

Exact Word-Map Distributions on Four-by-Four Unitriangular Groups

Alper Ferudun*

4 October 2026

Abstract

Let q be any prime power and let w be a group word in an arbitrary number of variables, without constants. We give its exact distribution on $\mathrm{UT}_4(\mathbb{F}_q)$. A nonzero exponent sum modulo the characteristic gives the uniform distribution. Otherwise the answer is determined by the rank of the quadratic coefficient matrix of the truncated noncommutative expansion of w and one bilinear zero count on its kernel. The formula includes characteristics two and three. It implies the Amit–Ashurst lower bound for every nonempty word fiber on these groups, and shows that each word image is the whole group, its derived subgroup, its center, or the identity. A separate elementary split-extension argument gives an identity-fiber bound on full unitriangular groups in every dimension. These statements concern full unitriangular groups and do not resolve the conjecture for arbitrary finite p -groups.

1 Scope and notation

For a finite group G and a word $w \in F_k$ in the free group on $x_1, \dots, x_k$, put

$$P_{w,G}(g) = \frac{|\{(g_1, \dots, g_k) \in G^k : w(g_1, \dots, g_k) = g\}|}{|G|^k}.$$

Thus the arguments are independent and uniform. We allow the trivial word and unused variables, and take $k \geq 1$. No constants from G are inserted into w .

Question 49, attributed to Amit in Abért’s list [1, p. 7], asks whether $P_{w,G}(1) \geq |G|^{-1}$ for every finite p -group. The Amit–Ashurst conjecture strengthens this to every g in the word image $w(G)$; see [4]. The qualifier $g \in w(G)$ is essential. The present note proves this stronger inequality for $G = \mathrm{UT}_4(\mathbb{F}_q)$ by an exact distribution formula. It does not claim a solution for arbitrary p -groups or their arbitrary matrix subgroups.

Write $p = \mathrm{char} \mathbb{F}_q$ and let E_{ij} be a matrix unit. In $U = \mathrm{UT}_4(\mathbb{F}_q)$ the derived subgroup and center are respectively

$$H = U' = \{g(s, t, h) := I + sE_{13} + tE_{24} + hE_{14} : s, t, h \in \mathbb{F}_q\}, \quad Z = \{g(0, 0, h) : h \in \mathbb{F}_q\}.$$

For completeness, the quotient by H is abelian, and commutators of elementary transvections give every E_{13} , E_{24} and E_{14} entry. This proves the assertion about U' . Commuting with the three adjacent elementary transvections gives the assertion about Z .

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: [AlperTheKing](https://github.com/AlperTheKing).

Expand the word in the free associative algebra truncated above degree three:

$$w(1 + X_1, \dots, 1 + X_k) = 1 + \sum_i \ell_i X_i + \sum_{i,j} C_{ij} X_i X_j + \sum_{i,j,l} D_{ijl} X_i X_j X_l \pmod{\text{degree} \geq 4}. \quad (1)$$

All coefficients lie in $\mathbb{F}_p \subseteq \mathbb{F}_q$. The expansion is computed by multiplication, using $(1 + X_i)^{-1} = 1 - X_i + X_i^2 - X_i^3$ in the truncated algebra. If e_i is the integer exponent sum of x_i , then $\ell_i = e_i \pmod p$. There is no division by two or three in this construction.

2 The exact formula

Suppose for the moment that every ℓ_i is zero. Set

$$r = \text{rank } C, \quad K = \ker C, \quad T(a, b, c) = \sum_{i,j,l} D_{ijl} a_i b_j c_l, \quad (2)$$

$$\tau = \frac{1}{|K|^2} |\{(a, c) \in K^2 : T(a, b, c) = 0 \text{ for all } b \in \mathbb{F}_q^k\}|. \quad (3)$$

This is an exact finite bilinear zero count. No polynomial-time bound for computing it is asserted.

Theorem 1. *Let q be any prime power and $w \in F_k$. If at least one ℓ_i in (1) is nonzero, then $P_{w,U}(g) = q^{-6}$ for every $g \in U$. If every ℓ_i is zero, put*

$$B = q^{-3}(1 + (q^2 - 1)q^{-r}).$$

The word takes values in H , and its distribution there is

$$P_{w,U}(g(s, t, h)) = \begin{cases} (1 - q^{-r})q^{-3}, & (s, t) \neq (0, 0), \\ B - \tau q^{-(2r+1)}, & (s, t) = (0, 0), h \neq 0, \\ B + (1 - q^{-1})\tau q^{-2r}, & s = t = h = 0. \end{cases} \quad (4)$$

The probability is zero outside H . The formula remains valid when $r = 0$, in which case the two central probabilities are $q^{-1}(1 - \tau)$ off the identity and $q^{-1} + (1 - q^{-1})\tau$ at the identity.

Lemma 2. *When $\ell = 0$, one has $C^T = -C$. In characteristic two the diagonal need not vanish. In particular $\ker C = \ker C^T$ and $\text{im } C = \text{im } C^T$.*

Proof. Abelianize (1). The resulting expression is $\prod_i (1 + X_i)^{e_i}$ in commuting variables. For $i \neq j$, its $X_i X_j$ coefficient is $e_i e_j = 0$ in $\mathbb{F}_q$, so $C_{ij} + C_{ji} = 0$. For odd p , its X_i^2 coefficient is $\binom{e_i}{2} = 0$. In characteristic two the equation $C_{ii} = -C_{ii}$ holds without imposing $C_{ii} = 0$. The binomial identities also hold for negative integer exponent sums in the truncated algebra. The kernel and image statements follow at once. $\square$

Proof of the zero-exponent case of Theorem 1. Write the six entries of the i th independent random input $I + N_i$ as

$$(N_i)_{12} = a_i, \quad (N_i)_{23} = b_i, \quad (N_i)_{34} = c_i, \quad (N_i)_{13} = d_i, \quad (N_i)_{24} = e_i, \quad (N_i)_{14} = f_i.$$

Here e_i denotes a matrix coordinate, not the integer exponent sum used above. Products of four strictly upper triangular 4 by 4 matrices vanish, so substituting these matrices into the truncated expansion gives an exact identity. Its only possibly nonzero entries above the diagonal are

$$s = a^T C b, \quad t = b^T C c, \quad h = a^T C e + d^T C c + T(a, b, c). \quad (5)$$

The f_i do not occur. In particular the output belongs to H .

Put $u = C^T a$ and $v = Cc$. By Lemma 2, these are independent uniform vectors in the same r -dimensional space $W = \text{im } C$. Unless $u = v = 0$, the expression for h is a nonzero affine linear function of the independent uniform vectors e, d . Consequently, after conditioning on a, b, c , each value of h has probability q^{-1} on this nonexceptional event.

The restriction to W of $z \mapsto b^T z$ is a uniform element of W^* , because the standard pairing on $\mathbb{F}_q^k$ is nondegenerate. This functional is zero with probability q^{-r} . When it is nonzero, its values on the independent u, v are independent uniform scalars. Hence

$$\mathbb{P}(u^T b = s, b^T v = t) = \begin{cases} (1 - q^{-r})q^{-2}, & (s, t) \neq (0, 0), \\ q^{-r} + (1 - q^{-r})q^{-2}, & (s, t) = (0, 0). \end{cases} \quad (6)$$

The exceptional event $u = v = 0$ has probability q^{-2r} and forces $s = t = 0$. Conditional on that event, a, c are independent uniform vectors in K . For each such pair, $b \mapsto T(a, b, c)$ is a linear functional. It is identically zero for a proportion τ of these pairs and otherwise is uniform on $\mathbb{F}_q$. Its probability at zero is therefore $\tau + (1 - \tau)/q$, and at any nonzero scalar is $(1 - \tau)/q$.

For $(s, t) \neq (0, 0)$ the exceptional event is absent, so multiplying the first line of (6) by q^{-1} proves the first line of (4). For $(s, t) = (0, 0)$, subtract q^{-2r} from the second line of (6), multiply by q^{-1} , and then add the exceptional contribution just computed. This gives the other two lines, including the case $r = 0$. $\square$

3 Nonempty fibers and examples

Corollary 3. *For every q , every word $w \in F_k$ and every $g \in w(U)$,*

$$P_{w,U}(g) \geq |U|^{-1} = q^{-6}.$$

If all ℓ_i vanish, the inequality is strict for every nonempty fiber. Every word image is one of $U, H, Z, \{1\}$, and the identity is always a point of maximum probability.

Proof. The nonzero-exponent case follows from uniformity, proved below. If $r \geq 1$, the first line of (4) is at least $(1 - q^{-1})q^{-3} > q^{-6}$. For the second line, let $x = q^{-r} \leq q^{-1}$. Since $\tau \leq 1$,

$$q^3 P_{w,U}(g(0, 0, h)) \geq 1 + (q^2 - 1)x - q^2 x^2 \geq 1 \quad (h \neq 0).$$

The last inequality holds because $q^2 - 1 - q^2 x \geq q^2 - q - 1 > 0$ for $q \geq 2$. The identity probability is also at least q^{-3} . Thus every element of H occurs in this case.

If $r = 0$ and $T = 0$, the word is a law. If $r = 0$ and $T \neq 0$, choose j such that the matrix $M = (D_{ijl})_{i,l}$ is nonzero. Here $K = \mathbb{F}_q^k$, and

$$1 - \tau \geq \mathbb{P}(a^T M c \neq 0) = (1 - q^{-\text{rank } M})(1 - q^{-1}) \geq (1 - q^{-1})^2. \quad (7)$$

Indeed $M c$ is nonzero with probability $1 - q^{-\text{rank } M}$, and then its pairing with uniform a is nonzero with probability $1 - q^{-1}$. All nonzero central values therefore occur with probability at least $(q - 1)^2/q^3 > q^{-6}$; the identity has probability at least q^{-1} . This proves the bound and the image classification.

Finally, in the zero-exponent case the identity probability minus a nonidentity central probability is $\tau q^{-2r} \geq 0$. Its difference from a probability with $(s, t) \neq (0, 0)$ is $q^{-r-1} + (1 - q^{-1})\tau q^{-2r} > 0$ whenever that stratum occurs. Uniformity handles the remaining case. $\square$

The diagonal issue in characteristic two is visible already for $w = x^2$. Then $C = (1)$, $r = 1$, $K = 0$ and $\tau = 1$; the distribution is supported on all of H . For $q = 2$, the identity has probability $7/16$, the nonidentity central element has probability $3/16$, and each of the other six values has probability $1/16$.

For $w = x^3$ in characteristic three, $C = 0$ and $T(a, b, c) = abc$. Thus $\tau = 2/q - 1/q^2$. At $q = 3$ the identity probability is $19/27$ and each of the two other central probabilities is $4/27$. For $w = [x, y] = xyx^{-1}y^{-1}$ in any characteristic, C has rank two and zero kernel, so $\tau = 1$. At $q = 2$, its identity probability is $1/4$, its other central probability is $3/16$, and each remaining value in H has probability $3/32$.

These examples also explain why a class-two calculation cannot simply be used in place of (5). If $x = I + E_{12}$, $y = I + E_{23}$, $z = I + E_{34}$, then $[xy, z]$ has an E_{14} contribution absent from $[x, z][y, z]$. The cubic term must be retained.

4 The split-extension calculation in every dimension

The following elementary argument finishes the uniformity case and also supplies an identity bound beyond dimension four.

Proposition 4. *Let $n \geq 2$ and $U_n = \text{UT}_n(\mathbb{F}_q)$. If an exponent sum of w is nonzero modulo p , its distribution on U_n is uniform. If all exponent sums vanish modulo p , then*

$$P_{w, U_n}(1) \geq q^{-(n-1)(n-2)/2} = |U'_n|^{-1}. \quad (8)$$

In particular the identity version of Amit's inequality holds on every full unitriangular group.

Proof. Write a matrix in blocks as (A, v) , with $A \in U_{n-1}$ and $v \in \mathbb{F}_q^{n-1}$. Multiplication and inversion are

$$(A, v)(B, z) = (AB, v + Az), \quad (A, v)^{-1} = (A^{-1}, -A^{-1}v).$$

For fixed $A_1, \dots, A_k$, the last column of the word evaluation is therefore a homogeneous linear map

$$L_A : (\mathbb{F}_q^{n-1})^k \longrightarrow \mathbb{F}_q^{n-1}.$$

Each positive occurrence of x_i contributes to its coefficient block a product of unitriangular matrices; a negative occurrence contributes the negative of such a product. The block is upper triangular and has constant diagonal ℓ_i .

If some $\ell_i \neq 0$, that block is invertible. Conditional on every choice of the top-left blocks, the output column is uniform. Induction, starting at the additive group U_2 , shows that every matrix output has probability $q^{-(n-1)}|U_{n-1}|^{-1} = |U_n|^{-1}$.

If all $\ell_i = 0$, every coefficient block is strictly upper triangular, so $\text{rank } L_A \leq n-2$. Above every top-left tuple whose word value is the identity, the zero-column fiber has probability $q^{-\text{rank } L_A} \geq q^{-(n-2)}$. There is no affine shift: the displayed block decomposition is a split extension. At $n = 2$ the word is a law. Multiplying the conditional lower bounds for $n, n-1, \dots, 3$ gives (8). The order of U'_n follows either from the elementary commutators or by counting the entries beyond the first superdiagonal. $\square$

Remark 5. The split hypothesis cannot be silently dropped. In the nonsplit extension $C_4 \rightarrow C_2$, written additively, the nonzero quotient element satisfies the equation $2x = 0$, but both lifts $1, 3 \in C_4$ satisfy $2x = 2 \neq 0$. Thus a quotient solution may have no identity lift. This is not a counterexample to Amit's inequality on C_4 . Similarly, representing a finite p -group inside U_n does not transfer our distribution formulas: uniform independent matrix coordinates are lost after restriction to a subgroup.

5 Relation to previous work and verification

Camina, Iñiguez and Thillaisundaram [3] give general-fiber bounds for class-two groups, including the Amit–Ashurst bound for two-variable words. Kishnani and Kulshrestha [5] obtain automorphic word normal forms and further bounds in class two. Their class-two hypothesis is not satisfied by U . The theorem of Camina, Cocke and Thillaisundaram [4] concerns finite p -groups with a cyclic maximal subgroup. The recent work of de las Heras, Toti and Vannacci [7] concerns Engel words in specified wreath products. Its kernel-and-image calculations are relevant to the conditional-linear viewpoint, but its cyclic-module hypotheses are different from those here. None of those theorems is invoked as an all-word theorem for U .

Bier [2] studies verbal subgroups and verbal width of unitriangular groups. The exact probabilities in (4) are a different datum from the subgroup generated by word values or the number of such values needed in a product. We do not claim a new general theory of verbal subgroups. Nor is the elementary split-extension observation of Proposition 4 claimed to be new. A bounded primary-literature search did not identify the specific rank-and-kernel formula (4); this is not a certification of novelty or absolute priority.

The accompanying Python checker compares the formula with direct matrix evaluation on 6,071,366 full input tuples across 35 word–field cases, including $\mathbb{F}_4$ and $\mathbb{F}_8$. It also performs 34,500 pointwise exact expansion checks and four larger reduced enumerations, explicitly distinguished from full group enumeration. A separately written JavaScript checker uses dense 4 by 4 matrices and verifies 13,081 full tuples in five cases. It also checks the class-two commutator defect and the nonsplit lifting obstruction above. These finite tests are regression checks, not proofs of the arbitrary-word or arbitrary-field quantifiers. The proofs of those assertions are the arguments in this note. The caution in [6, Section 4.1] about fixed-arity enumeration is particularly pertinent here.

This unrefereed note was developed with AI assistance in derivation, source checking, exact computation and drafting. The argument was self-audited; no independent specialist review or proof-assistant verification is claimed. The full conjecture in Abért’s Question 49, recorded as AMR-011-0049 in UnsolvedMath v1.6.0, remains unresolved by this work.

References

- [1] M. Abért, *Some questions*, 2 November 2010, Question 49. <https://www.renyi.hu/~abert/questions.pdf>.
- [2] A. Bier, The width of verbal subgroups in the group of unitriangular matrices over a field, *Internat. J. Algebra Comput.* **22** (2012), no. 3, 1250019. doi:10.1142/S0218196712500191.
- [3] R. D. Camina, A. Iñiguez and A. Thillaisundaram, Word problems for finite nilpotent groups, *Arch. Math.* **115** (2020), 599–609. doi:10.1007/s00013-020-01504-w.
- [4] R. D. Camina, W. L. Cocke and A. Thillaisundaram, *The Amit–Ashurst conjecture for finite metacyclic p -groups*, accepted manuscript, arXiv:2201.04860v3 (2023). <https://arxiv.org/abs/2201.04860v3>.
- [5] H. Kishnani and A. Kulshrestha, *Automorphic word maps and Amit–Ashurst conjecture*, arXiv:2309.09010v1 (2023). <https://arxiv.org/abs/2309.09010v1>.
- [6] B. S. Chlebus, W. Cocke and M.-C. Ho, Enumerating word maps in finite groups, *Int. J. Group Theory* **13** (2024), no. 3, 307–318. doi:10.22108/ijgt.2023.136972.1833.
- [7] I. de las Heras, T. Toti and M. Vannacci, *Engel probability in wreath products of p -groups*, arXiv:2607.11605v1 (2026). <https://arxiv.org/abs/2607.11605v1>.