Journal of Number Theory 133 (2013) 3578–3588

Contents lists available at SciVerse ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

# On a uniformly distributed phenomenon in matrix groups

Su Hu [a], Yan Li [b,*]

[a] *Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. West, Montréal, Québec, H3A 2K6, Canada*
[b] *Department of Applied Mathematics, China Agriculture University, Beijing 100083, China*

A R T I C L E  I N F O

*Article history:*
Received 25 February 2013
Revised 15 April 2013
Accepted 15 April 2013
Available online 5 July 2013
Communicated by Dinesh S. Thakur

*MSC:*
11C20
11T23

*Keywords:*
Matrices
Finite fields
Uniform distribution
Character sum

A B S T R A C T

We show that a classical uniformly distributed phenomenon for an element and its inverse in $(\mathbb{Z}/n\mathbb{Z})^*$ also exists in $\mathrm{GL}_n(\mathbb{F}_p)$. A $\mathrm{GL}_n(\mathbb{F}_p)$ analogy of the uniform distribution on modular hyperbolas has also been considered.



## 1. Introduction

The distance between an element $x \in (\mathbb{Z}/n\mathbb{Z})^*$ and its inverse $x^{-1} \pmod n$ has been studied by many authors [1,4,11,18,20–22]. Shparlinski [16] gave a survey of a variety of recent results about the distribution and some geometric properties of points $(x, y)$ on modular hyperbolas $xy \equiv a \pmod n$.

Denote by $\{x\}$ the fractional part of a real number $x$. Let

$$f_n : (\mathbb{Z}/n\mathbb{Z})^* \to [0,1] \times [0,1],$$
$$x \mapsto \left( \left\{ \frac{x}{n} \right\}, \left\{ \frac{x^{-1}}{n} \right\} \right).$$

* Corresponding author.
*E-mail addresses:* hus04@mails.tsinghua.edu.cn, hu@math.mcgill.ca (S. Hu), liyan_00@mails.tsinghua.edu.cn (Y. Li).

0022-314X/$ – see front matter 
http://dx.doi.org/10.1016/j.jnt.2013.04.020

By using the Erdös–Turán–Koksma inequality and the Weil–Estermann inequality for Kloosterman sum, Beck and Khan [1] gave an elegant proof for the following classical result.

**Theorem 1.1.** *Let $R \subset [0,1]^2$ be a measurable set having the following property that for every $\epsilon > 0$, there exist two finite collections of non-overlapping rectangles $R_1, \dots, R_k$ and $R^1, \dots, R^l$ such that $\bigcup_{i=1}^k R_i \subseteq R \subseteq \bigcup_{j=1}^l R^j$, $\operatorname{area}(R/\bigcup_{i=1}^k R_i) < \epsilon$ and $\operatorname{area}(\bigcup_{j=1}^l R^j/R) < \epsilon$. Then*

$$\lim_{n\to\infty} \frac{\operatorname{cardinality}(\operatorname{Image}(f_n) \cap R)}{\varphi(n)} = \operatorname{area}(R).$$

**Remark 1.2.** Notice that our statement of the above theorem is slightly different from the statement in Beck and Khan [1]. The statement in [1] is as follows:

"Let $R \subseteq [0,1]^2$ be a measurable set having the following property that for every $\epsilon > 0$, there exists a finite collection of non-overlapping rectangles $\{R_1, R_2, \dots, R_k\}$ such that $\bigcup_{i=1}^k R_i \subseteq R$ and $\operatorname{area}(R/\bigcup_{i=1}^k R_i) < \epsilon$. Then

$$\lim_{n\to\infty} \frac{\operatorname{cardinality}(\operatorname{Image}(f_n) \cap R)}{\varphi(n)} = \operatorname{area}(R)."$$

(See Theorem 2 of [1].)

The original assumption should be strengthened. Otherwise there is a counterexample as follows:

Let $R_1 = [0,1/2)^2$ and $R_2 = \{(x,y) \in [0,1]^2 \mid x, y \in \mathbb{Q}\}$. Denote by $R = R_1 \cup R_2$. Since $\operatorname{area}(R_2) = 0$, we have $\operatorname{area}(R) = \operatorname{area}(R_1) = 1/4$. So $R$ satisfies the conditions in the statement of Theorem 2 in [1]. Since the image of $f_n$ are rational points in $[0,1]^2$ and $R$ contains all the rational points in $[0,1]^2$, we have $\operatorname{Image}(f_n) \cap R = \varphi(n)$ for any positive integer $n$, thus

$$\lim_{n\to\infty} \frac{\operatorname{cardinality}(\operatorname{Image}(f_n) \cap R)}{\varphi(n)} = 1.$$

But $\operatorname{area}(R) = \operatorname{area}(R_1) = 1/4$, so

$$\lim_{n\to\infty} \frac{\operatorname{cardinality}(\operatorname{Image}(f_n) \cap R)}{\varphi(n)} \neq \operatorname{area}(R).$$

Notice that, the new conditions in Theorem 1.1 are quite natural. Numerous types of regions satisfy the conditions of Theorem 1.1 such as polygons, disks, annuli lying in the unit square.

Beck and Khan [1, p. 150] remarked that: "In all likelihood this theorem dates back to the late 20's and early 30's and was known to mathematicians such as Davenport, Estermann, Kloosterman, Salie."

Let $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{p-1}\}$ be the finite field with $p$ elements, $M_n(\mathbb{F}_p)$ be the set of all $n \times n$ matrices over $\mathbb{F}_p$, $\mathrm{GL}_n(\mathbb{F}_p)$, $\mathrm{SL}_n(\mathbb{F}_p)$ and $\mathcal{Z}_n(\mathbb{F}_p)$ be the group of invertible matrices, the group of matrices of determinant 1 and the set of singular matrices, respectively, where all matrices are from $M_n(\mathbb{F}_p)$.

In this paper, by using bounds of Ferguson, Hoffman, Ostafe, Luca and Shparlinski [3] for the matrix analogue of classical Kloosterman sums (see Lemma 2.1 below), we show that the above mentioned uniformly distributed phenomenon also exists in $\mathrm{GL}_n(\mathbb{F}_p)$.

For $A = (\overline{a_{ij}}) \in \mathrm{GL}_n(\mathbb{F}_p)$, $A^{-1} = (\overline{b_{ij}})$ denotes the inverse of $A$.

Let

$$g_p : \mathrm{GL}_n(\mathbb{F}_p) \to \underbrace{[0,1] \times [0,1] \times \cdots \times [0,1]}_{2n^2}, \tag{1.1}$$

$$A = (\overline{a_{ij}}) \mapsto \begin{pmatrix} \frac{a_{11}}{p}, & \ldots, & \frac{a_{1n}}{p}, & \frac{b_{11}}{p}, & \ldots, & \frac{b_{1n}}{p} \\ \frac{a_{21}}{p}, & \ldots, & \frac{a_{2n}}{p}, & \frac{b_{21}}{p}, & \ldots, & \frac{b_{2n}}{p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{a_{n1}}{p}, & \ldots, & \frac{a_{nn}}{p}, & \frac{b_{n1}}{p}, & \ldots, & \frac{b_{nn}}{p} \end{pmatrix}.$$

In fact, we prove the following theorem.

**Theorem 1.3.** *Let $R \subset [0,1]^{2n^2}$ be a measurable set having the following property that for every $\epsilon > 0$, there exist two finite collections of non-overlapping rectangles $R_1, \ldots, R_k$ and $R^1, \ldots, R^l$ such that $\bigcup_{i=1}^k R_i \subset R \subset \bigcup_{j=1}^l R^j$, $\text{area}(R/\bigcup_{i=1}^k R_i) < \epsilon$ and $\text{area}(\bigcup_{j=1}^l R^j/R) < \epsilon$. Then*

$$\lim_{p\to\infty} \frac{\text{cardinality}(\text{Image}(g_p) \cap R)}{\#\text{GL}_n(\mathbb{F}_p)} = \text{area}(R).$$

**Remark 1.4.** A $\text{GL}_n(\mathbb{F}_p)$ analogy of the uniform distribution on modular hyperbolas can also be established using the same procedure for the proof of the above theorem (see Remark 3.1 below).

Furthermore, let

$$h_p : \text{GL}_n(\mathbb{F}_p) \to \underbrace{[0,1] \times [0,1] \times \cdots \times [0,1]}_{n^2}, \tag{1.2}$$

$$A = (\overline{a_{ij}}) \mapsto \begin{pmatrix} \frac{a_{11}}{p}, & \ldots, & \frac{a_{1n}}{p} \\ \frac{a_{21}}{p}, & \ldots, & \frac{a_{2n}}{p} \\ \vdots & \vdots & \vdots \\ \frac{a_{n1}}{p}, & \ldots, & \frac{a_{nn}}{p} \end{pmatrix},$$

$$s_p : \text{SL}_n(\mathbb{F}_p) \to \underbrace{[0,1] \times [0,1] \times \cdots \times [0,1]}_{n^2}, \tag{1.3}$$

$$A = (\overline{a_{ij}}) \mapsto \begin{pmatrix} \frac{a_{11}}{p}, & \ldots, & \frac{a_{1n}}{p} \\ \frac{a_{21}}{p}, & \ldots, & \frac{a_{2n}}{p} \\ \vdots & \vdots & \vdots \\ \frac{a_{n1}}{p}, & \ldots, & \frac{a_{nn}}{p} \end{pmatrix}.$$

Using the same methods, and the bounds of Ferguson, Hoffman, Ostafe, Luca and Shparlinski [3] on character sums along $\mathcal{Z}_n(\mathbb{F}_p)$ and $\text{SL}_n(\mathbb{F}_p)$ (see Lemmas 2.2, 2.3 and 2.4 below), we can also obtain the following two results.

**Theorem 1.5.** *Let $R \subseteq [0,1]^{n^2}$ be a measurable set having the same property as in Theorem 1.3. Then*

$$\lim_{p\to\infty} \frac{\text{cardinality}(\text{Image}(h_p) \cap R)}{\#\text{GL}_n(\mathbb{F}_p)} = \text{area}(R).$$

**Theorem 1.6.** *Assumption as above, then*

$$\lim_{p\to\infty} \frac{\text{cardinality}(\text{Image}(s_p) \cap R)}{\#\text{SL}_n(\mathbb{F}_p)} = \text{area}(R).$$

## 2. Preliminaries

We need some lemmas to prove the main theorems.

First, we recall some results in [3].

Given two matrices $U = (u_{ij})$, $X = (x_{ij}) \in M_n(\mathbb{F}_p)$, their product is defined by

$$U \cdot X = \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} u_{ij} x_{ij}$$

(see [3, p. 503]).

Let $q$ be a power of a prime number, $\Psi$ be a fixed nonprincipal additive character of $\mathbb{F}_q$. For $\mathcal{M}, U, V \in M_n(\mathbb{F}_q)$, let

$$K\big(\mathrm{GL}_n(\mathbb{F}_q), U, V, \mathcal{M}\big) = \sum_{X \in \mathrm{GL}_n(\mathbb{F}_q)} \Psi\big(U \cdot X + V \cdot \big(\mathcal{M} X^{-1}\big)\big)$$

be the matrix analogue of classical Kloosterman sums (see [3, p. 505]) and

$$S\big(\mathrm{GL}_n(\mathbb{F}_q), U\big) = \sum_{X \in \mathrm{GL}_n(\mathbb{F}_q)} \Psi(U \cdot X),$$

$$S\big(\mathrm{SL}_n(\mathbb{F}_q), U\big) = \sum_{X \in \mathrm{SL}_n(\mathbb{F}_q)} \Psi(U \cdot X),$$

$$S\big(\mathcal{Z}_n(\mathbb{F}_q), U\big) = \sum_{X \in \mathcal{Z}_n(\mathbb{F}_q)} \Psi(U \cdot X).$$

The authors in [3] obtained the following results.

**Lemma 2.1.** *(See [3, Lemma 5].) Uniformly over all matrices $U, V \in M_n(\mathbb{F}_q)$ among which at least one is a nonzero matrix, and $\mathcal{M} \in \mathrm{GL}_n(\mathbb{F}_q)$, we have*

$$K\big(\mathrm{GL}_n(\mathbb{F}_q), U, V, \mathcal{M}\big) \ll q^{n^2 - 1/2},$$

*where the implied constant in the symbol "$\ll$" depends only on $n$.*

**Lemma 2.2.** *(See [3, Lemma 3].) Uniformly over all nonzero matrices $U \in M_n(\mathbb{F}_q)$, we have*

$$S\big(\mathcal{Z}_n(\mathbb{F}_q), U\big) = O\big(q^{n^2 - 5/2}\big),$$

*where the implied constant in the symbol "$O$" depends only on $n$.*

**Lemma 2.3.** *Uniformly over all nonzero matrices $U \in M_n(\mathbb{F}_q)$, we have*

$$S\big(\mathrm{GL}_n(\mathbb{F}_q), U\big) = O\big(q^{n^2 - 5/2}\big),$$

*where the implied constant in the symbol "$O$" depends only on $n$.*

**Proof.** For any nonzero matrix $U \in M_n(\mathbb{F}_q)$, $\bar{\Psi}(X) = \Psi(U \cdot X)$ is also a nontrivial additive character on $M_n(\mathbb{F}_q)$, so we have

$$S\big(\mathcal{Z}_n(\mathbb{F}_q), U\big) + S\big(\mathrm{GL}_n(\mathbb{F}_q), U\big) = \sum_{X \in M_n(\mathbb{F}_q)} \Psi(U \cdot X) = 0.$$

From Lemma 2.2, we obtain the result. □

**Lemma 2.4.** *(See [3, Lemma 4].) Uniformly over all nonzero matrices $U \in M_n(\mathbb{F}_q)$, we have*

$$S\big(\mathrm{SL}_n(\mathbb{F}_q), U\big) = O\big(q^{n^2-2}\big),$$

*where the implied constant in the symbol "O" depends only on n.*

**Remark 2.5.** Recently, we obtained explicit expressions of $S(\mathrm{GL}_n(\mathbb{F}_q), U)$ and $S(\mathrm{SL}_n(\mathbb{F}_q), U)$. (See [12, Theorems 2.1 and 2.4].) Such expressions only involve Gauss sums and Kloosterman sums. As a consequence, we got

$$S\big(\mathrm{GL}_n(\mathbb{F}_q), U\big) = O\big(q^{n^2-n}\big),$$
$$S\big(\mathrm{SL}_n(\mathbb{F}_q), U\big) = O\big(q^{n^2-n}\big) = O\big(\max\big\{q^{n^2-n-1}, q^{(n^2-1)/2}\big\}\big).$$

(See [12, Corollaries 2.2 and 2.5].) Our bounds improved the bounds in Lemmas 2.3 and 2.4. (See [12, Remarks 2.3 and 2.6].)

Next we recall Erdös–Turán–Koksma inequality for the discrepancy of sequences.

Let $B = [a_1, b_1) \times \cdots \times [a_k, b_k) \subseteq [0, 1)^k$ be a rectangle, $(\boldsymbol{x}_n)$ be a sequence in $[0, 1)^k$, and $A(B, N, \boldsymbol{x}_n)$ be the number of points $\boldsymbol{x}_n$, $1 \leqslant n \leqslant N$, such that $\boldsymbol{x}_n \in B$, i.e.

$$A(B, N, \boldsymbol{x}_n) = \sum_{n=1}^{N} \chi_B(\boldsymbol{x}_n),$$

where $\chi_B$ is the characteristic function of $B$.

**Definition 2.6.** (See [2, p. 4].) Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be a finite sequence of points in $[0, 1)^k$. Then the number

$$D_N = D_N(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = \sup_{B \subseteq [0,1)^k} \left| \frac{A(B, N, \boldsymbol{x}_n)}{N} - \mathrm{area}(B) \right|$$

is called the discrepancy of the given sequence, where $B$ runs over all rectangles located in $[0, 1)^k$.

Set $e(x) = \exp(2\pi i x)$ and denote the usual inner product in $\mathbb{R}^k$ by $\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{k} x_i y_i$.

The Erdös–Turán–Koksma inequality provides an upper bound for the discrepancy.

**Lemma 2.7.** *(See [2, p. 15] or [13, p. 63].) Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ be a finite sequence of points in $[0, 1)^k$ and H an arbitrary positive integer. Then*

$$D_N \leqslant \left(\frac{3}{2}\right)^k \left(\frac{2}{H+1} + \sum_{0<\|\boldsymbol{h}\|_\infty \leqslant H} \frac{1}{r(\boldsymbol{h})} \left|\frac{1}{N}\sum_{n=1}^{N} e(\boldsymbol{h}\cdot\boldsymbol{x}_n)\right|\right),$$

*where* $r(\boldsymbol{h}) = \prod_{i=1}^{k} \max\{1, |h_i|\}$ *and* $\|\boldsymbol{h}\|_\infty = \max\{|h_i| \mid 1 \leqslant i \leqslant k\}$*, for* $\boldsymbol{h} = (h_1, \ldots, h_k) \in \mathbb{Z}^k$.

## 3. Proofs of main results

**Proof of Theorem 1.3.** We only need to prove the case that $R = [a_1, b_1) \times \cdots \times [a_k, b_k) \subset [0, 1)^k$ is a rectangle. The reason is as follows:

Assume the theorem holds for $R$ being a rectangle. Let $R$ be a measurable set as in the assumptions. For every $\epsilon > 0$, let $R_1, \ldots, R_k$ and $R^1, \ldots, R^l$ be two finite collections of non-overlapping rectangles such that

$$\bigcup_{i=1}^{k} R_i \subseteq R \subseteq \bigcup_{j=1}^{l} R^j, \qquad \operatorname{area}\left(R \Big/ \bigcup_{i=1}^{k} R_i\right) < \epsilon \quad \text{and} \quad \operatorname{area}\left(\bigcup_{j=1}^{l} R^j \Big/ R\right) < \epsilon.$$

Then

$$\sum_{i=1}^{k} \operatorname{area}(R_i) \leqslant \operatorname{area}(R) \leqslant \sum_{j=1}^{l} \operatorname{area}\big(R^j\big),$$

$$\sum_{i=1}^{k} \frac{\#(\operatorname{Image}(g_p) \cap R_i)}{\#\mathrm{GL}_n(\mathbb{F}_p)} \leqslant \frac{\#(\operatorname{Image}(g_p) \cap R)}{\#\mathrm{GL}_n(\mathbb{F}_p)} \leqslant \sum_{j=1}^{l} \frac{\#(\operatorname{Image}(g_p) \cap R^j)}{\#\mathrm{GL}_n(\mathbb{F}_p)}. \tag{3.1}$$

Taking $p$ sufficiently large, the left hand (resp. right hand) sides of the above two inequalities are sufficiently close, i.e., there exists $M$ such that if $p > M$ then

$$\left|\sum_{i=1}^{k} \frac{\#(\operatorname{Image}(g_p) \cap R_i)}{\#\mathrm{GL}_n(\mathbb{F}_p)} - \sum_{i=1}^{k} \operatorname{area}(R_i)\right| < \epsilon,$$

$$\left|\sum_{j=1}^{l} \frac{\#(\operatorname{Image}(g_p) \cap R^j)}{\#\mathrm{GL}_n(\mathbb{F}_p)} - \sum_{j=1}^{l} \operatorname{area}\big(R^j\big)\right| < \epsilon. \tag{3.2}$$

Since

$$0 \leqslant \sum_{j=1}^{l} \operatorname{area}\big(R^j\big) - \sum_{i=1}^{k} \operatorname{area}(R_i) < 2\epsilon, \tag{3.3}$$

then from inequalities (3.1), (3.2), and (3.3), for $p > M$, we have

$$\left|\frac{\#(\operatorname{Image}(g_p) \cap R)}{\#\mathrm{GL}_n(\mathbb{F}_p)} - \operatorname{area}(R)\right| < 4\epsilon.$$

This implies that

$$\lim_{p\to\infty} \frac{\operatorname{cardinality}(\operatorname{Image}(g_p) \cap R)}{\#\mathrm{GL}_n(\mathbb{F}_p)} = \operatorname{area}(R).$$

Now we prove the fundamental case in which $R$ is the rectangle $[a_1, b_1) \times \cdots \times [a_k, b_k)$.

In Lemma 2.7, viewing the points $\boldsymbol{x}$ in $\mathrm{Image}(g_p)$ as a sequence in $[0,1)^k$, where $N = \#\mathrm{GL}_n(\mathbb{F}_p)$ and $k = 2n^2$, we have

$$\left|\frac{\mathrm{cardinality}(\mathrm{Image}(g_p)\cap R)}{\#\mathrm{GL}_n(\mathbb{F}_p)} - \mathrm{area}(R)\right| \\ \ll \frac{2}{H+1} + \sum_{0<\|\boldsymbol{h}\|_\infty \leqslant H} \frac{1}{r(\boldsymbol{h})}\left|\frac{1}{\#\mathrm{GL}_n(\mathbb{F}_p)} \sum_{\boldsymbol{x}\in \mathrm{Image}(g_p)} e(\boldsymbol{h}\cdot\boldsymbol{x})\right|, \tag{3.4}$$

where $r(\boldsymbol{h}) = \prod_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant n} \max\{1, |h_{ij}|\}$ for

$$\boldsymbol{h} = \begin{pmatrix} h_{11}, & \ldots, & h_{1n}, & h_{(n+1)1}, & \ldots, & h_{(n+1)n} \\ h_{21}, & \ldots, & h_{2n}, & h_{(n+2)1}, & \ldots, & h_{(n+2)n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{n1}, & \ldots, & h_{nn}, & h_{(2n)1}, & \ldots, & h_{(2n)n} \end{pmatrix}$$

in $\mathbb{Z}^{2n^2}$.

Since $\#\mathrm{GL}_n(\mathbb{F}_p) = p^{n^2} + O(p^{n^2-1})$, if $\boldsymbol{h}$ modulo $p$ is nonzero, then by Lemma 2.1 (taking $\mathcal{M} = I$, $U = (\overline{h_{ij}})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant n}$, $V = (\overline{h_{(n+i)j}})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant n}$, $X = A$), we will get

$$\frac{1}{r(\boldsymbol{h})}\left|\frac{1}{\#\mathrm{GL}_n(\mathbb{F}_p)} \sum_{\boldsymbol{x}\in \mathrm{Image}(g_p)} e(\boldsymbol{h}\cdot\boldsymbol{x})\right| \ll p^{-1/2}. \tag{3.5}$$

Notice that

$$\#\left\{\boldsymbol{h}\in\mathbb{Z}^{2n^2} \;\middle|\; 0 < \|\boldsymbol{h}\|_\infty \leqslant H\right\} \ll H^{2n^2}. \tag{3.6}$$

From (3.4), (3.5), (3.6), and taking $H = \lfloor p^{1/(2(2n^2+1))}\rfloor$, we get

$$\left|\frac{\mathrm{cardinality}(\mathrm{Image}(g_p)\cap R)}{\#\mathrm{GL}_n(\mathbb{F}_p)} - \mathrm{area}(R)\right| \ll \frac{2}{H+1} + H^{2n^2} p^{-1/2} \\ \ll p^{-1/(2(2n^2+1))}. \tag{3.7}$$

Letting $\lim_{p\to\infty}$ in (3.7), we get our result. □

**Remark 3.1.** For any fixed $C = (\overline{a_{ij}}) \in \mathrm{GL}_n(\mathbb{F}_p)$, we consider the matrix equation $BA = C$, where $A = (\overline{a_{ij}})$, $B = (\overline{b_{ij}})$ in $\mathrm{GL}_n(\mathbb{F}_p)$.

Let

$$\tilde{g}_p : \mathrm{GL}_n(\mathbb{F}_p) \to \underbrace{[0,1]\times[0,1]\times\cdots\times[0,1]}_{2n^2}, \tag{3.8}$$

$$A = (\overline{a_{ij}}) \mapsto \begin{pmatrix} \frac{a_{11}}{p}, & \ldots, & \frac{a_{1n}}{p}, & \frac{b_{11}}{p}, & \ldots, & \frac{b_{1n}}{p} \\ \frac{a_{21}}{p}, & \ldots, & \frac{a_{2n}}{p}, & \frac{b_{21}}{p}, & \ldots, & \frac{b_{2n}}{p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{a_{n1}}{p}, & \ldots, & \frac{a_{nn}}{p}, & \frac{b_{n1}}{p}, & \ldots, & \frac{b_{nn}}{p} \end{pmatrix}.$$

The same procedure can show Theorem 1.3 is also established for $\tilde{g}_p$. The only change is that we take $\mathcal{M} = C$, $U = (h_{ij})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant n}$, $V = (h_{(n+i)j})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant n}$ and $X = A$ in Lemma 2.1. This can be viewed as a $\mathrm{GL}_n(\mathbb{F}_p)$ analogy of the uniform distribution on modular hyperbolas.

**Remark 3.2.** As pointed out by the referee, we may also consider classes of sets for which the desired epsilon-approximation can be achieved with $O(\epsilon^{-\alpha})$ rectangles and obtain an explicit bound on the discrepancy, depending on $\alpha$. A result by Wolfgang M. Schmidt in [15] may give an explicit $\alpha$ for say convex sets.

**Proof of Theorem 1.5.** The method is exactly the same as in Theorem 1.3, except that we replace the estimation of $K(\mathrm{GL}_n(\mathbb{F}_p), U, V, \mathcal{M})$ by the estimation of $S(\mathrm{GL}_n(\mathbb{F}_p), U)$ (see Lemma 2.3). □

**Proof of Theorem 1.6.** By [14, Theorem 3.15 (iiib)], we have

$$\begin{aligned} \#\mathrm{SL}_n(\mathbb{F}_p) &= \frac{p^{n(n-1)/2}}{p-1} \cdot \prod_{j=1}^{n} (p^j - 1) \\ &= p^{n(n-1)/2}(p^2 - 1)\cdots(p^n - 1) \\ &= p^{n^2-1} + O(p^{n^2-3}). \end{aligned} \tag{3.9}$$

The method to prove this theorem is exactly the same as in Theorem 1.3, except that we replace the estimation of $K(\mathrm{GL}_n(\mathbb{F}_p), U, V, \mathcal{M})$ by the estimation of $S(\mathrm{SL}_n(\mathbb{F}_p), U)$ (see Lemma 2.4) and notice that formula (3.9) holds. □

## 4. Further discussions

In this section, we discuss the following questions:

Is the image of $\mathrm{SL}_n(\mathbb{F}_p)$ under $g_p$ (see (1.1)), uniformly distributed in $[0,1]^{2n^2}$?

The case $n = 1$ is trivial.

For $n = 2$, since

$$\begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}^{-1} = \begin{pmatrix} \bar{d} & -\bar{b} \\ -\bar{c} & \bar{a} \end{pmatrix}, \quad \text{where } \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{F}_p),$$

one can easily find a nonzero vector

$$\boldsymbol{h} = \begin{pmatrix} h_{11}, & h_{12}, & h_{13}, & h_{14} \\ h_{21}, & h_{22}, & h_{23}, & h_{24} \end{pmatrix} \in \mathbb{Z}^8$$

such that

$$\boldsymbol{x} \mapsto e(\boldsymbol{h}\cdot\boldsymbol{x}) = 1, \quad \text{for all } \boldsymbol{x} \in g_p\big(\mathrm{SL}_2(\mathbb{F}_p)\big).$$

For example, take arbitrary nonzero $\boldsymbol{h}$ with $h_{11} + h_{24} = 0$, $h_{12} - h_{14} = 0$, $h_{13} + h_{22} = 0$ and $h_{21} - h_{23} = 0$.

Hence

$$\lim_{p\to\infty} \frac{1}{\#\mathrm{SL}_2(\mathbb{F}_p)} \sum_{\boldsymbol{x}\in g_p(\mathrm{SL}_2(\mathbb{F}_p))} e(\boldsymbol{h}\cdot\boldsymbol{x}) \neq \int_{[0,1]^8} e(\boldsymbol{h}\cdot\boldsymbol{x})\, d\boldsymbol{x}.$$

This implies that the image of $\mathrm{SL}_2(\mathbb{F}_p)$ under $g_p$ is not uniformly distributed, that is, Theorem 1.3 does not hold for $\mathrm{SL}_2(\mathbb{F}_p)$.

For the case $n \geqslant 3$, we conjecture that elements and their inverses in $\mathrm{SL}_n(\mathbb{F}_p)$ are also uniformly distributed.

**Remark 4.1.** As in the proof of [3, Lemma 4], the above conjecture may be treated within a general theory of bounds of exponential sums along varieties. Thus, as pointed out by the referee, the estimations in [6–8,17,19] may be relevant to this consideration. In this remark, we shall mention and discuss some of the results in these works, and we hope the above conjecture may be solved in the future along this way.

Let $V \subset \mathbf{A}^n_{\mathbf{Z}}$ be a closed subscheme of dimension $\leqslant d$ defined by the vanishing of several polynomials, let $f(X_1, \ldots, X_n)$ be a polynomial defined as a function on $V$, let $\psi$ be a nontrivial additive character of $\mathbb{F}_p$, and let $\boldsymbol{h} = (h_1, \ldots, h_n) \in \mathbf{A}^n(\mathbb{F}_p)$. We set

$$S_{V,f}(\boldsymbol{h}; p) := \sum_{(x_1,\ldots,x_n)\in V(\mathbb{F}_p)} \psi\big(f(x_1, \ldots, x_n) + h_1x_1 + \cdots + h_nx_n\big).$$

The "good" bound which is expected is

$$S_{V,f}(\boldsymbol{h}; p) \leqslant C(V, f)p^{d/2},$$

which is essentially equivalent to the Riemann hypothesis for an appropriate $L$-function over the finite field $\mathbb{F}_p$.

In [10], using the formalism of perversity and the properties of the geometric Fourier transform, Katz and Laumon proved that, for $f \equiv 0$ and under an extremely mild hypothesis on $V$ (in particular, no smoothness assumptions are required), and for $p$ large enough (depending on $V$), the "good" bound holds for all $\boldsymbol{h}$ lying outside the set of $\mathbb{F}_p$-points of a codimension-one subscheme $X_1 \subset \mathbf{A}^n_{\mathbf{Z}}$.

Let $f = 0$. In [6], Fouvry investigated the structure of this exceptional locus $X_1$ on which the good bound does not hold, and in particular at the "size" of the intermediate subsets ($X_j$, say) of the $\boldsymbol{h}$ for which the bound for $S_{V,f}(\boldsymbol{h}; p)$ deviates from the good one by a factor of at least $p^{j/2}$. Furthermore, he also applied these results to prove several new results on the distribution of rational points of varieties over finite fields.

In the case of a general $f$, Fouvry and Katz [7] established the existence of a decreasing filtration $\cdots \subset X_j \subset X_{j-1} \subset \cdots \subset X_1 \subset X_0 := \mathbf{A}^n_{\mathbf{Z}}$, by closed subschemes of codimension $j$, such that for $p$ large enough and $\boldsymbol{h} \notin X_j(\mathbb{F}_p)$, the following bound holds:

$$S_{V,f}(\boldsymbol{h}; p) \leqslant C(V, f)p^{d/2+(j-1)/2}.$$

Note that in the above statement, there is essentially no assumption on the regularity of $V$ or $f$. When $f \equiv 0$, $V_{\mathbf{C}}$ is smooth, and an extra geometrical condition is satisfied (i.e. the non-vanishing of a certain "A-number" which is the rank of an $l$-adic sheaf attached to the situation), they also improved the above bound as

$$S_{V,f}(\boldsymbol{h}; p) \leqslant C(V, f)p^{\sup(d/2, d/2+(j-2)/2)},$$

for $\boldsymbol{h} \notin X_j(\mathbb{F}_p)$. Furthermore, they also presented several Diophantine applications, in particular, they improved some of previous results by Fouvry in [5].

Let $k$ be a finite field, let $X/k$ be a closed subscheme of the projective space $\mathbf{P}^N$ of dimension $n \geqslant 1$, defined by a set of polynomials of fixed degree $D_1, \ldots, D_r$, let $L \in H^0(X, \mathcal{O}(1))$ be a linear form on $X$ and for a fixed integer $d$, let $H \in H^0(X, \mathcal{O}(d))$ be a form of degree $d$; set $X \cap L = X \cap \{L = 0\}$ and $V = X - X \cap L$; and let $\psi : k \to \mathbb{C}^\times$ be a nontrivial additive character. Denote by

$$S := \sum_{x\in V(k)} \psi\big(\big(H/L^d\big)(x)\big).$$

In [8], Katz proved the sharp upper bound $|S| \leqslant C|k|^{n/2}$ (for some constant $C = C(N, d, D_1, \dots, D_r)$ depending on $N, d, D_1, \dots, D_r$ only), under the assumptions that char $k$ is coprime with $d$, that $X$, $X \cap L$ and $X \cap L \cap H$ are all nonsingular, and that $X \cap L$ and $X \cap L \cap H$ have codimension one and two, respectively, in $X$. In [9], he further removed many of the smoothness assumptions made above; more precisely, assuming either $\mathrm{H}_1$: "$X \otimes_k \bar{k}$ is irreducible and integral" or $\mathrm{H}'_1$: "$X$ is Cohen–Macaulay and equidimensional" and also assuming $\mathrm{H}_2$: "the scheme $X \cap L \cap H$ has codimension 2," he obtained the upper bound $|S| \leqslant C|k|^{(n+\delta+1)/2}$, where $\delta$ is the dimension of the singular locus of $X \cap L \cap H$, at least when char $k$ is large enough. He also proved a slightly weaker result, valid in any characteristic coprime with $d$ under the same assumptions ($\mathrm{H}_1$ or $\mathrm{H}'_1$, and $\mathrm{H}_2$). Furthermore, he also applied these general bounds to the case of hypersurfaces.

Shparlinski and Skorobogatov [17] estimated the modulus of exponential sums over the variety of dimension $n - s$ defined by a system of $s$ forms in $n$ variables, with a linear form in the exponent. They also applied the estimations to the study of the distribution of rational points on such a variety defined over a finite field or the field of rationals.

Using the results of Deligne's Weil Conjectures II and a generalization of Lefschetz hyperplane theorem to singular varieties, Skorobogatov [19] estimated exponential sums with additive character along an affine variety given by a system of homogeneous equations, with a homogeneous function in the exponent. He applied his estimate to obtain an upper bound for the number of integer solutions of a system of homogeneous equations in a box. Furthermore, he also applied his estimate to uniform distribution of solutions of a system of homogeneous congruences modulo a prime in the following sense: the portion of solutions in a box is proportional to the volume of the box, provided the box is not very small.

## Acknowledgments

The authors are grateful to the anonymous referee for his/her helpful comments.

Y. Li is supported by the National Natural Science Foundation of China (Grant No. 11101424 and Grant No. 11071277).

## References

[1] J. Beck, M.R. Khan, On the uniform distribution of inverse modulo $n$, Period. Math. Hungar. 44 (2002) 147–155.
[2] M. Damota, R.F. Tichy, Sequences, Discrepancies and Applications, Lecture Notes in Math., vol. 1652, Springer-Verlag, 1997.
[3] R. Ferguson, C. Hoffman, F. Luca, A. Ostafe, I.E. Shparlinski, Some additive combinatorics problems in matrix rings, Rev. Mat. Complut. 23 (2010) 501–513.
[4] K. Ford, M.R. Khan, I.E. Shparlinski, C.L. Yankov, On the maximal difference between an element and its inverse in residue rings, Proc. Amer. Math. Soc. 133 (2005) 3463–3468.
[5] E. Fouvry, Sur les propriétés de divisibilité des nombres de classes des corps quadratiques, Bull. Soc. Math. France 127 (1999) 95–113 (in French).
[6] E. Fouvry, Consequences of a result of N. Katz and G. Laumon concerning trigonometric sums, Israel J. Math. 120 (2000) 81–96.
[7] E. Fouvry, N. Katz, A general stratification theorem for exponential sums, and applications, J. Reine Angew. Math. 540 (2001) 115–166.
[8] N.M. Katz, Sommes exponentielles, course taught at the University of Paris, Orsay, Fall 1979, with a preface by Luc Illusie, notes written by Gérard Laumon, with an English summary, Astérisque, vol. 79, Société Mathématique de France, Paris, 1980 (in French).
[9] N.M. Katz, Estimates for "singular" exponential sums, Int. Math. Res. Not. IMRN 16 (1999) 875–899.
[10] N.M. Katz, G. Laumon, Transformation de Fourier et majoration de sommes exponentielles, Inst. Hautes Études Sci. Publ. Math. 62 (1985) 361–418 (in French); see also Erratum: "The Fourier transform and upper bounds for exponential sums", Inst. Hautes Études Sci. Publ. Math. 69 (1989) 233.
[11] M.R. Khan, I.E. Shparlinski, On the maximal difference between an element and its inverse modulo $n$, Period. Math. Hungar. 47 (2003) 111–117.
[12] Y. Li, S. Hu, Gauss sums over some matrix groups, J. Number Theory 132 (2012) 2967–2976.
[13] H. Niederreiter, A. Winterhof, Exponential sums for nonlinear recurring sequences, Finite Fields Appl. 14 (2008) 59–64.
[14] H.E. Rose, A Course on Finite Groups, Universitext, Springer-Verlag London, Ltd., London, 2009.
[15] W.M. Schmidt, Irregularities of distribution. IX, collection of articles in memory of Juriĭ Vladimirovič Linnik, Acta Arith. 27 (1975) 385–396.
[16] I.E. Shparlinski, Modular hyperbolas, Jpn. J. Math. 7 (2012) 235–294.

[17] I.E. Shparlinski, A.N. Skorobogatov, Exponential sums and rational points on complete intersections, Mathematika 37 (1990) 201–208.
[18] I.E. Shparlinski, A. Winterhof, On the number of distance between the coordinates of points on modular hyperbolas, J. Number Theory 128 (2008) 1224–1230.
[19] A.N. Skorobogatov, Exponential sums, the geometry of hyperplane sections, and some Diophantine problems, Israel J. Math. 80 (1992) 359–379.
[20] M.-T. Tsai, A. Zaharescu, On the action of permutations on distances between values of rational functions modulo $p$, Finite Fields Appl. 17 (2011) 481–487.
[21] W. Zhang, On the distribution of inverse modulo $n$, J. Number Theory 61 (1996) 301–310.
[22] W. Zhang, On the distribution of primitive roots modulo $p$, Publ. Math. Debrecen 53 (1998) 245–255.