ARITHMETIC CONCENTRATION AND GAUSSIAN FLUCTUATIONS FOR RANDOM WINDOWS OF ELLIPTIC-CURVE CONGRUENTIAL GENERATORS

ZIRAN LIU, CHUNG PANG MOK

Abstract. Let $E/\mathbb{F}_q$ have characteristic $p \geq 5$, let $G_0 \in E(\mathbb{F}_q)$ have order t, and consider the translation orbit $P_u = P_0 + [u]G_0$. We choose the initial index U uniformly in $\mathbb{Z}/t\mathbb{Z}$ and study length-N windows of a trace-based digital output, with $1 \leq N \leq t$. The central question is seed-wise: how many cyclic shifts produce an atypical Walsh bias or an atypically large discrepancy, even though the underlying orbit is periodic and has no mixing mechanism? We prove uniform mixed-correlation bounds of every total degree below p, exact Gaussian diagonal terms for even moments, and explicit arithmetic errors of size $N^D q^{1/2}/t$. These estimates yield finite-order subgaussian tails, confidence bounds for finite Walsh batteries, concentration of random-window discrepancy, and an effective concentration depth determined by the largest moment order for which the Gaussian diagonal dominates the arithmetic error. The restriction on the degree is genuine: a characteristic-p resonance occurs at degree p. In mesoscopic windows we prove joint complex-Gaussian limits for arbitrary fixed frequency families, including repeated- and inverse-frequency resonances, together with quantitative polynomial-moment and fixed-cumulant estimates. The mechanism is arithmetic square-root cancellation in complete elliptic subgroup sums, not dynamical mixing. We also identify precisely what additional all-order control would be required for moderate or large deviations.

Contents

2020 *Mathematics Subject Classification.* 11G20, 11K38, 11T23, 60E15, 60F05.

Key words and phrases. elliptic curves, congruential generators, Walsh functions, high moments, discrepancy, arithmetic concentration, Gaussian fluctuations.

1. Introduction and main results

1.1. **Three questions in arithmetic probability on elliptic dynamics.** Elliptic-curve pseudo-random constructions give rise to three related but mathematically distinct probability questions.

(QI) *Random seed on a deterministic orbit.* A deterministic elliptic-curve orbit is observed through a finite window, and its starting index is chosen uniformly. How many seeds produce an atypical empirical bias or discrepancy?

(QII) *Random affine evolution.* Independent increments are inserted at every step of an affine recursion $X_{n+1} = TX_n + B_{n+1}$. At what time does the one-time marginal mix, and what arithmetic controls cutoff or exceptional slow modes?

(QIII) *Empirical deviations after random evolution.* For the random affine walk, what concentration, moderate-deviation, and large-deviation laws govern path averages, uniformly over a growing family of elliptic state spaces?

The present paper gives a self-contained answer to (QI) for cyclic translation orbits. The companion mixing problem (QII) requires Fourier products, entropy obstructions, and post-wrap cancellation for power orbits. Question (QIII) additionally requires uniform control of tilted transition operators and their Perron data. These distinctions are structural: no Markov-chain mixing theorem is used anywhere in the proofs below.

1.2. **The precise problem treated in Paper I.** The probability space in this paper is the finite seed space, not a path-space noise law. Let $E/\mathbb{F}_q$ be an elliptic curve, let $G_0 \in E(\mathbb{F}_q)$ have order t, and let

$$P_u = P_0 + [u]G_0, \qquad u \in \mathbb{Z}/t\mathbb{Z}.$$

The orbit is completely deterministic.

(R) Randomness enters only through a uniformly chosen starting index $U \in \mathbb{Z}/t\mathbb{Z}$.

For a nonzero Walsh frequency k, let $Y_k(u)$ be the corresponding trace-digital Walsh observable, completed by the value 0 when $P_u = O$, and define

$$Z_{N,k}(U) = \sum_{n=0}^{N-1} Y_k(U+n), \qquad M_{N,k}(U) = \frac{Z_{N,k}(U)}{N}, \qquad \mu_k = \frac{1}{t}\sum_{u \bmod t} Y_k(u),$$

where $1 \le N \le t$. Paper I asks for a distributional theory of these length-N windows as the seed varies.

More precisely, Paper I resolves four linked questions.

(I.1) *Correlation and moment structure.* Given translated Walsh observables, which time assignments contribute a main term, and how large is every remaining correlation? The required answer must be uniform in the nonzero Walsh frequencies and must keep track of the characteristic.

(I.2) *Exceptional seeds.* For a prescribed moment order m, how small is

$$\mathbb{P}\big(|M_{N,k}(U) - \mu_k| \ge x\big),$$

and how does the best admissible m depend on N, p, and the arithmetic period ratio $t/q^{1/2}$? The aim is not merely an asymptotic statement, but a finite-parameter bound that identifies when Gaussian-scale seed tails are actually available.

(I.3) *Whole-window output quality.* Can the individual Walsh estimates be assembled, with all finite-grid and point-at-infinity corrections visible, into a bound for the extreme discrepancy of the complete digital window and therefore into a randomized quasi–Monte Carlo integration estimate?

(I.4) *Fluctuation law.* In a mesoscopic asymptotic regime, do finitely many normalized Walsh window sums converge jointly to a Gaussian vector, and which repeated- or inverse-frequency resonances survive in the covariance and pseudo-covariance?

These questions are not answered by a full-orbit discrepancy estimate. A full-orbit estimate controls one average over all t indices; it does not give the distribution, over U, of the strongly overlapping local windows $\{U, \dots, U+N-1\}$. Nor can one appeal to decay of temporal dependence: the map $u \mapsto u+1$ on $\mathbb{Z}/t\mathbb{Z}$ is periodic and has no mixing. The paper must therefore classify the complete high-order correlations generated by the window expansion itself.

The arithmetic mechanism is the following. A product of translated Walsh observables becomes an additive character of a rational function on E. At each occupied time shift one obtains a translated copy of the coordinate poles. If the coefficient vector at some shift is nonzero, the resulting phase has there a pole of order 2 or 3 with nonzero leading coefficient; all other translated terms are regular at that point. For total degree $D < p$, this pole order is prime to p, excludes the Artin–Schreier exceptional class, and permits square-root cancellation in a complete elliptic subgroup sum. Balanced assignments, by contrast, have zero phase and produce the combinatorial main terms. Thus the three structural restrictions

$$N \le t, \qquad D < p, \qquad \frac{q^{1/2}}{t}$$

enter for different reasons: distinct translated poles, exclusion of characteristic-p resonance, and the size of the arithmetic off-diagonal error, respectively.

The principal conclusions are stated next, before any technical setup. They answer (I.1)–(I.4) in that order and make explicit what is proved uniformly, what is asymptotic, and what remains outside the scope of Paper I.

1.3. **Main results.** The detailed digital notation is introduced in Section 2. We record the theorem architecture here. Put

$$C_{\mathrm{corr}} = 7, \qquad C_{\mathrm{ar}} = 14,$$

and let

$$\mathfrak{D}_m(N) = \#\big\{(\mathbf{n}, \mathbf{m}) \in \{0, \dots, N-1\}^{2m} : \{n_1, \dots, n_m\} = \{m_1, \dots, m_m\}\big\}.$$

Thus $\mathfrak{D}_m(N) = m!N^m + O_m(N^{m-1})$ and $\mathfrak{D}_m(N) \le m!N^m$.

Theorem A (Mixed moments and Wick pairings)**.** *Let $k \neq 0$ and let $a, b \ge 0$ satisfy $1 \le a + b < p$. If $a \neq b$, then*

$$\left|\mathbb{E} Z_{N,k}^{a} \overline{Z_{N,k}}^{b}\right| \le C_{\mathrm{corr}}(a+b) N^{a+b} \frac{q^{1/2}}{t}.$$

If $a = b = m$, then

$$\mathbb{E}\,|Z_{N,k}|^{2m} = \mathfrak{D}_m(N) + \mathcal{E}_m, \qquad |\mathcal{E}_m| \le \frac{m\mathfrak{D}_m(N)}{t} + C_{\mathrm{ar}} m N^{2m} \frac{q^{1/2}}{t}.$$

More generally, every fixed joint mixed moment of a finite frequency family equals its compatible Wick-pairing term plus

$$O_D\left(N^{D/2-1/2} + \frac{N^{D/2}}{t} + N^D \frac{q^{1/2}}{t}\right),$$

where $D < p$ is the total degree and the pairing rule records exactly repeated and inverse Walsh frequencies.

The one-frequency statement is Theorem 4.3; the general pairing statement is Theorem 4.7.

Theorem B (Finite-order arithmetic concentration)**.** *For $1 \le m < p/2$, define*

$$\mathfrak{a}_m(q,t) = \left(C_{\mathrm{ar}} m \frac{q^{1/2}}{t} \right)^{1/(2m)}.$$

Then, uniformly in $k \neq 0$,

$$\|M_{N,k} - \mu_k\|_{2m} \le \sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + C_{\mathrm{corr}} \frac{q^{1/2}}{t}.$$

Consequently, the probability of exceeding e *times the displayed scale is at most* e^{-2m}*. If $m \le N$ and*

$$C_{\mathrm{ar}} m \frac{q^{1/2}}{t} \le \left(\frac{m}{N}\right)^m,$$

then

$$\mathbb{P}\left(|M_{N,k} - \mu_k| \ge 3\mathrm{e}\sqrt{\frac{m}{N}} \right) \le \mathrm{e}^{-2m}.$$

This package is proved in Theorem 5.1 and Corollary 5.6.

Theorem C (Random-window discrepancy)**.** *Let $M = p^a$, $d = 2r$, and*

$$\Delta_{p,a,d} = 1 - (1 - M^{-1})^d.$$

There are explicit finite-grid Walsh weights $\rho_d(k)$ satisfying

$$W_d := \sum_{k \neq 0} \rho_d(k) \le (1 + 7 \log M)^d - 1.$$

If $D_{U,N}$ denotes the extreme discrepancy of the completed length-N window, then for $2m < p$,

$$\mathbb{P}\left(D_{U,N} > \Delta_{p,a,d} + W_d \left(C_{\mathrm{corr}} \frac{q^{1/2}}{t} + \frac{1}{N} \right) + \mathrm{e} W_d \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + C_{\mathrm{corr}} \frac{q^{1/2}}{t} \right] \right) \le \mathrm{e}^{-2m}.$$

The finite-grid inequality is Lemma 6.3, and the probabilistic conclusion is Theorem 6.4.

Theorem D (Joint Gaussian fluctuations)**.** *Consider a family for which $p_j \to \infty$, $N_j \to \infty$, and*

$$N_j^m \frac{q_j^{1/2}}{t_j} \longrightarrow 0 \quad \textit{for every fixed } m \ge 1.$$

For any fixed finite family of nonzero Walsh frequencies whose equality and inverse-frequency relations eventually stabilize, the centered vector

$$\left(\frac{Z_{N_j,k_{1,j}} - N_j \mu_{k_{1,j}}}{\sqrt{N_j}}, \dots, \frac{Z_{N_j,k_{s,j}} - N_j \mu_{k_{s,j}}}{\sqrt{N_j}} \right)$$

converges to the centered complex Gaussian vector whose covariance records equality and whose pseudo-covariance records inverse-frequency resonance. Fixed polynomial moments and fixed joint cumulants admit explicit finite-window errors.

The full joint limit and its quantitative refinements are Theorems 8.2, 8.3, and 8.8.

1.4. **Sharpness, scope, and what remains open.** The restriction $D < p$ is not a technical convenience. At degree p, the identity $Y_k(u)^p = 1$ at every finite orbit point creates a correlation of order one; Proposition 3.8 makes this obstruction explicit. The paper therefore proves a finite-order concentration theory whose available depth is jointly limited by the characteristic and by the ratio $t/q^{1/2}$.

The results do not constitute a speed-N large-deviation principle. Such a theorem would require all-order control at moment order comparable with N, or direct convergence of the two-dimensional logarithmic moment generating function. A further finite-sample obstruction is that every nonempty seed event has probability at least $1/t$.

The proof has six steps. First, the finite Walsh dual is identified explicitly with the additive-character dual of $\mathbb{F}_q^2$. Second, a translated-pole lemma treats all single- and multi-frequency correlations at once. Third, balanced assignments are counted, with pair blocks producing Wick's rule. Fourth, moment estimates are optimized into finite-order tails. Fifth, the one-dimensional finite-grid Walsh weights are bounded directly and tensorized to discrepancy. Sixth, the joint moments are normalized to obtain Gaussian limits and quantitative cumulant estimates.

1.5. **Relation to earlier work and organization.** The discrepancy background comes from Mok [5], Wang [6], and Liu–Mok [4]. The complete subgroup character-sum input is the untwisted specialization of Kohel–Shparlinski [10]. Related elliptic exponential sums and random walks appear in Lange–Shparlinski [7]. Classical concentration results for mixing systems derive their tails from decay of dependence [8]; the present periodic setting uses arithmetic cancellation instead. Probabilistic discrepancy bounds for independent Monte Carlo points are discussed in Aistleitner–Hofer [9].

Sections 2–4 establish the arithmetic moment input. Sections 5–7 derive concentration and discrepancy. Section 8 proves Gaussian fluctuation results. The final sections explain applications, the relation to the random affine program, and the precise additional input needed for moderate and large deviations.

2. Setup

2.1. **The cyclic orbit.** Let $q = p^{ar}$, where $a, r \geq 1$ and $p \geq 5$, and let $E/\mathbb{F}_q$ be an elliptic curve with identity O. Fix a Weierstrass model over $\mathbb{F}_q$ and write $x, y \in \mathbb{F}_q(E)$ for its coordinate functions. Fix $G_0 \in E(\mathbb{F}_q)$ of order

$$t = \operatorname{ord}(G_0),$$

and a point $P_0 \in E(\mathbb{F}_q)$. Put

$$P_u = P_0 + [u]G_0, \qquad u \in \mathbb{Z}/t\mathbb{Z},$$

and write $H = \langle G_0 \rangle$.

Remark 2.1 (Role of the characteristic hypothesis). The assumption $p \geq 5$ is used when an unbalanced correlation is shown to have a pole of order 2 or 3. These orders are prime to p, so Lemma 3.4 excludes the Artin–Schreier exceptional class. In characteristics 2 and 3 a separate analysis is necessary, and the frequency-uniform moment theorems below are not asserted.

2.2. **Digital coordinates and the finite Walsh group.** Fix an $\mathbb{F}_{p^a}$-basis $\{\lambda_1, \dots, \lambda_r\}$ of $\mathbb{F}_q$ and its trace-dual basis $\{\lambda_1', \dots, \lambda_r'\}$ for $\operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_{p^a}}$. Fix an $\mathbb{F}_p$-basis $\{\kappa_1, \dots, \kappa_a\}$ of $\mathbb{F}_{p^a}$ and its trace-dual basis $\{\kappa_1', \dots, \kappa_a'\}$ for $\operatorname{Tr}_{\mathbb{F}_{p^a}/\mathbb{F}_p}$. For $\eta \in \mathbb{F}_q$, set

$$\phi_j(\eta) = \sum_{\nu=1}^{a} \frac{\left\langle \operatorname{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\eta \lambda_j' \kappa_\nu') \right\rangle_p}{p^\nu}, \qquad 1 \leq j \leq r,$$

where $\langle\cdot\rangle_p \in \{0,\dots,p-1\}$ is the standard representative. Thus the a base-p digits of $\phi_j(\eta)$ are the $\mathbb{F}_p$-coordinates of the jth $\mathbb{F}_{p^a}$-coordinate of η. For $P \neq O$, define

$$\Phi(P) = \big(\phi_1(x(P)),\dots,\phi_r(x(P)),\phi_1(y(P)),\dots,\phi_r(y(P))\big) \in [0,1)^{2r}.$$

Put

$$\Delta_{2r} = \{0,1,\dots,p^a-1\}^{2r}, \qquad \Delta_{2r}^* = \Delta_{2r} \setminus \{0\}.$$

We identify every digit in $\{0,\dots,p-1\}$ with its image in the prime field $\mathbb{F}_p$. Write the jth component of $k \in \Delta_{2r}$ as

$$k_j = \sum_{\nu=1}^{a} k_{j,\nu} p^{\nu-1}, \qquad k_{j,\nu} \in \{0,\dots,p-1\}.$$

If $z_j = \sum_{\nu=1}^{a} z_{j,\nu} p^{-\nu}$ is a point of the digital grid, our finite Walsh convention is

$$w_k(z) = \exp\left(\frac{2\pi i}{p} \sum_{j=1}^{2r}\sum_{\nu=1}^{a} k_{j,\nu} z_{j,\nu}\right).$$

All grid coordinates are represented by their terminating a-digit expansions. The group operation $\oplus$ on Δ_{2r} is digitwise addition in $\mathbb{F}_p$, without carries; $\ominus k$ denotes the digitwise inverse.

Lemma 2.2 (Explicit Walsh-to-character conversion)**.** *For $k \in \Delta_{2r}$ define*

$$\eta_k = \sum_{j=1}^{r}\sum_{\nu=1}^{a} k_{j,\nu}\lambda'_j\kappa'_\nu, \qquad \widetilde{\eta}_k = \sum_{j=1}^{r}\sum_{\nu=1}^{a} k_{r+j,\nu}\lambda'_j\kappa'_\nu.$$

Then, for $P \neq O$,

$$w_k(\Phi(P)) = \psi_1\big(\eta_k x(P) + \widetilde{\eta}_k y(P)\big), \qquad \psi_1(z) = \exp\left(\frac{2\pi i}{p}\,\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(z)\right).$$

Proof. By trace duality and transitivity of trace, the νth digit of $\phi_j(\eta)$ is

$$\mathrm{Tr}_{\mathbb{F}_{p^a}/\mathbb{F}_p}\left(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{p^a}}(\eta\lambda'_j)\kappa'_\nu\right) = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(\eta\lambda'_j\kappa'_\nu).$$

Substituting these digits in the definition of w_k and using linearity of the absolute trace gives

$$\sum_{j,\nu} k_{j,\nu}\,\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x(P)\lambda'_j\kappa'_\nu) + \sum_{j,\nu} k_{r+j,\nu}\,\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(y(P)\lambda'_j\kappa'_\nu)$$

inside the exponential. Pulling the sums into the trace yields the claimed formula. □

Write

$$\vartheta(k) = (\eta_k, \widetilde{\eta}_k) \in \mathbb{F}_q^2.$$

Lemma 2.3 (Frequency coefficient map)**.** *The map*

$$\vartheta : (\Delta_{2r}, \oplus) \longrightarrow (\mathbb{F}_q^2, +)$$

is an isomorphism of $\mathbb{F}_p$-vector spaces. In particular,

$$\vartheta(k) = 0 \iff k = 0, \qquad \vartheta(\ominus k) = -\vartheta(k).$$

Proof. The displayed formulas for η_k and $\widetilde{\eta}_k$ show that ϑ is $\mathbb{F}_p$-linear. The elements $\lambda'_j\kappa'_\nu$ form an $\mathbb{F}_p$-basis of $\mathbb{F}_q$, so vanishing of both coordinates forces every digit $k_{j,\nu}$ to vanish. Thus the kernel is trivial. Both vector spaces have cardinality $p^{2ar} = q^2$, so ϑ is bijective. The inverse-frequency identity follows from linearity. □

For $k \in \Delta_{2r}^*$ put

$$Y_k(u) = \begin{cases} w_k(\Phi(P_u)), & P_u \neq O, \\ 0, & P_u = O, \end{cases} \qquad \mu_k = \frac{1}{t} \sum_{u \bmod t} Y_k(u).$$

The completion at O implies

$$Y_{\ominus k}(u) = \overline{Y_k(u)} \quad \text{for every } u.$$

2.3. **Random cyclic windows.** Let U be uniform on $\mathbb{Z}/t\mathbb{Z}$. For $1 \le N \le t$, define

$$Z_{N,k}(U) = \sum_{n=0}^{N-1} Y_k(U+n), \qquad M_{N,k}(U) = \frac{1}{N} Z_{N,k}(U).$$

All indices are taken modulo t. Unless stated otherwise, $\mathbb{E}$, $\mathbb{P}$, and L^s-norms refer to the uniform choice of U.

3. Complete elliptic correlation sums

The analytic input is a complete additive-character estimate over an elliptic-curve subgroup.

Lemma 3.1 (Complete subgroup sum)**.** *Let $J \le E(\mathbb{F}_q)$, let $f \in \mathbb{F}_q(E)$ be nonconstant, and let ψ be a nontrivial additive character of $\mathbb{F}_q$. Write*

$$\psi(z) = \psi_1(\alpha z) \qquad (\alpha \in \mathbb{F}_q^\times).$$

Assume

$$\alpha f \neq g^p - g + c \qquad \text{for every } g \in \overline{\mathbb{F}}_q(E),\ c \in \overline{\mathbb{F}}_q.$$

Then

$$\left| \sum_{\substack{P \in J \\ f(P) \neq \infty}} \psi(f(P)) \right| \le 2 \deg_\infty(f) q^{1/2},$$

where $\deg_\infty(f)$ is the degree of the polar divisor. The same estimate holds with $f(P)$ replaced by $f(P+Q)$ for any $Q \in E(\mathbb{F}_q)$.

Proof. Every nontrivial additive character of $\mathbb{F}_q$ has the form $\psi(z) = \psi_1(\alpha z)$ for a unique $\alpha \in \mathbb{F}_q^\times$. Multiplication by α does not change the polar divisor, so $\deg_\infty(\alpha f) = \deg_\infty(f)$. Write the polar divisor of f as

$$(f)_\infty = \sum_{i=1}^{s} n_i Q_i.$$

The complete twisted estimate of Kohel–Shparlinski, applied to the phase αf, states that for every group character ω of $E(\mathbb{F}_q)$,

$$\left| \sum_{\substack{P \in E(\mathbb{F}_q) \\ f(P) \neq \infty}} \omega(P) \psi_1(\alpha f(P)) \right| \le \left(\sum_{i=1}^{s} m_i \deg Q_i \right) q^{1/2},$$

where the Artin–Schreier conductor exponent at Q_i satisfies $m_i \le n_i + 1$. The nonexceptional hypothesis in the present lemma is what ensures that this additive character sheaf is nontrivial. Since every $n_i \ge 1$,

$$\sum_{i=1}^{s} m_i \deg Q_i \le \sum_{i=1}^{s} (n_i + 1) \deg Q_i \le 2 \sum_{i=1}^{s} n_i \deg Q_i = 2 \deg_\infty(f).$$

Consequently,

$$\left|\sum_{\substack{P\in E(\mathbb{F}_q)\\ f(P)\neq\infty}} \omega(P)\psi_1(\alpha f(P))\right| \le 2\deg_\infty(f)q^{1/2}. \tag{3.1}$$

We now write out the subgroup reduction. Let

$$J^\perp = \{\chi \in \widehat{E(\mathbb{F}_q)} : \chi(P) = 1 \text{ for every } P \in J\}.$$

Character orthogonality on the quotient $E(\mathbb{F}_q)/J$ gives, for every $P \in E(\mathbb{F}_q)$,

$$\mathbf{1}_J(P) = \frac{1}{|J^\perp|}\sum_{\chi\in J^\perp}\chi(P).$$

Therefore

$$\sum_{\substack{P\in J\\ f(P)\neq\infty}} \psi(f(P)) = \frac{1}{|J^\perp|}\sum_{\chi\in J^\perp}\sum_{\substack{P\in E(\mathbb{F}_q)\\ f(P)\neq\infty}} \chi(P)\psi_1(\alpha f(P)).$$

Applying (3.1) to each inner sum and then the triangle inequality proves the stated bound; the number $|J^\perp|$ cancels with the normalizing factor.

Finally, let $f_Q(P) = f(P+Q)$. Translation by Q is an $\mathbb{F}_q$-automorphism of E, so it carries the polar divisor of f bijectively to that of f_Q and preserves its degree. If $\alpha f_Q = g^p - g + c$, composition with translation by $-Q$ would give the same exceptional representation for αf. Thus the hypotheses and the bound are invariant under translation. □

Remark 3.2 (Artin–Schreier qualification). The nonexceptional hypothesis in Lemma 3.1 is essential: if the phase attached to ψ were of the form $g^p - g + c$, the additive character could be constant on all finite values and no square-root bound would hold. We have therefore written this qualification explicitly. In every application below it is verified locally by the existence of a pole of order 2 or 3, prime to p.

Remark 3.3 (Exact dependency scope). All moment, concentration, and Gaussian theorems use Lemma 3.1 only in its untwisted form. The discrepancy argument uses a finite Walsh inversion proved below together with the direct one-dimensional coefficient estimate in Lemma 6.1. No probabilistic mixing theorem enters the paper.

Lemma 3.4 (Pole-order criterion)**.** *If a rational function $F \in \overline{\mathbb{F}}_q(E)$ has a pole whose order is not divisible by p, then*

$$F \neq g^p - g + c$$

for every $g \in \overline{\mathbb{F}}_q(E)$ and $c \in \overline{\mathbb{F}}_q$.

Proof. If g is regular on the projective curve E, then g is constant and $g^p - g + c$ has no pole. Otherwise, let $d > 0$ be the pole order of g at a point Q. In a local parameter at Q, the leading term of g^p has order $-pd$, whereas the leading term of g has order $-d$. Since $pd > d$, the leading term of g^p cannot cancel with either $-g$ or the constant c. Hence every pole of $g^p - g + c$ has order divisible by p. □

For $k \in \Delta_{2r}^*$ put

$$f_k(P) = \eta_k x(P) + \widetilde{\eta}_k y(P).$$

At O, the function x has pole order 2 and y has pole order 3. Hence f_k has a unique pole: it has order 3 if $\widetilde{\eta}_k \neq 0$, and order 2 if $\widetilde{\eta}_k = 0$ (in which case $\eta_k \neq 0$).

Lemma 3.5 (Translated coordinate poles)**.** *For $Q \in E(\mathbb{F}_q)$, put*

$$x_Q(P) = x(P+Q), \qquad y_Q(P) = y(P+Q).$$

Then x_Q and y_Q have no poles except at $-Q$, where their pole orders are 2 and 3, respectively. Consequently, if $(\alpha,\beta) \neq (0,0)$, then

$$\alpha x_Q + \beta y_Q$$

has a pole of order 3 at $-Q$ when $\beta \neq 0$, and a pole of order 2 there when $\beta = 0$.

Proof. Translation $\tau_Q : P \mapsto P+Q$ is an automorphism of the smooth projective curve E. Pullback by an automorphism preserves valuations. Since x and y have unique poles at O of orders 2 and 3, the pullbacks $x \circ \tau_Q$ and $y \circ \tau_Q$ have unique poles at $\tau_Q^{-1}(O) = -Q$ with the same orders. In a local parameter z at $-Q$, one may therefore write

$$x_Q = c_2 z^{-2} + O(z^{-1}), \qquad y_Q = c_3 z^{-3} + O(z^{-2}),$$

with $c_2, c_3 \neq 0$. If $\beta \neq 0$, the coefficient of z^{-3} in $\alpha x_Q + \beta y_Q$ is $\beta c_3 \neq 0$, so no order-3 cancellation is possible. If $\beta = 0$, then $\alpha \neq 0$ and the coefficient of z^{-2} is $\alpha c_2 \neq 0$. □

For $\epsilon \in \{+1,-1\}$ define

$$Y_k^{[\epsilon]}(u) = \begin{cases} Y_k(u), & \epsilon = +1, \\ \overline{Y_k(u)}, & \epsilon = -1. \end{cases}$$

Lemma 3.6 (General translated-pole correlation bound)**.** *Let $1 \le D < p$. Choose nonzero frequencies $k_1, \dots, k_D$, signs $\epsilon_1, \dots, \epsilon_D \in \{+1,-1\}$, and shifts $0 \le n_\ell < N \le t$. For $0 \le n < N$, set*

$$v_n = (\alpha_n, \beta_n) = \sum_{\ell : n_\ell = n} \epsilon_\ell \vartheta(k_\ell) \in \mathbb{F}_q^2.$$

If $v_n \neq 0$ for at least one n, then

$$\left| \frac{1}{t} \sum_{u \bmod t} \prod_{\ell=1}^{D} Y_{k_\ell}^{[\epsilon_\ell]}(u + n_\ell) \right| \le C_{\mathrm{corr}} D \frac{q^{1/2}}{t} = 7D \frac{q^{1/2}}{t}.$$

Proof. Let

$$\mathcal{T} = \{n : \text{at least one occurrence has shift } n\}, \qquad \mathcal{U} = \{n : v_n \neq 0\}.$$

By hypothesis $\mathcal{U}$ is nonempty, and $|\mathcal{T}| \le D$, $|\mathcal{U}| \le D$. Write $P = [u]G_0 \in H$. Whenever none of the points $P + P_0 + [n]G_0$, $n \in \mathcal{T}$, equals O, the Walsh-to-character identity and the relation $Y_k^{[-1]} = \overline{Y_k}$ give

$$\prod_{\ell=1}^{D} Y_{k_\ell}^{[\epsilon_\ell]}(u + n_\ell) = \psi_1(F(P)),$$

where

$$F(P) = \sum_{n=0}^{N-1} \left[\alpha_n x(P + P_0 + [n]G_0) + \beta_n y(P + P_0 + [n]G_0) \right]. \tag{3.2}$$

For each n, put $Q_n = -P_0 - [n]G_0$. If $0 \le n < m < N \le t$, then $[m-n]G_0 \neq O$, so $Q_n \neq Q_m$. Fix $n \in \mathcal{U}$. By Lemma 3.5, the nth summand of (3.2) has at Q_n a pole of order 3 when $\beta_n \neq 0$ and of order 2 when $\beta_n = 0$. For $m \neq n$, one has

$$Q_n + P_0 + [m]G_0 = [m-n]G_0 \neq O,$$

so every term belonging to shift m is regular at Q_n. Hence the displayed pole cannot be cancelled by any other translated term. It follows that F is nonconstant and that its polar divisor is supported exactly among the points Q_n, $n \in \mathcal{U}$, with

$$\deg_\infty(F) = \sum_{n\in\mathcal{U}} \begin{cases} 3, & \beta_n \neq 0, \\ 2, & \beta_n = 0, \end{cases} \leq 3|\mathcal{U}| \leq 3D. \tag{3.3}$$

At least one of these poles has order 2 or 3. Since $p \geq 5$, that order is not divisible by p, and Lemma 3.4 shows that F is not of Artin–Schreier form. Lemma 3.1 and (3.3) therefore give

$$\left| \sum_{\substack{P\in H \\ F(P)\neq\infty}} \psi_1(F(P)) \right| \leq 6Dq^{1/2}. \tag{3.4}$$

It remains only to compare the sum in (3.4) with the completed product. If $n \in \mathcal{U}$ and $Q_n \in H$, then the complete sum omits Q_n because F has a pole there, while the completed product is 0 because at least one factor is evaluated at O; hence this point contributes no discrepancy between the two sums. If $n \in \mathcal{T} \setminus \mathcal{U}$ and $Q_n \in H$, then F is regular at Q_n but the completed product is 0, so the two sums differ there by the single term $\psi_1(F(Q_n))$, of modulus one. There are at most $|\mathcal{T} \setminus \mathcal{U}| \leq D$ such points, and they are distinct. Thus the absolute difference of the two sums is at most D. Combining this with (3.4), using $q^{1/2} \geq 1$, and dividing by t gives

$$\left| \frac{1}{t} \sum_{u \bmod t} \prod_{\ell=1}^{D} Y_{k_\ell}^{[\epsilon_\ell]}(u + n_\ell) \right| \leq \frac{6Dq^{1/2} + D}{t} \leq 7D\frac{q^{1/2}}{t}.$$

□

Corollary 3.7 (Single-frequency mixed correlation)**.** *Let $a, b \geq 0$ with $1 \leq a + b < p$, and let*

$$\mathbf{n} = (n_1, \dots, n_a), \qquad \mathbf{m} = (m_1, \dots, m_b),$$

where all shifts lie in $\{0, \dots, N-1\}$. If the two multisets of shifts are unequal, then, uniformly in $k \neq 0$,

$$\left| \frac{1}{t} \sum_{u \bmod t} \prod_{i=1}^{a} Y_k(u + n_i) \prod_{j=1}^{b} \overline{Y_k(u + m_j)} \right| \leq 7(a + b)\frac{q^{1/2}}{t}.$$

Proof. At shift n, the coefficient vector is

$$(\alpha_n - \beta_n)\vartheta(k),$$

where α_n and β_n are the two multiplicities. Since the multisets differ, some integer $c_n = \alpha_n - \beta_n$ is nonzero. The inequality $|c_n| \leq a + b < p$ shows that its image in $\mathbb{F}_p$ is nonzero. Since $\vartheta(k) \neq 0$, the corresponding vector is nonzero. Apply Lemma 3.6 with $D = a + b$. □

Proposition 3.8 (Characteristic-p resonance)**.** *Let $k \in \Delta_{2r}^*$ and $0 \leq n < N$. Then*

$$\frac{1}{t} \sum_{u \bmod t} Y_k(u + n)^p = 1 - \frac{\varepsilon_O}{t}, \qquad \varepsilon_O = \mathbf{1}_{\{O\in P_0+H\}}.$$

Consequently, along any family for which

$$\frac{t}{pq^{1/2}} \longrightarrow \infty,$$

there is no frequency-uniform degree-p estimate of the form $Cpq^{1/2}/t$. The restriction $D < p$ is therefore a genuine resonance barrier for this method.

Proof. At every finite orbit point,

$$Y_k(u+n)^p = \psi_1(f_k(P_{u+n}))^p = 1,$$

because ψ_1 takes values among the pth roots of unity. If the coset $P_0 + H$ contains O, there is exactly one value of u mod t for which $P_{u+n} = O$; at that value $Y_k(u+n) = 0$. If the coset does not contain O, there is no exceptional value. This proves the identity. Its left side tends to 1 in the displayed family, whereas $Cpq^{1/2}/t$ tends to 0. □

4. Mixed moments and compatible pairings

For $m, N \geq 1$, let

$$\mathfrak{D}_m(N) = \#\left\{(\mathbf{n}, \mathbf{m}) \in \{0, \dots, N-1\}^{2m} : \{n_1, \dots, n_m\} = \{m_1, \dots, m_m\}\right\}.$$

Equivalently,

$$\mathfrak{D}_m(N) = \sum_{r_0 + \cdots + r_{N-1} = m} \left(\frac{m!}{r_0! \cdots r_{N-1}!}\right)^2.$$

Lemma 4.1 (Diagonal combinatorics)**.** *For fixed m,*

$$\mathfrak{D}_m(N) = m! N^m + O_m(N^{m-1}),$$

and for all m, N,

$$\mathfrak{D}_m(N) \leq m! N^m.$$

Proof. Fix the first ordered tuple $\mathbf{n} = (n_1, \dots, n_m)$. Every ordered tuple $\mathbf{m}$ with the same multiset is obtained from $\mathbf{n}$ by a permutation of the m labels. Distinct permutations may coincide when $\mathbf{n}$ has repetitions, but there are never more than $m!$ resulting tuples. Summing over the N^m choices of $\mathbf{n}$ proves $\mathfrak{D}_m(N) \leq m! N^m$.

If $\mathbf{n}$ has pairwise distinct coordinates, all $m!$ permutations are distinct. The number of such tuples is

$$(N)_m = N(N-1)\cdots(N-m+1) = N^m + O_m(N^{m-1}).$$

For completeness, a union bound over pairs of positions gives

$$\#\{\mathbf{n} : \ n_i = n_j \text{ for some } i < j\} \leq \binom{m}{2} N^{m-1}.$$

Each repeated tuple has at most $m!$ partners. Hence the total contribution of all repeated tuples is $O_m(N^{m-1})$, while the distinct tuples contribute $m!(N)_m = m! N^m + O_m(N^{m-1})$. Adding the two classes proves the asymptotic. □

The next elementary observation records exactly what a balanced time assignment contributes; it removes all ambiguity caused by the convention $Y_k(O) = 0$.

Lemma 4.2 (Exact value of a balanced assignment)**.** *Use the notation of Lemma 3.6, and assume that $v_n = 0$ at every shift. Let*

$$T = \{n : \textit{at least one occurrence has shift } n\}.$$

Then

$$\frac{1}{t} \sum_{u \bmod t} \prod_{\ell=1}^{D} Y_{k_\ell}^{[\epsilon_\ell]}(u + n_\ell) = 1 - \varepsilon_O \frac{|T|}{t}.$$

In particular, the absolute difference from 1 is at most D/t.

Proof. At a starting index for which none of the points P_{u+n}, $n \in T$, equals O, the product is the additive character with phase coefficient $v_n = 0$ at every shift and hence equals 1. If $O \notin P_0 + H$, no exceptional starting index exists. If $O \in P_0 + H$, then for each $n \in T$ there is exactly one u mod t with $P_{u+n} = O$. These values of u are distinct because the shifts in $T \subset \{0, \dots, N-1\}$ are distinct modulo t. The completed product is 0 at precisely these $|T|$ starting indices. Averaging gives the formula. □

Theorem 4.3 (Mixed moments of one Walsh window)**.** *Let $a, b \geq 0$ and $1 \leq a + b < p$. If $a \neq b$, then*

$$\left| \mathbb{E}\left[Z_{N,k}^a \overline{Z_{N,k}}^b \right] \right| \leq 7(a+b) N^{a+b} \frac{q^{1/2}}{t}.$$

If $a = b = m$, then

$$\mathbb{E}\,|Z_{N,k}|^{2m} = \mathfrak{D}_m(N) + \mathcal{E}_m,$$

where

$$|\mathcal{E}_m| \leq \frac{m\mathfrak{D}_m(N)}{t} + 14mN^{2m}\frac{q^{1/2}}{t}.$$

In particular,

$$\mathbb{E}\,|Z_{N,k}|^{2m} \leq m!N^m + 14mN^{2m}\frac{q^{1/2}}{t}.$$

All bounds are uniform in $k \neq 0$.

Proof. Expand

$$Z_{N,k}^a \overline{Z_{N,k}}^b = \sum_{\mathbf{n} \in \{0,\dots,N-1\}^a} \sum_{\mathbf{m} \in \{0,\dots,N-1\}^b} \prod_{i=1}^{a} Y_k(U + n_i) \prod_{j=1}^{b} \overline{Y_k(U + m_j)}.$$

If $a \neq b$, the two shift multisets cannot be equal. Every one of the N^{a+b} summands is therefore bounded by Corollary 3.7, proving the first assertion.

Now let $a = b = m$. Exactly $\mathfrak{D}_m(N)$ assignments have equal shift multisets. Such an assignment is balanced, and Lemma 4.2 shows that its normalized correlation equals $1 - \varepsilon_O|T|/t$, with $|T| \leq m$. The total diagonal correction from $\mathfrak{D}_m(N)$ is therefore at most $m\mathfrak{D}_m(N)/t$ in absolute value. Every other assignment is unbalanced and has normalized correlation at most

$$7(2m)\frac{q^{1/2}}{t} = 14m\frac{q^{1/2}}{t}.$$

There are at most N^{2m} such assignments. This proves the error estimate. For the upper bound, note additionally that every balanced correlation is at most 1, and use Lemma 4.1. □

Definition 4.4 (Finite-order dissociativity)**.** Let $\mathbf{k} = (k_1, \dots, k_s)$ be nonzero Walsh frequencies and let $D < p$. We call $\mathbf{k}$ *D-dissociated* if

$$\sum_{j=1}^{s} c_j \vartheta(k_j) = 0, \qquad c_j \in \mathbb{Z}, \qquad \sum_{j=1}^{s} |c_j| \leq D,$$

implies $c_1 = \cdots = c_s = 0$. The integers act through their images in $\mathbb{F}_p$; the condition $D < p$ prevents a nonzero coefficient in this range from vanishing modulo p.

Theorem 4.5 (Joint moments for dissociated frequencies)**.** *Let $k_1, \dots, k_s \in \Delta_{2r}^*$, let $a_j, b_j \geq 0$, and put*

$$D = \sum_{j=1}^{s} (a_j + b_j), \qquad 1 \leq D < p.$$

Assume that $(k_1, \dots, k_s)$ is D-dissociated. If $a_j \neq b_j$ for at least one j, then

$$\left| \mathbb{E} \prod_{j=1}^{s} Z_{N,k_j}^{a_j} \overline{Z_{N,k_j}}^{b_j} \right| \leq 7DN^D \frac{q^{1/2}}{t}.$$

If $a_j = b_j = m_j$ for every j, then

$$\mathbb{E} \prod_{j=1}^{s} \left| Z_{N,k_j} \right|^{2m_j} = \prod_{j=1}^{s} \mathfrak{D}_{m_j}(N) + \mathcal{E},$$

where

$$|\mathcal{E}| \leq \frac{D}{t} \prod_{j=1}^{s} \mathfrak{D}_{m_j}(N) + 7DN^D \frac{q^{1/2}}{t}.$$

Proof. Expand the product and label all D occurrences. At a fixed shift n, let $c_{j,n}$ be the number of unbarred occurrences of k_j minus the number of barred occurrences of k_j. The coefficient vector at the translated pole is

$$v_n = \sum_{j=1}^{s} c_{j,n} \vartheta(k_j), \qquad \sum_j |c_{j,n}| \leq D.$$

By D-dissociativity, $v_n = 0$ if and only if every $c_{j,n} = 0$. Thus an assignment is balanced exactly when, for each j, the unbarred and barred shift multisets belonging to k_j agree.

If some $a_j \neq b_j$, no balanced assignment exists. Lemma 3.6 bounds each of the at most N^D assignments by $7Dq^{1/2}/t$.

If $a_j = b_j = m_j$ for all j, the balanced assignments are chosen independently for each frequency, so their number is $\prod_j \mathfrak{D}_{m_j}(N)$. Lemma 4.2 changes each balanced contribution from 1 by at most D/t. Every remaining assignment is bounded by Lemma 3.6. Summing the two classes gives the result. □

Definition 4.6 (Compatible Wick pairings)**.** Fix nonzero frequencies $k_1, \dots, k_s$ and nonnegative integers a_i, b_i. Form a labeled set of

$$D = \sum_{i=1}^{s} (a_i + b_i)$$

occurrences. Attach $\vartheta(k_i)$ to each of the a_i unbarred occurrences of k_i and $-\vartheta(k_i)$ to each of its b_i barred occurrences. A pair partition is *compatible* if the two vectors in every pair sum to zero. Let

$$\mathfrak{W}(\mathbf{a}, \mathbf{b}; \mathbf{k})$$

be the number of compatible pair partitions, with value 0 when D is odd.

Theorem 4.7 (General joint Wick asymptotics)**.** *Let $k_1, \dots, k_s \in \Delta_{2r}^*$ and $a_i, b_i \geq 0$. Put*

$$D = \sum_{i=1}^{s} (a_i + b_i), \qquad 1 \leq D < p.$$

Then

$$\mathbb{E} \prod_{i=1}^{s} Z_{N,k_i}^{a_i} \overline{Z_{N,k_i}}^{b_i} = \mathfrak{W}(\mathbf{a}, \mathbf{b}; \mathbf{k}) N^{D/2} + O_D\left(N^{D/2 - 1/2} + \frac{N^{D/2}}{t} + N^D \frac{q^{1/2}}{t} \right).$$

For odd D, the first term is 0.

Proof. Label the D occurrences and attach to each label the vector prescribed in Definition 4.6. Expanding the moment assigns to every label a time in $\{0,\dots,N-1\}$. The fibers of this assignment form a set partition π of the D labels. The assignment is balanced if and only if the sum of the attached vectors is zero in every block of π. If it is not balanced, Lemma 3.6 bounds its normalized correlation by $7Dq^{1/2}/t$. Since there are exactly N^D assignments, the total unbalanced contribution is

$$O_D\left(N^D\frac{q^{1/2}}{t}\right). \tag{4.1}$$

We now count balanced assignments. No balanced block can be a singleton, because every attached vector is nonzero. Thus a balanced partition with r blocks satisfies $r\le\lfloor D/2\rfloor$. If D is even and $r=D/2$, every block has size two, and balance says precisely that the partition is a compatible pair partition. Fix such a pair partition. Choosing distinct times for its $D/2$ labeled blocks gives

$$(N)_{D/2}=N^{D/2}+O_D(N^{D/2-1})$$

assignments. Different pair partitions give disjoint assignment sets, because the equal-time fibers recover the pair partition. Their total is therefore

$$\mathfrak{W}(\mathbf{a},\mathbf{b};\mathbf{k})N^{D/2}+O_D(N^{D/2-1}). \tag{4.2}$$

Every other balanced partition has at most $D/2-1$ blocks when D is even, and at most $(D-1)/2$ blocks when D is odd. For a fixed partition with r blocks, the number of assignments inducing it is the falling factorial $(N)_r\le N^r$. The number of set partitions of a fixed D-element set is the finite Bell number B_D. Hence all nonleading balanced assignments together number

$$O_D(N^{D/2-1})\quad(D\text{ even}),\qquad O_D(N^{(D-1)/2})\quad(D\text{ odd}),$$

which is uniformly $O_D(N^{D/2-1/2})$ for $N\ge1$.

Finally, Lemma 4.2 says that the normalized correlation of a balanced assignment differs from 1 by at most D/t. The total number of balanced assignments is at most $B_DN^{\lfloor D/2\rfloor}=O_D(N^{D/2})$, so all point-at-infinity corrections contribute

$$O_D\left(\frac{N^{D/2}}{t}\right). \tag{4.3}$$

Combining (4.1), (4.2), and (4.3) proves the theorem. When D is odd there is no pair partition, so the main term is zero by definition. □

Corollary 4.8 (Mean and second moment)**.** *For every $k\neq0$,*

$$|\mu_k|\le7\frac{q^{1/2}}{t},$$

and

$$\mathbb{E}\,|Z_{N,k}|^2=N+O\left(\frac{N}{t}+\frac{N^2q^{1/2}}{t}\right).$$

Proof. The mean estimate is Lemma 3.6 with $D=1$ and $N=1$. The second formula is Theorem 4.3 with $m=1$, for which $\mathfrak{D}_1(N)=N$. □

Proposition 4.9 (Finite-window covariance and pseudo-covariance)**.** *For nonzero frequencies k,ℓ,*

$$\mathbb{E}[Z_{N,k}\overline{Z_{N,\ell}}]=N\mathbf{1}_{\{k=\ell\}}+O\left(\frac{N}{t}+\frac{N^2q^{1/2}}{t}\right),$$

and

$$\mathbb{E}[Z_{N,k}Z_{N,\ell}] = N\mathbf{1}_{\{k=\ominus\ell\}} + O\left(\frac{N}{t} + \frac{N^2q^{1/2}}{t}\right).$$

With $\widetilde{Z}_{N,k} = Z_{N,k} - N\mu_k$,

$$\mathbb{E}[\widetilde{Z}_{N,k}\overline{\widetilde{Z}_{N,\ell}}] = N\mathbf{1}_{\{k=\ell\}} + O\left(\frac{N}{t} + \frac{N^2q^{1/2}}{t} + \frac{N^2q}{t^2}\right),$$

$$\mathbb{E}[\widetilde{Z}_{N,k}\widetilde{Z}_{N,\ell}] = N\mathbf{1}_{\{k=\ominus\ell\}} + O\left(\frac{N}{t} + \frac{N^2q^{1/2}}{t} + \frac{N^2q}{t^2}\right).$$

Proof. Expand each second moment over pairs of shifts. In the conjugated moment, a same-time pair has coefficient $\vartheta(k) - \vartheta(\ell)$ and is balanced exactly when $k = \ell$. In the unconjugated moment, the coefficient is $\vartheta(k) + \vartheta(\ell)$ and is balanced exactly when $k = \ominus\ell$. Lemma 4.2 gives $N + O(N/t)$ from the balanced same-time pairs. Every other pair is controlled by Lemma 3.6 with $D = 2$, yielding $O(N^2q^{1/2}/t)$. Centering subtracts $N^2\mu_k\overline{\mu_\ell}$ or $N^2\mu_k\mu_\ell$, whose modulus is $O(N^2q/t^2)$ by Corollary 4.8. □

Proposition 4.10 (Variance of a finite Walsh polynomial)**.** *Let* $\mathcal{K} \subset \Delta_{2r}^*$ *be finite, let* $(c_k)_{k\in\mathcal{K}} \subset \mathbb{C}$, *and define*

$$\mathcal{T}_{N,c} = \sum_{k\in\mathcal{K}} c_k(M_{N,k} - \mu_k).$$

Then

$$\mathbb{E}\,|\mathcal{T}_{N,c}|^2 = \frac{\|c\|_2^2}{N} + O\left(\|c\|_1^2\left[\frac{1}{Nt} + \frac{q^{1/2}}{t} + \frac{q}{t^2}\right]\right),$$

where

$$\|c\|_1 = \sum_{k\in\mathcal{K}} |c_k|, \qquad \|c\|_2^2 = \sum_{k\in\mathcal{K}} |c_k|^2.$$

Consequently, for $x > 0$,

$$\mathbb{P}(|\mathcal{T}_{N,c}| \geq x) \leq \frac{\|c\|_2^2/N + C\,\|c\|_1^2\,(1/(Nt) + q^{1/2}/t + q/t^2)}{x^2}.$$

Proof. Since $\mathcal{T}_{N,c} = N^{-1}\sum_k c_k\widetilde{Z}_{N,k}$,

$$\mathbb{E}\,|\mathcal{T}_{N,c}|^2 = \frac{1}{N^2}\sum_{k,\ell\in\mathcal{K}} c_k\overline{c_\ell}\,\mathbb{E}[\widetilde{Z}_{N,k}\overline{\widetilde{Z}_{N,\ell}}].$$

The diagonal main terms in Proposition 4.9 sum to $N^{-1}\|c\|_2^2$. Taking absolute values in the error terms costs at most $\sum_{k,\ell}|c_kc_\ell| = \|c\|_1^2$, and division by N^2 gives the stated error. Chebyshev's inequality proves the tail bound. □

5. Arithmetic concentration inequalities

For $m \geq 1$, define

$$\mathfrak{a}_m(q,t) = \left(C_{\mathrm{ar}}m\frac{q^{1/2}}{t}\right)^{1/(2m)} = \left(14m\frac{q^{1/2}}{t}\right)^{1/(2m)}.$$

Theorem 5.1 (Moment-norm concentration scale)**.** *Let $1 \le m < p/2$. Uniformly in $k \neq 0$,*

$$\|M_{N,k} - \mu_k\|_{2m} \le \sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t}.$$

Consequently,

$$\mathbb{P}\left(|M_{N,k} - \mu_k| \ge \mathrm{e}\left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t}\right]\right) \le \mathrm{e}^{-2m}.$$

Proof. The upper bound in Theorem 4.3 gives

$$\|Z_{N,k}\|_{2m}^{2m} \le m!N^m + 14mN^{2m}\frac{q^{1/2}}{t}.$$

For nonnegative A, B and $r \ge 1$, $(A+B)^{1/r} \le A^{1/r} + B^{1/r}$. Hence

$$\|Z_{N,k}\|_{2m} \le (m!)^{1/(2m)}N^{1/2} + N\left(14m\frac{q^{1/2}}{t}\right)^{1/(2m)}.$$

The elementary bound $m! \le m^m$ gives $(m!)^{1/(2m)} \le \sqrt{m}$. Dividing by N and applying Minkowski's inequality to $M_{N,k} - \mu_k = N^{-1}Z_{N,k} - \mu_k$, together with $|\mu_k| \le 7q^{1/2}/t$ from Corollary 4.8, proves the norm estimate. Markov's inequality at e times the L^{2m} norm yields

$$\mathbb{P}(|X| \ge \mathrm{e}\,\|X\|_{2m}) \le \mathrm{e}^{-2m}.$$

□

Corollary 5.2 (Simultaneous finite Walsh battery)**.** *Let $\varnothing \neq \mathcal{K} \subset \Delta_{2r}^*$ be finite and let $1 \le m < p/2$. Then*

$$\mathbb{P}\left(\max_{k\in\mathcal{K}}|M_{N,k} - \mu_k| \ge \mathrm{e}\left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t}\right]\right) \le |\mathcal{K}|\mathrm{e}^{-2m}.$$

If additionally $m \le N$ and

$$14m\frac{q^{1/2}}{t} \le \left(\frac{m}{N}\right)^m, \tag{5.1}$$

then

$$\mathbb{P}\left(\max_{k\in\mathcal{K}}|M_{N,k} - \mu_k| \ge 3\mathrm{e}\sqrt{\frac{m}{N}}\right) \le |\mathcal{K}|\mathrm{e}^{-2m}.$$

If moreover $2m > \log(t|\mathcal{K}|)$, the displayed bound holds for every cyclic starting index U; that is, its exceptional set is empty.

Proof. The first assertion follows from Theorem 5.1 and the union bound. Under (5.1),

$$\mathfrak{a}_m(q,t) \le \sqrt{\frac{m}{N}}.$$

Writing $x = m/N \le 1$, the same hypothesis gives

$$7\frac{q^{1/2}}{t} \le \frac{1}{2m}x^m \le \frac{1}{2}\sqrt{x}.$$

Thus the norm scale in Theorem 5.1 is at most $\frac{5}{2}\sqrt{m/N} < 3\sqrt{m/N}$. This proves the second assertion. Finally, the exceptional probability is a multiple of $1/t$. If $|\mathcal{K}|\mathrm{e}^{-2m} < 1/t$, it must be zero. □

Corollary 5.3 (Confidence-optimized battery)**.** *Let $\varnothing \neq \mathcal{K} \subset \Delta_{2r}^*$ and $0 < \delta < 1$, and put*

$$m_{\mathcal{K},\delta} = \left\lceil \frac{1}{2} \log \frac{|\mathcal{K}|}{\delta} \right\rceil.$$

Assume

$$1 \le m_{\mathcal{K},\delta} \le N, \qquad 2m_{\mathcal{K},\delta} < p,$$

and that (5.1) holds with $m = m_{\mathcal{K},\delta}$. Then

$$\mathbb{P}\left(\max_{k \in \mathcal{K}} |M_{N,k} - \mu_k| \ge 3\mathrm{e}\sqrt{\frac{1 + \log(|\mathcal{K}|/\delta)}{N}} \right) \le \delta.$$

If $2m_{\mathcal{K},\delta} > \log(t|\mathcal{K}|)$, the same bound holds deterministically for all $U \in \mathbb{Z}/t\mathbb{Z}$.

Proof. By the definition of the ceiling, $|\mathcal{K}|\mathrm{e}^{-2m_{\mathcal{K},\delta}} \le \delta$ and $m_{\mathcal{K},\delta} \le 1 + \frac{1}{2}\log(|\mathcal{K}|/\delta)$. Apply Corollary 5.2; enlarging the square-root numerator from $m_{\mathcal{K},\delta}$ to $1 + \log(|\mathcal{K}|/\delta)$ only weakens the threshold. The all-seed assertion is the last part of that corollary. □

Corollary 5.4 (Finite Walsh polynomials)**.** *Let $\mathcal{K} \subset \Delta_{2r}^*$ be finite and let $(c_k)_{k \in \mathcal{K}} \subset \mathbb{C}$. For*

$$\mathcal{T}_{N,c} = \sum_{k \in \mathcal{K}} c_k (M_{N,k} - \mu_k),$$

one has, for $1 \le m < p/2$,

$$\|\mathcal{T}_{N,c}\|_{2m} \le \|c\|_1 \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t} \right].$$

Consequently,

$$\mathbb{P}\left(|\mathcal{T}_{N,c}| \ge \mathrm{e}\,\|c\|_1 \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t} \right] \right) \le \mathrm{e}^{-2m}.$$

Proof. Minkowski's inequality gives

$$\|\mathcal{T}_{N,c}\|_{2m} \le \sum_{k \in \mathcal{K}} |c_k| \, \|M_{N,k} - \mu_k\|_{2m}.$$

Insert Theorem 5.1. Markov's inequality gives the tail estimate. □

Theorem 5.5 (Interpolation between ℓ^2 and ℓ^1 coefficient scales)**.** *Let $\mathcal{T}_{N,c}$ be as above and choose $2 \le m < p/2$. Put*

$$A_2(c) = \left[\frac{\|c\|_2^2}{N} + C\,\|c\|_1^2 \left(\frac{1}{Nt} + \frac{q^{1/2}}{t} + \frac{q}{t^2} \right) \right]^{1/2}$$

and

$$A_{2m}(c) = \|c\|_1 \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t} \right],$$

where the absolute constant C is chosen to dominate the error in Proposition 4.10. For $1 \le s \le m$, let

$$\theta_{s,m} = \frac{m - s}{s(m-1)}.$$

Then

$$\|\mathcal{T}_{N,c}\|_{2s} \le A_2(c)^{\theta_{s,m}} A_{2m}(c)^{1-\theta_{s,m}},$$

and

$$\mathbb{P}\left(|\mathcal{T}_{N,c}| \ge \mathrm{e} A_2(c)^{\theta_{s,m}} A_{2m}(c)^{1-\theta_{s,m}} \right) \le \mathrm{e}^{-2s}.$$

Proof. Proposition 4.10 gives $\|\mathcal{T}_{N,c}\|_2 \le A_2(c)$, and Corollary 5.4 gives $\|\mathcal{T}_{N,c}\|_{2m} \le A_{2m}(c)$. Since

$$\frac{1}{2s} = \frac{\theta_{s,m}}{2} + \frac{1-\theta_{s,m}}{2m},$$

log-convexity of L^p norms proves the interpolation estimate. Markov's inequality gives the last display. □

Corollary 5.6 (Finite-order subgaussian regime)**.** *Assume $1 \le m \le N$, $2m < p$, and (5.1). Then*

$$\mathbb{P}\left(|M_{N,k} - \mu_k| \ge 3\mathrm{e}\sqrt{\frac{m}{N}}\right) \le \mathrm{e}^{-2m}.$$

Proof. This is Corollary 5.2 with a one-element family. □

Corollary 5.7 (Subgaussian deviation window)**.** *There are absolute constants $c_0, c_1, C_0 > 0$ with the following property. Let $0 < x \le 1$, assume $x^2 N \ge C_0$, and put*

$$m = \lfloor c_0 x^2 N \rfloor.$$

If $2m < p$ and (5.1) holds, then

$$\mathbb{P}\left(|M_{N,k} - \mu_k| \ge x\right) \le 2\exp(-c_1 x^2 N).$$

Proof. Choose $c_0 > 0$ so that $3\mathrm{e}\sqrt{c_0} \le 1/2$. For C_0 sufficiently large, $m \ge 1$ and $m \ge (c_0/2)x^2 N$. The threshold in Corollary 5.6 is at most x, while

$$\mathrm{e}^{-2m} \le \exp(-c_0 x^2 N) \le 2\exp(-c_1 x^2 N)$$

for a suitable absolute c_1. □

Remark 5.8. The condition (5.1) compares the arithmetic off-diagonal term with the Gaussian diagonal at the selected moment order. For fixed x, subject also to $p \gg x^2 N$, it allows exponential seed tails only while the logarithmic period gap can support $m \asymp x^2 N$. The moment theorem remains valid beyond that range, but its arithmetic term limits the confidence obtainable by direct moment optimization.

6. Concentration inequality of random-window discrepancy

For a nonempty finite multiset $\mathcal{X} \subset [0,1)^d$, define its extreme discrepancy by

$$D(\mathcal{X}) = \sup_B \left| \frac{\#(\mathcal{X} \cap B)}{\#\mathcal{X}} - \lambda_d(B) \right|,$$

where B ranges over axis-parallel half-open boxes. Its star discrepancy $D^*(\mathcal{X})$ is obtained by restricting to boxes anchored at the origin. Thus $D^*(\mathcal{X}) \le D(\mathcal{X})$.

Put $M = p^a$ and

$$\mathcal{G}_M = \{0, 1/M, \dots, (M-1)/M\}.$$

The use of Walsh coefficients for discrepancy follows the general framework of Hellekalek [11]; for completeness, we prove below the exact finite-grid coefficient bound needed here.

For a grid interval $I^\sharp = [A/M, B/M)$ and a one-dimensional Walsh frequency $0 \le h < M$, define

$$\widehat{\mathbf{1}}_{I^\sharp}(h) = M^{-1} \sum_{x \in \mathcal{G}_M} \mathbf{1}_{I^\sharp}(x)\overline{w_h(x)}, \qquad \rho_1(h) = \sup_{I^\sharp} \left|\widehat{\mathbf{1}}_{I^\sharp}(h)\right|.$$

Lemma 6.1 (One-dimensional finite Walsh weights)**.** *Write*

$$h = \sum_{\nu=1}^{a} h_\nu p^{\nu-1}, \qquad h_\nu \in \{0, \dots, p-1\}.$$

Then $\rho_1(0) = 1$. If $h \ne 0$ and

$$g = g(h) = \max\{\nu : h_\nu \ne 0\},$$

then

$$\rho_1(h) \le \frac{4}{p^g \sin(\pi h_g/p)}.$$

Consequently, for $p \ge 5$,

$$\sum_{h=0}^{M-1} \rho_1(h) \le 1 + 7\log M.$$

Proof. The identity $\rho_1(0) = 1$ follows by taking the full interval. Let $h \neq 0$, let $g = g(h)$, and put $L = p^{a-g}$. As n runs through $0, \dots, M-1$, the value $\overline{w_h(n/M)}$ is constant on each consecutive block of L integers. Within any string of p consecutive such blocks for which the first $g-1$ digits are fixed, the block values form, up to a common factor, the progression

$$1, \omega^{-h_g}, \dots, \omega^{-(p-1)h_g}, \qquad \omega = e^{2\pi i/p},$$

and hence sum to 0.

Consider an arbitrary integer interval $[A, B)$. At its two endpoints there are at most two incomplete blocks of length L; their combined contribution has modulus at most $2L$. Between those endpoint pieces lies a consecutive interval of complete L-blocks. Split its block indices into maximal complete strings of p consecutive blocks and discard those strings, whose Walsh values sum to zero. The undiscarded complete blocks form at most two consecutive runs, each containing between 1 and $p-1$ blocks. For any such run,

$$\left| \sum_{v=s}^{s+r-1} \omega^{-h_g v} \right| = \left| \omega^{-h_g s} \frac{1-\omega^{-h_g r}}{1-\omega^{-h_g}} \right| \le \frac{2}{|1-\omega^{-h_g}|} = \frac{1}{\sin(\pi h_g/p)} \qquad (1 \le r \le p-1).$$

Each block has L equal Walsh values. The complete-block contribution is therefore at most $2L/\sin(\pi h_g/p)$. Adding the two incomplete endpoint blocks and using $\sin(\pi h_g/p) \le 1$ gives

$$\left| \sum_{n=A}^{B-1} \overline{w_h(n/M)} \right| \le 2L + \frac{2L}{\sin(\pi h_g/p)} \le \frac{4L}{\sin(\pi h_g/p)}.$$

Division by $M = p^a$ proves the coefficient bound.

For a fixed level g, there are p^{g-1} choices of the preceding digits and $p-1$ choices for h_g. Therefore

$$\sum_{h:g(h)=g} \rho_1(h) \le \frac{4}{p} \sum_{r=1}^{p-1} \frac{1}{\sin(\pi r/p)}.$$

By symmetry and the concavity estimate $\sin(\pi r/p) \ge 2r/p$ for $1 \le r \le p/2$,

$$\sum_{r=1}^{p-1} \frac{1}{\sin(\pi r/p)} \le p\, H_{\lfloor p/2 \rfloor} \le p(1+\log p).$$

Since $p \ge 5$, $4(1+\log p) \le 7\log p$. Summing over $g = 1, \dots, a$ yields

$$\sum_{h=0}^{M-1} \rho_1(h) \le 1 + 7a\log p = 1 + 7\log M.$$

□

For a grid box $B^\sharp = I_1^\sharp \times \cdots \times I_d^\sharp$ and $k = (k_1, \dots, k_d) \in \{0, \dots, M-1\}^d$, put

$$\widehat{\mathbf{1}}_{B^\sharp}(k) = M^{-d} \sum_{x \in \mathcal{G}_M^d} \mathbf{1}_{B^\sharp}(x) \overline{w_k(x)}, \qquad \rho_d(k) = \sup_{B^\sharp} \left| \widehat{\mathbf{1}}_{B^\sharp}(k) \right|.$$

Lemma 6.2 (Tensor factorization of the weights)**.** *For $k = (k_1, \dots, k_d)$,*

$$\rho_d(k) = \prod_{j=1}^{d} \rho_1(k_j).$$

Consequently,

$$\sum_{k \in \{0,\dots,M-1\}^d} \rho_d(k) = \left(\sum_{h=0}^{M-1} \rho_1(h) \right)^d \le (1 + 7 \log M)^d.$$

Since $\rho_d(0) = 1$,

$$W_d := \sum_{k \neq 0} \rho_d(k) \le (1 + 7 \log M)^d - 1.$$

Proof. The indicator and Walsh character both factor coordinatewise, so

$$\widehat{\mathbf{1}}_{B^\sharp}(k) = \prod_{j=1}^{d} \widehat{\mathbf{1}}_{I_j^\sharp}(k_j).$$

The intervals in the different coordinates can be chosen independently. Taking suprema of absolute values gives the first identity. Summing the product over all k gives the second identity, and Lemma 6.1 gives the bound. □

For a multiset $\mathcal{X} \subset \mathcal{G}_M^d$, write

$$S(w_k, \mathcal{X}) = \frac{1}{\#\mathcal{X}} \sum_{x \in \mathcal{X}} w_k(x).$$

Lemma 6.3 (Finite-grid Walsh discrepancy inequality)**.** *Let $\mathcal{X}$ be a nonempty finite multiset in $\mathcal{G}_M^d$. Then*

$$D(\mathcal{X}) \le \Delta_{p,a,d} + \sum_{k \neq 0} \rho_d(k) \, |S(w_k, \mathcal{X})| \,, \qquad \Delta_{p,a,d} = 1 - (1 - M^{-1})^d.$$

Proof. Let $B = \prod_{j=1}^{d} [\alpha_j, \beta_j)$ be an arbitrary half-open box and set

$$\alpha_j^\sharp = \frac{\lceil M\alpha_j \rceil}{M}, \qquad \beta_j^\sharp = \frac{\lceil M\beta_j \rceil}{M}, \qquad B^\sharp = \prod_{j=1}^{d} [\alpha_j^\sharp, \beta_j^\sharp).$$

For a grid point $x_j = n/M$, the conditions $x_j \ge \alpha_j$ and $x_j < \beta_j$ are equivalent to $x_j \ge \alpha_j^\sharp$ and $x_j < \beta_j^\sharp$, respectively. Hence $B \cap \mathcal{G}_M^d = B^\sharp \cap \mathcal{G}_M^d$.

Let $u_j = \beta_j - \alpha_j$ and $v_j = \beta_j^\sharp - \alpha_j^\sharp$. Then $|u_j - v_j| \le M^{-1}$. We use the elementary inequality

$$\left| \prod_{j=1}^{d} u_j - \prod_{j=1}^{d} v_j \right| \le 1 - (1 - \delta)^d \quad (0 \le u_j, v_j \le 1, \ |u_j - v_j| \le \delta). \tag{7.1}$$

To verify it, suppose without loss that $\prod u_j \ge \prod v_j$. Replacing v_j by $\max(u_j - \delta, 0)$ can only decrease the second product. The difference $\prod u_j - \prod (u_j - \delta)_+$ is nondecreasing in each u_j and is therefore at most its value at $u_1 = \cdots = u_d = 1$, namely $1 - (1 - \delta)^d$. Interchanging u and v proves the other case. Applying (7.1) with $\delta = M^{-1}$ gives

$$|\lambda_d(B) - \lambda_d(B^\sharp)| \le \Delta_{p,a,d}.$$

Finite Walsh inversion on the group $\mathcal{G}_M^d$ gives, for every grid point x,

$$\mathbf{1}_{B^\sharp}(x) = \sum_{k} \widehat{\mathbf{1}}_{B^\sharp}(k) w_k(x).$$

The zero coefficient is $\widehat{\mathbf{1}}_{B^\sharp}(0) = \lambda_d(B^\sharp)$. Averaging the inversion formula over $\mathcal{X}$, removing the zero frequency, and taking absolute values yield

$$\left| \frac{\#(\mathcal{X} \cap B^\sharp)}{\#\mathcal{X}} - \lambda_d(B^\sharp) \right| \le \sum_{k \neq 0} \rho_d(k) |S(w_k, \mathcal{X})|.$$

The grid intersections of B and $B^\sharp$ agree. Combining the last two displays and taking the supremum over B proves the lemma. □

Extend the output map by $\Phi^\dagger(O) = 0$, and define the completed window multiset

$$\mathcal{X}_{U,N} = \big(\Phi^\dagger(P_U), \dots, \Phi^\dagger(P_{U+N-1})\big), \qquad D_{U,N} = D(\mathcal{X}_{U,N}).$$

Theorem 6.4 (Random-window discrepancy concentration)**.** *Let $m \ge 1$ satisfy $2m < p$. Then*

$$\mathbb{P}\Bigg(D_{U,N} > \Delta_{p,a,d} + W_d\left(7\frac{q^{1/2}}{t} + \frac{1}{N}\right) + \mathrm{e} W_d\left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t}\right] \Bigg) \le \mathrm{e}^{-2m}.$$

If an occurrence of O is deleted instead of being mapped to 0, the resulting discrepancy differs from $D_{U,N}$ by at most $1/N$ whenever $N \ge 2$.

Proof. For $k \neq 0$, let

$$A_{N,k}(U) = \frac{1}{N} \sum_{n=0}^{N-1} w_k(\Phi^\dagger(P_{U+n})).$$

A length-N window contains O at most once because $N \le t$. At finite points the summand equals $Y_k(U+n)$, whereas at O the completed Walsh value is $w_k(0) = 1$ and Y_k is 0. Hence

$$|A_{N,k}(U) - M_{N,k}(U)| \le \frac{1}{N}.$$

Lemma 6.3, the triangle inequality, and $|\mu_k| \le 7q^{1/2}/t$ give the pointwise bound

$$\begin{aligned} D_{U,N} &\le \Delta_{p,a,d} + \sum_{k \neq 0} \rho_d(k) |A_{N,k}(U)| \\ &\le \Delta_{p,a,d} + W_d\left(7\frac{q^{1/2}}{t} + \frac{1}{N}\right) + \sum_{k \neq 0} \rho_d(k) |M_{N,k}(U) - \mu_k|. \end{aligned}$$

By Minkowski's inequality and Theorem 5.1, the L^{2m} norm of the final random sum is at most

$$W_d\left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t}\right].$$

Markov's inequality proves the probability estimate.

For the last assertion, suppose the completed empirical measure has N atoms and one of them is the completed point corresponding to O. If a box contains a of the other $N-1$ atoms and has indicator $\iota \in \{0,1\}$ at the completed point, then

$$\left| \frac{a+\iota}{N} - \frac{a}{N-1} \right| \le \frac{1}{N}.$$

Taking suprema over boxes proves the discrepancy comparison. □

Theorem 6.5 (Moment and expectation bounds for discrepancy)**.** *Put*

$$B_{0,N} = \Delta_{p,a,d} + W_d \left(7\frac{q^{1/2}}{t} + \frac{1}{N} \right).$$

For $2m < p$*,*

$$\|(D_{U,N} - B_{0,N})_+\|_{2m} \le W_d \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t} \right].$$

In particular,

$$\mathbb{E}D_{U,N} \le \Delta_{p,a,d} + W_d \left[\frac{1}{\sqrt{N}} + \mathfrak{a}_1(q,t) + 14\frac{q^{1/2}}{t} + \frac{1}{N} \right].$$

Proof. The pointwise estimate in the proof of Theorem 6.4 gives

$$(D_{U,N} - B_{0,N})_+ \le \sum_{k\neq 0} \rho_d(k) \, |M_{N,k} - \mu_k| \, .$$

Minkowski's inequality and Theorem 5.1 give the asserted L^{2m} estimate. For $m = 1$, use $\mathbb{E}X \le \|X\|_2$ for a nonnegative random variable X and add $B_{0,N}$; the two occurrences of $7q^{1/2}/t$ combine to $14q^{1/2}/t$. □

Corollary 6.6 (Subgaussian seed-quality bound)**.** *Assume* $1 \le m \le N$*,* $2m < p$*, and* (5.1)*. Then*

$$\mathbb{P}\left(D_{U,N} > \Delta_{p,a,d} + W_d \left[7\frac{q^{1/2}}{t} + \frac{1}{N} + 3\mathrm{e}\sqrt{\frac{m}{N}} \right] \right) \le \mathrm{e}^{-2m}.$$

Proof. Under (5.1), the random L^{2m} scale in Theorem 6.4 is at most $3W_d\sqrt{m/N}$ by the calculation in Corollary 5.2. □

Corollary 6.7 (Random-shift integration error)**.** *Let* $g : [0,1]^d \to \mathbb{R}$ *have bounded Hardy–Krause variation and put*

$$Q_{U,N}(g) = \frac{1}{N} \sum_{n=0}^{N-1} g(\Phi^\dagger(P_{U+n})).$$

Under the hypotheses of Theorem 6.4,

$$\mathbb{P}\left(\left| Q_{U,N}(g) - \int_{[0,1]^d} g(x)\, dx \right| > V_{\mathrm{HK}}(g)\, T_{m,N} \right) \le \mathrm{e}^{-2m},$$

where

$$\begin{aligned} T_{m,N} =& \Delta_{p,a,d} + W_d \left(7\frac{q^{1/2}}{t} + \frac{1}{N} \right) \\ &+ \mathrm{e}W_d \left[\sqrt{\frac{m}{N}} + \mathfrak{a}_m(q,t) + 7\frac{q^{1/2}}{t} \right]. \end{aligned}$$

Proof. The Koksma–Hlawka inequality bounds the integration error by $V_{\mathrm{HK}}(g)D^*(\mathcal{X}_{U,N})$; see, for example, [12, Chapter 2]. Since $D^* \le D$, Theorem 6.4 gives the result. □

7. Arithmetic concentration depth

The finite-order subgaussian estimate is controlled by the largest moment order for which the Gaussian diagonal dominates the arithmetic off-diagonal term. Put

$$R_{\mathrm{ar}} = \frac{t}{q^{1/2}}$$

and define

$$m_{\mathrm{eff}}(N,q,t,p) = \max\left\{1 \le m \le \min\left(N, \left\lfloor \frac{p-1}{2} \right\rfloor\right) : R_{\mathrm{ar}} \ge 14m\left(\frac{N}{m}\right)^m\right\},$$

with $m_{\mathrm{eff}} = 0$ when the set is empty.

Lemma 7.1 (Logarithmic form of the dominance condition)**.** *For $1 \le m \le N$,*

$$R_{\mathrm{ar}} \ge 14m(N/m)^m$$

is equivalent to

$$(m-1)\log\frac{N}{m} \le \log\frac{R_{\mathrm{ar}}}{14N}.$$

Proof. After division by $14N$,

$$\frac{14m(N/m)^m}{14N} = \frac{m}{N}\left(\frac{N}{m}\right)^m = \left(\frac{N}{m}\right)^{m-1}.$$

Both sides are positive, so taking logarithms gives the equivalence. □

Theorem 7.2 (Effective concentration profile)**.** *Assume $m_{\mathrm{eff}} \ge 1$. Uniformly in $k \ne 0$,*

$$\mathbb{P}\left(|M_{N,k} - \mu_k| \ge 3\mathrm{e}\sqrt{\frac{m_{\mathrm{eff}}}{N}}\right) \le \exp(-2m_{\mathrm{eff}}).$$

Moreover,

$$\mathbb{P}\left(D_{U,N} > \Delta_{p,a,d} + W_d\left[7\frac{q^{1/2}}{t} + \frac{1}{N} + 3\mathrm{e}\sqrt{\frac{m_{\mathrm{eff}}}{N}}\right]\right) \le \exp(-2m_{\mathrm{eff}}).$$

Proof. By definition, m_{eff} satisfies the characteristic, window-length, and arithmetic-dominance hypotheses of Corollaries 5.6 and 6.6. Apply those two corollaries. □

Lemma 7.3 (Lower bounds for the effective order)**.** *Let*

$$\Lambda = \log\frac{R_{\mathrm{ar}}}{14N}.$$

There are absolute constants $c, C > 0$ such that the following hold.

(i) *If $N \ge 8$, $C \le \Lambda \le N/4$, and*

$$p \ge C\frac{\Lambda}{\log(eN/\Lambda)},$$

then

$$m_{\mathrm{eff}} \ge c\frac{\Lambda}{\log(eN/\Lambda)}.$$

(ii) *If $N \ge 8$, $\Lambda \ge N/2$, and $p \ge N$, then*

$$m_{\mathrm{eff}} \ge \left\lfloor \frac{N}{8} \right\rfloor.$$

Proof. For (i), put

$$L = \log(eN/\Lambda), \qquad R = \frac{\Lambda}{L}.$$

Because $\Lambda > 0$, the order $m = 1$ satisfies the logarithmic dominance condition, and the standing assumption $p \geq 5$ makes it characteristic-admissible. Thus $m_{\mathrm{eff}} \geq 1$. If R is bounded by a sufficiently large absolute constant, this already implies $m_{\mathrm{eff}} \geq cR$ after choosing $c > 0$ small enough.

It remains to consider R larger than that constant. Set

$$m = \lfloor c_0 R \rfloor$$

with $c_0 > 0$ sufficiently small; then $m \geq c_0 R/2 \geq 1$. Since $\Lambda \leq N/4$, one has $L \geq \log(4e)$. Moreover,

$$\log \frac{N}{m} \leq \log \frac{eN}{\Lambda} + \log \frac{2L}{c_0} \leq C_1 L$$

for an absolute C_1. Therefore

$$(m-1) \log \frac{N}{m} \leq C_1 c_0 \Lambda \leq \Lambda$$

when $c_0 \leq 1/C_1$. The stated lower bound on p, with its absolute constant chosen sufficiently large relative to c_0, ensures $2m < p$. Lemma 7.1 makes m admissible, and $m \geq (c_0/2)R$ proves the claim after renaming the constant.

For (ii), take $m = \lfloor N/8 \rfloor$. If $8 \leq N < 16$, then $m = 1$ and the logarithmic left side is 0. If $N \geq 16$, then $m \geq N/16$, and hence

$$(m-1) \log \frac{N}{m} \leq \frac{N}{8} \log 16 < \frac{N}{2} \leq \Lambda.$$

Moreover $2m \leq N/4 < p$ because $p \geq N$. Thus m is admissible. □

Corollary 7.4 (Long-period logarithmic-window concentration)**.** *Fix $\eta > 0$ and suppose*

$$t \geq q^{1/2+\eta}.$$

There are constants $c_0, c_1, C > 0$ such that, for $0 < x \leq 1$, if

$$x^2 N \geq C, \qquad p \geq c_0 x^2 N, \qquad x^2 N \log(e/x) \leq c_1 \eta \log q,$$

then, for all sufficiently large q in terms of η,

$$\mathbb{P}(|M_{N,k} - \mu_k| \geq x) \leq 2 \exp(-c_1 x^2 N).$$

The same hypotheses give

$$\mathbb{P}\left(D_{U,N} > \Delta_{p,a,d} + W_d \left[7 \frac{q^{1/2}}{t} + \frac{1}{N} + x \right] \right) \leq 2 \exp(-c_1 x^2 N),$$

after changing the absolute constants.

Proof. Choose $m = \lfloor c_2 x^2 N \rfloor$ with $c_2 > 0$ small enough for the threshold in Corollary 5.6 to be at most x. The assumptions ensure $m \geq 1$, $m \leq N$, and $2m < p$. Since $R_{\mathrm{ar}} \geq q^{\eta}$,

$$\log R_{\mathrm{ar}} \geq \eta \log q.$$

On the other hand,

$$\log\left(14m(N/m)^m\right) = \log(14m) + m \log(N/m) \leq C_2 x^2 N \log(e/x).$$

For c_1 sufficiently small, the right side is at most $\eta \log q$, so the arithmetic dominance condition holds. Apply Corollary 5.7 for the Walsh statistic and Corollary 6.6 for discrepancy, adjusting constants to replace the latter's threshold by x. □

Corollary 7.5 (Uniform certification of every cyclic shift)**.** *Suppose an integer m satisfies $1 \le m \le N$, $2m < p$, the dominance condition (5.1), and*

$$2m > \log t.$$

Then every $U \in \mathbb{Z}/t\mathbb{Z}$ satisfies

$$|M_{N,k}(U) - \mu_k| < 3\mathrm{e}\sqrt{\frac{m}{N}},$$

and every U satisfies the discrepancy bound in Corollary 6.6.

Proof. Each exceptional probability is at most $\mathrm{e}^{-2m} < 1/t$. Under the uniform seed measure every event has probability in $t^{-1}\mathbb{Z}$, so each exceptional event is empty. □

Remark 7.6. The parameter m_{eff} is an arithmetic concentration depth, not a mixing time. It is limited both by the logarithmic period gap and by the characteristic ceiling $2m < p$. Thus an exponential-in-N seed tail requires not only $\log(t/q^{1/2}) \gg N$ but also $p \gg N$. When $0 < \Lambda \ll N$, the useful order is at least of size $\Lambda/\log(eN/\Lambda)$ under the corresponding characteristic hypothesis. If $\Lambda \le 0$, the Gaussian diagonal need not dominate even at the first moment level encoded by the definition.

8. Gaussian fluctuations in mesoscopic windows

Write $\mathcal{N}_{\mathbb{C}}(0,1)$ for the circular complex Gaussian law with density

$$\pi^{-1}e^{-|z|^2}\,dz$$

on $\mathbb{C} \simeq \mathbb{R}^2$. Its mixed moments are

$$\mathbb{E}[Z^a\overline{Z}^b] = \begin{cases} a!, & a = b, \\ 0, & a \ne b. \end{cases}$$

Let $(q_j, p_j, t_j, N_j, E_j, G_{0,j}, P_{0,j})$ be a sequence of data as above. For each j, choose an arbitrary nonzero Walsh frequency k_j for the corresponding digital output map.

Theorem 8.1 (Complex-Gaussian limit)**.** *Assume*

$$p_j \longrightarrow \infty, \qquad N_j \longrightarrow \infty,$$

and for every fixed $m \ge 1$,

$$N_j^m \frac{q_j^{1/2}}{t_j} \longrightarrow 0.$$

Then

$$\frac{Z_{N_j,k_j}(U) - N_j\mu_{k_j}}{\sqrt{N_j}} \Longrightarrow \mathcal{N}_{\mathbb{C}}(0,1).$$

Moreover, for every fixed $a, b \ge 0$,

$$\mathbb{E}\left[\left(\frac{Z_{N_j,k_j} - N_j\mu_{k_j}}{\sqrt{N_j}}\right)^a \left(\frac{\overline{Z_{N_j,k_j}} - N_j\overline{\mu_{k_j}}}{\sqrt{N_j}}\right)^b\right] \longrightarrow \begin{cases} a!, & a = b, \\ 0, & a \ne b. \end{cases}$$

Proof. Fix $a, b \ge 0$. For all sufficiently large j, one has $a + b < p_j$. If $a \ne b$, Theorem 4.3 gives

$$\left|\mathbb{E}\left[\left(\frac{Z_{N_j,k_j}}{\sqrt{N_j}}\right)^a \left(\frac{\overline{Z_{N_j,k_j}}}{\sqrt{N_j}}\right)^b\right]\right| \ll_{a,b} N_j^{(a+b)/2}\frac{q_j^{1/2}}{t_j} \longrightarrow 0.$$

If $a = b = m$, then Theorem 4.3 and Lemma 4.1 give

$$\mathbb{E}\left|\frac{Z_{N_j,k_j}}{\sqrt{N_j}}\right|^{2m} = m! + O_m(N_j^{-1}) + O_m\left(\frac{1}{t_j} + N_j^m \frac{q_j^{1/2}}{t_j}\right),$$

which tends to $m!$.

It remains to justify centering. Corollary 4.8 and the case $m = 1$ of the hypothesis imply

$$\sqrt{N_j}|\mu_{k_j}| \le C\sqrt{N_j}\frac{q_j^{1/2}}{t_j} \longrightarrow 0.$$

For each fixed total degree, the normalized uncentered mixed moments are uniformly bounded by the preceding estimates. Expanding

$$\left(\frac{Z_{N_j,k_j} - N_j\mu_{k_j}}{\sqrt{N_j}}\right)^a \left(\frac{\overline{Z_{N_j,k_j}} - N_j\overline{\mu_{k_j}}}{\sqrt{N_j}}\right)^b$$

by the binomial theorem therefore shows that every term containing at least one factor $\sqrt{N_j}\mu_{k_j}$ tends to zero. The centered mixed moments have the asserted limits.

Write the centered normalized variable as $X_j + iY_j$. Every real monomial $X_j^r Y_j^s$ is a finite linear combination of mixed monomials in the variable and its complex conjugate, so the preceding convergence gives convergence of all joint moments of (X_j, Y_j) to those of the two-dimensional centered Gaussian with covariance matrix $\frac{1}{2}I_2$. The uniform second-moment bound implies tightness: for $R > 0$,

$$\mathbb{P}(X_j^2 + Y_j^2 > R^2) \le R^{-2}\mathbb{E}(X_j^2 + Y_j^2).$$

Every subsequential weak limit therefore exists, and the uniform bounds for moments of degree larger than any fixed test degree give uniform integrability for that test degree. Hence each subsequential limit has all the Gaussian moments. The Gaussian law is moment-determinate (its moment generating function is finite in a neighborhood of the origin), so every subsequential limit is the same circular Gaussian law. Tightness then yields convergence of the full sequence. □

Theorem 8.2 (Joint Gaussian limit with resonances)**.** *Fix $s \ge 1$. For the jth member of an asymptotic family, let*

$$\mathbf{k}_j = (k_{1,j}, \dots, k_{s,j})$$

be nonzero Walsh frequencies. Assume that, for every pair i, h, the two indicators

$$\mathbf{1}_{\{k_{i,j}=k_{h,j}\}}, \qquad \mathbf{1}_{\{k_{i,j}=\ominus k_{h,j}\}}$$

are eventually constant as $j \to \infty$. Let $G = (G_1, \dots, G_s)$ be the centered complex Gaussian vector determined by

$$\mathbb{E}[G_i\overline{G_h}] = \lim_{j\to\infty} \mathbf{1}_{\{k_{i,j}=k_{h,j}\}}, \qquad \mathbb{E}[G_iG_h] = \lim_{j\to\infty} \mathbf{1}_{\{k_{i,j}=\ominus k_{h,j}\}}.$$

Such a vector is obtained by assigning one independent standard circular complex Gaussian to each eventual equivalence class of frequencies modulo $k \sim \ominus k$, identifying repeated frequencies, and assigning complex conjugates to inverse frequencies. Assume

$$p_j \to \infty, \qquad N_j \to \infty, \qquad N_j^m \frac{q_j^{1/2}}{t_j} \to 0 \quad \text{for every fixed } m \ge 1.$$

Then

$$\left(\frac{Z_{N_j,k_{1,j}} - N_j\mu_{k_{1,j}}}{\sqrt{N_j}}, \dots, \frac{Z_{N_j,k_{s,j}} - N_j\mu_{k_{s,j}}}{\sqrt{N_j}}\right) \Longrightarrow G.$$

Proof. Fix nonnegative powers of the s coordinates and their conjugates, and let D be the resulting total degree. For all sufficiently large j, $D < p_j$. After division by $N_j^{D/2}$, Theorem 4.7 gives the compatible-pairing number plus the three errors

$$O_D(N_j^{-1/2}), \qquad O_D(t_j^{-1}), \qquad O_D\left(N_j^{D/2}\frac{q_j^{1/2}}{t_j}\right).$$

The first two tend to zero. For the last one, choose an integer $m \geq D/2$; then $N_j^{D/2} \leq N_j^m$, and the assumed arithmetic condition forces convergence to zero. Wick's formula identifies the compatible-pairing number with the corresponding mixed moment of G because a compatible unbarred–barred pair records covariance, while a compatible pair with equal bar status records pseudo-covariance.

For every fixed coordinate i, Corollary 4.8 gives

$$\sqrt{N_j}|\mu_{k_{i,j}}| \ll \sqrt{N_j}\frac{q_j^{1/2}}{t_j} \longrightarrow 0.$$

A finite multinomial expansion therefore shows that centering leaves every fixed joint moment unchanged in the limit; all lower-degree normalized moments appearing in that expansion are uniformly bounded by the same joint-moment theorem.

To conclude distributional convergence, identify $\mathbb{C}^s$ with $\mathbb{R}^{2s}$. The coordinate second moments are uniformly bounded, so Markov's inequality gives tightness of the vector laws. Complex mixed-moment convergence is equivalent, by linear expansion of real and imaginary parts, to convergence of every real joint moment. Every subsequential limit has the moments of the stated real Gaussian vector, and that Gaussian is moment-determinate because its moment generating function is finite on all of $\mathbb{R}^{2s}$. Hence every subsequential limit is G, which proves the claimed convergence. □

Theorem 8.3 (Quantitative Gaussian approximation for polynomial tests)**.** *Fix nonzero Walsh frequencies $k_1, \ldots, k_s$ and let $G_{\mathbf{k}} = (G_1, \ldots, G_s)$ be the centered complex Gaussian vector with covariance and pseudo-covariance*

$$\mathbb{E}[G_i\overline{G_h}] = \mathbf{1}_{\{k_i = k_h\}}, \qquad \mathbb{E}[G_iG_h] = \mathbf{1}_{\{k_i = \ominus k_h\}}.$$

Let $P(z, \overline{z})$ be a polynomial in $(z_1, \ldots, z_s, \overline{z}_1, \ldots, \overline{z}_s)$ of total degree at most D, where $1 \leq D < p$. Put

$$W_{N,\mathbf{k}} = \left(\frac{Z_{N,k_1} - N\mu_{k_1}}{\sqrt{N}}, \ldots, \frac{Z_{N,k_s} - N\mu_{k_s}}{\sqrt{N}}\right)$$

and

$$\varepsilon_D(N, q, t) = N^{-1/2} + t^{-1} + N^{D/2}\frac{q^{1/2}}{t}.$$

If $\varepsilon_D(N, q, t) \leq 1$, then

$$\left|\mathbb{E}P(W_{N,\mathbf{k}}, \overline{W_{N,\mathbf{k}}}) - \mathbb{E}P(G_{\mathbf{k}}, \overline{G_{\mathbf{k}}})\right| \leq C_P\,\varepsilon_D(N, q, t),$$

where C_P depends only on the polynomial P (and hence on its fixed frequency dimension and degree), not on q, t, N or on the chosen frequencies.

Proof. It suffices to treat a monomial of total degree $d \leq D$. Before centering, Theorem 4.7, divided by $N^{d/2}$, gives the corresponding Wick moment of $G_{\mathbf{k}}$ with error

$$O_d\left(N^{-1/2} + t^{-1} + N^{d/2}\frac{q^{1/2}}{t}\right) = O_d(\varepsilon_D(N, q, t)).$$

(The estimate $N^{-1/2}$ is deliberately uniform in the parity of d; for even d the balanced lower-order contribution is in fact $O_d(N^{-1})$.)

Set $a_i = \sqrt{N}\,\mu_{k_i}$. Corollary 4.8 gives

$$|a_i| \ll \sqrt{N}\frac{q^{1/2}}{t} \le N^{D/2}\frac{q^{1/2}}{t} \le \varepsilon_D(N,q,t)$$

for $D \ge 1$. The uncentered normalized moments of every degree at most D are uniformly bounded when $\varepsilon_D \le 1$, by the same mixed-moment estimate. Expanding each centered monomial by the binomial theorem therefore changes its expectation by at most $O_d(\varepsilon_D)$. Summing over the finitely many monomials of P proves the result. □

Corollary 8.4 (Dissociated frequencies give independent limits)**.** *If, in addition, the frequency family is D-dissociated for every fixed D and all sufficiently large indices, then no two listed frequencies are eventually equal or inverse. The Gaussian vector in Theorem 8.2 therefore has independent standard circular complex Gaussian coordinates.*

Corollary 8.5 (Real linear statistics of a resonant Walsh family)**.** *Under the hypotheses of Theorem 8.2, fix coefficients $c_1,\dots,c_s \in \mathbb{C}$ and put*

$$L_j = \Re\sum_{i=1}^{s} c_i \frac{Z_{N_j,k_{i,j}} - N_j\mu_{k_{i,j}}}{\sqrt{N_j}}.$$

Then L_j converges in distribution to a centered real Gaussian variable of variance

$$\sigma_c^2 = \frac{1}{2}\Re\sum_{i,h} c_i\overline{c_h}\,C_{ih} + \frac{1}{2}\Re\sum_{i,h} c_i c_h\,P_{ih},$$

where

$$C_{ih} = \lim_j \mathbf{1}_{\{k_{i,j}=k_{h,j}\}}, \qquad P_{ih} = \lim_j \mathbf{1}_{\{k_{i,j}=\ominus k_{h,j}\}}.$$

The variance may vanish, in which case the limit is the point mass at zero.

Proof. The joint convergence in Theorem 8.2 and the continuous mapping theorem reduce the claim to the corresponding real linear functional of the limiting complex Gaussian vector. If $W = \sum_i c_i G_i$, then

$$\operatorname{Var}(\Re W) = \frac{1}{2}\mathbb{E}|W|^2 + \frac{1}{2}\Re\mathbb{E}W^2,$$

which gives the displayed formula from the covariance and pseudo-covariance matrices. □

Corollary 8.6 (Multivariate real Walsh-polynomial limit)**.** *Under the hypotheses of Theorem 8.2, fix an integer $q_0 \ge 1$ and coefficient arrays*

$$(c_{\alpha i})_{1\le\alpha\le q_0,\ 1\le i\le s} \subset \mathbb{C}.$$

For each j, define the real vector $L_j = (L_{1,j},\dots,L_{q_0,j})$ by

$$L_{\alpha,j} = \Re\sum_{i=1}^{s} c_{\alpha i}\frac{Z_{N_j,k_{i,j}} - N_j\mu_{k_{i,j}}}{\sqrt{N_j}}.$$

Then

$$L_j \Longrightarrow \mathcal{N}_{q_0}(0,\Sigma),$$

where

$$\Sigma_{\alpha\beta} = \frac{1}{2}\Re\sum_{i,h=1}^{s} c_{\alpha i}\overline{c_{\beta h}}\,C_{ih} + \frac{1}{2}\Re\sum_{i,h=1}^{s} c_{\alpha i}c_{\beta h}\,P_{ih},$$

and

$$C_{ih} = \lim_j \mathbf{1}_{\{k_{i,j}=k_{h,j}\}}, \qquad P_{ih} = \lim_j \mathbf{1}_{\{k_{i,j}=\ominus k_{h,j}\}}.$$

The covariance matrix Σ is allowed to be singular.

Proof. Theorem 8.2 gives convergence of the underlying complex vector to $G = (G_1, \dots, G_s)$. Apply the continuous real-linear map

$$(z_1, \dots, z_s) \longmapsto \left(\Re \sum_i c_{1i} z_i, \dots, \Re \sum_i c_{q_0 i} z_i \right).$$

The image of a complex Gaussian vector under a real-linear map is a real Gaussian vector. For $W_\alpha = \sum_i c_{\alpha i} G_i$,

$$\operatorname{Cov}(\Re W_\alpha, \Re W_\beta) = \frac{1}{2} \Re \mathbb{E}[W_\alpha \overline{W_\beta}] + \frac{1}{2} \Re \mathbb{E}[W_\alpha W_\beta],$$

which gives the displayed formula. □

Corollary 8.7 (Cumulant decay for finite real Walsh statistics)**.** *Under the hypotheses of Corollary 8.6, let*

$$L_j = (L_{1,j}, \dots, L_{q_0,j})$$

and let $\kappa_j(\alpha_1, \dots, \alpha_r)$ denote the joint cumulant of

$$L_{\alpha_1,j}, \dots, L_{\alpha_r,j}.$$

Then, for every fixed $r \geq 1$,

$$\kappa_j(\alpha_1, \dots, \alpha_r) \longrightarrow \begin{cases} 0, & r = 1, \\ \Sigma_{\alpha_1 \alpha_2}, & r = 2, \\ 0, & r \geq 3. \end{cases}$$

Thus all fixed higher joint cumulants vanish in the mesoscopic regime.

Proof. For fixed order r, a joint cumulant is a universal polynomial in joint moments of orders at most r. Theorem 4.7, followed by the same centering expansion used in Theorem 8.2, proves convergence of every joint moment of the coordinates of L_j of total degree at most r. Thus every moment entering the cumulant polynomial converges to the corresponding moment of $\mathcal{N}_{q_0}(0, \Sigma)$. Substitution in the finite cumulant polynomial proves the claim. Centered Gaussian cumulants vanish in all orders other than two. □

Theorem 8.8 (Quantitative cumulant approximation)**.** *Fix nonzero Walsh frequencies $k_1, \dots, k_s$ and coefficient arrays*

$$(c_{\alpha i})_{1 \leq \alpha \leq q_0,\ 1 \leq i \leq s} \subset \mathbb{C}.$$

For a single member of the family put

$$L_{\alpha,N} = \Re \sum_{i=1}^{s} c_{\alpha i} \frac{Z_{N,k_i} - N \mu_{k_i}}{\sqrt{N}}, \qquad 1 \leq \alpha \leq q_0,$$

and let G_α be the corresponding real linear statistics of the resonant Gaussian vector $G_{\mathbf{k}}$. Fix $r < p$ and indices $\alpha_1, \dots, \alpha_r$. If

$$\varepsilon_r(N, q, t) = N^{-1/2} + t^{-1} + N^{r/2} \frac{q^{1/2}}{t} \leq 1,$$

then

$$\left| \kappa\big(L_{\alpha_1,N}, \dots, L_{\alpha_r,N}\big) - \kappa\big(G_{\alpha_1}, \dots, G_{\alpha_r}\big) \right| \leq C_{r,\mathbf{c}}\, \varepsilon_r(N, q, t),$$

where κ denotes the joint cumulant and $C_{r,\mathbf{c}}$ depends only on r and the fixed coefficient array. In particular, for $r \geq 3$,

$$\left| \kappa\big(L_{\alpha_1,N}, \dots, L_{\alpha_r,N}\big) \right| \leq C_{r,\mathbf{c}}\, \varepsilon_r(N, q, t),$$

and for $r = 2$ the covariance differs from the limiting Gaussian covariance by $O_{\mathbf{c}}(\varepsilon_2(N, q, t))$.

Proof. A joint cumulant of order r is a universal polynomial, with integer coefficients depending only on r, in joint moments indexed by the set partitions of $\{1,\dots,r\}$. Each joint moment appearing in that polynomial is the expectation of a polynomial of total degree at most r in $W_{N,\mathbf{k}}$ and its conjugate. Theorem 8.3 therefore approximates each such moment by its Gaussian counterpart with error $O_{r,\mathbf{c}}(\varepsilon_r)$. Under $\varepsilon_r \le 1$, all of these finitely many moments are uniformly bounded, so substitution into the cumulant polynomial gives the displayed estimate. Gaussian cumulants vanish in every order $r \ge 3$. □

Proposition 8.9 (Real Walsh variance at finite window length)**.** *Fix nonzero frequencies $k_1,\dots,k_s$ and coefficients $c_1,\dots,c_s \in \mathbb{C}$. Put*

$$L_N = \Re \sum_{i=1}^{s} c_i \frac{\widetilde{Z}_{N,k_i}}{\sqrt{N}}, \qquad \widetilde{Z}_{N,k} = Z_{N,k} - N\mu_k,$$

and define

$$\sigma_c^2 = \frac{1}{2}\Re \sum_{i,h} c_i \overline{c_h} \mathbf{1}_{\{k_i = k_h\}} + \frac{1}{2}\Re \sum_{i,h} c_i c_h \mathbf{1}_{\{k_i = \ominus k_h\}}.$$

Then

$$\left|\mathrm{Var}(L_N) - \sigma_c^2\right| \ll \|c\|_1^2 \left(\frac{1}{t} + \frac{Nq^{1/2}}{t} + \frac{Nq}{t^2}\right).$$

In particular, if

$$\frac{Nq^{1/2}}{t} \longrightarrow 0,$$

then

$$\mathrm{Var}(L_N) \longrightarrow \sigma_c^2.$$

Proof. For a centered complex random variable W,

$$\mathrm{Var}(\Re W) = \frac{1}{2}\mathbb{E}|W|^2 + \frac{1}{2}\Re \mathbb{E}W^2.$$

Apply this identity with

$$W = N^{-1/2} \sum_i c_i \widetilde{Z}_{N,k_i}$$

and use Proposition 4.9 term by term. Summing the error terms with absolute values costs at most $\|c\|_1^2$. If $Nq^{1/2}/t \to 0$, then also $q^{1/2}/t \to 0$ and

$$\frac{Nq}{t^2} = \left(\frac{Nq^{1/2}}{t}\right)\left(\frac{q^{1/2}}{t}\right) \longrightarrow 0.$$

□

Corollary 8.10 (A convenient mesoscopic regime)**.** *Suppose that, along an asymptotic family, $N \to \infty$, $p \to \infty$, and for some fixed $\eta > 0$,*

$$\frac{t}{q^{1/2}} \ge q^{\eta}, \qquad N = q^{o(1)}.$$

Then Theorem 8.1 applies.

Proof. For every fixed $m \ge 1$, the hypothesis $N = q^{o(1)}$ gives $N^m = q^{o(1)}$. Therefore

$$N^m \frac{q^{1/2}}{t} \le q^{-\eta + o(1)} \longrightarrow 0.$$

Together with $N \to \infty$ and $p \to \infty$, these are exactly the hypotheses of Theorem 8.1. □

9. Applications and interpretation

9.1. **Seed certification.** For a prescribed moment order m, Corollary 5.2 bounds the fraction of cyclic seeds for which one or several selected Walsh tests are atypical. Corollary 7.5 also identifies a regime in which the exceptional probability is smaller than the mass $1/t$ of one seed, forcing the exceptional set to be empty. The resulting statements are exact finite-population confidence bounds: the generator is deterministic, and the probability is solely the uniform measure on its t cyclic starting indices.

9.2. **Randomly shifted quasi–Monte Carlo rules.** A cyclic shift is a natural randomization of a deterministic point set. Theorem 6.4 supplies a high-probability discrepancy bound for elliptic-curve windows under this randomization, while Corollary 6.7 converts it into an integration-error bound for functions of bounded Hardy–Krause variation. The deterministic grid term $\Delta_{p,a,d}$, the complete-orbit bias $q^{1/2}/t$, and the random-window fluctuation scale are kept separate throughout; this separation is useful when the base, dimension, period, and window length vary simultaneously.

9.3. **Finite batteries of output tests.** Individual Walsh coefficients are linear statistics of the digital output. Theorems 5.1 and 5.5 therefore give a hierarchy of finite-window tests, ranging from variance-sensitive ℓ^2 control to high-confidence ℓ^1 control. The joint Gaussian theorem identifies the limiting covariance of any fixed finite test battery and records precisely when repeated or inverse frequencies create resonances.

9.4. **Arithmetic concentration without dynamical mixing.** The map $u \mapsto u + 1$ on $\mathbb{Z}/t\mathbb{Z}$ is periodic and has no decay of correlations in the usual ergodic-theoretic sense. The concentration proved here comes from a different source: every unbalanced correlation becomes a complete additive-character sum of a non–Artin–Schreier rational function on an elliptic subgroup. Thus the saving is supplied by arithmetic cancellation after averaging over the seed space, not by temporal decorrelation along a random process.

10. Relation to random affine walks and empirical deviations

The probability model of this paper is a uniformly chosen cyclic shift of one deterministic translation orbit. It must be distinguished from the random affine recursion

$$X_{n+1} = TX_n + B_{n+1}$$

on a finite abelian group, in which fresh independent noise enters at every step. The latter belongs to the Chung–Diaconis–Graham line of mixing problems [1, 2, 3]. On a cyclic elliptic subgroup it is conjugate, at the level of the state variable, to a one-dimensional modular affine walk. Rank-two torsion groups, nonlinear elliptic outputs, and output-sensitive convergence are the first genuinely new elliptic directions.

Within the wider three-question program stated in the introduction, the division of labor is as follows.

(i) The present paper resolves the random-seed question for translation windows by complete elliptic correlation sums. Its main objects are the empirical window statistics $M_{N,k}$ and $D_{U,N}$.
(ii) The companion mixing problem asks for Fourier cancellation, entropy lower bounds, and post-wrap control of exceptional frequencies for random affine walks. Its main object is the one-time law of X_n and its distance from stationarity.
(iii) The empirical-deviation problem for the random affine walk asks for concentration, moderate deviations, and large deviations of path sums after burn-in. Its decisive additional input is uniform spectral perturbation of tilted transition operators as the state space grows.

These three mechanisms should not be conflated. A strong seed-wise bound in the present paper does not imply that a random affine walk mixes, and a state-space mixing theorem does not by itself identify a path-space rate function. Conversely, no result from either companion direction is used in the proofs of the theorems above; the present manuscript is logically self-contained apart from the explicitly cited subgroup-sum and standard analytic inputs.

11. Toward moderate and large deviations

The preceding results are finite-order concentration estimates and fixed-order Gaussian approximations. A moderate- or large-deviation theorem requires a specified joint asymptotic regime for (q, t, N, p) and uniform control of logarithmic moment generating functions at the proposed speed.

Put

$$S_{N,k}(U) = Z_{N,k}(U) - N\mu_k.$$

Because $S_{N,k}$ is complex-valued, the relevant normalized logarithmic moment generating function is two-dimensional:

$$\Lambda_N(\theta_1, \theta_2) = \frac{1}{N} \log \mathbb{E} \exp\left(\theta_1 \Re S_{N,k}(U) + \theta_2 \Im S_{N,k}(U)\right).$$

A local Gärtner–Ellis argument would require convergence and suitable regularity of Λ_N on a neighborhood of the origin. A full speed-N LDP requires control on the effective domain together with the hypotheses needed for the lower bound; see, for example, Dembo–Zeitouni [13]. Fixed-order cumulant convergence alone does not provide this uniform information.

There is also a finite-seed obstruction that every proposed asymptotic formulation must respect.

Proposition 11.1 (Finite-seed entropy ceiling)**.** *Let $\mathcal{A} \subset \mathbb{Z}/t\mathbb{Z}$ be nonempty. Under the uniform choice of U,*

$$\mathbb{P}(U \in \mathcal{A}) \geq \frac{1}{t}.$$

Consequently,

$$-\frac{1}{N} \log \mathbb{P}(U \in \mathcal{A}) \leq \frac{\log t}{N}.$$

Thus, for any sequence of nonempty deviation events, a speed-N exponential rate cannot exceed $\log t/N$. In particular, a nonzero finite speed-N rate for such events requires a regime in which $\log t/N$ does not vanish.

Proof. Every event in the uniform t-point seed space has probability in $t^{-1}\mathbb{Z}$. A nonempty event therefore has probability at least $1/t$, and taking logarithms gives the claim. □

The moment expansion isolates the additional arithmetic input that would be sufficient for a Cramér-type range of deviations.

Problem 11.2 (All-order arithmetic hypercontractivity)**.** Determine the largest range of integers $m = m(q, t, N)$ for which

$$\mathbb{E}\,|S_{N,k}(U)|^{2m} \leq (CmN)^m$$

holds uniformly in every nonzero Walsh frequency. Any proof relying only on the pole-multiplicity argument of the present paper also requires $2m < p$; reaching $m \asymp N$ by that mechanism therefore requires $p \gg N$.

If the estimate in Problem 11.2 held for every $m \leq cN$, then moment optimization would give

$$\mathbb{P}\big(|M_{N,k} - \mu_k| \geq x\big) \leq 2\exp(-c'Nx^2)$$

through a fixed range of x, subject to Proposition 11.1. The theorem proved in this paper leaves the absolute off-diagonal contribution

$$N^{2m}\frac{q^{1/2}}{t}.$$

This is effective for fixed and logarithmic moment orders, but not generally for $m \asymp N$. Closing that gap requires cancellation among off-diagonal configurations, an all-order spectral-flatness statement, or a different representation of the seed-space log-mgf.

A moderate-deviation principle is a potentially earlier target. It would suffice to obtain a cumulant or log-mgf expansion uniformly for a growing but sublinear order, with a remainder compatible with the selected moderate-deviation scale. The exact admissible range must simultaneously respect the arithmetic ratio $t/q^{1/2}$, the characteristic ceiling, and the finite-seed entropy ceiling.

12. Extensions and limits of transfer

The argument applies verbatim to every additively parametrized cyclic coset

$$Q_u = A + [u]R, \qquad u \in \mathbb{Z}/\operatorname{ord}(R)\mathbb{Z},$$

provided the random window is consecutive in this additive coordinate. Translation of the coset changes only the locations of the poles and leaves their orders unchanged.

A general affine elliptic-curve congruential generator may admit a set-theoretic or algebraic reparametrization by a cyclic coordinate. Such a reparametrization does not automatically preserve chronological windows. The theorems above therefore apply immediately to windows that are consecutive in the additive coset coordinate, but not automatically to chronological initial segments of an arbitrary affine recursion. The latter can produce multiplicative or affine index progressions and requires separate Fourier or spectral information.

The same translated-pole framework can treat fixed finite collections of rational observables, but only after an explicit nondegeneracy analysis. One must identify collisions among translated poles, possible cancellation of leading Laurent coefficients, and short linear relations among the resulting coefficient vectors. No serial or multi-observable high-moment theorem beyond the hypotheses explicitly proved in Sections 3 and 4 is asserted here.

Characteristics 2 and 3 also remain outside the scope of the frequency-uniform theorem. In those characteristics, pole orders 2 or 3 can be divisible by p, and the Artin–Schreier exclusion must be replaced by a separate local analysis rather than by a formal reuse of Lemma 3.4.

13. Conclusion

For random cyclic starting indices on an elliptic-curve translation orbit, this paper establishes the following complete chain:

$$\text{translated-pole cancellation} \Longrightarrow \text{mixed moments and Wick pairings} \Longrightarrow \text{finite-order concentration},$$
$$\text{finite-order concentration} \Longrightarrow \text{discrepancy bounds and Gaussian fluctuations}.$$

More concretely, it proves uniform mixed moments below the characteristic, an explicit degree-p resonance showing the sharpness of that range for the method, finite-order subgaussian seed tails, confidence bounds for finite Walsh batteries, direct finite-grid discrepancy estimates, an effective arithmetic concentration depth, and joint Gaussian limits with quantitative polynomial and cumulant errors. The paper also separates these results sharply from Markov-chain mixing and from triangular-array empirical large deviations.

The remaining large-deviation problem is not concealed inside the finite-order theorems: it is the genuinely stronger task of obtaining all-order arithmetic control or uniform convergence of the seed-space logarithmic moment generating function, in a regime compatible with both the characteristic and the finite number of seeds.

References

[1] F. R. K. Chung, P. Diaconis, and R. L. Graham, Random walks arising in random number generation, *Ann. Probab.* **15** (1987), 1148–1165.

[2] S. Eberhard and P. P. Varjú, Mixing time of the Chung–Diaconis–Graham random process, *Probab. Theory Related Fields* **179** (2021), 317–344.
[3] E. Breuillard and P. P. Varjú, Cut-off phenomenon for the $ax + b$ Markov chain over a finite field, *Probab. Theory Related Fields* **184** (2022), 85–113.
[4] Z. Liu and C. P. Mok, On discrepancy estimates for pseudorandom vectors constructed by the elliptic curve congruential generator, preprint, 2026, arXiv:2605.20627.
[5] C. P. Mok, Pseudorandom vector generation using elliptic curves and applications to Wiener processes, *Finite Fields Appl.* **85** (2023), 102129.
[6] X. Wang, On the distribution of pseudorandom vectors generated by elliptic curves, *Finite Fields Appl.* **93** (2024), 102318.
[7] T. Lange and I. E. Shparlinski, Certain exponential sums and random walks on elliptic curves, *Canad. J. Math.* **57** (2005), 338–350.
[8] J.-R. Chazottes and S. Gouëzel, Optimal concentration inequalities for dynamical systems, *Comm. Math. Phys.* **316** (2012), 843–889.
[9] C. Aistleitner and M. Hofer, Probabilistic discrepancy bound for Monte Carlo point sets, *Math. Comp.* **83** (2014), 1373–1381.
[10] D. R. Kohel and I. E. Shparlinski, On exponential sums and group generators for elliptic curves over finite fields, in *Algorithmic Number Theory*, Lecture Notes in Comput. Sci. **1838**, Springer, 2000, 395–404.
[11] P. Hellekalek, General discrepancy estimates: the Walsh function system, *Acta Arith.* **67** (1994), 209–218.
[12] H. Niederreiter, *Random Number Generation and Quasi–Monte Carlo Methods*, CBMS–NSF Regional Conference Series in Applied Mathematics, vol. 63, SIAM, Philadelphia, 1992.
[13] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Springer, New York, 1998.

(Z. Liu) Shanghai Institute for Mathematics and Interdisciplinary Sciences (SIMIS), Shanghai, China, 200433

(Z. Liu) Research Institute of Intelligent Complex Systems, Fudan University, Shanghai 200433, China
Email address: zliu@simis.cn

(C. P. Mok) Shanghai Institute for Mathematics and Interdisciplinary Sciences (SIMIS), Shanghai, China, 200433
Email address: cpmok@simis.cn