A Wasserstein rate for Kloosterman paths, with sharp small-ball estimates

Scott Schubr

Luminary Oddities

Abstract

Kowalski and Untrau asked for a quantitative Wasserstein-metric version of the convergence, due to Kowalski and Sawin, of Kloosterman paths in $C([0,1])$. We give one. For the two-parameter prime ensemble, the law μ_p of the polygonal Kloosterman path and the law μ of the Kloosterman process satisfy $W_1(\mu_p,\mu) = O(\sqrt{\log\log p/\log p})$ in the uniform metric, with explicit constants, while every path law supported on at most $(p-1)^2$ points lies at distance at least a constant times $(\log p)^{-1/2}$ from μ. The rate is therefore optimal up to the factor $\sqrt{\log\log p}$. The lower bound comes from a two-sided small-ball theorem for the limit process: around every admissible half-Hölder center, including boundary coefficients and endpoints, the uniform small-ball probability has negative logarithm of exact order ϵ^{-2}, and the upper estimate holds at every continuous center. The upper rate combines the Kloosterman sheaf, with all representation-degree dependence explicit, growing moments of dyadic Fourier blocks, and an unbiased tensor Bernstein approximation of path functionals. A direct polynomial transfer of rare events also shows that, for all sufficiently large primes (with an effective threshold), uniformly over a fixed Hölder class of centers, a proportion at least $\frac{1}{2}\exp[-c_L\sqrt{\log p}]$ of all Kloosterman paths lies within uniform distance $D_L(\log p)^{-1/4}$ of the center.

Keywords: Kloosterman paths; Kloosterman sums; Wasserstein distance; quantitative equidistribution; small-ball probabilities; random Fourier series.
Mathematics Subject Classification (2020): Primary 11T23; Secondary 60B10, 60G17, 42A61.

1 Introduction

Kowalski and Sawin [1] proved that the polygonal paths interpolating the partial sums of Kloosterman sums modulo a prime p, with both parameters uniformly distributed in $\mathbb{F}_p^\times$, converge in law in the Banach space $C([0,1])$ to a random Fourier series K with independent Sato–Tate coefficients. In their work on Wasserstein metrics and quantitative equidistribution of exponential sums, Kowalski and Untrau single out this theorem and ask for "a quantitative version of such a statement in terms of Wasserstein metrics" [4, Section 1.3(4)]. This paper gives such a version, together with a lower bound that determines the rate up to a factor $\sqrt{\log\log p}$.

Let μ_p be the law of the prime path $K_p(\cdot;A,B)$ defined in (6.1), and let μ be the law of K. For the Wasserstein distance W_1 with the uniform cost on $C([0,1])$, Theorem 6.1 and Corollary 6.2 prove, for every odd prime p and integers $2 \le M \le N < p/2$,

$$\frac{\sqrt{c}}{2\sqrt{3\log p}} \le W_1(\mu_p,\mu) \le D_B(p,M,N), \qquad W_1(\mu_p,\mu) = O\Big(\sqrt{\frac{\log\log p}{\log p}}\Big), \tag{1.1}$$

where D_B is the explicit finite bound (6.3) and c is the absolute constant of Theorem 1.1. For example $W_1(\mu_p, \mu) \le 20000\sqrt{\log\log p/\log p}$ once $\log p \ge 10^6$. The lower bound holds for every path law with at most $(p-1)^2$ atoms.

The upper bound uses the classical normalized Kloosterman sheaf, with trace, Euler–Poincaré and weight estimates in which every dependence on the representation degree is retained [14, 15, 13]. Growing moments control the dyadic Fourier blocks of the arithmetic path. An unbiased tensor Bernstein approximation, with degrees allocated according to the squared Fourier weights, then compares Lipschitz path functionals: conditional binomial concentration pays the approximation error, and a bounded-polynomial character expansion controls the arithmetic comparison.

The lower bound in (1.1) is a consequence of a two-sided small-ball theorem for the limit process, proved first and of independent interest. Around every admissible half-Hölder center the negative logarithm of the uniform small-ball probability has exact order ϵ^{-2} (Theorem 1.1). Its upper half holds at every continuous center, and a union bound over the atoms of μ_p converts it into the support-size bound. Transferred directly to the arithmetic paths by exponentially accurate polynomial sandwiches, the lower half shows that, for all sufficiently large primes, a proportion at least $\frac{1}{2}\exp[-c_L\sqrt{\log p}]$ of all pairs (A, B) have paths within $D_L(\log p)^{-1/4}$ of any prescribed center in a fixed Hölder class (Theorem 6.5). This radius improves the radius $O_L((\log\log p)^{-1/2})$ obtained from the Wasserstein bound and a Lipschitz bump (Corollary 6.7).

The current author manuscript of Kowalski–Untrau [4] records that this Wasserstein question is currently studied by S. Pitte. Rahimi [5] proves Wasserstein rates of all orders for the Sato–Tate law of the normalized Kloosterman sums, that is, for the endpoint $K_p(1; A, B)$; the estimates here concern the full path in the uniform metric.

1.1 The Kloosterman process and the small-ball theorem

Let $(X_h)_{h\in\mathbb{Z}}$ be independent real random variables of density

$$f(x) = \frac{\sqrt{4-x^2}}{2\pi}\,\mathbf{1}_{[-2,2]}(x). \tag{1.2}$$

Set $v_0(t) = t$, and, for $h \ne 0$, set

$$v_h(t) = \frac{\mathrm{e}^{2\pi iht} - 1}{2\pi ih}, \qquad 0 \le t \le 1.$$

The Kloosterman process is

$$K(t) = \sum_{h\in\mathbb{Z}} X_h v_h(t), \tag{1.3}$$

where the nonzero frequencies are summed symmetrically. Almost sure convergence with a continuous limit is [1, Theorem 1.1]; uniform convergence is proved in Section 2. We use $\|w\|_\infty = \sup_{0\le t\le 1}|w(t)|$ for complex-valued paths, and $\widehat{u}(h) = \int_0^1 u(t)\mathrm{e}^{-2\pi iht}\,dt$ for Fourier coefficients.

For $L \ge 1$, let $\mathcal{H}_L$ be the class of continuous $g : [0,1] \to \mathbb{C}$ with the following properties:

1. $g(0) = 0$ and $a = g(1)$ belongs to the real interval $[-2, 2]$;
2. $|g(t) - g(s)| \le L\sqrt{|t-s|}$ for all $s, t \in [0,1]$;
3. with $u(t) = g(t) - at$, the numbers

$$\beta_h = 2\pi ih\widehat{u}(h), \qquad h \in \mathbb{Z}\setminus\{0\}, \tag{1.4}$$

are real and belong to $[-2, 2]$.

The endpoint and Fourier conditions are the established support characterization of Kowalski–Sawin [2, Theorem 1.1]; we add a uniform Hölder bound.

Theorem 1.1. *For every $L \geq 1$, every $g \in \mathcal{H}_L$, and $0 < \epsilon \leq 1/4$,*

$$\mathbb{P}(\|K - g\|_\infty < \epsilon) \geq \exp\{-26000(L+2)^2\epsilon^{-2}\}. \tag{1.5}$$

Let

$$\varpi = 1 - \pi^{-1}, \qquad \delta = \frac{\varpi}{128\pi^2}, \qquad \epsilon_0 = \sqrt{\delta/2}, \qquad c = \frac{\varpi\delta}{16}. \tag{1.6}$$

Then, for every deterministic $g \in C([0,1], \mathbb{C})$,

$$\mathbb{P}(\|K - g\|_\infty < \epsilon) \leq \exp\{-c\epsilon^{-2}\}, \qquad 0 < \epsilon \leq \epsilon_0. \tag{1.7}$$

Thus, for fixed L, the negative logarithm of the probability has order ϵ^{-2}, uniformly on $\mathcal{H}_L$. Both $a = \pm 2$ and individual $\beta_h = \pm 2$ are allowed. The upper bound needs none of the endpoint, Fourier or Hölder restrictions imposed on the lower bound.

The closest quantitative predecessor is Kyselov [3, Theorem 1.2], whose version 1 gives

$$\mathbb{P}(\|K - g\|_\infty < \epsilon) \geq \exp\{-C_L\epsilon^{-2}\log^2(2/\epsilon)\}$$

for closed centers $g(0) = g(1) = 0$ with $\beta_h \in [-1, 1]$ and the same Hölder exponent. The lower estimate (1.5) removes both logarithmic factors, extends the coefficient interval to its full permitted range, and allows every permitted endpoint. Estimate (1.7) shows that the power ϵ^{-2} cannot be improved uniformly, or even at an individual admissible center, to a smaller logarithmic order.

Our contribution is specific to this bounded-coefficient Fourier process and this path norm. Shifted small-deviation estimates for non-Gaussian series form an established subject; for example, the p-exponential framework in [6] uses unbounded coefficient laws. The classical comparison ingredients below are instances of Prékopa marginal log-concavity [7], product peakedness associated with Kanter [8, 9], and Gaussian translation comparison [10]. The point at which coefficient boundaries matter is Section 4: the Fejér weights supply different interior margins at different frequencies, and their product retains a factorial.

Other non-Gaussian small-deviation results use different norms. Uniform-series priors in [11, Section 4.2] use a coordinate-supremum sequence topology rather than the uniform norm of this Fourier image. The non-Gaussian Fourier estimates in [12, Section 2.6.4] concern L^2 small deviations; an L^2 lower bound does not by itself give a uniform-norm lower bound.

1.2 Organization and companion results

Section 2 proves convergence of the series and a centered comparison principle. Sections 3, 4 and 5 prove Theorem 1.1. Section 6 proves the Wasserstein rate, its support-size lower bound and the shrinking-target counts. A companion paper [16] uses Theorem 1.1 and the graded Fejér margins of Section 4 to determine the worst-case small-ball scale $\epsilon^{-\max(2,1/\gamma)}$ for every Hölder exponent γ, together with the roughness–density thresholds for other bounded symmetric unimodal coefficient laws.

2 Convergence and a centered comparison

Lemma 2.1. *If the independent coefficients W_h are either all distributed as (1.2) or all distributed as $N(0,4)$, the symmetric partial sums $\sum_{|h|\leq H} W_h v_h$ converge almost surely uniformly on $[0,1]$, and in expected uniform norm. The limit is continuous. A finite set of coefficients is independent of the continuous series formed from its complement.*

Proof. For integers $n \geq 1$ and $n \leq H \leq 2n$, put

$$P_{n,H}(t) = \sum_{n<|h|\leq H} \frac{W_h \mathrm{e}^{2\pi i h t}}{2\pi i h}.$$

A centered variable in $[-2,2]$ satisfies $\mathbb{E}\mathrm{e}^{\lambda X} \leq \cosh(2\lambda) \leq \mathrm{e}^{2\lambda^2}$: convexity bounds $\mathrm{e}^{\lambda x}$ by its endpoint chord, whose expectation is $\cosh(2\lambda)$, and $\log\cosh y \leq y^2/2$ follows by integrating $|\tanh y| \leq |y|$. This is the elementary bounded-variable exponential estimate. The last bound is an equality for $N(0,4)$. Independence and $\sum_{n<|h|\leq 2n} h^{-2} \leq C/n$ therefore show that both real and imaginary parts of $P_{n,H}(t)$ have subgaussian tails of variance parameter at most C/n. Exponential Markov inequality, followed by a union over the two parts, gives

$$\mathbb{P}(|P_{n,H}(t)| > v) \leq 4\mathrm{e}^{-c_1 n v^2} \tag{2.1}$$

with absolute constants, uniformly in t, H.

For completeness, a deterministic grid may be chosen with only polynomially many points. If P is a trigonometric polynomial of degree d, each Fourier coefficient is at most $\|P\|_\infty$ in absolute value, whence

$$\|P'\|_\infty \leq 2\pi \sum_{|k|\leq d} |k| \, \|P\|_\infty = 2\pi d(d+1) \, \|P\|_\infty .$$

A grid of spacing at most $[4\pi d(d+1)]^{-1}$ gives $\|P\|_\infty \leq 2 \max_{t \text{ on grid}} |P(t)|$, by applying the derivative bound between a maximizer and a grid point. For $d = 2n$ this uses $O(n^2)$ points. Combining (2.1) for these points and all H gives

$$\mathbb{P}\left(\max_{n\leq H\leq 2n} \|P_{n,H}\|_\infty > 2v \right) \leq \min(1, C_1 n^3 \mathrm{e}^{-c_1 n v^2}).$$

Integrating this inequality in v yields

$$\mathbb{E} \max_{n\leq H\leq 2n} \|P_{n,H}\|_\infty \leq C_2 \sqrt{\frac{\log(2n)}{n}}.$$

Subtracting $P_{n,H}(0)$ changes the bound by at most a factor two and produces the blocks of $\sum W_h v_h$. On $n = 1, 2, 4, \ldots$, the expected maximal block norms have a finite sum. Tonelli's theorem gives almost sure summability of these maxima, hence convergence of complete and incomplete symmetric blocks. The same estimates give convergence in expected norm. The finite low-frequency terms cause no difficulty. Measurability as a uniform limit and independence of disjoint coefficient families give the last assertion. □

Write $Z \succeq Y$ for symmetric scalar variables when $\mathbb{P}(|Z| \leq r) \geq \mathbb{P}(|Y| \leq r)$ for all $r \geq 0$. This is the usual central peakedness order in one dimension.

Lemma 2.2 (Classical peakedness comparison in the required form)**.** *Let $Z_1, \ldots, Z_m$ be independent symmetric variables with log-concave densities, and let Y_j be independent $N(0,4)$ variables. If $Z_j \succeq Y_j$ for each j, then for every real linear map $T : \mathbb{R}^m \to \mathbb{R}^d$ and convex symmetric Borel set $A \subseteq \mathbb{R}^d$,*

$$\mathbb{P}(TZ \in A) \geq \mathbb{P}(TY \in A).$$

The conclusion applies to the inverse image of a closed uniform-norm ball under a finite continuous-function map.

Proof. Replace one coordinate at a time. For the coordinate being replaced, the conditional acceptance function is

$$F(a) = \int \mathbf{1}_A \left(aT(e_j) + \sum_{k \neq j} z_k T(e_k) \right) \prod_{k \neq j} f_k(z_k)\, dz.$$

The integrand is jointly log-concave: its indicator is that of a convex inverse image, with the extended convention for zero values. Prékopa's marginal theorem [7, Theorem 6] makes F log-concave. Reflection of all coordinates makes F even. An even log-concave function is nonincreasing in absolute value, since every $a \in [-b, b]$ is a convex combination of b and $-b$. Its strict superlevel sets are central intervals. Scalar peakedness and layer cake therefore give $\mathbb{E}F(Z_j) \geq \mathbb{E}F(Y_j)$. Each intermediate product still has symmetric log-concave factors, so iteration proves the result. No full-rank assumption on T is used. For a function map, its closed-ball inverse image in coefficient space is convex and symmetric, so the same proof applies directly. □

If X has density (1.2), its central mass on $[0, 2]$ is concave in the radius and joins 0 to 1. Thus

$$\mathbb{P}(|X| \leq r) \geq r/2 \quad (0 \leq r \leq 2).$$

For $Y \sim N(0, 4)$, the density bound gives $\mathbb{P}(|Y| \leq r) \leq r/\sqrt{2\pi} \leq r/2$. For $r \geq 2$ the first probability is one. Also

$$\frac{d^2}{dx^2} \log f(x) = -\frac{4 + x^2}{(4 - x^2)^2} < 0 \quad (|x| < 2),$$

so X has the symmetry, log-concavity and peakedness required in Lemma 2.2. A uniform variable on $[-r_j, r_j]$, $0 < r_j \leq 1/2$, has the same properties and dominates Y: for $0 \leq s \leq r_j$ its central mass is $s/r_j \geq 2s$, and beyond r_j it is one. All these laws have densities, so open or closed central intervals give the same probabilities.

Lemma 2.3 (The bounded symmetric unimodal case)**.** *Let $Z_1, \ldots, Z_m$ be independent variables whose densities are even, bounded, supported on $[-2, 2]$, and nonincreasing on $[0, 2)$. Let Y_j be independent $N(0, 4)$ variables. For every real linear map T and convex symmetric Borel set A,*

$$\mathbb{P}(TZ \in A) \geq \mathbb{P}(TY \in A).$$

The conclusion also applies to the inverse image of a closed uniform-norm ball under a finite continuous-function map.

Proof. This is a special case of classical unimodal product peakedness [9, Theorem 3.6]; the following mixture argument records the hypotheses needed here. Choose the right-continuous monotone representative of one density f on $[0, 2)$ and set $f(2) = 0$. Let μ be its positive Stieltjes measure on $(0, 2]$, so that $\mu((x, 2]) = f(x)$ for $0 \leq x < 2$. Put $\nu(ds) = 2s\mu(ds)$. Tonelli's theorem gives

$$\nu((0, 2]) = 2 \int_0^2 \mu((x, 2])\, dx = 2 \int_0^2 f(x)\, dx = 1.$$

Furthermore,

$$f(x) = \int_{(0,2]} \frac{\mathbf{1}_{\{|x| < s\}}}{2s}\, \nu(ds) \quad \text{for almost every } x.$$

Thus each coordinate is a scale mixture of centered uniforms of half-width at most two. Choose its mixing scale independently of the other coordinates. Conditional on all scales, the variables are independent, symmetric and log-concave. A uniform variable on $[-s, s]$, $0 < s \leq 2$, is more peaked than $N(0, 4)$: its central mass is at least $r/2$ for $0 \leq r \leq 2$, with value one once $r \geq s$, whereas the Gaussian central mass is at most $r/\sqrt{2\pi}$. Lemma 2.2 therefore gives the claimed inequality conditionally on every scale vector. Integrating over that vector proves the result. The finite path-ball assertion follows from its convex symmetric inverse image exactly as before. □

3 A Gaussian reference and uniform tail control

Let Y_h be independent $N(0,4)$ variables and put $G = \sum_h Y_h v_h$, using Lemma 2.1.

Lemma 3.1. *For $0 < r \le 1$,*

$$\mathbb{P}(\|G\|_\infty \le r) \ge \exp(-64/r^2). \tag{3.1}$$

For every finite $I \subset \mathbb{Z}$, with $G_I = \sum_{h\in I} Y_h v_h$,

$$\mathbb{P}(\|G_I\|_\infty \le r) \ge \mathbb{P}(\|G\|_\infty \le r). \tag{3.2}$$

Consequently, for every $M \ge 0$ and $R_M = \sum_{|h|>M} X_h v_h$,

$$\mathbb{P}(\|R_M\|_\infty \le r) \ge \exp(-256/r^2), \qquad 0 < r \le 1. \tag{3.3}$$

Proof. For $h \ge 1$ put $U_h = (Y_h + Y_{-h})/\sqrt{2}$ and $V_h = (Y_h - Y_{-h})/\sqrt{2}$. These are mutually independent $N(0,4)$ variables, independent of Y_0. Direct expansion gives

$$\operatorname{Re} G(t) = tY_0 + \sum_{h\ge1} \frac{U_h \sin(2\pi ht)}{\sqrt{2}\pi h},$$
$$\operatorname{Im} G(t) = \sum_{h\ge1} \frac{V_h(1-\cos(2\pi ht))}{\sqrt{2}\pi h}.$$

Consider the continuous centered Gaussian process

$$B(t) = \sum_{h\ge1} \frac{U_h \sin(2\pi ht) + V_h(\cos(2\pi ht)-1)}{\sqrt{2}\pi h}.$$

The preceding convergence proof also applies to its partial sums. Its covariance is

$$\frac{2}{\pi^2} \sum_{h\ge1} \frac{\cos(2\pi h(t-s)) - \cos(2\pi hs) - \cos(2\pi ht) + 1}{h^2}.$$

The identity

$$\sum_{h\ge1} \frac{\cos(2\pi hx)}{h^2} = \pi^2(x^2 - x + 1/6), \qquad 0 \le x \le 1, \tag{3.4}$$

follows by computing the Fourier coefficients of the displayed quadratic; the absolutely convergent series represents that continuous periodic function. For $0 \le s \le t \le 1$, substitution gives $\operatorname{Cov}(B(s), B(t)) = 4s(1-t)$. Thus B is a variance-four Brownian bridge, since a centered Gaussian law is determined by its covariances. Its sine and cosine-minus-one parts are respectively $(B(t) - B(1-t))/2$ and $(B(t) + B(1-t))/2$, each of norm at most $\|B\|_\infty$. Hence

$$\|G\|_\infty \le |Y_0| + 2\,\|B\|_\infty . \tag{3.5}$$

Here is an explicit confinement estimate, with its normalization recorded. Variance-four Brownian motion has generator $2\,d^2/dx^2$ and free density $1/\sqrt{8\pi}$ at $(1;0,0)$. The Dirichlet eigenfunction expansion on $(-b,b)$ gives its killed density there as

$$p_D(1;0,0) = \frac{1}{b} \sum_{j\ge0} \exp\left\{ -\frac{(2j+1)^2\pi^2}{2b^2} \right\}.$$

Indeed the odd indexed eigenfunctions have squared value $1/b$ at zero, and the other ones vanish there. Dividing by the free endpoint density is the probability that the bridge stays in $(-b, b)$. Keeping the first positive term yields

$$\mathbb{P}(\|B\|_\infty < b) \geq \frac{\sqrt{8\pi}}{b} \exp\{-\pi^2/(2b^2)\} \geq \exp\{-\pi^2/(2b^2)\}, \qquad 0 < b \leq 1.$$

On $[-r/3, r/3]$, the $N(0,4)$ density is at least $\mathrm{e}^{-1/72}/(2\sqrt{2\pi})$ for $r \leq 1$. Independence, with $b = r/3$, and (3.5) give

$$\mathbb{P}(\|G\|_\infty \leq r) \geq c_0 r \exp\{-9\pi^2/(2r^2)\}, \qquad c_0 = \frac{\mathrm{e}^{-1/72}}{3\sqrt{2\pi}}.$$

Since $9\pi^2/2 < 45$, $\log(1/c_0) < 4$, and $\log(1/r) \leq r^{-2}$ on this range, (3.1) follows.

To prove (3.2), first evaluate paths on a finite time grid and identify complex values with a real vector. Its radius-r acceptance set A is a product of disks, hence convex and symmetric. If z is the finite Gaussian coefficient vector of G_I with evaluation map T, then

$$F(y) = \int \mathbf{1}_A(Tz + y)\phi(z)\, dz$$

is log-concave by Prékopa and even by reflection. Therefore $F(y) \leq F(0)$. Conditioning on the independent evaluation vector of $G - G_I$ proves that adding this complement cannot increase acceptance. Increasing the grid to a countable dense set proves the assertion for closed uniform balls, by path continuity and continuity of probability from above. This argument allows singular evaluation maps and an infinite independent complement; it is the required instance of the classical Gaussian translation comparison.

Finally let $S_{M,H} = \sum_{M<|h|\leq H} X_h v_h$. Lemma 2.2 and (3.2) show

$$\mathbb{P}(\|S_{M,H}\|_\infty \leq r') \geq \mathbb{P}(\|G\|_\infty \leq r').$$

For $r' < r$, almost sure uniform convergence implies

$$\limsup_{H\to\infty} \{\|S_{M,H}\|_\infty \leq r'\} \subseteq \{\|R_M\|_\infty \leq r\}$$

outside a null set. Reverse Fatou for bounded indicators gives the probability of this limsup at least the limsup of the probabilities. Take $r' = r/2$ and apply (3.1) to obtain (3.3). The radius slack avoids any assumption about atoms on a norm-sphere. □

4 Fejér approximation with graded margins

Fix $g \in \mathcal{H}_L$, $a = g(1)$ and $u = g - at$. The function u vanishes at both endpoints and has Hölder constant at most $L + 2$ on $[0, 1]$. As a periodic function, it satisfies

$$|u(t) - u(s)| \leq \sqrt{2}(L + 2)\sqrt{d(t - s, \mathbb{Z})}. \tag{4.1}$$

For a wrapped pair this follows by passing through the common zero endpoint and using $\sqrt{x}+\sqrt{y} \leq \sqrt{2(x + y)}$; otherwise use the ordinary Hölder estimate.

The degree-M Fejér kernel has integral one, is nonnegative, and obeys

$$F_M(t) \leq M + 1, \qquad F_M(t) \leq \frac{1}{4(M + 1)d(t, \mathbb{Z})^2}.$$

These bounds follow from $F_M(t) = \sin^2((M+1)\pi t)/((M+1)\sin^2(\pi t))$ and $\sin(\pi d) \geq 2d$ for $0 \leq d \leq 1/2$. Splitting the integral at $1/(M+1)$ gives

$$\int_0^1 F_M(t)\sqrt{d(t,\mathbb{Z})}\,dt \leq \frac{7}{3\sqrt{M+1}}, \qquad M \geq 1. \tag{4.2}$$

In fact, the inner part is at most $4/(3\sqrt{M+1})$, and the outer part is at most $1/\sqrt{M+1}$, by direct integration of the two kernel bounds. If $\sigma_M u = F_M * u$, equations (4.1) and (4.2) imply

$$\|\sigma_M u - \sigma_M u(0) - u\|_\infty \leq \frac{14\sqrt{2}(L+2)}{3\sqrt{M+1}}.$$

For $0 < \epsilon \leq 1/4$ choose

$$M = \left\lceil 4096(L+2)^2\epsilon^{-2} \right\rceil. \tag{4.3}$$

The last approximation error is at most $\epsilon/8$. Set

$$b_0 = (1-\epsilon/8)a, \qquad x_h = \frac{|h|}{M+1}, \qquad b_h = (1-x_h)\beta_h \quad (0 < |h| \leq M). \tag{4.4}$$

The pinned polynomial

$$q_M(t) = b_0 t + \sigma_M u(t) - \sigma_M u(0) = \sum_{|h|\leq M} b_h v_h(t)$$

satisfies

$$\|q_M - g\|_\infty \leq \epsilon/4 + \epsilon/8 = 3\epsilon/8. \tag{4.5}$$

The equality follows directly from the Fourier formula for a Fejér mean and (1.4).

Lemma 4.1. *For $T_M z = \sum_{|h|\leq M} z_h v_h$ and the coefficients in (4.4),*

$$\mathbb{P}(\|T_M(X-b)\|_\infty \leq r) \geq c_e \epsilon^{3/2} \exp\{-6M - 64/r^2\}, \qquad 0 < r \leq 1, \tag{4.6}$$

where $c_e = \sqrt{3/8}/(8\pi) > \mathrm{e}^{-5}$.

Proof. For $0 < |h| \leq M$ put $r_h = \min(1/2, x_h)$ and let Z_h be independent uniform variables on $[-r_h, r_h]$. If $|z| \leq r_h$, then

$$|b_h + z| \leq 2 - 2x_h + r_h \leq 2 - x_h, \qquad f(b_h+z) \geq \frac{\sqrt{3x_h}}{2\pi}.$$

As $2r_h \geq x_h$, comparison of this shifted density with the uniform density pays a factor at least $c_* x_h^{3/2}$, where $c_* = \sqrt{3}/(2\pi)$. Multiplying over both signs of each frequency gives

$$c_*^{2M}\left(\frac{M!}{(M+1)^M}\right)^3 \geq \mathrm{e}^{-6M}. \tag{4.7}$$

Indeed,

$$\begin{aligned}\log\frac{M!}{(M+1)^M} &\geq \int_1^M \log t\,dt - M\log(M+1)\\ &= -M + 1 - M\log(1+1/M) \geq -M,\end{aligned}$$

and $3+2\log(2\pi/\sqrt{3})<6$. Retaining the differing margins x_h in this product is essential: replacing all of them by their smallest value would introduce an unnecessary $M\log M$ loss.

For $h=0$, take $r_0=\epsilon/8$ and Z_0 uniform on $[-r_0,r_0]$. The inequalities

$$|b_0+z|\le 2-\epsilon/8, \qquad f(b_0+z)\ge \frac{\sqrt{3\epsilon/8}}{2\pi}$$

hold on this interval, including when $a=\pm 2$. Its density factor is at least $c_e\epsilon^{3/2}$. Integrating the finite product density only on the box $\prod_{|h|\le M}[-r_h,r_h]$ gives

$$\mathbb{P}(\|T_M(X-b)\|_\infty\le r)\ge c_e\epsilon^{3/2}\mathrm{e}^{-6M}\mathbb{P}(\|T_MZ\|_\infty\le r).$$

All uniform widths are positive and at most $1/2$. Lemma 2.2, followed by (3.2) and (3.1), bounds the last probability below by e^{-64/r^2}. This proves (4.6). The shifts are handled by a finite density comparison, before the centered comparison is applied. □

5 Proof of the two-sided bounds

Proof of the lower bound in Theorem 1.1. The low-frequency event in Lemma 4.1 with $r=\epsilon/4$ has probability at least

$$c_e\epsilon^{3/2}\exp\{-6M-1024\epsilon^{-2}\}.$$

The independent high-frequency event $\|R_M\|_\infty\le\epsilon/4$ has probability at least $\exp(-4096\epsilon^{-2})$ by (3.3). Together with (4.5), these events imply

$$\|K-g\|_\infty\le\epsilon/4+\epsilon/4+3\epsilon/8=7\epsilon/8<\epsilon.$$

Their product gives the strict-ball lower bound $c_e\epsilon^{3/2}\exp\{-6M-5120\epsilon^{-2}\}$. For the allowed parameters,

$$M\le 4097(L+2)^2\epsilon^{-2}, \qquad \log\frac{1}{c_e\epsilon^{3/2}}\le 7\epsilon^{-2}.$$

Therefore its negative logarithm is at most

$$[24582(L+2)^2+5127]\epsilon^{-2}\le 26000(L+2)^2\epsilon^{-2},$$

which proves (1.5). □

Proof of the upper bound in Theorem 1.1. For an arbitrary $g\in C([0,1],\mathbb{C})$, set $a=g(1)$, $u=g-at$ and $\beta_h=2\pi ih\widehat{u}(h)$, which may now be complex. On the strict small-ball event, $K(1)=X_0$ gives $|X_0-a|<\epsilon$. Consequently

$$D(t)=K(t)-tX_0-u(t)=(K(t)-g(t))+t(a-X_0)$$

has norm less than 2ϵ. Uniform convergence justifies its nonzero Fourier coefficients $\widehat{D}(h)=(X_h-\beta_h)/(2\pi ih)$. Parseval, retaining only positive frequencies through N, gives

$$\sum_{h=1}^{N}|X_h-\beta_h|^2\le N^2\sum_{h=1}^{N}\frac{|X_h-\beta_h|^2}{h^2}\le 16\pi^2N^2\epsilon^2. \tag{5.1}$$

Since $f \le 1/\pi$, an interval of length one has probability at most $1/\pi$. The independent indicators $I_h = \mathbf{1}_{\{|X_h - \operatorname{Re}\beta_h| \ge 1/2\}}$ have success probabilities at least $\varpi = 1 - 1/\pi$. On the small-ball event,

$$\sum_{h=1}^{N} I_h \le 4 \sum_{h=1}^{N} |X_h - \operatorname{Re}\beta_h|^2 \le 4 \sum_{h=1}^{N} |X_h - \beta_h|^2 \le 64\pi^2 N^2 \epsilon^2.$$

With $N = \lfloor \delta \epsilon^{-2} \rfloor$ and $\delta = \varpi/(128\pi^2)$, the right-hand side is at most $\varpi N/2$. For clarity, exponential Markov inequality at $\lambda = \log 2$ yields

$$\begin{aligned} \mathbb{P}\left(\sum I_h \le \varpi N/2\right) &\le \exp\{\lambda \varpi N/2 + \varpi N(\mathrm{e}^{-\lambda} - 1)\} \\ &= \exp\{-\varpi N(1 - \log 2)/2\} \le \exp(-\varpi N/8). \end{aligned}$$

Here each success probability is at least ϖ, so $\mathbb{E}\mathrm{e}^{-\lambda I_h} \le \exp(\varpi(\mathrm{e}^{-\lambda} - 1))$. If $\epsilon^2 \le \delta/2$, then $N \ge \delta/(2\epsilon^2)$, giving (1.7) with the constants (1.6). □

6 The Wasserstein rate and shrinking targets

For an odd prime p and $(A, B) \in (\mathbb{F}_p^\times)^2$, let $K_p(\cdot; A, B)$ be the polygonal path with vertices

$$K_p\left(\frac{j}{p-1}; A, B\right) = p^{-1/2} \sum_{1 \le x \le j} \exp\left(\frac{2\pi i(Ax + B\overline{x})}{p}\right), \qquad 0 \le j \le p-1. \tag{6.1}$$

Here $\overline{x}$ is the inverse of x modulo p, and both parameters are chosen independently and uniformly nonzero. Write μ_p for this path law and $\mu = \mathrm{Law}(K)$. On probability measures with finite first uniform moment, W_1 is the Wasserstein distance for the cost $\|x - y\|_\infty$, equivalently the supremum of differences of integrals of real 1-Lipschitz functions.

The qualitative convergence $\mu_p \Rightarrow \mu$ is known [1, Theorem 1.5]. We quantify it for this two-parameter ensemble. The sheaf-theoretic machinery below is classical: in particular the rank, puncture and Swan trace estimate is the curve estimate explained in [13, Remark 4.2]. The proof records its dependence on growing representation degree. Growing-moment bounds on dyadic Fourier blocks then sharpen the path approximation, before an unbiased tensor Bernstein comparison of path functionals. Theorem 6.1 supplies the quantitative Wasserstein form of this convergence asked for by Kowalski–Untrau [4, Section 1.3(4)].

For $M \ge 1$ put

$$a_0 = 1, \qquad a_h = \frac{1}{\pi|h|} \ (0 < |h| \le M), \qquad S_M = \sum_{|h| \le M} a_h = 1 + \frac{2}{\pi} \sum_{h=1}^{M} \frac{1}{h}.$$

For integers $2 \le M \le N < p/2$, put $d = 2M + 1$, $r = \lceil \log(16M) \rceil$, and allocate degrees

$$n_0 = M, \qquad n_h = \left\lceil \frac{M^2}{h^2} \right\rceil \quad (0 < |h| \le M). \tag{6.2}$$

Define

$$\begin{aligned} D_B(p, M, N) = 1152\sqrt{r/M} + 2048N^{-1/4} + 384r(N^2/p)^{1/(4r)} \\ + p^{-1/2} + \frac{8M}{p} + \frac{8S_M(d+1)}{\sqrt{p}} \prod_{|h| \le M} (n_h + 2)^2. \end{aligned} \tag{6.3}$$

The last product is at most $\exp(16M)$.

Theorem 6.1. *For every odd prime p and integers $2 \le M \le N < p/2$,*

$$W_1(\mu_p, \mu) \le D_B(p, M, N).$$

In particular,

$$W_1(\mu_p, \mu) = O\left(\sqrt{\frac{\log\log p}{\log p}}\right). \tag{6.4}$$

A coarse explicit version is $W_1(\mu_p, \mu) \le 20000\sqrt{\log\log p/\log p}$ for $\log p \ge 10^6$.

Corollary 6.2 (Support-size obstruction)**.** *For any path law ν supported on at most $q \ge 1$ points,*

$$W_1(\mu, \nu) \ge \tfrac{1}{2} \min\left\{\epsilon_0, \sqrt{c/\log(2q)}\right\},$$

where c, ϵ_0 are the constants in (1.6). In particular for every odd prime p,

$$W_1(\mu_p, \mu) \ge \tfrac{1}{2} \min\left\{\epsilon_0, \sqrt{c/\log(2(p-1)^2)}\right\} \ge \frac{\sqrt{c}}{2\sqrt{3\log p}}.$$

Proof. Let S contain the at most q support points and put $r = \min\{\epsilon_0, \sqrt{c/\log(2q)}\}$. The universal-center upper bound of Theorem 1.1 and a union bound show that the union of their open radius-r balls has μ-mass at most $qe^{-c/r^2} \le 1/2$. Outside it the distance to S is at least r. Every coupling with ν therefore has mean cost at least $r/2$. The arithmetic law has at most $(p-1)^2$ atoms, regardless of collisions. Finally $\log(2(p-1)^2) \le 3\log p$ and $c/\epsilon_0^2 = (1-\pi^{-1})/8 < 1$ imply the last inequality for $p \ge 3$. □

Thus the upper rate in Theorem 6.1 and the support-size lower order agree up to the factor $\sqrt{\log\log p}$.

6.1 Path approximation with a uniform expected modulus

For $I \subseteq \mathbb{F}_p^\times$, $H = |I|$, set

$$T_I = p^{-1/2} \sum_{x \in I} \exp(2\pi i(Ax + B/x)/p).$$

Orthogonality over all $A, B \in \mathbb{F}_p$ gives

$$\sum_{A,B \in \mathbb{F}_p} |T_I|^4 = N_I,$$

where N_I counts quadruples with equal sums and equal inverse sums. If the common sum is nonzero, the inverse-sum condition forces equal products; each first ordered pair has at most two second pairs. If it is zero, both pairs are opposite and there are at most H^2 quadruples. Thus $N_I \le 3H^2$. Restricting this nonnegative sum to nonzero parameters yields

$$\mathbb{E}_{A,B \ne 0}|T_I|^4 \le \frac{3H^2}{(p-1)^2}. \tag{6.5}$$

This use of both parameters is essential to the present argument.

Define Q_p on the mesh j/p using the same partial sums through $j = p-1$, and make its last mesh segment constant. Then $K_p(t) = Q_p((p-1)t/p)$ and

$$\|K_p - Q_p\|_\infty \le p^{-1/2}, \tag{6.6}$$

because Q_p is deterministically $\sqrt{p}$-Lipschitz. For $\delta = |t-s| < p^{-1}$ this bound gives $|Q_p(t) - Q_p(s)|^4 \le \delta^2$. For $\delta \ge p^{-1}$, decompose an increment into its complete interior mesh intervals and at most two fractional endpoint increments. The interior set has at most $p\delta$ elements, and each endpoint increment is at most $p^{-1/2}$. Using (6.5), $p/(p-1) \le 3/2$ and $|z_1+z_2+z_3|^4 \le 27\sum |z_j|^4$ gives

$$\mathbb{E}|Q_p(t) - Q_p(s)|^4 \le 1024\delta^2. \tag{6.7}$$

The same estimate holds for K. Indeed

$$\sum_{h\in\mathbb{Z}} |v_h(t) - v_h(s)|^2 = \delta$$

by (3.4). The bounded-variable exponential estimate in Lemma 2.1, applied to the real and imaginary parts and then integrated, gives the displayed fourth-moment bound. Finite truncations converge in L^4 at fixed times, so the estimate passes to the series.

Here is the expected-modulus consequence with constants. For a continuous X satisfying (6.7) and $X(0) = 0$, let A_j be the largest adjacent dyadic increment of level j. The fourth moment and a union bound give $\mathbb{E}A_j \le 6\,2^{-j/4}$. If $\delta \le 2^{-m} < 2\delta$, dyadic chaining and continuity give

$$\omega_X(\delta) \le A_m + 2\sum_{j>m} A_j, \qquad \mathbb{E}\omega_X(\delta) \le 100\delta^{1/4}.$$

Also $\mathbb{E}|X(1)| < 6$. Pin $u_X(t) = X(t) - tX(1)$ and extend it periodically. For a wrapped pair, split at the common zero endpoint. The expected circular modulus is at most $208\delta^{1/4}$.

For the ordinary Fejér kernel, direct integration of its two bounds in Section 4 gives

$$\int_{-1/2}^{1/2} |t|^{1/4} F_M(t)\,dt \le 3(M+1)^{-1/4}.$$

With $\mathcal{P}_M X = tX(1) + F_M * u_X - (F_M * u_X)(0)$, it follows that

$$\mathbb{E}\,\|X - \mathcal{P}_M X\|_\infty \le 2048 M^{-1/4}, \qquad X = Q_p,\ K. \tag{6.8}$$

The dyadic sums also prove finite first uniform moments.

6.2 Arithmetic coefficients and character correlations

Put

$$W_h = p^{-1/2} \sum_{x\in\mathbb{F}_p^\times} \exp(2\pi i((A-h)x + B/x)/p) = \mathrm{Kl}_2(B(A-h); p).$$

These variables are real and $|W_h| \le 2$, including the singular value $W_h = -p^{-1/2}$ when $A = h$. Integrating each mesh derivative against $\mathrm{e}^{-2\pi i h t}$ shows that the normalized pinned coefficient of Q_p is $c_{p,h}W_h$, where

$$c_{p,h} = \mathrm{e}^{\pi i h/p}\frac{\sin(\pi h/p)}{\pi h/p}, \qquad |c_{p,h} - 1| \le 2\pi|h|/p \quad (0 < |h| < p/2).$$

The endpoint is W_0. On $[-2,2]^{2M+1}$ define

$$\mathcal{A}_M(x) = x_0 t + \sum_{0<|h|\le M} \left(1 - \frac{|h|}{M+1}\right) x_h v_h(t).$$

Then

$$\|\mathcal{P}_M Q_p - \mathcal{A}_M(W)\|_\infty \le 8M/p, \qquad \mathcal{P}_M K = \mathcal{A}_M((X_h)_{|h|\le M}). \tag{6.9}$$

For the first inequality each nonzero frequency contributes at most $2|c_{p,h} - 1|/(\pi|h|) \le 4/p$.

Lemma 6.3. *Define $\chi_k(2\cos\theta) = \sin((k+1)\theta)/\sin\theta$ for $k \geq 0$, using continuous endpoint values. Let h_i be distinct modulo an odd prime p, and let $k_i \geq 0$ not all vanish. Put $s = \#\{i : k_i > 0\}$ and $D = \prod_i (k_i + 1)$. For each fixed $B \neq 0$,*

$$\left| \mathbb{E}_{A \neq 0} \prod_i \chi_{k_i}(W_{h_i}) \right| \leq \frac{4(s+1)D}{\sqrt{p}}. \tag{6.10}$$

The corresponding expectation for independent Sato–Tate variables is zero. The trivial character has zero discrepancy.

Proof. The polynomial χ_k is the trace character of the kth symmetric power of the standard SL_2 representation. It has dimension $k+1$ and $|\chi_k(x)| \leq k+1$ on $[-2,2]$. Also

$$\mathbb{E}\chi_k(X) = \frac{2}{\pi} \int_0^\pi \sin((k+1)\theta) \sin\theta \, d\theta = \mathbf{1}_{\{k=0\}}.$$

We now specify the classical sheaf inputs and their quantitative use. The normalized rank-two Kloosterman sheaf is pure of weight zero, has trivial determinant, trace $-\mathrm{Kl}_2$, geometric and arithmetic monodromy SL_2, tame nontrivial unipotent ramification at zero, and sole infinity break $1/2$, of Swan conductor one; see [14, proof of Proposition 3.2]. The change $y = BA$ reduces the active factors to distinct translations $y \mapsto y - Bh_i$. Their direct sum has full product geometric monodromy SL_2^s by the Goursat–Kolchin–Ribet argument in [1, Lemma 2.4]. Distinct finite singularities exclude rank-one-twisted geometric isomorphisms: nontrivial unipotent local monodromy cannot become unramified by a scalar twist.

Tensor the symmetric powers with positive k_i. The resulting rank-D sheaf $\mathcal{F}$ on $U = \mathbb{A}^1 \setminus \{Bh_i : k_i > 0\}$ has no geometric invariants: each nontrivial symmetric power is an irreducible nontrivial representation of its own factor of the product. Irreducibility here is in characteristic zero; the usual raising and lowering operators connect its monomial basis. At finite punctures it is tame. Every infinity upper ramification group of index greater than $1/2$ acts trivially on every rank-two factor, hence on tensor powers, symmetric-power subquotients and their tensor product. Thus $\mathrm{Swan}_\infty(\mathcal{F}) \leq D/2$, independently of the degrees except for the explicit rank.

The trace/Euler/weight formulas [15, Theorems 4.1, 4.2, 4.6 and (4.2)], also summarized in the curve estimate [13, Remark 4.2], give

$$H_c^0 = H_c^2 = 0, \qquad \dim H_c^1 = D(s-1) + \mathrm{Swan}_\infty(\mathcal{F}) \leq Ds.$$

The sheaf is pure of weight zero, so the full U trace sum has absolute value at most $Ds\sqrt{p}$ by Deligne's theorem. Its trace is $(-1)^{\sum k_i} \prod_i \chi_{k_i}(W_{h_i})$; this sign does not affect the zero-main-term bound.

At most $s+1$ points are changed when passing from $U(\mathbb{F}_p)$ to $\mathbb{F}_p^\times$. Each product character there is at most D in absolute value, including at singular values $W_h = -p^{-1/2}$. Therefore

$$\left| \mathbb{E}_{A \neq 0} \prod_i \chi_{k_i}(W_{h_i}) \right| \leq D \, \frac{s\sqrt{p} + s + 1}{p-1} \leq \frac{4(s+1)D}{\sqrt{p}}.$$

The proof is uniform in B and the distinct shifts. This is an explicit use of the classical character trace bound, rather than an extrapolation of a fixed-degree implicit constant. □

6.3 Growing moments and a sharper path approximation

For distinct residues h, complex scalars b_h, and an integer $r \ge 1$, put $\sigma^2 = \sum_h |b_h|^2$ and $\ell = \sum_h |b_h|$. Then

$$\left(\mathbb{E} \left| \sum_h b_h W_h \right|^{2r} \right)^{1/(2r)} \le 8\sqrt{r}\,\sigma + 8\ell p^{-1/(4r)}. \tag{6.11}$$

Here and below the arithmetic expectation averages both nonzero parameters. To see the degree dependence explicitly, the character $x^m = \chi_1(x)^m$ decomposes as $\sum_k c_{m,k}\chi_k(x)$ with nonnegative integer multiplicities and $\sum_k c_{m,k}(k+1) = 2^m$. This follows by induction from $\chi_1\chi_k = \chi_{k+1} + \chi_{k-1}$ for $k \ge 1$ and $\chi_1\chi_0 = \chi_1$. Thus Lemma 6.3 bounds the discrepancy of any monomial of total degree $2r$ by $4(2r+1)2^{2r}p^{-1/2}$: the trivial character cancels, and every other term involves at most $2r$ active shifts. Expanding the absolute $2r$th power gives

$$\mathbb{E}|\sum_h b_h W_h|^{2r} \le \mathbb{E}|\sum_h b_h X_h|^{2r} + 4(2r+1)(2\ell)^{2r}p^{-1/2}.$$

For the independent variables, the bounded-variable exponential inequality gives $\mathbb{P}(|\sum b_h X_h| > u) \le 4\exp[-u^2/(16\sigma^2)]$. Integration bounds its $2r$th moment by $4(16\sigma^2)^r r!$. Taking roots, using $r! \le r^r$ and $2[4(2r+1)]^{1/(2r)} \le 8$, proves (6.11). The second term is absent for independent coefficients. This argument does not require a sheaf with all available shifts active simultaneously.

For a trigonometric polynomial T of degree at most H, the Bernstein inequality $\|T'\|_\infty \le 2\pi H\|T\|_\infty$ and the grid of $J = \lceil 4\pi H \rceil \le 14H$ equally spaced points give

$$\|T\|_\infty \le 2 \max_{0 \le j < J} |T(j/J)|.$$

For complex T, the Bernstein inequality follows from its real form by rotating the value of the derivative at a maximizing point and taking the real part. Applying (6.11) at the grid points, with no independence assumption on those points, yields

$$\mathbb{E}\|T - T(0)\|_\infty \le 32(14H)^{1/(2r)}(\sqrt{r}\,\sigma + \ell p^{-1/(4r)}), \quad T(t) = \sum_{|h| \le H} b_h W_h \mathrm{e}^{2\pi i h t}. \tag{6.12}$$

Write $w_J(h) = (1 - |h|/(J+1))_+$. The difference $\mathcal{P}_N Q_p - \mathcal{P}_M Q_p$ has zero endpoint term and coefficients

$$b_h = \frac{[w_N(h) - w_M(h)]c_{p,h}}{2\pi i h}, \qquad 0 < |h| \le N,$$

in front of $W_h(\mathrm{e}^{2\pi i h t} - 1)$. In particular $|c_{p,h}| \le 1$. Split the frequencies into $1 \le |h| \le M$ and the subsequent blocks $H_{j-1} < |h| \le \min(H_j, N)$, where $H_j = 2^j M$ and $1 \le j \le J_* := \lceil \log_2(N/M) \rceil$. The first block has $|b_h| \le 1/(2\pi M)$, hence $\ell \le 1/\pi < 1$ and $\sigma^2 \le 1/(2\pi^2 M) < 1/M$. Each later block has $|b_h| \le 1/(\pi H_j)$ and at most $2H_j$ coefficients, so $\ell < 1$ and $\sigma^2 < 1/H_j$ again. The last block has degree at most $H_{J_*} < 2N$, but its nonzero frequencies still lie in $[-N, N]$ and are distinct modulo p.

Set $r = \lceil \log(16M) \rceil \ge 4$. The first terms in (6.12) sum to at most

$$32\sqrt{r}\,(14M)^{1/(2r)} M^{-1/2} \sum_{j \ge 0} 2^{j(-1/2+1/(2r))} \le 512\sqrt{r/M},$$

since $(14M)^{1/(2r)} \le \sqrt{\mathrm{e}} < 2$ and the sum is at most $(1 - 2^{-3/8})^{-1} < 5$. The second terms sum to at most $384r(N^2/p)^{1/(4r)}$, using $(1 - 2^{-1/(2r)})^{-1} \le 6r$ and $28^{1/(2r)} < 2$. The former inequality follows

from $1-\mathrm{e}^{-x} \geq x/2$ for $0 \leq x \leq 1$, with $x = \log 2/(2r)$. Apply (6.8) only at the larger cutoff N to obtain

$$\mathbb{E}\|Q_p - \mathcal{P}_M Q_p\| \leq 512\sqrt{r/M} + 2048N^{-1/4} + 384r(N^2/p)^{1/(4r)}. \tag{6.13}$$

For K the block proof has no arithmetic error. Letting $N \to \infty$ and using the L^1 convergence in (6.8) gives

$$\mathbb{E}\|K - \mathcal{P}_M K\| \leq 512\sqrt{r/M}. \tag{6.14}$$

6.4 Unbiased Bernstein approximation of path functionals

Use the degrees in (6.2), and retain $d = 2M+1$ and $r = \lceil \log(16M) \rceil$. For $x \in [-2,2]^d$, conditionally choose independent binomial variables

$$B_h \sim \mathrm{Bin}\left(n_h, \frac{x_h+2}{4}\right), \quad Y_h = 4B_h/n_h - 2, \quad \Delta_h = Y_h - x_h.$$

For a real 1-Lipschitz path functional φ with $\varphi(0) = 0$, write $G(x) = \varphi(\mathcal{A}_M(x))$ and define its tensor Bernstein polynomial $P(x) = \mathbb{E}[G(Y) \mid x]$. This is a polynomial of degree at most n_h in coordinate h, and $|P(x)| \leq 2S_M$ on the cube.

Lemma 6.4 (Uniform functional approximation)**.** *For every such φ,*

$$\sup_{x \in [-2,2]^d} |P(x) - G(x)| \leq 64\sqrt{r/M}.$$

Proof. The usual binomial representation of Bernstein polynomials is used; we give the concentration argument to retain all dimension dependence. For a Bernoulli variable of mean q, the second derivative of its centered log moment-generating function is a Bernoulli variance and is at most $1/4$. Its value and derivative at zero vanish, so its log moment-generating function is at most $u^2/8$. Consequently, uniformly over every conditioning vector x,

$$\mathbb{E}[e^{u\Delta_h} \mid x] \leq e^{2u^2/n_h}, \qquad \mathbb{E}[\Delta_h \mid x] = 0.$$

This includes $x_h = \pm 2$, where the coordinate is deterministic.

Put $b_h = (1 - |h|/(M+1))/(2\pi i h)$ for $h \neq 0$ and $T(t) = \sum_{0<|h|\leq M} b_h \Delta_h e^{2\pi i h t}$. The variance proxy satisfies

$$\sigma^2 = \sum_{0<|h|\leq M} \frac{|b_h|^2}{n_h} \leq \frac{1}{2\pi^2 M} < \frac{1}{M}.$$

The real and imaginary exponential estimates therefore give

$$\Pr(|T(t)| > u \mid x) \leq 4e^{-u^2/(16\sigma^2)}.$$

Use the trigonometric Bernstein grid with $J = \lceil 4\pi M \rceil \leq 14M$, for which

$$\|T\|_\infty \leq 2 \max_{j<J} |T(j/J)|.$$

If $V = \max_{j<J} |T(j/J)|$, integrate its union tail at $a = 4\sigma\sqrt{\log(4J)}$ to obtain

$$\mathbb{E}[V \mid x] \leq 4\sigma\sqrt{\log(4J)} + \frac{2\sigma}{\sqrt{\log(4J)}} \leq 6\sigma\sqrt{\log(4J)}.$$

Indeed $\int_a^\infty e^{-u^2/(16\sigma^2)}du \le 8\sigma^2 a^{-1}e^{-a^2/(16\sigma^2)}$. Pinning adds at most a factor two, so $\mathbb{E}\|T - T(0)\|_\infty \le 24\sigma\sqrt{\log(56M)}$. Also $\mathbb{E}|\Delta_0| \le 2/\sqrt{M}$. Since $\log(56M) \le 2r$ and $r \ge 4$, these estimates imply

$$\mathbb{E}\|\mathcal{A}_M(Y - x)\|_\infty \le 64\sqrt{r/M}.$$

The Lipschitz property and Jensen's inequality for absolute values prove the asserted approximation bound, uniformly in x. No independence of arithmetic coefficients or of process increments is needed: independence here is only that of the auxiliary binomials under a fixed x. □

Lift $x_h = 2\cos\theta_h$. Since P is bounded by $2S_M$ on the cube, every complex Fourier coefficient of its even angular lift has modulus at most $2S_M$. Its degree in coordinate h is at most n_h. Define $C_j(2\cos\theta) = \cos(j\theta)$. The exact identities

$$C_0 = \chi_0, \qquad C_1 = \chi_1/2, \qquad C_j = (\chi_j - \chi_{j-2})/2 \quad (j \ge 2)$$

show that the absolute character mass weighted by representation dimension is at most $j + 1$ for C_j. Grouping Fourier sign orbits into cosine products costs no more than summing the original complex Fourier modes. It follows that

$$P(x) = \sum_{\mathbf{k}} b_{\mathbf{k}} \prod_h \chi_{k_h}(x_h), \qquad \sum_{\mathbf{k}} |b_{\mathbf{k}}| \prod_h (k_h + 1) \le 2S_M \prod_h (n_h^2 + 3n_h + 1) \le 2S_M \prod_h (n_h + 2)^2.$$

This coefficient estimate uses the bounded polynomial on the cube, not its generally large monomial coefficients. The latter are not inserted into the sheaf estimate.

The allocated degrees have only exponential total character cost:

$$\prod_h (n_h + 2)^2 \le 4M^2 \frac{4^{4M} M^{8M}}{(M!)^8} \le e^{16M}.$$

Indeed $n_0 + 2 \le 2M$ and $n_h + 2 \le 4M^2/h^2$ for $h \ne 0$. Using $\log M! \ge M\log M - M$ bounds the logarithm by $(8 + 4\log 4)M + 2\log M + \log 4 \le 16M$ for $M \ge 2$. For an integer-only coarser bound, $e < 3$ and $4M^2 \le 4^M$ give $\prod_h (n_h + 2)^2 \le 2^{23M}$, since $4 \cdot 6^8 < 2^{23}$.

Applying Lemma 6.3, with the trivial character discrepancy omitted as before, yields

$$|\mathbb{E}G(W) - \mathbb{E}G(X)| \le 128\sqrt{r/M} + \frac{8S_M(d+1)}{\sqrt{p}} \prod_h (n_h + 2)^2.$$

Adding (6.13), (6.14), (6.6) and (6.9), and taking the supremum over φ, gives

$$W_1(\mu_p, \mu) \le D_B(p, M, N),$$

which proves the finite part of Theorem 6.1. The previously established finite first moments justify Kantorovich–Rubinstein duality on the separable path space. All high-degree dependence remains in the explicit rank-weighted character bound of Lemma 6.3.

Rate optimization. For $t = \log p \ge 10^6$, choose

$$M = \lfloor t/128 \rfloor, \qquad N = \lfloor p^{1/8} \rfloor, \qquad r = \lceil \log(16M) \rceil.$$

Then $M \ge t/256$, $r \le \log t$, $2 \le M \le N < p/2$ and $16M \le t/8$. The principal term is at most $18432\sqrt{\log t/t}$. The other terms are bounded by

$$2500e^{-t/32} + 384\log t\, e^{-3t/(16\log t)} + 9te^{-t/2} + 8t^2e^{-3t/8} \le \sqrt{\log t/t}.$$

For the last inequality each term is at most one quarter of the right side at $t = 10^6$, and its ratio decreases thereafter. For the second ratio its logarithmic derivative is

$$\frac{1}{2t} + \frac{1}{2t\log t} - \frac{3(\log t - 1)}{16(\log t)^2} < 0;$$

the other three derivatives are immediately negative. Thus

$$W_1(\mu_p, \mu) \le 20000\sqrt{\frac{\log\log p}{\log p}} \quad (\log p \ge 10^6),$$

and in particular $W_1 = O(\sqrt{\log\log p/\log p})$.

This proves (6.4) with the stated coarse constant.

6.5 Direct transfer of rare path events

The preceding Wasserstein bound is useful for general Lipschitz functionals. For small balls whose probabilities are exponentially small in ϵ^{-2}, an additive Lipschitz approximation error becomes too large. We instead approximate an indicator by a bounded polynomial with exponentially small noise tails. The auxiliary noise is independent only conditionally on its coefficient vector; the arithmetic coefficients remain dependent.

Put

$$C_L = 26000(L+2)^2, \quad \Lambda_L = 64C_L, \quad C_0 = 2^{36}, \quad D_L = [1024C_0(\Lambda_L+1)]^{1/4}.$$

Theorem 6.5. *Fix $L \ge 1$. For every sufficiently large prime p, uniformly for $g \in \mathcal{H}_L$, with $\epsilon_p = D_L(\log p)^{-1/4}$,*

$$\mu_p(B(g, \epsilon_p)) \ge \tfrac{1}{2}\exp[-\Lambda_L \epsilon_p^{-2}]. \tag{6.15}$$

Consequently the number of pairs $(A, B) \in (\mathbb{F}_p^\times)^2$ whose polygonal paths lie in this open uniform ball is at least

$$\tfrac{1}{2}(p-1)^2 \exp[-\Lambda_L D_L^{-2}\sqrt{\log p}].$$

The threshold is effective. Centers may depend deterministically on p while remaining in one fixed class.

We prove an explicit finite inequality first, so that all probability errors and degree restrictions are visible.

For integers $2 \le M \le N < p/2$ and $r \ge \max(4, \lceil\log(16M)\rceil)$, define

$$\begin{aligned} B_r &= 512\sqrt{r/M} + 384r(N^2/p)^{1/(4r)}, \\ T_{\mathrm{lim}} &= \left[\frac{4096\sqrt{r}}{\epsilon\sqrt{M}}\right]^{2r}, & T_{\mathrm{ar}} &= \frac{32768N^{-1/4}}{\epsilon} + \left[\frac{16B_r}{\epsilon}\right]^{2r}, \\ T_{\mathrm{noise}} &= (56M+2)\exp(-M\epsilon^2/65536), & T_{\mathrm{char}} &= \frac{4(2M+2)}{\sqrt{p}}\exp(16M). \end{aligned} \tag{6.16}$$

Proposition 6.6. *For $0 < \epsilon \le 1/4$ and these integer parameters, if*

$$p^{-1/2} + 8M/p \le \epsilon/16, \tag{6.17}$$

then

$$\inf_{g\in\mathcal{H}_L} \mu_p(B(g,\epsilon)) \ge e^{-\Lambda_L\epsilon^{-2}} - T_{\mathrm{lim}} - T_{\mathrm{ar}} - 2T_{\mathrm{noise}} - T_{\mathrm{char}}. \tag{6.18}$$

A negative right side is permitted.

Proof. First upgrade the growing-moment block estimate to a path-norm moment. For a polynomial with the coefficients of (6.12), the L^{2r} norm of its maximum on a J-point grid is at most $J^{1/(2r)}$ times the largest pointwise L^{2r} norm: sum the $2r$th powers before taking roots. Grid reconstruction and pinning contribute a factor four. Thus (6.12) remains valid with its left side replaced by the L^{2r} norm of the uniform norm. Minkowski then sums the dyadic blocks used above. The same geometric sums, now for arbitrary $r \geq \max(4, \lceil \log(16M) \rceil)$, give

$$\left(\mathbb{E}\,\|\mathcal{P}_N Q_p - \mathcal{P}_M Q_p\|_\infty^{2r}\right)^{1/(2r)} \leq B_r, \tag{6.19}$$

$$\left(\mathbb{E}\,\|K - \mathcal{P}_M K\|_\infty^{2r}\right)^{1/(2r)} \leq 512\sqrt{r/M}. \tag{6.20}$$

Indeed $(14M)^{1/(2r)} < 2$, the first geometric sum is at most five, and the second at most $6r$, as in the previous proof. For the limit process the arithmetic term is absent. Apply the finite-block bound through any larger cutoff; the L^1 convergence of $\mathcal{P}_N K$ to K supplies an almost surely convergent subsequence, and Fatou proves (6.20). No L^{2r} convergence is assumed. Markov at $\epsilon/8$ gives

$$\mathbb{P}(\|K - \mathcal{P}_M K\|_\infty > \epsilon/8) \leq T_{\mathrm{lim}}.$$

Split $Q_p - \mathcal{P}_M Q_p$ at $\mathcal{P}_N Q_p$, allocating $\epsilon/16$ to each part. Equations (6.8) and (6.19) give

$$\mathbb{P}(\|Q_p - \mathcal{P}_M Q_p\|_\infty > \epsilon/8) \leq T_{\mathrm{ar}}.$$

Use the allocated degrees (6.2) and the conditional binomial variables Y_h of Lemma 6.4. Their centered differences $\Delta_h = Y_h - x_h$ obey the same exponential bound uniformly in x. At the endpoint,

$$\mathbb{P}(|\Delta_0| > \epsilon/16 \mid x) \leq 2e^{-M\epsilon^2/2048}.$$

For $T(t) = \sum_{h\neq 0} b_h \Delta_h e^{2\pi i h t}$ as in that lemma, the variance proxy is $< 1/M$, so each grid value has tail $4e^{-Mu^2/16}$. There are at most $14M$ grid points and $\|T - T(0)\|_\infty \leq 4 \max_{\mathrm{grid}} |T|$. At $u = \epsilon/64$, the nonzero-frequency part exceeds $\epsilon/16$ with probability at most $56Me^{-M\epsilon^2/65536}$. Adding the endpoint shows

$$\mathbb{P}(\|\mathcal{A}_M(Y - x)\|_\infty > \epsilon/8 \mid x) \leq T_{\mathrm{noise}}. \tag{6.21}$$

This includes boundary coordinates, where the auxiliary binomial is deterministic.

Fix g and set

$$F(y) = \mathbf{1}_{\{\|\mathcal{A}_M(y) - g\|_\infty < \epsilon/2\}}, \qquad P(x) = \mathbb{E}[F(Y) \mid x].$$

Despite the discontinuity of F, its conditional average is an ordinary tensor Bernstein polynomial: it is a finite sum of fixed grid values times binomial probability polynomials. Its degree in coordinate h is at most n_h, and $0 \leq P \leq 1$ everywhere on the cube. By (6.21),

$$\begin{aligned} P(x) &\geq 1 - T_{\mathrm{noise}} && \text{if } \|\mathcal{A}_M(x) - g\|_\infty < \epsilon/4, \\ P(x) &\leq T_{\mathrm{noise}} && \text{if } \|\mathcal{A}_M(x) - g\|_\infty \geq 3\epsilon/4. \end{aligned}$$

The radius slack avoids any norm-sphere atom assumption.

The angular-lift character calculation above applies to every polynomial bounded by one with these separate degrees. It gives rank-weighted character mass at most $\prod_h (n_h + 2)^2 \leq e^{16M}$. Lemma 6.3, with its trivial-character discrepancy zero and at most $2M + 1$ active shifts, therefore gives

$$|\mathbb{E}P(W) - \mathbb{E}P(X)| \leq T_{\mathrm{char}}.$$

The coefficients of the discontinuous F are never inserted into the trace estimate; only the bounded polynomial P is used.

Put $v = e^{-\Lambda_L \epsilon^{-2}}$. Theorem 1.1 at radius $\epsilon/8$, and the limit residual bound, yield

$$\mathbb{P}(\|\mathcal{A}_M(X) - g\|_\infty < \epsilon/4) \geq v - T_{\mathrm{lim}}.$$

If $T_{\mathrm{noise}} \leq 1$, the two polynomial implications and the character comparison imply

$$\begin{aligned}\mathbb{P}(\|\mathcal{A}_M(W) - g\|_\infty < 3\epsilon/4) &\geq (1 - T_{\mathrm{noise}})(v - T_{\mathrm{lim}}) - T_{\mathrm{char}} - T_{\mathrm{noise}} \\ &\geq v - T_{\mathrm{lim}} - 2T_{\mathrm{noise}} - T_{\mathrm{char}}.\end{aligned}$$

If $T_{\mathrm{noise}} > 1$, the final lower bound is nonpositive and trivial. Outside an event of probability at most T_{ar}, (6.6), (6.9) and (6.17) give $\|K_p - \mathcal{A}_M(W)\|_\infty \leq 3\epsilon/16$. The finite-coefficient ball of radius $3\epsilon/4$ therefore lands inside the strict final ball, since $3/4 + 3/16 = 15/16 < 1$. Subtract T_{ar} to prove (6.18). Every estimate is uniform in g. □

Proof of Theorem 6.5. Set $t = \log p$, $\epsilon = D_L t^{-1/4}$ and choose

$$r = \lceil 4(\Lambda_L + 1)\epsilon^{-2} \rceil, \qquad M = \lceil C_0 r \epsilon^{-2} \rceil, \qquad N = \lfloor p^{1/8} \rfloor.$$

Then r is asymptotic to a positive constant times $\sqrt{t}$, and $M \sim t/256$. Eventually $M \leq t/128$, all finite hypotheses hold, and (6.17) holds. Write $v = e^{-\Lambda_L \epsilon^{-2}}$. We verify that every error in (6.18) is $o(v)$.

Since $M \geq C_0 r \epsilon^{-2}$, $T_{\mathrm{lim}} \leq 64^{-2r}$, and $r \geq 4(\Lambda_L + 1)\epsilon^{-2}$ makes its ratio to v tend to zero. The first part of T_{ar} is a fixed polynomial in t times $e^{-t/32}$, up to the harmless floor constant in N, and is also $o(v)$. The first part of B_r is at most $\epsilon/512$; its second part is at most

$$384 r \exp[-3t/(16r)] = o(\epsilon).$$

Thus eventually $B_r \leq \epsilon/256$, so the second part of T_{ar} is at most $16^{-2r} = o(v)$. Furthermore

$$M\epsilon^2/65536 \geq 2^{20} r,$$

which makes $T_{\mathrm{noise}}/v \to 0$ despite its polynomial prefactor. Finally $16M \leq t/8$ gives $T_{\mathrm{char}} \leq 4(2M + 2)e^{-3t/8} = o(v)$. Their sum, including the factor two on T_{noise}, is at most $v/2$ eventually. Proposition 6.6 proves (6.15). All comparisons are explicit elementary inequalities in t, giving an effective threshold for fixed L. □

The radius gain is paid by probability-scale accuracy: r of order ϵ^{-2} controls the genuine and auxiliary tails, M of order $r\epsilon^{-2} = \epsilon^{-4}$ pays uniform path concentration, and the character budget allows M of order $\log p$.

6.6 Uniform arithmetic neighborhoods

Put $A_L = 104000(L + 2)^2$. A bump equal to one on the closed $\epsilon/2$-ball, linear on the annulus and zero outside the open ϵ-ball has Lipschitz constant $2/\epsilon$. Theorem 1.1 and Theorem 6.1 yield

$$\inf_{g \in \mathcal{H}_L} \mu_p(B(g, \epsilon)) \geq \exp(-A_L \epsilon^{-2}) - \frac{2D_B(p, M, N)}{\epsilon}, \qquad 0 < \epsilon \leq 1/4. \tag{6.22}$$

This is a finite-prime inequality, even when its right side is negative. The balls are open; using a closed inner ball needs no norm-sphere atom assumption.

Corollary 6.7. *Fix $L \geq 1$ and $0 < \kappa < 1/2$, and put*

$$\epsilon_p = \sqrt{\frac{A_L}{\kappa \log\log p}}.$$

For all sufficiently large primes, with an effective threshold,

$$\inf_{g \in \mathcal{H}_L} \#\{(A,B) \in (\mathbb{F}_p^\times)^2 : \|K_p(\cdot; A, B) - g\|_\infty < \epsilon_p\} \geq \tfrac{1}{2}(p-1)^2 (\log p)^{-\kappa}.$$

In particular $\kappa = 1/4$ gives $\epsilon_p = \sqrt{4A_L/\log\log p}$.

Proof. The main term in (6.22) is $(\log p)^{-\kappa}$. The error is $O_L((\log p)^{-1/2} \log\log p)$, eventually at most half the main term. The explicit estimate (6.3) and $\epsilon_p \leq 1/4$ give an effective threshold. □

The bounds permit deterministic prime-dependent centers in one fixed class, with explicit thresholds.

The following fixed-radius consequence uses only weak convergence. Its compactness argument also applies to the Hölder classes of the companion paper [16].

Corollary 6.8. *Fix $L \geq 1$ and $0 < \epsilon \leq 1/4$. Then*

$$\liminf_{\substack{q \to \infty \\ q \text{ prime}}} \inf_{g \in \mathcal{H}_L} \frac{1}{(q-1)^2} \#\{(a,b) \in (\mathbb{F}_q^\times)^2 : \|K_q(\cdot; a, b) - g\|_\infty < \epsilon\} \geq \exp\{-104000(L+2)^2 \epsilon^{-2}\}. \tag{6.23}$$

In particular the same bound holds for any deterministic choice of centers $g_q \in \mathcal{H}_L$.

Proof. The class $\mathcal{H}_L$ is compact in uniform norm. Indeed $|g(t)| \leq L\sqrt{t} \leq L$ and its common Hölder bound gives equicontinuity, hence relative compactness by Arzelà–Ascoli. It is closed: endpoints and Hölder inequalities pass to uniform limits, and for each fixed $h \neq 0$ the map

$$g \longmapsto 2\pi i h \int_0^1 [g(t) - t g(1)] \mathrm{e}^{-2\pi i h t}\, dt$$

is continuous. Its constraint to the closed real interval $[-2, 2]$ is therefore preserved. The intersection of these constraints is closed.

Choose a finite $\epsilon/4$-net $g_1, \ldots, g_J$ from the class. For every $g \in \mathcal{H}_L$ some j satisfies $B(g_j, \epsilon/2) \subset B(g, \epsilon)$, with open uniform balls. It follows that

$$\inf_{g \in \mathcal{H}_L} \mu_q(B(g, \epsilon)) \geq \min_{1 \leq j \leq J} \mu_q(B(g_j, \epsilon/2)).$$

Portmanteau and Theorem 1.1 give, for each j,

$$\liminf_q \mu_q(B(g_j, \epsilon/2)) \geq \mu(B(g_j, \epsilon/2)) \geq \exp\{-104000(L+2)^2 \epsilon^{-2}\}.$$

For finitely many sequences, the liminf of their minimum is at least the minimum of their liminfs: take the largest of their finitely many eventual starting indices. This proves (6.23). The deterministic moving-center assertion follows from the infimum bound. □

Statements and Declarations

Funding. This research was self-funded. No external funding or material support was received.

Competing interests. The author owns Luminary Oddities.

Data availability. All mathematical arguments supporting the results are contained in this article.

Author contribution. Scott Schubr directed the research and preparation of the manuscript and is responsible for its mathematical content and attribution.

AI assistance. OpenAI ChatGPT/Codex and Anthropic Claude supported argument development, literature discovery, proof checking and drafting. The named author is responsible for the mathematical arguments and the manuscript.

References

[1] E. Kowalski and W. F. Sawin, Kloosterman paths and the shape of exponential sums, *Compos. Math.* **152** (2016), no. 7, 1489–1516. https://doi.org/10.1112/S0010437X16007351; Preprint https://arxiv.org/abs/1410.7892v1; theorem numbers follow arXiv:1410.7892v1.

[2] E. Kowalski and W. Sawin, *On the support of the Kloosterman paths*, arXiv:1709.05192, version 1 (2017). https://arxiv.org/abs/1709.05192v1.

[3] D. Kyselov, *Space-filling curves and small-ball probabilities for the Kloosterman process*, arXiv:2610.22645, version 1, October 15, 2026. https://arxiv.org/abs/2610.22645v1.

[4] E. Kowalski and T. Untrau, *Wasserstein metrics and quantitative equidistribution of exponential sums over finite fields*, arXiv:2505.22059v2 (July 28, 2025), and author manuscript dated May 26, 2026, Section 1.3(4). https://arxiv.org/abs/2505.22059v2; https://people.math.ethz.ch/~kowalski/wasserstein.pdf.

[5] Amir M. Rahimi, *Wasserstein rates of all orders for the Sato–Tate law of Kloosterman sums*, arXiv:2610.16169, version 2, November 3, 2026, Theorem A. https://arxiv.org/abs/2610.16169v2.

[6] S. Agapiou, M. Dashti and T. Helin, *Rates of contraction of posterior distributions based on p-exponential priors*, arXiv:1811.12244, version 4, October 8, 2020. https://arxiv.org/abs/1811.12244v4.

[7] A. Prékopa, On logarithmic concave measures and functions, *Acta Sci. Math. (Szeged)* **34** (1973), 335–343, Theorem 6. https://acta.bibl.u-szeged.hu/14411/1/math_034_335-343.pdf.

[8] M. Kanter, Unimodality and dominance for symmetric random vectors, *Trans. Amer. Math. Soc.* **229** (1977), 65–85. https://doi.org/10.1090/S0002-9947-1977-0445580-7.

[9] G. Paouris and P. Pivovarov, *Randomized isoperimetric inequalities*, arXiv:1607.00519, version 1, Theorem 3.6 and its proof. https://arxiv.org/abs/1607.00519v1.

[10] T. W. Anderson, The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities, *Proc. Amer. Math. Soc.* **6**, no. 2 (1955), 170–176. https://doi.org/10.1090/S0002-9939-1955-0069229-1.

[11] C. Clason, T. Helin, R. Kretschmann and P. Piiroinen, Generalized modes in Bayesian inverse problems, *SIAM/ASA J. Uncertain. Quantif.* **7** (2019), 652–684. https://doi.org/10.1137/18M1191804; preprint https://arxiv.org/abs/1806.00519v2.

[12] F. Aurzada, *Small deviation probabilities of some stochastic processes*, doctoral dissertation, Friedrich-Schiller-Universität Jena, 2006, Section 2.6.4, printed pp. 49–51. https://d-nb.info/989385213/34.

[13] L. Fu, Y.-K. Lau and P. Xi, *An effective Deligne's equidistribution theorem*, arXiv:2406.09633, version 2, June 27, 2024, Remark 4.2. https://arxiv.org/abs/2406.09633v2.

[14] E. Fouvry, S. Ganguly, E. Kowalski and P. Michel, Gaussian distribution for the divisor function and Hecke eigenvalues in arithmetic progressions, *Comment. Math. Helv.* **89** (2014), 979–1014. https://doi.org/10.4171/CMH/342; preprint https://arxiv.org/abs/1301.0214v1, proof of Proposition 3.2.

[15] E. Fouvry, E. Kowalski, P. Michel and W. Sawin, *Lectures on applied ℓ-adic cohomology*, extended notes for the 2016 Arizona Winter School, Theorems 4.1, 4.2, 4.6 and equation (4.2). https://people.math.ethz.ch/~kowalski/applied-ell-adic.pdf.

[16] S. Schubr, *Roughness and boundary-density thresholds for small balls of the Kloosterman process*, companion paper, Luminary Research Archive (2026).