_{p_{11}(h^{-1}Mh)=0}) = \mathbb{E}_{N\in E}(\mathbf{1}_{p_{12}(N)p_{21}(M)=0}).$$

Given a matrix $N \in E$, the condition $p_{12}(N)p_{21}(M) = 0$ holds if and only if each *row* of $p_{12}(N)$ lies in the left kernel $\mathscr{L} \subseteq \mathbb{F}_q^{n-k}$ of the $(n-k)\times k$ matrix $p_{21}(M)$. If $p_{21}(M) \neq 0$, then $\mathscr{L}$ is orthogonal to a non-zero column of $p_{21}(M)$. Hence, one of the coordinates of $\mathscr{L}$ is a linear function of the others. Thus, one of the columns of $p_{12}(N)$ is uniquely determined by others. Since each column of $p_{12}(N)$ has k entries, it follows that

$$\mathbb{E}_{h\in 1+E}(\mathbf{1}_{p_{11}(h^{-1}Mh)=0}) \leqslant q^{-k}.$$

On the other hand, if $p_{21}(M)=0$, then

$$\mathbb{E}_{h\in 1+E}(\mathbf{1}_{p_{11}(h^{-1}Mh)=0}) = 1$$

and $(p_{11}, p_{21})(h^{-1}Mh) = 0$ for all $h \in 1+E$, by (5.6) and (5.7). Now (1) follows.

(2): As in the proof of (1), we may assume that $(p_{11}, p_{22})(M) = 0$. Then

$$p_{22}(h^{-1}Mh) = p_{21}(M)p_{12}(N) \tag{5.8}$$

for $h = 1+N$, by (5.5). The map

$$\Phi\colon \mathbb{F}_q^{k\times(n-k)} \to \mathsf{M}_k(\mathbb{F}_q)\times \mathsf{M}_{n-k}(\mathbb{F}_q), \quad A \mapsto (Ap_{21}(M), p_{21}(M)A) \tag{5.9}$$

is linear in the entries of A. By (5.6) and (5.8), we have

$$\mathbb{E}_{h\in 1+E}(\mathbf{1}_{(p_{11},p_{22})(h^{-1}Mh)=0}) = \mathbb{E}_{N\in E}(\mathbf{1}_{p_{12}(N)\in\ker\Phi}),$$

If $A \in \ker\Phi$, then the *rows* of A lie in the left kernel of $p_{21}(M)$, and the *columns* of A lie in the right kernel of $p_{21}(M)$. If $p_{21}(M) \neq 0$ and $A \in \ker\Phi$, then it follows that there exist a column C and row R of A such that the entries of A are uniquely determined by the entries in $A \setminus (C \cup R)$. Since $\#(C \cup R) = k + (n-k) - 1 = n-1$, it follows that

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{22})(h^{-1}Mh)=0}) \leqslant q^{-(n-1)} \tag{5.10}$$

if $p_{21}(M) \neq 0$. On the other hand, if $p_{21}(M) = 0$, then

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{22})(h^{-1}Mh)=0}) = 1$$

and $(p_{11}, p_{21}, p_{22})(h^{-1}Mh) = 0$ for all $h \in 1 + E$, by (5.6), (5.7), and (5.8). Hence (2) follows.

(3): As in the proof of (1), we may assume that $(p_{11}, p_{12}, p_{22})(M) = 0$. Then

$$p_{12}(h^{-1}Mh) = -p_{12}(N)p_{21}(M)p_{12}(N) \tag{5.11}$$

for $h = 1 + N$, by (5.5). By (5.6), (5.11), and (5.8), we have

$$\mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{12},p_{22})(h^{-1}Mh)=0}) = \mathbb{E}_{N \in E}(\mathbf{1}_{p_{12}(N) \in \ker \Phi}),$$

where Φ is defined as in (5.9). However, since $M \neq 0$, we must have $p_{21}(M) \neq 0$. Therefore, (3) follows from (5.10). □

Henceforth, assume $p \geqslant n$, so that the following simple combinatorial lemma holds.

Lemma 5.2 *Let $1 \leqslant k \leqslant n-1$. Suppose $x \in \mathbb{F}_q^n$ is non-zero. Then $\sum_{i \in I} x_i \neq 0$ for some k-element subset $I \subseteq \{1, 2, \ldots, n\}$.*

Proof If $x_{n-1} \neq x_n$, say, then either $x_1 + \cdots + x_{k-1} + x_{n-1}$ or $x_1 + \cdots + x_{k-1} + x_n$ must be non-zero. By symmetry, it remains to consider the case $x_1 = \cdots = x_n$. Then $x_1 \neq 0$, since $x \neq 0$. Thus $x_1 + \cdots + x_k = kx_1 \neq 0$, because $p \geqslant n > k$. □

Lemma 5.3 *Let $M \in \mathrm{M}_n(\mathbb{F}_q)$, with $M \neq 0$. Let $1 \leqslant k \leqslant n-1$. Then the following hold:*

(1) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21})(g^{-1}Mg)=0}) \ll_n q^{-k}$.
(2) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_{11}(g^{-1}Mg)=0}) \ll_n q^{-k}$.
(3) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{12},p_{22})(g^{-1}Mg)=0}) \ll_n q^{-(n-1)}$.
(4) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21},p_{22})(g^{-1}Mg)=0}) \ll_n q^{-(n-1)}$.
(5) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{22})(g^{-1}Mg)=0}) \ll_n q^{-(n-1)}$.
(6) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_{22}(g^{-1}Mg)=0}) \ll_n q^{-(n-k)}$.
(7) $\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(p_{11}(g^{-1}Mg))=0}) \ll_n q^{-1}$.

Proof Although (2) is stronger than (1), and (5) is stronger than both (3) and (4), it turns out that proving (1) first will help in proving (2), and likewise for (3) and (4).

(1): Let $e_1, \ldots, e_n \in \mathbb{F}_q^n$ be the usual coordinate basis vectors. The condition $(p_{11}, p_{21})(M) = 0$ holds if and only if $Me_j = 0$ for all $1 \leqslant j \leqslant k$. Thus

$$\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21})(g^{-1}Mg)=0}) = \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{g^{-1}Mge_1=\cdots=g^{-1}Mge_k=0}).$$

Note that $(ge_1, \dots, ge_k)$ is equally likely to be any list of k linearly independent vectors. Since the *right* kernel of M has dimension at most $n-1$, and the number of lists of k linearly independent vectors in $\mathbb{F}_q^d$ is $\mathscr{V}(d) = \prod_{1 \leqslant i \leqslant k}(q^d - q^{i-1}) \leqslant (q^d)^k$, it follows that

$$\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21})(g^{-1}Mg)=0}) \leqslant \frac{\mathscr{V}(n-1)}{\mathscr{V}(n)} \leqslant \prod_{1 \leqslant i \leqslant k} \frac{q^{n-1}}{q^n - q^{i-1}} \ll_n q^{-k}.$$

(2): We have

$$\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_{11}(g^{-1}Mg)=0}) = \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} \mathbb{E}_{h \in 1+E}(\mathbf{1}_{p_{11}((gh)^{-1}Mgh)=0}). \tag{5.12}$$

Applying Lemma 5.1(1) with $g^{-1}Mg$ in place of M, we get

$$\begin{aligned}\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_{11}(g^{-1}Mg)=0}) &\leqslant q^{-k} + \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21})(g^{-1}Mg)=0}) \\ &\ll_n q^{-k},\end{aligned}$$

by (1).

(3): Mimicking (5.12), we have

$$\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{12},p_{22})(g^{-1}Mg)=0}) = \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} \mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{12},p_{22})((gh)^{-1}Mgh)=0}).$$

Applying Lemma 5.1(3) with $g^{-1}Mg$ in place of M, we get (3).

(4): Immediate from (3) with (M^t, g^t) in place of (M, g).

(5): Mimicking (5.12), we have

$$\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{22})(g^{-1}Mg)=0}) = \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} \mathbb{E}_{h \in 1+E}(\mathbf{1}_{(p_{11},p_{22})((gh)^{-1}Mgh)=0}).$$

Applying Lemma 5.1(2) with $g^{-1}Mg$ in place of M, we get

$$\begin{aligned}\mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{22})(g^{-1}Mg)=0}) &\leqslant q^{-(n-1)} + \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{(p_{11},p_{21},p_{22})(g^{-1}Mg)=0}) \\ &\ll_n q^{-(n-1)},\end{aligned}$$

by (4).

(6): This follows from (2) after replacing k with $n-k$, and conjugating M by a suitable permutation matrix.

(7): Assume $q \gg_n 1$. By the $k=1$ case of (2), we may conjugate M to assume that its top left entry is non-zero. Then by Lemma 5.2, there exist k diagonal entries of M whose total is non-zero. By a further conjugation, we may assume $\mathrm{tr}(p_{11}(M)) \neq 0$. Therefore, the subvariety $\mathrm{tr}(p_{11}(g^{-1}Mg)) = 0$ of $\mathrm{GL}_n(\overline{\mathbb{F}}_q)$ is either empty or of codimension one, since $\mathrm{GL}_n(\overline{\mathbb{F}}_q)$ is irreducible. The Lang–Weil bound now implies (7). □

We are finally ready to bound the quantity $L(V, M)$ from (5.3).

Lemma 5.4 *Let $V, M \in \mathsf{M}_n(\mathbb{F}_q)$. Let $f_V(t) \in \mathbb{F}_q[t]$ be the* radical *of the characteristic polynomial of V. Then*

$$L(V, M) \ll_n q^{1-\deg f_V} \mathbf{1}_{\mathrm{tr}(M)=0} + \mathbf{1}_{M=0}. \tag{5.13}$$

Proof We use strong induction on $n \geqslant 1$. Since $1 \in C(gVg^{-1})$ for all g, we may assume $\mathrm{tr}(M) = 0$, or else $L(V, M) = 0$. Moreover, if $\deg f_V = 1$ or $M = 0$, then the trivial bound $L(V, M) \leqslant 1$ suffices. Therefore, we may assume from now on that $\mathrm{tr}(M) = 0$, $M \neq 0$, and $\deg f_V \geqslant 2$.

By definition, a *constructible set* is a finite union of locally closed sets. By the inclusion–exclusion principle, the usual Lang–Weil estimate (for varieties) extends to constructible sets. By the Lang–Weil estimate applied to the constructible set

$$\begin{aligned}&\{g \in \mathrm{GL}_n(\overline{\mathbb{F}}_q) : C(gVg^{-1}) \subseteq M^\perp\}\\ &\quad = \bigcup_{0 \leqslant d \leqslant n^2} \{g \in \mathrm{GL}_n(\overline{\mathbb{F}}_q) : \dim(C(gVg^{-1})) = \dim(C(gVg^{-1}) \cap M^\perp) = d\},\end{aligned}$$

it suffices to prove the desired inequality (5.13) under the assumption that f_V splits completely in $\mathbb{F}_q$. By conjugation, we may assume that V is in rational canonical form, which coincides with Jordan normal form because f_V splits completely. Since $\deg f_V \geqslant 2$, the matrix V has at least 2 distinct eigenvalues. After permuting Jordan blocks if necessary, we may assume that

$$V = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix},$$

where $V_1 \in \mathsf{M}_k(\mathbb{F}_q)$ and $V_2 \in \mathsf{M}_{n-k}(\mathbb{F}_q)$, with $1 \leqslant k \leqslant n-1$, such that V_1 and V_2 share no eigenvalues. Then, in particular,

$$C(V) = \begin{bmatrix} C(V_1) & 0 \\ 0 & C(V_2) \end{bmatrix}. \tag{5.14}$$

Let $p_1 := p_{11}$ and $p_2 := p_{22}$. Let

$$H := \begin{bmatrix} \mathrm{GL}_k(\mathbb{F}_q) & 0 \\ 0 & \mathrm{GL}_{n-k}(\mathbb{F}_q) \end{bmatrix} \subseteq \mathrm{GL}_n(\mathbb{F}_q). \tag{5.15}$$

Since $C(ghVh^{-1}g^{-1}) = gC(hVh^{-1})g^{-1}$, and $g^{-1}M^\perp g = (g^{-1}Mg)^\perp$ by the conjugation-invariance of $\mathrm{tr}(\cdot)$, we have

$$\begin{aligned} L(V, M) &= \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} \mathbb{E}_{h \in H} (\mathbf{1}_{C(ghV(gh)^{-1}) \subseteq M^\perp}) \\ &= \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} \mathbb{E}_{h \in H} (\mathbf{1}_{C(hVh^{-1}) \subseteq (g^{-1}Mg)^\perp}) \\ &= \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} L(V_1, p_1(g^{-1}Mg)) L(V_2, p_2(g^{-1}Mg)), \end{aligned} \tag{5.16}$$

where the last step uses (5.14), (5.15), and the block matrix identity

$$\mathrm{tr}(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} M) = \mathrm{tr}(Ap_1(M)) + \mathrm{tr}(Bp_2(M))$$

for $(A, B) \in \mathsf{M}_k(\mathbb{F}_q) \times \mathsf{M}_{n-k}(\mathbb{F}_q)$. By the inductive hypothesis applied to the two factors of L in (5.16), we get

$$L(V, M) \ll_n \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)} \prod_{1\leqslant i\leqslant 2} (q^{1-\deg f_{V_i}} \mathbf{1}_{\mathrm{tr}(p_i(g^{-1}Mg))=0} + \mathbf{1}_{p_i(g^{-1}Mg)=0})$$
$$\leqslant q^{2-\deg f_V}\mathcal{P}_0 + q^{1-\deg f_{V_1}}\mathcal{P}_1 + q^{1-\deg f_{V_2}}\mathcal{P}_2 + \mathcal{P}_3,$$

where

$$\mathcal{P}_0 := \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(p_1(g^{-1}Mg))=0}\mathbf{1}_{\mathrm{tr}(p_2(g^{-1}Mg))=0}),$$
$$\mathcal{P}_1 := \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(p_1(g^{-1}Mg))=0}\mathbf{1}_{p_2(g^{-1}Mg)=0}),$$
$$\mathcal{P}_2 := \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(p_2(g^{-1}Mg))=0}\mathbf{1}_{p_1(g^{-1}Mg)=0}),$$
$$\mathcal{P}_3 := \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_1(g^{-1}Mg)=0}\mathbf{1}_{p_2(g^{-1}Mg)=0}).$$

We now wish to bound the probabilities $\mathcal{P}_i$. Since

$$0 = \mathrm{tr}(M) = \mathrm{tr}(g^{-1}Mg) = \mathrm{tr}(p_1(g^{-1}Mg)) + \mathrm{tr}(p_2(g^{-1}Mg)),$$

the probabilities $\mathcal{P}_i$ for $0 \leqslant i \leqslant 2$ simplify as follows:

$$\mathcal{P}_0 = \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(p_1(g^{-1}Mg))=0}),$$
$$\mathcal{P}_1 = \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_2(g^{-1}Mg)=0}),$$
$$\mathcal{P}_2 = \mathbb{E}_{g\in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{p_1(g^{-1}Mg)=0}).$$

By parts (7), (6), (2), and (5) of Lemma 5.3, respectively, we have $\mathcal{P}_0 \ll_n q^{-1}$, $\mathcal{P}_1 \ll_n q^{-(n-k)} \leqslant q^{-\deg f_{V_2}}$, $\mathcal{P}_2 \ll_n q^{-k} \leqslant q^{-\deg f_{V_1}}$, and $\mathcal{P}_3 \ll_n q^{1-n} \leqslant q^{1-\deg f_V}$. Thus

$$L(V, M) \ll_n q^{1-\deg f_V} + q^{1-\deg f_{V_1}} q^{-\deg f_{V_2}} + q^{1-\deg f_{V_2}} q^{-\deg f_{V_1}} + q^{1-\deg f_V}$$
$$= 4q^{1-\deg f_V},$$

since $f_{V_1} f_{V_2} = f_V$. □

Plugging Lemma 5.4 into (5.4), we get

$$\Sigma(M) \ll_n q^{n^2} \sum_{V\in K_n(\mathbb{F}_q)} (q^{1-\deg f_V} \mathbf{1}_{\mathrm{tr}(M)=0} + \mathbf{1}_{M=0}).$$

But $K_n(\mathbb{F}_q)$ is in bijection with the set of rational canonical forms on $\mathsf{M}_n(\mathbb{F}_q)$, where we recall that the *rational canonical form* of $M \in \mathsf{M}_n(\mathbb{F}_q)$ consists of a partition $\lambda_\phi = (\lambda_{\phi,1} \geqslant \lambda_{\phi,2} \geqslant \cdots \geqslant 0)$ for each monic irreducible polynomial $\phi \in \mathbb{F}_q[t]$, such that the action of M on $\mathbb{F}_q^n$ is isomorphic to multiplication by t on the vector space $\bigoplus_\phi \bigoplus_i (\mathbb{F}_q[t]/\phi^{\lambda_{\phi,i}}\mathbb{F}_q[t])$, where $\sum_\phi \deg(\phi)|\lambda_\phi| = n$. Thus, by the prime number theorem in $\mathbb{F}_q[t]$ and the fact that n has only finitely many partitions, we deduce that

$$\#\{V \in K_n(\mathbb{F}_q) : \deg f_V = d\} \ll_n q^d$$

for $1 \leqslant d \leqslant n$. Summing over $1 \leqslant d \leqslant n$, we get

$$\Sigma(M) \ll_n q^{n^2+1}\mathbf{1}_{\operatorname{tr}(M)=0} + q^{n^2+n}\mathbf{1}_{M=0}.$$

Finally, we conclude from (5.1) that Theorem 2.2 holds.

Acknowledgements The authors are very grateful to Alina Ostafe, Matthew Satriano and Igor Shparlinski for drawing their attention to this problem and for useful comments, and to Michael Larsen and Peter Sarnak for their helpful correspondence. We also thank the referee for their valuable input. While working on this paper the first author was supported by a FWF grant (DOI 10.55776/P36278), the second author by a Sloan Research Fellowship, and the third author by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 101034413.

Funding Open access funding provided by Institute of Science and Technology (IST Austria).

References

1. Basili, R.: On the irreducibility of varieties of commuting matrices. Pure Appl. Algebra **149**, 107–120 (2000)
2. Bonolis, D., Kowalski, E., Woo, K.: Stratification theorems for exponential sums in families. *Preprint* (2025). arXiv:2506.18299
3. Bonolis, D., Pierce, L.B., Woo, K.: Counting integral points in thin sets of type II: singularities, sieves, and stratification. *Preprint*, (2025). arXiv:2505.11226
4. Browning, T.D., Heath-Brown, D.R.: Rational points on quartic hypersurfaces. J. Reine Angew. Math. **629**, 37–88 (2009)
5. Castryck, W., Cluckers, R., Dittmann, P., Nguyen, K.H.: The dimension growth conjecture, polynomial in the degree and without logarithmic factors. Algebra Number Theory **14**, 2261–2294 (2020)
6. Feit, W., Fine, N.J.: Pairs of commuting matrices over a finite field. Duke Math. J. **27**, 91–94 (1960)
7. Fouvry, E., Katz, N.: A general stratification theorem for exponential sums, and applications. J. Reine Angew. Math. **540**, 115–166 (2001)
8. Fujiwara, M.: Upper bounds for the number of lattice points on hypersurfaces. Number theory and combinatorics (Japan, 1984), 89–96, World Scientific (1985)
9. Fujiwara, M.: Distribution of rational points on varieties over finite fields. Mathematika **35**, 155–171 (1988)
10. Hartshorne, R.: Algebraic Geometry. Grad. Texts in Math. **52**, Springer (1977)
11. Heath-Brown, D.R.: The density of rational points on non-singular hypersurfaces. Proc. Indian Acad. Sci. (Math. Sci.), **104**, 13–29 (1994)
12. Knutson, A.: Some schemes related to the commuting variety. J. Algebraic Geom. **14**, 283–294 (2005)
13. Knutson, A., Zinn-Justin, P.: A scheme related to the Brauer loop model. Adv. Math. **214**, 40–77 (2007)
14. Larsen, M., Lu, Z.: Flatness of the commutator map over ${}_n$. IMRN, 5605–5622 (2021)
15. Matsumura, H.: Commutative Ring Theory. Cambridge University Press, Cambridge (1986)
16. Motzkin, T.S., Taussky, O.: Pairs of matrices with property L, II. Trans. Am. Math. Soc. **80**, 387–401 (1955)
17. Neubauer, M.G.: The variety of pairs of matrices with $\operatorname{rank}(AB - BA) \leqslant 1$. Proc. Am. Math. Soc. **105**, 787–792 (1989)

18. Poonen, B.: Rational Points on Varieties. Graduate Studies in Math. **186**, American Mathematical Society, Providence, RI (2017)
19. Salberger, P.: Counting rational points on projective varieties. Proc. Lond. Math. Soc. **126**, 1092–1133 (2023)
20. Shoda, K.: Einige Sätze über Matrizen. Japan. J. Math. **13**, 361–365 (1937)
21. Shparlinski, I.E., Skorobogatov, A.N.: Exponential sums and rational points on complete intersections. Mathematika **37**, 201–208 (1990)
22. Skorobogatov, A.N.: Exponential sums, the geometry of hyperplane sections, and some Diophantine problems. Israel J. Math. **80**, 359–379 (1992)
23. The Stacks project authors, *The Stacks project*. https://stacks.math.columbia.edu (2024)
24. Vermeulen, F.: Dimension growth for affine varieties. IMRN, 11464–11483 (2024)
25. Young, H.-W.V.: On matrix pairs with diagonal commutators. J. Algebra **570**, 437–451 (2021)