Kloosterman shrinking targets and cyclotomic trace geometry

Scott Schubr
Luminary Oddities

November 2026

Abstract

We determine how often Kloosterman and hyper-Kloosterman paths enter small balls around prescribed centers. For every fixed completed rank $R \geq 2$, supported half-Hölder centers have arithmetic path-ball probability $\exp(-\Theta(\varepsilon^{-2}))$ down to radius of order $(\log P)^{-1/2}$, for complete primes, long progressions, most translates of arbitrary growing parameter sets and all prime-power depths, with L^1 upper bounds for every center and an optimal uniform coverage radius. At fixed primes the residue-conditioned counts start at $J^{-1/2}$ with $J = \min(n, \sqrt{n \log p})$, and weighted residue masks give exact deep covering radii with logarithmic rates. For the traces themselves we determine the geometry of cyclotomic unit lattices: sharp Wasserstein and disk orders $q^{-1/\varphi(d)}$ and $q^{-2/\varphi(d)}$, effective arcosh depth laws with discrepancy bounds over prime windows, boundary-scale disk masses for odd rank, fixed-prime boundary exponents with exact exceptional dimensions, and microscopic counts at mixed cyclotomic corners including the exact law of the first count above threshold.

Contents

1 Introduction **3**

2 Kloosterman shrinking targets - revised **4**
2.1 Statements . . . 4
2.2 Fourier representation and the comparison budget . . . 5
2.3 Limiting uniform balls . . . 6
2.4 Finite-norm lower bounds . . . 7
2.5 Uniform-norm lower bounds . . . 8
2.6 Removing the gate at large depth . . . 10
2.7 An upper bound in L^1 . . . 10
2.8 Finite support and the radius . . . 12
.1 Prime character comparison . . . 13
.2 Prime-power cosine comparison . . . 14
.3 Pinned Fourier tails . . . 17

A Exact uniform prime Kloosterman shrinking targets **17**
A.1 Statements . . . 18
A.2 Fourier and arithmetic estimates . . . 19
A.3 Localized tests of controlled energy . . . 19
A.3.1 The strongly log-concave cutoff estimate . . . 19
A.3.2 Sample rows and the even-rank test . . . 20

A.3.3 Dense localization in odd rank . . . 21
A.3.4 High-band energy from trajectory discrepancies . . . 22
A.4 Finite propagation and the signed polynomial . . . 24
A.5 Transfer of the lower bound . . . 25
A.6 An upper bound in L^1 . . . 26
A.7 Sharpness for uniform coverage . . . 27
.1 Additive twists and character norms . . . 28

A Exact Kloosterman shrinking targets at prime powers 29
A.1 Statements . . . 30
A.2 Completion, angle modes and comparison . . . 31
A.3 A killed-heat construction on the angle space . . . 32
A.3.1 A dense torus test by coupled Brownian packets . . . 33
A.3.2 Trajectory discrepancy and high bands . . . 34
A.3.3 The complete label law . . . 35
A.4 Wave localization and transfer . . . 37
A.5 The upper bound in L^1 . . . 39
A.6 Coverage and the missing progression . . . 40
A.6.1 An obstruction from opposite trace labels . . . 42
.1 Last-digit and depth comparisons . . . 43

A Higher-rank Kloosterman shrinking targets 44
A.1 Path families and estimates . . . 45
A.2 Trace image and interior Haar mass . . . 46
A.3 Small balls for the independent Haar trace path . . . 47
A.4 The arithmetic comparison window . . . 48
A.5 Reversible heat observations . . . 49
A.6 Exact-mean tilts in finite norms . . . 50
A.7 An upper bound in L^1 . . . 51
.1 Additive twists and character norms . . . 53

A Weighted residue masks for Kloosterman paths 54
A.1 Statements . . . 55
A.2 The exact covering radius . . . 56
A.3 The complete Laplace product . . . 57
A.4 Logarithmic rate and prefactor . . . 58
A.5 Long residue runs at moderate depths . . . 60

B Boundary-scale hyper-Kloosterman traces 61
B.1 Statements . . . 61
B.2 Dimension-weighted comparisons . . . 62
B.3 Uniform Haar mass at a moving boundary . . . 63
B.4 Character costs at the boundary . . . 65

C Cyclotomic lattice transport at prime powers 67
C.1 Phase lattices and trace disks . . . 67
C.2 The lattice coupling . . . 68
C.3 Aligned tips and exact lattice shells . . . 69

C.3.1 Exact counts for $d = 2, 3, 4, 6$ . . . 69
C.3.2 Microscopic counts as depth varies . . . 70
C.3.3 An explicit critical-depth law at order five . . . 72
C.4 Moving along the boundary at prime order . . . 72

D Effective quartic cyclotomic trace laws **74**
D.1 Trace laws and a unit coordinate . . . 75
D.2 The critical ideal energy . . . 76
D.3 Effective depth averaging . . . 77
D.4 Prime averaging with a growing common depth . . . 78

E Smooth boundary trace disks at fixed prime **79**
E.1 Boundary counts and exceptional targets . . . 79
E.2 Rectangles in the cyclotomic angle lattice . . . 80
E.3 The boundary chart and the disk scale . . . 82
E.4 Dimension of the exceptional set . . . 83
E.5 A fixed set of split primes . . . 84

F Composite cyclotomic boundary arcs and corners **84**
F.1 The boundary of the sum . . . 84
F.2 Trace counts along a smooth arc . . . 86
F.3 The law at a mixed corner . . . 88
F.4 Rectangular comparison and exclusion . . . 88

G Microscopic trace counts at mixed cyclotomic corners **90**
G.1 Angular coordinates and the local quadratic form . . . 90
G.2 The torus and the fixed-prime law . . . 91
G.3 Averaging over split primes . . . 92
G.4 An unfolded mean . . . 94
G.5 The angular projection at fixed prime . . . 95
G.6 Full torus laws with several fixed primes . . . 95
G.7 Exact first radii . . . 97

H The first count above a cyclotomic corner threshold **98**
H.1 The onset law . . . 98
H.2 Norm-one candidates and exact multiplicity . . . 99
H.3 The volume just above the minimum . . . 100
H.4 Depth dependence from the local unit quotient . . . 102
H.5 A fixed set of split primes . . . 103

1 Introduction

The shrinking-target question asks how the probability that a random path lies within ε of a fixed center decays, and down to which radius the arithmetic family follows the limiting process. Small-ball estimates for the Kloosterman process were obtained in [1]. Section 2 proves the relative small-ball scale for Kloosterman paths in every finite L^r norm, and Sections A–A make it uniform in the prime, extend it to prime powers and higher ranks, and determine weighted residue masks.

Sections B–H study the trace geometry at prime powers: boundary-scale disks, cyclotomic lattices, effective depth laws, boundary exponents, composite corners and microscopic corner counts.

2 Kloosterman shrinking targets - revised

We count prime and prime-power Kloosterman paths in shrinking balls about a class of Hölder centers. For each fixed finite L^r norm, the negative logarithm of the ball probability has order ϵ^{-2} down to a constant times $(\log p)^{-1/2}$ as the base prime p grows, uniformly in the fixed inverse coefficient and the prime-power depth. For prime moduli this also holds on most translates of any growing parameter set, with its size in place of p. The uniform-norm lower bound reaches radius of order $(\log\log p/\log p)^{1/2}$, with the analogous replacement for translates. For depths $n \geq C\log p(\log\log p)^2$, the arithmetic uniform-norm lower bound also reaches the radius $(\log p)^{-1/2}$. At fixed prime in a residue class, the upper and finite-norm lower bounds reach order $n^{-1/4}$ in depth n, and the uniform lower bound reaches $n^{-1/4}\sqrt{\log n}$. The fixed-prime lower bounds require centers whose coefficients respect the square gate. The proofs use polynomial tilts, beta posterior moments, and an L^1 upper bound.

Keywords Kloosterman paths, shrinking targets, polynomial tilting, discrepancy, arcsine law.
Mathematics Subject Classification 11L05, 11T23, 60G17, 60G50.

2.1 Statements

Write $e(t) = \exp(2\pi i t)$ and $e_Q(t) = e(t/Q)$. For $Q = p^n$ with p odd, $K_Q(A, B)$ is the polygon with successive increments $Q^{-1/2}e_Q(Ax + B/x)$ at the increasing representatives x of the units modulo Q, with equal time steps on $[0, 1]$. Keep B a unit. For $n = 1$ let A range uniformly over $\mathbb{F}_p$ or over a translate $S + u$ of a deterministic L-point set S. For $n \geq 2$ let A be uniform on the units. At fixed p we also condition on a fixed nonzero residue a_0 of A modulo p and allow any lift B of a fixed nonzero residue b_0. The square gate at index h in this last family is

$$I_h = \mathbf{1}_{\{b_0(a_0-h) \text{ is a nonzero square in } \mathbb{F}_p\}}.$$

For $H > 0$, let $\mathcal{H}_H$ be the class of paths g starting at zero with half-Hölder seminorm at most H, endpoint $g(1) \in [-2, 2]$, and normalized Fourier coefficients

$$\beta_h(g) = 2\pi i h\,\widehat{g - tg(1)}(h) \in [-2, 2] \qquad (h \neq 0).$$

For a gated family use the subclass with $\beta_h(g) = 0$ at every inactive $h \neq 0$ and $g(1) = 0$ if $I_0 = 0$. We denote the norm on $[0, 1]$ by $\|\cdot\|_r$ for $1 \leq r \leq \infty$.

Theorem 2.1. *In the full prime and prime-power families write $P = p$, and for the translated prime family write $P = L$. For each fixed $1 \leq r < \infty$ and $H > 0$ there are positive constants such that, uniformly for sufficiently large P, all admissible centers, and $D_{H,r}/\sqrt{\log P} \leq \epsilon \leq \epsilon_*$,*

$$e^{-C_{H,r}\epsilon^{-2}} \leq \mathbb{P}\{\|K_Q - g\|_r < \epsilon\} \leq e^{-c\epsilon^{-2}}.$$

The upper bound holds for all centers in L^1 with its L^1 norm and for all $\epsilon \geq d/\sqrt{\log P}$ in the stated small-radius range. For $r = \infty$ the lower bound holds for $\epsilon \geq D_H\sqrt{\log\log P/\log P}$. All bounds are uniform in B and in $n \geq 2$ when the base prime grows. For any fixed B, S, the assertions about translates hold outside a common $O_{H,r}(L^{-c'})$ fraction of translates, independent of the centers and radii.

Theorem 2.2. *For each fixed odd p, families with $A \equiv a_0 \pmod p$ and $B \equiv b_0 \pmod p$ satisfy the finite-norm inequalities of Theorem 2.1 down to $D_{p,H,r}n^{-1/4}$, for gated centers in the lower bound. The L^1 upper bound for all centers holds down to $d_p n^{-1/4}$. In the uniform norm the lower bound holds for $\epsilon \geq D_{p,H}n^{-1/4}\sqrt{\log n}$. The constants are uniform in the residue labels and lifts at the fixed prime.*

The upper and lower radii differ in the uniform norm.

Theorem 2.3. *Put $J = \min(n, \sqrt{n \log p})$. Uniformly in odd p, residue labels and lifts, the residue-conditioned lower bounds hold for compatible centers at $\epsilon \geq D_{H,r}/\sqrt{J}$ in each fixed finite L^r, and at $\epsilon \geq D_H\sqrt{\log(2J)/J}$ in the uniform norm, for sufficiently large J. Both are of the form $\exp(-C_{H,r}\epsilon^{-2})$ in the same small-radius range as above.*

For arbitrary admissible centers and complete averaging over unit A, the uniform-norm lower bound $\exp(-C_H\epsilon^{-2})$ holds from

$$D_H \max\big((\log p)^{-1/2}, (\log(2J)/J)^{1/2}\big)$$

for sufficiently large p and J. In particular if $n \geq C \log p(\log\log p)^2$, the uniform upper and lower bounds both hold at radius $D_H/\sqrt{\log p}$, uniformly in n and the fixed unit B.

Corollary 2.4. *Fix the completed rank $R \geq 2$ of the twisted hyper-Kloosterman family in Appendix .1, a nonzero inverse-scale parameter a, and an arbitrary L-point set $S \subseteq \mathbb{F}_p$. For sufficiently large L, outside an $O_R(L^{-c_R})$ fraction of translates, the L^1 upper bound $\exp(-c'_R\epsilon^{-2})$ holds simultaneously for all centers and $d_R/\sqrt{\log L} \leq \epsilon \leq \epsilon_*$.*

Kowalski and Sawin established the prime limiting process [7] and characterized its support [8, Theorem 1.1]. For prime powers, Ricotta and Royer [10] proved fixed-parameter finite-dimensional convergence for each fixed depth $n \geq 2$ and two-parameter functional convergence. Ricotta, Royer and Shparlinski [11] proved fixed-parameter functional convergence for fixed $n \geq 31$. Milićević and Zhang [12] treated the fixed-prime depth limit and its square-gated law, and Milićević, Robinson and Shupe [14] refined product estimates using the alignment of critical points.

The Archive edition [2] gives the sharp limiting Sato–Tate exponent and a direct two-parameter arithmetic uniform-norm lower bound from radius of order $(\log p)^{-1/4}$. Kyselov [3, Theorem 1.2 and Corollary 5.1] gives a limiting lower bound with a squared logarithmic loss for closed Hölder loops with normalized coefficients in $[-1, 1]$, and a fixed-radius arithmetic liminf for his constructed loop. The finite-norm lower bound below uses the mean of an exactly centered polynomial tilt. The uniform lower bound uses conditional Fourier moments under a beta posterior and needs a lower count in the limiting process. We prove that lower count for the square-gated and mixed arcsine processes.

2.2 Fourier representation and the comparison budget

Extend the summand by zero at nonunits and use the mesh j/Q, with mesh path $Q_*(t)$, $0 \leq t \leq 1$. Then $\|K_Q - Q_*\|_\infty \ll Q^{-1/2}$, and at mesh vertices

$$Q_*(j/Q) = (j/Q)W_0 + \sum_{0<|h|\leq(Q-1)/2} b_{Q,h}W_h(e(hj/Q) - 1), \qquad b_{Q,h} = \frac{1}{Q(1 - e(-h/Q))}. \quad (1)$$

Here W_h is the complete normalized sum with linear coefficient $A - h$. The formula follows by summing its discrete Fourier expansion. The time comparison follows because the jth unit has

position $j + \lfloor (j-1)/(p-1) \rfloor$, which divided by Q differs from $j/\varphi(Q)$ by $O(Q^{-1})$. The mesh path is $\sqrt{Q}$-Lipschitz. For $0 < |h| < Q/2$, $|b_{Q,h}| \ll 1/|h|$ and $b_{Q,h} = (2\pi i h)^{-1} + O(Q^{-1})$. When using the actual Fourier coefficient of the pinned polygon, integration of the piecewise constant derivative gives

$$2\pi i h \, \widehat{Q_* - tW}_0(h) = c_{Q,h} W_h, \qquad c_{Q,h} = e(h/(2Q)) \operatorname{sinc}(\pi h/Q), \quad \operatorname{sinc}(x) = \sin(x)/x. \tag{2}$$

Let R_N denote the polynomial in (1) through $|h| \le N$. We use the following arithmetic inputs, proved in Appendices .1–.3. Complete primes compare with independent Sato–Tate traces, all depths at large base prime compare with independent traces of law $(\delta_0 + \sigma)/2$, where σ is the arcsine law on $[-2, 2]$, and fixed-prime residue families compare with independent traces of law σ at active indices and zero elsewhere. In the large-base mixed model the fair square gates are included in the comparison.

For large base prime or translated sets use $N = \lfloor P^{1/8} \rfloor$, $M \le a \log P$ for sufficiently small a. For the depth aspect take $N = \lfloor e^J \rfloor$ and $M \le aJ$, with J as in Theorem 2.3. All tests in the proof below are polynomials with separate degrees $O(M^2 + s)$ for a moment order $2s$, with at most $O(M + s)$ active shifts in any character or cosine term. Their weighted character or cosine norm is $\exp(O(M))$ times at most $\exp(o(\log P))$ or $\exp(o(J))$, respectively. We use s fixed or $s = O(\log M)$. The discrepancy is $O(P^{-\gamma})$ for some $\gamma > 0$ in the large-base case, in L^2 over u for translates, after taking a small. It is $O(\exp(-CJ))$ in depth aspect, uniformly in odd p, where C can be made large by reducing a. The large-base terminal tail satisfies $\mathbb{E} \, \|Q_* - R_N\|_\infty = O(P^{-\gamma})$ for complete ensembles and in mean over translates. The depth-aspect tail is $O(\exp(-\gamma J))$. Thus one may divide the errors by a normalization v of size at least $\exp(-C_1 M)$ by taking a small enough. This assertion for translates uses Markov's inequality, which gives a common good set for the terminal estimate independent of center and radius.

2.3 Limiting uniform balls

We use the following quantitative signing bound. Spencer's theorem also gives exponentially many colorings at a fixed constant times $\sqrt{N}$ (see [5, Section 1] for a discussion). We need a bound that improves when the allowed discrepancy increases.

Lemma 2.5. *Let D be a real matrix with N columns, at most $C_0 N$ rows, and entries bounded by one. For independent uniform signs ξ and $\lambda \ge 1$,*

$$\mathbb{P}\{\|D\xi\|_\infty \le C\lambda\sqrt{N}\} \ge \exp\{-CN \exp(-c\lambda^2)\},$$

where $C, c > 0$ depend only on C_0.

Proof. A row of $D\xi/\sqrt{N}$ has subgaussian tails. Round each normalized row to the nearest multiple of λ. The resulting integer Y has $\mathbb{E}|Y| \le Ce^{-c_1\lambda^2}$. Testing its entropy against the probability mass proportional to $\exp(-c_1\lambda^2|j|)$ gives $\mathrm{Ent}(Y) \le Ce^{-c_2\lambda^2}$. By subadditivity there is a common rounding fiber E of cube measure at least e^{-b}, where $b = CNe^{-c_2\lambda^2}$.

Hamming distance from ξ to $-E$ is one-Lipschitz in each of the N coordinates. The bounded-difference lower tail and the probability it is zero give a mean at most $\sqrt{Nb/2}$. Its upper tail is at most $e^{-b}/2$ at distance $d = \lceil C\sqrt{N(b+1)} \rceil$ for large enough C. Thus at least half of $x \in E$ are at distance at most d from $-E$. For each such x, take $y \in E$ with $\mathrm{dist}(x, -y) \le d$. The partial signing $(x - y)/2$ has at most d zeros and discrepancy at most $\lambda\sqrt{N}$. Complete the zeros using Spencer's real bounded-entry matrix theorem [4], padding to a square of side $O(N)$. The added discrepancy

is $O(\sqrt{N})$. Each resulting signing lies within distance d of x. The fraction of resulting signings is at least

$$\frac{e^{-b}}{2\sum_{j\le d}\binom{N}{j}}.$$

For λ larger than a sufficiently large multiple of $\sqrt{\log(2N)}$, the row tail union bound already proves the result. Below that threshold weaken c_2 so $b \ge 1$. For λ larger than a fixed constant, $d/N \ll e^{-c_2\lambda^2/2}$ and the logarithm of the binomial sum is $O(N\sqrt{d/N})$. For the remaining bounded λ a single Spencer signing suffices after enlarging C. These bounds prove the claim. □

Lemma 2.6. *For the independent Sato–Tate, mixed-arcsine or fixed-prime gated Fourier process,*

$$X_0 t + \sum_{h\neq 0} \frac{X_h}{2\pi i h}(e(ht) - 1),$$

the uniform ball of sufficiently small radius ϵ about an admissible center has probability at least $\exp(-C_H\epsilon^{-2})$. In the gated case the center has the corresponding inactive zeros.

Proof. For a dyadic band of size U, condition on all magnitudes and write each symmetric trace as magnitude times an independent sign. The real and imaginary parts on a Bernstein mesh are at most CU rows with multipliers $O(U^{-1})$. Lemma 2.5, followed by the Bernstein mesh estimate and pinning at zero, gives band norm at most $C\lambda U^{-1/2}$ with probability at least $\exp(-CUe^{-c\lambda^2})$. Above frequency M, take $U = 2^j M$ and $\lambda = 2^{j/4}$. Radii sum to $O(M^{-1/2})$ and the negative logarithms of probabilities sum to $O(M)$. The bands are independent. Uniform convergence of the independent series follows directly from the pointwise subgaussian bound and Bernstein meshes, whose expected block maxima are $O(\sqrt{\log(2U)/U})$.

For $0 < |h| \le M$ use Fejér means $c_h = (1-|h|/(M+1))\beta_h(g)$, and at zero $c_0 = (1-M^{-1/2})g(1)$. The resulting pinned polynomial is within $O_H(M^{-1/2})$ of g, by the half-Hölder estimate for the Fejér kernel. At active indices let the residual be the difference of two independent uniforms of width $c|h|/M$ about zero, shifted by c_h, and use width $cM^{-1/2}$ for the endpoint. It is supported inside $[-2,2]$. For these widths w_h the trace law dominates this triangular probability multiplied by $cw_h^{3/2}$, using the Fejér margin in the Sato–Tate case, the lower arcsine density, and its half weight in the mixed case. Inactive coordinates are zero. The product cost is at most CM in negative logarithms. The two one-copy centered low polynomials each have pointwise variance proxy $O(M^{-1})$. Quantize their real and imaginary parts on a Bernstein mesh of size $O(M)$ at spacing $M^{-1/2}$. The vector entropy is $O(M)$ by the same individual tail argument as in Lemma 2.5. Two independent copies agree with probability at least $\exp(-O(M))$ by Jensen's inequality applied to their common atom probabilities. On agreement their difference has pinned uniform norm $O(M^{-1/2})$. The endpoint residual is always $O(M^{-1/2})$. Taking M a sufficiently large constant times ϵ^{-2} finishes the proof. □

2.4 Finite-norm lower bounds

Let λ be Sato–Tate or arcsine on $[-2,2]$. With $a \in [-2,2]$ and integer $D \ge 1$, define

$$f_{D,a}(x) = (1-(x-a)^2/32)^D.$$

Both laws are uniformly doubling on their closed support. Indeed the mass at a point a of radius $0 < r \le 1$ is comparable to $r(2-a+r)^\beta(2+a+r)^\beta$, with $\beta = 1/2$ for Sato–Tate and $\beta = -1/2$ for arcsine. This follows by integration near either boundary and direct comparison in the interior. In particular the mass is at least $cr^{3/2}$. Set $\delta = D^{-1/2}$. On $|x-a| \le \delta$ the polynomial has an

absolute positive lower bound, giving $\mathbb{E}_\lambda f_{D,a} \geq cD^{-3/4}$. On the whole interval it is bounded by $\exp(-cD|x-a|^2)$. Use dyadic shells about a and doubling to find that under the tilted law with density $f_{D,a}/\mathbb{E}_\lambda f_{D,a}$, the tails of $|X-a|$ are at most $C\exp(-cDt^2)$. Its mean lies within $C\delta$ of a and varies continuously with a. By the intermediate value theorem every c in $[-2+C\delta, 2-C\delta]$ is a tilted mean. The centered variable has subgaussian proxy $O(1/D)$, by Jensen with an independent copy and the exponential series for the symmetric difference.

Use the coefficients c_h of the Fejér approximant from Lemma 2.6, with $D_h = \lceil CM^2/h^2 \rceil$ for $0 < |h| \leq M$ and $D_0 = \lceil CM \rceil$. For sufficiently large absolute C, the Fejér margins allow the choice of a_h giving mean c_h at each active slot. Form the weight F as their product over $|h| \leq M$. In the large-base mixed model require in addition that all these square gates be active. In residue-conditioned ensembles omit inactive slots, which vanish in both arithmetic and model. The normalized tilted model has independent centered residuals with subgaussian proxies $O(h^2/M^2)$ for $0 < |h| \leq M$, $O(1/M)$ at zero and $O(1)$ above M. Its weight has mean $v \geq e^{-C_1 M}$ because $\sum_{|h|\leq M} \log(D_h+2) = O(M)$, including the factor $2^{-(2M+1)}$ for prescribed mixed gates.

At each time the residual of R_N from its tilted mean has proxy $O(1/M)$, by $|b_{Q,h}| \ll 1/|h|$. Its integrated $2s$th moment is $O((Cs/M)^s)$. Its mean differs from the Fejér approximant by $O(M/Q)$. The angular lift for each scalar factor uses separate degree $2D_h$, cube bound one and weighted character norm $O((1+D_h)^2)$. Its product has norm $e^{O(M)}$. Multiplying by the residual power and integrating costs at most $(C\log(2N))^{2s}$ by submultiplicativity and the bound on the coefficient sum. Each term uses at most $2M+1+2s$ active shifts and has separate degree $O(M^2+s)$. Apply Section 2.2 to this test and to F, taking fixed $2s \geq r$ and $M = \lfloor K_{H,r}\epsilon^{-2} \rfloor$. For complete families the arithmetic weight mean v_a satisfies $v_a \geq v/2$, and under weight F/v_a the arithmetic integrated moment is at most $2(Cs/M)^s$ plus an error smaller than any fixed power of M^{-1}. The same holds off a small exceptional set for translates, using Chebyshev for the L^2 comparison and then dividing by v_a. The terminal bound of Section 2.2 after division by v_a still gives $o(\epsilon)$ expected error. In particular it includes the value at zero subtracted in pinning. The mesh error is negligible. Take $K_{H,r}$ large and apply Markov under the tilt. At least half the tilted paths lie in the open ball. Since $F \leq 1$, the arithmetic ball has probability at least $v_a/2 \geq e^{-C_{H,r}\epsilon^{-2}}$ after adjusting the exponent constant. The choice of $D_{H,r}$ makes M small enough for the comparison window, including $M \leq aJ$ in the joint aspect. The same proof gives the residue-conditioned finite-norm lower in Theorem 2.3.

2.5 Uniform-norm lower bounds

Here M is the full allowed window, namely $\lfloor a\log P \rfloor$ or $\lfloor aJ \rfloor$. Choose an even smooth cutoff γ supported on $[-1,1]$, with $0 \leq \gamma \leq 1$, equal to $(1+3x^2)^{-1}$ for $|x| \leq 1/2$ and strictly below one at $x \neq 0$. Write $\ell = 3$ for Sato–Tate and $\ell = 1$ for arcsine. At active coordinates $0 < |h| \leq M$, let $\nu_h = \ell\gamma(h/M)/(1-\gamma(h/M))$ and draw an integer D_h between the two neighbors of ν_h, independently of the input, with $\mathbb{E}[D_h/(D_h+\ell)] = \nu_h/(\nu_h+\ell)$. At zero use $D_0 = M$. Conditional on an input x_h and degree D_h, draw $k_h \sim \mathrm{Bin}(D_h, (x_h+2)/4)$ and set $z_h = (4k_h - 2D_h)/(D_h+\ell)$. In the mixed model expose the square gate as part of the observation. For inactive indices and outside the window put $z_h = 0$. In arithmetic use the same conditional kernel. Write Y for all observations.

In the model, active x_h corresponds to a $\mathrm{Beta}(\ell/2, \ell/2)$ variable after rescaling to $[0,1]$. Its posterior given Y is $\mathrm{Beta}(k_h + \ell/2, D_h - k_h + \ell/2)$, with trace mean z_h. Coordinates remain independent. The centered posterior has subgaussian proxy $O((D_h+1)^{-1})$ uniformly. For the Sato–Tate posterior, and the arcsine posterior when $1 \leq k_h \leq D_h - 1$, the negative log-density has curvature at least $c(D_h+1)$ on the trace interval. Under every exponential tilt it has the same curvature. Integration by parts gives $\mathrm{Cov}(U, V'(U)) = 1$ for the tilted variable and its tilted negative log-density V. By comparing independent copies the curvature gives $\mathrm{Var}(U) \leq$

$1/(c(D_h+1))$. Integrating this bound for the second derivative of the cumulant gives the proxy. The boundary arcsine posterior is, up to reflection, $\mathrm{Beta}(1/2, D_h+1/2)$. Its upper tail at $t \in [0,1]$ is $O(e^{-cD_h t})$, by bounding its density by a constant times $\sqrt{D_h}x^{-1/2}e^{-cD_h x}$ and integrating, with bounded degrees handled directly. This also gives Gaussian tails on scale $D_h^{-1/2}$ and the centered proxy by symmetrization.

Conversely $\mathbb{E}[z_h \mid x_h] = \gamma(h/M)x_h$ for nonzero h in the window, and $\mathbb{E}[z_0 \mid x_0] = Mx_0/(M+\ell)$. The conditional fluctuation is subgaussian with proxy $O((1+\nu_h)^{-1})$ by Hoeffding first for the binomial and then for the two possible conditional means. At zero its proxy is $O(1/M)$. These statements include inactive slots. For

$$T(Y)(t) = tz_0 + \sum_{0<|h|\le M} b_{Q,h} z_h (e(ht) - 1)$$

the expected conditional supremum fluctuation is $O(\sqrt{\log(2M)/M})$ uniformly in the input: each time has variance proxy $O(1/M)$ and a Bernstein mesh of $O(M)$ points bounds the polynomial maximum. Its conditional mean in the independent model is the smooth cutoff of the limit path up to $O(M/Q + 1/M)$. The cutoff on the pinned periodic part has uniformly bounded L^1 kernel and absolute circular half-moment $O(M^{-1/2})$. In fact twice summing by parts bounds the kernel by $CM/(1+M^2\operatorname{dist}(t,\mathbb{Z})^2)$. Thus it approximates every admissible center to $O_H(M^{-1/2})$, accounting for the affine part separately.

Put $F(Y) = \mathbf{1}_{\{\|T(Y)-g\|_\infty < \epsilon/2\}}$. When the independent limiting path is within a small constant multiple of ϵ of g, its conditional mean reference is within $\epsilon/4$. Markov and the noise bound imply $v = \mathbb{E}F \ge e^{-C_H \epsilon^{-2}}$ by Lemma 2.6, for ϵ a sufficiently large constant times $\sqrt{\log(2M)/M}$. In particular $v \ge e^{-o(M)}$ uniformly.

For a bounded function of Y, the conditional expectation given the input and any prescribed gates is a polynomial bounded on the cube by the function's supremum. Its separate degrees are at most $d_0 = M$, $d_h = \lceil \nu_h \rceil$ for $0 < |h| \le M$, and $d_h = 0$ elsewhere. The separate-degree norm bound gives $e^{O(M)}$ for a cube bound one, since $\sum_{|h|\le M} \log(2+d_h) = O(M)$. Gate masks cost a further factor 2^{2M+1} in the mixed family. Inserting a product of $2s$ residuals $x_h - z_h$ increases the sum of these degree bounds by at most $2s$, and the cube bound by at most 4^{2s}. The separate-degree product therefore increases by at most C^{2s} apart from changing the absolute constant C. In particular $v_a = \mathbb{E}_a F \ge v/2$, using Chebyshev and discarding a power-saving fraction for each test in the translated case.

Take $s = \lceil \log(16M) \rceil$. In the unpinned residual of each high band of size U between M and N, integrate the F-weighted $2s$th moment over time. Only tuples with the first frequency sum equal to the second remain, of total absolute coefficient weight at most C^{2s}/U, since fixing all but the last index determines the last. In the low band use instead $(C\log(2M))^{2s}$. The supremum of a band to power $2s$ is at most $(2sU+1)$ times this integrated power, by applying Parseval and Cauchy–Schwarz to the sth power. The relative arithmetic comparison errors after taking roots and summing are negligible. More explicitly, the factor $1/U$ cancels the U in the supremum estimate, so the error roots are at most a polylogarithmic factor times $P^{-\gamma/(Cs)}$ or $\exp(-cJ/s)$ per band, including normalization by v_a. In the model each conditional high band has pointwise variance proxy $O(1/U)$, while the low band and endpoint have proxy $O(1/M)$. Their supremum moment roots after comparison have model terms bounded by $(2sU+1)^{1/(2s)} C\sqrt{s/U}$ in a high band, the analogous bound with $U = M$ at low frequencies, and $C\sqrt{s/M}$ for the endpoint. The prefactor from normalizing is bounded by $v/v_a \le 2$ before taking roots. Summation, pinning and the terminal estimate yield

$$\mathbb{E}_a[\|K_Q - T(Y)\|_\infty \mid F = 1] \ll \sqrt{s/M}.$$

For translates, Chebyshev can be applied to each integrated discrepancy after dividing out its stated coefficient weight. A union over the $O(\log L)$ bands still leaves a power-saving exceptional fraction for each center and radius. Increase the radius constant and apply conditional Markov; the resulting lower bound has the claimed mass.

A common good set of translates. The terminal good set is independent of the center and radius. For the polynomial tests, first use geometric test radii and nets chosen from the admissible class at each test radius ϵ with supremum mesh a small constant times ϵ. The log cardinality of such a net is $O_H(\epsilon^{-2})$: round values on a time grid of spacing $c_H\epsilon^2$, and count codes using that the initial value is zero and each successive value has only a bounded number of possible rounded increments. A maximal separated set injects into these codes for suitable mesh constants. In the stated lower ranges, taking the minimum radius constant large makes the net cardinality a sufficiently small power of L for the union bound. After the union, increase the minimum constant once more to pass from the geometric test radii and net centers to arbitrary balls by inclusion. The same reasoning applies to the finite-r centers, since a supremum net controls their L^r distances.

2.6 Removing the gate at large depth

The constants of Lemma 2.6 in the lower bound are uniform in the deterministic mask: inserting zero high-band columns is allowed in Lemma 2.5, and omitting inactive low factors only improves the density comparison. The binomial posterior proof uses the fixed arcsine law on active slots and zero elsewhere. The arithmetic estimates of Section 2.2 are uniform in the joint aspect. They therefore prove the compatible-center lower bound in Theorem 2.3 in the entire joint window $M = \lfloor aJ \rfloor$.

For an unrestricted admissible center g, choose $K = \lceil C_H\epsilon^{-2} \rceil$ so that its pinned Fejér polynomial g_K through $|h| \le K$, with the original affine endpoint, is within $\epsilon/4$ of g. The Fejér convolution preserves the circular half-Hölder seminorm of the pinned part, and subtracting its value at zero restores pinning. Thus g_K has uniformly bounded seminorm depending only on H. Its normalized coefficients are in $[-2, 2]$ and zero outside the cutoff.

Prescribe that every gate at $|h| \le K$, including zero, is active. Among nonzero residues of A the fraction of this mask is

$$2^{-2K-1} + O(K/\sqrt{p}).$$

Indeed expand $\prod_{|h|\le K}(1 + \chi(B(A - h)))/2$ using the quadratic character χ. The distinct shifts give squarefree products for nonempty subsets and the sum of the expansion coefficients is one. The Weil bound and deleting the singular and excluded residues give the error uniformly in B. By taking D_H large in $\epsilon \ge D_H/\sqrt{\log p}$, we have $K \le c_0 \log p$ with c_0 small, so this fraction is at least $\exp(-C'_H\epsilon^{-2})$.

Every residue in the mask makes g_K compatible. If $\epsilon/2 \ge D'_H\sqrt{\log(2J)/J}$, its conditioned ball at radius $\epsilon/2$ has probability at least $\exp(-C''_H\epsilon^{-2})$ by the first part of the theorem. This ball is inside the requested ball. The residue of a uniform unit modulo p^n is uniform among nonzero residues modulo p, so multiplication gives the unconditional lower bound. For $n \ge C \log p(\log\log p)^2$, one has $J \ge c \log p \log\log p$. Since $\log(2J)/J$ decreases in this range, it is $O(1/\log p)$, completing the depth assertion. This is a sufficient depth condition.

2.7 An upper bound in L^1

Take $m = \lfloor k\epsilon^{-2} \rfloor$ with $k > 0$ small. Choose a smooth real multiplier ϕ supported inside $(1, 2)$ and bounded below by a positive constant on a fixed subinterval. Let T_m have Fourier multipliers $\phi(h/m)$, and take $J = 8m$ time samples $t_j = j/J$. The kernel bound $Cm/(1 + m^2 \operatorname{dist}(t, \mathbb{Z})^2)$

follows by two summations by parts, and its average absolute value over the displaced sample grid is bounded uniformly in the displacement. Hence for every periodic L^1 function (viewing functions on $[0,1]$ this way),

$$\frac{1}{J}\sum_j |T_m f(t_j)| \ll \|f\|_1 . \tag{3}$$

Use a fixed finite η-net for the bounded endpoint W_0. If $\|K_Q - g\|_1 < \epsilon$, some net point w is within η of W_0. Write

$$y_j = \sqrt{m}\, T_m(Q_* - tW_0)(t_j), \qquad z_j = \sqrt{m}\, T_m(g - tw)(t_j).$$

Then the average of $|y_j - z_j|$ is $O(\sqrt{m}(\epsilon + Q^{-1/2}) + \eta)$. For the first term use (3) on $Q_* - g$. For the correction by $w - W_0$, the Fourier coefficient of t at nonzero h is $-1/(2\pi i h)$, and sampled Parseval bounds the mean square of its normalized band by $O(|w - W_0|^2)$. No endpoint of g is needed.

By (2), $y = TW$ where the column of T at h has entries

$$\sqrt{m}\, \phi(h/m) c_{Q,h} e(ht_j)/(2\pi i h).$$

Distinct columns are orthogonal, each has bounded norm, and their square norms sum to at least $c_0 m$. In the fixed-prime case only use active columns, which have positive periodic density; the core subinterval still contains $\gg_p m$ active indices. Every row has square norm $O(1)$. If U restricts to the coordinates outside a set I of at most αJ deleted samples, then $\|UT\|_{2\to 2} = O(1)$ and the sum of column square norms is still at least $c_0 m/2$, uniformly for α below a sufficiently small fixed constant.

For independent bounded model coefficients and any v in the Euclidean ball of radius $C_1\sqrt{m}$, the map $x \mapsto Z = \|UTx - v\|_2$ is convex and Lipschitz with a bounded constant, treating complex coefficients as two real coordinates within one independent block. Talagrand's convex-distance inequality [6, Theorem 4.1.1] gives subgaussian tails around its median. Indeed, for any sublevel set of a convex Lipschitz function, Jensen bounds the excess above the level by the Lipschitz constant times the Euclidean distance to any convex mixture of points from the set. The distance to the mixture is at most twice the coefficient bound times the Euclidean norm of the vector of coordinate-block mismatch proportions. Taking the infimum over mixtures gives the required comparison with convex distance. The usual upper and lower median tails follow by applying the inequality to the median sublevel and to a sublevel below the median, respectively. Now $\mathbb{E}Z^2 \geq c_2 m$ from independent mean-zero coefficients with variance one (variance two at active fixed-prime slots). On the other hand $\mathbb{E}Z^2 \leq 2(\operatorname{med} Z)^2 + O(1)$. Thus $\mathbb{P}(Z^2 < c_3 m) \leq \exp(-c_4 m)$.

On the coefficient cube UTW lies in the Euclidean ball of radius $C_1\sqrt{m}$ for a suitable fixed C_1. Project Uz there and use a net of mesh $\delta\sqrt{m}$ contained in the ball, of size at most $(1 + 2C_1/\delta)^{2J}$. If $|y_j - z_j| \leq \delta$ at each retained coordinate, the selected net point v has $\|UTW - v\|_2 \leq C\delta\sqrt{m}$. For sufficiently large fixed C_2 the polynomial

$$F_{I,v}(x) = \left(1 - \frac{\|UTx - v\|_2^2}{C_2 m}\right)^m$$

lies in $[0,1]$ on the coefficient cube (or product of trace disks). Its value on this event is at least $\exp(-C_3\delta^2 m)$. Its model mean is at most $\exp(-c_3 m/C_2) + \exp(-c_4 m)$. Its total degree is at most $2m$, and its weighted character norm is $\exp(O(m))$ by the total-degree estimate in Appendix .1, or by the cosine analogue there. Section 2.2 pays this comparison cost with a remainder smaller than the model bound, by requiring $m \leq (k/d^2)\log P$ or $m \leq (k/d_p^2)\sqrt{n}$. For translates, Chebyshev and a union over all deleted masks, their nets and all positive integer values of m costs at most

$(1 + \log L) \exp(C_\delta k \log L/d^2)$. Choose d large, leaving an $O(L^{-c'})$ fraction exceptional. The bound therefore holds for all these candidate polynomials on one good set.

Choose δ sufficiently small that the lower certificate loses at most one fourth of the model exponent. For fixed g and rounded endpoint only one net point per mask is used. Choose α small so the sum of binomial coefficients through αJ costs less than a quarter of the resulting exponential rate. The constants in the model concentration were uniform for all smaller α. Choose η, k so the average discrepancy from (3) is at most $\alpha\delta$, then at most αJ coordinates need deletion. The finite endpoint net costs a fixed factor. Union over endpoints and masks proves the claimed upper bound after reducing ϵ_* to absorb fixed factors and the integer part. Since $\|f\|_1 \le \|f\|_r$ for every $1 \le r \le \infty$, this finishes all upper bounds, including Corollary 2.4.

2.8 Finite support and the radius

For clarity, finite support itself prevents a smaller uniform coverage radius for the prime and translated prime ensembles. It also gives a lower cutoff of order $n^{-1/2}$ in the fixed-prime case, without proving sharpness of the depth comparison range.

Here is a packing inside any fixed $\mathcal{H}_H$ (and the gated subclass at fixed prime). Put normalized coefficients of a fixed small amplitude on active indices h in $[m, 2m]$, with independent choices of signs. There are $\asymp m$ such indices in the two prime cases and $\asymp_p m$ in the gated case. Apply Lemma 2.5 to the real and imaginary parts of the unpinned band on a Bernstein grid, with matrix coefficients rescaled to be bounded by one. With a sufficiently large fixed λ, a set of at least $\exp(cm)$ choices has unpinned sup norm $O(m^{-1/2})$ (constants may depend on p in the gated case). Pinning preserves this order. Bernstein's inequality gives derivative bound $O(m^{1/2})$, so the minimum of the amplitude and the derivative times an increment gives a common half-Hölder bound. Scaling the amplitude handles any prescribed $H > 0$. Hamming greedy packing leaves exponentially many choices separated in a fixed fraction of coefficients. Parseval on the nonzero Fourier modes gives squared L^2 distance $\gg_H m^{-1}$, while their L^∞ distance is $O_H(m^{-1/2})$. It follows from $\|f\|_2^2 \le \|f\|_1 \|f\|_\infty$ that their L^1 distances are $\gg_H m^{-1/2}$. The same separation holds in every L^r norm. For a family supported on J_0 paths take m a sufficiently large multiple of $\log J_0$, so there are more packed centers than paths. At least one center has an empty ball of radius $c_H(\log J_0)^{-1/2}$. For full and translated primes $J_0 \le p$ or $J_0 \le L$, and for the fixed-prime family $J_0 \le p^{n-1}$. No depth-uniform base-prime lower cutoff for coverage follows from this support count.

A depth-independent obstruction from the gates. Fix $0 < \delta < 1/4$. For every sufficiently large p there is a single admissible center with half-Hölder bound H whose uniform distance from all unit-parameter paths at all depths $n \ge 2$ is at least $c_{H,\delta} p^{-1/8-\delta/2}$. To construct it, take $m = \lceil p^{1/4+\delta} \rceil$ and use one of the signed-band centers just constructed, with normalized coefficients of amplitude $a_H > 0$ at every $m \le h < 2m$. The Burgess bound [13] for the quadratic character on any such translated interval gives

$$\sum_{m \le h < 2m} \chi(B(A - h)) = o(m)$$

uniformly in the unit residues. Thus at least $m/4$ shifts in the band have inactive gates. Their exact completed traces vanish for every $n \ge 2$. By (2), the normalized pinned Fourier coefficients of the mesh path also vanish there. Parseval over these shifts bounds the pinned distance below by $ca_H/\sqrt{m}$. Pinning costs only a factor of two in the uniform norm, and the mesh error is $O(p^{-n/2})$.

.1 Prime character comparison

For an odd prime p and $j \geq 1$, write $\mathrm{Kl}_j(c;p)$ for the normalized rank-j hyper-Kloosterman sum

$$p^{-(j-1)/2} \sum_{x_1 \cdots x_j = c} e_p(x_1 + \cdots + x_j), \qquad e_p(x) = \exp(2\pi i x/p),$$

with all variables nonzero. In particular $\mathrm{Kl}_1(c;p) = e_p(c)$. Fix the completed coefficient rank $R \geq 2$. For $a \in \mathbb{F}_p^\times$ and $b \in \mathbb{F}_p$, set

$$\varphi_{a,b}(x) = \mathrm{Kl}_{R-1}((ax)^{-1};p)e_p(bx) \quad (x \neq 0), \qquad \varphi_{a,b}(0) = 0.$$

Let $K_p(a,b)$ be the polygon with vertices $p^{-1/2}\sum_{x=1}^{j} \varphi_{a,b}(x)$ at times $j/(p-1)$, $0 \leq j \leq p-1$. At $R = 2$ these are ordinary Kloosterman paths with inverse coefficient a^{-1} and linear coefficient b.

The normalized summands are bounded by a constant depending on R by the Kloosterman sheaf trace bound. Their prime mesh identity and pinned Fourier coefficients are therefore as in (1)–(2), with mesh error $O_R(p^{-1/2})$. Haar traces of $G_R = \mathrm{USp}_R$ for even R and $G_R = \mathrm{SU}_R$ for odd R have mean zero, second absolute moment one, and modulus at most R by irreducibility of the standard representation and character orthogonality. At $R = 2$ the Haar trace has the Sato–Tate law.

The completed coefficients of the normalized increments are

$$W_h = p^{-1/2} \sum_{x \neq 0} \varphi_{a,b}(x)e_p(-hx).$$

For $b \neq h$, introducing the last multiplicative coordinate $(b-h)x$ gives exactly $\mathrm{Kl}_R((b-h)/a;p)$. At $b = h$ the sum factors into $R-1$ complete nonzero additive sums, giving $(-1)^{R-1}p^{-(R-1)/2}$. This small real number belongs to the compact trace image: use conjugate eigenvalue pairs, and for odd R also an eigenvalue one. Thus compact group characters may be evaluated at a chosen class with this trace at the exceptional point. At ordinary points use the semisimple Frobenius class of the normalized Kloosterman sheaf. For even R multiply its trace class by the central element $-I$, which preserves the Haar model. For odd R the trace convention already has positive sign.

For distinct active shifts, let $F(b)$ be a product of nontrivial irreducible compact characters of those classes, with d_0 active shifts and product of dimensions D. We claim

$$\left| \sum_{b \in \mathbb{F}_p} F(b)e_p(kb) \right| \ll_R (d_0+1)D\sqrt{p} \qquad (k \in \mathbb{F}_p). \tag{4}$$

The normalized single-factor sheaf has arithmetic and geometric group Sp_R or SL_R, tame regular unipotent monodromy at zero and a single infinity break $1/R$, by Katz's results as recorded in [9, Prop. 4.6]. In the product over translates, inertia at $b = h$ acts nontrivially and unipotently in that factor alone. Projection onto each factor is surjective, and the normal closure of a noncentral unipotent in its connected simple group is the full factor. Hence the geometric group is the product. Applying the irreducible representations and tensoring gives rank D, no invariants, and infinity breaks at most $1/R$. For a nontrivial additive twist the break at infinity is one, since the original breaks are strictly smaller. In particular the twisted tensor has no invariants. All finite ramification is tame. The Euler characteristic on the affine line with the active punctures removed gives $\dim H_c^1 \leq d_0 D$ and $H_c^0 = H_c^2 = 0$. Purity and the trace formula bound the sum there by $d_0 D\sqrt{p}$. The exceptional points contribute at most $d_0 D$. This proves (4). The all-trivial product has the exact full or restricted expectation one.

We use the norm $\|P\|_{\mathrm{ch}}$ of an expansion into products of characters, with absolute coefficients weighted by product dimensions. It is submultiplicative, and both the trace and its complex conjugate have norm R. For a polynomial in traces and dual traces of total degree r on d factors, with bound U on the product of compact groups,

$$\|P\|_{\mathrm{ch}} \le U \exp(C_R r \log(eC_R(d+r)/r)) \quad (r \ge 1). \tag{5}$$

Indeed Haar character coefficients are bounded by U by orthonormality. Each possible highest weight in one factor has weight norm bounded by C_R times its trace degree, since it occurs in a tensor product of the standard representation and its dual. The total highest-weight norm is at most $C_R r$. Stars and bars count the products of possible highest weights, and the product of dimensions is at most $\exp(C_R r)$ by the Weyl dimension formula. For separate degrees r_h, the same argument bounds the norm by $U \prod_h [C_R(1+r_h)^{C_R}]$ on active factors, with factor one when $r_h = 0$. These estimates also apply to real and imaginary components expressed using the trace and its dual.

Let $m_S(k) = L^{-1} \sum_{b \in S} e_p(kb)$ and $\widehat{F}(k) = \sum_b F(b) e_p(-kb)$. For each nontrivial product F, Parseval over translates gives

$$\left\| \mathbb{E}_{b \in S+u} F(b) - \mathbb{E}_{b \in \mathbb{F}_p} F(b) \right\|_{L^2_u}^2 = \frac{1}{p^2} \sum_{k \ne 0} |\widehat{F}(k)|^2 |m_S(k)|^2.$$

Use (4) and $\sum_k |m_S(k)|^2 = p/L$, and add the error of the full average relative to the independent Haar characters. Thus for a test on d shifts,

$$\left\| \mathbb{E}_{b \in S+u} P(W) - \mathbb{E} P(X) \right\|_{L^2_u} \ll_R (d+1) L^{-1/2} \|P\|_{\mathrm{ch}}. \tag{6}$$

The complete-family discrepancy in (4) has the same form with p in place of L. For scalar Sato–Tate or arcsine comparisons one can also use the angular lift $x = 2\cos\theta$ of a polynomial. Fourier coefficients are bounded by the cube supremum and each sign class gives one cosine mode. For Sato–Tate, $\cos(k\theta) = (\chi_k - \chi_{k-2})/2$ for $k \ge 2$ and $\cos\theta = \chi_1/2$, so the dimension-weighted cost of one mode is at most $k+1$. For arcsine use weight $k+1$ directly; the product formula for cosines shows submultiplicativity. Summing over all modes through separate degrees d_h gives the bound $\prod_h (d_h+2)^2$ times the cube supremum. For total degree r on d coordinates the number of mode vectors with sum at most r is $\binom{d+r}{r}$ and the product of weights is at most 2^r. The possible sign classes contribute at most another 2^r. This yields the stated total-degree cosine bound.

.2 Prime-power cosine comparison

In this appendix write $q = p^n$ for the modulus.

Put

$$W_h = q^{-1/2} \sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} e_q((A-h)x + B/x).$$

When $A-h$ is a unit, change variables to identify it with $q^{-1/2} \sum_{x \in (\mathbb{Z}/q\mathbb{Z})^\times} e_q(x + C/x)$ for $C = B(A-h)$. If $A-h$ is not a unit, the original sum vanishes by translating x by a multiple of p^{n-1} and using the nonzero derivative modulo p. The transformed sum with nonunit C vanishes by the same argument, since $1 - C/x^2$ is a unit.

The transformed sum is zero unless C is a nonzero square modulo p. In that case let s be either Hensel root of C modulo q. For even $n = 2k$, partition x modulo p^k and sum the higher digits. A class contributes only when $1 - C/x^2$ vanishes modulo p^k, and the two normalized terms

are $e_q(2s)$ and its conjugate. For odd $n = 2k+1$, first partition modulo p^{k+1}, which requires stationarity modulo p^k. Around a root s, use the p representatives $s + p^k y$, y modulo p. The phase is $2s + p^{2k} y^2/s$ modulo q, since $3k \geq 2k+1$. After normalization the root contributes $e_q(2s)G(s^{-1})$, where $G(d) = p^{-1/2} \sum_{y \bmod p} e_p(dy^2)$ has modulus one for d a unit. The other term is its conjugate. This is the usual prime-power stationary-phase formula (compare [12, Lemma 2]). Thus always $|W_h| \leq 2$, and on the square condition $W_h = 2\cos\theta_h$, where the additional Gauss phase depends just on the root modulo p.

Write $C_k(x) = \cos(k \arccos(x/2))$. A product of these polynomials on the square coordinates expands into exponential sums with phase $2\sum_i k_i s_i(A)$ for signed nonzero integers k_i on the active shifts. The additional Gauss factor is constant on each residue class of A modulo p. At nonsquare coordinates, $C_k(W_h) = \cos(k\pi/2)$ is deterministic. For the moment of a monomial in the W_h one may also first impose that all its shifts are square, since the monomial is zero otherwise.

For distinct active shifts $h_1, \ldots, h_t$, take formal square roots $s_i(A)^2 = B(A - h_i)$. Over $\overline{\mathbb{F}}_p(A)$ the square classes of the $B(A-h_i)$ are independent, as shown by their valuations at the distinct places $A = h_i$. In the multiquadratic field, define

$$N(A) = \operatorname{Norm}\left(\sum_{j=1}^{t} k_j \prod_{i \neq j} s_i(A) \right),$$

where the norm is the product over the 2^t sign changes. This is nonzero if each k_i is nonzero in the field. Indeed the rational expression $\sum k_i/s_i$ is nonzero, by subtracting its image under a single sign change. The norm is a polynomial of degree at most $(t-1)2^{t-1}$, since expanding the product and cancelling the odd powers of each square root leaves that degree bound.

Lemma .7. *For s shifts with distinct reductions modulo p, and integers $1 \leq k_j < p$, the discrepancy between the arithmetic expectation of $\prod_{j=1}^{s} C_{k_j}(W_{h_j})$ and its expectation for independent $(\delta_0 + \sigma)/2$ variables is*

$$O(sp^{-1/2} + s3^s/p),$$

uniformly over all $n \geq 2$ and B.

Proof. On each class of A modulo p^{n-1}, average the last digit. A lift of a chosen root changes by $Bp^{n-1}u/(2s_i)$ when A changes by $p^{n-1}u$. Thus a nonempty square subset I in the product contributes only when a signed relation $\sum_{i \in I} \pm k_i/s_i = 0$ holds modulo p. As the roots are nonzero, its presence forces a zero of the norm polynomial in A modulo p. The exceptional sets over all subsets have total size $O(s3^s)$, since a subset of size t has norm degree at most $(t-1)2^{t-1}$. The exceptional residues $A = h_i$ cost another $O(s/p)$. On the event all coordinates are nonsquares, the product is identically $\prod \cos(k_i\pi/2)$. Its probability is $2^{-s} + O(s/\sqrt{p})$ by expansion of the product of $(1 - \chi(B(A-h_i)))/2$, with the quadratic character χ, using Weil's bound for squarefree products and excluding the at most $s+1$ singular or nonunit residues. For the independent law the positive-degree cosine means are $\cos(k_i\pi/2)/2$, which gives the matching main term. □

Lemma .8. *Suppose t active shifts are distinct modulo p^u, with $u \geq 1$, and the signed coefficients k_i are nonzero integers. Put*

$$k = 2 + \lceil \log_p(2t) \rceil, \qquad V = \lfloor n/2 \rfloor - \min_i v_p(k_i) - (t-1)u - kt.$$

If $V > 0$, the normalized signed phase contribution on the square mask is $O(tp^{-V/(t-1)})$ for $t \geq 2$, and zero for $t = 1$. The same estimate holds in any specified unit residue class of A modulo p. Constant factors in the phase on each residue class, of absolute value at most one, may be included.

Proof. Sum the top digits of A in each class modulo $p^{\lceil n/2 \rceil}$. The Taylor expansion of the square roots in that class is linear modulo p^n, since the higher-order terms vanish. Orthogonality gives a zero contribution unless the derivative sum $D(A) = \sum_i k_i / s_i(A)$ is zero modulo $p^{\lfloor n/2 \rfloor}$. We bound the fraction of points satisfying this condition inside each class modulo p^k on which the gate is active. The condition $V > 0$ guarantees $k \le \lfloor n/2 \rfloor$.

Fix such a class centered at $A = a$. Let $c_i = k_i / s_i(a)$, $z_i = (a - h_i)^{-1}$, and $M_j = \sum_i c_i z_i^j$. The z_i are units and $v_p(z_i - z_l) = v_p(h_i - h_l) < u$. For i minimizing $v_p(k_i)$, the coefficients of $\prod_{l \ne i}(z - z_l)$ give a linear combination of $M_0, \dots, M_{t-1}$ equal to $c_i \prod_{l \ne i}(z_i - z_l)$. Consequently

$$b = \min_{0 \le j < t} v_p(M_j) \le \min_i v_p(k_i) + (t-1)u.$$

All later M_j have valuation at least b, by the integral recurrence with characteristic polynomial $\prod_i (z - z_i)$.

Expand on $A = a + p^k z$. The jth coefficient of $D(a + p^k z)$ is $p^{kj} \binom{-1/2}{j} M_j$. The binomial is integral for odd p and its valuation for $j \ge 1$ is the number of carries in $j + j$ in base p, at most $1 + \lfloor \log_p(2j) \rfloor$. For $j = 0$ its valuation is zero. Thus some index $j_0 < t$ attaining b gives a rescaled coefficient of valuation strictly below $b + kt$, while all $j \ge t$ have valuation at least $b + kt$. Divide the series by p^κ, where κ is its minimum coefficient valuation. The resulting restricted power series has last unit coefficient at some index $d \le t - 1$, and $\kappa \le b + kt$.

If $d = 0$ this series is a unit everywhere. Otherwise, setting $L = \lceil V/d \rceil$, suppose it takes values of valuation at least V at $d + 1$ points distinct modulo p^L. Their divided difference of order d then has valuation at least $V - d(L-1) > 0$, by the interpolation formula. On monomials of degree below d this divided difference is zero, at degree d it is one, and above d it is an integral homogeneous symmetric polynomial in the points. The last identity follows by expanding $\prod_i (1 - z_i X)^{-1}$. Passing to the restricted series shows that the divided difference equals the unit coefficient at d modulo p, a contradiction. There are thus at most d solution classes modulo p^L in the z variable. Here $L \le n - k$. The relative count per class of A is at most $d p^{-V/d}$, which implies the result after averaging. □

The proof of Lemma .7 also allows a prescribed gate mask. Partition into patterns if needed. A cosine mode of positive degree on an active coordinate has zero independent mean under arcsine. If any such mode is active, its last-digit sum vanishes off the same exceptional set by the signed derivative test. Otherwise the remaining cosine modes are constants at zero and the mask probability is compared by expanding products of $(1 \pm \chi(B(A - h)))/2$. The same error estimate applies for any specific mask, with s now counting all tested and prescribed shifts. Summing over arbitrary low gate masks in an observation costs at most 2^{2M+1}, as used above.

For the joint depth aspect, take $u = 1 + \lceil \log_p(4N) \rceil$ with $N = \lfloor e^J \rfloor$. Then all shifts tested are distinct when reduced modulo p^u. For support $t = O(M + s)$ and degree $O(M^2 + s)$, s fixed or $O(\log M)$, the loss in V is bounded as follows:

$$(t-1)u = O(a(J + J^2/\log p)) + o(n), \qquad kt = O(a(J + J\log(2J)/\log p)) + o(n).$$

Indeed $J \le n$, $J^2 \le n \log p$, and $J \log(2J)/\log p = O(n)$, by distinguishing $n \le \log p$ from $n > \log p$. The errors from the $O(\log M)$ terms and the valuations of the signed degrees are $o(n)$ uniformly as $J \to \infty$. Thus for small a the remaining V is at least $n/4$. By Lemma .8, conditioned on each residue the error of a nontrivial signed term is $O(tp^{-cn/t})$. The local divided-difference count is a proportion within each residue ball and incurs no factor p from the residue conditioning. Since $n \log p \ge J^2$, the error is at most $\exp(-CJ)$ with any fixed C by reducing a, after paying the norm loss of Section 2.2.

.3 Pinned Fourier tails

For a dyadic band $U/2 < |h| \le U$ in (1), put $T(t) = \sum b_{Q,h} W_h e(ht)$. By Parseval on T^2 its fourth supremum power is at most $(4U+1)\int |T|^4$. The integrated moment has only tuples $h_1 + h_2 = h_3 + h_4$. Their total absolute coefficient product is $O(U^{-1})$, since fixing three determines the last. If two distinct selected integer shifts agree modulo J, the marked pair has $O(U^2/J)$ choices (a collision implies $2U \ge J$), one other shift has $O(U)$ choices, and the frequency equality determines the last. The colliding contribution is then $O((JU)^{-1})$. Under any of the independent bounded centered models the fourth block supremum moment root is $O(U^{-1/4})$, from pointwise fourth moment $O(U^{-2})$ and the same Parseval inequality.

For complete primes use the character estimate for the at most four distinct shifts, which are distinct modulo p throughout the symmetric complete range. The discrepancy in every balanced tuple is $O(p^{-1/2})$. The fourth block supremum moment root is $O(U^{-1/4} + p^{-1/8})$. The complete prime expectation also governs the integral of expectation over translated sets by Fubini.

For large-base prime powers with $2 \le n < 400$ use Lemma .7 on good tuples modulo p and the colliding bound with $J = p$. For $n \ge 400$ (including in fixed-prime residues for large n) use $J = p^{\lfloor n/200 \rfloor}$. On good tuples fix the gate pattern via the residue class of A. The zero signed phases give the independent square-gated model in that residue class; the remaining terms have support at most four and signed degrees at most four. Lemma .8 pays $O(Q^{-c})$. The model fourth bound is uniform in that gate pattern. After using the collision estimate and taking fourth roots, all prime-power cases have the bound $O(U^{-1/4} + Q^{-c_0})$ for some absolute $c_0 > 0$, reducing it to cover bounded depths. For the joint window with $n \to \infty$ the terminal constants are uniform in odd p. In Lemma .8 at $t \le 4$ the extra carry and signed-valuation losses are bounded absolutely, including $p = 3$. Partial blocks satisfy the same estimates.

At mesh points pinning costs a factor of at most two. Summing dyadic blocks beyond N gives

$$\mathbb{E}\,\|Q_* - R_N\|_\infty \ll N^{-1/4} + (1+\log Q)Q^{-c_0} + N/Q.$$

The last term follows by interpolating R_N, whose Lipschitz constant is $O(N)$. This proves the terminal assertions of Section 2.2 uniformly, since the comparison window there has $N \ll Q$.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, proof drafting and internal checks. No data or numerical certificates are required for the proofs.

A Exact uniform prime Kloosterman shrinking targets

For each fixed inverse coefficient, classical prime Kloosterman paths have uniform-ball probabilities of order $\exp(-\Theta(\epsilon^{-2}))$ about supported half-Hölder centers down to radius $D/\sqrt{\log p}$. This is the optimal order of the radius for uniform coverage. The result extends to hyper-Kloosterman paths of every fixed completed rank, long progressions, and most translates of every growing set. The proof projects a localized function onto an ellipsoid of compact-group eigenvalues. Finite propagation gives a signed character polynomial whose independent mean is positive and whose character norm is exponential in ϵ^{-2}.

Keywords Kloosterman paths, shrinking targets, finite propagation, character comparison.
MSC 2020 11L05, 11T23, 60G17, 58J45.

A.1 Statements

Write $e(t) = \exp(2\pi i t)$ and $e_p(t) = e(t/p)$. For $j \geq 1$, define

$$\mathrm{Kl}_j(c; p) = p^{-(j-1)/2} \sum_{x_1 \cdots x_j = c} e_p(x_1 + \cdots + x_j)$$

with all summation variables nonzero. Fix a completed rank $R \geq 2$, an odd prime p, and $a \in \mathbb{F}_p^\times$. For $b \in \mathbb{F}_p$ let

$$\varphi_{a,b}(x) = \mathrm{Kl}_{R-1}((ax)^{-1}; p) e_p(bx) \quad (x \neq 0), \qquad \varphi_{a,b}(0) = 0.$$

Let $K_p(a, b)$ linearly interpolate the vertices $p^{-1/2} \sum_{x=1}^{j} \varphi_{a,b}(x)$ at times $j/(p-1)$, $0 \leq j \leq p-1$. At $R = 2$ the inverse coefficient is a^{-1} and the linear coefficient is b.

For a continuous complex path g starting at zero set

$$\beta_0 = g(1), \qquad \beta_h = 2\pi i h \, \widehat{g - tg(1)}(h) \quad (h \neq 0),$$

with Fourier coefficients on $[0, 1]$ and convention $e(-ht)$. Write $\mathcal{H}_H$ for paths with half-Hölder seminorm at most H and all $\beta_h \in D_R$, the trace image of $G = \mathrm{USp}_R$ for even R and $G = \mathrm{SU}_R$ for odd R. In rank two this is the half-Hölder class in the usual limiting support [8].

Theorem A.1. *Fix $R \geq 2$, $H > 0$ and, in the progression case, $\delta > 0$. Sample b uniformly in one of the following families.*

1. *The complete field $\mathbb{F}_p$, with $P = p$.*
2. *A nonzero-step progression of length at least $p^{1/2+\delta}$, with $P = p$.*
3. *A translate $S + u$ of any L-point set $S \subseteq \mathbb{F}_p$, with $P = L$. At most $O(L^{-c_0})$ of the translates, as a fraction, must be excluded.*

There are constants $D, C, c, \epsilon_*, c_0 > 0$ such that for sufficiently large P and every $g \in \mathcal{H}_H$,

$$\exp(-C\epsilon^{-2}) \leq \mathbb{P}_b\{\|K_p(a, b) - g\|_\infty < \epsilon\} \leq \exp(-c\epsilon^{-2}) \qquad \left(\frac{D}{\sqrt{\log P}} \leq \epsilon \leq \epsilon_*\right).$$

The upper bound in this radius range also holds with $\|\cdot\|_1$ for every L^1 center. The excluded translates can depend on a, S, and the same set works for all centers and radii. All constants are uniform in a. The lower radius has optimal order for uniform coverage of $\mathcal{H}_H$.

Kowalski and Sawin [7, 8] identify the classical limiting process and its support. The November 21 edition of [2] proves a sharp model exponent and transfers the lower bound to prime arithmetic balls from radius of order $(\log p)^{-1/4}$. Kowalski and Untrau [15] develop dimension-weighted Wasserstein comparison, including higher-rank scalar counts near submersive targets in their Proposition 4.6. Here a fixed inverse coefficient suffices, and the test gives relative path-ball probabilities at the logarithmic coverage order. The construction uses concentration for strongly log-concave measures [16] and the finite speed of the wave equation [17]. Carruth, Elkies, Gonçalves and Kelly [18] study signed band-limited box minorants for Lebesgue measure. Our certificate uses a weighted compact-group spectrum and controls its negative contribution under the product Haar law.

A.2 Fourier and arithmetic estimates

Put $W_h = p^{-1/2}\sum_{x\neq 0}\varphi_{a,b}(x)e_p(-hx)$. These coefficients are traces of compact classes indexed by h modulo p. Appendix .1 gives their character comparison with independent Haar elements of G. For a character polynomial F, let $\|F\|_{\mathrm{ch}}$ be the sum of the absolute coefficients times their product dimensions. If every term uses at most j shifts, the comparison error is

$$\begin{cases} O_R((j+1)p^{-1/2}\,\|F\|_{\mathrm{ch}}) & \text{on the complete field,}\\ O_R((j+1)p^{-\delta}\log p\;\|F\|_{\mathrm{ch}}) & \text{on long progressions,}\\ O_R((j+1)L^{-1/2}\,\|F\|_{\mathrm{ch}}) & \text{in } L^2 \text{ over translates.}\end{cases}$$

These estimates use distinct residues for the shifts in each character term.

For the time mesh at j/p write Q_p for the path with increments $p^{-1/2}\varphi_{a,b}(x)$, extended by zero at the last mesh interval. The bounded normalized summand gives $\|Q_p - K_p\|_\infty = O_R(p^{-1/2})$. Fourier inversion at mesh vertices gives

$$Q_p(t) = tW_0 + \sum_{0<|h|\le(p-1)/2} b_{p,h}W_h(e(ht)-1), \qquad b_{p,h} = \frac{1}{p(1-e(-h/p))}.$$

In particular $b_{p,h} \ll |h|^{-1}$ and $b_{p,h} = (2\pi i h)^{-1} + O(p^{-1})$. Retain $|h| \le N = \lfloor P^{1/8}\rfloor$ with those last multipliers. Replacing the coefficients and interpolating the retained polynomial changes the bound at all times by $O_R(N/p+p^{-1/2})$, by summing the coefficient errors and bounding the retained derivative by $O_R(N)$. For later use, integration of the piecewise constant derivative also gives

$$2\pi i h\,\widehat{Q_p - tW_0}(h) = c_{p,h}W_h, \qquad c_{p,h} = e(h/(2p))\,\mathrm{sinc}(\pi h/p),\quad \mathrm{sinc}\,x = \sin x/x$$

for $0 < |h| \le (p-1)/2$.

We record the dyadic tail bounds. In a band of frequency size U, integrating the $2r$th power of the modulus of its unpinned sum leaves tuples with $h_1+\cdots+h_r = h_{r+1}+\cdots+h_{2r}$. Their total absolute coefficient weight is at most C_R^{2r}/U, since all but one index determine the last. By Parseval and Cauchy–Schwarz on the rth power, its supremum to power $2r$ is at most CrU times this integral. The Haar traces are bounded, centered and have variance one by character orthogonality. Pointwise concentration therefore gives the independent-model supremum moment bound

$$C_R(rU)^{1/(2r)}\sqrt{r/U} \tag{7}$$

after taking the $2r$th root. This applies also to partial bands. For the unweighted terminal tail $|h| > N$ through $(p-1)/2$, take $r = 2$, using $b_{p,h}$ and the preceding arithmetic comparison on four shifts. The mean supremum tail, including pinning, is $O_R(P^{-\eta})$ for some $\eta > 0$ on the complete field and long progressions. In detail the block bound is $O_R(U^{-1/4}+p^{-\eta'})$, with η' reduced to absorb the progression logarithm. For translates, the translate mean of this terminal expectation is its full-field expectation by averaging. Markov then bounds the terminal expectation by $O_R(P^{-\eta''})$ off a power-saving exceptional fraction, independently of centers and radii. Interpolated tail contributions have norm bounded by their maximum at the time mesh.

A.3 Localized tests of controlled energy

A.3.1 The strongly log-concave cutoff estimate

For even $R = 2\ell$ let ρ be the Haar trace density. Two properties of ρ enter the even-rank proof:

$$\rho \text{ is even and strongly log-concave,} \qquad |(\log\rho)'(x)| \ll_R (R-|x|)^{-1} \quad \text{a.e. on } (-R,R). \tag{8}$$

Indeed the Weyl density in the ordered variables $u_i = \cos\theta_i$ is proportional to $\prod_{i<j}(u_i-u_j)^2 \prod_i \sqrt{1-u_i^2}$ on $-1 < u_1 < \cdots < u_\ell < 1$ [19]. Its negative logarithm has a positive lower Hessian bound from the one-variable terms, and the gap terms are convex there. The trace $x = 2\sum_i u_i$ is linear. Marginalization preserves strong log-concavity, as follows by splitting into the trace direction and its orthogonal complement and applying Prékopa's theorem [16]. The central element $-I$ gives evenness. For the right edge put $t = \ell - x/2$ and $1 - u_i = tv_i$. On the ordered unit simplex the density integral is a constant times

$$t^{\ell^2+\ell/2-1} \int \prod_{i<j}(v_i - v_j)^2 \prod_i \sqrt{v_i(2-tv_i)}\, dv,$$

where the measure is the fixed simplex measure in dimension $\ell - 1$. The integral and its reciprocal are bounded and its derivative is bounded for $0 < t \le 1$. This gives the derivative estimate at the right edge, then at the left by symmetry. On interior compact subintervals log-concavity supplies the remaining bound.

Choose an even log-concave function ϕ supported on $[-1,1]$, equal to one on $[-1/2,1/2]$, and continuously differentiable with bounded derivative. For example, exponentiate the negative of the even convex function $((|v| - 1/2)_+)^2/(1-|v|)$ on $(-1,1)$ and extend by zero. We will use the following elementary relative estimate. Suppose a reference density has an even, strongly convex negative logarithm with Hessian bounded below by κI in the sense of strong convexity. Let $s_i(x)$ be linear forms of uniformly bounded Euclidean row norm, and put $f = \prod_i \phi(s_i/T_i)$. Removing any one factor i from f^2 leaves an even strongly log-concave probability measure after normalization. Under it each s_i has uniform subgaussian tails, by applying the Brascamp–Lieb variance bound to its linear tilts and integrating twice. The same bound holds on convex support by approximation. Consequently the ratio of the squared norms with and without each factor is at least $1 - Ce^{-cT_i^2}$. The relative integral of $(\phi'(s_i/T_i))^2 \prod_{j\ne i}\phi(s_j/T_j)^2$ against the reference law, divided by the squared norm of f, is $O(e^{-cT_i^2})$ when T_i is large enough. Both facts follow since the derivative and the departure from one require $|s_i| \ge T_i/2$. They also give a bounded ratio for fixed positive thresholds if the reference law is a Gaussian of sufficiently small fixed variance. With A the matrix of the forms, the gradient is $A^{\mathsf{T}}(a_i/T_i)_i$ where $a_i = \phi'(s_i/T_i)\prod_{j\ne i}\phi(s_j/T_j)$. Its relative squared integral is thus at most $C\,\|A\|_{2\to 2}^2 \sum_i T_i^{-2} e^{-cT_i^2}$ in the large-threshold case.

A.3.2 Sample rows and the even-rank test

Fix a large integer m and a center $g \in \mathcal{H}_H$. Set $c_h = (1 - |h|/(m+1))\beta_h$ for $0 < |h| \le m$, $c_0 = (1 - m^{-1/2})\beta_0$, $w_h = |h|/(m+1)$ and $w_0 = m^{-1/2}$. The Fejér kernel applied to the pinned periodic path, and this endpoint change, approximate g uniformly with error $O_{R,H}(m^{-1/2})$. This follows from the bound $O(m^{-1/2})$ on the circular half-moment of the Fejér kernel. The trace edge distance of c_h is at least Rw_h. For trace variables x_h write $z_h = (x_h - c_h)/w_h$. Form the real and imaginary samples y_i of

$$\sqrt{m} \sum_{0<|h|\le m} \frac{x_h - c_h}{2\pi i h} e(ht)$$

on an equally spaced grid of $O(m)$ points and include the real and imaginary parts of the endpoint row $\sqrt{m}(x_0 - c_0)$. The grid is fine enough that a polynomial's supremum is at most twice the maximum modulus of its samples by Bernstein's inequality. In the variables z_h the sample matrix has bounded Euclidean operator norm by discrete orthogonality. Its nonendpoint coefficients in the complex representation have modulus $O(m^{-1/2})$ before sampling.

Use a small-fixed-variance product Gaussian density $G_0(z)$ and take a product $f_d(z)$ of cutoffs in the samples at a fixed threshold and in each z_h at threshold $1/2$. Both collections have joint operator norm $O(1)$. The relative energy estimate gives a Gaussian Dirichlet quotient $O_R(m)$, and f_d is nonzero near zero. Transfer to the product trace law in z coordinates, of density $\rho_d(z) = \prod_h w_h \rho(c_h + w_h z_h)$, by putting

$$u_d = f_d \sqrt{G_0/\rho_d}$$

on its support and zero elsewhere. On this support the squared gradient of $\log(G_0/\rho_d)$ is $O_R(m)$ by (8). Its relative Euclidean Dirichlet energy under ρ_d is therefore $O_R(m)$. On the compact group take a fixed bi-invariant metric. Its trace has a bounded gradient. With Laplacian speed $\sigma_h = w_h$ on these factors, differentiation with respect to x_h cancels the scale in z_h. The angular Dirichlet quotient is again $O_R(m)$. All the functions pull back by trace to class functions. Compact support within the scaled coordinate boxes and the cutoff regularity put the dense test in the form domain.

Use further dyadic bands through a dyadic endpoint $B_m \asymp C_* m \log(2m)$, with C_* fixed large. In a band of scale $U > m$ use real and imaginary samples of the unpinned block multiplied by $\sqrt{U}$, with a Bernstein mesh of size $O(U)$. The matrix again has bounded operator norm. Let $\tau_U = A_0(U/m)^{1/4}$ with fixed A_0 large and take Laplacian speed $\sigma_h = \sqrt{\tau_U}$ on the band. Use the product of sample cutoffs at threshold τ_U in even rank. In even rank the reference law is the centered product of ρ. The relative energy estimate and the bounded trace gradient give angular quotient $O_R(Ue^{-c_R \tau_U^2})$, whose sum is $O_R(m)$. In either parity, take the tensor product f of the dense and high tests and normalize it to norm one. Its Rayleigh quotient will be $O_R(m)$ for

$$\mathcal{L} = \sum_{|h| \le B_m} \sigma_h^2 \mathcal{L}_h,$$

the nonnegative Laplace–Beltrami operator on the correspondingly scaled product of G. Both parity constructions use the same high-band speeds and sample scales. The dyadic sum of the corresponding unpinned block bounds above m is $O_R(\sum_U \tau_U/\sqrt{U}) = O_R(m^{-1/2})$.

A.3.3 Dense localization in odd rank

For odd R the trace image of SU_R is the filled R-cusped hypocycloid [20]. We use the following quantitative point about its interior. For $c \in (1-\eta)D_R$ with $0 < \eta \le 1$, there is a maximal-torus preimage of c with simple eigenvalues, at which the least target singular value of the trace differential and the least pairwise eigenvalue gap are both $\gg_R \eta^{C_R}$.

We first verify regularity. On the determinant-one torus the real differential is onto the complex trace plane whenever there are at least three distinct eigenvalues, since it contains the real span of $i(z_j - z_R)$, as used by Kowalski and Untrau [15, Lemma 4.7]. One eigenvalue, or two eigenvalues with multiplicities $R-1, 1$, gives a point of the hypocycloid boundary. For two distinct eigenvalues u, v of multiplicities $k_1, k_2 \ge 2$, take $u_s = ue^{is/k_1}$, $v_s = ve^{-is/k_2}$ and split a pair in each cluster by opposite angles. The resulting traces are

$$k_1 u_s + k_2 v_s - r u_s - t v_s, \qquad r, t \ge 0$$

near $r = t = s = 0$. An opposite pair with angles $\pm\alpha$ gives $r = 2(1 - \cos\alpha)$. The derivative in s is $i(u - v)$, while the projections of $-u, -v$ on the normal $u - v$ are nonzero and opposite. The implicit function theorem solves for s, r in terms of t with $r'(0) = 1$, preserving the trace and giving at least three distinct eigenvalues for small positive t. This pair-splitting mechanism also appears in the trace geometry discussion of Bryant [21]. A regular preimage can be made simple: choose an angle direction with zero sum in each cluster of equal eigenvalues and pairwise distinct entries

within each repeated cluster. This direction lies in the kernel of the trace differential. A curve with this tangent in the regular trace fiber splits every repeated cluster to first order and leaves distinct clusters apart.

For the quantitative assertion, the boundary is parametrized by $\gamma(t) = (R-1)e^{it} + e^{-i(R-1)t}$. Its argument winds once, is strictly increasing away from cusps, and has derivative

$$\frac{(R-1)(R-2)(1-\cos Rt)}{|\gamma(t)|^2}.$$

Since $|\gamma| \geq R-2$ and the derivative vanishes only to order two, the radial boundary function is 1/3-Hölder. Radial contraction by $1-\eta$ gives distance $\gg_R \eta^3$ from the boundary: perturbing a contracted trace by $a\eta^3$ changes the boundary radius by at most $O_R(a^{1/3}\eta)$ away from the origin, and near the origin there is a fixed interior margin. Now maximize over the torus fiber the minimum of the gap and the least target singular value. This is a semialgebraic function of the trace, since in the algebraic torus all these quantities use polynomial entries and quantification over a compact fiber. It is positive and locally bounded below in the interior by the simple regular local sections. Its infimum over traces with boundary distance at least $b > 0$ is positive and semialgebraic in b. Polynomial growth of its reciprocal [22, Proposition 2.11] gives the asserted bound at c.

In a small orthonormal angle chart v at such a preimage of c_h, write

$$\Psi(v) = (\Re(\operatorname{tr}(v) - c_h),\ \Im(\operatorname{tr}(v) - c_h),\ \Pi_{\ker d\operatorname{tr}(0)} v).$$

The inverse differential at zero has norm $O_R(1/s_h)$ with s_h the least target singular value. The second derivatives of Ψ are bounded. Choose a polynomial scale $\delta_h = c_R w_h^{C_R}$, at most a sufficiently small constant times each of $w_h, s_h^2, s_h g_h$, where g_h is the eigenvalue gap. Quantitative inverse mapping gives a chart $v = \Psi^{-1}(\delta_h z)$ on a fixed size box, with $\|v\| + \|D_z v\| \ll_R \delta_h/s_h$ and $\|D_v z\| \ll_R \delta_h^{-1}$. The trace there is exactly $c_h + \delta_h(z_1 + iz_2)$. The logarithmic derivative of the Weyl density in the original angles is $O_R(1/g_h)$, because the chart is well inside the simple-spectrum set. For the inverse Jacobian its logarithmic derivative in v is $O_R(1/s_h)$. Thus the logarithmic derivative of the full chart density in z is $O_R(\delta_h/(s_h g_h) + \delta_h/s_h^2) = O_R(1)$.

Use the same product Gaussian and coordinate and sample cutoffs in these charts as in even rank. The complex trace coordinates are linear in z and $\delta_h \leq w_h$, so the sample operator is still bounded on Euclidean space. Divide by the square root of the chart density and insert the factor $\sqrt{G_0}$ as before. The angular Dirichlet quotient costs $O_R(m)$ with speeds $\sigma_h = \delta_h$. Weyl conjugate copies of each chart are disjoint by the gap bound. Copy the test to them and extend by zero, then lift it from the torus as a class function. The Dirichlet quotient is unchanged, since on regular classes the full group gradient of a class function is its angular gradient and Haar integration uses the same Weyl density. Scaling changes the heat partition budget only by a fixed rank-dependent factor, as $\sum_h \log(1/\delta_h) = O_R(m)$. The global trace gradients and $\delta_h \leq w_h$ give the same dense propagation bounds.

A.3.4 High-band energy from trajectory discrepancies

In odd rank use a Dirichlet test obtained from stationary Brownian paths on the group. We give the argument because the odd-rank trace image need not be convex. On the factors in a high band of size U, use generator $-\sigma^2 \mathcal{L}_h$ with $\sigma = \sqrt{\tau_U}$. For a single stationary group trajectory Z_v, $0 \leq v \leq 1$, write $X(v) = \operatorname{tr}(Z_v)$. Its martingale trace increments have brackets bounded by $C_R\sigma^2\Delta$ on time intervals of length Δ, and drift bounded by $C_R\sigma^2\Delta$, by Itô's formula and bounded trace

gradient and Laplacian. Boundedness of the trace controls intervals with $\Delta \geq \sigma^{-2}$. A union bound on adjacent dyadic increments of length 2^{-j}, followed by dyadic telescoping, gives

$$\mathbb{P}\{[X]_{1/4} > K\} \ll_R e^{-c_R \tau_U^{1/2}}, \qquad K = C_R \tau_U^{3/4}.$$

Indeed on the short increments the drift divided by $\Delta^{1/4}$ is $O_R(\sigma^{1/2})$, and the martingale tail at threshold a constant times $K\Delta^{1/4}$ is at most $C_R \exp(-c_R K^2 \sigma^{-2} \Delta^{-1/2})$. Summing 2^j such probabilities over all j gives the assertion.

Condition each coordinate trajectory to obey this Hölder bound, at product probability cost at least $\exp(-C_R U e^{-c_R \tau_U^{1/2}})$. This conditioning is invariant under multiplication by the center of SU_R; the central trace action is by Rth roots of unity. Sample independent conditioned latent trajectories, then multiply them by independent uniform central roots. The distribution stays the conditioned product law. With the latent trajectories fixed, require that all Fourier sample rows in the band remain at scale τ_U throughout diffusion time. Besides the samples at time zero, use parent-to-child increments on the dyadic diffusion-time grid. On multiplying the normalized rows by $\sqrt{U}$, at level j there are $O_R(U2^j)$ real rows with individual entries bounded by $C_R K 2^{-j/4}$, including the level zero rows after changing the constant. The entries depend on one of R colors at each column, and average to zero across colors.

Here are the counting details for these rows. For every common cyclic shift of the color labels, round each shifted row sum at level j to the nearest multiple of width $\lambda K \sqrt{U}\, 2^{-j/8}$, where $\lambda = \tau_U / K \asymp_R \tau_U^{1/4}$. The tail bound for independent bounded centered colors gives row entropy $O_R(\exp(-c_R \lambda^2 2^{j/4}))$, by comparison with a geometric law on the integer bin. The joint entropy is at most $b = C_R U e^{-c_R \lambda^2}$. This bound for the countable vector follows by monotone finite truncation. A common fiber E has probability at least e^{-b}. The Hamming distances to its cyclic shifts are one-Lipschitz, zero with probability at least e^{-b}. Bounded differences show that at least half the points of E are within $O_R(\sqrt{U(b+1)})$ of every such shift. For such a point choose nearby representatives in these shifted fibers, and average the representative colors coordinatewise to a fractional coloring. Its row vector is bounded by the rounding widths: the inverse shifts of the row vectors at any common point of E average to zero, and the simultaneous quantizations give each comparison. The fractional coloring agrees with the original point except at $O_R(\sqrt{U(b+1)})$ columns.

We round the fractional columns to colors preserving any fixed vertices. Approximate each probability vector by dyadic weights with sum one. At each precision divide the odd-numerator weights into pairs within a column, and add one unit to one member and subtract one from the other so the numerators become even. At that precision there are at most $C_R U$ sign choices and the row entries for the choices are bounded by $B_j = C_R K 2^{-j/4}$. For completeness, on n free choices round the sign sums to nearest multiples of widths $C B_j \sqrt{n(1 + j + \log(C_R U/n))}$. Summing the entropies over $O_R(U2^j)$ rows of level j gives less than cn for small fixed c, by taking C large. A largest fiber contains two signings at Hamming distance at least $n/4$, since a set of diameter less than $n/4$ fits in a Hamming ball of radius $n/4$ and cannot have that mass. Half their difference assigns those signs with a row loss at most the chosen width. Iterate on the remaining choices. The summed width is $O_R(B_j \sqrt{U(1+j)})$, by geometrically decreasing n. For bounded n direct assignment satisfies the same bound. At a dyadic precision the actual error has an additional factor equal to the dyadic unit. Summing all precisions and passing through finite truncations gives the same row bound for the rounded colors in all rows simultaneously. All temporal row widths, and the rounding errors, are summable in j.

An output coloring differs from the original fiber point in at most $O_R(\sqrt{U(b+1)})$ positions. In the range $U \leq B_m$ one may take $b \geq 1$ for large m since $\tau_U = O_R((\log(2m))^{1/4})$. The logarithm of

the corresponding Hamming ball size is $O_R(Ue^{-c'_R\lambda^2})$, by taking A_0 large. Counting the outputs and using continuity extends the constraints to all diffusion times. We conclude that the product Brownian trajectory stays in the desired row bounds with probability at least

$$\exp(-C_R U e^{-c_R \tau_U^{1/2}}).$$

Choose an open region Ω with row bounds a fixed factor larger than those just obtained. The killed semigroup on Ω has survival $Z = \langle 1_\Omega, e^{-\mathcal{L}_D} 1_\Omega \rangle$ at least that probability. The supported test $u_U = e^{-\mathcal{L}_D/2} 1_\Omega$, extended by zero, has squared norm Z. Its normalized spectral tail at energy t is at most e^{-t}/Z, so its relative Dirichlet energy is at most $1 + \log(1/Z)$. It is class invariant since the domain, operator and initial function are class invariant. The sum of these quotients is $O_R(m)$ by the dyadic sum of $Ue^{-c_R\tau_U^{1/2}} + 1$. The propagation and character norm bounds use the same speeds $\sigma = \sqrt{\tau_U}$ as in even rank.

A.4 Finite propagation and the signed polynomial

Every dense row function has Lipschitz constant $O_R(1)$ in the scaled product metric, by the bounded group trace gradient and the row norm estimate in z coordinates. On a high band the Lipschitz constant is $O_R(\sqrt{\tau_U})$. It follows that a fixed sufficiently small closed metric neighborhood of $\operatorname{supp} f$, of width $\delta_0 > 0$, satisfies $|y_i| \le T_i/2$ for thresholds T_i of fixed order in the dense case and of order τ_U in the high bands. The thresholds can be chosen independently of the center for each m and row.

Put $k = \lceil K\sqrt{m} \rceil$ for sufficiently large fixed K and define

$$S(\lambda) = \left(\operatorname{sinc} \frac{\delta_0 \sqrt{\lambda}}{2k} \right)^{2k}.$$

It is an average of $\cos(t\sqrt{\lambda})$ over a symmetric probability measure on $[-\delta_0, \delta_0]$ by convolution of $2k$ uniform measures. Finite propagation for the wave equation [17] shows that $u = S(\mathcal{L})f$ is supported on that closed neighborhood. By the Rayleigh bound at least half the squared norm of f lies below Cm for suitable fixed C. Choose K so the sinc argument there is at most $1/2$. It follows that $\|u\|_2 \ge e^{-C_1\sqrt{m}}$. On $\lambda > Bm$, the sinc bound $|S(\lambda)| \le (2k/(\delta_0\sqrt{\lambda}))^{2k}$ is at most $e^{-C_2\sqrt{m}}$ with C_2 as large as desired by taking fixed B large. Project onto eigenvalues at most Bm and normalize this projection to obtain P_m with $\|P_m\|_2 = 1$. After scaling u by the same scalar we have

$$\|P_m - u\|_2 \ll e^{-c_R\sqrt{m}}. \tag{9}$$

The operators preserve class invariance. By the compact-group character basis, P_m has support in products $\prod_h \chi_{\pi_h}$ with $\sum_h \sigma_h^2 C(\pi_h) \le Bm$, where C is the Casimir eigenvalue. On each factor it grows quadratically in the dominant highest weight length plus a linear term. The number of nontrivial factors is $O_R(m)$, since the dense factors have that cardinality and each nontrivial high factor has cost bounded below.

For a fixed bi-invariant metric the Weyl dimension formula and the quadratic Casimir give

$$\sum_\pi (\dim \pi)^2 e^{-\sigma^2 C(\pi)} = \begin{cases} O_R(\sigma^{-C_R}), & 0 < \sigma \le 1, \\ 1 + O_R(e^{-c_R\sigma^2}), & \sigma \ge 1. \end{cases}$$

They follow by polynomial dimension and lattice-point bounds and by the positive least nonzero eigenvalue. Cauchy–Schwarz on the character coefficients and separation of variables give

$$\|P_m\|_{\mathrm{ch}}^2 \le e^{Bm} \prod_h \sum_\pi (\dim \pi)^2 e^{-\sigma_h^2 C(\pi)} = e^{O_R(m)}. \tag{10}$$

For the dense part use $\sum_{1\le h\le m}\log((m+1)/h)=O(m)$; for the high part use $\sum_U Ue^{-c_R\tau_U}=O_R(m)$.

Let

$$b_m=\sum_i (y_i/T_i)^{2q},\qquad F_m=|P_m|^2(1-b_m),\qquad V_m=|P_m|^2,\qquad q=\lceil C_q\log(2m)\rceil.$$

The number of rows is $O_R(m\log(2m))$, so $b_m=o(1)$ on the support of u for large fixed C_q. Its supremum and its character norm are $\exp(O_R(\log^2(2m)))$ because each row is linear in traces with coefficient absolute sum $O_R(\sqrt{B_m})$ and a similarly bounded constant term. Equation (9), with Cauchy–Schwarz for the difference of squared moduli, gives $\mathbb{E}b_m|P_m|^2=o(1)$ under Haar. Thus $\mathbb{E}F_m=1-o(1)$ and $\mathbb{E}V_m=1$. Pointwise $F_m\le V_m$, and $F_m\le 0$ whenever any $|y_i|>T_i$. Both polynomials have character norm $\exp(O_R(m))$ by (10) and submultiplicativity.

We also need a basis independent of the center. Both F_m and V_m lie in the span of the characters in an enlarged ellipsoid $\sum_h \sigma_h^2 C(\pi_h)\le C_R'm$ of cardinality $\exp(O_R(m))$, by the same partition bound. In a tensor product any occurring highest weight ν is dominated by the sum of the input highest weights $\lambda+\lambda'$. The Casimir is monotone in this dominance order on dominant weights, and $\sqrt{C(\lambda+\lambda')}\le\sqrt{C(\lambda)}+\sqrt{C(\lambda')}$ by the quadratic plus positive linear formula. Hence multiplying two ellipsoid-supported polynomials adds their ellipsoid square-root radii. For a monomial of b_m, its at most $2q$ trace factors cost total square-root radius $O_R(q\max_h\sigma_h)=O_R((\log(2m))^{9/8})$, which is $o(\sqrt{m})$. This proves the basis assertion and also shows that each monomial of both certificates uses $O_R(m)$ nontrivial factors.

A.5 Transfer of the lower bound

Use integers m in the large range $m\le a_0\log P$, where $a_0>0$ is a sufficiently small constant. Then $B_m<N$ for sufficiently large P. On the complete field and long progressions the character comparison gives $\mathbb{E}_bF_m\ge 3/4$ and $\mathbb{E}_bV_m=1+o(1)$. The error is uniform because $\|F_m\|_{\mathrm{ch}}+\|V_m\|_{\mathrm{ch}}\le e^{C_Rm}$ and a_0 is small.

For translates, divide each basis character product in the enlarged ellipsoid by its product dimension. Each normalized basis element has character norm one and uses $O_R(m)$ shifts. The L^2 comparison and Chebyshev bound its error by $L^{-1/4}$ off a fraction $O_R((m+1)^2L^{-1/2})$. Union over the basis and over integers $m\le a_0\log L$ costs a power-saving exceptional fraction for small a_0. Expansion of both polynomials and (10) then give the same expectation bounds. This construction is independent of g.

For the blocks from B_m to N compare their integrated $2r$th moments weighted by V_m, with $r=\lceil\log(16m)\rceil$ throughout. In the independent law V_m is independent of a later block and has mean one. The integrated block moment is a character polynomial of norm at most C_R^{2r}/U by the balanced tuple bound, and each term uses at most $2r$ shifts. The resulting supremum moment comparison loses the factor $O(rU)$ from the Parseval bound, which cancels the $1/U$ in the comparison error. After taking roots, the discrepancy summed over $O(\log P)$ blocks is negligible, by taking a_0 small. For translates apply the L^2 estimate to each normalized low basis element times the *whole integrated moment* of a later block. The error threshold is $L^{-1/4}$ times C_R^{2r}/U. The additional union over $O(\log L)$ blocks leaves a power-saving exceptional fraction. It applies simultaneously to the weighted moments for every g. From (7), the sum of the model moment roots is

$$O_R(\sqrt{r/B_m}(rB_m)^{1/(2r)})=O_R(m^{-1/2}).$$

Taking expectation with the nonnegative weight V_m and using $\mathbb{E}_bV_m=1+o(1)$ therefore bounds the middle tail expectation by $O_R(m^{-1/2})$. The terminal tail expectation under V_m is negligible

since its unweighted bound is a power saving and $\|V_m\|_\infty \le \exp(O_R(m))$. The deterministic mesh error is also negligible.

Let the tail-good event require the combined tail bound at a sufficiently large constant times $m^{-1/2}$, including pinning and interpolation. By Markov, $\mathbb{E}_b[V_m \mathbf{1}_{\text{tail bad}}] \le 1/4$. Then

$$\mathbb{E}_b[F_m \mathbf{1}_{\text{tail good}}] \ge \tfrac{1}{2}.$$

On failure of any row bound $F_m \le 0$, and its value on the joint success event is at most $V_m \le e^{C_R m}$. It follows that this joint success has probability at least $\frac{1}{2}e^{-C_R m}$. The row bounds control the endpoint and all retained blocks by Bernstein and dyadic summation. Together with the Fejér approximation and the tail this puts the path within $C_{R,H} m^{-1/2}$ of g. Taking $m = \lceil (2C_{R,H}/\epsilon)^2 \rceil$ fits the comparison range for $\epsilon \ge D/\sqrt{\log P}$ and large D. A small fixed ϵ_* ensures m is large enough for the estimates. This proves the lower bound in Theorem A.1.

A.6 An upper bound in L^1

Take $m = \lfloor k\epsilon^{-2} \rfloor$ with $k > 0$ small. Choose a smooth real multiplier ϕ supported inside $(1, 2)$ and bounded below by a positive constant on a fixed subinterval. Let T_m have Fourier multipliers $\phi(h/m)$, and take $J = 8m$ time samples $t_j = j/J$. The kernel bound $Cm/(1 + m^2 \operatorname{dist}(t, \mathbb{Z})^2)$ follows by two summations by parts, and its average absolute value over the displaced sample grid is bounded uniformly in the displacement. Hence for every periodic L^1 function (viewing functions on $[0, 1]$ this way),

$$\frac{1}{J} \sum_j |T_m f(t_j)| \ll \|f\|_1 . \tag{11}$$

Use a fixed finite η-net for the bounded endpoint W_0. If $\|K_p - g\|_1 < \epsilon$, some net point w is within η of W_0. Write

$$y_j = \sqrt{m}\, T_m(Q_p - tW_0)(t_j), \qquad z_j = \sqrt{m}\, T_m(g - tw)(t_j).$$

Then the average of $|y_j - z_j|$ is $O(\sqrt{m}(\epsilon + p^{-1/2}) + \eta)$. For the first term use (11) on $Q_p - g$. For the correction by $w - W_0$, the Fourier coefficient of t at nonzero h is $-1/(2\pi i h)$, and sampled Parseval bounds the mean square of its normalized band by $O(|w - W_0|^2)$. No endpoint of g is needed.

By the mesh Fourier coefficient identity in Section A.2, $y = TW$ where the column of T at h has entries

$$\sqrt{m}\, \phi(h/m) c_{p,h} e(ht_j)/(2\pi i h).$$

Distinct columns are orthogonal, each has bounded norm, and their square norms sum to at least $c_0 m$. Every row has square norm $O(1)$. If U restricts to the coordinates outside a set I of at most αJ deleted samples, then $\|UT\|_{2\to 2} = O(1)$ and the sum of column square norms is still at least $c_0 m/2$, uniformly for α below a sufficiently small fixed constant.

For independent bounded model coefficients and any v in the Euclidean ball of radius $C_1\sqrt{m}$, the map $x \mapsto Z = \|UTx - v\|_2$ is convex and Lipschitz with a bounded constant, treating complex coefficients as two real coordinates within one independent block. Talagrand's convex-distance inequality [6, Theorem 4.1.1] gives subgaussian tails around its median. Indeed, for any sublevel set of a convex Lipschitz function, Jensen bounds the excess above the level by the Lipschitz constant times the Euclidean distance to any convex mixture of points from the set. The distance to the mixture is at most twice the coefficient bound times the Euclidean norm of the vector of coordinate-block mismatch proportions. Taking the infimum over mixtures gives the required comparison with convex distance. The usual upper and lower median tails follow by applying the inequality to the median sublevel and to a sublevel below the median, respectively. Now $\mathbb{E}Z^2 \ge c_2 m$ from

independent mean-zero coefficients with variance one. On the other hand $\mathbb{E}Z^2 \le 2(\operatorname{med} Z)^2 + O(1)$. Thus $\mathbb{P}(Z^2 < c_3 m) \le \exp(-c_4 m)$.

On the bounded trace cube UTW lies in the Euclidean ball of radius $C_1\sqrt{m}$ for a suitable fixed C_1. Project Uz there and use a net of mesh $\omega\sqrt{m}$ contained in the ball, of size at most $(1+2C_1/\omega)^{2J}$. If $|y_j - z_j| \le \omega$ at each retained coordinate, the selected net point v has $\|UTW - v\|_2 \le C\omega\sqrt{m}$. For sufficiently large fixed C_2 the polynomial

$$F_{I,v}(x) = \left(1 - \frac{\|UTx - v\|_2^2}{C_2 m}\right)^m$$

lies in $[0,1]$ on the bounded trace cube. Its value on this event is at least $\exp(-C_3\omega^2 m)$. Its model mean is at most $\exp(-c_3 m/C_2) + \exp(-c_4 m)$. Its total degree is at most $2m$, and its weighted character norm is $\exp(O(m))$ by the total-degree estimate in Appendix .1. Section A.2 pays this comparison cost with a remainder smaller than the model bound, by taking $m \le (k/d^2)\log P$ with d large. For translates, Chebyshev and a union over all deleted masks, their nets and all positive integer values of m costs at most $(1+\log L)\exp(C_\omega k \log L/d^2)$. Choose d large, leaving an $O(L^{-c'})$ fraction exceptional. The bound therefore holds for all these candidate polynomials on one good set.

Choose ω sufficiently small that the lower certificate loses at most one fourth of the model exponent. For fixed g and rounded endpoint only one net point per mask is used. Choose α small so the sum of binomial coefficients through αJ costs less than a quarter of the resulting exponential rate. The constants in the model concentration were uniform for all smaller α. Choose η, k so the average discrepancy from (11) is at most $\alpha\omega$, then at most αJ coordinates need deletion. The finite endpoint net costs a fixed factor. Union over endpoints and masks proves the claimed upper bound after reducing ϵ_* to absorb fixed factors and the integer part. This proves the L^1 upper bound and hence the uniform upper bound in Theorem A.1.

A.7 Sharpness for uniform coverage

On a frequency band $m \le h < 2m$, take $\beta_h = d_H\varepsilon_h$ for signs $\varepsilon_h \in \{-1, 1\}$ and a fixed sufficiently small $d_H > 0$, with endpoint and other coefficients zero. There are $\exp(cm)$ signs for which the unpinned trigonometric polynomial of the normalized coefficients has supremum $O_H(m^{-1/2})$. Here is the counting argument. On a Bernstein mesh apply it to the real and imaginary parts of the unpinned sum after multiplying by m/d_H. The resulting matrix has $O(m)$ rows with bounded entries. Round each row sum to the nearest multiple of a width $\lambda\sqrt{m}$ with sufficiently large fixed λ. The subgaussian row bound and the geometric-reference entropy bound on integers give total entropy at most ηm, for arbitrarily small fixed $\eta > 0$. Some common fiber E therefore contains at least $2^m e^{-\eta m}$ signs. The Hamming distance from a uniform sign to $-E$ is one-Lipschitz in each coordinate and is zero with probability at least $e^{-\eta m}$. The bounded-difference inequality first bounds its mean by $O(\sqrt{\eta}\, m)$ and then shows that at least half of E is at distance $O(\sqrt{\eta}\, m)$ from $-E$. For such a sign x, take $y \in E$ with $-y$ this close to x. The partial signed vector $(x-y)/2$ has row sums bounded by the quantization width, and it agrees with x except at $O(\sqrt{\eta}\, m)$ zero entries. Spencer's theorem [4] fills the zeros with added row discrepancy $O(\sqrt{m})$ by padding to a bounded square matrix. The output differs from x in at most $O(\sqrt{\eta}\, m)$ places. Each output has at most the size of that Hamming ball many preimages. For small η the number of outputs is $\exp(cm)$, as claimed.

Bernstein bounds the derivative of these unpinned sums by $O_H(\sqrt{m})$; the minimum of the derivative increment bound and twice the uniform bound gives a common half-Hölder constant.

Choose d_H so the pinned paths lie in $\mathcal{H}_H$. A greedy Hamming packing leaves $\exp(c'm)$ signs separated by a positive fraction of indices. Parseval on the pinned differences gives L^2 separation $\gg_H m^{-1/2}$, hence uniform separation of that order. Each arithmetic family is supported on at most P paths. Taking m to be a sufficiently large fixed multiple of $\log P$ yields more separated centers than paths. At least one center has an empty ball of radius $c_H/\sqrt{\log P}$. This proves the asserted optimal order.

If the complete ensemble is restricted to $b \neq 0$, deleting the possible path at $b = 0$ changes the probability by $O(1/p)$. Taking the constant D larger makes this smaller than the lower bound by a power of p. The same argument applies when one removes $b = 0$ from a long progression or a translate, using its length in place of p. In rank two averaging the fixed-inverse estimate also gives the two-parameter result.

.1 Additive twists and character norms

The completed coefficients of the normalized increments are

$$W_h = p^{-1/2} \sum_{x \neq 0} \varphi_{a,b}(x) e_p(-hx).$$

For $b \neq h$, introducing the last multiplicative coordinate $(b-h)x$ gives exactly $\mathrm{Kl}_R((b-h)/a; p)$. At $b = h$ the sum factors into $R-1$ complete nonzero additive sums, giving $(-1)^{R-1} p^{-(R-1)/2}$. This singular value belongs to the trace image in both parities, as is seen by taking conjugate eigenvalue pairs and, for odd R, one eigenvalue 1. Choose any class with that trace there. At ordinary points use the semisimple Frobenius class of the normalized Kloosterman sheaf, multiplied by the central element $-I$ in even rank. This gives the desired trace and preserves Haar.

For distinct active shifts, let $F(b)$ be a product of nontrivial irreducible compact characters of those classes, with d_0 active shifts and product of dimensions D. We claim

$$\left| \sum_{b \in \mathbb{F}_p} F(b) e_p(kb) \right| \ll_R (d_0 + 1) D \sqrt{p} \qquad (k \in \mathbb{F}_p). \tag{12}$$

The normalized single-factor sheaf has arithmetic and geometric group Sp_R or SL_R, tame regular unipotent monodromy at zero and a single infinity break $1/R$, by Katz's results as recorded in [9, Prop. 4.6]. In the product over translates, inertia at $b = h$ acts nontrivially and unipotently in that factor alone. Projection onto each factor is surjective, and the normal closure of a noncentral unipotent in its connected simple group is the full factor. Hence the geometric group is the product. Applying the irreducible representations and tensoring gives rank D, no invariants, and infinity breaks at most $1/R$. For a nontrivial additive twist the break at infinity is one, since the original breaks are strictly smaller. In particular the twisted tensor has no invariants. All finite ramification is tame. The Euler characteristic on the affine line with the active punctures removed gives $\dim H^1_c \leq d_0 D$ and $H^0_c = H^2_c = 0$. Purity and the trace formula bound the sum there by $d_0 D \sqrt{p}$. The exceptional points contribute at most $d_0 D$. This proves (12). The all-trivial product has the exact full or restricted expectation one.

We use the norm $\|P\|_{\mathrm{ch}}$ of an expansion into products of characters, with absolute coefficients weighted by product dimensions. It is submultiplicative, and both the trace and its complex conjugate have norm R. For a polynomial in traces and dual traces of total degree r on d factors, with bound U on the product of compact groups,

$$\|P\|_{\mathrm{ch}} \leq U \exp(C_R r \log(eC_R(d+r)/r)) \quad (r \geq 1). \tag{13}$$

Indeed Haar character coefficients are bounded by U by orthonormality. Each possible highest weight in one factor has weight norm bounded by C_R times its trace degree, since it occurs in a tensor product of the standard representation and its dual. The total highest-weight norm is at most $C_R r$. Stars and bars count the products of possible highest weights, and the product of dimensions is at most $\exp(C_R r)$ by the Weyl dimension formula. For separate degrees r_h, the same argument bounds the norm by $U \prod_h [C_R(1+r_h)^{C_R}]$ on active factors, with factor one when $r_h = 0$. These estimates also apply to real and imaginary components expressed using the trace and its dual.

Let $m_S(k) = L^{-1} \sum_{b \in S} e_p(kb)$ and $\widehat{F}(k) = \sum_b F(b) e_p(-kb)$. For each nontrivial product F, Parseval over translates gives

$$\left\| \mathbb{E}_{b \in S+u} F(b) - \mathbb{E}_{b \in \mathbb{F}_p} F(b) \right\|_{L^2_u}^2 = \frac{1}{p^2} \sum_{k \neq 0} |\widehat{F}(k)|^2 |m_S(k)|^2.$$

Use (12) and $\sum_k |m_S(k)|^2 = p/L$, and add the error of the full average relative to the independent Haar characters. Thus for a test with at most d active shifts in each character term,

$$\left\| \mathbb{E}_{b \in S+u} P(W) - \mathbb{E} P(X) \right\|_{L^2_u} \ll_R (d+1) L^{-1/2} \left\| P \right\|_{\mathrm{ch}}. \tag{14}$$

For a progression of step $d \neq 0$ and length L, the geometric sum for $k \neq 0$ gives

$$|m_I(k)| \ll \min(1, p/(L \operatorname{dist}(kd, p\mathbb{Z}))).$$

It follows that $\sum_k |m_I(k)| \ll p \log p / L$. Fourier inversion gives the pointwise bound

$$|\mathbb{E}_{b \in I} P(W) - \mathbb{E} P(X)| \ll_R (d+1)(\sqrt{p} \log p / L) \left\| P \right\|_{\mathrm{ch}}. \tag{15}$$

Both extend to series convergent absolutely in this norm.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including development and checking of the wave certificate, and with manuscript preparation. No code or data are required for the proof.

A Exact Kloosterman shrinking targets at prime powers

We count paths with completed hyper-Kloosterman coefficients over p^n in uniform balls about supported half-Hölder centers. At a fixed residue the lower probability is $\exp(-O(\epsilon^{-2}))$ down to radius of order $J^{-1/2}$, where $J = \min(n, \sqrt{n \log p})$. Complete families attain radius of order $(\log p)^{-1/2}$ at every tame rank and depth. At sufficient depth a single progression of missing frequencies forces the sharp worst-case coverage order $(\log p)/p$ in the circle case, and order $1/p$ in finite L^r norms. We construct a wave-localized test using coupled torus diffusions, and a commuting single-site refresh filter for the finite labels.

Keywords hyper-Kloosterman paths, prime powers, shrinking targets, finite propagation.
MSC 2020 11L05, 11L07, 60G17, 58J45.

A.1 Statements

Fix $R \geq 2$, an odd prime $p > R$, $n \geq 2$ and $q = p^n$. For a unit c put

$$\mathrm{Kl}_j(c; q) = q^{-(j-1)/2} \sum_{x_1 \cdots x_j = c} e_q(x_1 + \cdots + x_j), \qquad e_q(v) = \exp(2\pi i v/q),$$

summing over units. Use zero at nonunits. Given units a, b, form the path $K_q(a, b)$ from the increments

$$q^{-1/2}\, \mathrm{Kl}_{R-1}((ax)^{-1}; q) e_q(bx)$$

in increasing order over the units x, with equal time steps. The complete ensemble samples b uniformly among all units, keeping a fixed. The residue ensemble conditions on $b = b_0 \pmod p$, where $b_0 \neq 0$. We put $d = \gcd(R, p-1)$, $L = (p-1)/d$ and $J = \min(n, \sqrt{n \log p})$.

Here are the coefficient supports needed to state the result. Let $\mathcal{T}_d$ be the Haar torus in $\mathbb{T}^d$ with relations $\prod_j Z_j^{m_j} = 1$ whenever $\sum_j m_j \zeta_d^j = 0$, where ζ_d is a primitive complex dth root and $m_j \in \mathbb{Z}$. On a residue ensemble, write $c_h = (b_0 - h)/(a \bmod p)$. The coefficient support at index h is $\{0\}$ if c_h is not a nonzero dth power. On the other indices it is the image of $\mathcal{T}_d$ by the following functions. Write χ for the quadratic character of $\mathbb{F}_p^\times$, $\varepsilon_p = 1$ or i according as $p \equiv 1$ or $3 \pmod 4$, and $\kappa = \varepsilon_p^{R-1} \chi(2^{-(R-1)} R)$.

case	active coefficient function
n even	$\sum_j Z_j$
n, R odd	$\kappa \sum_j Z_j$
n odd, R even, L odd	$\kappa \sum_j (-1)^j Z_j$
n odd, R even, L even	$\kappa\, s_h \sum_j Z_j$.

Here $s_h = 1$ on $(2d)$th powers and -1 on the remaining dth powers. The complete independent model has active probability $1/d$ at each index, with independent fair signs when the last row applies, and independent Haar angles. Its coefficient support is the union of those possible images, and includes zero as an inactive value when $d \geq 2$. When $d = 1$ the complete coefficient support is the unit circle and at a residue the only omitted progression is $h \equiv b_0 \pmod p$.

For a center $g(0) = 0$ let $\beta_0 = g(1)$ and $\beta_h = 2\pi i h\, \widehat{g - tg(1)}(h)$ for $h \neq 0$, using Fourier convention $e(t) = \exp(2\pi i t)$. Call it supported in an ensemble if the corresponding supports contain every β_h. All lower bounds are for supported centers with half-Hölder seminorm at most H, for fixed $H > 0$.

Theorem A.1. *Fix R, H, and choose constants $D, C, c, \epsilon_* > 0$ depending at most on them. The following hold uniformly in the fixed unit a, for sufficiently large J in the residue comparison and sufficiently large p in the complete comparison.*

1. *In a residue ensemble,*

$$\mathbb{P}\{\|K_q - g\|_\infty < \epsilon\} \geq \exp(-C\epsilon^{-2}) \qquad (D/\sqrt{J} \leq \epsilon \leq \epsilon_*).$$

   For the lower bound, compatibility of β_h is needed only through $|h| \leq \lceil C\epsilon^{-2} \rceil$.

2. *In the complete ensemble, for $D/\sqrt{\log p} \leq \epsilon \leq \epsilon_*$ the upper bound is $\exp(-c\epsilon^{-2})$ for every L^1 center, even in the L^1 norm. The matching uniform lower holds at every tame rank and every depth $n \geq 2$.*

3. *In residue ensembles the L^1 upper holds at radii $\epsilon \geq D/\sqrt{J}$ if $d = 1$. For arbitrary fixed p it holds at those radii in a sufficiently small fixed interval of ϵ, with constants allowed to depend on p.*

Theorem A.2. *Suppose $d = 1$. In the complete ensemble the uniform lower of Theorem A.1 holds down to $D_H \log p/p$ if $J \geq C_H p^2/(\log p)^2$. For each fixed $1 \leq r < \infty$ the corresponding L^r lower reaches $D_{H,r}/p$ if $J \geq C_{H,r} p^2$. Both ranges have the matching exponential upper. The coverage orders are sharp for H above a fixed constant: some supported center has an empty ball of radius $c \log p/p$ in the uniform norm for every $n \geq 2$, and one of radius c/p in L^1 for every $n \geq 3$, for sufficiently large p.*

At odd depths there is a second sharp coverage regime. If R is even, d is not a power of two, and L is even, some supported center has an empty L^1 ball of radius $c_H/\sqrt{\log p}$ for sufficiently large fixed H. One example is $R = 6$ and $p \equiv 1 \pmod{12}$. Theorem A.1 gives the matching coverage order in all the indicated norms. We prove this in Section A.6.

The qualitative rank-two path laws were obtained by Ricotta and Royer, Ricotta, Royer and Shparlinski, and Milievi and Zhang [10, 11, 12]. Milievi, Robinson and Shupe [14] give finer alignment estimates for rank-two shifted products. Tame stationary phase and scalar cyclotomic trace models are classical [23, 24, 25]. The new ingredients are a coupling whose finite color average of the trace has an exact deterministic drift, sparse trajectory discrepancy for label jumps, and a pair of commuting filters with controlled support. The missing progression also yields a deterministic obstruction to uniform coverage.

A.2 Completion, angle modes and comparison

Extend the increments by zero at nonunits, and let Q_q be the path on the time mesh j/q. A unit's original time differs from its mesh time by $O(1/q)$, and the summand is bounded by a rank constant, so $\|Q_q - K_q\|_\infty = O_R(q^{-1/2})$. The completed coefficient W_h is $\mathrm{Kl}_R((b-h)/a; q)$ for $b-h$ a unit, by rescaling the last linear variable. If $b - h$ is a nonunit, the derivative in that variable is a unit in the inverse part, and summing the last digit gives zero. At mesh vertices the Fourier representation and coefficient identity are

$$Q_q(t) = tW_0 + \sum_{0<|h|\leq(q-1)/2} \frac{W_h(e(ht) - 1)}{q(1 - e(-h/q))}, \tag{16}$$

$$2\pi i h\, \widehat{Q_q - tW_0}(h) = e(h/(2q)) \operatorname{sinc}(\pi h/q) W_{h \bmod q} \qquad (h \neq 0). \tag{17}$$

The first follows by summing Fourier inversion and the second by integrating the piecewise constant derivative. The second applies to all integer frequencies. Keeping $|h| \leq N$ and using the continuous multiplier $(2\pi i h)^{-1}$ instead of the mesh multiplier costs $O_R(N/q)$ including interpolation.

For a unit c, the critical points of the complete phase $\sum_{i=1}^{R-1} x_i + c/\prod_i x_i$ have all coordinates equal to a root $s^R = c$. At a critical point the Hessian is $s^{-1}(I + \mathbf{1}\mathbf{1}^{\mathsf{T}})$, of determinant $Rs^{-(R-1)}$, a unit. The even-depth last half digits give the contribution $e_q(Rs)$ per root. At odd depth an additional quadratic Gauss sum for the half-Hessian gives $\kappa\chi(s)^{R-1} e_q(Rs)$ per root, by diagonalizing that nonsingular form. These usual tame computations follow by expanding at the lifted critical points and summing the top digits, as in [23, 24]. The roots exist exactly on the nonzero dth-power label. They are $s\zeta^j$ in a fixed indexing, with ζ the primitive dth root lifted to $\mathbb{Z}_p$. If L is odd and a square branch is needed, half of the roots have that property. If L is even and R even, all roots have the same sign. Writing $s = g^i$ and $c = g^{dj}$ in a multiplicative generator gives $j \equiv (R/d)i \pmod{L}$. Since $\gcd(R/d, L) = 1$, their parity agrees and gives exactly the stated sign label. For

each branch the tuple $(e_q(Rs\zeta^j))_j$ belongs to $\mathcal{T}_d$. The Haar coordinate characters are distinct and nontrivial. Thus active coefficients are centered with second absolute moment d and bound R, and the complete model has second absolute moment one. The cyclotomic torus is the image of a $\varphi(d)$-dimensional torus by integral Laurent powers [25].

We use the following comparison of sparse tests, with proofs of the phase bounds in Appendix .1. A mode at an active coordinate is $e_q(Rs\gamma)$, where $\gamma = \sum_j m_j\zeta^j$ and the integers m_j specify its torus character. Its Haar mean is zero unless $\gamma = 0$. An unused independent torus angle may be adjoined at an inactive coordinate. Finite labels on nonzero residues are multiplicative classes of order d or $2d$. A finite product test expands in gated characters, with coefficient norm given by the sum of absolute coefficients. We use constant mode 1 on a coordinate without requiring any gate there. The norm is submultiplicative, and the trace and its conjugate have norm $O_R(1)$.

1. In complete averaging, with distinct shifts modulo p, any such monomial with t participating shifts has discrepancy $O_R(p^{-1/2}e^{C_R t})$ whenever its nonzero γ are units. In particular, for large p this applies to modes of size a fixed power of $\log p$. For rank two any individual cosine index of absolute value strictly below p is also allowed. At $d = 1$ fill a singular phase arbitrarily with a circle point. Its added error in a test with t shifts is $O(t/p)$.

2. For a fixed residue, take $N = \lfloor e^J/16 \rfloor$. A test through $|h| \le N$ using $O_R(m)$ active shifts in each term, with mode indices bounded by $O_R(m^C)$ and norm at most $e^{C'm}$, has discrepancy at most e^{-Am} for any fixed A when $m \le a_{R,C,C',A}J$ and m is sufficiently large. In this estimate the labels are deterministic. The complete counterpart through $N = \lfloor p^{1/8} \rfloor$ has power saving if $m \le a \log p$ for small a and all the modes are permitted in part 1.

Here and below constants in a budget $m \le aJ$ may also depend on H after the analytic test is chosen. A bounded polynomial in traces and conjugates of total degree k on $O(k)$ shifts has norm $\exp(O_R(k))$. To see this for the varying-label law, expand on the product of the tori and a finite label set, which can be Fourier expanded as a finite cyclic set. A nontrivial mode evaluates to zero on an inactive factor after averaging its auxiliary angle. At singular factors retain a separate inactive symbol. Coefficients are bounded by orthogonality, the total torus degree is $O_R(k)$, and only $O(k)$ labels can be present. Counting modes with stars and bars gives the bound, also when a separate gate in the expanded monomial is evaluated as a character indicator. For fixed residue the same argument uses only torus frequencies. These facts also hold for ordinary real trace cosines in rank two.

A.3 A killed-heat construction on the angle space

We first construct tests on products of tori for fixed residue labels, and on product circles for the complete $d = 1$ law. Then we add the finite label dynamics for every complete mixed model. The analytic goal at an integer scale m is an L^2-normalized function supported where the retained Fourier sample rows are small, whose relative Dirichlet quotient is $O_{R,H}(m)$.

Separate the dense frequencies $|h| \le m$ from dyadic bands of size U through the first dyadic endpoint $B_m \asymp C_* m \log(2m)$. A band is sampled on $O(U)$ equally spaced times fine enough for the Bernstein maximum bound. Its rows are the real and imaginary parts of the unpinned polynomial times $\sqrt{U}$. For the dense block subtract a fixed finite approximation to the center before sampling and use the factor $\sqrt{m}$, including the real and imaginary endpoint errors. In high bands use the scale $\tau_U = A_0(U/m)^{1/4}$ with fixed large A_0 and angular speed $\sigma_h = \sqrt{\tau_U}$. Their row bounds $O_R(\tau_U)$ imply pinned norm $O_R(m^{-1/2})$ for their sum. An inactive factor in a conditioned ensemble is a point and contributes zero to the rows.

A.3.1 A dense torus test by coupled Brownian packets

Choose a smooth even multiplier ϕ supported on $[-1,1]$, between zero and one, with $\phi(v) = 1 - cv^2$ near zero and $\phi(v) < 1$ off zero. Put $\rho_h = \phi(h/m)$, $w_h = \sqrt{1-\rho_h}$ for $h \neq 0$, and $\rho_0 = 1 - m^{-1}$, $w_0 = m^{-1/2}$. For nonzero dense indices $w_h \asymp |h|/m$. Set $\sigma_h = a_1|h|/m$ off zero and $\sigma_0 = a_1 m^{-1/2}$ with a sufficiently small fixed $a_1 > 0$. On the flat torus choose a metric invariant under the cyclic action rotating the cyclotomic coordinates. Their characters have a common positive Laplacian eigenvalue λ. Write

$$\Psi(u) = d^{-1} \sum_{j=0}^{d-1} \Re Z_j(u).$$

It has mean zero, maximum one at the identity, and negative definite Hessian there because the characters span the torus. In the orbit of u under cyclic rotations and inversion the mean of each coordinate character is $\Psi(u)$.

For any $0 \le \rho < 1$ we construct four packets, to be used with $\rho = \rho_h$. Put $w = \sqrt{1-\rho}$. Near $\rho = 1$ choose a reference point on a fixed regular ray with $\Psi = \rho$, and take four copies. Its gradient has size comparable to w. Away from one choose regular points u, v such that $\Psi(u) + \Psi(v) = 2\rho$, and reference packets (u, u, v, v). Such choices exist because the range of Ψ contains $[-c_R, 1]$ and its critical values are finite. Finiteness follows from Sard and semialgebraicity on the algebraic torus. The inverse charts can be chosen uniformly by a finite cover of any compact interval away from one. Near one rescale angles by w and $1 - \Psi$ by w^2 to obtain uniform charts from the positive quadratic limit. Start the four packets U_j in small coordinate boxes of the hypersurface

$$\sum_{j=1}^{4} \Psi(U_j) = 4\rho.$$

A smooth free-coordinate box law on this hypersurface has each packet marginal density $O_R(w^{-k})$ relative to Haar, $k = \varphi(d)$. Indeed the projection onto any full packet is a submersion by a nonzero derivative in another packet. In the rescaled charts all relevant derivative ratios and inverse Jacobians are uniformly bounded. Integrating over the remaining free variables gives the bound in the rescaled box, and rescaling gives w^{-k}. Choose the initial boxes inside smaller neighborhoods of the references, with room for motions of order w.

Let $g_j = \nabla\Psi(U_j)$, $r_j = |g_j|$ and $v_j = g_j/r_j$. Near reference lengths (r, r, s, s) there are smooth planar unit vectors e_j with $\sum_j r_j e_j = 0$. To see this, fix a diagonal of length $\min(r, s)$ at the reference and use two triangles with opposite diagonal vectors. The triangle inequalities remain strict in a small neighborhood, including after rescaling near one, and the cosine rule gives the vectors smoothly. With shared planar Brownian motion B and independent flat Brownian motions B_j, until the first chart exit use

$$dU_j = \sqrt{2}\,\sigma_h [v_j (e_j \cdot dB) + (I - v_j v_j^{\mathsf{T}})\, dB_j].$$

Each packet has marginal quadratic variation $2\sigma_h^2 I\, dt$ and zero coordinate drift. After exit extend it by Brownian increments of the same speed. Lévy's characterization, in the filtration including the initial packets, gives the exact flat Brownian marginal conditional on its initial point. While the packets remain in the chart, Itô's formula gives a zero summed martingale term by the closed quadrilateral. Since Ψ has eigenvalue λ, the aggregate satisfies

$$\frac{1}{4} \sum_j \Psi(U_j(v)) = \rho\, e^{-\lambda\sigma_h^2 v} \qquad (0 \le v \le 1).$$

For small a_1, the event of no exit and quarter-Hölder bounds $O_R(w)$ for each packet displacement has probability bounded below by a positive constant. Use the Brownian marginals, the initial chart slack of order w, and the maximal and dyadic increment bounds of a Brownian motion of speed $\sigma_h \ll_R a_1 w$.

At an active frequency choose z_h^0 with trace β_h under its specified active function. The colors are the four packets with their signed cyclic orbits translated by z_h^0. Each color marginal is Brownian of speed σ_h with initial density $O_R(w_h^{-k})$, because the orbit maps and translation are isometries. On the good event their equal-color mean trace is

$$\rho_h e^{-\lambda \sigma_h^2 v} \beta_h.$$

Each color's residual from this mean has size $O_R(w_h)$ and quarter-Hölder increments in v with constant $O_R(w_h)$. Near the identity this follows from the chart size and trace Lipschitz bound. Elsewhere w_h is bounded below and the compact bounds suffice.

Take the packets independently across the dense frequencies. Their joint good event costs $\exp(-O_R(m))$. Conditional on their trajectories use initially fractional equal color weights. On the fine Fourier sampling grid, and for the parent-to-child increments on the dyadic diffusion-time grid at level j, the centered normalized rows have $O(m2^j)$ entries across rows of size at most $C_R m^{-1/2} 2^{-j/4}$ per variable, since $\sqrt{m} w_h/|h| = O(m^{-1/2})$. Lemma A.3 below rounds the weights to one color per frequency with row discrepancies $O_R(2^{-j/4}\sqrt{1+j})$. The sum over j converges. Thus independent uniform colors give at least $\exp(-O_R(m))$ probability of simultaneous row success. The normalized endpoint row is bounded directly as $\sqrt{m} w_0 = 1$.

At time v the mean pinned path approximates g with error $O_{R,H}(m^{-1/2})$. The multiplier $\phi(h/m) e^{-\lambda a_1^2 (h/m)^2 v}$ has a uniformly smooth compact profile for $v \in [0,1]$ and its periodic kernel has integral one and circular half-moment $O_R(m^{-1/2})$. The endpoint change has that same bound. In forming the slabs use the fixed center with coefficients $\rho_h \beta_h$, corresponding to $v = 0$. The difference of the two unpinned nonzero-frequency polynomials has mean zero. Its supremum is at most twice that of its pinned difference, so the fixed slabs retain the successful trajectories after increasing the widths by a fixed H-dependent amount. For an inactive coefficient use the zero center and no motion.

For each frequency, unconditionally the uniform color choice has the Brownian path law with initial density bounded by $C_R w_h^{-k}$ relative to stationary Brownian. Product domination costs $\exp(O_R(m))$ since $\sum_{h \le m} \log(m/h) = O(m)$. Hence independent stationary torus Brownian motion survives throughout these fixed dense sample slabs with probability at least $\exp(-C_{R,H} m)$. On a slightly larger open slab region Ω take the killed heat test

$$u = e^{-\mathcal{L}_\Omega/2} 1_\Omega.$$

Its squared norm is the stationary survival at time one. If that norm is at least e^{-B}, the normalized spectral energy tail is at most e^{B-t} at energy t. The relative Dirichlet quotient is therefore at most $B+1$. Zero extension gives a supported element of the global form domain.

A.3.2 Trajectory discrepancy and high bands

Lemma A.3. *Suppose there are $O(U)$ columns, each with a bounded number of colors, and at most $CU2^j$ real rows at level $j \ge 0$, with all entries bounded by B_j. A fractional color assignment can be rounded to colors with row discrepancies at most $C' B_j \sqrt{U(1+j)}$, preserving fixed vertices. If the color entries have mean zero in each column and row, and $B_j = K 2^{-j/4}$, independent uniform colors have probability at least $\exp(-C' U e^{-c'\lambda^2})$ of row bounds $C' \lambda K \sqrt{U}\, 2^{-j/8}$ simultaneously, for λ sufficiently large in the range $U e^{-c'\lambda^2} \ge 1$.*

Proof. The proof uses entropy partial coloring in the spirit of Spencer [4] and the dyadic multicolor reduction [28]. For the rounding it suffices to sign $O(U)$ pairs of colors by pairing odd numerators in dyadic approximations of each fractional vector. For n' free pairs, round the sign sums at level j to nearest multiples of $CB_j\sqrt{n'(1+j+\log(CU/n'))}$. Hoeffding's inequality and the entropy bound from a geometric law on integers give total entropy less than cn' for large C and small fixed c. A common fiber contains two signs at Hamming separation at least $n'/4$, by comparison with the size of a Hamming ball. Half their difference assigns those signs. Iterate on the remaining pairs to obtain the row bound $O(B_j\sqrt{U(1+j)})$ by geometric summation. At each dyadic precision the cost carries the dyadic unit. Sum over precisions and take a limit in the finite set of final assignments. The countable row argument follows either by finite truncation followed by compactness or by the summable entropy estimates.

For the count, let the number of colors be fixed and pad or repeat labels as needed when applying this to a fixed color count. Simultaneously round all row sums after all common cyclic shifts of the colors at the stated widths, with symmetric bins. The entropy of the joint vector is at most $b = CUe^{-c\lambda^2}$, since level j contributes at most $CU2^j e^{-c\lambda^2 2^{j/4}}$. A common fiber has mass at least e^{-b}. By bounded differences for the distance to each shifted fiber, at least half its points are within $O(\sqrt{U(b+1)})$ of every shifted fiber. Average representatives from these shifted fibers around such a point. The fractional average agrees with it outside that many coordinates and its row sum is within the quantization widths of zero. This holds because all inverse cyclic shifts of the row sums at a common fiber point have zero average, and the fiber shares all those quantizations. Round by the first part. The added row discrepancies fit the stated widths. The number of preimages per final coloring is at most the Hamming ball size, with logarithm $O(Ue^{-c'\lambda^2})$ after decreasing c' in the relevant range. This proves the count. □

On an active high-band torus factor choose a fixed cyclic group of translations on which every cyclotomic coordinate character is nontrivial. To obtain it use a large prime ℓ with the character vectors nonzero modulo ℓ and a cocharacter outside the union of the finitely many corresponding hyperplanes. Such a cyclic group can be fixed at each rank and divisor d. Translation preserves the stationary Brownian law, and the average of the weighted trace over the colors is zero. The trace martingale bracket on a time interval of length Δ is $O_R(\sigma_h^2\Delta)$ and drift has the same bound, by the bounded trace gradients and Laplacians. Its quarter-Hölder constant exceeds $K = C_R\tau_U^{3/4}$ with probability $O_R(e^{-c_R\tau_U^{1/2}})$. To check this, for $\Delta \geq \sigma_h^{-2}$ use boundedness, and at shorter adjacent dyadic intervals use a martingale exponential bound $C_R e^{-c_R K^2\sigma_h^{-2}\Delta^{-1/2}}$, which is summable over the dyadic increments after paying their count. The same drift bound is absorbed at short scales.

Condition simultaneously on the bound for all finitely many translated characters, so the condition is color-invariant. Conditioning independently at each frequency pays $\exp(-C_R U e^{-c_R\tau_U^{1/2}})$, then use independent uniform color translations on the conditioned trajectories. On the Fourier sample grid and dyadic diffusion-time increments, after multiplying the normalized rows by $\sqrt{U}$, Lemma A.3 applies with $\lambda \asymp \tau_U/K \asymp_R \tau_U^{1/4}$. The resulting row bounds sum in time and give survival in the high slabs with probability at least $\exp(-C_R U e^{-c_R\tau_U^{1/2}})$. In the retained range $U \leq B_m$ the range condition in that lemma holds for large m, since $\tau_U \ll_R (\log(2m))^{1/4}$. The killed test thus costs relative energy $O_R(1 + Ue^{-c_R\tau_U^{1/2}})$. Its sum over bands is $O_R(m)$.

A.3.3 The complete label law

In the complete product model take a nominal label and an independent Haar torus at every frequency, retaining an unused angle at an inactive label. The nominal label probabilities are

bounded below by a rank constant on their finite support. In the dense frequencies, pick one label at each frequency whose trace support contains its center coefficient. The dense test lives on this label pattern, with the angles evolving by the preceding torus construction. Freeze the dense labels in the generator. The relative initial energy is unchanged, including when the test is constant as a function of the unused angle at an inactive label. All retained torus factors have positive angle speed.

On a high band at scale U, let the torus move at speed $\sqrt{\tau_U}$ and let each label refresh independently at rate τ_U to an independent sample of its nominal stationary law. The joint process is stationary and reversible, with continuous and label generators acting on different tensor factors. The finite torus translations used above give centered trace colors for every label, even during a refresh. Require the quarter-Hölder bound simultaneously on all the finitely many translated coordinate-character trajectories. This event is invariant under the color translations and has the same probability estimates as before by a finite union. It is independent of the label process.

Write $\tau = \tau_U$ and $K = C_R\tau^{3/4}$. Stop the temporal dyadic grid at $j_* = \lceil 3\log_2 U \rceil$. With probability at least $1-O(U^{-2})$ every adjacent dyadic interval of level $j \le j_*$ has at most $C_R(U\tau 2^{-j} + \log U)$ refreshes in the entire band. Indeed the count has Poisson mean $O(U\tau 2^{-j})$. The exponential Markov bound with a large multiple of $\log U$ handles the union over $O(U^3)$ intervals. This event is independent of the angular paths. At all dyadic sample times split each trace increment, for angle z and label l, as

$$[f_{l(t)}(z(t)) - f_{l(t)}(z(s))] + [f_{l(t)}(z(s)) - f_{l(s)}(z(s))].$$

The first is its continuous increment with final label held fixed. The second is supported on columns with at least one refresh. In the unnormalized row accounting of Lemma A.3, the continuous entries are bounded by $C_R K 2^{-j/4}$, and the jump entries are bounded by C_R on at most

$$D_j = \min(C_R U, C_R(U\tau 2^{-j} + \log U))$$

columns. The Fourier rows at time zero have bounded entries. For any fixed label all colors have mean trace zero, so both split rows are centered. Include time one either as the first increment or the corresponding endpoint row.

In the common shifted-fiber proof round both sets of increments at widths $C_R\tau\sqrt{U}\,2^{-j/8}$ with centered bins. The continuous entropy bound is unchanged. For jump rows, the squared width divided by the Hoeffding variance bound is at least a constant times

$$\min\left(\tau 2^{3j/4},\ \tau^2 U\, 2^{-j/4}/\log U\right).$$

The latter branch is at least $c\,\tau^2 U^{1/4}/\log U$ through j_*. It pays the total row count in this range, while the former branch gives an entropy sum $O_R(Ue^{-c_R\sqrt{\tau}})$ by dyadic summation. Thus both sets of rows and the fixed number of color shifts have that same total entropy.

In the partial-coloring step for a jump row with n' free sign choices, use width

$$C_R\sqrt{\min(n', D_j)(1 + j + \log(C_R U/n'))}.$$

The constants account for the bounded number of color pairs at a column. The entropy accounting works since the actual variance proxy uses only those nonzero entries. As the free set shrinks geometrically, the total row widths before normalization are bounded by

$$C_R\sqrt{D_j}\,(1 + \log(C_R U/D_j))\sqrt{1 + j + \log(C_R U/D_j)}.$$

Above D_j this follows by counting the geometric stages, and below D_j by the convergent square-root sum. Since $D_j \ge c_R U 2^{-j}$, the sum of these bounds over $j \le j_*$ divided by $\sqrt{U}$ is $O_R(\sqrt{\tau} +$

$(\log U)^3/\sqrt{U})$. The continuous rounding cost is $O_R(K)$ by the original rounding argument. Both fit the allowed $O_R(\tau)$ row size. The dyadic precision factors for fractional colors sum as in Lemma A.3. Counting the repaired color vectors by Hamming multiplicity gives probability at least $\exp(-C_R U e^{-c_R\sqrt{\tau}})$. Between the finest grid times there are at most $C_R \log U$ refreshes in a cell, and bounded angular oscillation by the quarter-Hölder assumption, so the remaining normalized row error is

$$O_R\big((\log U)/\sqrt{U} + \sqrt{U} K U^{-3/4}\big).$$

Thus the row bounds hold throughout the time interval, including at jumps. On multiplying by the probabilities of the angle and Poisson conditions we still obtain survival at least $\exp(-C_R U e^{-c_R\sqrt{\tau}})$.

The killed semigroup of this joint jump and diffusion process gives a supported test of relative energy $O_R(1 + U e^{-c_R\sqrt{\tau_U}})$ per high band, by the same spectral tail calculation. Its Dirichlet form is the sum of the continuous form and the label-refresh form, with zero extension from the allowed states. The high quotients sum to $O_R(m)$. The only label modes of zero energy outside the constants occur in the dense factors.

A.4 Wave localization and transfer

For both comparison models, multiply the dense and high tests, normalize the product f to unit L^2 norm, and let $\mathcal{L}$ be the nonnegative Laplacian with the indicated product speeds and the finite label operator. The product form gives relative energy $O_{R,H}(m)$. With labels fixed, dense sample rows have bounded Lipschitz constants in the scaled angle metric because $\sigma_h \ll_R w_h = O(|h|/m)$ off zero and the phase-to-trace maps have bounded derivatives. The normalized endpoint has the same property. High sample rows are $O_R(\sqrt{\tau_U})$-Lipschitz, by the row Euclidean bound on weights. Thus after fixed enlargements of the original support we can use thresholds T_i with $T_i = O_{R,H}(1)$ in the dense part and $T_i = O_R(\tau_U)$ in the high part.

For a complete mixed model, write $\mathcal{L} = \mathcal{L}_c + \mathcal{L}_g$ for the commuting angle and label parts. In a residue model take $\mathcal{L}_g = 0$. Denote the label conditional expectation at site h by E_h. Then $\mathcal{L}_g = \sum_{|h|>m} \tau_h (I - E_h)$, where τ_h is the scale of its band at sites subject to refresh and is zero otherwise. Let

$$k = \lceil K\sqrt{m} \rceil, \qquad A = Dm + \sum_{|h|>m} \tau_h, \qquad \ell = \lfloor km/A \rfloor.$$

The sum is zero when labels are frozen throughout. Take D large compared with the relative energy constant and K fixed. We have $\ell \gg_{R,H} \sqrt{m}/(\log(2m))^{5/4}$ in the retained frequency window. Consider the filter

$$(I - \mathcal{L}_g/A)^k \left(\operatorname{sinc} \frac{\delta_m \sqrt{\mathcal{L}_c}}{2\ell}\right)^{2\ell}, \qquad \delta_m = c_0 \ell/\sqrt{m}$$

with sufficiently small fixed c_0 and large m. The label factor is a positive combination of at most k successive single-site refreshes, since $I - \mathcal{L}_g/A$ is a convex combination of the identity and the E_h. Its support changes at most k labels. A single change moves any normalized high row at scale U by $O_R(U^{-1/2})$, and the dense rows are unchanged. Hence the label factor enlarges the high row support by $O_R(k/\sqrt{U}) = O_R(1)$. The angle multiplier is a cosine transform of a probability measure supported on $[-\delta_m, \delta_m]$. The finite propagation property of the product wave equation [26, Theorem 3.4] enlarges the angles by at most this distance while leaving the labels fixed. Equivalently it follows from the Davies–Gaffney estimate of the scaled flat torus Dirichlet form. The combined filter of f is therefore supported in the widened row slabs $|y_i| \le T_i/2$ after choosing T_i by a fixed enlargement of the initial dense and high scales.

At least half the initial squared mass is in joint energies $\lambda_c + \lambda_g \le Cm$ by the relative energy bound. On this region the sinc factor is at least $e^{-C'\ell}$ for small fixed c_0, while $\lambda_g/A \le 1/2$ and $(1-\lambda_g/A)^k \ge e^{-C'\ell}$ after changing C'. For total energy greater than Bm one of the two energies exceeds $Bm/2$. The corresponding factor is at most $e^{-Bkm/(2A)}$ for the label part, or at most $(2\ell/(\delta_m\sqrt{Bm/2}))^{2\ell}$ for the angle part. Both factors have modulus at most one. With fixed B sufficiently large these estimates give relative error $e^{-c\ell}$ after projecting the filtered function to joint spectral energy at most Bm and normalizing to obtain P_m. Write u for the filtered function with the same rescaling. Then $\|P_m - u\|_2 \ll e^{-c\ell}$, which dominates any factor $\exp(O(\log^2 m))$.

The joint basis is the product of flat torus characters and orthonormal finite label functions. The constant label function has norm one. Every other label basis function has a bounded expansion in the finite gated symbols, uniformly at fixed rank since each nominal label has positive mass bounded below. Counting lattice points gives polynomial growth for the number of torus modes below an energy cutoff. On high factors the first nonconstant eigenvalue of either part is bounded below by $c_R\tau_h$. There are only $O(m)$ dense factors with zero label energy. Thus Cauchy–Schwarz in the joint basis and the exponential partition bound show that the squared arithmetic norm of P_m is at most

$$e^{Bm}\prod_h\left(\sum_\nu w_\nu^2 e^{-\lambda_{h,\nu}}\right) = \exp(O_{R,H}(m)).$$

Here w_ν bounds its one-factor absolute norm, and $\lambda_{h,\nu}$ is its one-factor energy. At dense sites the logarithms are $O_R(1+\log(1/w_h))$ by the lattice sum and bounded label multiplicity. At high sites each factor is $1 + O_R(e^{-c_R\tau_h})$. Both sums of logarithms are $O_R(m)$ by the factorial and dyadic estimates. Each term has only $O(m)$ nontrivial sites because every high site costs energy bounded below, and the one-site torus indices are bounded by $O_R(m^{3/2})$ from the smallest angle speed.

Let

$$V_m = |P_m|^2, \qquad b_m = \sum_i (y_i/T_i)^{2s_0}, \qquad F_m = V_m(1-b_m), \qquad s_0 = \lceil C_0\log(2m)\rceil.$$

The norm and supremum of the penalty are at most $\exp(O_{R,H}(\log^2(2m)))$ by the linear trace representation of each row. On the support of u the penalty is $o(1)$ for large fixed C_0, since there are $O(m\log(2m))$ rows. The projection error beats the supremum of this penalty, so $\mathbb{E}V_m = 1$ and $\mathbb{E}F_m = 1-o(1)$ in the independent model. Both tests have arithmetic norm at most $\exp(O_{R,H}(m))$. They have $O_{R,H}(m)$ frequencies in each term after expansion. Always $F_m \le V_m$, and $F_m \le 0$ if any row bound $|y_i| \le T_i$ fails.

The arithmetic comparison applies directly to the angle and finite label products. At an inactive shift adjoin a Haar angle so that testing a nontrivial mode there has mean zero. When the complete nominal law has an inactive label, use it also for singular shifts. In the complete $d = 1$ family a singular low coefficient can be filled with an independent arbitrary fixed-model angular point to define the test. In the lower bound discard parameter residues with any singular retained frequency $|h| \le B_m$. Their total weighted mass is at most $\exp(Cm)O(B_m/p) = o(1)$ when $m \le a\log p$ for small a. At any other frequency use the actual zero coefficient when singular. In a participating moment tuple it contributes exactly the singular-trace exception already covered by the comparison estimate. Residue-conditioned phases are lifted without any dummy at active indices, and inactive slots are omitted.

Choose $m \le aJ$ conditionally or $m \le a\log p$ for the complete comparison, where a is sufficiently small after fixing the analytic constants. By the phase estimates, $\mathbb{E}_b V_m = 1 + o(1)$ and $\mathbb{E}_b F_m \ge 3/4$, with auxiliary phases included when needed. We explain why later Fourier contributions are also small under the weight V_m. In a dyadic frequency block of scale U let T_U be its unpinned

trigonometric sum with multiplier $O(1/U)$. Frequency balance shows

$$\int_0^1 |T_U(t)|^{2r} dt$$

has monomial absolute multiplier sum at most C_R^{2r}/U, by fixing all but the last shift. Parseval and Cauchy–Schwarz on its rth power bound the supremum to power $2r$ by CrU times this integral. For independent bounded centered coefficients the model supremum moment root is at most $C_R(rU)^{1/(2r)}\sqrt{r/U}$ from pointwise subgaussian moments. This also holds for fixed gates. Between B_m and N set $r = \lceil \log(16m) \rceil$ on every block. In the model the weight V_m is independent of these bands and of mean one. Each integrated moment tuple adds at most $2r$ shifts and mode degree $O(r)$. The arithmetic error under V_m remains exponentially small in $\log p$ in the complete case and in J in the conditioned case. The factor U^{-1} cancels the U in the supremum estimate. The sum of the model roots is $O_R(\sqrt{r/B_m}(rB_m)^{1/(2r)}) = O_R(m^{-1/2})$. Taking roots in the arithmetic comparison and summing costs a negligible error, uniformly over partial final blocks.

The unweighted terminal tail above N is exponentially small in $\log p$ or J, respectively. Here only the fixed fourth moment is needed. For large n work conditionally, allowing a collision between distinct shifts modulo $P_0 = p^{\lfloor n/C_4 \rfloor}$ for a sufficiently large fixed C_4. On balanced tuples such a collision costs absolute weight $O_R(1/(UP_0))$: choose the colliding pair in $O(U^2/P_0)$ ways, fix the next free slot, then determine the last by balance. When $2U < P_0$ no such collision occurs. The good tuples pay q^{-c_R} by the local estimate of Appendix .1, with bounded mode degrees. Model moments are $O_R(U^{-2})$ before the supremum factor. For bounded n in complete averaging take $P_0 = p$ and use the complete last-digit estimate at fixed moment degree. The resulting bound after fourth roots and dyadic summation is

$$O_R(N^{-1/4} + (1 + \log q)q^{-c_R} + N/q)$$

in the stated respective regimes, with $c_R > 0$ sufficiently small. In the bounded-depth complete case the large-base saving is of this form since n is bounded. In the conditioned case J large ensures n large. Partial blocks and interpolation obey the same bounds by (16). The terminal contribution under the weight is negligible because $\|V_m\|_\infty \le e^{Cm}$. Thus the weighted expected pinned tail after B_m is $O_{R,H}(m^{-1/2})$.

By Markov the V_m-weighted mass of tail failure at a sufficiently large fixed multiple of $m^{-1/2}$ is at most $1/4$. In complete $d = 1$ one can also discard the low singular positions at negligible weighted cost. Therefore on the joint tail-good, nonsingular event the signed expectation is still at least a fixed positive constant. Since $F_m \le 0$ off its retained row bounds and $F_m \le V_m \le \exp(Cm)$ everywhere, the probability of all good bounds is at least $\exp(-C'm)$. The row, tail and mesh bounds imply distance at most $C_{R,H}m^{-1/2}$ from g, using the dense multiplier approximation. Take m a fixed multiple of ϵ^{-2} with $m \le aJ$ or $m \le a \log p$ as appropriate. This proves both the residue lower and the complete lower at every depth. Auxiliary angular lifts do not change the original path probabilities.

A.5 The upper bound in L^1

We give the argument for all complete families, and for residue families with uniform positive active density in the relevant band. In the latter case this holds for the single missing progression and also if $m \ge C_R p$. Take $m = \lfloor k_0 \epsilon^{-2} \rfloor$ with $k_0 > 0$ small. Choose a smooth multiplier $\vartheta(h/m)$ supported in $(1, 2)$, positive on a fixed core subinterval. At $M = 8m$ equally spaced samples, the mean absolute size of this projection of a function is bounded by $C\,\|f\|_1$. Indeed its kernel is bounded by

$Cm/(1+m^2\operatorname{dist}(t,\mathbb{Z})^2)$ by summation by parts, and the mean of the translated absolute kernel on the grid is uniformly bounded.

If $\|K_q - g\|_1 < \epsilon$, round the endpoint W_0 to a finite η-net point v_0 in the bounded trace disc. Let y be the sampled projection of $Q_q - tW_0$, multiplied by $\sqrt{m}$, and z the corresponding vector for $g - tv_0$. Their mean absolute difference is $O_R(\sqrt{m}(\epsilon + q^{-1/2}) + \eta)$. The affine correction follows by sampled Parseval and $\widehat{t}(h) = -1/(2\pi i h)$ for $h \neq 0$. No endpoint of g is required. Let T be the linear map from W to y using (17). It has mutually orthogonal complex columns with bounded Euclidean norms, total squared norm $\gg_R m$ on the core active columns, and each row has bounded squared norm. When conditioning, use only the active columns. Thus if U restricts to all but at most αM sample coordinates, for small fixed α the restricted map still has total column squared norm $\gg_R m$ and bounded operator norm.

For independent bounded model traces of variance one in the complete case and variance d on active indices in the conditioned case, $\|UTX - v\|_2$ is convex and uniformly Lipschitz for any v. Talagrand's convex-distance inequality [6, Theorem 4.1.1] gives subgaussian deviations about its median, treating each complex coefficient as one independent two-real-dimensional block. In fact the Euclidean distance to a convex mixture of a sublevel set is bounded by twice the coefficient bound times the Euclidean norm of the coordinate mismatch proportions. Convexity and Lipschitz continuity compare the excess above the level with convex distance. Applying the convex-distance inequality to the median sublevel and a sublevel below the median gives both median tails. Since the second moment of the norm is at least $c_R m$ by orthogonality and variance, its median is $\gg_R \sqrt{m}$, and

$$\mathbb{P}\{\|UTX - v\|_2^2 \le c_1 m\} \le e^{-c_2 m}$$

for suitable fixed $c_1, c_2 > 0$, uniformly in v.

Project the target Uz onto a Euclidean ball of radius $C_R\sqrt{m}$ containing all UTW, and call the resulting point v. On the trace cube the polynomial

$$\left(1 - \frac{\|UTW - v\|_2^2}{C_3 m}\right)^m$$

lies between zero and one for large fixed C_3. Its model mean is at most $e^{-c_1 m/C_3} + e^{-c_2 m}$. It has total degree $2m$ on $O(m)$ coordinates and absolute character norm $\exp(O_R(m))$, so the phase comparisons pay this exponential bound with room for arbitrary small $m/\log p$ or m/J. If all retained coordinates have $|y_i - z_i| \le \omega$ the polynomial is at least $e^{-C_R\omega^2 m}$, since deletion of coordinates is fixed for this test and projection is nonexpansive. Choose ω small. Next choose α so the union over deleted masks costs less than a fixed fraction of the resulting exponential upper. Choose k_0, η small so that mean absolute discrepancy at most $\alpha\omega$ follows from the original L^1 event, and hence one may delete at most αM coordinates. Union over the masks and the fixed finite endpoint net proves the bound. The comparison is uniform in every projected target, because its polynomial norm estimate is independent of its location. The residue argument has constants uniform when $d = 1$, and at fixed p one can choose a small enough upper radius for $m \ge C_R p$.

A.6 Coverage and the missing progression

Suppose $d = 1$ and write $f = g - t\beta_0$, extended periodically. For $b_0 \neq 0$ let

$$P_{b_0} f(t) = p^{-1} \sum_{j=0}^{p-1} e(-b_0 j/p) f(t + j/p).$$

It is the projection onto frequencies $h \equiv b_0 \pmod p$, with uniform and circular Hölder operator norms at most one. Then

$$g_{b_0} = g - P_{b_0} f + P_{b_0} f(0)$$

is compatible with the residue law and has half-Hölder bound $O(H)$. Restrict to the positive fraction of residues with least absolute representative at distance at least cp from zero. On each such progression a smooth low pass through scale p^2 has absolute coefficient sum $O(\log p/p)$, since $|\beta_h| = 1$ and the denominators grow in steps of p. The smooth remainder costs $O_H(p^{-1})$ in sup norm by the half-Hölder approximation at degree p^2. Thus the uniformly pinned deletion costs $O_H(\log p/p)$.

For finite r, after modulation by $e(-b_0 t)$ the deleted polynomial series is a series in pt with absolute coefficients bounded by $C/[p(1 + |k|)]$. Inverse Hausdorff–Young for $r \geq 2$ and L^2 for $r < 2$ give $\|P_{b_0} f\|_r \leq C_r/p$. To control pinning, subtract the smooth low pass at scale cp that has no frequencies in any of these far progressions. The remainder u satisfies $\|u\|_\infty \ll_H p^{-1/2}$, and

$$\sum_{b_0 \bmod p} |P_{b_0} u(0)|^2 = p^{-1} \sum_{j=0}^{p-1} |u(j/p)|^2 \ll_H p^{-1}.$$

For a positive fraction of the far residues $|P_{b_0} u(0)| \ll_H p^{-1}$. Hence the pinned deletion costs $O_{H,r}(p^{-1})$ there.

Apply the conditioned uniform lower to g_{b_0} at radius $\epsilon/2$ and average over the stated positive fraction. This gives the two lower bounds in Theorem A.2, using $\|\cdot\|_r \leq \|\cdot\|_\infty$ for the ball about g_{b_0}. Their upper bounds follow by averaging the conditioned L^1 upper. The sufficient depth condition in the uniform case is equivalent, for large p, to $n \geq C_H p^4/(\log p)^5$, and in finite norms to $n \geq C_{H,r} p^4/\log p$.

We construct matching obstructions. Take Rudin–Shapiro signs ξ_b with

$$\sup_t \left| \sum_{b \in I} \xi_b e(bt) \right| \ll \sqrt{|I|} \tag{18}$$

for any interval I in the nonnegative integers. The complementary-polynomial recursion bounds the norm of each aligned power-of-two block by C times the square root of its length, and an interval splits into at most two such blocks at each scale. This proves (18) from the Rudin–Shapiro construction [27]. For $0 \leq b < p$ and $|k| \leq p$ use the Fourier coefficients

$$\beta_{b+kp} = \xi_b e(-bk/p) \operatorname{sgn}_+(k) \qquad (b + kp \neq 0),$$

where $\operatorname{sgn}_+(k)$ is 1 for $k \geq 0$ and -1 otherwise. Fill the exterior positive and negative frequencies with ordinary Rudin–Shapiro signs and set the endpoint to 1. All coefficients are supported.

We verify the common half-Hölder bound. On any rectangle with b in an interval in $[0, p-1]$ and k in a consecutive interval of length $L \leq p$ on one sign side, the unweighted sum has uniform norm $O(\sqrt{pL})$. Sum in k first and write $v = pt - b/p$. On the intervals where $\operatorname{dist}(v, \mathbb{Z}) \leq 1/L$, factor out the starting modulation of the Dirichlet sum. The remaining weight has supremum and total variation $O(L)$ on these intervals by its derivative bound $O(L^2)$. There are $O(p/L)$ terms on each. Partial summation with (18) gives $O(\sqrt{pL})$. On a dyadic annulus $\operatorname{dist}(v, \mathbb{Z}) \asymp \delta$, split the geometric numerator into its two modulations. The weight $(1 - e(v))^{-1}$ has supremum and variation $O(1/\delta)$ on each of $O(1)$ intervals with $O(p\delta)$ terms. The modulation changes only the real parameter in (18). This contributes $O(\sqrt{p/\delta})$, and summing the annuli gives the rectangle estimate.

On a dyadic band of positive $k \asymp K \geq 1$, partial summation for $1/(b+kp)$ gives norm $O((pK)^{-1/2})$ for the unpinned path contribution. For example use $1/(b+kp) = \int_0^\infty e^{-v(b+kp)}dv$ and partial summation with the two monotone weights on rectangles. Negative bands with $|k| \geq 2$ follow using $|b+kp| \geq p(|k|-1)$, reversing the weight in b. The exceptional $k = 0, -1$ coefficients are modulated Rudin–Shapiro signs and use the usual dyadic grouping in $|h|$. The exterior frequencies obey the same bounds by (18), including the partial boundary bands. Each block at frequency scale A has norm $O(A^{-1/2})$ and derivative $O(A^{1/2})$ by Bernstein. Summing the minimum of its amplitude and the derivative increment bound shows a uniform half-Hölder constant. The additive endpoint and pinning preserve that property.

Apply a nonnegative smooth multiplier equal to one for $|h| \leq p^2/4$ and zero for $|h| \geq p^2/2$, denoting the operator by S. Its uniform operator norm is bounded. At a time with $pt = b/p \bmod 1$, every contributing term of $P_b S f(t)$ aligns with a common phase, since $\operatorname{sgn}_+(k)/(b+kp) > 0$. The terms $1 \leq k \leq p/8$ alone give $|P_b S f(t)| \gg (\log p)/p$ for every $b \neq 0$. For the pinned arithmetic mesh all frequencies in that progression are zero by (17). Applying the bounded projection and multiplier to the pinned difference gives the uniform obstruction after subtracting the mesh error $O_R(q^{-1/2})$, which is negligible already at $n = 2$.

For the L^1 obstruction the three normalized coefficients at $h = b, b+p, b-p$ contain a pair differing by at least $\sqrt{2}$, because $\max(|1-e(-b/p)|, |1+e(b/p)|) \geq \sqrt{2}$. At all these frequencies the normalized *ordinary* mesh coefficients are $-W_0$, while those of the center are $\beta_h - \beta_0$. Subtract two comparisons to cancel the affine terms. At least one ordinary coefficient difference has magnitude $\gg 1/p$ after division by $2\pi i h$. This bounds the L^1 distance below, up to the mesh error, which is negligible for $n \geq 3$. It also bounds every fixed finite L^r from below. The half-Hölder bound is absolute for these examples. This proves the sharpness statements of Theorem A.2.

A.6.1 An obstruction from opposite trace labels

Suppose the last sign case of the support table applies and d is not a power of two. The unweighted torus trace $S_d = \sum Z_j$ is real because $Z_{j+d/2} = \overline{Z_j}$. Its maximum is d and its minimum is strictly greater than $-d$. Indeed equality at the minimum would require all $Z_j = -1$, incompatible with the relation given by the cyclotomic polynomial Φ_d, whose coefficient sum $\Phi_d(1)$ is one since d is even and not a prime power. Thus the targets $\pm\kappa d$ are attained in the union of the two active supports, and either has fixed positive distance from the opposite label and from zero.

Take M a sufficiently large fixed multiple of $\log p$, choose $\beta_h = \kappa d\, \epsilon_h$ for $M \leq h < 2M$ with signs ϵ_h, and set the remaining coefficients and endpoint to zero. The complete model supports them. There is a sign choice whose unweighted trigonometric polynomial has norm $O_R(\sqrt{M})$ and which mismatches every arithmetic residue label sequence in at least cM positions. Indeed the exponential coloring count in Lemma A.3, with all later levels empty, on a Bernstein grid at large fixed λ, puts probability at least $e^{-\eta M}$ on the norm bound for arbitrarily small fixed η. For every fixed residue pattern, independent signs have probability at most $e^{-c'M}$ of fewer than cM mismatches, with inactive slots counted as mismatches. Union over at most p patterns and choose $M/\log p$ large. Bernstein gives a uniform half-Hölder bound for the resulting pinned center.

For the L^1 comparison, pair the normalized ordinary Fourier coefficient difference with $\epsilon_h \overline{\kappa}$ and take the real sum. Each matching active trace is bounded by the target in this ordering, and each wrong label loses a fixed amount. The total is $\gg_R M$, up to $O_R(M^2/q)$ from the sinc mesh phases and $O_R(\sqrt{M})$ from the affine term, since $\sum_h \epsilon_h = O_R(\sqrt{M})$. The dual test in L^∞ is a constant multiple of the derivative of the sign polynomial and has norm $O_R(M^{3/2})$ by Bernstein. Hence the L^1 distance is $\gg_R M^{-1/2}$, after the negligible mesh difference. This proves the claimed signed-label obstruction.

.1 Last-digit and depth comparisons

For a nontrivial active torus mode let $\gamma = \sum_j m_j \zeta^j \neq 0$. The valuation bound

$$v_p(\gamma) \leq \log_p |\operatorname{Norm}_{\mathbb{Q}(\zeta_d)/\mathbb{Q}}(\gamma)| \ll_R \log_p(2 + \max_j |m_j|)$$

follows because all conjugates are algebraic integers and the absolute norm is at most a fixed power of the maximum coefficient size. In the rank-two case $d = 2$, the mode lattice is one-dimensional and γ is an integer multiple of the corresponding root exponent, a unit when its nonzero absolute size is less than p. Chebyshev modes of the active real trace decompose into the two opposite exponents at the same nonzero index, with unit-modulus prefactors determined by the initial residue.

First use distinct shift residues modulo p. Fix the active labels at the slots that carry nontrivial phase modes and average the last digit of b. The phase survives only if $\sum_i \gamma_i / s_i^{R-1} = 0 \pmod p$, since its derivative is a^{-1} times this sum. When each nontrivial γ_i is a unit, this forces a zero of a nonzero polynomial of degree $O_R(tR^t)$, with t the number of nontrivial phase slots. Over $\overline{\mathbb{F}}_p(b)$ the Rth roots of the linear functions $(b - h_i)/a$ generate a Kummer extension with group of order R^t. The distinct simple zeros give independence by valuations. Rotating just one chosen root changes its term by a nontrivial root of unity, because $\gcd(R, R-1) = 1$, so the rational derivative sum is nonzero. Clearing the denominator and taking the norm gives the degree and nonvanishing. This accounts for all possible choices of branches.

For terms with trivial torus modes the comparison reduces to the labels at the tested shifts. The character of order d, or $2d$ in the sign-labeled case, gives their joint product probabilities up to $O_R(p^{-1/2} e^{C_R t})$. The trivial character terms give the independent main term after discarding the at most $t + 1$ singular and excluded residues. Each other term has a distinct linear factor with exponent not divisible by the character order, so the multiplicative Weil bound applies. If $d = 1$, discard the same exceptional residues and all other labels are active. For terms with a nontrivial torus mode, the choice of unused Haar variables on inactive coordinates makes those cases vanish. If all the nontrivial modes are active, the derivative argument bounds the arithmetic contribution while the Haar main term is zero. Hence the complete-family discrepancy for an individual Fourier mode on the product of labels and tori is $O_R(p^{-1/2} e^{C_R t})$ uniformly in depth. Here t can count all participating shifts. This applies whenever all nontrivial γ_i are units, which includes modes of size $O((\log p)^C)$ for any fixed C as p grows.

For the depth comparison we use arbitrary nonzero modes, retaining their valuations. Suppose t active shifts are distinct modulo p^u, $u \geq 1$, and the mode coefficients $\gamma_i \in \mathbb{Z}_p$ are nonzero. Put

$$k_1 = 2 + \lceil \log_p(Rt + R) \rceil, qquad V = \lfloor n/2 \rfloor - \min_i v_p(\gamma_i) - (t-1)u - k_1 t.$$

If $V > 0$, the normalized contribution of the corresponding product phase, conditional on a fixed active residue class, is $O(tp^{-V/(t-1)})$ when $t \geq 2$ and is zero when $t = 1$.

To see this, first sum the top digits after fixing b modulo $p^{\lceil n/2 \rceil}$. Taylor expansion is linear in the summed digits, so only points where $D(b) = \sum_i \gamma_i / s_i(b)^{R-1}$ is divisible by $p^{\lfloor n/2 \rfloor}$ contribute. On a ball $b = b_1 + p^{k_1} z$ set $c_i = \gamma_i / s_i(b_1)^{R-1}$, $y_i = (b_1 - h_i)^{-1}$ and $M_j = \sum_i c_i y_i^j$. There is an index $0 \leq j < t$ with

$$v_p(M_j) \leq \min_i v_p(\gamma_i) + (t-1)u.$$

Indeed, for i attaining the minimum, use the coefficients of $\prod_{\ell \neq i}(X - y_\ell)$ as a linear combination of the first t moments. It takes value $c_i \prod_{\ell \neq i}(y_i - y_\ell)$, and $v_p(y_i - y_\ell) = v_p(h_i - h_\ell) < u$. If $b_2 = \min_{j<t} v_p(M_j)$, all later M_j have valuation at least b_2 by the integral recurrence with roots y_i.

The coefficient of z^j in the restricted expansion of D is $p^{k_1 j}\binom{-(R-1)/R}{j}M_j$. The binomial is p-integral and its valuation is at most $\log_p(Rj+R)$: in the numerator product of the positive progression $R-1, 2R-1, \dots, Rj-1$, the number of multiples of each p^i differs by at most one from the denominator $j!$, and the difference is nonnegative. Some index $j < t$ attaining b_2 has a coefficient valuation strictly smaller than b_2+k_1t, while all later indices have valuation at least b_2+k_1t. Divide by the smallest coefficient valuation. The restricted series then has last unit coefficient at an index $d_1 \le t-1$, and surviving points are zeros modulo at least p^V of this series. If $d_1 = 0$ there are none. Otherwise they occupy at most d_1 residue classes modulo $p^{\lceil V/d_1 \rceil}$. Indeed d_1+1 such distinct classes would give a divided difference of valuation at least $V - d_1(\lceil V/d_1 \rceil - 1) > 0$. But on the restricted power series this divided difference is the unit coefficient at d_1 modulo p: monomials of lower degree vanish, and higher divided differences are integral symmetric polynomials, so their nonunit coefficients vanish modulo p. This contradiction bounds the proportion within each ball by $d_1 p^{-V/d_1}$. Averaging such balls inside the specified residue incurs no additional factor. This is the same local power-series mechanism as in high-depth estimates for rank-two shifts [12, 14].

Take any sufficiently large $m \le a_2 J$ for sufficiently small a_2. In the depth aspect we can compare tests involving shifts up to $N = \lfloor e^J/16 \rfloor$ by choosing $u = 1 + \lceil \log_p(4N) \rceil$. For a term with $O(m)$ active shifts and mode size a fixed power of m, the valuations cost $O(\log(2J)/\log p)$ and the other losses in V are $O(a_2 n) + o(n)$ uniformly in p, since

$$J \le n, \qquad J^2 \le n \log p, \qquad J \log(2J)/\log p = O(n).$$

For a_2 small and J large, $V \ge n/4$. The discrepancy is exponentially small in $n \log p/m \ge J/a_2$, which pays any fixed exponential cost in m. The model here is the deterministic gate and label law on each residue; zero modes give its exact main terms.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including the packet coupling, mixed wave localization, and obstruction arguments, and with drafting and checking the manuscript. No data or numerical certificates are required.

A Higher-rank Kloosterman shrinking targets

We count paths of twisted exponential sums as the inverse parameter is held fixed. Away from the singular parameter their completed Fourier coefficients are hyper-Kloosterman traces of fixed rank. For each fixed finite L^r norm the negative logarithm of the ball probability has order ϵ^{-2} down to radius of order $(\log p)^{-1/2}$. In the uniform norm the lower radius is of order $(\log\log p/\log p)^{1/2}$ and the upper bound reaches $(\log p)^{-1/2}$. The results hold for complete averages and for progressions of length at least $p^{1/2+\delta}$, with $\delta > 0$ fixed. They also hold on most translates of any growing set, with its size in place of p. The center's Fourier coefficients may lie anywhere in the compact trace image, including its cusps. We combine small balls for independent Haar traces with reversible heat observations and dimension-weighted character comparison.

Keywords hyper-Kloosterman paths, shrinking targets, Haar trace, heat semigroup, discrepancy.
Mathematics Subject Classification 11L05, 11T23, 60G17, 60B15.

A.1 Path families and estimates

Write $e(t) = \exp(2\pi i t)$. For an odd prime p and $j \geq 1$, write $\mathrm{Kl}_j(c;p)$ for the normalized rank-j hyper-Kloosterman sum

$$p^{-(j-1)/2} \sum_{x_1 \cdots x_j = c} e_p(x_1 + \cdots + x_j), \qquad e_p(x) = \exp(2\pi i x/p),$$

with all variables nonzero. In particular $\mathrm{Kl}_1(c;p) = e_p(c)$. Fix the completed coefficient rank $R \geq 2$. For $a \in \mathbb{F}_p^\times$ and $b \in \mathbb{F}_p$, set

$$\varphi_{a,b}(x) = \mathrm{Kl}_{R-1}((ax)^{-1};p)e_p(bx) \quad (x \neq 0), \qquad \varphi_{a,b}(0) = 0.$$

Let $K_p(a,b)$ be the polygon with vertices $p^{-1/2}\sum_{x=1}^{j} \varphi_{a,b}(x)$ at times $j/(p-1)$, $0 \leq j \leq p-1$. At $R = 2$ these are ordinary Kloosterman paths with inverse coefficient a^{-1} and linear coefficient b.

Set $G_R = \mathrm{USp}_R$ for even R and $G_R = \mathrm{SU}_R$ for odd R. Let D_R be the full trace image of G_R, a real interval in even rank and a planar domain in odd rank. For a path g starting at zero define

$$\beta_0 = g(1), \qquad \beta_h = 2\pi i h\, \widehat{g - tg(1)}(h) \quad (h \neq 0).$$

We use the class of such curves with half-Hölder seminorm at most H and every $\beta_h \in D_R$. No uniform distance from the boundary of D_R is required.

Theorem A.1. *Fix $R \geq 2$, $H > 0$ and $1 \leq r < \infty$. Keep $a \in \mathbb{F}_p^\times$ fixed and sample b uniformly in one of the following ensembles.*

1. *For $\mathbb{F}_p$, take $P = p$.*
2. *For a cyclic interval or nonzero-step progression of length at least $p^{1/2+\delta}$, with fixed $\delta > 0$, take $P = p$.*
3. *For $S + u$, where S is any deterministic L-point subset of $\mathbb{F}_p$ and $L \to \infty$, take $P = L$ and exclude at most an $O_{R,H,r}(L^{-c})$ fraction of u.*

Uniformly in these ensembles for sufficiently large P there are positive constants such that

$$e^{-C\epsilon^{-2}} \leq \mathbb{P}_b\{\|K_p(a,b) - g\|_r < \epsilon\} \leq e^{-c'\epsilon^{-2}} \qquad (D/\sqrt{\log P} \leq \epsilon \leq \epsilon_*).$$

The upper bound holds in L^1 for every center in L^1 and every $\epsilon \geq d/\sqrt{\log P}$ in the same small-radius range. The lower bound in the uniform norm holds for every admissible center when $\epsilon \geq D'\sqrt{\log\log P/\log P}$. In the translated case one exceptional set works for all centers and radii. The constants are uniform in a, with an exceptional set allowed to depend on a and S. In the progression case constants may also depend on δ.

The lower finite-norm radius is optimal in order for uniform coverage of the indicated center class. The uniform-norm lower and upper radii differ by a factor of order $\sqrt{\log\log P}$.

Kowalski and Sawin [7, 8] established the classical prime limiting process and its support, and the Archive estimate [2] supplies a uniform-ball exponent in rank two. Kowalski and Untrau [15, Proposition 4.6] count scalar odd-rank hyper-Kloosterman traces near a fixed submersive target at radius of order $p^{-1/[3(R^2-1)]}$, using a compact-group Wasserstein bound. We count paths, using the geometry of the trace image to choose a growing number of coefficient targets and a relative estimate for their Fourier sum. Borda and Cuenin [29, Section 4.3] use compact-group heat smoothing in transport. Here the heat kernel is a reversible observation. Its conditional mean is a scalar multiple of the trace in both directions under Haar, and the character expansion controls relative errors for the observed path event.

A.2 Trace image and interior Haar mass

Lemma A.2. *For a fixed R, the trace image D_R has the following properties.*

1. *For $0 < \eta \le 1$, the radial shrink $(1-\eta)D_R$ is at distance at least $c_R\eta^3$ from the complement of D_R in its real ambient dimension $k \in \{1, 2\}$.*
2. *At every $c \in (1-\eta)D_R$ the Haar trace law dominates $c_R\eta^K$ times Lebesgue measure on a ball of radius $c_R\eta^K$ centered at c, for constants $c_R, K > 0$.*
3. *At every $z \in D_R$, the Haar trace law assigns mass at least $c_R a^{\dim G_R}$ to the ball of radius $0 < a \le 1$ centered at z.*

Proof. For USp_R the trace image is $[-R, R]$, and every interior trace has a submersive preimage on the maximal torus. Use all cosines equal to the target divided by R, with phases away from 0 and π.

For SU_R its trace image is the filled hypocycloid [20, Section 3]. Kaiser also gives the explicit SU_3 Haar trace density in his equation (9). We give an argument that also supplies a submersive preimage for every interior trace. On the determinant-one torus the differential has full real target rank whenever there are at least three distinct eigenvalues, since the tangent image contains the real span of $i(z_j - z_R)$. Kowalski and Untrau [15, Lemma 4.7] use the submersion at such a point to obtain a trace ball lower bound there. A single eigenvalue or a pair with one multiplicity one gives a point of the hypocycloid. For two distinct eigenvalues u, v of multiplicities $k_1, k_2 \ge 2$, vary them to $u_s = ue^{is/k_1}$ and $v_s = ve^{-is/k_2}$ and split a pair in each group by opposite angles. This realizes traces

$$F(s, r, t) = k_1 u_s + k_2 v_s - r u_s - t v_s \qquad (0 \le r, t \ll 1).$$

For example, splitting the first pair by $\pm\alpha$ subtracts ru_s with $r = 2(1 - \cos\alpha)$. Bryant [21] uses single pair splits to show interiority in this multiplicity case. Here the derivative $F_s = i(u - v)$ and the projections of $-u, -v$ onto the real normal $u - v$ are nonzero and opposite. The implicit function theorem solves $F(s, r, t) = F(0, 0, 0)$ for s, r near zero as functions of t, with $r'(0) = 1$. Taking $t > 0$ small gives $r > 0$ and at least three distinct eigenvalues at the same trace. Thus any point off the hypocycloid curve in the trace image has a submersive preimage.

The curve $\gamma(t) = (R-1)e^{it} + e^{-i(R-1)t}$ winds once about zero, since the factor $(R-1) + e^{-iRt}$ has zero winding. The derivative of its argument is

$$\frac{(R-1)(R-2)(1-\cos(Rt))}{|\gamma(t)|^2}.$$

It is positive except at finitely many cusps, and $|\gamma(t)| \ge R - 2$. Thus the curve is a radial graph. Its interior contains trace zero, realized by equally spaced eigenvalues with a common determinant adjustment. Since the compact trace image contains the curve and has no boundary off it, the trace image consists precisely of the filled region. Integrating the angular derivative over any sufficiently short arc of length b gives a lower bound $c_R b^3$. Therefore the radial boundary function is 1/3-Hölder. For $w = (1-\eta)z$ there is radial gap at least $c_R\eta$. Away from zero, perturbing w by $O(c\eta^3)$ changes the boundary radius by at most $O_R(c^{1/3}\eta)$, so for sufficiently small c the perturbed point stays in the image. Points near zero have a fixed margin. This proves the first assertion in both groups.

Let $s(z)$ be the largest least-target singular value of the trace differential at a point of the fiber above z. In a fixed Lie algebra basis translated from the identity, the differential depends algebraically on the matrix entries. The compact real algebraic group and Tarski–Seidenberg imply that s is semialgebraic. It is positive and locally bounded below in the interior, by the local

section at a submersive preimage. The infimum of $s(z)$ at distance at least a from the boundary is therefore positive on its nonempty range, and semialgebraic in a. Polynomial growth of its reciprocal at $1/a \to \infty$ [22, Proposition 2.11] gives $s(z) \gg_R \operatorname{dist}(z, \partial D_R)^K$ for some K.

In an exponential chart at a preimage with least-target singular value s, choose k orthonormal coordinate directions on which the inverse derivative has norm at most $1/s$. Uniform C^2 bounds on trace in translated charts make the derivative variation at most $s/2$ on a box of radius cs in those directions and $c's^2$ in the others. For each fixed value of the other coordinates the contraction proof of inverse mapping gives an image ball of radius $c''s^2$. Its center moves by only $O_R(c's^2)$ from z. There is thus a common ball of radius $\gg_R s^2$ about z in the images. The determinant in the first k coordinates is bounded above and the chart Haar density is bounded below. Integrating over the remaining coordinates yields density domination by $\gg_R s^{2(\dim G_R - k)}$ on that ball. Combining with the first assertion and the semialgebraic bound proves the second. The third follows from a group ball about any preimage, by uniform Haar volume and the Lipschitz bound of the trace. □

A.3 Small balls for the independent Haar trace path

Lemma A.3. *Let A be a complex matrix with N columns, $O(N)$ rows and bounded entries. Let ξ_i be independent uniform qth roots of unity, with $q \geq 2$ fixed. For all $\lambda \geq 1$,*

$$\mathbb{P}(\|A\xi\|_\infty \leq C\lambda\sqrt{N}) \geq \exp(-CNe^{-c\lambda^2}),$$

where the constants may depend on q and the matrix bounds.

Proof. For each normalized row sum $N^{-1/2}(A\xi)_i$, its real and imaginary parts have subgaussian tails. Quantize them into bins of width λ. A rounded integer Z satisfies $\mathbb{E}|Z| \ll e^{-c_0\lambda^2}$, so its entropy is $O(e^{-c_1\lambda^2})$ by testing against the mass proportional to $e^{-c_0\lambda^2|j|}$. By subadditivity a fiber E of the vector quantization has mass at least e^{-b}, where $b = CNe^{-c_1\lambda^2}$.

Write $\zeta = e^{2\pi i/q}$. Each Hamming distance to $\zeta^{-j}E$ is one-Lipschitz in each coordinate and has mass at least e^{-b} at zero. By the bounded-difference inequality its mean is at most $\sqrt{Nb/2}$ and its upper tail is at most $e^{-b}/(2q)$ at $d = \lceil C_q\sqrt{N(b+1)}\rceil$. Thus for at least half the $x \in E$, there are $x^{(j)} \in E$ with $\zeta^{-j}x^{(j)}$ within distance d of x for every j. The vector

$$z = q^{-1}\sum_{j=0}^{q-1} \zeta^{-j}x^{(j)}$$

is in the root polygon in every coordinate and agrees with x outside at most qd coordinates. Since $\sum_j \zeta^{-j} = 0$ and the vectors $Ax^{(j)}$ share their bin vector, $\|Az\|_\infty \ll \lambda\sqrt{N}$.

We round the remaining fractional coordinates to root values with added discrepancy $O_q(\sqrt{N})$, preserving any coordinate already a vertex. Write each as convex weights on the roots and approximate the weights by dyadics of precision 2^{-T} preserving sums and vertices. At each halving of precision, pair the weights with odd fine-grid numerators within each coordinate and add one unit to one numerator in each pair and subtract one from the other. The odd count is even because the sum is one. Choose the signs of the changes by Spencer's bounded real matrix theorem [4], applied to real and imaginary row coordinates with padding to a square of side $O_q(N)$. At precision 2^{-j} this costs $O_q(2^{-j}\sqrt{N})$ in discrepancy. Sum the errors and pass to a subsequence of the final discrete outputs. Deterministic multicolour discrepancy bounds are developed by Doerr and Srivastav [28]. For binary inputs Spencer's original result also counts exponentially many low-discrepancy colorings at a fixed threshold, as discussed by Cai et al. [5]. The cyclic rounding here serves to count outputs as the threshold varies.

Each full output differs from its original x in at most qd coordinates, so its possible preimages are bounded by the volume of a Hamming ball of that radius. For $\lambda \le C_q\sqrt{\log(2N)}$ weaken c_1 so that $b \ge 1$. Above a suitable fixed λ, we have $qd/N \ll_q e^{-c_1\lambda^2/2}$ and the logarithm of the Hamming ball volume is $O_q(Ne^{-c_2\lambda^2})$. For bounded λ rounding the zero vector alone gives a solution with probability at least q^{-N}. Above $C_q\sqrt{\log(2N)}$ the original row tail union bound gives probability at least $1 - O(Ne^{-c\lambda^2})$, which yields the asserted exponential form. This proves the count. □

Let X_h be independent Haar traces of G_R. Irreducibility and Haar character orthogonality give $\mathbb{E}X_h = 0$ and $\mathbb{E}|X_h|^2 = 1$. These bounded coefficients define the continuous path

$$K_R(t) = tX_0 + \sum_{h\neq 0} \frac{X_h}{2\pi i h}(e(ht) - 1).$$

The series converges uniformly almost surely, since the expected maxima of the dyadic blocks, including partial blocks, are $O_R(\sqrt{\log(2U)/U})$ at degree U by pointwise concentration and Bernstein meshes. The center of G_R acts on this trace law by sign for even R and by the Rth roots of unity for odd R. Represent $X_h = V_h\xi_h$ with independent uniform rotations ξ_h independent of V_h whose law is the original trace law. Conditional on V_h in a band of degree U, the Fourier sample matrix has $O(U)$ rows and columns and entries $O(U^{-1})$. Lemma A.3 and the Bernstein mesh give pinned band norm at most $C_R\lambda U^{-1/2}$ with probability at least $\exp(-C_R U e^{-c_R\lambda^2})$. Above a cutoff M, take $U = 2^j M$ and $\lambda = 2^{j/4}$. Independence and summation bound the high part by $O_R(M^{-1/2})$ with probability $e^{-O_R(M)}$.

For $0 < |h| \le M$ use Fejér means $c_h = (1 - |h|/(M+1))\beta_h$, and at zero use $c_0 = (1 - M^{-1/2})\beta_0$. Write η_h for the amounts of contraction. Their pinned approximation error is $O_{R,H}(M^{-1/2})$ by the circular half-moment of the Fejér kernel. By Lemma A.2, at each c_h the Haar trace dominates a shifted triangular law of width $c_R\eta_h^K$ (use the difference of two uniform real intervals or coordinate boxes) times $c_R\eta_h^{K'}$. Take $K \ge 1$. The product cost of domination is $e^{-O_R(M)}$ by the factorial estimate. The two one-copy Fourier band polynomials from the triangular law have pointwise variance proxy $O_R(M^{-1})$ because their widths are $O_R(|h|/M)$. Quantize their real and imaginary parts on a Bernstein mesh of size $O(M)$ at spacing $M^{-1/2}$. The vector entropy is $O_R(M)$, so two independent copies have equal bin vectors with probability at least $e^{-O_R(M)}$ by Jensen applied to the atom probabilities. On this event the pinned difference is $O_R(M^{-1/2})$, as is the endpoint error. Combining low and high, then taking M a sufficiently large constant times ϵ^{-2}, proves

$$\mathbb{P}(\|K_R - g\|_\infty < \epsilon) \ge e^{-C_{R,H}\epsilon^{-2}} \tag{19}$$

for all admissible centers.

A.4 The arithmetic comparison window

Write W_h for the complete normalized coefficient of the twisted summand at frequency h, and use the mesh polygon Q_p at times j/p after extending the summand by zero at zero. Its distance from K_p is $O_R(p^{-1/2})$. At a mesh vertex its exact Fourier sum has endpoint tW_0 and nonzero-frequency terms

$$b_{p,h}W_h(e(ht) - 1), \qquad b_{p,h} = \frac{1}{p(1 - e(-h/p))}, \quad 0 < |h| \le (p-1)/2.$$

Indeed summing the discrete Fourier inversion gives this identity, with $|b_{p,h}| \ll |h|^{-1}$ and $b_{p,h} = (2\pi i h)^{-1} + O(p^{-1})$. The actual Fourier coefficient of $Q_p - tW_0$ satisfies

$$2\pi i h\,\widehat{Q_p - tW_0}(h) = c_{p,h}W_h, qquad c_{p,h} = e(h/(2p))\,\mathrm{sinc}(\pi h/p), \quad \mathrm{sinc}(x) = \sin(x)/x.$$

This follows by integrating its piecewise constant derivative on the mesh. Here and below partial Fourier sums include the subtraction of their value at zero.

Appendix .1 gives a character comparison with error $O_R((d+1)p^{-1/2}\|F\|_{\mathrm{ch}})$ on the complete family, $O_R((d+1)p^{-\delta}\log p\,\|F\|_{\mathrm{ch}})$ on a long progression, and $O_R((d+1)L^{-1/2}\|F\|_{\mathrm{ch}})$ in L^2 of translates. Here each character term in F uses at most d active shifts, and $\|F\|_{\mathrm{ch}}$ weights absolute expansion coefficients by product dimensions. Set $N=\lfloor P^{1/8}\rfloor$ and allow a window $M\le a_0\log P$. This comparison pays tests of norm $e^{O_R(M)}$ times $e^{o(\log P)}$ with $d=O(M+\log M)$, by choosing a_0 small. The error is $O(P^{-\theta})$ for some $\theta>0$ before normalization. It can be divided by an event mass e^{-CM} for sufficiently small a_0.

The unweighted terminal tail above N is $O(P^{-\theta'})$ in expected supremum norm for full families and progressions, and in mean over translates. To verify it for full families and progressions, integrate the fourth power of an unpinned dyadic Fourier block of degree U. Only four-tuples with $h_1+h_2=h_3+h_4$ remain, with total coefficient weight $O(U^{-1})$. Parseval on the square of the block bounds the fourth supremum power by $(4U+1)$ times that integrated power. The independent trace fourth moment is $O_R(U^{-2})$ at each time. Each tuple discrepancy is $O_R(p^{-1/2})$ or $O_R(p^{-\delta}\log p)$, since all actual shifts are distinct modulo p unless equal. Thus the block norm is $O_R(U^{-1/4}+p^{-\delta'/4})$ in expectation for some $\delta'>0$. Sum above N, pin the high part, and interpolate the retained polynomial at cost $O_R(N/p)$. For translates the complete expectation is exactly its translate mean by Fubini. In this case Markov on the terminal expectation removes a power-saving fraction of translates independently of targets. Dividing this terminal bound by an event mass or weight of size at least e^{-CM} still leaves a power saving by taking a_0 small. These facts also cover partial blocks.

A.5 Reversible heat observations

Choose a bi-invariant Riemannian metric on G_R and write P_t for its heat semigroup. Let χ be the standard trace and $\lambda_0>0$ its eigenvalue for the negative generator. With heat started at g_0 the mean of the trace at time t is $e^{-\lambda_0 t}\chi(g_0)$. Its deviation from the mean in every real direction is subgaussian with proxy $O_R(\min(t,1))$, uniformly in g_0. Indeed Itô's formula for the directional eigenfunction gives a martingale M_s with bracket rate $O_R(1)$ because the gradient is bounded. The centered deviation is $\int_0^t e^{-\lambda_0(t-s)}\,dM_s$, whose bracket has deterministic bound $O_R(\min(t,1))$. The exponential martingale inequality gives the assertion. In the limiting model, heat started from Haar is stationary and reversible. Given a heat observation Y at time t, the conditional original group element has the heat law at time t started at Y, so its trace has mean $z=e^{-\lambda_0 t}\chi(Y)$ and the same proxy. Passing to observed conjugacy class preserves the trace law because it is independent of the representative.

Take $M=\lfloor a_0\log P\rfloor$. Choose a smooth even function γ supported on $[-1,1]$, taking values in $[0,1]$, equal to $(1+3x^2)^{-1}$ on $|x|\le 1/2$ and below one at nonzero x. Use independent heat observations at each $0<|h|\le M$ with times t_h satisfying $e^{-2\lambda_0 t_h}=\gamma(h/M)$, and at zero use $t_0=1/M$. At time ∞ use independent Haar and set the predictor to zero. In arithmetic apply the same kernel to the compact conjugacy classes attached to W_h. Given inputs, the predictor means are $\gamma(h/M)W_h$ at nonzero h and $e^{-2\lambda_0/M}W_0$ at zero. The forward predictor and the model posterior residual have directional proxies $O_R((h/M)^2)$ and $O_R(M^{-1})$ at zero. For small h/M this follows from $t_h\asymp_R (h/M)^2$, and otherwise from bounded traces. The endpoint drift is $O_R(M^{-1})$ and its expected noise is $O_R(M^{-1/2})$.

The reference path $T(Y)$ uses these predictors with the exact mesh weights $b_{p,h}$ in Section A.4, including the affine endpoint. Its conditional forward mean is the smooth cutoff of the limiting path, up to $O_R(M/p+1/M)$, and the expected supremum of its conditional noise is $O_R(\sqrt{\log(2M)/M})$.

For the latter, use pointwise variance proxy $O_R(1/M)$ and a Bernstein mesh. The cutoff has uniformly bounded L^1 kernel and its circular half-moment is $O(M^{-1/2})$ by two summations by parts. Thus if the full independent path is within a small constant multiple of ϵ of g, the reference mean is within $\epsilon/4$ of g, for $\epsilon \geq D\sqrt{\log(2M)/M}$. With $F(Y) = \mathbf{1}_{\{\|T(Y)-g\|_\infty < \epsilon/2\}}$, Markov and (19) give

$$v = \mathbb{E}F(Y) \geq e^{-C_{R,H}\epsilon^{-2}}. \tag{20}$$

For the arithmetic comparison let $H_1(Y)$ be any bounded function of the observation classes. The G_R character version of Peter–Weyl and the Casimir eigenvalues give the class heat kernel

$$\sum_\pi e^{-t\lambda_\pi}\chi_\pi(g)\overline{\chi_\pi(y)}$$

against Haar class measure. The coefficient of a product character in H_1 is bounded by $\|H_1\|_\infty$ by orthonormality. The rank-weighted character norm after the heat kernels is therefore bounded by

$$\|H_1\|_\infty \prod_{|h|\leq M} \sum_\pi d_\pi e^{-t_h\lambda_\pi} \leq \|H_1\|_\infty e^{C_R M}. \tag{21}$$

The highest-weight parametrization counts a polynomial number of weights in a ball; the Weyl dimension formula gives polynomial dimension growth and the Casimir grows quadratically. Thus each sum is $O_R(1 + t_h^{-C_R})$, or one for $t_h = \infty$. The factorial estimate for $t_h \gg_R (h/M)^2$ and $t_0 = 1/M$ yields (21). Absolute convergence is uniform on input classes, so the dimension-weighted character comparison applies to these infinite expansions. Inserting a product of $2s$ residual factors costs C_R^{2s} and adds at most $2s$ active shifts, by expanding each factor as an input trace minus a bounded function of Y and applying (21). Conjugated factors obey the same estimate.

Use $s = \lceil \log(16M) \rceil$ and retain all shifts through N as in Section A.4. Under the model conditioned on Y, residual coefficients are independent and centered. The proxies above give $O_R(1/M)$ for a pointwise low-band residual and $O_R(1/U)$ for a high-band residual of degree U. The endpoint has proxy $O_R(1/M)$. For a high band integrate its unpinned $2s$th power with weight $F(Y)$. The balanced frequency tuples have total coefficient weight $O(C^{2s}/U)$, fixing all but one. In the low band use $(C\log(2M))^{2s}$. The weighted arithmetic moment discrepancy from (21) and the character comparison is bounded by $P^{-\theta}$ times these coefficient weights after making a_0 small and absorbing $e^{O_R(M)}$. For translates this first holds in L^2_u; Chebyshev after division by the coefficient weights gives a power-saving exceptional fraction for the $O(\log P)$ moment tests and for v. Thus on the good sets $v_a = \mathbb{E}_a F \geq v/2$ by (20). Divide the weighted moments by v_a so the model terms carry factor $v/v_a \leq 2$.

Parseval on the sth power of a band of degree U bounds its $2s$th supremum moment by $(2sU + 1)$ times the integrated one. This cancels the $1/U$ in the high-band error, and after roots the discrepancy terms sum to a negligible quantity. The model terms are bounded by $C_R(2sU + 1)^{1/(2s)}\sqrt{s/U}$ and sum over the dyadic bands to $O_R(\sqrt{s/M})$. The low band and the endpoint obey that same bound. Pinning, the terminal tail divided by v_a, and the mesh interpolation cost therefore give

$$\mathbb{E}_a[\|K_p - T(Y)\|_\infty \mid F = 1] \ll_R \sqrt{\log(2M)/M}.$$

Increase the radius constant and use Markov. Removing the conditioning gives mass at least $v_a/2$ in the uniform ball, proving its lower bound.

A.6 Exact-mean tilts in finite norms

For $a \in D_R$ and integer $D \geq 1$ take

$$f_{D,a}(x) = (1 - |x-a|^2/(8R^2))^D.$$

This lies in $[0,1]$ on the trace image. The third assertion of Lemma A.2 gives $\mathbb{E} f_{D,a} \gg_R D^{-\dim G_R/2}$. Its value decays as $e^{-c_R D|x-a|^2}$. In the tilted law the resulting tails are Gaussian on scale $w_D = C_R\sqrt{\log(2D)/D}$, uniformly in a, by absorbing the polynomial normalizer for $|x-a|$ larger than a constant times w_D. The deviations from the mean have the same directional subgaussian proxy $O_R(w_D^2)$ by independent-copy symmetrization. The tilted mean $m_D(a)$ is continuous and within w_D of a, with C_R large enough. If the closed ball in the real ambient space of radius $2w_D$ about c lies in D_R, the map $a \mapsto c + a - m_D(a)$ takes the ball into itself. Brouwer's fixed point theorem gives a weight parameter whose mean is exactly c.

For the Fejér target coefficients c_h let $D_h = \lceil C_R \eta_h^{-K_1} \rceil$ with a fixed $K_1 > 6$ and C_R large. Then w_{D_h} is at most $O_R(\eta_h)$ and the ball of radius $2w_{D_h}$ about c_h is inside the trace image by the first assertion of Lemma A.2. Form the product weight in the independent model and in arithmetic, with the exact means prescribed by these parameters. It has model mass $v \geq e^{-C_R M}$ and character norm $e^{O_R(M)}$, since $\sum \log(2 + D_h) = O_R(M)$ and the separate-degree character bound applies equally to traces and dual traces. Under the independent tilt the residual coefficients are centered and independent, with proxies $O_R(\eta_h^2)$ in the window and $O_R(1)$ elsewhere. Thus the pointwise proxy of the retained path is $O_R(1/M)$.

Take $2s \geq r$ fixed and $M = \lfloor K_{R,H,r} \epsilon^{-2} \rfloor$. For the integrated $2s$th moment of the retained residual the comparison costs the weight norm times $(C_R \log(2N))^{2s}$ from the linear forms. The shift count in each term is at most $2M + 1 + 2s$. Comparing both this test and the weight, then dividing by the arithmetic weight mass $v_a \geq v/2$, gives an integrated moment at most $2(Cs/M)^s$ plus a negligible error under the tilt. This holds on the good sets obtained as above, with no exceptional set for complete families or long progressions. The mean approximates g uniformly within $O_{R,H}(M^{-1/2})$ by Fejér and the mesh coefficient error. The terminal pinned tail still has negligible expected norm under the tilt, since the weight is bounded by one and its mass is at least $e^{-C_R M}/2$. A sufficiently large $K_{R,H,r}$ and Markov give ball probability at least $v_a/2$ after removing the weight. Take D in Theorem A.1 large so this choice of M fits the allowed window.

For both lower bounds on translates, choose a geometric list of radii and supremum nets chosen from the admissible class at each radius, with mesh a fixed small fraction of the radius. Their log cardinality is $O_{R,H}(\epsilon^{-2})$ by coding rounded values on a uniform grid of spacing $c_{R,H}\epsilon^2$, with a bounded number of choices for each successive increment. A maximal separated set injects into the rounded codes. With initial radius constants large each net size is a sufficiently small power of L to preserve the union bound over radii and centers. Increase the minimum radius constants once more for inclusion from the smaller test balls. This gives the common exceptional set for all admissible centers and radii.

A.7 An upper bound in L^1

Take $m = \lfloor k\epsilon^{-2} \rfloor$ with $k > 0$ small. Choose a smooth real multiplier ϕ supported inside $(1,2)$ and bounded below by a positive constant on a fixed subinterval. Let T_m have Fourier multipliers $\phi(h/m)$, and take $J = 8m$ time samples $t_j = j/J$. The kernel bound $Cm/(1 + m^2 \operatorname{dist}(t,\mathbb{Z})^2)$ follows by two summations by parts, and its average absolute value over the displaced sample grid is bounded uniformly in the displacement. Hence for every periodic L^1 function (viewing functions on $[0,1]$ this way),

$$\frac{1}{J}\sum_j |T_m f(t_j)| \ll \|f\|_1 . \tag{22}$$

Use a fixed finite η-net for the bounded endpoint W_0. If $\|K_p - g\|_1 < \epsilon$, some net point w is within η of W_0. Write

$$y_j = \sqrt{m}\, T_m(Q_p - tW_0)(t_j), \qquad z_j = \sqrt{m}\, T_m(g - tw)(t_j).$$

Then the average of $|y_j - z_j|$ is $O(\sqrt{m}(\epsilon + p^{-1/2}) + \eta)$. For the first term use (22) on $Q_p - g$. For the correction by $w - W_0$, the Fourier coefficient of t at nonzero h is $-1/(2\pi i h)$, and sampled Parseval bounds the mean square of its normalized band by $O(|w - W_0|^2)$. No endpoint of g is needed.

By the mesh Fourier coefficient identity in Section A.4, $y = TW$ where the column of T at h has entries

$$\sqrt{m}\, \phi(h/m) c_{p,h} e(ht_j)/(2\pi i h).$$

Distinct columns are orthogonal, each has bounded norm, and their square norms sum to at least $c_0 m$. Every row has square norm $O(1)$. If U restricts to the coordinates outside a set I of at most αJ deleted samples, then $\|UT\|_{2\to 2} = O(1)$ and the sum of column square norms is still at least $c_0 m/2$, uniformly for α below a sufficiently small fixed constant.

For independent bounded model coefficients and any v in the Euclidean ball of radius $C_1\sqrt{m}$, the map $x \mapsto Z = \|UTx - v\|_2$ is convex and Lipschitz with a bounded constant, treating complex coefficients as two real coordinates within one independent block. Talagrand's convex-distance inequality [6, Theorem 4.1.1] gives subgaussian tails around its median. Indeed, for any sublevel set of a convex Lipschitz function, Jensen bounds the excess above the level by the Lipschitz constant times the Euclidean distance to any convex mixture of points from the set. The distance to the mixture is at most twice the coefficient bound times the Euclidean norm of the vector of coordinate-block mismatch proportions. Taking the infimum over mixtures gives the required comparison with convex distance. The usual upper and lower median tails follow by applying the inequality to the median sublevel and to a sublevel below the median, respectively. Now $\mathbb{E}Z^2 \geq c_2 m$ from independent mean-zero coefficients with variance one. On the other hand $\mathbb{E}Z^2 \leq 2(\operatorname{med} Z)^2 + O(1)$. Thus $\mathbb{P}(Z^2 < c_3 m) \leq \exp(-c_4 m)$.

On the coefficient cube UTW lies in the Euclidean ball of radius $C_1\sqrt{m}$ for a suitable fixed C_1. Project Uz there and use a net of mesh $\delta\sqrt{m}$ contained in the ball, of size at most $(1 + 2C_1/\delta)^{2J}$. If $|y_j - z_j| \leq \delta$ at each retained coordinate, the selected net point v has $\|UTW - v\|_2 \leq C\delta\sqrt{m}$. For sufficiently large fixed C_2 the polynomial

$$F_{I,v}(x) = \left(1 - \frac{\|UTx - v\|_2^2}{C_2 m}\right)^m$$

lies in $[0, 1]$ on the coefficient cube (or product of trace disks). Its value on this event is at least $\exp(-C_3\delta^2 m)$. Its model mean is at most $\exp(-c_3 m/C_2) + \exp(-c_4 m)$. Its total degree is at most $2m$, and its weighted character norm is $\exp(O(m))$ by the total-degree estimate in Appendix .1. Section A.4 pays this comparison cost with a remainder smaller than the model bound, by taking $m \leq (k/d^2)\log P$ with d large. For translates, Chebyshev and a union over all deleted masks, their nets and all positive integer values of m costs at most $(1 + \log L)\exp(C_\delta k \log L/d^2)$. Choose d large, leaving an $O(L^{-c'})$ fraction exceptional. The bound therefore holds for all these candidate polynomials on one good set.

Choose δ sufficiently small that the lower certificate loses at most one fourth of the model exponent. For fixed g and rounded endpoint only one net point per mask is used. Choose α small so the sum of binomial coefficients through αJ costs less than a quarter of the resulting exponential rate. The constants in the model concentration were uniform for all smaller α. Choose η, k so the average discrepancy from (22) is at most $\alpha\delta$, then at most αJ coordinates need deletion. The finite endpoint net costs a fixed factor. Union over endpoints and masks proves the claimed upper

bound after reducing ϵ_* to absorb fixed factors and the integer part. Since $\|f\|_1 \le \|f\|_r$ for every $1 \le r \le \infty$, this finishes all upper bounds in Theorem A.1.

Sharpness of the finite-norm coverage radius. The trace image contains a real segment about zero. On a frequency band $m \le h < 2m$ take normalized coefficients of fixed sufficiently small amplitude and arbitrary signs in that segment, with other coefficients and endpoint zero. Lemma A.3 for $q = 2$, with a sufficiently large fixed λ, supplies e^{cm} sign choices with pinned and unpinned path norm $O(m^{-1/2})$. Bernstein gives derivative bound $O(m^{1/2})$, and taking the minimum of the resulting increment bound and twice the uniform bound gives a common half-Hölder constant. Scaling allows any fixed $H > 0$. Greedily choose a Hamming packing with relative distance bounded below. Parseval on nonzero Fourier modes and the uniform bound yield $\|g - g'\|_1 \ge \|g - g'\|_2^2 / \|g - g'\|_\infty \gg m^{-1/2}$ between distinct centers, and hence the same separation in every norm in Theorem A.1. Taking m a sufficiently large multiple of $\log P$ makes the packing larger than the number of sampled paths. At least one admissible center then has an empty ball at radius $c/\sqrt{\log P}$.

.1 Additive twists and character norms

The completed coefficients of the normalized increments are

$$W_h = p^{-1/2} \sum_{x \ne 0} \varphi_{a,b}(x) e_p(-hx).$$

For $b \ne h$, introducing the last multiplicative coordinate $(b-h)x$ gives exactly $\mathrm{Kl}_R((b-h)/a; p)$. At $b = h$ the sum factors into $R-1$ complete nonzero additive sums, giving $(-1)^{R-1} p^{-(R-1)/2}$. This small real number belongs to the compact trace image: use conjugate eigenvalue pairs, and for odd R also an eigenvalue one. Thus compact group characters may be evaluated at a chosen class with this trace at the exceptional point. At ordinary points use the semisimple Frobenius class of the normalized Kloosterman sheaf. For even R multiply its trace class by the central element $-I$, which preserves the Haar model. For odd R the trace convention already has positive sign.

For distinct active shifts, let $F(b)$ be a product of nontrivial irreducible compact characters of those classes, with d_0 active shifts and product of dimensions D. We claim

$$\left| \sum_{b \in \mathbb{F}_p} F(b) e_p(kb) \right| \ll_R (d_0 + 1) D \sqrt{p} \qquad (k \in \mathbb{F}_p). \tag{23}$$

The normalized single-factor sheaf has arithmetic and geometric group Sp_R or SL_R, tame regular unipotent monodromy at zero and a single infinity break $1/R$, by Katz's results as recorded in [9, Prop. 4.6]. In the product over translates, inertia at $b = h$ acts nontrivially and unipotently in that factor alone. Projection onto each factor is surjective, and the normal closure of a noncentral unipotent in its connected simple group is the full factor. Hence the geometric group is the product. Applying the irreducible representations and tensoring gives rank D, no invariants, and infinity breaks at most $1/R$. For a nontrivial additive twist the break at infinity is one, since the original breaks are strictly smaller. In particular the twisted tensor has no invariants. All finite ramification is tame. The Euler characteristic on the affine line with the active punctures removed gives $\dim H_c^1 \le d_0 D$ and $H_c^0 = H_c^2 = 0$. Purity and the trace formula bound the sum there by $d_0 D \sqrt{p}$. The exceptional points contribute at most $d_0 D$. This proves (23). The all-trivial product has the exact full or restricted expectation one.

We use the norm $\|P\|_{\mathrm{ch}}$ of an expansion into products of characters, with absolute coefficients weighted by product dimensions. It is submultiplicative, and both the trace and its complex conjugate have norm R. For a polynomial in traces and dual traces of total degree r on d factors, with bound U on the product of compact groups,

$$\|P\|_{\mathrm{ch}} \le U \exp(C_R r \log(eC_R(d+r)/r)) \quad (r \ge 1). \tag{24}$$

Indeed Haar character coefficients are bounded by U by orthonormality. Each possible highest weight in one factor has weight norm bounded by C_R times its trace degree, since it occurs in a tensor product of the standard representation and its dual. The total highest-weight norm is at most $C_R r$. Stars and bars count the products of possible highest weights, and the product of dimensions is at most $\exp(C_R r)$ by the Weyl dimension formula. For separate degrees r_h, the same argument bounds the norm by $U \prod_h [C_R(1+r_h)^{C_R}]$ on active factors, with factor one when $r_h = 0$. These estimates also apply to real and imaginary components expressed using the trace and its dual.

Let $m_S(k) = L^{-1} \sum_{b \in S} e_p(kb)$ and $\widehat{F}(k) = \sum_b F(b) e_p(-kb)$. For each nontrivial product F, Parseval over translates gives

$$\left\| \mathbb{E}_{b \in S+u} F(b) - \mathbb{E}_{b \in \mathbb{F}_p} F(b) \right\|_{L^2_u}^2 = \frac{1}{p^2} \sum_{k \ne 0} |\widehat{F}(k)|^2 |m_S(k)|^2.$$

Use (23) and $\sum_k |m_S(k)|^2 = p/L$, and add the error of the full average relative to the independent Haar characters. Thus for a test with at most d active shifts in each character term,

$$\left\| \mathbb{E}_{b \in S+u} P(W) - \mathbb{E} P(X) \right\|_{L^2_u} \ll_R (d+1) L^{-1/2} \|P\|_{\mathrm{ch}}. \tag{25}$$

For a progression of step $d \ne 0$ and length L, the geometric sum for $k \ne 0$ gives

$$|m_I(k)| \ll \min(1, p/(L \operatorname{dist}(kd, p\mathbb{Z}))).$$

It follows that $\sum_k |m_I(k)| \ll p \log p / L$. Fourier inversion gives the pointwise bound

$$|\mathbb{E}_{b \in I} P(W) - \mathbb{E} P(X)| \ll_R (d+1)(\sqrt{p} \log p / L) \|P\|_{\mathrm{ch}}. \tag{26}$$

Both extend to series convergent absolutely in this norm.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, proof drafting and internal checks. No data or numerical certificates are required for the proofs.

A Weighted residue masks for Kloosterman paths

At a fixed prime, the attainable Fourier coefficients of a Kloosterman path over high powers of that prime depend on the residue class of the linear parameter. In an endpoint-and-Fourier Hilbert metric we compute the limiting covering radius exactly as a weighted minimum of these masks. The fraction of residues with weighted loss at most t has asymptotic $\sqrt{t/(\pi I)} e^{-I/t}$ when $t \to 0$ and $t \log p \to \infty$, with $I = \eta(1/2)^2/\pi$. The logarithmic asymptotic holds even when $\liminf t \log p > 4I$. Long quadratic-residue runs show that, below the logarithmic radius, improved coverage for square inverse scales can coexist with failure of the complete-family exponential upper bound for nonsquare inverse scales.

Keywords Kloosterman paths, prime powers, quadratic residues, small deviations.
MSC 2020 11L05, 11T23, 60F10, 60G17.

A.1 Statements

Fix an odd prime p, $n \geq 2$, and a unit a modulo p^n, with fixed residue modulo p when the depth varies. For units b form $K_{p^n}(a,b)$ by joining with equal time steps the partial sums of

$$p^{-n/2} e_{p^n}\big(bx + (ax)^{-1}\big)$$

over the units x in increasing order. Write $e_q(v) = e^{2\pi i v/q}$ and χ for the Legendre symbol modulo p. For each nonzero residue b_0 put

$$S_{p,a,b_0}(h) = \mathbf{1}_{\{\chi((b_0-h)/a)=1\}} \qquad (h \in \mathbb{Z}).$$

This is the mask of allowed Fourier indices, including the endpoint index $h = 0$. On an allowed index the limiting coefficient has support $[-2,2]$, and on the other indices it is zero. This follows from the fixed-residue path law of Milievi and Zhang [30, Proposition 1 and Theorem 1].

For a path g starting at zero let $\beta_0(g) = g(1)$ and $\beta_h(g) = 2\pi i h\, \widehat{g - tg(1)}(h)$ at $h \neq 0$, using Fourier convention $e(-ht)$. Let $\mathcal{H}_H$ consist of such paths with half-Hölder seminorm at most H and every $\beta_h(g)$ in $[-2,2]$. Define the Hilbert metric on these paths by

$$d_2(g,k)^2 = |g(1) - k(1)|^2 + \left\| u - \int_0^1 u(t)\, dt \right\|_2^2, \qquad u(t) = g(t) - k(t) - t(g(1) - k(1)).$$

The endpoint and the mean-zero part of the pinned curve determine the path, since its initial value is given. Thus this is a metric, dominated by a constant times the uniform metric. Let $r_{p,n,a}(H)$ be the supremum over $g \in \mathcal{H}_H$ of its d_2 distance from the nearest arithmetic path $K_{p^n}(a,b)$ for b a unit.

For residues b_0 with allowed endpoint write

$$B_{p,a}(b_0) = \sum_{h \neq 0} \frac{1 - S_{p,a,b_0}(h)}{\pi^2 h^2} = \frac{1}{p^2} \sum_{\substack{1 \leq r < p \\ \chi((b_0 - r)/a) \neq 1}} \csc^2(\pi r/p).$$

The same infinite sum defines $B_{p,a}(b_0)$ also at other residues.

Theorem A.1. *There is an absolute H_0 such that, for fixed p and $H \geq H_0$,*

$$\lim_{n \to \infty} r_{p,n,a}(H)^2 = \min_{\substack{b_0 \in \mathbb{F}_p^\times \\ S_{p,a,b_0}(0) = 1}} B_{p,a}(b_0).$$

If $r_{p,a}$ denotes the square root of this minimum, then

$$r_{p,a} - O(p^{-n/2}) \leq r_{p,n,a}(H) \leq r_{p,a} + O_{p,H}(J^{-1/2}), \qquad J = \min(n, \sqrt{n \log p}),$$

the upper bound using the quantitative conditioned shrinking-target theorem Section A.

Theorem A.2. *Let $I = \eta(1/2)^2/\pi$, where $\eta(s)$ is the Dirichlet eta function. Uniformly in a, if $t \to 0$, $p \to \infty$ and $\liminf t \log p > 4I$, then*

$$t \log \mathbb{P}_{b_0 \in \mathbb{F}_p^\times}\{S_{p,a,b_0}(0) = 1,\ B_{p,a}(b_0) \leq t\} \longrightarrow -I.$$

If in addition $t \log p \to \infty$, the probability is asymptotic to

$$\sqrt{\frac{t}{\pi I}} \exp(-I/t).$$

Theorem A.2 counts the fraction of residue supports within distance $\sqrt{t}$ of the extremal center in Theorem A.1. The probability is over residue classes, rather than over the active Fourier angles.

Theorem A.3. *There are infinitely many primes p with $M \asymp \log p \log\log\log p$ for which both of the following hold.*

1. *For every square inverse scale $a \bmod p$ and $J \geq C_H M$, every $g \in \mathcal{H}_H$ has complete-family uniform ball probability at least $\exp(-C_H \epsilon^{-2})$ for $D_H/\sqrt{M} \leq \epsilon \leq \epsilon_*$.*

2. *For every nonsquare inverse scale and $n \geq C \log p$ there are radii $\epsilon = o((\log p)^{-1/2})$ with*

$$\mathbb{P}_b\{\|K_{p^n}(a,b)\|_\infty < \epsilon\} \geq \exp(-(1+o(1))\log p), \qquad \epsilon^{-2}/\log p \longrightarrow \infty.$$

This refutes a uniform bound $\exp(-c\epsilon^{-2})$ at these radii, also for the complete two-parameter average.

Theorem A.1 uses the Fourier support of the fixed-prime depth law [30]. Its worst center can be taken from the Rudin–Shapiro construction. The independent-series argument behind Theorem A.2 is a lower-tail saddle point in the sense of Lifshits [31]. The arithmetic comparison here keeps the entire Laplace product, with unweighted character norm one on the real axis. Theorem A.3 uses the least-nonresidue theorem of Graham and Ringrose [35], together with a variant of the prime-power wave estimate with no low active frequencies.

A.2 The exact covering radius

Extend the Kloosterman increments by zero at nonunits and use the j/p^n mesh to form Q from the same increments. The time changes from the unit mesh are $O(p^{-n})$ and the summands have modulus one. Thus $\|Q - K_{p^n}\|_\infty = O(p^{-n/2})$ and their endpoints agree. Put $q = p^n$ and let W_h be the complete shifted coefficient, indexed modulo q. The last-digit computation for Kloosterman sums shows $W_h = 0$ whenever $S_{p,a,b_0}(h) = 0$, and $|W_h| \leq 2$ everywhere. Integrating the piecewise constant derivative gives, for every integer $h \neq 0$,

$$\beta_h(Q) = e(h/(2q))\,\mathrm{sinc}(\pi h/q) W_{h \bmod q}, \qquad \beta_0(Q) = W_0.$$

Hence all inactive modes vanish, including outside a single set of mesh residues, and all normalized coefficients of Q are bounded by two. In this metric Parseval gives

$$d_2(g,Q)^2 = |\beta_0(g) - W_0|^2 + \sum_{h \neq 0} \frac{|\beta_h(g) - \beta_h(Q)|^2}{4\pi^2 h^2}.$$

For each b_0 set the inactive coefficients of g to zero, retaining the other coefficients, and pin the result with its chosen endpoint. The periodic multiplier for the nonzero coefficients is a sum of the finitely many residue projections

$$f \longmapsto p^{-1} \sum_{j=0}^{p-1} e(-rj/p) f(t + j/p).$$

It is bounded on periodic half-Hölder functions by $O(p)$. Thus this clipped center has half-Hölder seminorm $O_p(H)$, is compatible with the residue, and has d_2 distance from g at most

$$\sqrt{4(1 - S_{p,a,b_0}(0)) + B_{p,a}(b_0)}.$$

Apply the conditioned lower bound from Section A at a sufficiently large multiple of $J^{-1/2}$. Since the ball has positive mass, an arithmetic path is within that distance in the uniform norm and hence in d_2, up to a constant. The endpoint is allowed for some residues, and $B \le \sum_{h\neq 0} 1/(\pi^2 h^2) = 1/3$, so a residue with inactive endpoint never minimizes the last display.

The qualitative upper limit can also be obtained from fixed-window equidistribution, without the quantitative path estimate. For fixed p, b_0 and finitely many active shifts, the shifted stationary phases converge to independent circle phases as $n \to \infty$ by the fixed-order shifted-root comparison of [30, Section 4.2, Theorem 4]. The averages in its residue classes apply here since the conditioning on b_0 is invariant under translations by the higher power of p used in that comparison. Equivalently, the real completed coefficients converge to independent arcsine variables on those shifts. A finite net in their support cube can be covered by open product neighborhoods of positive limiting probability. All those neighborhoods are hit for sufficiently large n. The Fourier tails above a fixed cutoff N contribute at most $O(1/N)$ to the square of d_2 by the deterministic bound on $\beta_h(Q)$. This suffices uniformly over compatible centers. The same fixed-order comparison is uniform in any lifts of the fixed inverse residue.

For the matching lower bound choose the endpoint 2 and every nonzero β_h of modulus 2, with Rudin–Shapiro signs on positive and negative indices [27]. The exponential sums of these signs on intervals of length U are $O(\sqrt{U})$ uniformly. This follows from the complementary-polynomial norm bound on aligned dyadic intervals and the decomposition of each interval into at most two such pieces at every scale. Partial summation on dyadic frequency intervals gives path amplitude $O(U^{-1/2})$ and Bernstein gives a derivative of size $O(U^{1/2})$. Summing the minimum of these bounds for increments gives a uniform half-Hölder seminorm. This center belongs to $\mathcal{H}_H$ for $H \ge H_0$. Its inactive coefficient discrepancies attain the envelope above, up to the mesh error. Finally, summing $1/(r+kp)^2$ over $k \in \mathbb{Z}$ by the standard cotangent partial fractions gives $\pi^2/(p^2 \sin^2(\pi r/p))$. When the endpoint is allowed the entire residue class $h \equiv 0$ is allowed as well. This proves Theorem A.1.

A.3 The complete Laplace product

Let ξ_h, $h \in \mathbb{Z} \setminus \{0\}$, be independent fair Bernoulli variables and put

$$Y = \sum_{h\neq 0} \frac{\xi_h}{\pi^2 h^2}, \qquad \Lambda(z) = \mathbb{E}e^{-zY}, \qquad F(s) = -\log \Lambda(s) \quad (s > 0).$$

For the arithmetic measure write

$$\Lambda_p(z) = \mathbb{E}_{b_0 \in \mathbb{F}_p^\times} \left[\mathbf{1}_{\{S_{p,a,b_0}(0)=1\}} e^{-zB_{p,a}(b_0)} \right].$$

When the endpoint is allowed, the weight of a missing residue $1 \le r < p$ is $w_r = p^{-2} \csc^2(\pi r/p)$. For $z = s$ or $z = s(1 - i\xi)$ with $|\xi| \le R_0$ fixed, put

$$A_r = \tfrac{1}{2}(1 + e^{-zw_r}), \qquad B_r = \tfrac{1}{2}(1 - e^{-zw_r}).$$

At nonsingular $b_0 - r$, the factor from r is $A_r + B_r \chi((b_0 - r)/a)$. Multiply by the endpoint factor $(1 + \chi(b_0/a))/2$. If $b_0 = r$, the missing-index factor should be e^{-zw_r}, whereas the displayed factor is A_r. Correcting it costs at most $|B_r|$, since the other actual factors have modulus at most one. Averaged over all b_0, this costs $O_{R_0}((1 + \sqrt{s})/p)$, using

$$\sum_{r=1}^{p-1} \min(sw_r, 1) \ll 1 + \sqrt{s}.$$

At real $z = s$ the sum of the absolute expansion coefficients of the whole product, including the endpoint, is one, because A_r, B_r are nonnegative and sum to one. The sum of expansion coefficients weighted by the number of participating shifts is at most $1 + \sum_r B_r \ll 1 + \sqrt{s}$. Each nontrivial term is a product of quadratic characters at distinct shifts, whose complete sum is $O(j\sqrt{p})$ for j participating shifts by Weil's bound [32]. Removing $b_0 = 0$ changes the normalized comparison by $O(1/p)$ per term. Thus

$$\Lambda_p(s) = \tfrac{1}{2} \prod_{r=1}^{p-1} A_r + O((1 + \sqrt{s})p^{-1/2}).$$

For complex $z = s(1 - i\xi)$, $|A_r| + |B_r|$ lies between 1 and $1 + C_{R_0} \min(sw_r, 1)$. The total coefficient norm is $\exp(O_{R_0}(1 + \sqrt{s}))$. Weighting by the participating shifts costs just an extra factor $O_{R_0}(1 + \sqrt{s})$ by expanding the weighted sum one site at a time. Consequently

$$\Lambda_p(z) = \tfrac{1}{2} \prod_{r=1}^{p-1} A_r + O(p^{-1/2} \exp(C_{R_0}(1 + \sqrt{s}))) \tag{27}$$

uniformly in the indicated complex sector and a.

For a least absolute representative h of r, $w_r = 1/(\pi^2 h^2) + O(p^{-2})$. The missing infinite independent weights beyond $(p - 1)/2$ have sum $O(p^{-1})$. Along the indicated sector, $\frac{d}{dw} \log(1 + e^{-zw}) = O_{R_0}(s)$ uniformly for $w \geq 0$. To see this, its denominator is bounded away from zero uniformly for ξ in the fixed compact interval and $sw \geq 0$, by continuity, the limit at infinity, and strict positive real part of z when $w > 0$. The logarithm starts from $\log 2$ at $w = 0$. Thus the finite independent product equals $\Lambda(z) \exp(O_{R_0}(s/p))$. In the ranges used below $s = O(\log^2 p)$, so this grouping error is negligible. The bound $|\Lambda(z)| \leq \Lambda(s)$ will be useful for normalized comparisons.

For the model, with $v = \sqrt{s}/\pi$ and $f(u) = \log 2 - \log(1 + e^{-1/u^2})$,

$$F(s) = 2 \sum_{h \geq 1} f(h/v) = D\sqrt{s} - \log 2 + O(s^{-1/2}), \qquad D = \frac{2}{\pi} \int_0^\infty f(u)\, du.$$

Here $f(0) = \log 2$, the derivatives extend at zero, and its second derivative is integrable, so the trapezoid form of Euler–Maclaurin gives the displayed remainder. Substituting $x = u^{-2}$ and integrating by parts gives

$$\int_0^\infty f(u)\, du = \int_0^\infty \frac{x^{-1/2}}{e^x + 1}\, dx = \sqrt{\pi}\, \eta(1/2).$$

The last integral is the standard Mellin form of the alternating zeta series, also obtained by integrating its bounded alternating geometric partial sums. Thus $D^2/4 = I$. The factor $\frac{1}{2}$ for an allowed endpoint cancels the constant term of F. For real $s \to \infty$ with $D\sqrt{s} \leq (1/2 - \delta) \log p$ for fixed $\delta > 0$ we obtain

$$\Lambda_p(s) = (1 + o(1))e^{-D\sqrt{s}}. \tag{28}$$

The error estimates are uniform in a.

A.4 Logarithmic rate and prefactor

Put $A_p(t) = \mathbb{P}(S(0) = 1, B_{p,a} \leq t)$. Chernoff's bound at $s = D^2/(4t^2)$ gives $A_p(t) \leq e^{st}\Lambda_p(s)$. If $\liminf t \log p > D^2$, equation (28) applies with a fixed margin. For the reverse logarithmic inequality let $0 < \theta < 1/4$ be fixed and small and use $s = D^2/[4(1 - \theta)^2 t^2]$. Tilt the endpoint-restricted arithmetic measure by $e^{-sB_{p,a}}/\Lambda_p(s)$. In this measure $B_{p,a} \in [(1-2\theta)t, t]$ with probability

$1 - o(1)$. Indeed both tails follow by applying Chernoff with parameters $\pm\gamma s$ and comparing $\Lambda_p(s(1 \mp \gamma))/\Lambda_p(s)$ for sufficiently small fixed $\gamma > 0$. To see the signs, the upper logarithmic bound divided by st has expansion $-\gamma + (1-\theta)\gamma + O(\gamma^2)$. The lower tail bound has expansion $(1-2\theta)\gamma - (1-\theta)\gamma + O(\gamma^2)$. Both are negative at small γ. The real transform bound holds at all these parameters if θ, γ are chosen small relative to the strict threshold margin. Undoing the tilt gives at least $(1 - o(1))\Lambda_p(s)e^{s(1-2\theta)t}$. Let $\theta \to 0$ after the limit to obtain the logarithmic rate.

Now suppose $t \log p \to \infty$, and set $s = (D/(2t))^2$. Under the independent tilt proportional to e^{-sY}, let $Z = s(Y - t)$. The Bernoulli variables remain independent, with success probabilities $(1 + e^{s/(\pi^2 h^2)})^{-1}$. Riemann sums for the tilted moments give

$$\mathbb{E}_s Z = O(1), \qquad \mathrm{Var}_s(Z) = V\sqrt{s} + O(1), \qquad V = D/4.$$

Indeed the first and second tilted cumulants of sY use the profiles $x/(1+e^x)$ and $x^2e^x/(1+e^x)^2$ at $x = s/(\pi h)^2$. They extend smoothly at both ends of the Riemann sum and have bounded variation. Integrating by parts in x gives respective leading coefficients $D/2$ and $D/4$ times $\sqrt{s}$. The sum of third absolute centered moments of the summands in sY is $O(\sqrt{s})$ by the same Riemann bound for a dominating profile $Cx^3/(1+e^x)$. Lyapunov's theorem now gives $Z/s^{1/4} \Rightarrow N(0, V)$.

For each fixed R_0, the characteristic function of Z under the independent tilt satisfies

$$|\mathbb{E}_s e^{i\xi Z}| \le e^{-c_{R_0}\sqrt{s}\,\xi^2}, \qquad |\xi| \le R_0.$$

Use just the $\gg_{R_0} \sqrt{s}$ terms with $x = s/(\pi h)^2$ in a compact positive interval $[a_0, b_0]$ with $R_0 b_0 < \pi/2$. On these terms the tilt probabilities are bounded away from zero and one, and the squared single factor modulus is $1 - 2q_h(1-q_h)(1-\cos(\xi x))$. This gives the bound.

The characteristic function of the tilted arithmetic variable $s(B_{p,a} - t)$ uses $\Lambda_p(s(1 - i\xi))/\Lambda_p(s)$ and a unit scalar. By (27) and the grouping comparison it differs from the model tilted function by

$$p^{-1/2} \exp(O_{R_0}(1 + \sqrt{s})) + O_{R_0}(s/p),$$

uniformly for $|\xi| \le R_0$. Because $\sqrt{s} = o(\log p)$, this tends to zero even after multiplication by $s^{1/4}$. It follows by Fourier inversion, the substitution $\xi = u/s^{1/4}$ and dominated convergence that for every continuous integrable G of fixed compact Fourier support,

$$s^{1/4}\mathbb{E}_{p,s}\, G(s(B_{p,a} - t)) \longrightarrow \frac{1}{\sqrt{2\pi V}} \int_{\mathbb{R}} G(x)\, dx. \tag{29}$$

Indeed the model Fourier integrand after substitution is dominated by a Gaussian, and the arithmetic Fourier integral of the discrepancy is negligible.

Apply this to upper and lower bandlimited approximations to $\psi(x) = e^x \mathbf{1}_{\{x \le 0\}}$. One construction integrates the Selberg upper and lower interval functions for $[u, 0]$ against $e^u\, du$ on $u \le 0$. The interval functions have Fourier support in $[-R_0, R_0]$ and integrals $|u| \pm O(1/R_0)$ in the chosen Fourier convention, as in [33]. Their one-sided bounds hold at interval endpoints by continuity from inside for the majorant and outside for the minorant. The integrated functions bound ψ, since $\int_{-\infty}^{0} e^u \mathbf{1}_{[u,0]}(x)\, du = \psi(x)$. The L^1 norms are integrable in u under that weight, since the distance in L^1 from the interval indicator is $O(1/R_0)$. The approximation integrals are $1 \pm O(1/R_0)$. Direct extremizers for truncated exponentials are also given by Carneiro and Littmann [34, Theorem 2]. Taking the limit in (29) for fixed R_0 and then $R_0 \to \infty$ yields the same limit with ψ and integral one.

Undoing the change of measure now gives

$$A_p(t) = \Lambda_p(s)e^{st}\mathbb{E}_{p,s}\psi(s(B_{p,a} - t)) \sim e^{-D^2/(4t)}\, \frac{s^{-1/4}}{\sqrt{2\pi V}} = \sqrt{\frac{t}{\pi I}}e^{-I/t}.$$

This completes the proof of Theorem A.2.

A limitation of the moment comparison. For the independent finite-residue mask with weights w_r, the same real product calculation gives logarithmic probability $-I/t + o(t^{-1})$ at $t = A/\log p$ for every fixed $A > 0$. If $A < 2I$, that probability including the endpoint condition is $o(p^{-1/2})$. Removing the event and renormalizing changes every bounded monomial expectation by $o(p^{-1/2})$. Thus bounds of the Weil comparison's size on individual mask moments alone do not guarantee any point in that smaller event. This observation is about the moment information, independent of whether the actual arithmetic mask has a point there.

A.5 Long residue runs at moderate depths

Graham and Ringrose [35, Theorem 1] show that infinitely many p have least quadratic nonresidue at least $c \log p \log\log\log p$. Choose M a sufficiently small fixed multiple of that expression, rounded down, and put $b_0 = 2M$. Then $b_0 - h$ is a nonzero square for all $|h| \le M$.

If a is a square, this residue allows all those frequencies. The conditioned lower theorem of Section A requires compatibility only through its dense cutoff $m = O_H(\epsilon^{-2})$. Hence it applies to every $g \in \mathcal{H}_H$ at $\epsilon \ge D_H/\sqrt{M}$ and $J \ge C_H M$ with suitable constants. This one residue has mass $1/(p-1)$, so its lower bound gives the complete lower $\exp(-C'_H\epsilon^{-2})$ as long as $\epsilon \le D'_H/\sqrt{\log p}$. The larger radii follow from the complete wave theorem. Since $M/\log p \to \infty$, this gives coverage at a smaller order on these primes.

If a is a nonsquare, the same b_0 suppresses all modes $|h| \le M$, including zero. Put $T = M/\log p$ and choose $K = A(\log T)^2$ for a sufficiently large fixed A. We show that already at $n \ge C \log p$ the conditioned zero-center ball at radius $C_1 K/\sqrt{M}$ has mass at least $\exp(-C_2 m_*)$, where

$$m_* = \lceil M e^{-c_1\sqrt{K}} \rceil = o(\log p)$$

for an appropriate fixed $c_1 > 0$. Notice that $m_* = M^{1-o(1)}$ since $T \asymp \log\log\log p$. This leaves a large range of spectral degrees even though all the lowest frequencies are absent.

We describe the budget adjustment in the conditioned torus wave proof of Section A. Its high-band killed test on a block of scale U with normalized sample rows and threshold τ has relative energy $O(1 + Ue^{-c\sqrt{\tau}})$ at angular speed $\sqrt{\tau}$. This follows from the surviving path bound for independent circle Brownian angles and cyclic trace colors by the trajectory discrepancy estimate of that proof. No density of active slots is needed for this lower bound. In our case there is no dense term to pay. Retain bands from $U \asymp M$ to a dyadic cutoff $B \asymp M\log(2M)$, and take

$$\tau_U = K(U/M)^{1/4}.$$

The sum of relative energies is $O(m_*)$ by dyadic summation and a decrease in c_1. The fixed propagation radius still fits the sample widening since row Lipschitz constants are $O(\sqrt{\tau_U})$. The spectral logarithmic partition sum is $O(\sum_U Ue^{-c\tau_U}) = O(m_*)$. Use the usual compactly supported sinc filter and project to eigenvalues at most a sufficiently large constant times m_*. The relative L^2 leakage is $e^{-c\sqrt{m_*}}$, which beats the $\exp(O(\log^2 M))$ growth of the even-power sample penalty. The resulting positive weight and signed test have Fourier absolute norm $\exp(O(m_*))$. After the penalty and the intermediate moment insertions each term uses $O(m_* + \log M)$ shifts of polynomial mode degree. The sum of the retained block radii is $O(K/\sqrt{M})$.

For clarity, the conditioned local phase comparison used here is the one in the appendix of Section A. Through $N = \lfloor e^J/16 \rfloor$ its discrepancy per participating mode is at most

$$\exp\big(-cn\log p/(m_* + \log M)\big)$$

when $m_* + \log M \le a_0 J$. The hypotheses hold uniformly for $J \ge C \log p$ on the chosen sequence. The depth condition $n \ge C' \log p$ ensures this after increasing C'. The inserted balanced moments of order $2r$ with $r = \lceil \log(16M) \rceil$ above B are independent of the weight in the reference model. Their summed supremum roots are $O(M^{-1/2})$, since $B \asymp M \log M$. The arithmetic errors remain negligible after the roots and the at most $O(J)$ blocks. Indeed $n \log p \ge J^2$, $m_* + \log M = o(\log p)$ and the Fourier weight costs $\exp(O(m_*))$. The fixed-fourth-moment terminal tail past N is $\exp(-cJ)$ in expectation and absorbs that same weight. Therefore the signed expectation and Markov transfer apply with m_* in the probability cost, proving the stated conditioned lower at radius $C_1 K/\sqrt{M}$.

On averaging over all unit residues its probability is at least $\exp(-C_2 m_*)/(p-1)$. But

$$\frac{\epsilon^{-2}}{\log p} \asymp \frac{T}{(\log T)^4} \longrightarrow \infty, \qquad \frac{m_*}{\log p} \longrightarrow 0$$

by the choice of A. This proves the second assertion of Theorem A.3. All nonsquare lifts have the same bound, and the nonsquare scales form half the units. The two-parameter conclusion follows by averaging. Since a uniform ball is contained in the ball of the same radius for any finite L^r norm, the obstruction also applies to those complete-family upper bounds.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, proof checking, and manuscript preparation. No data or numerical certificates are required for the results.

B Boundary-scale hyper-Kloosterman traces

We estimate the number of normalized odd-rank hyper-Kloosterman traces in disks about the boundary of the compact trace image. The two-sided comparison is uniform as the target approaches a cusp. At fixed targets we obtain Haar asymptotics with power slack. The Haar mass has three regimes, determined by the radius and the separation of the two distinct eigenvalues at the target. Writing Δ for the eigenvalue separation, the sufficient radius is a constant times the larger of $p^{-1/(R^2-1)}$ and $p^{-1/R^2}\Delta^{2/R^2}$. The proof uses anisotropic estimates for the Weyl denominator. The latter exponent is sharp for dimension-weighted character moment bounds alone. The same comparison gives localized counts at arbitrary fixed interior traces.

Keywords hyper-Kloosterman sums, shrinking targets, Weyl characters, trace figures, cusp.
Mathematics Subject Classification 11L05, 11T23, 22E46, 60B15.

B.1 Statements

For odd primes p and fixed rank $R \ge 2$, put

$$\mathrm{Kl}_R(a;p) = p^{-(R-1)/2} \sum_{x_1 \cdots x_R = a} \exp(2\pi i (x_1 + \cdots + x_R)/p) \qquad (a \in \mathbb{F}_p^\times),$$

with all summation variables nonzero. We take p sufficiently large in terms of R. Let G be SU_R for odd R and USp_R for even R. Write μ for the Haar trace law and μ_p for the probability measure of $\mathrm{Kl}_R(a;p)$ as a is uniform on $\mathbb{F}_p^\times$. In even rank the conventional sign change between the sheaf trace and the sum is implemented by the central element $-I$ and preserves Haar. For $z \in \mathbb{C}$ let $B(z,\epsilon)$ be the open disk, or its intersection with the line when the trace is real.

When R is odd let $d = R^2 - 1$ and $k = R - 1$. The trace image of SU_R has the hypocycloid boundary

$$z = ku + v, \qquad u = e^{it}, \quad v = e^{-ikt}.$$

Let $\Delta = |u - v|$ and $\alpha = (k^2 + 1)/2$. The case $\Delta = 0$ is a cusp. For fixed sufficiently small positive constants c, c' define a representative of the mass scale

$$V(\epsilon, \Delta) = \begin{cases} \epsilon^{\alpha} \Delta^{2k-\alpha}, & \epsilon \le c\Delta^3, \\ \epsilon \Delta^{d-2}, & c\Delta^3 < \epsilon < c'\Delta^2, \\ \epsilon^{d/2}, & \epsilon \ge c'\Delta^2. \end{cases}$$

Here c may be taken sufficiently smaller than c' so the regimes are ordered for $0 \le \Delta \le 2$. At a cusp use the last expression. Changing the break constants by fixed factors changes the scale only by fixed factors.

Theorem B.1. *For each fixed odd $R \ge 3$, uniformly in boundary targets z and sufficiently small $\epsilon > 0$,*

$$\mu(B(z, \epsilon)) \asymp_R V(\epsilon, \Delta).$$

For a sufficiently large constant C_R, whenever

$$\epsilon^{R^2/2} \ge C_R p^{-1/2} (\Delta + \sqrt{\epsilon}), \tag{30}$$

one also has $\mu_p(B(z, \epsilon)) \asymp_R V(\epsilon, \Delta)$. At a fixed noncusp point, if $\epsilon \to 0$ and $\epsilon \ge p^{-1/R^2 + \eta}$ for any fixed $\eta > 0$, the ratio $\mu_p(B(z, \epsilon))/\mu(B(z, \epsilon))$ tends to one. At fixed cusps the analogous threshold is $\epsilon \ge p^{-1/d+\eta}$.

Proposition B.2. *For either parity of R, let $d = \dim_{\mathbb{R}} G$ and let b be the real target dimension (one in even rank and two in odd rank). Suppose z has a submersive trace preimage in G. Then $\mu_p(B(z, \epsilon)) \gg_{R,z} \epsilon^b$ down to $\epsilon \ge C_{R,z} p^{-1/(d+b)}$. If every preimage is submersive, the count is asymptotic to Haar when $\epsilon \to 0$ and $\epsilon \ge p^{-1/(d+b)+\eta}$. At a central scalar cusp in either group the ball count is of order $\epsilon^{d/2}$ down to $C_R p^{-1/d}$, with the corresponding asymptotic at power slack.*

Kowalski and Untrau [15, Proposition 4.6 and Lemma 4.7] use a submersive trace preimage and the compact-group Wasserstein bound to obtain counts at fixed odd-rank interior targets down to radius of order $p^{-1/(3d)}$. Fouvry and Michel [37, Sections II.3–II.5 and equation (2.21)] give the dimension-weighted localization exponent at trace zero through a Weyl–Sobolev estimate. Proposition B.2 extends that order to arbitrary fixed submersive targets and gives relative asymptotics at regular values. Fu, Lau and Xi [39, Theorem 1.2] prove absolute discrepancy bounds for small alcove boxes, whose lifts have disjoint Weyl translates. Here the trace ball has mass varying with both position and radius, and the relevant estimate is relative to that mass. Kaiser [40, Sections 2–3 and equation (9)] determines the trace image and gives the normalized invariant integration formula for SU_3. For its trace density see also Evans and Pugh [36, Theorem 7.2]. In rank two Rahimi [41, Theorem A(3) and Proposition 5.2] gives the arithmetic edge count using Feldheim and Sodin's semicircle moment discrepancy [38, Proposition 5]. The boundary comparison here concerns higher rank and moving trace targets.

B.2 Dimension-weighted comparisons

For a smooth central function F on G, write $\widehat{F}(\pi) = \langle F, \chi_\pi \rangle$ and $\|F\|_{\mathrm{ch}} = \sum_\pi d_\pi |\widehat{F}(\pi)|$, where π ranges over irreducible representations, χ_π is its character, and d_π its dimension. Katz's monodromy and the local ramification of the hyper-Kloosterman sheaf, as recorded for example in [9,

Proposition 4.6], give

$$\left|\mathbb{E}_a F(g_a) - \int_G F\, dg\right| \ll_R p^{-1/2}\, \|F\|_{\mathrm{ch}}\,, \tag{31}$$

where g_a is its compact trace class. Indeed a nontrivial irreducible representation applied to the sheaf has no geometric invariants, rank d_π and break at infinity at most $1/R$, because the higher ramification acts trivially on the original factor. It is tame at zero. The compactly supported first cohomology on $\mathbb{G}_m$ thus has dimension $O_R(d_\pi)$ by the Euler characteristic, and purity and the trace formula give the bound. The sum of the character estimates converges absolutely for the functions below.

Choose a bi-invariant metric, with eigenvalues λ_π for the nonnegative Casimir. For $0 < \rho \le 1$ and an integer $s > d/4$,

$$\|F\|_{\mathrm{ch}} \ll_G \rho^{-d/2} \left\|(1 + \rho^2(-\Delta_G))^s F\right\|_2 . \tag{32}$$

By Cauchy–Schwarz the factor squared is bounded by $\sum_\pi d_\pi^2 (1+\rho^2\lambda_\pi)^{-2s}$. Highest weights lie in a lattice of rank $\ell = \mathrm{rank}(G)$, the Weyl dimension polynomial has degree the number m_+ of positive roots, and $d = \ell + 2m_+$. On dyadic shells the Casimir grows quadratically. This proves that the sum is $O_G(\rho^{-d})$.

For Proposition B.2, choose a smooth compactly supported bump in a submersion chart at a preimage of z and multiply it by a cutoff of $(\mathrm{tr} - z)/\epsilon$. Its Haar mass is $\gg \epsilon^b$, and each derivative of order j has L^2 norm $O(\epsilon^{b/2-j})$. Average under conjugation, which preserves trace and Haar and commutes with the bi-invariant Laplacian. The function still lies below the indicator in question and has character norm $O(\epsilon^{(b-d)/2})$ by (32) with $\rho = \epsilon$. Equation (31) gives the lower bound. If z is a regular value, the compact fiber has a finite submersion chart cover, and the pushforward Haar density is smooth and positive near z. For fixed fractional collar width the same estimates apply to trace cutoffs bracketing the ball on the full group. Relative character errors tend to zero at power slack. Sending the collar width to zero gives the asserted asymptotic.

At a scalar central element ζI one has

$$\|g - \zeta I\|_{\mathrm{F}}^2 = 2(R - \Re(\overline{\zeta}\,\mathrm{tr}\, g)) \le 2|\,\mathrm{tr}\, g - R\zeta|.$$

The trace has zero differential at this point. In exponential coordinates $g = \zeta e^X$ its deviation equals $-\zeta\,\|X\|_{\mathrm{F}}^2/2 + O_R(\|X\|_{\mathrm{F}}^3)$. The Haar trace ball has mass $c_G \epsilon^{d/2}(1 + O_R(\sqrt{\epsilon}))$ by the group volume in these coordinates. Smooth radial cutoffs at radius of order $\sqrt{\epsilon}$ bracket the trace ball with any fixed relative collar once ϵ is small. With $\rho = \sqrt{\epsilon}$ their scaled Sobolev norms are $O(\epsilon^{d/4})$, hence their character norms are $O(1)$. Equation (31) proves the cusp part of the proposition, and with fixed collar then vanishing width gives the cusp asymptotic.

The scale at the cusp cannot be improved from (31) for individual characters alone. Remove from Haar the trace ball of radius $cp^{-1/d}$ and place its mass on any conjugacy class outside that ball. The mass moved is $O_G(c^{d/2}p^{-1/2})$, and the change in each character average is at most twice this mass times d_π. The resulting conjugacy-invariant probability has the same form of dimension-weighted character estimates at all degrees and no mass in the ball.

B.3 Uniform Haar mass at a moving boundary

The derivative of trace on the determinant-one torus is spanned over $\mathbb{R}$ by $i(z_j - z_R)$. Thus at least three distinct eigenvalues make the preimage submersive, as in [15, Lemma 4.7]. If there are two distinct eigenvalues u, v of multiplicities $k_1, k_2 \ge 2$, varying their common phases to $u_s = ue^{is/k_1}$ and $v_s = ve^{-is/k_2}$ and splitting one equal pair in each block realizes traces

$$k_1 u_s + k_2 v_s - r u_s - t v_s, \qquad 0 \le r, t \ll 1.$$

Bryant's argument [21] uses the two single splits to exhibit interiority. Here the s derivative is $i(u-v)$, while $-u$ and $-v$ have opposite nonzero normal projections onto $u-v$. The implicit function theorem allows s, r to vary with $t>0$ keeping the original trace, with $r'(0)=1$. The resulting group element has at least three distinct eigenvalues. Thus no boundary trace can have such a preimage. Every noncusp boundary target therefore has exactly one eigenvalue of multiplicity $R-1$ and one simple eigenvalue. The parameter of the unique class is determined by the boundary point, since the curve winds once and its argument derivative is

$$\frac{(R-1)(R-2)(1-\cos(Rt))}{|(R-1)e^{it}+e^{-i(R-1)t}|^2},$$

strictly positive except at isolated cusps. The same argument shows every interior trace has a submersive preimage, so Proposition B.2 applies there. At odd rank all traces with $|z|<1$ are regular values, since unequal multiplicities in a preimage with at most two eigenvalues force $|z|\geq 1$.

Near a noncusp class of eigenvalues u of multiplicity k and v simple, use torus coordinates

$$ue^{i(s/k+x_j)}\ (1\leq j\leq k),\quad ve^{-is},\qquad \sum_{j=1}^{k} x_j=0.$$

The trace deviation has expansion

$$i(u-v)s-\tfrac{1}{2}u\sum_j x_j^2+O_R(s^2+\|x\|^3+|s|\,\|x\|^2). \tag{33}$$

The first term is tangent and the second has nonzero normal projection of magnitude $\Delta\sum_j x_j^2/4$. Away from cusps this confines a trace disk to $|s|\ll\epsilon$, $\|x\|\ll\sqrt{\epsilon}$. Rescale $s=\epsilon S$, $x=\sqrt{\epsilon}X$. Weyl's integration formula supplies one squared Vandermonde factor for each within-cluster pair of angles, a contribution $\epsilon^{k(k-1)/2}$. The cross factors are separated from zero. The coordinate volume is $\epsilon^{1+(k-1)/2}$. Dominated convergence on the rescaled bounded domain gives the Haar asymptotic $c_{G,z}\epsilon^{\alpha}(1+o(1))$. The limiting boundary has measure zero and the limiting weighted volume is positive. Equivalently, integrating out the linear tangent coordinate and rescaling X in (33) gives $c_{G,z}=c_R\Delta^{2k-\alpha}$ for fixed noncusps. These estimates are uniform when Δ is bounded away from zero.

For small Δ rotate the cusp to R. The unitary norm identity in the previous section localizes all group elements with trace close to the target to a neighborhood of the identity. Write their torus eigenangles as $r\omega$, with $\sum\omega_j=0$ and $\|\omega\|_2=1$. Let $A=R-\Re\operatorname{tr} g$. Uniformly in ω,

$$A=r^2/2+O_R(r^4),\qquad \partial_r A=r+O_R(r^3).$$

The ratio $r(A,\omega)/\sqrt{2A}$ is smooth up to $A=0$ by solving for r^2. Therefore the normalized imaginary part $H_A(\omega)=A^{-3/2}\Im\operatorname{tr} g$ has a smooth limit, with uniform fixed-order angular derivatives:

$$H_0(\omega)=-2^{3/2}\sum_j\omega_j^3/6.$$

On the mean-zero unit sphere the cubic has extrema at configurations with at most two values by Lagrange multipliers. For multiplicities $j, R-j$ the magnitude is $|R-2j|/\sqrt{Rj(R-j)}$, maximized for a singleton. At the maximizing pattern of the positive cubic with singleton $a>b$ and repeated value b, the constrained Hessian on tangent directions of the repeated group is $3(b-a)$ times the identity. Thus it is strictly negative. The sign-reversed pattern applies to H_0. Permutation symmetry of the repeated group forces the same directions to be critical for H_A. Strict local concavity and

the gap away from these finitely many directions imply that, for small A, the maximum persists there and the normalized imaginary deficit is comparable to squared angular distance near the maxima. Use conjugation for the lower boundary. The actual upper boundary as a function of A is $A^{3/2} \max H_A$, with derivative $O_R(\sqrt{A})$.

For the target let $A_0 = R - \Re z$, so $A_0 \asymp_R \Delta^2$ near the cusp. If $\epsilon \le c\Delta^3$, membership in the trace ball requires $|A - A_0| \le \epsilon$ and angular distance at most $C_R\sqrt{\epsilon/\Delta^3}$ from one maximizing direction. Conversely smaller constants in both inequalities place the trace in the disk, by the boundary derivative bound. In polar torus variables the Weyl density is comparable to

$$r^{R(R-1)} \prod_{i<j} |\omega_i - \omega_j|^2 \, r^{R-2} \, dr \, d\omega.$$

Changing the radial variable to A yields a radial factor comparable to $\Delta^{d-2} \, dA$. For a small angular cap of radius b_1 at the maximizing direction, its weighted angular measure is of order $b_1^{k^2-1}$. Indeed the tangent dimension is $k-1$ and the within-cluster squared Vandermonde vanishes to order $k(k-1)$, with a matching lower on a fixed subcone. Multiplying the radial shell and angular contributions gives

$$\epsilon\Delta^{d-2}(\epsilon/\Delta^3)^{(k^2-1)/2} = \epsilon^{\alpha}\Delta^{2k-\alpha}.$$

When ϵ lies between constants times Δ^3 and Δ^2, a fixed smaller angular cap gives a lower fraction and the entire angular sphere gives the upper. The radial shell then gives order $\epsilon\Delta^{d-2}$. Once ϵ is comparable to Δ^2 or larger, the group lies in a radial ball of radius $O_R(\sqrt{\epsilon})$, and a smaller shell (in the transition) or a smaller radial ball supplies the lower of order $\epsilon^{d/2}$. This completes the model estimate in Theorem B.1.

B.4 Character costs at the boundary

Let F be a smooth trace cutoff, bounded in $[0,1]$, supported on the disk of radius a fixed multiple of ϵ, and equal to one on a concentric smaller disk. Use such functions to upper and lower bracket the indicator. The preceding section shows that their Haar integrals are of order V. We will show

$$\frac{\|F\|_{\mathrm{ch}}}{V} \ll_R (\Delta + \sqrt{\epsilon})\epsilon^{-R^2/2}. \tag{34}$$

The constants may depend on the fixed collar parameters.

First take ϵ at least a small constant times Δ^2. Near a cusp the first derivatives of trace are then $O_R(\sqrt{\epsilon})$ on the support and higher derivatives bounded. Equation (32) with $\rho = \sqrt{\epsilon}$ gives norm $O_R(1)$ since each scaled derivative has L^2 norm $O_R(\sqrt{V})$ and $V \asymp \epsilon^{d/2}$. In the middle regime take $\rho = \epsilon/\Delta$. At every point in the trace collar the first trace derivative is $O_R(\Delta)$, the gradient vanishing at the central scalar. Higher derivatives are bounded and ρ^2/ϵ is bounded. Equation (32), using support Haar mass $O_R(V)$, yields

$$\frac{\|F\|_{\mathrm{ch}}}{V} \ll_R \frac{(\Delta/\epsilon)^{d/2}}{\sqrt{\epsilon\Delta^{d-2}}} = \Delta\epsilon^{-R^2/2}.$$

Both estimates remain valid at constant overlaps of the regimes.

It remains to treat $\epsilon \ll \Delta^3$ and small $\Delta > 0$. Use torus coordinates (s, x) from (33) at the finitely many permutations of the target. The angular cap confines $\|x\| \ll_R \sqrt{\epsilon/\Delta}$, and the preliminary radial and angular constraints imply

$$|s| \ll_R \Delta\sqrt{\epsilon/\Delta^3} + \epsilon/\Delta + \Delta^3.$$

By taking the thin-regime constant and the cusp neighborhood small, this is a sufficiently small multiple of Δ. The real-trace derivative in s then has magnitude comparable to Δ, since at zero it equals $\Re(i(u-v))$ and the eigenvalues are near one. At $s=0$ the real trace shift caused by x is $O_R(\|x\|^2)$. Hence $|s| \ll_R \epsilon/\Delta^2$. For each fixed x the support slice in s nevertheless has length $O_R(\epsilon/\Delta)$ by monotonicity of the real trace. Thus the torus Lebesgue volume of the support is

$$O_R((\epsilon/\Delta)(\sqrt{\epsilon/\Delta})^{k-1}).$$

Following the Weyl conversion used by Fouvry and Michel [37, Section II.5], write A_ρ for the Weyl denominator, with ρ the half sum of positive roots in this notation. The character coefficient of central F equals the torus Fourier coefficient of FA_ρ at the weight $\lambda+\rho$, for highest weight λ. Indeed the Weyl character and integration formulas first express it as $|W|^{-1}$ times the signed sum over $w \in W$ of these Fourier coefficients at $w(\lambda+\rho)$, where W is the Weyl group. Anti-invariance of FA_ρ makes the summands equal after their signs.

Use derivative scales

$$\rho_s = \epsilon/\Delta, \qquad \rho_x = \sqrt{\epsilon\Delta}.$$

For a local piece of FA_ρ, each fixed-order derivative scaled by these lengths is bounded on its support by

$$O_R(\Delta^k(\sqrt{\epsilon/\Delta})^{k(k-1)/2}).$$

For F this follows from $\partial_s \operatorname{tr} g = O_R(\Delta)$, $\partial_x \operatorname{tr} g = O_R(\|x\|)$ and bounded higher derivatives. Products of at least two differential scales divided by ϵ stay bounded since $\epsilon \ll \Delta^3$. In the denominator the k cross factors are of order Δ, and each within-cluster factor is $O_R(\sqrt{\epsilon/\Delta})$. The scaled derivatives preserve these bounds. One can partition the torus near the distinct permuted centers using cutoffs with derivatives at scale $c_R\Delta$; both lengths above are smaller, so the same estimates apply. With the torus support volume, the scaled Sobolev L^2 norm of each piece is $O_R(\sqrt{V})$.

For each piece apply Cauchy–Schwarz to the weighted Fourier coefficients with reciprocal scaled Sobolev weights. The weight lattice has $O_R(B_s B_x^{k-1})$ points in a box of dual lengths $B_s = \Delta/\epsilon$, $B_x = (\epsilon\Delta)^{-1/2}$, as follows by thickening with a bounded fundamental cell. The same bound times the dilation to the power k holds for dilates. The Weyl dimension polynomial has $k(k-1)/2$ root factors that see only the x directions and at most k cross factors that cost $O_R(B_s)$ on this box, since $B_s \gg B_x$. Taking a fixed sufficiently high Sobolev order to sum the dilates gives

$$\|F\|_{\mathrm{ch}} \ll_R \sqrt{V}\,(B_s B_x^{k-1})^{1/2} B_x^{k(k-1)/2} B_s^k.$$

Root products in absolute value are invariant under the finite coordinate permutations used in the partition. Substitute the thin-regime expression for V. The ratio is $O_R(\Delta\epsilon^{-R^2/2})$, which is (34). Away from cusps the same argument uses scales $\rho_s \asymp_R \epsilon$, $\rho_x \asymp_R \sqrt{\epsilon}$ in (33), with Δ separated from zero. This also supplies (34) there.

Combine (34) with (31) and the fixed smooth brackets. Under (30) their comparison errors are below a sufficiently small fraction of V. This proves the uniform two-sided count. At a fixed noncusp the asymptotic Haar mass $c_{G,z}\epsilon^\alpha$ lets the fractional collar width tend to zero after taking the limit with power slack, since the arithmetic error for each fixed collar is then $o(V)$. At cusps the same argument uses the radial volume asymptotic already proved.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, proof drafting and internal checks. No data or computational certificates are required.

C Cyclotomic lattice transport at prime powers

Cyclotomic phase vectors modulo $q = p^n$ have exact Wasserstein order $q^{-1/\varphi(d)}$, including W_∞, even when the parameter is restricted to units. This gives shrinking trace-disk counts at the sharp radius $q^{-2/\varphi(d)}$ at every aligned tip. For prime d we obtain uniform boundary estimates through the transition to the cusps. In degree two, Gaussian and Eisenstein lattice shells give exact eventual counts at the microscopic radius q^{-1}. A trace-dual ideal description also gives the law of microscopic counts as the depth varies at a fixed prime. The results apply to tame hyper-Kloosterman sums whenever the stationary trace has a common Gauss factor.

Keywords cyclotomic lattices, Gaussian periods, hyper-Kloosterman sums, shrinking targets.
MSC 2020 11L05, 11H30, 11R18, 60B15.

C.1 Phase lattices and trace disks

Fix an integer $d > 1$, a prime $p > d$ with $d \mid p-1$, and $q = p^n$ with $n \geq 2$. Let $k = \varphi(d)$ and let ζ be a primitive dth root in $\mathbb{Z}_p$. Write ζ_d for a primitive complex dth root. In the power basis of $\mathbb{Z}[\zeta_d]$, let $v_j \in \mathbb{Z}^k$ be the coefficient vector of ζ_d^j for $0 \leq j < d$. On $\mathbb{T}^k = \mathbb{R}^k/\mathbb{Z}^k$ set

$$T_d(\theta) = \sum_{j=0}^{d-1} e(v_j \cdot \theta), \qquad e(x) = \exp(2\pi i x).$$

Let $\nu_{q,d}$ be the law of $\theta_q(s) = s(1, \zeta, \dots, \zeta^{k-1})/q \bmod \mathbb{Z}^k$ with s uniform among units modulo q. Thus $T_d(\theta_q(s)) = \sum_{j=0}^{d-1} e_q(s\zeta^j)$, where $e_q(x) = e(x/q)$. Taking all s instead gives the full grid, and $s \equiv s_0 \pmod p$ gives an affine subgrid. Use the Euclidean quotient metric on the torus, and write m_d for Haar measure there.

Theorem C.1. *Uniformly in p, n, the Wasserstein distance of $\nu_{q,d}$ to m_d has order $q^{-1/k}$ for every $1 \leq s \leq \infty$. The same order holds for the full grid. In a fixed residue class of the parameter modulo p the order is $(q/p)^{-1/k}$.*

The arithmetic interpretation is as follows. Fix $R \geq 2$, $p > R$, and $d = \gcd(R, p-1)$. A normalized rank-R hyper-Kloosterman sum at a unit c is

$$\mathrm{Kl}_R(c; q) = q^{-(R-1)/2} \sum_{x_1 \cdots x_R = c} e_q(x_1 + \cdots + x_R).$$

If c is not a dth power modulo p the sum vanishes. Otherwise the stationary points are $x_1 = \cdots = x_{R-1} = u$ with $u^R = c$. The Hessian at each point in the $R-1$ free variables is $u^{-1}(I + \mathbf{1}\mathbf{1}^{\mathsf{T}})$, with determinant $Ru^{-(R-1)}$. At even depth each root contributes $e_q(Ru)$. At odd depth an additional Gaussian factor $\varepsilon_p^{R-1}\chi(2^{-(R-1)}R)\chi(u)^{R-1}$ comes from the half-Hessian, where χ is the Legendre symbol and $\varepsilon_p = 1$ or i according as $p \equiv 1$ or $3 \pmod 4$. This is the tame stationary-phase evaluation described by Smith, Dbrowski–Fisher and Van Order [23, 24, 44]. When R is odd the Gaussian factor is independent of the root. Consequently, at even depth or at odd R, the law on active c is the pushforward of $\nu_{q,d}$ by T_d up to a common unit factor. Indeed one may take the root uniform among units, since each active c has d roots, and multiplication of that root by R permutes the units. For a small disk about a nonzero aligned tip the complete arithmetic probability is $1/d$ times the disk probability conditional on active parameters. A fixed inverse coefficient in a path does not change the endpoint law under complete averaging in the linear coefficient. For residue-conditioned active c one may use one root branch and the affine lattice for that residue, with its fixed Gaussian factor when that factor is common to the branches.

Call $d\omega$ an aligned tip if all the characters can equal ω. This occurs exactly when $\omega^{\Phi_d(1)} = 1$, where Φ_d is the cyclotomic polynomial. Set

$$C_d = \frac{\mathrm{vol}(B^k)\sqrt{|\operatorname{disc}(\mathbb{Q}(\zeta_d))|}}{(2\pi^2 d)^{k/2}},$$

where B^k is the Euclidean unit ball.

Theorem C.2. *For every aligned tip and all small $\epsilon > 0$,*

$$\nu_{q,d}\{|T_d - d\omega| < \epsilon\} = C_d\epsilon^{k/2} + O_d\Big(\epsilon^{k/2+1} + q^{-1/k}\epsilon^{(k-1)/2} + q^{-1}\Big).$$

This probability is zero when $\epsilon < c_d q^{-2/k}$ for a sufficiently small $c_d > 0$, and has order $\epsilon^{k/2}$ when $\epsilon \geq D_d q^{-2/k}$ for a sufficiently large D_d. It is asymptotic to the displayed main term if $\epsilon \to 0$ and $\epsilon q^{2/k} \to \infty$.

Below we give exact microscopic counts for $d = 2, 3, 4, 6$, a limiting depth distribution of microscopic counts for higher degree, and moving-boundary counts at prime d.

Duke, Garcia and Lutz [42, Theorems 1.1 and 6.3] and Garcia, Hyde and Lutz [43] identify Gaussian period images, including the hypocycloid at prime order. Untrau [25, Theorem A, Proposition B and Section 5.3] gives the Laurent torus and qualitative equidistribution for scalar sums over admissible prime powers, with full additive and unit parameter families among the cases. Kowalski and Untrau [15, Proposition 3.4] obtain the W_1 rate for a full prime-ideal residue field by a norm bound and a Fourier cutoff. The same norm geometry here leads to a lattice coupling at the W_∞ scale. The local trace behavior and the integer counts at the smallest scale then require separate geometric arguments.

C.2 The lattice coupling

Let $v = (1, \zeta, \ldots, \zeta^{k-1})$ represented modulo q, and put $\Lambda = \mathbb{Z}^k + q^{-1}v\mathbb{Z}$. Its dual Λ^* consists of the integer vectors m with $m \cdot v \equiv 0 \pmod q$. It has index q in $\mathbb{Z}^k$ because the first coordinate of v is one. Such a vector m corresponds to the cyclotomic integer $\alpha = \sum_{i=0}^{k-1} m_i\zeta_d^i$. This is nonzero for $m \neq 0$. Under the embedding into $\mathbb{Q}_p$ given by the chosen root ζ, its p-adic valuation is at least n. All conjugates are integral, so q divides its absolute norm. Since that norm is $O_d(\|m\|^k)$, the shortest nonzero vector in Λ^* has length $\gg_d q^{1/k}$.

By Minkowski's second theorem [46], all k successive minima of Λ^* have order $q^{1/k}$. Here is a direct consequence for the primal lattice. Choose k independent dual vectors of norm $O_d(q^{1/k})$. They span a sublattice of index $O_d(1)$ in Λ^* by the determinant bound. Their dual basis has norms $O_d(q^{-1/k})$ by the cofactor formula, and multiplying it by that index gives vectors spanning a sublattice of Λ. Its fundamental parallelepiped has diameter $O_d(q^{-1/k})$. Thus the covering radius of Λ has that upper bound. Also any vector of Λ shorter than $c_d q^{-1/k}$ pairs to zero with each chosen independent dual vector and is itself zero. The same estimates apply to $\Lambda_0 = \mathbb{Z}^k + (q/p)^{-1}v\mathbb{Z}$ at scale $(q/p)^{-1/k}$. The class of v in the definition of Λ_0 is reduced modulo q/p.

Voronoi quantization couples Haar on the torus to the uniform full grid $\Lambda/\mathbb{Z}^k$ within its covering radius. To delete the points of $\Lambda_0/\mathbb{Z}^k$, choose $\xi \in \Lambda \setminus \Lambda_0$ of norm $O_d(q^{-1/k})$. For sufficiently large p a shortest nonzero fine vector is outside Λ_0, because the coarse shortest length is larger by a factor $\gg_d p^{1/k}$. At the bounded remaining primes choose any nontrivial coarse coset in Λ and subtract a nearest coarse point. The coarse covering radius still has order $q^{-1/k}$ with a rank-dependent constant in this finite set. The class of ξ generates $\Lambda/\mathbb{Z}^k$, since its index in the original cyclic

parametrization is prime to p. In its cyclic order partition the full grid into successive groups of p points, with the deleted point first in each group. Number these positions $j = 0, \dots, p-1$. With original mass one per point, send mass $1 - j/(p-1)$ from j to $j+1$ for $0 \le j < p-1$, retaining the rest in place. This is a one-step coupling: the deleted point has final mass zero, and each other point has final mass $1 + 1/(p-1)$. The cost is at most $\|\xi\|$. Rescaling gives the unit law. The same Voronoi quantization for a translated coarse grid gives the fixed-residue assertion.

For lower bounds, the union of balls of radius $cM^{-1/k}$ about M points covers at most half of Haar for small fixed c. In any coupling at least half its mass must then move that far. Apply this to the support sizes of the grids. It proves Theorem C.1 for all Wasserstein orders.

C.3 Aligned tips and exact lattice shells

The integer relations among the phase coordinates are exactly the multiples of Φ_d in the polynomial ring of integer exponents. Thus a common value ω is possible if and only if $\omega^{\Phi_d(1)} = 1$. A trace of modulus d has just one phase tuple over it, by equality in the triangle inequality. It corresponds to one point of $\mathbb{T}^k$ because the first k phases are the torus coordinates. The number of aligned tips is $\Phi_d(1)$, which equals a prime ℓ for d an ℓ-power and equals one otherwise.

Write θ for the local offset from that point and $t_j = 2\pi v_j \cdot \theta$. Then $\sum_j t_j = 0$, since the dth roots sum to zero. Set $S_i = \sum_j t_j^i$. The real defect and imaginary part of the trace, after rotation by $\overline{\omega}$, give

$$|T_d - d\omega| = \frac{S_2}{2} - \frac{S_4}{24} + \frac{S_3^2}{36 S_2} + O_d(\|\theta\|^6) \qquad (\theta \neq 0). \tag{35}$$

Here the imaginary part starts in degree three. Expanding the modulus around real defect $S_2/2$ gives the displayed degree-four terms with uniform remainder. In particular the quartic correction is strictly negative for every nonzero sufficiently small offset: $S_3^2 \le S_2 S_4$ gives correction at most $-S_4/72$, of order $-\|\theta\|^4$ since the forms span. This sign will determine membership at exact quadratic shell radii.

Put $Q(\theta) = S_2/2 = 2\pi^2 \sum_j (v_j \cdot \theta)^2$. The disk is squeezed by ellipsoids $Q(\theta) < \epsilon(1 + O_d(\epsilon))$, with the corresponding inner bound. All its points are in the tip chart for small ϵ, by uniqueness of the fiber. If E is the embeddings matrix of the power basis, and V has columns v_j, the rows of EV are the distinct primitive exponential rows on the dth roots of unity. Hence $(EV)(EV)^* = dI$ and

$$\det(VV^{\mathsf{T}}) = d^k / |\operatorname{disc}(\mathbb{Q}(\zeta_d))|.$$

The leading Haar volume of the ellipsoid is $C_d \epsilon^{k/2}$ as stated. The error from (35) is $O_d(\epsilon^{k/2+1})$.

By Theorem C.1, the angular movement in coupling to the unit lattice is at most $\delta = O_d(q^{-1/k})$. In the norm $Q^{1/2}$ the coupling changes radii by at most $C'_d \delta$. The ellipsoid bounds therefore give the stated arithmetic error, including the $O_d(q^{-1})$ term when the inner radius disappears. To see the empty disk below that scale, every tip point has torus coordinates of denominator dividing $\Phi_d(1)$, a fixed integer coprime to p. Multiply its distance to a unit lattice point by that denominator. Any local difference then represents a nonzero fine lattice vector because the unit point has order q in the torus. The shortest-vector bound forces angular distance $\gg_d q^{-1/k}$. The positive quadratic defect and the tip chart finish the proof of Theorem C.2.

C.3.1 Exact counts for $d = 2, 3, 4, 6$

Let $N_{q,\omega}(\lambda)$ count the *root parameters* s modulo q that are units and satisfy $|T_d(\theta_q(s)) - d\omega| < \lambda/q^{2/k}$. In the table, $\mathcal{O} + \alpha$ is an affine lattice in $\mathbb{R}$ or $\mathbb{C}$, and c is the constant multiplying squared modulus.

For all sufficiently large q, at every fixed $\lambda > 0$ the count is

$$\#\{z \in \mathcal{O} + \alpha : c|z|^2 \le \lambda\} - \#\{z \in \mathcal{O} + \alpha : c\,p^{2/k}|z|^2 \le \lambda\}. \tag{36}$$

The inequalities on the right include equality even though the trace disk is open.

d	$d\omega$	$\mathcal{O}$	α	c
2	2	$\mathbb{Z}$	0	$4\pi^2$
2	-2	$\mathbb{Z}$	$1/2$	$4\pi^2$
3	3	$\mathbb{Z}[\omega_3]$	0	$4\pi^2$
3	$3\omega_3,\ 3\omega_3^2$	$\mathbb{Z}[\omega_3]$	$(1-\omega_3)/3$	$4\pi^2$
4	4	$\mathbb{Z}[i]$	0	$4\pi^2$
4	-4	$\mathbb{Z}[i]$	$(1-i)/2$	$4\pi^2$
6	6	$\mathbb{Z}[\omega_6]$	0	$8\pi^2$

Here $\omega_j = e(1/j)$. In particular the first scaled radii for $d = 3$ converge to $4\pi^2$ at the real tip and $4\pi^2/3$ at the other two. The corresponding nonzero limits for $d = 2, 4, 6$ are respectively $4\pi^2, \pi^2$, then $4\pi^2, 2\pi^2$, then $8\pi^2$. In the hyper-Kloosterman application the number of active parameters is $N_{q,\omega}(\lambda)/d$ after any common rotation.

For $d = 2$ the coordinate is the ordinary circle angle, the trace is $2\cos(2\pi\theta)$, and the fine and coarse grids have spacings q^{-1} and p/q. The tip coordinates are zero and one half, with the latter mapping to the half-integer coset after scaling either grid because p is odd. This gives (36) by (35).

For $d = 3, 4, 6$, let ω_d be the complex primitive root and use $z = x - \omega_d y$ for the two power-basis angles. For $d = 3$ the quadratic form is $4\pi^2(x^2 + xy + y^2)$, for $d = 4$ it is $4\pi^2(x^2 + y^2)$, and for $d = 6$ it is $8\pi^2(x^2 - xy + y^2)$. These are $c|z|^2$ in the table. Under this map the fine lattice is the fractional ideal $\mathcal{O} + q^{-1}(1 - \zeta\omega_d)\mathbb{Z}$. Indeed write $\Phi_d(X) = X^2 - uX + 1$ with $u = -1, 0, 1$ for $d = 3, 4, 6$ respectively. Multiplication of $1 - \zeta\omega_d$ by ω_d gives $\zeta + (1 - u\zeta)\omega_d$, congruent to $\zeta(1 - \zeta\omega_d)$ modulo $q\mathcal{O}$ by $\Phi_d(\zeta) \equiv 0 \pmod q$. Hence this lattice is stable under $\mathcal{O}$. The coarse lattice is the same construction with q/p. In the Eisenstein and Gaussian rings these ideals are principal. Their generators have squared moduli $1/q$ and p/q by covolumes relative to $\mathcal{O}$. Their inverse generators are integral and prime to 2 or 3 respectively, as the inverse ideal norms are powers of p.

The real positive tip is centered at zero. For the other two $d = 3$ tips the offsets are $\alpha = (1-\omega_3)/3$ and 2α. The quotient $\alpha\mathcal{O}/\mathcal{O}$ has three elements, and the inverse generators of the fine and coarse ideals induce permutations of its two nonzero classes. The two classes have equal radial counts by negation. For $d = 4$ the negative tip is centered at $(1 - i)/2$ and the corresponding quotient has just one nonzero class, fixed by each inverse generator. Thus the fine and coarse lattices centered at a tip become the indicated translates after rescaling and rotating. The first squared norm is $1/3$ in the nontrivial Eisenstein translate and $1/2$ in the nontrivial Gaussian translate. This follows by multiplying back by the indicated α, which has that squared norm, and observing that the nonzero congruence class contains a unit.

Every contributing lattice point is in the unique tip chart, and the error in the scaled radius from (35) is $O_d(q^{-2/k})$ for bounded λ. A point on an exact quadratic shell belongs to the strict trace disk by the negative quartic term. All other points have a gap from the fixed radius in the finite set of bounded quadratic shells. For varying p the coarse nonzero shells exceed the fixed radius except at a finite set of primes. This proves the eventual count including equality.

C.3.2 Microscopic counts as depth varies

The lattice also describes critical-scale depth statistics outside degrees one and two. Fix p, d with $k = \varphi(d) > 2$. Write $K = \mathbb{Q}(\zeta_d)$, $\mathcal{O}_K = \mathbb{Z}[\zeta_d]$, let $\mathfrak{p}$ be the prime selected by the embedding into

$\mathbb{Q}_p$, and let $\mathfrak{D}$ be the different. The real trace pairing identifies the fine and deleted lattices with

$$\mathfrak{p}^{-n}\mathfrak{D}^{-1} \quad \text{and} \quad \mathfrak{p}^{1-n}\mathfrak{D}^{-1}$$

in the Minkowski space of K. In fact the dual kernel in the power basis is $\mathfrak{p}^n$, and the trace dual of an ideal $\mathfrak{a}$ is $\mathfrak{a}^{-1}\mathfrak{D}^{-1}$. Under this identification an element γ has angles $\theta_j = \mathrm{Tr}(\gamma\zeta_d^j)$ for the first k basis coordinates. The quadratic form near the positive tip is

$$Q(\theta) = 2\pi^2 d \sum_{\sigma} |\sigma(\gamma)|^2$$

by orthogonality of the root powers, pairing conjugate complex embeddings. The sum is over all embeddings.

Choose $h > 0$ such that $\mathfrak{p}^h = (\varpi)$ and write $n = hm + t$, $0 \le t < h$. There is one log coordinate for each complex conjugate pair. In the hyperplane of coordinate sum zero let

$$x_i = \frac{h}{k}\log p - \log|\sigma_i(\varpi)|, \qquad \mathbb{T}_K = \left\{ y \in \mathbb{R}^{k/2} : \sum_i y_i = 0 \right\} / \mathcal{U},$$

where $\mathcal{U}$ is the lattice of $\log|\sigma_i(u)|$ for units u. This is a compact real torus by Dirichlet's unit theorem. Let $\mathbb{T}_x$ be the closure of the cyclic orbit of x in $\mathbb{T}_K$. For fixed $\lambda > 0$ define $M_t(y)$ to count elements $\gamma \in \mathfrak{p}^{-t}\mathfrak{D}^{-1} \setminus \mathfrak{p}^{1-t}\mathfrak{D}^{-1}$ with

$$2\pi^2 d\, p^{2t/k} \sum_{\sigma} e^{2y_{[\sigma]}} |\sigma(\gamma)|^2 \le \lambda, \tag{37}$$

where $[\sigma]$ denotes the conjugate pair. This count depends only on the log-torus class because the units preserve both ideals. It is uniformly bounded on a compact fundamental domain.

Theorem C.3. *As n ranges over depths up to N and $N \to \infty$, the empirical distribution of $N_{p^n,1}(\lambda)$ converges to the equal mixture over $0 \le t < h$ of the pushforwards of Haar on $\mathbb{T}_x$ by M_t. This is for the period model. For the arithmetic application the same description uses the depths on which the common Gaussian factor applies, with the tip rotated by that factor.*

To prove it, write an element of either depth-n ideal as ϖ^{-m} times an element of its fixed depth-t ideal, and normalize angles by $q^{1/k}$. The resulting radial energies are (37) at $y = mx$. Rotation by the phases of the archimedean factors does not affect their quadratic norms. The depth points become equidistributed on $\mathbb{T}_x$ by the character criterion for a cyclic orbit in a compact torus. The orbit is not torsion when $k > 2$. Otherwise a positive power of ϖ times a unit would have all conjugate absolute values equal. Its product with its complex conjugate would then be rational. Its principal ideal is supported only on $\mathfrak{p}$ and its conjugate, whereas a rational ideal divisible by p is supported at all k primes over p. This is impossible.

On every positive-dimensional component of $\mathbb{T}_x$, the equality in (37) has measure zero. In local linear lifts there are only finitely many candidate elements. For each nonzero element its energy is strictly convex in any nonzero linear direction, since all its embeddings are nonzero. Its analytic level set therefore has measure zero on the component. The scaled Taylor error in replacing the trace radius by the quadratic energy is $O(q^{-2/k})$ in a bounded microscopic disk. The count can differ only at points approaching the finitely many shell hypersurfaces in local patches, which have vanishing Haar mass in shrinking neighborhoods. Equidistribution and then a neighborhood limit prove the claimed empirical law. If only even depths are used one can take h even and retain only the even residue classes t, with equal weights. The argument for the orbit remains unchanged.

C.3.3 An explicit critical-depth law at order five

In order $d = 5$, put $\varphi = (1 + \sqrt{5})/2$, $L = \log \varphi$, and

$$C = \frac{8\pi^2}{\sqrt{5}}.$$

For fixed $p \equiv 1 \pmod 5$ define the nearest scaled distance to the positive tip by

$$A_n = q^{1/2} \min_{s \in (\mathbb{Z}/q\mathbb{Z})^\times} |T_5(\theta_q(s)) - 5|, \qquad q = p^n.$$

As depth varies, its empirical distribution has distribution function

$$\begin{cases} 0, & v < C, \\ L^{-1} \operatorname{arcosh}(v/C), & C \le v \le 4\pi^2, \\ 1, & v > 4\pi^2. \end{cases} \tag{38}$$

In particular these are its lower and upper limiting radii. The same holds for the minimum over the full rank-five Kloosterman family, since the common Gaussian factor at odd depth is $\chi(5) = 1$ and zero is far from this tip.

Here is the computation. The field $\mathbb{Q}(\zeta_5)$ has discriminant 125, and its class number is one by the Minkowski bound $(4/\pi)^2(4!/4^4)\sqrt{125} < 2$. Its different is principal with generator $\Phi_5'(\zeta_5)$. Put $\gamma_0 = 1/\Phi_5'(\zeta_5)$, so $|\operatorname{Norm}(\gamma_0)| = 1/125$. In the real quadratic subfield the positive units are powers of φ. One way to see minimality is that a unit v with $1 < v < \varphi$ has integral trace $v \pm v^{-1}$, strictly between 2 and $\sqrt{5}$ in the plus case and between 0 and 1 in the minus case. No such trace is integral. Its totally positive units are thus even powers of φ. For a unit u of the cyclotomic field, $u\overline{u}$ is a totally positive unit in this subfield. It follows that the log unit period in either complex embedding is L, since φ itself realizes this period.

Here $h = 1$ is permitted in Theorem C.3. The non-torsion orbit is dense in the one-dimensional log torus. For each torus point, the smallest quadratic energy among unit multiples of γ_0 is

$$C \cosh\left(2 \operatorname{dist}(y - y_0, L\mathbb{Z})\right)$$

for a fixed offset y_0. Indeed the two positive embedding weights after balancing have product $|\operatorname{Norm}(\gamma_0)|$, and the two conjugate pairs give this sum of opposite exponentials. The balanced value is $2\pi^2 \cdot 5 \cdot 4\,|\operatorname{Norm}(\gamma_0)|^{1/2} = C$. Any nonunit multiplier has absolute algebraic norm at least two, so its scaled energy is at least $\sqrt{2}\,C$ by the arithmetic-geometric mean inequality. The unit envelope never exceeds $C \cosh L = 4\pi^2 < \sqrt{2}\,C$. Hence a unit always minimizes, and its point survives the deletion of the coarse ideal. The scaled trace error is $o(1)$ uniformly for these bounded shortest radii by (35). Finally the distance to $L\mathbb{Z}$ for uniform y mod L is uniform on $[0, L/2]$. This gives (38).

C.4 Moving along the boundary at prime order

Let $d \ge 3$ be prime, $k = d - 1$, and $z_0 = ku + v$ with $u \in \mathbb{T}$ and $v = u^{-k}$. Put $\Delta = |u - v|$. The phase relation is exactly the determinant-one relation and the trace image is the filled hypocycloid [25, Proposition 5.3]. For the degree-two case $d = 3$, Evans and Pugh [45, Section 2.6] give the explicit unweighted torus trace density. The following estimates concern unweighted torus Haar measure in every prime order, and its quantitative transfer to the unit lattice.

Theorem C.4. *Uniformly along that boundary and for small ϵ the Haar disk mass $V_\Delta(\epsilon) = m_d\{|T_d - z_0| < \epsilon\}$ has order*

$$\begin{cases} (\epsilon/\Delta)^{(k+1)/2}, & 0 < \epsilon \lesssim \Delta^3, \\ \epsilon\Delta^{k-2}, & \Delta^3 \lesssim \epsilon \lesssim \Delta^2, \\ \epsilon^{k/2}, & \epsilon \gtrsim \Delta^2. \end{cases}$$

The expressions agree in order on the transition regions, and the last applies at $\Delta = 0$. Set $\delta = q^{-1/k}$. If $\epsilon \geq C_d\delta(\Delta + \sqrt{\epsilon})$, the unit-lattice probability has the same order. If $\epsilon \to 0$ with $\delta(\Delta + \sqrt{\epsilon}) = o(\epsilon)$, its ratio to $V_\Delta(\epsilon)$ tends to one, uniformly over such moving targets. The same conclusions hold on an affine residue grid with $\delta = (q/p)^{-1/k}$.

Away from the cusps the noncusp boundary fibers have a single repeated-eigenvalue class up to permutation. To verify this, the trace differential on the determinant-one torus is surjective at points with at least three distinct phases, since differences of three distinct unit circle points span the real plane. If two distinct phases occur with both multiplicities at least two, split a pair in each cluster by opposite small angles and vary their center arguments with reciprocal multiplicities to preserve determinant. The normal projections of the two pair splits are nonzero with opposite signs, while the center-angle variation is tangent. The implicit function theorem then produces a point of the same trace with both pair splits positive, hence a regular point. Thus only the one-versus-rest boundary fiber is possible. The boundary parametrization $\gamma(t) = (d-1)e^{it} + e^{-i(d-1)t}$ winds once about zero. Its argument derivative is $(d-1)(d-2)(1-\cos(dt))/|\gamma(t)|^2$, positive except at the cusps. Thus the parametrization has no repetitions apart from its period.

At such a fiber write its k repeated phases as $ue^{i(s/k+w_j)}$, $\sum_j w_j = 0$, and the singleton as ve^{-is}. The local trace increment is

$$i(u-v)s - \tfrac{1}{2}u\sum_j w_j^2 + O_d(s^2 + |s|\,\|w\|^2 + \|w\|^3).$$

For Δ bounded below the two displayed directions are independent over the reals. The tangential and normal projections force $|s| \ll_d \epsilon$ and $\|w\| \ll_d \sqrt{\epsilon}$, with reverse inclusion for sufficiently small constants. This gives the appropriate volume $\epsilon^{(k+1)/2}$ and a fixed-target leading coefficient by rescaling.

Near a cusp rotate it to the positive tip and use mean-zero angles $t = r\theta$ with $\sum\theta_j = 0$ and $\sum\theta_j^2 = 1$. Set $A = d - \Re T_d$. Then

$$A = r^2/2 + O_d(r^4), \qquad \partial_r A = r + O_d(r^3).$$

For small A the function $H_A(\theta) = \Im T_d/A^{3/2}$ on the angle sphere is smooth in A and converges in C^2 to a nonzero multiple of $\sum_j \theta_j^3$. The extrema of that cubic have the one-versus-rest form, up to permutations. The multiplier equations for the mean and norm constraints give at most two coordinate values. With multiplicities $j, d-j$ the absolute cubic value is $|d-2j|/\sqrt{dj(d-j)}$, largest at $j = 1, d-1$. The Hessian on the tangent to the sphere at an extremum is a definite nonzero scalar times the Euclidean form on the repeated cluster, by the multiplier equation. The same extremal directions persist for H_A: permutation symmetry fixes them as critical points and the C^2 estimates and compact gap away from the critical neighborhoods preserve their unique local maximality or minimality. Use conjugation for the lower branch. The gap from the maximum is comparable to squared angular distance near the finite extremal set.

For a target on that branch let $A_0 = d - \Re z_0 \asymp \Delta^2$. The upper imaginary boundary is $A^{3/2}\max H_A$ and has derivative $O_d(\sqrt{A})$. If $\epsilon \lesssim \Delta^3$, a point in the disk must have $|A - A_0| \lesssim \epsilon$

and angular distance at most $C_d(\epsilon/\Delta^3)^{1/2}$ from the extremal set. Conversely the same bounds with small constants put it in the disk. This follows by comparing its imaginary part with the boundary at A and using the derivative bound. The unweighted radial volume is of order $\Delta^{k-2}\,dA$. An angular cap of the indicated radius has volume of its radius to the power $k-1$, giving the first formula. For $\Delta^3 \lesssim \epsilon \lesssim \Delta^2$ use a fixed small angular cap for the lower estimate and all directions for the upper. The radial factor is unchanged. If ϵ has order Δ^2 the shell estimate reaches the $\epsilon^{k/2}$ order, and if ϵ is larger by a sufficiently large fixed factor the disk contains a smaller cusp ball and is contained in a larger one. This proves the three orders.

We record the relative continuity in radius needed for the moving asymptotic. If $\eta \to 0$ and $\epsilon \to 0$, the disk collar between radii $(1-\eta)\epsilon$ and $(1+\eta)\epsilon$ has Haar measure $o(V_\Delta(\epsilon))$ uniformly. Here is a rescaling proof. Along any sequence with Δ bounded away from zero, use $(s/\epsilon, w/\sqrt{\epsilon})$ in the preceding normal form and take a subsequence of the target. The limiting disk boundary has zero Lebesgue measure, since each fixed normal slice has at most two tangential positions on the boundary. For $\Delta = 0$ use the last rescaling below directly. If $\Delta > 0$ tends to zero, pass to subsequences for the two ratios ϵ/Δ^3 and ϵ/Δ^2. When the first tends to zero, use radial coordinate $(A - A_0)/\epsilon$ and angular coordinates scaled by $(\epsilon/\Delta^3)^{1/2}$ about an extremal direction. The limiting squared disk distance is a radial coordinate squared plus the square of a positive definite angular quadratic form, so its level one is null. If the first ratio has finite positive limit, the whole angular sphere is kept with the radial coordinate $(A - A_0)/\epsilon$. The limiting normal displacement is a fixed multiple of $H_0(\theta) - \max H_0$, with null boundary by radial slices. If the first ratio tends to infinity and the second to zero, the angular displacement vanishes in the normalized disk test, which becomes a radial slab. If the second ratio has positive limit or tends to infinity, divide ordinary angles by $\sqrt{\epsilon}$. The limiting disk is described by $|Q_0(y) - a| \le 1$, for the positive quadratic form Q_0 in those angles for a finite $a = \lim A_0/\epsilon \ge 0$. Its boundary is again null. In each rescaling the normalized Jacobians stay bounded and converge on bounded sets (or on the full compact sphere), and all disk contributions lie in those sets by the volume proof. Dominated convergence gives the negligible collar fraction. Since this holds on a subsequence of every possible sequence, the claim is uniform.

Under the coupling of Theorem C.1, the angular displacement is $O_d(\delta)$. Near a cusp every preimage of the relevant disk is within $O_d(\Delta + \sqrt{\epsilon})$ of the cusp angle by its real deficit. The trace differential vanishes at the cusp. The change of trace over the angular displacement is therefore

$$O_d\big(\delta(\Delta + \sqrt{\epsilon} + \delta)\big).$$

Away from the cusps the ordinary Lipschitz estimate gives the same bound since Δ is bounded below. The transfer condition absorbs the δ^2 term and sandwiches the arithmetic disk between Haar disks of fixed smaller and larger radii. This proves comparability. With the strict separation the widening is $o(\epsilon)$ and the collar estimate gives the relative asymptotic. The affine residue-grid result is identical.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including the lattice coupling and the shell calculations, and with drafting and checking the proof. No numerical data are needed for the results.

D Effective quartic cyclotomic trace laws

For each quartic cyclotomic order $d = 5, 8, 10, 12$, we study trace values at distance $O(q^{-1/2})$ from the aligned tip d over split prime powers $q = p^n$. The angular lattice identifies each

bounded list of scaled trace distances with energies in one real unit coordinate. The nearest trace has an explicit arcosh distribution at fixed prime as depth varies, and the full bounded-radius counts have effective empirical laws. A second averaging theorem gives the corresponding laws over split primes, uniformly for common depths as large as $\exp(c\log X/\log\log X)$ on a prime window $[X, 2X]$. The two averages use respectively a logarithmic Diophantine estimate and uniform prime-ideal sums of infinite-order Hecke characters.

Keywords: cyclotomic periods, trace disks, Hecke characters, linear forms in logarithms.
MSC 2020: 11L05, 11R18, 11R42, 11J86.

D.1 Trace laws and a unit coordinate

Fix $d \in \{5, 8, 10, 12\}$ and a prime $p \equiv 1 \pmod d$. Put $q = p^n$ and let ζ_q be a primitive dth root modulo q. For unit $s \bmod q$ define

$$T_{d,q}(s) = \sum_{j=0}^{d-1} \exp(2\pi i s\zeta_q^j/q), \qquad Z_{p,n}(s) = \sqrt{q}\,|T_{d,q}(s) - d|.$$

Let $M_{p,n}$ be the least of these scaled distances. For each fixed $v > 0$ let $B_{p,n}(v)$ be the number of units with $Z_{p,n}(s) < v$. These laws are independent of the choice of ζ_q, since a primitive root change permutes its powers. Multiplicities in $B_{p,n}$ are those of root parameters s.

Write $K = \mathbb{Q}(\zeta_d)$, $D = |\operatorname{disc}(K)|$, and let σ_1, σ_2 represent its two complex embedding pairs. With $\varphi = (1+\sqrt{5})/2$, define

d	D	L	$C_d = 8\pi^2 d/\sqrt{D}$	$C_d \cosh L$
5	125	$\log\varphi$	$8\pi^2/\sqrt{5}$	$4\pi^2$
8	256	$\log(1+\sqrt{2})$	$4\pi^2$	$4\sqrt{2}\,\pi^2$
10	125	$\log\varphi$	$16\pi^2/\sqrt{5}$	$8\pi^2$
12	144	$\frac{1}{2}\log(2+\sqrt{3})$	$8\pi^2$	$8\sqrt{3/2}\,\pi^2$

The positive generator L of the absolute unit logarithms is proved below. Fix the degree-one prime $\mathfrak{p}$ over p specified by evaluation at ζ_q. Write $\mathfrak{p} = (\pi)$ and the different $\mathfrak{D} = (\delta)$, using principality. Put

$$\tau_{p,n} = \tau_0 + nx_p \pmod L, \qquad \tau_0 = \tfrac{1}{4}\log D - \log|\sigma_1\delta|, \qquad x_p = \tfrac{1}{4}\log p - \log|\sigma_1\pi|. \tag{39}$$

Both classes modulo L are independent of the generators. For $\gamma \in \mathcal{O} \setminus \{0\}$ set

$$E_\tau(\gamma) = C_d\sqrt{|N_{K/\mathbb{Q}}\gamma|}\cosh\left(2\tau + \log\frac{|\sigma_1\gamma|}{|\sigma_2\gamma|}\right). \tag{40}$$

For a fixed radius v let $B_\tau^{(\mathfrak{p})}(v) = \#\{\gamma \in \mathcal{O} \setminus \mathfrak{p} : E_\tau(\gamma) < v\}$ and let $B_\tau^*(v)$ use $\gamma \in \mathcal{O} \setminus \{0\}$. These functions on the unit circle of length L are well defined because shifting τ by L reindexes γ by multiplication by a unit.

Theorem D.1. *Uniformly in the split prime and depth,*

$$M_{p,n} = C_d\cosh(2\operatorname{dist}(\tau_{p,n}, L\mathbb{Z})) + O_d(q^{-1/2}). \tag{41}$$

Over $n = 1, \dots, N$ the empirical distribution function of $M_{p,n}$ is within $O_d(N^{-c_d/\log p})$ of

$$F_d(v) = \begin{cases} 0, & v < C_d, \\ L^{-1}\operatorname{arcosh}(v/C_d), & C_d \le v \le C_d\cosh L, \\ 1, & v > C_d\cosh L. \end{cases} \tag{42}$$

Here $c_d > 0$ is effective. The same estimate holds when sampling only even depths $2, 4, \dots, 2N$. More generally for any fixed finite list of radii $v_1, \dots, v_j$, the empirical law of $(B_{p,n}(v_i))_{i\le j}$ is within $O_{d,\mathbf{v}}(N^{-c_d/\log p})$ in total variation of the law of $(B_U^{(\mathfrak{p})}(v_i))_{i\le j}$ for U Haar on $\mathbb{R}/L\mathbb{Z}$.

Theorem D.2. *Fix $A > 0$ and let p be uniform over $p \equiv 1 \pmod d$ in $[X, 2X]$. Uniformly for common depths*

$$1 \le n \le \exp(c_{d,A} \log X/\log\log X),$$

$M_{p,n}$ has distribution function $F_d + O_{d,A}((\log X)^{-A})$, uniformly in the argument. For any fixed finite list of radii, the joint count law is within $O_{d,A,\mathbf{v}}((\log X)^{-A})$ of the law of $(B_U^*(v_i))_{i\le j}$. Both statements also hold on any fixed reduced prime progression compatible with $p \equiv 1 \pmod d$, with constants allowed to depend on its modulus.

For even depths $n \ge 2$ these statements apply to tame hyper-Kloosterman endpoint traces when $\gcd(R, p-1) = d$. The normalized rank-R completed sum at a unit parameter c is zero if c is inactive, and on the active gate it is the sum of stationary phases $\exp(2\pi i Rs/q)$ over the Rth roots s of c. The standard tame stationary-phase formula gives these weights at even depth [23, 50]; multiplication by R permutes the unit roots when $p \nmid R$. Each active parameter has d roots, so divide the microscopic root count by d for the number of active parameters. For odd depths $n \ge 3$ the same interpretation applies when one common Gauss factor rotates the entire active trace law, for example at odd rank. For fixed R the condition $\gcd(R, p-1) = d$ selects a fixed union of prime progressions, so the final assertion of Theorem D.2 applies when this union is nonempty.

The qualitative trace images of Gaussian periods were studied by Duke, Garcia and Lutz [47] and Untrau [25]. Kowalski and Untrau [15, Proposition 3.4] give the cyclotomic norm method for finite-field Wasserstein estimates. The trace-dual lattice comparison below keeps the ideal at the critical scale, where a fixed torus weak limit alone does not retain the counts. The quantitative arithmetic inputs are Matveev's logarithm bound [48] and the infinite-order Hecke prime-ideal estimate in Ghidelli [49, Lemma 5.4], the latter used in a range with growing archimedean powers.

D.2 The critical ideal energy

The degree-four fields in the table have class number one. By the Minkowski bound, each ideal class has an integral representative of norm at most $(4/\pi)^2 \cdot 4!/4^4 \cdot \sqrt{D}$. This is less than two for $D = 125, 144$ and less than three for $D = 256$. In the latter field there is a unique prime of norm two, and $1 - \zeta_8$ has norm two, since $\Phi_8(1) = 2$ and the polynomial reduces to $(X-1)^4$ modulo two. This proves principality in the remaining case.

For a unit u of the CM field K, the real-subfield unit $u\bar{u}$ is totally positive and has norm one. Its possible values form the positive powers of the smallest totally positive unit greater than one, and their inverses. In $\mathbb{Q}(\sqrt{5})$ this generator is φ^2 by the integral trace bound: a totally positive norm-one unit greater than one has trace at least three, achieved by φ^2. In $\mathbb{Q}(\sqrt{2})$ it is $3+2\sqrt{2}$: its trace is $2a$ for $a \in \mathbb{Z}$, and $a = 2$ is excluded by $a^2 - 2b^2 = 1$, while $a = 3, b = 2$ is possible. In $\mathbb{Q}(\sqrt{3})$ the smallest is $2 + \sqrt{3}$ by the same trace argument with $a = 2, b = 1$. The logarithms $\log|\sigma_1 u|$ are therefore in $L\mathbb{Z}$ with L as in the table. Equality of the lattices is realized by the real units φ and $1 + \sqrt{2}$ for $d = 5, 10$ and 8. For $d = 12$, $1 - \zeta_{12}$ is a unit because $\Phi_{12}(1) = 1$, and its squared absolute value at one place is $2 - \sqrt{3} = (2 + \sqrt{3})^{-1}$, yielding the asserted half-log period. Since the different is invariant under Galois and exchanging the complex pairs sends τ_0 to $-\tau_0$, one has $2\tau_0 \in L\mathbb{Z}$. With the cyclotomic different generator $\Phi_d'(\zeta_d)$ the class of τ_0 is $L/2$ for $d = 5, 10$ (the absolute values in the two pairs have ratio φ or its inverse), and zero for $d = 8, 12$ (their absolute values agree).

Use power-basis angular coordinates $\theta \in \mathbb{R}^4/\mathbb{Z}^4$ corresponding to the characters $1, \zeta_d, \zeta_d^2, \zeta_d^3$. The fine grid of all residues s has as dual in the integral power basis the kernel of evaluation modulo q, namely $\mathfrak{p}^n$. Its dual under the trace pairing is $\mathfrak{p}^{-n}\mathfrak{D}^{-1}$. Indeed a trace-dual ideal to $\mathfrak{a}$ is $\mathfrak{a}^{-1}\mathfrak{D}^{-1}$, and the real angular coordinates of an element ξ of that ideal are $(\operatorname{Tr}(\xi\zeta_d^j))_{j=0}^3$. For nonunits s the coarse lattice is $\mathfrak{p}^{1-n}\mathfrak{D}^{-1}$, so the projected unit grid is the difference of the fine and coarse lattices. In the local chart around the angle zero one can therefore write

$$\xi = \pi^{-n}\delta^{-1}\gamma, \qquad \gamma \in \mathcal{O} \setminus \mathfrak{p}.$$

Changing the representative of $\tau_{p,n}$ by a unit corresponds to reindexing γ and leaves this condition intact.

Write $t_j = 2\pi \operatorname{Tr}(\xi\zeta_d^j)$. The sum $\sum_{j=0}^{d-1} t_j$ is zero. At the aligned tip, the real deficit has quadratic term $\sum_j t_j^2/2$ and the imaginary part starts in degree three. Expanding its modulus gives

$$|T - d| = \tfrac{1}{2}\sum_j t_j^2 + O_d(\|\theta\|^4)$$

in the local chart, since the displayed quadratic is positive definite. The phase sum has a unique fiber above d: equality in the triangle inequality at that value forces all its constituent phases to be one, hence also its first four coordinates. More directly, $d - \Re T$ controls each of the first four deficits $1 - \cos(2\pi\theta_j)$. Thus scaled distances bounded by a fixed v have $\|\theta\| \ll_{d,v} q^{-1/4}$. Conversely a bounded scaled quadratic energy also lies in that chart for sufficiently large q.

Orthogonality of the root powers in the sum over j gives $\sum_j (\operatorname{Tr}(\xi\zeta_d^j))^2 = d\sum_\sigma |\sigma\xi|^2$. Multiplying the quadratic energy by $\sqrt{q}$ and pairing the complex embeddings gives

$$4\pi^2 d\sqrt{q}\left(|\sigma_1(\pi^{-n}\delta^{-1}\gamma)|^2 + |\sigma_2(\pi^{-n}\delta^{-1}\gamma)|^2\right) = E_{\tau_{p,n}}(\gamma),$$

because $|\sigma_1(\pi^{-n}\delta^{-1})||\sigma_2(\pi^{-n}\delta^{-1})| = (qD)^{-1/2}$. With a bounded τ representative, bounded $E_\tau(\gamma)$ puts both embedding magnitudes of γ in a fixed compact box. The image of $\mathcal{O}$ is a lattice there, so only finitely many candidates occur for any fixed bound, uniformly in p, n. In these bounded lists, the normalized trace distances and the ideal energies differ by $O_{d,v}(q^{-1/2})$. This is a two-sided comparison, since bounded ideal energies give distinct local angle representatives and conversely all bounded trace distances have such representatives.

For units γ the energies in (40) range over $C_d \cosh(2(\tau + jL))$ with $j \in \mathbb{Z}$. The least unit energy is the right side of (41) without the error. Every unit survives deletion by $\mathfrak{p}$. A nonunit has absolute norm at least two, so its energy is at least $\sqrt{2}C_d$, whereas each table entry $C_d \cosh L$ is at most this value. This includes the tie of the two bounds at $d = 8$. Thus the least ideal energy is always given by a unit. Comparing the actual trace in a bounded energy interval proves (41). Under Haar in the unit coordinate, $\operatorname{dist}(U, L\mathbb{Z})$ is uniform on $[0, L/2]$, giving (42).

D.3 Effective depth averaging

The rotation x_p/L is irrational. If $jx_p = mL$ for integers $j \neq 0, m$, choose a unit with the corresponding logarithm and multiply $\pi^{|j|}$ by its appropriate power. This produces ω generating $\mathfrak{p}^{|j|}$ with equal absolute values at both complex pairs. Then $\omega\bar{\omega}$ lies in the real subfield with equal conjugates, hence is rational. Its principal ideal is supported on the two distinct primes $\mathfrak{p}, \bar{\mathfrak{p}}$, which is impossible for a rational number because p is completely split into four primes, all with identical valuations on rational numbers.

Choose the generator π balanced by a unit so that $|x_p| \leq L/2$. This makes its four archimedean magnitudes comparable to $p^{1/4}$ uniformly, and its absolute logarithmic height is $O_d(\log p)$. The

positive real algebraic numbers $a = p^{1/4}/|\sigma_1\pi|$ and $b = e^L$ have bounded joint degree and heights $O_d(\log p)$ and $O_d(1)$, respectively. For $j \geq 1$ choose the nearest integer m to jx_p/L, so $|m| \leq j+1$. Matveev's bound for a nonzero form in logarithms [48] gives

$$\|jx_p/L\|_{\mathbb{T}} \gg_d p^{-C'_d} j^{-C'_d \log p}.$$

Indeed the form $j \log a - m \log b$ is nonzero by irrationality, and its logarithmic lower bound is $-O_d((\log p)(1 + \log j))$. If $m = 0$, the corresponding one-logarithm estimate is also immediate by the elementary height bound. These arguments remain valid for $2x_p/L$.

For $1 \leq j \leq H$, geometric summation bounds the mean exponential of njx_p/L over N terms by $\min(1, 1/(N\|jx_p/L\|_{\mathbb{T}}))$. The Erdős–Turán inequality therefore bounds the discrepancy on intervals by $O(1/H + p^{C'_d} H^{1+C'_d \log p}/N)$. Taking H a small power $N^{c_d/\log p}$ gives $O_d(N^{-c_d/\log p})$ for $N \geq p^{C''_d}$ with a fixed sufficiently large C''_d. For smaller N the trivial bound gives the same estimate with an adjusted uniform factor. Thus the depth orbit is quantitatively uniform also on its even subsequence.

To pass to the trace values, discard at most $O_{d,\mathbf{v}}(1 + \log N/\log p)$ initial depths to make the scaled Taylor error $O(N^{-2})$. For the minimum in (41), a collar of width $O(N^{-2})$ has inverse image under the hyperbolic cosine expression of Haar length $O_d(N^{-1})$ uniformly in the target. For fixed count radii, the candidate set in the balanced chart is bounded. Each function $E_\tau(\gamma)$ has positive second derivative and its collar of width η has Haar length $O_{d,\mathbf{v}}(\sqrt{\eta})$; all its sublevel sets and collars are unions of a bounded number of intervals there. The joint counts are piecewise constant on the finitely many resulting intervals. Combining these comparisons with the uniform interval discrepancy, and the negligible fraction of discarded depths, proves Theorem D.1.

D.4 Prime averaging with a growing common depth

For any integral ideal $\mathfrak{a} = (z)$ of K define

$$\chi_m(\mathfrak{a}) = \exp\left(\frac{2\pi i m}{L}\left(\tfrac{1}{4}\log N\mathfrak{a} - \log|\sigma_1 z|\right)\right), \qquad m \in \mathbb{Z}.$$

Changing z by a unit does not alter this expression. It is a Hecke character of conductor one, since on principal ideals its continuous archimedean type is $|z_1|^{-i\pi m/L}|z_2|^{i\pi m/L}$. At $m \neq 0$ the opposite nonzero exponents show that it is of infinite order and has a non-norm, non-real infinity type, so the real exceptional character does not occur. The logarithm of its archimedean size in Ghidelli's character lattice is $O_d(\log(2 + |m|))$. The unweighted prime-ideal bound of [49, Lemma 5.4] gives

$$\left|\sum_{N\mathfrak{p} \leq Y} \chi_m(\mathfrak{p})\right| \ll_d Y \exp\left(-\frac{c'_d \log Y}{\log(2 + |m|) + \sqrt{\log Y}}\right).$$

Primes of degree at least two contribute $O_d(\sqrt{Y})$ and the ramified primes are finite. The degree-one primes in this Galois field lie over completely split primes, four per rational prime. By the prime ideal theorem (or the ordinary prime number theorem in the class 1 modulo d), there are $\asymp_d X/\log X$ split rational primes in $[X, 2X]$.

Apply the Erdős–Turán inequality to the multiset of angles $(\tau_0 + nx_{\mathfrak{p}})/L$ over the degree-one ideals in this window, where $x_{\mathfrak{p}} = (\log N\mathfrak{p})/4 - \log|\sigma_1\pi|$. Its Fourier mode of order m is the mean of χ_{mn} times a constant unit phase. Use $H = \lceil (\log X)^{A+1} \rceil$. If $n \leq \exp(c_{d,A} \log X/\log\log X)$ with sufficiently small $c_{d,A} > 0$, the prime-ideal estimate at all $1 \leq m \leq H$ gives a saving of at least $(\log X)^{-A-4}$ before normalization. After normalization and the weighted sum over m, the interval

discrepancy is $O_{d,A}((\log X)^{-A})$. This proves uniformity for a common depth throughout the prime window.

The count for each rational prime is repeated exactly four times among its ideals, because the primitive root choice just permutes the original phase sum. For the ideal comparison, when p is large in terms of the fixed radii, the deleted nonzero multiples of π have $|N\gamma| \geq p$ and energy at least $C_d\sqrt{p}$, above every threshold. Therefore the limiting count model uses B_U^*, independent of the prime. The Taylor collars have total Haar length $O_{d,\mathbf{v}}(X^{-n/4})$, and their prime-averaged frequency carries in addition the interval discrepancy. The same argument works uniformly at every threshold for the minimum, whose inverse collar lengths are bounded uniformly. This proves Theorem D.2.

To impose a fixed compatible reduced progression, use orthogonality of Dirichlet characters at its fixed modulus on the rational prime norms. The resulting Hecke characters are $\chi_{mn}(\mathfrak{a})\psi(N\mathfrak{a})$ with a fixed finite list of Dirichlet characters ψ. Their conductor divides a fixed ideal. The non-real infinite type is unchanged, and the same prime sum estimate and the prime number theorem in the combined progression apply. This proves the claimed restricted version.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including the ideal and discrepancy calculations, and with drafting and checking the proof. No data are needed for the results.

E Smooth boundary trace disks at fixed prime

Fix a prime order $d \geq 3$ and a split prime p. For the order-d period traces over $q = p^n$, disks about a smooth hypocycloid boundary point have the expected Haar asymptotic above radius $q^{-2/d+\eta}$ for every $\eta > 0$. At a boundary point with algebraic angle coordinates, disks below $q^{-2/d-\eta}$ are eventually empty. We also determine the Hausdorff dimension of boundary points admitting exceptionally close traces infinitely often. The proof controls the cyclotomic lattice in rationally oriented boxes of unequal side lengths using the subspace theorem at the real and p-adic places, then applies mass transference to boundary intervals.

Keywords: cyclotomic periods, shrinking disks, subspace theorem, Hausdorff dimension.
MSC 2020: 11L05, 11J87, 11K60, 28A78.

E.1 Boundary counts and exceptional targets

Fix a prime $d \geq 3$, put $k = d - 1$, and fix a prime $p \equiv 1 \pmod{d}$. Choose a primitive dth root $\zeta \in \mathbb{Z}_p$ and write $q = p^n$. For a unit $b \bmod q$, set

$$T_q(b) = \sum_{j=0}^{d-1} \exp(2\pi i b\zeta^j/q),$$

using residues modulo q in the exponent. Its angle tuple is $b(1, \zeta, \ldots, \zeta^{k-1})/q$ in $\mathbb{T}^k$. The trace on the whole angle torus is therefore

$$T(\theta) = \sum_{j=1}^{k} e^{2\pi i\theta_j} + e^{-2\pi i(\theta_1+\cdots+\theta_k)}.$$

Write m_k for Haar on $\mathbb{T}^k$. The image is the filled hypocycloid with boundary $z(u) = ku + v$, where $|u| = 1$ and $v = u^{-k}$ [25, Proposition 5.3]. Set $\Delta(u) = |u - v|$. A compact smooth boundary arc

has Δ bounded away from zero. The angle parameter is t in $u = e^{2\pi i t}$. Its being algebraic over $\mathbb{Q}$ is independent of the choice of lift of t.

Theorem E.1. *Let $0 < \eta < 2/d$. Uniformly on every compact smooth arc, as $n \to \infty$ and $r \to 0$ with $r \geq q^{-2/d+\eta}$,*

$$\begin{aligned}\mathbb{P}_b\{|T_q(b) - z(u)| < r\} &= m_k\{|T - z(u)| < r\}\big(1 + O(q^{-\kappa})\big) \\ &= C_d \left(\frac{r}{\Delta(u)}\right)^{d/2} \big(1 + O(\sqrt{r} + q^{-\kappa})\big).\end{aligned}$$

Here $\kappa > 0$ depends on d, η, and the depth threshold may depend on p and the compact arc. The constant is

$$C_d = \frac{d\sqrt{k}\,|S^{k-2}|}{2\pi k\,(\pi^2)^{(k-1)/2}} \int_0^1 t^{(k-1)/2-1}\sqrt{1-t^2}\,dt.$$

The probability may use all units or any nonempty union of nonzero residue classes modulo p. Every point of the full boundary is within $q^{-2/d+\eta}$ of a trace from each such class for all sufficiently large n. For each fixed smooth point with an algebraic angle parameter and any $\eta > 0$, the disk of radius $q^{-2/d-\eta}$ is eventually empty. Its distance to the unit trace set is thus $q^{-2/d+o(1)}$.

Theorem E.2. *For $a > 0$ let $\mathcal{E}_a$ be the set of smooth boundary points $z(u)$ satisfying $|T_q(b) - z(u)| < q^{-a}$ for some unit b at infinitely many depths. In every nonempty open smooth arc its Hausdorff dimension is*

$$\dim_{\mathrm{H}} \mathcal{E}_a = \begin{cases} 1, & a \leq 2/d, \\ \dfrac{1}{a} - \dfrac{d-2}{2}, & 2/d < a < 2/(d-2), \\ 0, & a \geq 2/(d-2), \end{cases}$$

with the convention that the empty set has dimension zero. It is empty on the smooth boundary for $a > 2/(d-2)$. At almost every smooth target the nearest unit trace distance is $q^{-2/d+o(1)}$.

These statements apply to tame hyper-Kloosterman packets of size $d = \gcd(R, p-1)$ when the normalized rank-R stationary trace has a common Gaussian rotation. This includes even $n \geq 2$ and, at odd rank, all $n \geq 2$. The stationary phases at an active parameter give $T_q(Rb)$ up to that rotation, with d roots per active parameter [23, 50]. Counts over unit root parameters divide by d for counts over active parameters. The infinitely-often and almost-everywhere assertions hold also on the even subsequence: the per-depth covering holds eventually throughout that subsequence.

Untrau's Proposition B and Proposition 5.3 give the qualitative period trace law and hypocycloid support. Kowalski and Untrau [15, Proposition 3.4] give the cyclotomic norm argument for a global finite-field Wasserstein bound. At the smooth boundary the angular scales are r in one rational direction and $\sqrt{r}$ in the remaining directions. Their product controls the comparison at fixed prime by Schlickewei's subspace theorem [51]. We pass from lattice coverings to dimension with Beresnevich and Velani [53].

E.2 Rectangles in the cyclotomic angle lattice

Put $V = (1, \zeta, \dots, \zeta^{k-1}) \in \mathbb{Z}_p^k$. Its coordinates are linearly independent over $\mathbb{Q}$ since d is prime. Let $\Lambda_j \subset \mathbb{R}^k$ be the lattice lifting the cyclic angle grid with denominator p^j. Its dual is

$$\Lambda_j^* = \{m \in \mathbb{Z}^k : L(m) = m \cdot V \equiv 0 \pmod{p^j}\}.$$

For b in a prescribed nonzero residue class modulo p, the angles at modulus q form an affine translate of $\Lambda_{n-1}/\mathbb{Z}^k$.

We use the subspace theorem at ∞, p in the following form. Given $l \geq 2$ and l linearly independent algebraic linear forms at each place, integer l-vectors whose product of all local absolute values is $O(\|x\|^{-c})$ for fixed $c > 0$ lie in finitely many proper rational subspaces. The primitive-vector statement implies this integer-vector form. Dividing by the positive coordinate gcd g divides the product by $(g|g|_p)^l \geq 1$ and does not increase the height [51, 52].

Lemma E.3. *Fix an invertible rational coordinate map $B : \mathbb{R}^k \to \mathbb{R}^k$ and $0 < \sigma < 1$. Whenever $0 < \ell_i \leq 1$ and $\prod_i \ell_i \geq q^{-1+\sigma}$, Haar couples to any affine residue grid above with displacement $O(q^{-\kappa}\ell_i)$ in each coordinate i of B, uniformly in the side scales and translate. The exponent $\kappa > 0$ depends only on k, σ, and the depth threshold can depend on p, B, σ. An angular box of these side scales in an injective chart contains its Haar volume times grid cardinality up to relative error $O(q^{-\kappa})$.*

Proof. For $D_\ell = \operatorname{diag}(\ell_i)$, the dual of $D_\ell^{-1} B \Lambda_{n-1}$ is $D_\ell B^{-\mathsf{T}} \Lambda_{n-1}^*$. Suppose a nonzero dual vector had length at most q^κ, where $0 < \kappa < \sigma/(4k)$. For the original $m \in \Lambda_{n-1}^*$ we have $|(B^{-\mathsf{T}} m)_i| \leq q^\kappa/\ell_i$, with product at most $q^{1-\sigma+k\kappa}$ and height $O_B(q^{1-\sigma+\kappa})$. At p the form L contributes at most p/q and can be completed by coordinate forms bounded by one on integer vectors. The product over both places is $O_p(q^{-\sigma+k\kappa})$ and the height is $O(q)$. The subspace theorem applies with a fixed positive exponent and gives finitely many proper rational subspaces.

In a rational subspace of dimension $l \geq 2$, take an integer basis of its integral points. Select l independent restrictions of the archimedean coordinate forms. Their product has the same bound because the discarded coordinate upper bounds are at least one. The restriction of L is nonzero on every nonzero rational vector, so it can be completed at p by coordinate forms in the integral basis. The same inequality holds with new fixed constants and height still $O(q)$, and the theorem descends to finitely many rational lines. On a line $m = tm_0$ with fixed primitive m_0 the congruence forces $p^{n-1-O(1)} \mid t$, since $L(m_0) \neq 0$. This contradicts the real height bound for large n. Thus the first minimum of the scaled dual exceeds q^κ.

Lattice transference bounds the covering radius of the scaled primal by a constant depending on k divided by this first minimum [46]. Voronoi cells in the scaled metric project to equal-volume cells on the original torus, yielding the desired Haar coupling, also for affine translates. Comparing rectangles to coordinatewise enlargements and contractions proves the last assertion. □

Lemma E.4. *Let $v_1, \dots, v_l \in \mathbb{Z}_p$ be algebraic over $\mathbb{Q}$ and linearly independent over $\mathbb{Q}$, and fix $D \in \mathbb{Z} \setminus \{0\}$. For every $\sigma > 0$, for large n there is no $A \in \mathbb{Z}^l$ satisfying*

$$A \equiv Db(v_1, \dots, v_l) \pmod{q}, \qquad b \in \mathbb{Z}_p^\times,$$
$$|A_i| \ll q\ell_i, \qquad 0 < \ell_i \leq 1, \qquad \prod_i \ell_i \leq q^{-1-\sigma}.$$

The threshold is uniform in the side scales with the other data fixed.

Proof. For $l = 1$ the real bound gives $A_1 = 0$, contrary to the congruence for large n. Otherwise use the $l - 1$ independent p-adic forms $A_i - (v_i/v_1)A_1$, $i \geq 2$, and complete by A_1. Their product is $O(q^{-(l-1)})$. The real coordinate product is $O(q^l \prod_i \ell_i)$, so the product over both places is $O(q^{-\sigma})$ with height $O(q)$. The subspace theorem puts the vectors into finitely many proper rational subspaces. None contains the p-adic line spanned by (v_i), since the coordinates are linearly independent over $\mathbb{Q}$. On every exceptional subspace the annihilating forms thus give an injective map with a fixed inverse bound. They force $|A_1|_p = O(q^{-1})$, contrary to $v_p(A_1) = v_p(Dv_1)$ for large n. □

E.3 The boundary chart and the disk scale

Write the phases as $z_1, \dots, z_d$ with product one. At a boundary fiber the differential has rank less than two, so the differences $z_j - z_d$ span at most a real line. Thus there are at most two distinct phases. If both values u, v have multiplicity at least two, changing the cluster means gives a nonzero tangent along $i(u - v)$. A split within either cluster yields a quadratic displacement along $-u$ or $-v$, with opposite projections transverse to the tangent. After solving the tangential coordinate by the implicit function theorem, these splits give a transverse interval around the point. This is an interior image. Every smooth boundary fiber therefore has one singleton and a cluster of k repeated phases. The boundary parameter is unique: equality of $ku + u^{-k}$ and $ku' + (u')^{-k}$ for $u \neq u'$ would make the geometric divided difference of the last powers attain the triangle bound k, forcing $u/u' = 1$. There are exactly d preimages given by the singleton positions.

Place the singleton last. For offsets of the remaining k angles set $\theta_j - \theta_{0,j} = s + w_j$ with $\sum w_j = 0$. The singleton offset is $-ks$. The s coordinate and $k-1$ independent split coordinates can be taken rational-linear in the original angle offsets. In the Euclidean metric on the split hyperplane, Haar has Jacobian $\sqrt{k}\, ds\, dw$. The Taylor expansion is

$$T - z(u) = 2\pi i k(u - v)s - 2\pi^2 u\|w\|^2 + O_d(s^2 + \|w\|^3 + |s|\|w\|^2). \tag{43}$$

When $u \neq v$ the two displayed directions are real-linearly independent. The tangent vector has magnitude $2\pi k\Delta$, and the orthogonal component of the quadratic coefficient has magnitude $\pi^2\Delta$. Thus $|T - z(u)|$ is comparable to $|s| + \|w\|^2$ near the fiber with constants uniform on a compact smooth arc. Compactness and the fiber description ensure that all small disks on such an arc come from these separate charts.

Rescale $s = ra, w = \sqrt{r}\, b$. The normalized Taylor error is $O(\sqrt{r})$ in bounded rescaled variables. The model sublevel volume is homogeneous in its disk radius, of degree $(k+1)/2$. For fixed $\|b\|^2$ the permitted interval of a has length $2(2\pi k\Delta)^{-1}\sqrt{1 - (\pi^2\Delta\|b\|^2)^2}$. Integrating radially in the split space, including the Haar Jacobian and the d fibers, gives the constant C_d and a relative error $O(\sqrt{r})$ for Haar.

We also need the actual Haar collar bound uniformly on the arc. Its relative mass for a relative radial width $0 < \delta \ll 1$ is $O(\delta)$, with no $\sqrt{r}$ term. To see this, dilate s, w to $ts, \sqrt{t}\, w$. For points with original radius comparable to r and t near one, the derivative in t of the trace distance is the magnitude of the leading term of (43) plus $O(r^{3/2})$, by the differentiated Taylor expansion. It is at least a constant times r. Dilations by $1 \pm C\delta$ then sandwich the disks of radii $(1 \pm \delta)r$ by the corresponding dilates of the radius-r domain. Points at distance at most $r/2$ remain in the inner disk by the uniform bound on the differentiated expansion. The dilation Jacobian gives the stated collar estimate.

Apply Lemma E.3 to each singleton chart with scales $r, \sqrt{r}, \dots, \sqrt{r}$, whose product is $r^{d/2} \geq q^{-1+\eta d/2}$. Near a radius-$r$ disk the derivative in the split directions is $O(\sqrt{r})$, so the coupled angular displacement changes the trace by $O(rq^{-\kappa})$. Each chart can use its own coupling because the preimage disks lie in separate chart interiors. The actual Haar collars give the first identity of Theorem E.1, and the model calculation gives the second. At any boundary point, including a cusp, the split derivative at the boundary tuple vanishes. Therefore a box of sufficiently small fixed multiples of $r, \sqrt{r}, \dots, \sqrt{r}$ lies in its disk, uniformly over the boundary, by (43). Lemma E.3 yields an angle point in it from each nonzero residue class when $r = q^{-2/d+\eta}$.

If the fixed target angle parameter is algebraic, every lift θ_0 of its fiber angles has algebraic real coordinates. An angle point lifted near it is M/q, with $M \in \mathbb{Z}^k$ and $M \equiv bV \pmod q$ for a unit b. Apply the subspace theorem to the $(k+1)$ integer coordinates (M, q). At the real place use the k chart coordinates of $M - q\theta_0$ and the last coordinate q. For a hit within radius r their product

is $O(q^{k+1}r^{d/2})$. At p use $M_i - \zeta^{i-1}M_1$ for $2 \le i \le k$, together with q and M_1. Their product is at most q^{-k}. Both systems are linearly independent and algebraic. When $r = q^{-2/d-\eta}$ the combined product is bounded by $O(q^{-\eta d/2})$ with height $O(q)$, so the integer vectors lie in finitely many proper rational subspaces. On such a subspace take a nonzero integer relation $c \cdot M + c_0 q = 0$. If $c = 0$, it is impossible since $q > 0$. Otherwise reduction modulo q forces $p^n \mid c \cdot V$, since b is a unit, contrary to the fixed nonzero algebraic value $c \cdot V$ for large n. This excludes the hit and completes Theorem E.1.

E.4 Dimension of the exceptional set

Use a compact arc inside a slightly larger smooth arc and a bi-Lipschitz local real parameter there. Write $r = q^{-a}$ with $2/(k+1) < a < 2/(k-1)$. Any trace lying in some radius-r disk on this arc has angles in one of a bounded number of rationally oriented tubes with bounded tangent length and split width $O(\sqrt{r})$, by the chart. Cover these with finitely many boxes of volume $O(r^{(k-1)/2})$, which is above q^{-1} by a fixed power. Lemma E.3 gives $O(qr^{(k-1)/2})$ possible trace parameters, including multiplicities. Each contributes an interval of arc targets of length $O(r)$. Summing over the levels $q = p^n$ proves the upper dimension bound, since the t-dimensional covering sum converges when $at > 1 - a(k-1)/2$. This also gives a convergent sum of Lebesgue lengths for each such a, proving the almost-everywhere conclusion by the covering statement.

For the lower bound fix $0 < t < 1/a - (k-1)/2$ and choose $\sigma > 0$ so small that $at < 1 - a(k-1)/2 - \sigma$. Use a box about each target angle with sufficiently small split width $c\sqrt{r}$ and tangent width comparable to

$$q^{-1+\sigma}r^{-(k-1)/2}.$$

Its product of side scales is at least $q^{-1+\sigma/2}$ for sufficiently large n. Lemma E.3 gives an angle from a fixed nonzero residue class for every target. Moving this angle's tangent coordinate alone follows the boundary. Along that whole line the split derivative vanishes, so the resulting trace is at most $r/4$ from the larger arc if c is small. Its closest projection to that arc is within $O(q^{-1+\sigma}r^{-(k-1)/2})$ in parameter of the original target. One closest projection per trace parameter suffices, because the constructed points stay away from the larger arc's endpoints.

Place intervals of radius $c'r$ about the projections of all trace points within $r/4$ of the larger arc, with $c' > 0$ small enough that every target in one interval is within r of its trace. Inflating the radii to $(c'r)^t$ covers the compact arc at every sufficiently large level by the strict exponent inequality. The local one-dimensional mass transference principle [53, Theorem 2] gives dimension at least t for the original interval limsup. Let t increase to $1/a - (k-1)/2$ and exhaust the open smooth arc by compact subarcs. The dimensions at the two endpoints follow by inclusion and the intermediate bounds. Below the first exponent the covering statement already gives all smooth targets.

For $a > 2/(k-1)$ there are eventually no hits on a compact smooth arc. Project the angular numerator $q\theta$ onto the $k-1$ rational split coordinates, using differences of repeated angles. Take a common lift for those angles along the fiber, so this map annihilates the whole fiber line in each local chart. Changing a chosen lift by an integer only adds multiples of q to the projected numerator. The fixed matrix has independent integer rows. The resulting integer vector A is congruent to the unit root parameter times the image of V under that matrix. Those $k-1$ entries are linearly independent over $\mathbb{Q}$. All real side widths are $O(\sqrt{r})$, whose product is strictly below q^{-1} by a power. Lemma E.4 rules out this vector. The conclusion holds on each compact smooth arc, and therefore rules out infinitely many hits at every fixed smooth target. This completes Theorem E.2.

E.5 A fixed set of split primes

The same statements hold for the order-d period trace over a fixed finite set of split primes $\mathcal{P}$, with $q = \prod_{p\in\mathcal{P}} p^{n_p} \to \infty$ and $n_p \geq 1$. Choose a fixed primitive dth root in each $\mathbb{Z}_p$ and use their reductions by the Chinese remainder theorem to define ζ_q. For conditioned counts use a unit residue modulo $P = \prod_{p\in\mathcal{P}} p$, or a nonempty union of such residues. Constants and the threshold may now depend on $\mathcal{P}$. The infinitely-often statements hold along any unbounded collection of these moduli.

Indeed, in the dual proof the local form for each p contributes $p^{-(n_p-1)}$ for the affine grid, giving P/q jointly. Use the subspace theorem at ∞ and all $p \in \mathcal{P}$, completing the independent systems at each local place as before. The primitive-vector reduction still holds because $g\prod_{p\in\mathcal{P}} |g|_p \geq 1$ for positive integers g. On a fixed rational line a common multiplier is divisible by $\prod_p p^{n_p - O_p(1)}$, so the same height contradiction applies. In the primal exclusion the product from the $l-1$ local annihilators is $q^{-(l-1)}$. On any exceptional rational subspace the projected root line fails to lie in the subspace at every p, by the same rational independence. At least one local depth tends to infinity along any unbounded subsequence after passage to a further subsequence. This gives the same contradiction with the bounded valuation of a coordinate. For algebraic targets, the forms including the last coordinate q contribute q^{-k} jointly, and a fixed rational exceptional relation would bound every n_p already by its local nonzero value. Thus the local counting and coverage estimates persist. Finally $\sum_q q^{-c} < \infty$ over such moduli for each $c > 0$, by a product of geometric sums, and the lower-dimension proof uses coverage at every sufficiently large chosen modulus. This proves the metric statements in the larger setting.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including the rectangular lattice argument, and with drafting and checking the proof. No data are needed for the results.

F Composite cyclotomic boundary arcs and corners

The trace image at order $d = \ell^b$ is a Minkowski sum of ℓ^{b-1} filled hypocycloids. We determine the boundary near every point at which this image contacts a side of its convex hull. The adjoining arcs are translated copies of a single hypocycloid arc, and at $\ell = 3$ these give the entire boundary. The smooth arcs have one linear angular direction and otherwise quadratic directions, yielding shrinking disk laws, an algebraic target exclusion, and an exact dimension for exceptional targets. At a mixed corner all directions are quadratic and the scaled trace measure has an explicit density in the cone generated by two inward rays.

Keywords: cyclotomic periods, Minkowski sums, hypocycloids, shrinking trace disks.
MSC 2020: 11L05, 11J87, 28A78, 52A10.

F.1 The boundary of the sum

Let $\ell \geq 3$ be an odd prime, $b \geq 2$, $m = \ell^{b-1}$ and $k = m(\ell-1)$. We write H_ℓ for the filled hypocycloid bounded by $(\ell-1)u + u^{-(\ell-1)}$, $|u| = 1$. The angular trace at order $d = \ell^b$ is a sum of m traces on independent determinant-one ℓ-phase blocks. To see the grouping, use

$$\Phi_{\ell^b}(X) = 1 + X^m + \cdots + X^{(\ell-1)m}.$$

For each residue of the power exponent modulo m, the first $\ell - 1$ powers give independent phase coordinates and the last is minus their sum in the angular coordinates. Thus the image is mH_ℓ and

angle Haar is product Haar on the blocks. This decomposition is due to Garcia, Hyde and Lutz [54, Corollary 1] and is used by Untrau [25, Proposition 5.4] for the period laws. Garcia, Hyde and Lutz also illustrate the threefold deltoid sum in their Figure 11.

Rotate by a unit complex number so that the two adjacent cusp rays about a selected side have angles $\pm\alpha$, where $\alpha = \pi/\ell$. Use outward and tangential coordinates X, Y. Put $M = \ell\cos\alpha$, $A = \ell\sin\alpha$, and denote the rotated unit cusp directions by $\omega_- = e^{-i\alpha}$ and $\omega_+ = e^{i\alpha}$. The support line of the sum on this side is $X = mM$. Put $Y_j = (2j - m)A$ for $0 \le j \le m$.

Theorem F.1. *The points of mH_ℓ on $X = mM$ are exactly (mM, Y_j). In the gap between Y_j and Y_{j+1} the outer boundary near either endpoint is the translated arc*

$$X = (m-1)M + h\big(Y - (2j - m + 1)A\big), \tag{44}$$

where h is the graph of

$$X = (\ell-1)\cos t - \cos((\ell-1)t), \qquad Y = (\ell-1)\sin t + \sin((\ell-1)t), \qquad |t| \le \alpha.$$

On these open arcs each fiber has one moving block, j upper cusp blocks and $m-1-j$ lower cusp blocks. Its number of labeled angular preimages is $\ell m\binom{m-1}{j}$. At an interior contact $1 \le j \le m-1$, the two incident inward tangent rays are $-\omega_-, -\omega_+$, making opening angle 2α.

When $\ell = 3$, equation (44) holds across every entire gap. The full boundary consists of $3m$ translated deltoid arcs. Its perimeter is $16m$ and the enclosed area is

$$\frac{27\sqrt{3}}{4}m^2 - \left(\frac{27\sqrt{3}}{4} - 2\pi\right)m.$$

Proof. The support of one block is $X \le M$, with equality just at the two adjacent cusps. In fact on the rotated boundary the critical points of X satisfy either $\cos(\ell t/2) = 0$ (a cusp) or $\sin((\ell-2)t/2) = 0$. In the latter case $X = (\ell-2)\cos t \le \ell - 2 < M$. The cusp maxima are exactly the two at $\pm\alpha$. The sum then meets the common support line at the listed $m+1$ points.

On the connecting arc, $Y' = 2(\ell-1)\cos(\ell t/2)\cos((\ell-2)t/2) > 0$ in the interior and $h' = \tan((\ell-2)t/2)$. In particular h is strictly convex, $h(\pm A) = M$ and its limiting endpoint slopes are $\pm\cot\alpha$. This arc is the rightmost point in H_ℓ at each fixed $Y \in [-A, A]$. Indeed the complete hypocycloid is a radial graph. For the unrotated curve written $e^{it}(\ell - 1 + e^{-i\ell t})$, the angle derivative is

$$\frac{(\ell-1)(\ell-2)(1-\cos(\ell t))}{|\ell - 1 + e^{-i\ell t}|^2},$$

positive except at cusp parameters. The connecting arc has positive X and spans the angular sector between its endpoints. If a point with the same Y had greater X, it would lie on a ray strictly inside that sector and beyond its radial boundary, because the boundary's $|Y|$ increases with its absolute angular direction there.

In a sufficiently small collar $X > M - \varepsilon_0$ the points of one block lie near one of these two cusps, by compactness and the support equality case. They satisfy $|Y| \le A$. For instance at the upper cusp the local angle expansion is $\ell\omega_+ - \omega_+ Q + O(\|\theta\|^3)$ with a positive definite real quadratic form Q; its imaginary part decreases from A. The lower cusp is analogous. Near a contact Y_j the candidate in (44) lies in mH_ℓ and has normal deficit less than ε_0. Every maximizing tuple therefore lies in the collar in each block. It remains to minimize $\sum_i (M - h(Y_i))$ for fixed $\sum_i Y_i$, with $|Y_i| \le A$. Strict concavity of $M - h$ forces at most one interior coordinate: varying any two interior coordinates in opposite directions strictly concavifies along a segment. Hence the maximizing tuple has the

specified endpoint allocations, and each of its blocks is rightmost at its own Y_i. This proves the contact arcs and the formula. The moving block has ℓ preimages, one for each singleton position. Indeed, at a boundary fiber the differential forces at most two distinct phases. Two clusters of multiplicity at least two would permit opposite transverse quadratic displacements under within-cluster splitting and give an interior image. The boundary parameter with one singleton is unique by the strict geometric-sum triangle inequality. A cusp has a unique all-aligned angle tuple. Counting positions proves the stated multiplicity. The endpoint slopes give the two inward tangent rays. These rightmost graphs are boundary arcs because every graph point lies in the sum and a point directly outward from it lies outside the sum.

For $\ell = 3$ the convex hull is a triangle with $|Y| \le A$ everywhere, so the same optimization works for the whole side. The filled deltoid is star-shaped by its radial graph, and its Minkowski sum remains star-shaped. On a translated arc with $X = (m-1)M + h(y)$ and $Y = Y_0 + y$, the quantity $X - Yh'(y)$ is positive, because $h(y) - yh'(y) > 0$ by strict radial monotonicity and $(m-1)M - Y_0h'(y) \ge 0$. Thus these rightmost arcs cover every ray between the two outer cusps; their rotations cover the other sides. A point farther on the same ray would violate the rightmost bound at its Y: the angle of the rightmost point increases with Y across a side. Hence there is no additional rim. The deltoid has perimeter 16 and area 2π by integration of its parametrization, and its circumscribed triangle has area $27\sqrt{3}/4$. The loss of signed area when replacing one chord by its arc is invariant under translation. Green's formula, applied along the three sides with m translated arcs each, gives the area and perimeter in the theorem. □

F.2 Trace counts along a smooth arc

Fix $p \equiv 1 \pmod d$, put $q = p^n$, and let $T_q(b) = \sum_{h=0}^{d-1} e^{2\pi i b\zeta^h/q}$ with ζ a primitive dth root modulo q. Let μ be the pushforward of angular Haar, and μ_q the trace law over unit b, or over any nonzero residue class modulo p and unions of such classes. Statements below concern the smooth arcs verified in Theorem F.1, and compact subarcs for uniform estimates.

Theorem F.2. *For any $\eta > 0$ and $r \to 0$ with $r \ge q^{-2/(k+1)+\eta}$, uniformly on a compact subarc,*

$$\mu_q(B(z,r)) = \mu(B(z,r))(1 + O(q^{-c})), \qquad \mu(B(z,r)) = C(z)r^{(k+1)/2}(1 + O(\sqrt{r}))$$

with $c > 0$ and $C(z) > 0$. If the moving block's normalized torus angle $t/(2\pi)$ is real algebraic, then $B(z, q^{-2/(k+1)-\eta})$ is eventually empty of unit traces. At almost every point the nearest unit trace has distance $q^{-2/(k+1)+o(1)}$.

The Hausdorff dimension on any open subarc of targets within q^{-a} of a unit trace at infinitely many depths is

$$\begin{cases} 1, & a \le 2/(k+1), \\ 1/a - (k-1)/2, & 2/(k+1) < a < 2/(k-1), \\ 0, & a \ge 2/(k-1). \end{cases}$$

For $a > 2/(k-1)$ the set is empty. The same conclusions hold on the even subsequence.

Proof. For the moving block put its singleton last and vary the repeated angles by $s + w_i$ with $\sum_i w_i = 0$. The other blocks remain at their cusp centers when only s varies. Use coordinates W formed by the split coordinates of the moving block and all angle offsets in the other blocks. The real tangent s and these $k-1$ coordinates form a rational linear chart (an orthonormal split basis can be used for Euclidean volumes). The trace difference is

$$T - z = A_z s + Q_z^{\mathbb{C}}(W) + O(s^2 + \|W\|^3 + |s|\|W\|^2),$$

where A_z is the nonzero tangent vector and $Q_z^{\mathbb{C}}$ is a sum of quadratic forms oriented along the negative repeated phase in the moving block and the negative cusp phases in the frozen blocks. With unit outward normal proportional to $(1, -h')$, the cusp projections of the positive phase vectors are positive, because $\cos\alpha \mp h' \sin\alpha > 0$ on the open arc. For the repeated phase of the moving block the corresponding dot product before normalizing is

$$\cos t - h' \sin t = \frac{\cos(\ell t/2)}{\cos((\ell-2)t/2)} > 0.$$

Thus the inward normal quadratic form Q_z on W is positive definite. The size of the trace difference is comparable to $|s| + \|W\|^2$. All constants are uniform on a compact subarc by the finite fiber description.

Put $s = ra$, $W = \sqrt{r}\, w$. In normalized coordinates the Taylor error is $O(\sqrt{r})$, and the model sublevel disk volume is homogeneous in its radius with degree $(k+1)/2$. If J is the constant Haar Jacobian in these coordinates and F the number of labeled fibers, one can write the leading coefficient as

$$C(z) = \frac{FJ|S^{k-2}|}{|A_z|\sqrt{\det Q_z}} \int_0^1 y^{(k-1)/2-1} \sqrt{1-y^2}\, dy.$$

Indeed at fixed w the allowed a-interval has length $2|A_z|^{-1}\sqrt{1 - Q_z(w)^2}$, and radial integration in Q_z gives the formula. Here $\det Q_z$ denotes the determinant of its matrix in the chosen Euclidean split and cusp coordinates. The actual radius-r collar of relative width δ has relative Haar mass $O(\delta)$, by dilating (s, W) to $(ts, \sqrt{t}W)$. In the collar the derivative in t of the trace distance is the leading homogeneous magnitude, at least a constant times r, plus $O(r^{3/2})$. Dilation and its Jacobian compare the actual disks at relative radii $1 \pm \delta$.

The rectangular covering lemma proved below applies with $r, \sqrt{r}, \dots, \sqrt{r}$ as side scales, whose product is at least $q^{-1+\eta(k+1)/2}$. The split derivative is $O(\sqrt{r})$ in the disk, so the resulting coupled angular errors move the trace by $O(rq^{-c})$. Apply the actual collar bound at each fiber, adding their contributions. For algebraic targets the angular fiber coordinates are real algebraic (all the frozen phases have rational angles). The algebraic exclusion in the lemma below gives the hole at scale $q^{-2/(k+1)-\eta}$ from the product of those same scales.

For the dimension bound let $r = q^{-a}$ in the strict range between the two exponents. All hit points lie in finitely many tubes with tangent length $O(1)$ and split widths $O(\sqrt{r})$. The rectangular lemma bounds their point count by $O(qr^{(k-1)/2})$. Each contributes an interval of targets of length $O(r)$ in an arc parameter, giving the upper bound by a geometric covering sum over $q = p^n$. The same sum for lengths gives the almost-everywhere lower bound on nearest distances. Below the first exponent, the covering statement from the disk comparison gives every target eventually.

For the lower dimension bound use instead split widths $c_0\sqrt{r}$ and tangent width $q^{-1+\sigma}r^{-(k-1)/2}$ with arbitrarily small fixed $\sigma > 0$. The rectangle has volume above q^{-1} by a power. Starting at any target on a compact subarc, the rectangle contains an angle configuration. Its trace lies within $r/4$ of a slightly larger arc if c_0 is small: the split derivative vanishes at every point of the tangent fiber line. A closest projection of the trace to the larger arc is within $O(q^{-1+\sigma}r^{-(k-1)/2})$ in parameter from the original target. Intervals of radius $c_1 r$ around these projections are contained in the actual hit sets. Their enlarged intervals of radius $(c_1 r)^t$ cover the compact subarc at every large level whenever $at < 1 - a(k-1)/2 - \sigma$. The local mass transference principle [53, Theorem 2] gives dimension at least t. Letting t approach the claimed dimension, and using inclusion at the endpoints, proves the formula. Finally if $a > 2/(k-1)$, project off the tangent line after clearing the fixed rational cusp offsets. Its $k-1$ real scales have product $r^{(k-1)/2} < q^{-1}$ by a power. The primal exclusion of the lemma below shows that no point can lie in this tube at sufficiently large depths. This proves emptiness for infinitely-often hits. All arguments apply along even depths. □

F.3 The law at a mixed corner

Keep the rotated coordinate system of Theorem F.1, and let $z_j = mM + iY_j$ for $1 \le j \le m-1$. Write $k_- = (m-j)(\ell-1)$, $k_+ = j(\ell-1)$, so $k_- + k_+ = k$.

Theorem F.3. *For every bounded Borel continuity set $E \subset \mathbb{C}$ of the measure on the right,*

$$\lim_{\epsilon \downarrow 0} \epsilon^{-k/2} \mu\{T : (z_j - T)/\epsilon \in E\} = C_j \int_{\substack{a,b \ge 0 \\ a\omega_- + b\omega_+ \in E}} a^{k_-/2-1} b^{k_+/2-1}\, da\, db,$$

where

$$C_j = \binom{m}{j} \frac{\ell^{-m/2}(2\pi)^{-k/2}}{\Gamma(k_-/2)\Gamma(k_+/2)}.$$

The arithmetic law μ_q has the same limit at a fixed split prime when $\epsilon \to 0$ and $\epsilon q^{2/k} \to \infty$. The nearest unit trace is at distance of order $q^{-2/k}$, including in each nonzero residue class modulo p.

Proof. The support equality forces all blocks to be aligned cusps, with exactly j upper. There are $\binom{m}{j}$ such angle fibers. At a cusp direction ω, the complex deficit from that block is $2\pi^2\omega\,(\sum_{h=1}^{\ell-1} \theta_h^2 + (\sum_h \theta_h)^2) + O(\|\theta\|^3)$. Its real projection on the supporting direction is positive definite since $\cos\alpha > 0$. The total deficit of a configuration in a bounded scaled set therefore localizes to offsets $O(\sqrt{\epsilon})$ near the listed fibers. Divide those offsets by $\sqrt{\epsilon}$. In the limit the two groups give $a\omega_- + b\omega_+$ with a, b the quadratic radii in dimensions k_-, k_+. Diagonalizing the block forms introduces the Jacobian $\ell^{-1/2}$ per block, since $I + \mathbf{1}\mathbf{1}^{\mathsf{T}}$ has determinant ℓ. Changing each group radius to its squared radius times $2\pi^2$ gives factor $(2\pi)^{-k_\pm/2}/\Gamma(k_\pm/2)$. The error tends to zero on bounded sets and dominated convergence on the localized angle coordinates gives the claimed limit, using the continuity set hypothesis.

For the arithmetic comparison the angle lattice in a unit residue class has covering radius $O_{d,p}(q^{-1/k})$. Indeed its dual consists of integral power-basis vectors whose evaluations at the chosen p-adic root vanish modulo q/p. If the vector is nonzero its cyclotomic norm is a nonzero integer divisible by q/p and bounded above by a constant times its height to the power k. Thus its dual first minimum is $\gg_d (q/p)^{1/k}$. Transference and Voronoi coupling give the asserted covering radius and a Haar coupling. First derivatives vanish at the cusp fibers. The local trace displacement is therefore $O(\sqrt{\epsilon}\, q^{-1/k} + q^{-2/k})$, which is $o(\epsilon)$ at the asserted scales. Bounding a continuity set by fixed small enlargements and contractions, and then sending their width to zero, yields the arithmetic limit.

These angle fibers have fixed rational coordinates with denominators dividing ℓ. If a fine unit grid point approaches a fiber angle, multiplying its difference by that fixed denominator produces a nonzero fine lattice point modulo the torus lattice: the unit vector has additive order q, and the denominator is a unit modulo q. The fine lattice has shortest nonzero vector $\gg_d q^{-1/k}$. For completeness, its dual determinant is q and its first minimum $\gg_d q^{1/k}$ by the same norm. Minkowski's second theorem bounds its last minimum by $O_d(q^{1/k})$. A primal vector shorter than a sufficiently small constant times $q^{-1/k}$ would have zero integer pairings with k independent such dual vectors and would vanish. Thus every unit angle stays at distance $\gg_d q^{-1/k}$ from each cusp fiber. The supporting real quadratic form gives the lower trace radius $\gg_d q^{-2/k}$. Conversely the covering mesh in any fixed residue class gives an angle at distance $O_{d,p}(q^{-1/k})$ from the chosen fiber and hence a trace within $O_{d,p}(q^{-2/k})$. This proves the last assertion. □

F.4 Rectangular comparison and exclusion

Here is the arithmetic device used in Theorem F.2. Let V denote the tuple of the first k cyclotomic powers in the block ordering. Its entries are $\mathbb{Q}$-linearly independent. The affine angle lattice of one

unit residue class is a translate of the torus grid with mesh vector $V/(q/p)$ and size q/p. Its dual is the integer lattice on which the algebraic linear form $L(u) = u \cdot V$ is divisible by q/p.

Fix an invertible rational coordinate map and side scales $0 < s_i \le 1$ with $\prod_i s_i \ge q^{-1+\sigma}$, $\sigma > 0$. Inverting the side scales on the primal lattice multiplies the dual coordinates by the side scales in the dual rational frame. A putative scaled dual vector of length at most q^c has archimedean coordinate bounds q^c/s_i, of product at most $q^{1-\sigma+kc}$, and height $O(q^{1-\sigma+c})$. At p the factor $|L(u)|_p \le p/q$ can be completed by bounded coordinate forms to independent forms. The product bound is a fixed negative power of the height when c is small. Schlickewei's subspace theorem [51] puts these vectors in finitely many proper rational subspaces. Within each such subspace choose independent restrictions of the same real-coordinate forms and supplement L at p in an integral basis. Discarding real bounds at least one cannot worsen the product bound, and L is nonzero on every nonzero rational vector. The argument descends to finitely many rational lines. On a line the integer multiplier must be divisible by $p^{n-O(1)}$, contrary to the height bound. Thus the scaled primal covering radius is $O(q^{-c})$ by transference, giving Haar-to-grid coupling with coordinate errors $O(q^{-c}s_i)$. Comparing rectangular enlargements and contractions gives the box count. The form of the subspace theorem for primitive vectors suffices, since dividing by the coordinate gcd decreases the local product at ∞, p and does not increase the height.

For exclusion with scales whose product is at most $q^{-1-\sigma}$, first fix an algebraic angle target θ_0. A candidate angle M/q has $M \equiv bV \pmod q$ with unit b. Apply the subspace theorem to the $k+1$ coordinates (M, q), using the fixed anisotropic real-coordinate map applied to $M - q\theta_0$ and the last coordinate q. At p use q, the first coordinate M_1 (where the basis power is 1), and the $k-1$ forms annihilating the line V. The product over both places is $O(q \prod s_i)$. An exceptional rational relation $c \cdot M + c_0 q = 0$ would force $c \cdot V$ divisible by q, impossible for fixed nonzero c; if $c = 0$ then $q > 0$ contradicts the relation. For exclusion in a thin tube about an affine rational tangent line, project onto its $k-1$ rational split coordinates after clearing the fixed denominators of the translation. The resulting integer tuple is congruent to a fixed integer times b times the projected V. The entries of this projected vector remain $\mathbb{Q}$-linearly independent, because the projection has independent rational rows. In these $k-1$ variables the real coordinate product and the $k-2$ small p-adic factors again give $O(q \prod s_i)$. Every proper rational exceptional subspace misses the corresponding p-adic line, so the small forms force all its coordinates to have valuation at least $n - O(1)$, contrary to the fixed valuation of a nonzero projected coordinate. In dimension one the real bound itself forces the integer coordinate to vanish, again impossible. This gives the exclusions used above.

For single-prime powers the period statements transfer to tame stationary hyper-Kloosterman traces with root packet size d when there is a common Gaussian rotation, by the same root multiplicity factor d [23, 50]. The lattice arguments for the periods also work for moduli supported on any one fixed finite set of split primes. Use the CRT primitive roots, unit classes modulo the product of the primes, and the subspace theorem at all of those places. The small local factors multiply to the corresponding powers of $1/q$ up to a constant, and on a rational exceptional line the multipliers forced by the local congruences multiply. Geometric sums of q^{-c} over the permitted moduli still converge, so the dimension conclusions apply along any unbounded collection. For the fixed-corner angular mesh the norm is divisible by the product modulus after removing the fixed affine residue denominator.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, including the boundary fiber and disk calculations, and with drafting and checking the arguments. No data are needed for the results.

G Microscopic trace counts at mixed cyclotomic corners

For odd prime-power order ℓ^b with $b \geq 2$, we determine limiting count laws for cyclotomic periods in disks of radius $\lambda q^{-2/k}$ about mixed exposed corners, where $q = p^n$ and k is the cyclotomic degree. At a fixed split prime the empirical law over depths is obtained by counting points of explicit affine lattices along an orbit on a compact torus. Averaging over split primes gives full torus Haar measure, with a mixture determined by the power map on a finite ray class group. The prime-window comparison has arbitrary logarithmic power saving for common depths growing up to $\exp(c \log X/\log\log X)$. With suitably chosen local roots, independently varying the depth at $k/2$ prescribed split primes also gives full torus Haar measure. We give an unfolded norm-sum formula for the mean and prove continuity and nondegeneracy of the limiting nearest-distance distributions. At order nine the prime model is depth-independent and its first possible radius is $2\pi^2/3$.

Keywords: cyclotomic periods, affine lattice counts, Hecke characters, shrinking targets.
MSC 2020: 11L05, 11R44, 11J71.

G.1 Angular coordinates and the local quadratic form

Fix $d = \ell^b$ for an odd prime ℓ and $b \geq 2$. Put $m = \ell^{b-1}$, $k = m(\ell-1) = 2r$, and $K = \mathbb{Q}(\zeta_d)$. If $p \equiv 1 \pmod{d}$, choose an order-d root ρ modulo $q = p^n$ and write

$$T_{p,n}(a) = \sum_{j=0}^{d-1} \exp(2\pi i a \rho^j/q), \qquad a \in (\mathbb{Z}/q\mathbb{Z})^\times.$$

The choice of ρ permutes the summands. Angular coordinates $\theta_j \in \mathbb{R}/\mathbb{Z}$ for $0 \leq j < d$ satisfy $\sum_{v=0}^{\ell-1} \theta_{u+vm} = 0$ for each $0 \leq u < m$. The first k powers give free coordinates. This is the determinant-one block decomposition of Garcia, Hyde and Lutz [54, Corollary 1]; see also Untrau [25, Proposition 5.4]. The trace $T(\theta)$ sums the d phases. In each block, equality on a supporting line of the convex hull between consecutive ℓth roots forces all ℓ phases to equal one of the two consecutive roots ν_-, ν_+. Indeed each block image is the filled ℓ-cusped hypocycloid, and checking its standard boundary curve $((\ell-1)e^{it} + e^{-i(\ell-1)t})$ gives exactly those two contacts. Equivalently one may use the explicit support computation in Section F.

Choose $0 < j_0 < m$ and the corner $z = (m-j_0)\ell\nu_- + j_0\ell\nu_+$. Its angular fiber is finite and consists of the $\binom{m}{j_0}$ assignments of aligned blocks. The coordinates of each fiber point belong to $\ell^{-1}\mathbb{Z}/\mathbb{Z}$. Use $K_\mathbb{R} = K \otimes_\mathbb{Q} \mathbb{R} \simeq \mathbb{C}^r$, under one embedding from each complex conjugate pair, and the real isomorphism

$$L(y) = (\mathrm{Tr}(y\zeta_d^j))_{0 \leq j < k}.$$

Here the trace extends $\mathbb{R}$-linearly. Complete these to d real coordinates by the block sum zero rule, and denote the completed tuple by $L^*(y)$. The lattice $L(\mathfrak{D}^{-1})$ is $\mathbb{Z}^k$, since $(1, \zeta_d, \ldots, \zeta_d^{k-1})$ is an integral basis and $\mathfrak{D}$ is the different. Write $\omega = \zeta_d^m$ and put $\mathfrak{f} = (1-\omega)\mathcal{O}_K$. The lift for the fiber labeled by $c_a \in \{0, \ldots, \ell-1\}$, with aligned phase ω^{c_a} on block a, can be chosen as

$$dx = -\omega \sum_{a=0}^{m-1} c_a \zeta_d^{-a}. \tag{45}$$

Indeed $\mathrm{Tr}(-\omega\zeta_d^{j-a}) = m$ if $j \equiv a \pmod{m}$ and is zero otherwise for $0 \leq j < k$, by the prime-power Ramanujan sum. Also $\Phi_d'(\zeta_d) = d\zeta_d^{-1}/(\omega-1)$, so $d\mathfrak{D}^{-1} = \mathfrak{f}$ and $\mathfrak{f}x \subset \mathfrak{D}^{-1}$. The sole prime dividing

$\mathfrak{f}$ is the prime above ℓ. Each mixed-fiber lift is nonzero modulo $\mathfrak{D}^{-1}$ at that prime, because it has at least one nonintegral angular coordinate. Its quadratic form is

$$Q_x(y) = 2\pi^2 \sum_{u=0}^{m-1} \nu_u \sum_{v=0}^{\ell-1} L^*_{u+vm}(y)^2, \tag{46}$$

where ν_u is that fiber's aligned phase in block u. Indeed $z - T(L(x) + L(y)) = Q_x(y) + O_d(\|y\|^3)$ since the linear term sums to zero within each block. The projection of Q_x onto the outward normal is a positive definite real quadratic form: the normal makes an angle π/ℓ with each $\nu_\pm$. In particular $|Q_x(y)| \asymp_d \|y\|^2$. Compactness and the description of the fiber imply that a critical disk of scaled radius λ contains only offsets $\|y\| \ll_{d,\lambda} q^{-1/k}$ for large q.

The root in the definition of $T_{p,n}$ selects a prime $\mathfrak{p}$ over p, and its unique lift to $\mathbb{Z}_p$ evaluates ζ_d at that prime. In the free angular coordinates the complete grid lifts to $\mathfrak{p}^{-n}\mathfrak{D}^{-1}$, with excluded nonunits lifting to $\mathfrak{p}^{1-n}\mathfrak{D}^{-1}$. For example if $y \in \mathfrak{p}^{-n}\mathfrak{D}^{-1}$ then the fractional part of $\mathrm{Tr}(y\zeta_d^j)$ is its p-primary fractional part and only the embedding at $\mathfrak{p}$ contributes modulo $\mathbb{Z}_p$. It therefore agrees with $a\rho^j/q \pmod 1$, where a/q is the local fractional part of y. The index of the grid over $\mathfrak{D}^{-1}$ is q and the index of the excluded grid is q/p, giving the asserted equalities.

Denote by $N_{p,n}(\lambda)$ the number of unit parameters with $|T_{p,n}(a) - z| \le \lambda q^{-2/k}$, and by $M_{p,n}$ the nearest trace distance multiplied by $q^{2/k}$. Parameters are counted with multiplicity. All statements hold with simultaneous fixed positive radii.

G.2 The torus and the fixed-prime law

Let $F = \{g \in (\mathbb{C}^\times)^r : \prod_i |g_i|^2 = 1\}$, $U = \{u \in \mathcal{O}_K^\times : u \equiv 1 \pmod{\mathfrak{f}}\}$, and $G = F/\sigma(U)$, where σ denotes the chosen archimedean embeddings. By the unit theorem, G is a compact connected torus of dimension $k-1$. The unit logarithms form a lattice in the zero-sum hyperplane, and a compact logarithmic fundamental domain times the compact argument torus maps onto G. For $\alpha \in K^\times$, write $g_\alpha = |\,\mathrm{N}\,\alpha|^{1/k}/\sigma(\alpha) \in F$, with the ratio taken coordinatewise. We also write g_α for its class in G.

For any choice of U-invariant affine sets $\Lambda_x \subset K$ indexed by the fiber lifts, and a real scale $s > 0$, define

$$\mathcal{N}_{\Lambda,s}(g,\lambda) = \sum_x \#\{\gamma \in \Lambda_x : \ |Q_x(sg\sigma(\gamma))| \le \lambda\}. \tag{47}$$

We only use affine lattices and differences of affine lattices not containing zero. The model is independent of the lift of g. On a finite set of compact local lifts, only finitely many γ can contribute at bounded λ.

Theorem G.1. *Fix a split p and a prime $\mathfrak{p}$ above it. Let h be the order of $[\mathfrak{p}]$ in the ray class group modulo $\mathfrak{f}$, and choose $\pi \equiv 1 \pmod{\mathfrak{f}}$ with $(\pi) = \mathfrak{p}^h$. Let H be the cyclic orbit closure of g_π in G. For $0 \le t < h$ put*

$$\Lambda_{x,t} = (\mathfrak{p}^{-t}\mathfrak{D}^{-1} - x) \setminus (\mathfrak{p}^{1-t}\mathfrak{D}^{-1} - \pi x).$$

The empirical laws as $n \to \infty$ through $n \equiv t \pmod h$ of the vector of counts $N_{p,n}(\lambda)$ converge to those of $\mathcal{N}_{\Lambda_t,p^{t/k}}(g,\lambda)$ for Haar-random $g \in H$. The nearest-distance law is the minimum of the same model radii and is atomless. For all depths take the average of the h laws. An arithmetic progression of depths is handled by using a corresponding multiple of h.

Proof. Write $n = hN + t$, $N \ge 1$. Multiply each angular displacement by π^N. The fine coset becomes $\mathfrak{p}^{-t}\mathfrak{D}^{-1} - x$ because $(\pi^N - 1)x \in \mathfrak{D}^{-1}$. The excluded coset becomes $\mathfrak{p}^{1-t}\mathfrak{D}^{-1} - \pi x$ because $(\pi^N - \pi)x \in \mathfrak{p}^h\mathfrak{D}^{-1}$. The same congruence shows invariance under multiplication by U. Thus the

normalized angular displacement in $\mathbb{C}^r$ is exactly $p^{t/k} g_\pi^N \sigma(\gamma)$. In a fixed bounded region the model trace differs from the exact normalized trace by $O_{d,\lambda}(q^{-1/k})$. Each local grid point corresponds to one parameter and different fibers give disjoint neighborhoods for large n. Therefore the exact counts are squeezed between the model counts at radii $\lambda \pm O_{d,\lambda}(q^{-1/k})$.

The projection of g_π to the logarithmic unit torus is nontorsion. Otherwise π^a times a unit, for some nonzero a, would have all conjugates of the same absolute value. Its product with its complex conjugate would then be rational, yet its fractional ideal is supported at just $\mathfrak{p}$ and $\overline{\mathfrak{p}}$ among the $k \geq 6$ primes above p. This is impossible. The identity component of H therefore has a Lie direction with a nonzero logarithmic part. In a lift to F some log speed is positive because their sum is zero. Along the lifted line, $\|g\sigma(\gamma)\|$ is unbounded for every nonzero γ, since none of its embeddings vanishes. Consequently the real analytic function $|Q_x(p^{t/k} g\sigma(\gamma))|^2$ is nonconstant on every local lift of every component of H. Every positive level is Haar-null, and the zero level is empty. Compactness and local finiteness then give Haar-null shells for the model. Equidistribution of powers on their compact abelian orbit closure and the radius squeeze imply the result.

The candidate affine sets have nonzero points in every fiber, and their scaled distance minima vary continuously, as follows from local finiteness. They are bounded above uniformly on compact H, and bounded below by a positive constant because zero never belongs to the fine coset: a mixed fiber has a nontrivial ℓth root in at least one block. The nearest-distance convergence follows from the count laws. If the depths occupy a specified residue class, enlarge the period h to a common multiple and use a power of π; the log argument still holds. □

G.3 Averaging over split primes

Write $\mathcal{C}$ for the ray class group of K modulo $\mathfrak{f}$ and e for its exponent. For $C \in \mathcal{C}$ take any integral ideal $\mathfrak{b}$ in C and prime to ℓ, and put

$$\Lambda_x(C) = \mathfrak{b}^{-1}\mathfrak{D}^{-1} - x, \qquad s(C) = (\mathrm{N}\,\mathfrak{b})^{1/k}.$$

Let $\mathcal{L}_C$ be the joint law of the model counts $\mathcal{N}_{\Lambda(C),s(C)}(g,\lambda)$ for Haar-random $g \in G$. The same notation includes the associated nearest-distance law when used for that variable.

Theorem G.2. *Fix any finite list of positive radii. For each $A > 0$ there is $c > 0$ such that, for uniform $p \in [X, 2X]$ with $p \equiv 1 \pmod{d}$, the counting law at common depth n is within $O((\log X)^{-A})$ in total variation of*

$$\frac{1}{|\mathcal{C}^n|} \sum_{C \in \mathcal{C}^n} \mathcal{L}_C, \qquad \mathcal{C}^n = \{C^n : C \in \mathcal{C}\}, \tag{48}$$

uniformly for $1 \leq n \leq \exp(c \log X / \log\log X)$. The model is independent of ideal representatives. The value of $\gcd(n, e)$ determines the mixture.

The same comparison holds for the distribution function of the nearest distance at each fixed positive radius. It also holds uniformly in radius on bounded intervals. Each limiting nearest-distance law is atomless and nondegenerate.

Proof. Choose $\mathfrak{b}$ for each ray class once. For a degree-one $\mathfrak{p}$ over p with $[\mathfrak{p}]^n = C$, write $\mathfrak{p}^n = (\alpha)\mathfrak{b}$, with $\alpha \equiv 1$ at the prime dividing $\mathfrak{f}$ in the ray sense. At the primes dividing ℓ, $(\alpha - 1)x \in \mathfrak{D}^{-1}$ locally. Away from ℓ, $x \in \mathfrak{D}^{-1}$ locally and the valuation of $\alpha - 1$ is at least $-v(\mathfrak{b})$. Hence

$$\alpha(\mathfrak{p}^{-n}\mathfrak{D}^{-1} - x) = \mathfrak{b}^{-1}\mathfrak{D}^{-1} - x.$$

The same congruence proves U-invariance of the right-hand side. The deleted coset is $\mathfrak{p}\mathfrak{b}^{-1}\mathfrak{D}^{-1}-\alpha x$. When p avoids the finite set dividing the chosen $\mathfrak{b}$ or ℓ, each deleted γ belongs locally to $\mathfrak{p}\mathfrak{D}^{-1}$ at $\mathfrak{p}$. A fixed integer J clearing all the denominators of the fine lattices can be chosen prime to p for all sufficiently large p. Then $J\gamma$ is an integer of K divisible by $\mathfrak{p}$. No candidate equals zero because $x \notin \mathfrak{D}^{-1}$ locally at a prime over ℓ. Therefore $|\,\mathrm{N}(J\gamma)| \geq p$. A bounded archimedean normalized offset, even with an arbitrary torus lift, bounds $|\,\mathrm{N}\,\gamma|$ since the balancing factor has norm one. Thus the deleted coset is absent at bounded radii for all sufficiently large X.

The normalized offsets are exactly $s(C)g_\alpha\sigma(\gamma)$, up to the local Taylor comparison in Theorem G.1. If $\mathfrak{b}' = (\eta)\mathfrak{b}$ represents the same class with $\eta \equiv 1$ at the ray modulus, multiplication by η^{-1} carries the old affine fine lattice onto the new one, since $(\eta^{-1}-1)x \in (\mathfrak{b}')^{-1}\mathfrak{D}^{-1}$ by the same local argument. After the prefactors $s(C)$ are included, the identification translates the torus variable by g_η, which preserves Haar.

Here are details of the quantitative angular comparison. A character χ of G gives a character $(a) \mapsto \chi(g_a)$ on the group of ray-principal ideals. It extends to the group of all fractional ideals prime to the modulus, because that subgroup has finite index and every element of the unit circle has all roots. Choose extensions of a basis once and use their powers, so the extensions of frequency vector $v \in \mathbb{Z}^{k-1}$ are fixed Hecke characters of finite conductor dividing $\mathfrak{f}$ and with archimedean parameters $O_d(\|v\|)$. This is the usual angular Hecke character setup, as in [56, §1.1]. For nonzero v the infinity type is nontrivial and norm-balanced. One way to see this is that generators satisfying the ray congruence are dense at the archimedean places by weak approximation, and the lift of a nontrivial torus character is nontrivial and trivial on positive scalars. The extension and its finite ray twists are therefore neither trivial nor pure norm twists.

In a fixed ray class B for $\mathfrak{p}$, with $C = B^n$, the Fourier coefficient of the empirical balanced shapes is a constant unit factor times the prime-ideal mean of the extension at frequency nv. Indeed $\chi(g_\alpha) = \widehat{\chi}(\mathfrak{p})^n\widehat{\chi}(\mathfrak{b})^{-1}$. Ray-class orthogonality reduces the restricted prime-ideal sum to finitely many finite twists of the same infinity type. By Ghidelli's prime-ideal estimate [55, Lemma 5.4] each nonzero frequency has unnormalized sum over $X \leq \mathrm{N}\,\mathfrak{p} \leq 2X$ bounded by

$$\ll_d X \exp\Bigl(-\frac{c_d \log X}{\sqrt{\log X} + \log(2 + n\|v\|)}\Bigr).$$

Here the logarithmic size parameter in that estimate is bounded by the log of the archimedean parameters plus a fixed constant. Removing higher residue degrees costs $O_d(\sqrt{X})$ terms. The fixed-modulus prime ideal theorem gives $\mathrm{Li}(2X) - \mathrm{Li}(X)$ divided by $|\mathcal{C}|$, with arbitrary logarithmic power relative saving, as the number in each class. The Erdős–Turán–Koksma inequality [57] now yields any prescribed negative power of $\log X$ in box discrepancy. Namely use a cutoff $\|v\|_\infty \leq (\log X)^L$ with L sufficiently large for the prescribed power, and take the constant c in the depth range sufficiently small. The prime shape average is on all degree-one ideals in each class. Once all classes are combined, this is exactly the desired rational-prime average of arithmetic counts because the k primes above each split p all give the same trace multiset.

For completeness we pass from boxes to trace disks. On compact local lifts to F and for radii in a compact interval, only a fixed finite set of candidates needs to be examined. Each shell has a power bound for its collars, uniformly in its radius. To show this, choose the real logarithmic tangent $w = (r-1, -1, \ldots, -1)$ and zero angular speeds. Locally $|Q_x(sge^{tw}\sigma(\gamma))|^2$ is a finite linear combination of e^{at} with a fixed set of exponents a determined by r. The leading coefficient for $a = 4(r-1)$ is positive by the outward positivity of Q_x and the nonzero first embedding of γ. Among the derivatives of orders 1 through the number of distinct nonzero exponents, at each point at least one is bounded away from zero on compact sets, by the Vandermonde matrix on those exponents. Cover by finitely many boxes on which one specified derivative is bounded below in

absolute value. The one-dimensional sublevel estimate in that direction then gives measure $O(\delta^{\eta})$ for the set where the squared radius differs from a specified level by at most δ, with $\eta > 0$ depending only on d. The sublevel estimate follows from divided differences: if a sublevel set had length much more than $(\delta/c_0)^{1/j}$ where c_0 bounds the jth derivative below, one could pick $j+1$ points with that much separation and contradict the mean value formula for the divided difference.

Use a fixed torus coordinate box and subdivide it into boxes of side comparable to $(\log X)^{-L'}$. Periodic seams can be included in the boundary boxes. The shells crossing boxes are contained in collars of comparable width by bounded derivatives. The box discrepancy times the number of subdivision boxes and the total volume of the collars can both be made $O((\log X)^{-A})$ by first choosing L' and then the discrepancy saving. The Taylor error $O(X^{-1/k})$ is negligible in the same collars. This gives the model comparison for each specified joint count event and each specified nearest-distance threshold. Counts on bounded radii take values in a fixed finite set, so this gives total variation for the count vector. The derivative argument, and the finite candidate sets, give uniformity in the nearest-distance radius on bounded intervals. The minima are continuous and bounded away from zero on G, by compactness and local finiteness. The shell bound excludes atoms, ensuring nondegeneracy. The class weights are equal for B and therefore equal on the image $\mathcal{C}^n$. In a finite abelian group of exponent e this image depends only on $\gcd(n, e)$. □

G.4 An unfolded mean

Give F any multiplicative Haar measure and let R_U be the volume of a fundamental domain for $\sigma(U)$. Write

$$V_x(t) = \mathrm{vol}_F\{g : \ |Q_x(g)| \le t\}.$$

In this formula g is regarded as a vector in $\mathbb{C}^r$. The volume is finite, since a bounded norm-one vector has all coordinates bounded away from zero. It vanishes below the positive minimum of $|Q_x|$ on F.

Corollary G.3. *For any ray class C, the mean of its prime model at radius λ is*

$$\frac{1}{R_U} \sum_{x} \sum_{[\gamma] \in \Lambda_x(C)/U} V_x\Big(\frac{\lambda}{s(C)^2 |\,\mathrm{N}\,\gamma|^{2/k}}\Big).$$

The sum is finite at fixed λ. Just above the smallest possible radius of any point in a class model, the count has positive probability of being zero and positive probability of being nonzero. The same conclusion holds for the equal-weight prime mixture, at its smallest possible radius.

Proof. For each fiber, write its affine lattice as a disjoint union of free U-orbits. Integrating the orbit counts against Haar on G unfolds the integral to F with factor $1/R_U$. The balance $\sigma(\gamma)/|\,\mathrm{N}\,\gamma|^{1/k}$ belongs to F, so an invariant change of variable followed by quadratic homogeneity gives the displayed term. For finite sums, the support of V_x bounds $|\,\mathrm{N}\,\gamma|$, and a fixed denominator makes the resulting bounded-norm elements integral. Their principal ideals range over finitely many ideals, with only finitely many U-orbits per ideal.

The least radius is attained for a finite set of norm classes and for some element of F, by positivity and properness of $|Q_x|$ on F. At that radius the mean is zero by shell nullity. At any larger radius an open positive-measure set of shapes contains a point. By continuity of the mean and its integer-valued integrand, on a sufficiently small interval above the least radius the mean is strictly between zero and one. This bounds the probability of a point strictly between zero and one. The mixture is handled by taking the minimum over its finitely many ray classes. □

G.5 The angular projection at fixed prime

The fixed-prime theorem uses the actual orbit closure. Its angular projection can nevertheless be identified. Let w_K be the order of the root-of-unity group of K. For every unit u in this CM field, $u/\bar{u}$ is a root of unity: the ratio is a unit and all of its conjugates have absolute value one. Hence on G there is a well-defined angular projection to $(S^1)^r$ given by

$$g \longmapsto \left((g_i/|g_i|)^{2w_K}\right)_{i=1}^{r}.$$

The projection of the orbit closure of g_π is all of $(S^1)^r$. To see this, an annihilating character with integer frequencies a_i would give

$$\prod_{i=1}^{r} \left(\sigma_i(\pi)/\overline{\sigma_i(\pi)}\right)^{w_K a_i} = 1.$$

The embeddings are automorphisms of this cyclotomic field, and the $2r$ ideals obtained from $\mathfrak{p}$ under these automorphisms and their conjugates are distinct by complete splitting. Their valuations force $a_i = 0$ for every i. The same argument applies to every nonzero power of g_π. Thus the identity component of the orbit closure has dimension at least r.

G.6 Full torus laws with several fixed primes

The full torus also occurs without averaging over primes. Let $p_1, \dots, p_s$ be distinct split primes, choose $\mathfrak{p}_i$ above p_i, and let h_i be the respective ray orders. Choose $(\pi_i) = \mathfrak{p}_i^{h_i}$ with $\pi_i \equiv 1$ at the ray modulus. Put $q = \prod_i p_i^{n_i}$, with $n_i = h_i N_i + t_i$, $0 \le t_i < h_i$, and choose ρ by combining the primitive roots specified by the ideals. Define the period, counts and distances by the earlier formulas with this q. The powers of ρ form the chosen diagonal cyclic subgroup of order d. The local roots need not give the same diagonal subgroup under independent reindexing at different primes.

Theorem G.4. *For each fixed tuple (t_i), as the N_i vary independently and uniformly in $\{1, \dots, M\}$ with $M \to \infty$, the counts and nearest-distance laws converge to the models in (47) with*

$$I = \left(\prod_{i=1}^{s} \mathfrak{p}_i^{-t_i}\right) \mathfrak{D}^{-1}, \qquad s_0 = \prod_{i=1}^{s} p_i^{t_i/k},$$
$$\Lambda_x = (I - x) \setminus \bigcup_{i=1}^{s} (\mathfrak{p}_i I - \pi_i x).$$

The shape is Haar-random on the closed group generated by the g_{π_i}. The same limits hold for the uniform average over these moduli $q \le X$ as $X \to \infty$. If $s \ge r = k/2$, for any prescribed distinct split primes it is possible to choose one ideal above each so that this group is G.

Proof. The trace-grid argument at each prime and the Chinese remainder theorem give the fine ideal $(\prod_i \mathfrak{p}_i^{-n_i})\mathfrak{D}^{-1}$. Nonunit parameters are the union of its subideals obtained by multiplying by one $\mathfrak{p}_i$. On multiplying offsets by $\mu = \prod_i \pi_i^{N_i}$, the translated fine ideal becomes $I - x$. For the ith deleted ideal, $(\mu - \pi_i)x \in \mathfrak{p}_i I$. At the ray modulus this follows from $\mathfrak{f}x \subset \mathfrak{D}^{-1}$ and the ray congruences. At $\mathfrak{p}_i$ both μ and π_i are divisible by $\mathfrak{p}_i^{h_i}$. At all other places the needed inclusion follows from integrality relative to the different. This gives the stated affine sets. They are invariant under U, and the balanced shape is the product of the $g_{\pi_i}^{N_i}$. The model includes nonempty fine points after excluding the deleted sets. For example, the fine torus grid has unit parameters by the

Chinese remainder theorem, and the affine cosets in the normalized coordinates are independent of N_i.

As in Theorem G.1, the strictly positive local quadratic bound gives uniform Taylor comparison for bounded radii. On the independent depth boxes the Fourier means of the shape distribution are products of geometric-series means. They converge to the Fourier coefficients of Haar on the closed group generated by the g_{π_i}. For the uniform average over $q \le X$, the depth tuples occupy the lattice points in a weighted simplex with side lengths comparable to $\log X$. Translation by a fixed integer vector changes only $O((\log X)^{s-1})$ points among $\asymp (\log X)^s$. At a character nontrivial on a generator, translating the simplex by the corresponding coordinate vector shows that its Fourier mean tends to zero. This gives the same Haar law. The group has a Lie direction with nonzero log projection because it contains the first generator's orbit closure. Thus the same shell proof applies. The error tends uniformly to zero off a set of depth tuples of asymptotically zero proportion, since for any fixed bound on q, the proportion of such tuples tends to zero.

For the assertion about choices, first fix a single split prime and allow all its conjugate prime ideals. The resulting shape subgroup is $\mathrm{Gal}(K/\mathbb{Q})$-stable. Its identity component has full angular tangent space in the negative eigenspace of complex conjugation. Indeed the angular projection is already full for a single generator, and conjugation acts by inversion on arguments and trivially on logarithms. We show that this identity component also contains the full logarithmic tangent space.

Write K^+ for the maximal real subfield. The rational unit-log module for $\mathrm{Gal}(K^+/\mathbb{Q})$ is the augmentation submodule of the cyclic regular representation. To see this, its real span is the zero-sum embeddings hyperplane with the permutation action, and the rational representation of a cyclic group is determined by that real character. Consequently each nontrivial simple rational factor occurs once. The tangent space of a connected logarithmic subtorus is rational relative to the unit-log lattice, so if the logarithmic projection of the identity component is proper, it omits a rational factor. Project onto that factor by its rational idempotent and clear denominators to obtain an integral group-ring element E. As an endomorphism of the log torus it maps the whole closed subgroup to a finite subgroup.

For the single ideal $\mathfrak{p}$ and its generator π, put $\gamma = (\pi\bar{\pi})^r/p^h \in K^+$. Its logarithmic vector is $-k$ times the logarithmic vector of g_π. The valuations of γ above p in K^+ form the vector $h(r\delta - \mathbf{1})$, where δ is the coordinate vector of the prime below $\mathfrak{p}$. Application of E gives a nonzero vector because a coordinate vector in the cyclic regular representation projects nontrivially onto every factor. On the other hand, torsion of E applied to the log class of g_π would give a positive power of the E-product of the conjugates of γ with the same logarithms as a unit. Since all embeddings are real and the product is positive, this would make it a unit after squaring, contradicting the valuations. This proves the full logarithmic projection and hence full generation by the conjugate choices for any one split prime.

Now begin with one ideal above the first prescribed prime. Its cyclic closure has identity component of dimension at least r. If the group generated so far is proper, take a nontrivial torus character vanishing on its identity component. Above the next prescribed prime, at least one conjugate choice gives a nontorsion value of that character. Otherwise all the finitely many values would generate a finite subgroup, contrary to full generation by the conjugate choices. The new generator increases the dimension of the joint identity component by at least one. After at most r primes it has dimension $2r - 1$, so the subgroup is G. Adding any remaining prescribed primes preserves this property. □

Multiplicity. Each parameter count, exact or modeled, is a multiple of d. For exact periods the free action is $a \mapsto \rho a$. In the model, multiplication by ζ_d permutes the fiber labels cyclically and

carries γ to $\zeta_d\gamma$. If x' is the new lift then $\zeta_d x - x' \in \mathfrak{D}^{-1}$, which proves preservation of the fine coset. For a deleted coset, the same difference is multiplied by π_i and belongs to the corresponding coarse ideal. The cyclic permutation gives $Q_{x'}(\zeta_d y) = Q_x(y)$, and the action is free because $\gamma \neq 0$. The laws and the mean formula therefore also apply to counts of parameter orbits after division by d.

G.7 Exact first radii

For a mixed fiber let $R_x = \min_{g\in F} |Q_x(g)|$. This minimum is the same for every fiber and equals

$$R_x = 2\pi^2 dk \cos(\pi/\ell). \tag{49}$$

Indeed discrete Parseval in all d trace coordinates gives $\sum_j L_j^*(y)^2 = 2d\sum_i |y_i|^2$, by expanding the trace in the k primitive embedding frequencies. Rotate the normal to the real axis. Equation (46) gives

$$\operatorname{Re} Q_x(g) = 4\pi^2 d\cos(\pi/\ell)\sum_i |g_i|^2 \;\;\geq\;\; 4\pi^2 dr\cos(\pi/\ell) \qquad (g \in F)$$

in the rotated coordinates, by the arithmetic–geometric mean inequality. Equality in the bound for the real part holds throughout the phase torus. Within this torus the points $\sigma(\zeta_d^a)$ have trace coordinates with support on the single block $j \equiv -a \pmod m$, by the Ramanujan sum. Choose two such points whose block phases lie on opposite sides of the outward normal. The imaginary parts of the rotated quadratic form have opposite signs there. A path in the phase torus joins the points and gives an imaginary part of zero somewhere, proving (49).

For any prime-model class C represented by $\mathfrak{b}$, $d\gamma \in \mathfrak{b}^{-1}$ for every fine candidate by (45) and $d\mathfrak{D}^{-1} = \mathfrak{f}$. Thus $(d\gamma)\mathfrak{b}$ is a nonzero integral ideal. Equations (49) and (47) give a lower bound for the first radius,

$$\lambda_0 = 2\pi^2 \frac{k}{d}\cos(\pi/\ell). \tag{50}$$

This bound is attained if and only if $\mathfrak{b}^{-1}$ has a generator among the $d\gamma$ of some fiber. In the principal ray class with $\ell \nmid j_0$ the bound is attained for every adjacent pair. To prove this with $\mathfrak{b} = \mathcal{O}_K$, encode the two phases as $c, c+1$ (integers may be reduced modulo ℓ), replacing j_0 by $m - j_0$ if their labels are reversed, and assign consecutive higher-phase blocks $0, \dots, j_0 - 1$. Write

$$\xi = \zeta_d^{-1}, \qquad S = 1 + \xi + \cdots + \xi^{m-1}, \qquad B = 1 + \xi + \cdots + \xi^{j_0-1}.$$

The residue of $-dx$ is $\omega(cS + B)$ modulo $\mathfrak{f}$. Both B and $v = 1 + \xi S$ are global units, being ratios of differences of primitive roots $(1 - \xi^a)/(1 - \xi)$ with $a = j_0, m + 1$ respectively. In the residue ring modulo $\mathfrak{f}$ we have $(\xi - 1)S = 0$, $BS = j_0 S$ and $S^2 = mS = 0$. Thus $Bv^u \equiv B + uj_0 S$ for any integer u. Take $uj_0 \equiv c \pmod \ell$ to obtain a unit in the desired residue class after multiplication by ω. This supplies a fine candidate of minimal norm.

Corollary G.5. *At $d = 9$ the prime model is the same at every depth and its first possible radius is $2\pi^2/3$ at every mixed corner. The full-torus models of Theorem G.4 also have this first radius at $d = 9$.*

Proof. For $K = \mathbb{Q}(\zeta_9)$ the discriminant has absolute value 3^9, so the Minkowski bound for the class group is less than 5. The prime 2 is inert, since it has order 6 modulo 9, and the unique ideal of norm 3 is principal. It follows that the ordinary class group is trivial. With $t = \zeta_9 - 1$, the ray residue ring is $\mathcal{O}_K/\mathfrak{f} = \mathbb{F}_3[t]/(t^3)$. Its units lift to global units: the units ζ_9 and $-(1 + \zeta_9)$ map to $1 + t$ and $1 - t$, with product $1 - t^2$. These generate the nine units with constant term 1, and -1

supplies the other sign. Here $1+\zeta_9$ is a unit by $\Phi_9(-1)=1$. Hence the ray class group is also trivial. As $m=3$, the two mixed choices $j_0=1,2$ are both prime to $\ell=3$. This gives the claimed prime-model radius from (50).

For the model over several fixed primes take $\mathfrak{b}=\prod_i \mathfrak{p}_i^{t_i}$. Its ray class is trivial, so before deletion its full-Haar model is the principal model and the first radius has a candidate with $(d\gamma)\mathfrak{b}=\mathcal{O}_K$. Any deleted γ would have $d\gamma$ divisible locally by $\mathfrak{p}_i\mathfrak{b}^{-1}$ for some i. The term $d\pi_i x$ is divisible by $\mathfrak{p}_i$ there, and the other terms of the deleted coset lie in $\mathfrak{p}_i\mathfrak{b}^{-1}\mathfrak{f}$, which gives the claimed divisibility. This would contradict $(d\gamma)\mathfrak{b}=\mathcal{O}_K$, so the minimal candidate survives. In fact each deleted candidate has $\mathrm{N}((d\gamma)\mathfrak{b})\geq p_i$ for some i, so by (49) the order-nine CRT count laws agree with the prime model at all radii below $(2\pi^2/3)(\min_i p_i)^{2/k}$. □

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, independent checks of the arithmetic reductions, and manuscript drafting and verification. No datasets are needed for the results.

H The first count above a cyclotomic corner threshold

For cyclotomic periods of odd prime-power order $d=\ell^b$ with $b\geq 2$, we determine the first nonzero count law near a mixed corner formed by one exceptional block. At the critical scale $q^{-2/\varphi(d)}$ there is a fixed initial gap. Just above it the limiting count is either zero or d, with probability of the second outcome asymptotic to an explicit regulator and planar-walk density constant times the relative excess radius to the power $\varphi(d)/4$. The prime-window probability has a ray-class weight depending on the common depth. This gives genuine depth dependence at every order under consideration except nine. At order twenty-five the two depth regimes differ by an exact factor of five.

Keywords: cyclotomic periods, ray class groups, shrinking targets, planar random walks.
MSC 2020: 11L05, 11R18, 11R44, 60G50.

H.1 The onset law

Let ℓ be an odd prime, $b\geq 2$, $d=\ell^b$, $m=\ell^{b-1}$, $g=(\ell-1)/2$, $r=gm$ and $k=2r=\varphi(d)$. Write $K=\mathbb{Q}(\zeta_d)$, $\omega=\zeta_d^m$ and $\mathfrak{f}=(1-\omega)\mathcal{O}_K$. Let $\mathcal{C}$ be the ray class group modulo $\mathfrak{f}$, and let $\mathcal{C}^n=\{B^n : B\in\mathcal{C}\}$. For a split prime $p\equiv 1 \pmod d$, $q=p^n$ and an order-d root ρ modulo q, use the period

$$T_{p,n}(a)=\sum_{j=0}^{d-1} e^{2\pi i a\rho^j/q}, \qquad a\in(\mathbb{Z}/q\mathbb{Z})^\times.$$

Let ν_-,ν_+ be two adjacent ℓth roots of unity. Fix the target $z=\ell((m-1)\nu_-+\nu_+)$. The same results hold interchanging the two adjacent phases. The target is a mixed corner of the sum of m hypocycloids described by Garcia–Hyde–Lutz [54, Corollary 1] and Untrau [25, Proposition 5.4]. Let $N_{p,n}(\lambda)$ count the unit parameters whose trace differs from z by at most $\lambda q^{-2/k}$.

For the constants, set

$$A=2\pi^2\frac{k}{d}\cos(\pi/\ell), \qquad F=\{y\in(\mathbb{C}^\times)^r : \prod_i |y_i|=1\}.$$

Use one complex embedding from each conjugate pair. Let $U_\mathfrak{f}$ be the group of units congruent to one modulo $\mathfrak{f}$, and let $R_{\log}$ be the Euclidean covolume of its logarithms in the zero-sum hyperplane

of $\mathbb{R}^r$. Let V be the sum of the squared lengths of g independent planar walks, each made of m independent unit steps with uniform directions, and write f_V for its density. Define

$$C_d = \frac{\mathrm{vol}(B_r)}{R_{\log}} \left(\frac{r}{2}\right)^{(r-1)/2} \frac{\sqrt{2}}{\tan(\pi/\ell)} \frac{gm^2}{2} f_V\left(\frac{gm^2}{2}\right), \tag{51}$$

where B_r is the Euclidean unit ball in $\mathbb{R}^r$.

Theorem H.1. *In the prime-window limit for the microscopic counts, uniformly over the finitely many depth classes, the limiting law at $\lambda = A(1+\delta)$ is supported on $\{0, d\}$ for sufficiently small positive δ. Its probability of d is*

$$\frac{C_d}{|\mathcal{C}^n|}\, \delta^{k/4}(1+o(1)) \qquad (\delta \downarrow 0), \tag{52}$$

where $C_d > 0$. This is also the limiting probability that the scaled nearest trace lies within the disk.

Fix $B > 0$. If $\delta = \delta(X) \downarrow 0$ with $\delta \geq (\log X)^{-B}$, then the proportion of split primes $p \in [X, 2X]$ for which $N_{p,n}(A(1+\delta)) = d$ is asymptotic to (52), uniformly for a common $1 \leq n \leq \exp(c_B \log X/\log\log X)$. For all sufficiently large X the exact counts at these radii are in $\{0, d\}$, so the same asymptotic holds for an arithmetic nearest trace within the disk.

At $d = 9$, $\mathcal{C}$ is trivial and $C_9 = \sqrt{2}\pi f_V(9/2)/R_{K^+}$, where R_{K^+} is the standard regulator of the real cubic subfield. In this case the absolute remainder in (52) is $O(\delta^{5/2})$.

At every other order the near-threshold profile distinguishes at least two depth classes. The proof in Section H.4 gives the exact distinction at $d = 25$.

H.2 Norm-one candidates and exact multiplicity

We use the arithmetic comparison of Section G. We recall its model and give the local reduction needed here. Write $\mathfrak{D}$ for the different of K. A fiber is specified by its exceptional block location $a_* \in \{0, \dots, m-1\}$. For the corresponding assignment of aligned phases ω^{c_a}, let

$$dx = -\omega \sum_{a=0}^{m-1} c_a \zeta_d^{-a}, \qquad Q_x(y) = 2\pi^2 \sum_{a=0}^{m-1} \omega^{c_a} \sum_{v=0}^{\ell-1} \mathrm{Tr}(y\zeta_d^{a+vm})^2.$$

The trace is extended real-linearly on the archimedean embeddings. These formulas give an angular lift x and the quadratic corner deficit, and $d\mathfrak{D}^{-1} = \mathfrak{f}$. For a ray class C represented by an integral ideal $\mathfrak{b}$ prime to $\mathfrak{f}$, the prime model draws h with Haar measure on $F/\sigma(U_\mathfrak{f})$. It counts one point for each fiber x and $\gamma \in \mathfrak{b}^{-1}\mathfrak{D}^{-1} - x$ with

$$|Q_x((\mathrm{N}\,\mathfrak{b})^{1/k} h\sigma(\gamma))| \leq \lambda.$$

The model is unchanged in law by changing the representative. In the prime window the mixture is uniform on $\mathcal{C}^n$. The comparison with the arithmetic count has error $O((\log X)^{-L})$ at each bounded radius with any fixed $L > 0$, for a common depth in the corresponding range. The error can be taken uniform over radii in a bounded interval, as follows from the uniform analytic shell collar and finite candidate set in the proof of Section G.

We check where a point can first occur. Discrete Parseval gives $\sum_{j=0}^{d-1} \mathrm{Tr}(y\zeta_d^j)^2 = 2d\sum_i |y_i|^2$. Projection of Q_x onto the supporting normal is therefore $4\pi^2 d\cos(\pi/\ell)\sum_i |y_i|^2$. By the arithmetic–geometric mean inequality, the model radius of a candidate γ is at least $A(\mathrm{N}((d\gamma)\mathfrak{b}))^{2/k}$. Here $(d\gamma)\mathfrak{b}$ is a nonzero integral ideal. The bound is attainable in the unrestricted shape variable when

its norm is one. At equal embedding moduli, trace energy can be confined to either of the two phase types by taking a suitable root-of-unity vector. Interpolating in the phase torus then makes the imaginary projection vanish, while the normal projection stays constant.

For each singleton fiber, the residue of $-dx$ modulo $\mathfrak{f}$ is the residue of a global unit. To see this, encode the majority and exceptional phases as $\omega^c, \omega^{c+\epsilon}$ with $\epsilon = \pm 1$. With $\xi = \zeta_d^{-1}$ and $S = \sum_{a=0}^{m-1} \xi^a$, this residue is $\omega(cS + \epsilon\xi^{a_*})$. The global unit $v = 1 + \xi S = (1 - \xi^{m+1})/(1 - \xi)$ satisfies $v \equiv 1 + S$ modulo $\mathfrak{f}$, since $(\xi - 1)S = 0$ in that ring. Also $S^2 = mS = 0$ there. Thus $\omega\epsilon\xi^{a_*}v^u$ has the desired residue when $u\epsilon \equiv c \pmod{\ell}$. In the principal class, the norm-one candidates are consequently exactly one $U_{\mathfrak{f}}$ orbit of numerators $d\gamma$ per fiber. In any other class a norm-one candidate would generate $\mathfrak{b}^{-1}$, and division by a global unit in the required residue class would give a ray generator for $\mathfrak{b}^{-1}$. Hence all nonprincipal models have no points below $2^{2/k}A$.

The torsion subgroup of $U_{\mathfrak{f}}$ is $\langle\omega\rangle$, of order ℓ. In fact the roots of unity in K are $\pm\zeta_d^a$. The condition $\zeta_d^a \equiv 1 \pmod{\mathfrak{f}}$ holds exactly if m divides a, by the valuations at $1 - \zeta_d$, and a negative sign fails already modulo $1 - \zeta_d$. Multiplication by ω permutes trace coordinates within each block, leaving Q_x fixed.

Choose $\delta > 0$ small enough that only norm-one candidates can occur. Near the arithmetic–geometric mean minimum the balanced logarithmic moduli have Euclidean size $O(\sqrt{\delta})$. This follows from the positive definite Hessian at the minimum on the zero-sum hyperplane and properness on that hyperplane. The log unit lattice is discrete, so for each fiber at most one logarithmic translate of its unit orbit lies in this neighborhood for each h, for sufficiently small δ. It contributes either zero or ℓ candidates because its ℓ torsion variants give the same value. Multiplication by ζ_d acts on all the affine candidates and sends the exceptional location cyclically to the next location, preserving the radius. Thus the total is either zero or $d = m\ell$.

Let $P_0(\delta)$ be the probability of d in the principal-class model. Use on F the product measure of Euclidean logarithmic Lebesgue measure and probability Haar measure on the phase torus. For one fixed fiber, the unit orbit can be unfolded against Haar on $F/\sigma(U_{\mathfrak{f}})$. Its mean contribution is the volume of $|Q_x(y/d)| \le A(1 + \delta)$ divided by $R_{\log}/\ell$, since the log fundamental domain of volume $R_{\log}$ has an ℓ-fold angular quotient. The balanced phase of the representative numerator only translates the variable y. Multiplying by the m fibers gives the mean of the whole count, and division by d gives

$$P_0(\delta) = \frac{1}{R_{\log}} \operatorname{vol}_F\{y : |Q_x(y/d)| \le A(1 + \delta)\}. \tag{53}$$

H.3 The volume just above the minimum

Write $y_i = e^{x_i + i\theta_i}$ in F, with $\sum_i x_i = 0$. The symbol x_i here denotes a real logarithmic coordinate, while the subscript on Q_x still indicates the angular fiber. Use embedding representatives with exponents $c + \ell t$ for $1 \le c \le g$ and $0 \le t < m$. If $E_a(y) = \sum_v \operatorname{Tr}(y\zeta_d^{a+vm})^2$ is the energy of block a, the discrete Fourier transform in v gives

$$E_a(y) = 2\ell \sum_{c=1}^{g} \left| \sum_{t=0}^{m-1} y_{c,t} \zeta_d^{a(c+\ell t)} \right|^2, \qquad E_{\mathrm{all}}(y) = 2d \sum_i |y_i|^2. \tag{54}$$

In fact the sum in the trace separates into frequencies c mod ℓ, and conjugate frequencies give the same squared coefficient. Rotating the two adjacent phases around their outward normal, set

$$S(x) = \frac{1}{r} \sum_i e^{2x_i}, \qquad \eta = \frac{2E_{a_*}(y)}{E_{\mathrm{all}}(y)} - 1.$$

The absolute radius takes the exact form

$$\frac{|Q_x(y/d)|}{A} = S(x)\sqrt{1+\tan^2(\pi/\ell)\eta^2}. \tag{55}$$

At log coordinate $x = 0$, (54) gives the law $\eta = 2V/(gm^2) - 1$ under independent uniform phases. At nearby x, it gives independent walks in the g groups with positive step lengths e^{x_i}, followed by the deterministic normalization $2/(m\sum_i e^{2x_i})$ of their summed squared lengths.

The conditional density $f_x(u)$ of η is jointly continuous in (x,u) near $(0,0)$ and $f_0(0) > 0$. For $m \geq 5$, an individual planar walk with step lengths close to one has characteristic function $\prod_t J_0(e^{x_{c,t}}|v|)$ for $v \in \mathbb{R}^2$. The estimate $J_0(s) = O((1+s)^{-1/2})$ makes this product integrable in the plane, uniformly in x. Fourier inversion gives a jointly continuous planar density with uniform compact spatial support. If a rotationally invariant planar random vector has density $D(z)$, its squared length has density $\pi D(\sqrt{u})$ for $u > 0$, with the radial density on the right interpreted along any ray. These squared-length densities are bounded with uniform compact support. Convolution across groups and the normalization therefore give the stated joint continuity near the positive target. This also follows directly by dominated convergence, since the possible change of formula at the boundary $u = 0$ affects only a null set in each convolution. For positivity, within a walk choose one step at angle zero and the remaining steps in $\pm\theta$ pairs. The resultant can have length $m/\sqrt{2}$ with nonzero derivative in θ, since $1 + (m-1)\cos\theta$ takes that value in the interior of its range. A neighborhood of that configuration contributes positive density near $m^2/2$. Convolution proves positivity of f_V at $gm^2/2$.

For $m = 3$ necessarily $d = 9$ and $g = 1$. The only critical values of the squared length at equal radii are $0, 1, 9$. At nonzero resultant a stationary phase configuration requires each step to be parallel or antiparallel to the resultant. Thus the target $9/2$ is a regular value. This persists in a neighborhood when the lengths vary by compactness. Local coarea in phase coordinates gives a smooth density $f_x(u)$ for (x,u) near $(0,0)$ and the same positivity assertion. In all cases

$$f_0(0) = \frac{gm^2}{2} f_V(gm^2/2).$$

On the log plane, $S(x) = 1 + (2/r)\|x\|^2 + O(\|x\|^3)$. In the sublevel set of (55), both $\|x\|$ and $|\eta|$ are $O(\sqrt{\delta})$. Divide these coordinates by $\sqrt{\delta}$ and use dominated convergence with the density just established. The limiting ellipsoid is

$$\frac{2}{r}\|x\|^2 + \frac{\tan^2(\pi/\ell)}{2}\eta^2 \leq 1$$

in the space of dimension $r = (r-1)+1$. Its volume is $\mathrm{vol}(B_r)(r/2)^{(r-1)/2}\sqrt{2}/\tan(\pi/\ell)$. Substitution into (53) gives $P_0(\delta) = C_d\delta^{r/2}(1+o(1))$. Multiplying by the principal mixture weight $1/|\mathcal{C}^n|$ proves (52). The arithmetic statement follows by applying the uniform microscopic comparison with a log saving greater than $Bk/4$. That error is $o(\delta^{k/4})$. The exact count is a multiple of d by the free action $a \mapsto \rho a$. For every individual prime, the local Taylor comparison in Section G also bounds it from above by the undeleted class model at $A(1+\delta)+O(q^{-1/k})$, evaluated at that prime's shape. For small enough δ and large enough X this radius stays in the range where every class-model count is at most d. Hence the exact arithmetic count is also at most d, and the nearest hit is exactly the event of a count of d.

For $d = 9$, the smoother density gives an error $O(\delta^{5/2})$ as follows. By the Morse lemma, in smooth local coordinates on the log plane one has $S - 1 = \|X\|^2$. Put $Y = \mathrm{sgn}(\eta)\sqrt{\sqrt{1+3\eta^2}-1}$, a smooth coordinate near zero, and then $Z = \sqrt{S}Y$. The left side of (55) less one is exactly

$\|X\|^2 + Z^2$. In these coordinates the smooth density has a constant term, a linear term integrating to zero over the centered three-ball, and a quadratic remainder of integrated size $O(\delta^{5/2})$.

We record the regulator factor. All CM unit logarithms here are logarithms of units in the real subfield K^+. Indeed $u/\bar{u}$ is a root of unity by Kronecker's theorem. Since the ratio is one modulo $1-\zeta_d$, it lies in the odd-order group $\langle\zeta_d\rangle$ and is a square there. Removing a root-of-unity factor makes u real. Put $I = |\operatorname{im}(\mathcal{O}_K^\times \to (\mathcal{O}_K/\mathfrak{f})^\times)|$. The torsion reduction image has order $2m$, so the index on logarithms is $I/(2m)$. If R_{K^+} is the usual regulator obtained by deleting one real embedding coordinate,

$$R_{\log} = \sqrt{r}\, R_{K^+} \frac{I}{2m}. \tag{56}$$

For $d = 9$, $I = 18$ as verified below. Thus $R_{\log} = 3\sqrt{3}R_{K^+}$ and (51) gives the value of C_9 in Theorem H.1. It can also be written using the complete elliptic integral $\mathbf{K}$ of the first kind as

$$C_9 = \frac{\mathbf{K}\Big(\sqrt{(v+1)^3(3-v)/(16v)}\Big)}{\sqrt{2}\pi R_{K^+}\sqrt{v}}, \qquad v = 3/\sqrt{2}.$$

This follows from the formula for the three-step radial density in [58, Eq. (1.2)] and the change from length to squared length.

H.4 Depth dependence from the local unit quotient

Let $\mathfrak{l} = (1-\zeta_d)\mathcal{O}_K$. It is the unique prime above ℓ, and $\mathfrak{f} = \mathfrak{l}^m$. Put $s = (\zeta_d - 1)/(\zeta_d + 1)$ in the residue ring $\mathcal{R} = \mathcal{O}_K/\mathfrak{f}$. This ring is isomorphic to $\mathbb{F}_\ell[s]/(s^m)$: it has characteristic ℓ, uniformizer s and residue field $\mathbb{F}_\ell$, and the two rings have the same size. Conjugation sends s to $-s$. The invariant units are exactly the even polynomials with nonzero constant. The homomorphism $a \mapsto a/\bar{a}$ on $\mathcal{R}^\times$ therefore has image of order $\ell^{(m-1)/2}$. On reductions of global units its image is exactly the group generated by the reduction of ζ_d, of order m. Indeed all global ratios are roots of unity of odd order as above, and $\zeta_d/\overline{\zeta_d} = \zeta_d^2$ generates the same residue group.

For an explicit description, the local ratio group is the product of cyclic groups generated by

$$\frac{1+s^j}{1-s^j}, \qquad 1 \le j < m, \quad j \text{ odd}, \quad \ell \nmid j.$$

The jth factor has order ℓ^{a_j}, where a_j is the least positive integer with $j\ell^{a_j} \ge m$. Independence follows by writing the exponents in base ℓ: the first nonzero digit in one factor gives a term of degree $j\ell^a$ with nonzero coefficient. In a nontrivial product the least of these initial degrees occurs only once. Then the number of factors with multiplicities accounts for the $(m-1)/2$ odd exponents below m, proving the product description by cardinality. The $j = 1$ factor is exactly the image of the global ratios. The quotient by that factor is a quotient of $\mathcal{R}^\times/\operatorname{im}(\mathcal{O}_K^\times)$, which in turn injects into $\mathcal{C}$ by the ray class exact sequence.

In particular $|\mathcal{C}|$ is divisible by $\ell^{(m-1)/2-(b-1)}$. The exponent is positive for the orders considered here except $d = 9$. Thus the near-onset probabilities at $n = 1$ and at a depth multiple of the ray class exponent are distinct. This proves genuine depth dependence, even on arbitrarily small right neighborhoods of A.

For $d = 9$, the ordinary class number is one by the Minkowski bound, which is less than five since the discriminant is 3^9. The prime 2 is inert, and the unique ideal of norm 3 is principal. The units ζ_9 and $-(1+\zeta_9)$ reduce modulo $\mathfrak{f}$ to $1+t$ and $1-t$, where $t = \zeta_9 - 1$ and $t^3 = 0$ in characteristic three. Their product $1-t^2$ and -1 show that reduction onto all eighteen units is surjective. This proves $I = 18$ and the triviality of $\mathcal{C}$.

For $d = 25$ the ordinary class number is again one by the classification of Masley and Montgomery [59]. The invariant local units have order 100. They are generated by reductions of real global units. In fact the real cyclotomic units $u_a = \sum_{j=-(a-1)/2}^{(a-1)/2} \zeta_d^j$, for $a = 3, 9$, give

$$u_3 = 3 + 4s^2 + 4s^4, qquadu_9 = 4 + 2s^4 \quad \text{in } \mathbb{F}_5[s]/(s^5).$$

These expansions follow from substituting $\zeta_d = (1 + s)/(1 - s)$. They are units globally because each is a root of unity times $(1 - \zeta_d^a)/(1 - \zeta_d)$. Dividing by their constant terms and using the truncated logarithm in $s^2\mathbb{F}_5[s]/(s^5)$ gives the independent vectors $(3, 1)$ and $(0, 3)$ in degrees two and four. Fourth powers of the original units therefore generate the principal invariant units, while the constant term 3 generates $\mathbb{F}_5^\times$. On the ratio group the image of global units has order five. Their full image in $\mathcal{R}^\times$ has order $100 \cdot 5 = 500$, while $\mathcal{R}^\times$ has order 2500. Consequently $\mathcal{C}$ is cyclic of order five. At $d = 25$ the weight $1/|\mathcal{C}^n|$ is exactly $1/5$ for $5 \nmid n$ and one for $5 \mid n$. The ratio of limiting probabilities at every sufficiently small fixed positive δ is therefore exactly five. In finite prime windows it is asymptotic to five in the regime of Theorem H.1.

H.5 A fixed set of split primes

The same onset law applies to the fixed-prime-set models of Section G when their shape closure is the full torus. The period uses a chosen diagonal cyclic subgroup of order d modulo the product q of powers of the split primes. Fix one ideal $\mathfrak{p}_i$ per prime, its ray order h_i, and the residue depths $t_i \pmod{h_i}$. The normalized affine model before deleting nonunits is the model of the ideal $\mathfrak{b} = \prod_i \mathfrak{p}_i^{t_i}$ above. A deleted candidate would have $(d\gamma)\mathfrak{b}$ divisible by some $\mathfrak{p}_i$ and hence could not have norm one. Thus the near-onset law is the principal profile exactly when $[\mathfrak{b}] = 1$, and is identically zero otherwise. These statements concern fixed relative excess radii in the limit over depth boxes or $q \le X$.

If all residue depths are averaged, their weights are asymptotically uniform. This is immediate for independent depth boxes. For $q \le X$ the tuples of exponents are lattice points in a weighted simplex of side lengths proportional to $\log X$, and all the finitely many residue shifts have the same leading volume. The class product is then uniform in the subgroup generated by the $[\mathfrak{p}_i]$. The factor $1/|\mathcal{C}^n|$ in the prime-window formula is replaced by the reciprocal order of this subgroup.

Statements and disclosures

OpenAI Codex assisted with mathematical exploration, internal independent reviews, and drafting. No datasets are needed for the proofs.

Statements and Declarations

Funding. This research was self-funded.

Competing interests. The author owns Luminary Oddities and has no other relevant competing interests.

AI assistance. OpenAI Codex assisted with derivations, proof development, source checks and independent internal review of each part. Anthropic Claude edited, consolidated and cross-referenced the parts and checked computations. Scott Schubr directed the research and is responsible for the content.

References

[1] S. Schubr, A Wasserstein rate for Kloosterman paths, with sharp small-ball estimates, Luminary Research Archive, edition 6 (2026). `https://research.luminaryoddities.org/small-ball-estimates-holder-centers-kloosterman-process`.

[2] S. Schubr, *A Wasserstein rate for Kloosterman paths, with sharp small-ball estimates*, Luminary Oddities Research Archive, edition 6, November 21, 2026.

[3] D. Kyselov, *Space-filling curves and small-ball probabilities for the Kloosterman process*, arXiv:2610.22645v1.

[4] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. **289** (1985), 679–706, doi:10.1090/S0002-9947-1985-0784009-0.

[5] D. Cai, X. Chen, W. Shu, H. Wang, and G. Zou, *Revisit the partial coloring method: Prefix Spencer and sampling*, arXiv:2408.13756.

[6] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publ. Math. Inst. Hautes Études Sci. **81** (1995), 73–205.

[7] E. Kowalski and W. Sawin, *Kloosterman paths and the shape of exponential sums*, Compos. Math. **152** (2016), 1489–1516.

[8] E. Kowalski and W. Sawin, *On the support of the Kloosterman paths*, arXiv:1709.05192.

[9] E. Kowalski, P. Michel, and W. Sawin, *Bilinear forms with Kloosterman sums and applications*, Ann. of Math. **186** (2017), 413–500.

[10] G. Ricotta and E. Royer, *Kloosterman paths of prime powers moduli*, Comment. Math. Helv. **93** (2018), 493–532.

[11] G. Ricotta, E. Royer, and I. E. Shparlinski, *Kloosterman paths of prime powers moduli, II*, Bull. Soc. Math. France **148** (2020), 173–188.

[12] D. Milićević and S. Zhang, *Distribution of Kloosterman paths to high prime power moduli*, Trans. Amer. Math. Soc. Ser. B **10** (2023), 636–669, doi:10.1090/btran/98.

[13] D. A. Burgess, *On character sums and L-series. II*, Proc. Lond. Math. Soc. **13** (1963), 524–536, doi:10.1112/plms/s3-13.1.524.

[14] D. Milićević, C. Robinson, and C. Shupe, *Sums of products of Kloosterman sums to prime power moduli*, Anal. Math. **52** (2026), 1135–1173, doi:10.1007/s10476-026-00197-7.

[15] E. Kowalski and T. Untrau, *Wasserstein metrics and quantitative equidistribution of exponential sums over finite fields*, author version.

[16] H. J. Brascamp and E. H. Lieb, *On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation*, J. Funct. Anal. **22** (1976), 366–389, doi:10.1016/0022-1236(76)90004-5.

[17] J. Cheeger, M. Gromov and M. Taylor, *Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds*, J. Differential Geom. **17** (1982), 15–53, doi:10.4310/jdg/1214436699.

[18] J. Carruth, N. Elkies, F. Gonçalves and M. Kelly, *The Beurling–Selberg box minorant problem via linear programming bounds*, arXiv:1702.04579v2.

[19] D. Kohel and Y.-D. Shieh, *On Sato–Tate distributions, extremal traces, and real multiplication in genus 2*, arXiv:2012.10805, Theorem 1.

[20] N. Kaiser, *Mean eigenvalues for simple, simply connected, compact Lie groups*, arXiv:math-ph/0609082.

[21] R. Bryant, answer to *Boundary of the image of a compact manifold in the complex plane*, MathOverflow, May 4, 2017.

[22] M. Coste, *An introduction to semialgebraic geometry*, RAAG Network School notes, 2002.

[23] R. A. Smith, *On n-dimensional Kloosterman sums*, J. Number Theory **11** (1979), 324–343, doi:10.1016/0022-314X(79)90006-4.

[24] R. Dbrowski and B. Fisher, *A stationary phase formula for exponential sums over $\mathbb{Z}/p^m\mathbb{Z}$ and applications to GL(3)-Kloosterman sums*, Acta Arith. **80** (1997), 1–48, doi:10.4064/aa-80-1-1-48.

[25] T. Untrau, *Equidistribution of exponential sums indexed by a subgroup of fixed cardinality*, Math. Proc. Cambridge Philos. Soc. **176** (2024), 65–94, doi:10.1017/S0305004123000361.

[26] T. Coulhon and A. Sikora, *Gaussian heat kernel upper bounds via the Phragmén–Lindelöf theorem*, Proc. Lond. Math. Soc. **96** (2008), 507–544, arXiv:math/0609429.

[27] W. Rudin, *Some theorems on Fourier coefficients*, Proc. Amer. Math. Soc. **10** (1959), 855–859, doi:10.1090/S0002-9939-1959-0116184-5.

[28] B. Doerr and A. Srivastav, *Multicolour discrepancies*, Combin. Probab. Comput. **12** (2003), 365–399, doi:10.1017/S0963548303005662.

[29] B. Borda and J.-C. Cuenin, *Smoothing inequalities for transport metrics in compact spaces*, arXiv:2510.21380.

[30] D. Milievi and S. Zhang, *Distribution of Kloosterman paths to high prime power moduli*, Trans. Amer. Math. Soc. Ser. B **10** (2023), 636–669, doi:10.1090/btran/98.

[31] M. A. Lifshits, *On the lower tail probabilities of some random series*, Ann. Probab. **25** (1997), 424–442, doi:10.1214/aop/1024404294.

[32] A. Weil, *On some exponential sums*, Proc. Natl. Acad. Sci. USA **34** (1948), 204–207, doi:10.1073/pnas.34.5.204.

[33] J. D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. **12** (1985), 183–216, doi:10.1090/S0273-0979-1985-15349-2.

[34] E. Carneiro and F. Littmann, *Entire approximations for a class of truncated and odd functions*, J. Fourier Anal. Appl. **19** (2013), 967–996, doi:10.1007/s00041-013-9273-6.

[35] S. W. Graham and C. J. Ringrose, *Lower bounds for least quadratic non-residues*, in *Analytic Number Theory*, Progr. Math. **85**, Birkhäuser, 1990, 269–309, doi:10.1007/978-1-4612-3464-7_18.

[36] D. E. Evans and M. Pugh, *Spectral measures and generating series for nimrep graphs in subfactor theory*, arXiv:0906.4314.

[37] É. Fouvry and P. Michel, *À la recherche de petites sommes d'exponentielles*, Ann. Inst. Fourier **52** (2002), 47–80, doi:10.5802/aif.1876.

[38] O. N. Feldheim and S. Sodin, *One more proof of the Erdős–Turán inequality, and an error estimate in Wigner's law*, Contemp. Math. **545** (2011), 69–75, arXiv:0901.1620.

[39] L. Fu, Y.-K. Lau, and P. Xi, *A quantitative Deligne's equidistribution theorem*, Monatsh. Math. **209** (2026), 263–289, doi:10.1007/s00605-025-02148-0.

[40] N. Kaiser, *Mean eigenvalues for simple, simply connected, compact Lie groups*, J. Phys. A **39** (2006), 15287–15298, arXiv:math-ph/0609082.

[41] A. M. Rahimi, *Wasserstein rates of all orders for the Sato–Tate law of Kloosterman sums*, arXiv:2610.16169v2.

[42] W. Duke, S. R. Garcia and B. Lutz, *The graphic nature of Gaussian periods*, arXiv:1212.6825.

[43] S. R. Garcia, T. Hyde and B. Lutz, *Gauss' hidden menagerie: from cyclotomy to supercharacters*, arXiv:1501.07507.

[44] J. Van Order, *Dirichlet twists of* GL_n*-automorphic L-functions and hyper-Kloosterman Dirichlet series*, Ann. Fac. Sci. Toulouse Math., doi:10.5802/afst.1687.

[45] D. E. Evans and M. Pugh, *Spectral measures and generating series for nimrep graphs in subfactor theory II:* SU(3), arXiv:1002.2348.

[46] J. W. S. Cassels, *An Introduction to the Geometry of Numbers*, Grundlehren Math. Wiss. **99**, Springer, 1959.

[47] W. Duke, S. R. Garcia and B. Lutz, *The graphic nature of Gaussian periods*, Proc. Amer. Math. Soc. **143** (2015), 1849–1863, arXiv:1212.6825.

[48] E. M. Matveev, *An explicit lower bound for a homogeneous rational linear form in the logarithms of algebraic numbers. II*, Izv. Math. **64** (2000), 1217–1269, doi:10.1070/IM2000v064n06ABEH000314.

[49] L. Ghidelli, *Arbitrarily long gaps between the values of positive-definite cubic and biquadratic diagonal forms*, J. Lond. Math. Soc. **102** (2020), 43–68, arXiv:1910.05070.

[50] R. Dąbrowski and B. Fisher, *A stationary phase formula for exponential sums over* $\mathbb{Z}/p^m\mathbb{Z}$ *and applications to GL(3)-Kloosterman sums*, Acta Arith. **80** (1997), 1–48, doi:10.4064/aa-80-1-1-48.

[51] H. P. Schlickewei, *On linear forms with algebraic coefficients and diophantine equations*, J. Number Theory **9** (1977), 370–380, doi:10.1016/0022-314X(77)90073-7.

[52] J.-H. Evertse, *The p-adic Subspace Theorem*, lecture notes, 2011, author version.

[53] V. Beresnevich and S. Velani, *A Mass Transference Principle and the Duffin–Schaeffer conjecture for Hausdorff measures*, Ann. of Math. **164** (2006), 971–992, doi:10.4007/annals.2006.164.971.

[54] S. R. Garcia, T. Hyde and B. Lutz, *Gauss's hidden menagerie: From cyclotomy to supercharacters*, Notices Amer. Math. Soc. **62** (2015), 878–888, doi:10.1090/noti1269.

[55] L. Ghidelli, *Arbitrarily long gaps between the values of positive-definite cubic and biquadratic diagonal forms*, J. London Math. Soc. **102** (2020), 43–68, doi:10.1112/jlms.12312.

[56] T. Khale, C. O'Kuhn, A. Panidapu, A. Sun and S. Zhang, *A Bombieri–Vinogradov theorem for primes in short intervals and small sectors*, J. Number Theory **236** (2022), 267–293, doi:10.1016/j.jnt.2021.04.004.

[57] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Pure and Applied Mathematics, Wiley-Interscience, New York, 1974.

[58] J. M. Borwein, A. Straub, J. Wan and W. Zudilin, *Densities of short uniform random walks*, Canad. J. Math. **64** (2012), 961–990, doi:10.4153/CJM-2011-079-2.

[59] J. M. Masley and H. L. Montgomery, *Cyclotomic fields with unique factorization*, J. Reine Angew. Math. **286/287** (1976), 248–256, doi:10.1515/crll.1976.286-287.248.