hexagon:2610.00634v2

A lower bound for the sixth moment of the Riemann zeta function

Niklas Fraccaro

Abstract

We establish a new lower bound for the sixth moment of the Riemann zeta function on the critical line, improving the previous constant of 34.4 obtained by Durkan and Page to 41.008, close to the conjectured value of 42 predicted by the random matrix model of Keating and Snaith. Our approach extends the two-piece amplifier method by allowing independent profiles, centered derivatives, and finite linear combinations of two families of projected test functions.

2020 Mathematics Subject Classification. Primary 11M06.
Keywords. Riemann zeta function, sixth moment, amplified moments, twisted fourth moment.

Introduction

Understanding the behavior of the Riemann zeta function $\zeta(s)$ on the critical line $\Re(s) = \frac{1}{2}$ is a central problem in analytic number theory. In particular, determining the size of its largest values remains a challenging open problem. A natural approach is to study the moments of the zeta function, defined by

$$M_k(T) := \int_0^T \left|\zeta\left(\frac{1}{2} + it\right)\right|^{2k} \mathrm{d}t, \qquad k > 0.$$

Although moments are themselves difficult to analyze, they provide a more accessible way of studying the distribution of values of ζ on the critical line. Moreover, as k increases, the contribution from large values of $|\zeta(\frac{1}{2} + it)|$ becomes increasingly significant. Thus, higher moments provide valuable information about the extreme behavior of the zeta function.

Asymptotic formulas are known for the second and fourth moments. In particular,

$$M_1(T) \sim T \log T$$

and

$$M_2(T) \sim \frac{1}{2\pi^2} T (\log T)^4$$

as $T \to \infty$. However, for higher even moments, no corresponding asymptotic formula has yet been established. Keating and Snaith [6] conjectured that, for each fixed positive integer k,

$$M_k(T) \sim C_k T (\log T)^{k^2},$$

where $C_k > 0$ is an explicit constant.

The next natural case beyond the fourth moment is therefore the sixth moment, for which the conjecture predicts

$$M_3(T) \sim C_3 T (\log T)^9.$$

To be more precise, let

$$a_3 := \prod_p \left(1 - \frac{1}{p}\right)^4 \left(1 + \frac{4}{p} + \frac{1}{p^2}\right),$$

where the product runs over all primes. Consistent with the earlier conjecture of Conrey and Ghosh [2], the Keating–Snaith conjecture [6] predicts that

$$C_3 = \frac{42 a_3}{9!}.$$

Theorem 1. *We establish the lower bound*

$$\liminf_{T\to\infty} \frac{M_3(T)}{\frac{a_3}{9!} T (\log T)^9} > 41.008.$$

Outline of the Proof

The main idea is to construct a test function that approximates $\zeta(s)^3$ well enough to obtain a strong lower bound for its sixth moment using Cauchy–Schwarz.

We begin with two families of test functions, built from Dirichlet polynomials and centered derivatives of the zeta function. Together, these generalize the two-piece amplifier used by Durkan and Page. The functional equation allows us to project these functions onto the same real line in $\mathbb{C}$ as $\zeta(s)^3$. This projection preserves their correlation with $\zeta(s)^3$ while not increasing their norm. We then take finite real linear combinations of these projected functions. Applying Cauchy–Schwarz reduces the problem to finding a vector of mixed moments and a Gram matrix describing the norms and interactions of the test functions.

The next step is to evaluate these moments. For the numerator, the two families require different analytic results. For the first family, we extend the twisted fourth-moment argument of Bettin, Bui, Li and Radziwiłł [1] to Dirichlet polynomials of different lengths. This gives the range $3a + 5b < 2$. For the second family, we use the fourth-moment formula of Durkan and Page [4], with polynomial lengths below $T^{1/4}$. Although the two calculations are different, we show that their main terms can be expressed using the same kernel, with a reflection in one variable for the second family.

For the denominator, we expand the products of the projected test functions and use the functional equation to reduce them to balanced and unbalanced moments involving two zeta factors. The condition on the polynomial lengths allows us to control the off-diagonal terms. This gives explicit formulas for all entries of the Gram matrix, including the mixed terms between the two families.

Once these formulas are established, the remaining problem is finite-dimensional. We first work with point evaluations of the kernels, which are easier to compute exactly, and then approximate them by admissible smooth or polynomial profiles. This ensures that the resulting bound applies to genuine test functions.

Finally, we choose 250 explicit tests and evaluate the corresponding quadratic expressions using exact rational arithmetic. The resulting quotient is strictly greater than 41.008. Applying Cauchy–Schwarz and combining the estimates over finitely many dyadic intervals removes the smooth weight and gives the claimed lower bound for the sixth moment.

1 Notation and Setup

Let $T > 1$, $L = \log T$ and $s = \frac{1}{2} + it$. We use the functional equation in the following convention

$$\zeta(s) = \chi(s)\zeta(1-s), \qquad \chi(s) = \pi^{s-1/2} \frac{\Gamma((1-s)/2)}{\Gamma(s/2)}.$$

On the critical line, $1-s=\bar{s}$ and $|\chi(s)|=1$, so that $\zeta(s)^3=\chi(s)^3\overline{\zeta(s)^3}$.

For a real-valued, compactly supported function f with $\operatorname{supp} f\subseteq[0,a]$, define

$$A_{f,T}(w):=\sum_{n\geq 1} f\Big(\frac{\log n}{L}\Big)n^{-w},$$

which is a Dirichlet polynomial of length at most T^a, since $f(\log n/L)$ vanishes for $n>T^a$.

For $p\in\mathbb{R}[x]$, we put

$$\begin{aligned} C_{p,T}(w)&=p\Big(\frac{1}{2}+\frac{1}{L}\frac{\mathrm{d}}{\mathrm{d}w}\Big)\zeta(w)\\ &=p(\partial_z)\big[e^{z/2}\zeta(w+z/L)\big]\Big|_{z=0}, \qquad\qquad p^-(x)=p(-x).\end{aligned}$$

To simplify notation, we suppress T when fixed. Since the profiles and polynomial coefficients are real,

$$A_f(1-s)=\overline{A_f(s)},\qquad C_p(1-s)=\overline{C_p(s)}.$$

All profiles and test polynomials remain fixed as $T\to\infty$.

We also define

$$R_T(t,z):=e^z\frac{\chi(s+z/L)}{\chi(s)}.$$

The functional equation then gives the exact identity

$$e^{z/2}\zeta(s+z/L)=\chi(s)R_T(t,z)e^{-z/2}\zeta(1-s-z/L).$$

1.1 Test family construction

To give an idea where our approach starts, we begin with the two-piece amplifier test used by Durkan and Page [4, Section 1.3 and equation (1.9)],

$$\zeta(s)\big(A(s)+\chi(s)A(1-s)\big)^2.$$

Expanding the square gives

$$\begin{aligned}&\zeta(s)\big(A(s)+\chi(s)A(1-s)\big)^2\\ &\quad=\zeta(s)A(s)^2+2\chi(s)\zeta(s)A(s)A(1-s)+\chi(s)^2\zeta(s)A(1-s)^2.\end{aligned}$$

Although this expansion contains three terms, there are only two essentially different types of products: same-side products, such as $A(s)^2$ and $A(1-s)^2$, and mixed products of the form $A(s)A(1-s)$.

The two same-side terms are related by the functional equation of the Riemann zeta function. This suggests working with a projection that automatically produces the reflected terms.

For a complex-valued function $H(t)$, define

$$\Pi H(t)=\frac{H(t)+\chi(s)^3\overline{H(t)}}{2}.$$

Since $|\chi(s)|=1$ on the critical line, this is pointwise the orthogonal projection, with respect to the real inner product $\Re(z\bar{w})$, onto the line

$$\{w\in\mathbb{C}: w=\chi(s)^3\bar{w}\}.$$

By the functional equation, this line also contains $\zeta(s)^3$.

As we said before, A has real coefficients, so that $\overline{A(s)} = A(1-s)$ on the critical line. Using

$$\overline{\zeta(s)} = \chi(s)^{-1}\zeta(s), \qquad \overline{\chi(s)} = \chi(s)^{-1},$$

we obtain

$$\chi(s)^3\overline{\zeta(s)A(s)^2} = \chi(s)^2\zeta(s)A(1-s)^2.$$

Thus the projection of the first same-side term produces both same-side contributions:

$$2\Pi\big(\zeta(s)A(s)^2\big) = \zeta(s)A(s)^2 + \chi(s)^2\zeta(s)A(1-s)^2.$$

For the mixed term, we instead have

$$\chi(s)^3\overline{\chi(s)\zeta(s)A(s)A(1-s)} = \chi(s)\zeta(s)A(s)A(1-s).$$

Hence the mixed term is already invariant under the reflection, and

$$2\Pi\big(\chi(s)\zeta(s)A(s)A(1-s)\big) = 2\chi(s)\zeta(s)A(s)A(1-s).$$

Consequently, the entire two-piece test can be recovered by projecting just two terms:

$$\begin{aligned}2\Pi\Big(\zeta(s)A(s)^2 + \chi(s)\zeta(s)A(s)A(1-s)\Big)\\ = \zeta(s)\big(A(s) + \chi(s)A(1-s)\big)^2.\end{aligned}$$

This observation motivates our two test families. To further expand the test space, we generalize the zeta factor from $\zeta(s)$ to $C_p(s)$ and allow the two amplifier profiles to vary independently. We define

$$U_{p;f,g}(t) = \chi(s)C_p(s)A_f(s)A_g(1-s),$$

and

$$V_{p;f,g}(t) = C_p(s)A_f(s)A_g(s).$$

The family U corresponds to the mixed products, whereas V corresponds to the same-side products.

For the moment calculations we use the following identities:

$$\begin{aligned}\zeta(s)^3\overline{\Pi H} &= \Re\big(\zeta(s)^3\overline{H}\big),\\ (\Pi H)\overline{\Pi J} &= \frac{1}{2}\Re\big(H\overline{J} + \overline{\chi(s)^3}HJ\big),\\ |\Pi H| &\le |H|.\end{aligned} \tag{1.1}$$

Although both test families are projected using the same operator Π, the moment estimates used to analyze their numerators impose different restrictions on the admissible profiles. We therefore specify these restrictions before turning to the analytic estimates.

1.2 Admissible profiles

For the two test families, we impose the following conditions on the profiles:

$$\begin{aligned}U:&\quad f \in C_c^\infty((0,a)),\quad g \in C_c^\infty((0,b)),\quad 0 < a \le b,\quad 3a+5b<2;\\ V:&\quad f,g \text{ are real polynomials on } [0,\eta], \text{ extended by zero},\quad 0<\eta<\tfrac{1}{4}.\end{aligned}$$

When considering finitely many V-tests, we use a common cutoff η. For the pairwise moment calculations, we additionally require the V-profiles to vanish at both endpoints, so that their extensions by zero are continuous.

These restrictions imply that the two Dirichlet-polynomial factors in each test have a combined length exponent strictly less than $1/2$. Consequently, any pair of tests, including a mixed pair, has total length exponent less than one.

1.3 Finite test combinations

We now consider finite real linear combinations of the two test families. Let $H_1, \dots, H_m$ be admissible tests from U and V, and let $c = (c_1, \dots, c_m) \in \mathbb{R}^m$. We define

$$F_c(t) := \Pi\left(\sum_{j=1}^{m} c_j H_j(t)\right) = \sum_{j=1}^{m} c_j \Pi H_j(t).$$

For the moment calculations, let $\Phi \in C_c^\infty((1,2))$ be real-valued and write

$$\widehat{\Phi}(0) := \int_{\mathbb{R}} \Phi(x)\, \mathrm{d}x.$$

The moment formulas below also apply to signed weights Φ. For Cauchy–Schwarz, we take $\Phi \geq 0$ with $\widehat{\Phi}(0) > 0$. Whenever the denominator is nonzero, Cauchy–Schwarz gives

$$\int_{\mathbb{R}} \Phi(t/T)|\zeta(s)|^6\, \mathrm{d}t \geq \frac{\left|\int_{\mathbb{R}} \Phi(t/T)\zeta(s)^3\overline{F_c(t)}\, \mathrm{d}t\right|^2}{\int_{\mathbb{R}} \Phi(t/T)|F_c(t)|^2\, \mathrm{d}t}.$$

The correlation in the numerator is linear in c, whereas the squared norm in the denominator is quadratic in c. Their normalized leading terms will be described by a real vector k and a real symmetric Gram matrix G, giving rise to the expressions

$$c^\top k \qquad \text{and} \qquad c^\top G c.$$

We will determine these coefficients from the individual correlations and pairwise products of the projected tests in the following sections.

2 Analytic Foundations of the Lower Bound

The goal of this section is to determine the numerator vector k and the denominator Gram matrix G through explicit limiting kernels, thereby reducing the problem to the optimization of

$$\frac{(c^\top k)^2}{c^\top G c}.$$

We begin with a general nine-edge diagonal limit, whose underlying arithmetic structure also enters the numerator calculation, while the full limiting formula is needed for the denominator. We then derive the numerator asymptotics before turning to the denominator and its associated Gram matrix.

2.1 Nine-edge Diagonal Limit

Lemma 2 (Nine-edge diagonal limit)**.** *Let F be bounded, measurable and compactly supported on $\mathbb{R}^3_{\geq 0} \times \mathbb{R}^3_{\geq 0}$, and suppose that its discontinuity set has zero measure for the limiting measure below. Writing $\log \mathbf{n}/L = (\log n_1/L, \dots, \log n_3/L)$, we have*

$$\frac{1}{L^9} \sum_{n_1 n_2 n_3 = m_1 m_2 m_3} \frac{F(\log \mathbf{n}/L, \log \mathbf{m}/L)}{n_1 n_2 n_3} \longrightarrow a_3 \int_{x_{ij} \geq 0} F\left(\left(\sum_j x_{ij}\right)_i, \left(\sum_i x_{ij}\right)_j\right) \prod_{i,j=1}^{3} \mathrm{d}x_{ij}. \tag{2.1}$$

Proof. Rather than working directly with a general test function F, we first compute the Laplace transforms of the associated measures. This amounts to considering exponential test functions of the form

$$F_{\alpha,\beta}(\mathbf{x},\mathbf{y}) = \exp\left(-\sum_{i=1}^{3}\alpha_i x_i - \sum_{j=1}^{3}\beta_j y_j\right),$$

where $\alpha_i, \beta_j > 0$.

We first group the diagonal sum according to the common product $k = n_1n_2n_3 = m_1m_2m_3$. For this purpose, define

$$A_{\mathbf{u}}(k) = \sum_{n_1n_2n_3=k}\prod_{i=1}^{3} n_i^{-u_i}.$$

Since these coefficients are multiplicative, the diagonal Dirichlet series admits an Euler product:

$$\sum_{\prod_i n_i=\prod_j m_j} \frac{\prod_i n_i^{-u_i}\prod_j m_j^{-v_j}}{\prod_i n_i} = \sum_{k\geq 1}\frac{A_{\mathbf{u}}(k)A_{\mathbf{v}}(k)}{k} = \prod_p H_p(\mathbf{u},\mathbf{v}),$$

where the local factor at a prime p is

$$H_p(\mathbf{u},\mathbf{v}) = \sum_{r\geq 0} p^{-r}\left(\sum_{\substack{a_1+a_2+a_3=r\\ a_i\geq 0}} p^{-\sum_i u_i a_i}\right)\left(\sum_{\substack{b_1+b_2+b_3=r\\ b_j\geq 0}} p^{-\sum_j v_j b_j}\right).$$

The contribution from $r = 0$ is 1. For $r = 1$, exactly one exponent on each side equals 1, giving $3 \cdot 3 = 9$ possible pairs. Consequently,

$$H_p(\mathbf{u},\mathbf{v}) = 1 + \sum_{i,j=1}^{3} p^{-(1+u_i+v_j)} + O(p^{-2+\varepsilon}),$$

uniformly for sufficiently small shifts.

These nine first-order terms coincide with those arising from the local Euler factors of the nine zeta functions

$$\prod_{i,j=1}^{3}\zeta(1+u_i+v_j).$$

We therefore define the local remainder

$$E_p(\mathbf{u},\mathbf{v}) = H_p(\mathbf{u},\mathbf{v})\prod_{i,j=1}^{3}\left(1-p^{-(1+u_i+v_j)}\right).$$

The first-order terms cancel, so

$$E_p(\mathbf{u},\mathbf{v}) = 1 + O(p^{-2+\varepsilon}).$$

Choosing $\varepsilon < 1$, the Euler product defining E converges locally uniformly near the origin and hence defines a holomorphic function there. Thus

$$D(\mathbf{u},\mathbf{v}) = E(\mathbf{u},\mathbf{v})\prod_{i,j=1}^{3}\zeta(1+u_i+v_j).$$

Finally, at $\mathbf{u} = \mathbf{v} = \mathbf{0}$, the number of nonnegative solutions of $a_1 + a_2 + a_3 = r$ is $\binom{r+2}{2}$. Hence

$$\begin{aligned} E_p(\mathbf{0}, \mathbf{0}) &= (1 - p^{-1})^9 \sum_{r \geq 0} \binom{r+2}{2}^2 p^{-r} \\ &= (1 - p^{-1})^4 (1 + 4p^{-1} + p^{-2}). \end{aligned}$$

Taking the product over all primes gives $E(\mathbf{0}, \mathbf{0}) = a_3$.

Substituting $u_i = \alpha_i / L$ and $v_j = \beta_j / L$, with positive real parts, and normalizing by L^{-9}, yields the limiting Laplace transform

$$\lim_{L \to \infty} L^{-9} D(\alpha/L, \beta/L) = a_3 \prod_{i,j} (\alpha_i + \beta_j)^{-1}.$$

The limiting Laplace transform can be identified by considering the map that sends the nine nonnegative variables x_{ij} to their row and column sums. Indeed,

$$a_3 \int_{x_{ij} \geq 0} e^{-\sum_{i,j} (\alpha_i + \beta_j) x_{ij}} \prod_{i,j} \mathrm{d}x_{ij} = a_3 \prod_{i,j} (\alpha_i + \beta_j)^{-1}.$$

Thus, the limit is precisely the Laplace transform of the pushforward of a_3 times Lebesgue measure under this map.

Although the measures involved are not finite, multiplying them by a fixed positive exponential weight makes them finite. Their Laplace transforms still converge, so the continuity theorem for Laplace transforms yields weak convergence of the weighted measures.

Finally, since F has compact support, the exponential weight is bounded away from zero on its support. We can therefore remove the weight and use weak convergence for bounded measurable functions whose discontinuity sets have limiting measure zero. This proves the assertion. □

2.2 Numerator

The numerator is determined by the correlations of the two test families, U and V, with $\zeta(s)^3$. Although these correlations lead to different types of twisted moments, their limiting expressions can be described by a single kernel. The purpose of this section is to identify this common kernel and derive the corresponding numerator formulas for both families.

We begin by introducing the kernel and stating the main result. For the U-family, we use an asymmetric twisted fourth moment, which allows the two amplifier profiles to be different and to appear on opposite sides of the moment. Its six main terms can then be combined and reduced to a simplex integral. The V-family requires a different approach, based on the fourth-moment formula of Durkan–Page. Finally, we establish a reflection identity connecting the two expressions. This shows that both families are governed by the same kernel, with a reflected argument in the V-case, and completes the evaluation of the numerator.

2.2.1 The common numerator kernel and the main result

We first introduce this common kernel and state the corresponding limits, before turning to their derivation in the following subsections.

Let $X = (X_0, X_1, X_2, X_3)$ be uniformly distributed on the three-dimensional simplex

$$X_j \geq 0, \qquad \sum_{j=0}^{3} X_j = 1.$$

We normalize the associated expectation by

$$\mathbb{E}F(X) = 6\int_{\substack{x_1,x_2,x_3\geq 0\\ x_1+x_2+x_3\leq 1}} F(1-x_1-x_2-x_3,x_1,x_2,x_3)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3, \qquad \mathbb{E}1 = 1.$$

For $a, b \geq 0$, define

$$\begin{aligned} P(u,v) &= \frac{uv(1-u)(1-v)(1-u-v)}{12},\\ q(a,b,t) &= \Big(-\frac{1}{2},\, a-t-\frac{1}{2},\, \frac{1}{2}-b+t,\, \frac{1}{2}\Big),\\ K_z(a,b) &= \int_0^{\min(a,b)} P(b-t,a-t)\,\mathbb{E}e^{zq(a,b,t)\cdot X}\,\mathrm{d}t. \end{aligned} \tag{2.2}$$

Here a and b represent logarithmic positions in the amplifier profiles, rather than fixed support bounds. For each fixed a, b, the function $K_z(a,b)$ is entire in z.

The parameter z allows us to incorporate the polynomial weights appearing in the numerator. For a real polynomial p, we define

$$\begin{aligned} K_p(a,b) &= p(\partial_z)K_z(a,b)\big|_{z=0}\\ &= \int_0^{\min(a,b)} P(b-t,a-t)\,\mathbb{E}p\big(q(a,b,t)\cdot X\big)\,\mathrm{d}t. \end{aligned} \tag{2.3}$$

Thus, the kernels associated with different polynomial weights are obtained from the same generating function K_z.

We can now state the numerator limits for both test families in a common form.

Proposition 3 (Numerator limits)**.** *For an admissible U- or V-test and a real polynomial p of degree at most four, the corresponding limiting numerator coefficients are given by*

$$k_U(p;f,g) = \int_0^\infty\int_0^\infty f(a)g(b)K_p(a,b)\,\mathrm{d}a\,\mathrm{d}b, \tag{2.4}$$

$$k_V(p;f,g) = \int_0^\eta\int_0^\eta f(a)g(b)K_p(a,1-b)\,\mathrm{d}a\,\mathrm{d}b. \tag{2.5}$$

No additional projection factor appears in these formulas.

2.2.2 An asymmetric twisted fourth moment

The U-numerator contains one Dirichlet polynomial on each side of the critical line. Their lengths need not be equal, and imposing the same upper bound on both would unnecessarily restrict the available tests. We therefore retain the two lengths separately in the argument of Bettin–Bui–Li–Radziwiłł [1]. The main terms remain unchanged; the point is to preserve the finer length dependence in their off-diagonal estimates.

We first specify the moment and its complete main term. We then collect the ingredients from BBLR before applying them to the asymmetric problem. Throughout this subsection, references to equations and page numbers in [1] refer to arXiv:1609.02539v1.

For $\lambda = (\alpha,\beta,\gamma,\delta)$, define

$$\begin{aligned} I_\lambda(T;f,g) = \int_{\mathbb{R}} \Phi(t/T)\zeta(s+\alpha)\zeta(s+\beta)\zeta(\bar{s}+\gamma)\zeta(\bar{s}+\delta)\\ \cdot A_f(s)A_g(\bar{s})\,\mathrm{d}t, \qquad s=\tfrac{1}{2}+it. \end{aligned} \tag{2.6}$$

Put

$$\tau_{u,v}(n) = \sum_{de=n} d^{-u}e^{-v}, \qquad \lambda^* = (\gamma,\delta,\alpha,\beta),$$

and, for coprime positive integers h, k, let

$$\mathcal{A}_\lambda = \frac{\zeta(1+\alpha+\gamma)\zeta(1+\alpha+\delta)\zeta(1+\beta+\gamma)\zeta(1+\beta+\delta)}{\zeta(2+\alpha+\beta+\gamma+\delta)},$$

$$\mathcal{B}_\lambda(h) = \prod_{\ell^\nu \| h} \frac{\sum_{r\geq 0} \tau_{\alpha,\beta}(\ell^r)\tau_{\gamma,\delta}(\ell^{r+\nu})\ell^{-r}}{\sum_{r\geq 0} \tau_{\alpha,\beta}(\ell^r)\tau_{\gamma,\delta}(\ell^r)\ell^{-r}},$$

$$Z_\lambda(h,k) = \mathcal{A}_\lambda \mathcal{B}_\lambda(h)\mathcal{B}_{\lambda^*}(k).$$

Here and below, ℓ denotes a prime when it occurs in an Euler product. Writing $d = (h,k)$, define

$$\mathcal{M}_\lambda(T; f, g) = \sum_{h,k\geq 1} f\Big(\frac{\log h}{L}\Big) g\Big(\frac{\log k}{L}\Big) \frac{d}{hk} Z_\lambda(h/d, k/d). \tag{2.7}$$

This is a finite coefficient sum and does not depend on t.

The six shift configurations and their phase exponents are

j	$\lambda^{(j)}$	$e_j(\lambda)$
0	$(\alpha,\beta,\gamma,\delta)$	0
1	$(-\gamma,-\delta,-\alpha,-\beta)$	$-\alpha-\beta-\gamma-\delta$
2	$(-\gamma,\beta,-\alpha,\delta)$	$-\alpha-\gamma$
3	$(-\delta,\beta,\gamma,-\alpha)$	$-\alpha-\delta$
4	$(\alpha,-\gamma,-\beta,\delta)$	$-\beta-\gamma$
5	$(\alpha,-\delta,\gamma,-\beta)$	$-\beta-\delta.$

(2.8)

The same table defines $\sigma^{(j)}$ and $e_j(\sigma)$ for scaled shifts:

$$(\sigma/L)^{(j)} = \sigma^{(j)}/L, \qquad e_j(\sigma/L) = e_j(\sigma)/L.$$

We write

$$\mathfrak{M}_\lambda(T; f, g) = \int_{\mathbb{R}} \Phi(t/T) \sum_{j=0}^{5} \Big(\frac{t}{2\pi}\Big)^{e_j(\lambda)} \mathcal{M}_{\lambda^{(j)}}(T; f, g)\, \mathrm{d}t. \tag{2.9}$$

Although individual summands may have poles when shifts coincide, their sum has a removable continuation. This continuation, justified below, is the meaning of (2.9) at such configurations.

Lemma 4 (Asymmetric twisted fourth moment)**.** *Let $0 < a \leq b$ and $3a + 5b < 2$. Let $f \in C_c^\infty((0,a))$ and $g \in C_c^\infty((0,b))$ be real-valued, and let $\Phi \in C_c^\infty((1,2))$ be real-valued. For every compact $K \subset \mathbb{C}^4$, with $\lambda = \sigma/L$, we have*

$$I_\lambda(T; f, g) = \mathfrak{M}_\lambda(T; f, g) + o(T) \tag{2.10}$$

uniformly for $\sigma \in K$. More precisely, for every $\varepsilon > 0$ the error is

$$O_{K,f,g,\Phi,\varepsilon}\Big(T^{1/2+a+b+\varepsilon} + T^{3/4+(3a+5b)/8+\varepsilon}\Big). \tag{2.11}$$

The same statements hold with f and g interchanged in both the moment and the main term. Thus the shorter polynomial may occur on either side of the critical line.

Part I. Ingredients from BBLR.

BBLR already allows two independent coefficient sequences. We use that freedom without modification. The additional work will be to apply the intermediate estimates in their proof without replacing the two polynomial lengths by a common upper bound.

For the moment, suppose that the four signed scaled shifts

$$-\sigma_1, \quad -\sigma_2, \quad \sigma_3, \quad \sigma_4 \tag{2.12}$$

are pairwise distinct. All preliminary estimates are uniform on compact sets satisfying this condition. The restriction will be removed in Part II.

Write $\lambda^\dagger = (-\gamma, -\delta, -\alpha, -\beta)$. We choose an even entire function G, with $G(0) = 1$, which decays rapidly in vertical strips and vanishes at

$$w = \pm\frac{\alpha+\gamma}{2}, \quad \pm\frac{\alpha+\delta}{2}, \quad \pm\frac{\beta+\gamma}{2}, \quad \pm\frac{\beta+\delta}{2}.$$

One possible choice on the separated shift set is

$$G(w) = e^{w^2} \prod_{u\in\{\alpha,\beta\}} \prod_{v\in\{\gamma,\delta\}} \left(1 - \frac{4w^2}{(u+v)^2}\right). \tag{2.13}$$

This choice is unchanged by permutations within either shift pair and by $\lambda \mapsto \lambda^\dagger$. Its coefficients grow at most as a fixed power of L on compact separated scaled shift sets. Such logarithmic factors will be absorbed into T^ε; no uniform choice of G at a collision is needed.

Define

$$\begin{aligned}
\mathcal{G}_\lambda(w,t) &= \prod_{u\in\{\alpha,\beta\}} \frac{\Gamma((s+u+w)/2)}{\Gamma((s+u)/2)} \prod_{v\in\{\gamma,\delta\}} \frac{\Gamma((\bar{s}+v+w)/2)}{\Gamma((\bar{s}+v)/2)},\\
V_\lambda(x,t) &= \frac{1}{2\pi i}\int_{(1)} \frac{G(w)}{w}\mathcal{G}_\lambda(w,t)\pi^{-2w}x^{-w}\,\mathrm{d}w,\\
\mathcal{X}_\lambda(t) &= \chi(s+\alpha)\chi(s+\beta)\chi(\bar{s}+\gamma)\chi(\bar{s}+\delta).
\end{aligned}$$

Here (c) denotes the upward-oriented line $\Re w = c$. The approximate functional equation [1, Lemma 2.1 and equation (3)] is

$$\begin{aligned}
&\zeta(s+\alpha)\zeta(s+\beta)\zeta(\bar{s}+\gamma)\zeta(\bar{s}+\delta)\\
&\quad= \sum_{m,n\ge1} \frac{\tau_{\alpha,\beta}(m)\tau_{\gamma,\delta}(n)}{\sqrt{mn}}\left(\frac{m}{n}\right)^{-it} V_\lambda(mn,t)\\
&\qquad+ \mathcal{X}_\lambda(t)\sum_{m,n\ge1} \frac{\tau_{-\gamma,-\delta}(m)\tau_{-\alpha,-\beta}(n)}{\sqrt{mn}}\left(\frac{m}{n}\right)^{-it} V_{\lambda^\dagger}(mn,t) + O_N(T^{-N})
\end{aligned} \tag{2.14}$$

for every fixed $N > 0$, uniformly for $T \le t \le 2T$. We call the two sums the direct and dual sums.

The estimates needed for these weights follow from Stirling's formula; see [1, equations (4)–(6)]. In particular,

$$\mathcal{X}_\lambda(t) = \left(\frac{t}{2\pi}\right)^{-\alpha-\beta-\gamma-\delta}(1+O(T^{-1})), \tag{2.15}$$

and, in a fixed vertical strip with $|\Im w| \le \log T$,

$$\pi^{-2w}\mathcal{G}_\lambda(w,t) = \left(\frac{t}{2\pi}\right)^{2w}\left(1 + O\left(\frac{(1+|w|)^C}{T}\right)\right).$$

The Gaussian in (2.13) controls the complementary part of the contour. For any fixed $B, i, j \ge 0$, these estimates also give

$$x^i t^j \frac{\partial^{i+j}}{\partial x^i \partial t^j} V_\lambda(x,t) \ll_{B,i,j} L^C(1+x/T^2)^{-B}. \tag{2.16}$$

The same bounds apply to the dual weight.

It is convenient to introduce the simpler Mellin weight

$$V^*(x) = \frac{1}{2\pi i}\int_{(1)} \frac{G(w)}{w}(2\pi)^{-2w}x^{-w}\,\mathrm{d}w. \tag{2.17}$$

Its dependence on the shifts is suppressed in the notation. By the preceding Stirling estimate, for every $\varepsilon > 0$,

$$V_\lambda(x,t) = V^*(x/t^2) + O(T^{-1+\varepsilon}) \qquad (x \geq 1,\ T \leq t \leq 2T). \tag{2.18}$$

The same replacement holds for $V_{\lambda^\dagger}$, with the same function V^*.

The off-diagonal equation obtained from (2.14) is

$$hm_1m_2 - kn_1n_2 = \pm r, \qquad r > 0.$$

We record the estimate for a dyadic block in the form that retains its dependence on the two amplifier lengths. Let

$$h \asymp A, \quad k \asymp B, \quad m_i \asymp M_i, \quad n_i \asymp N_i, \quad r \asymp H,$$

and put

$$M = M_1M_2, \qquad N = N_1N_2, \qquad P = ABMN.$$

Let u_h, v_k be bounded coefficient sequences supported on the indicated intervals, and let $W_0, \dots, W_4$ be smooth weights supported in $[1,2]$. Consider

$$\begin{aligned} S_\pm = \sum_{hm_1m_2 - kn_1n_2 = \pm r} u_hv_k\, W_0(r/H)W_1(m_1/M_1)W_2(m_2/M_2) \\ \cdot W_3(n_1/N_1)W_4(n_2/N_2). \end{aligned} \tag{2.19}$$

Small-power losses in the coefficients and in finitely many derivatives of the weights are allowed and are included in T^ϵ below. All parameters are assumed to be bounded by fixed powers of T.

Suppose, for a sufficiently small auxiliary $\eta > 0$, that

$$\begin{aligned} &AM \asymp BN, \qquad MN \gg (AB)^{\delta_0}, \qquad M_1 \ll M_2P^\eta, \qquad N_1 \ll N_2P^\eta, \\ &\qquad\qquad BN_1 \leq AM_1, \qquad H \ll (AB)^{1/2+\eta}, \end{aligned} \tag{2.20}$$

where $\delta_0 > 0$ is fixed. The zero-frequency term in [1, Proposition 3.1] is

$$\begin{aligned} \mathcal{Z} = \sum_{\substack{h,k,m_1,n_1,d,q \geq 1 \\ (hm_1,kn_1)=d}} u_hv_kW_0(dq/H)W_1(m_1/M_1)W_3(n_1/N_1) \\ \cdot \int_0^\infty W_2\Big(\frac{kn_1x}{dM_2}\Big)\,W_4\Big(\frac{hm_1x}{dN_2}\Big)\,\mathrm{d}x. \end{aligned} \tag{2.21}$$

Thus the shift is $r = dq$. Notice that $\mathcal{Z}$ is the same for the two signs in (2.19).

The proof of that proposition splits the nonzero frequencies according to $d = (hm_1, kn_1)$. More precisely, the large-d range is

$$d > D_0, \qquad D_0 = \min\left\{(AB)^{1/2}(AM)^{-100\eta}, \Big(\frac{AM_1N_1H}{BM_2}\Big)^{1/3}\right\},$$

and the small-d range is its complement. Equation (15) of BBLR is used in the first range and equation (16) in the second. After integration and summation over d, but before the final symmetrization in A, B, the displayed estimates on pages 16–17 give

$$E_{\mathrm{I}} \ll_{\varepsilon} T^{\varepsilon}(PH^2)^{1/4}\left(AB + H^{1/4}A^{3/8}B^{1/8}P^{1/8}\right), \tag{2.22}$$

$$E_{\mathrm{II}} \ll_{\varepsilon} T^{\varepsilon}(PH^2)^{1/4}\left(AB + (ABH)^{1/4}P^{1/8}\right). \tag{2.23}$$

For a prescribed ε, the auxiliary η is chosen sufficiently small to absorb the small-power losses in these estimates. The preliminary replacement of the shifted weight in the Poisson calculation contributes $O(T^{\varepsilon}H^2)$; see [1, equation (14) and the preceding calculation]. Consequently,

$$S_{\pm} = \mathcal{Z} + O\Big(E_{\mathrm{I}} + E_{\mathrm{II}} + T^{\varepsilon}H^2\Big). \tag{2.24}$$

These are intermediate estimates from the proof, not the symmetrized error bound in the statement of Proposition 3.1.

To use (2.20), the two factors of each divisor function must first be ordered smoothly. Following [1, Section 4, pp. 21–25], choose ω with

$$\omega(x) + \omega(1/x) = 1$$

and with rapid decay as $x \to \infty$. Its Mellin transform is required to vanish at $\pm(\alpha-\beta)/2$ and $\pm(\gamma-\delta)/2$. For example, one may define ω by inverse Mellin transformation of

$$\widehat{\omega}(w) = \frac{e^{w^2}}{w}\left(1 - \frac{4w^2}{(\alpha-\beta)^2}\right)\left(1 - \frac{4w^2}{(\gamma-\delta)^2}\right).$$

The pole at zero has residue one, and the remaining numerator is even. Changing w to $-w$ and moving the contour across zero gives the required identity for ω. The identity

$$\big(\omega(m_1/m_2) + \omega(m_2/m_1)\big)\big(\omega(n_1/n_2) + \omega(n_2/n_1)\big) = 1 \tag{2.25}$$

therefore produces four ordered contributions. The Mellin zeros ensure that removing this ordering later introduces no residues at the two shift differences.

The weights in the moment are not initially products of five separate functions. After dyadic localization, Mellin inversion separates their dependence on the variables. Repeated integration by parts in the Mellin transforms gives an integrable bound in all imaginary directions. One can therefore apply (2.24) to the separated weights and then undo the Mellin inversions. This is the separation and recombination argument of [1, Section 4, equation (21)]. Only finitely many derivative bounds are needed; small powers arising from the localization are included in T^{ε}.

We spell out the resulting main-term formula, since this identifies the four crossed terms in our moment. Put

$$c_{\ell}(x,y,z) = \frac{(1-\ell^{-1-x})(1-\ell^{-1-y})}{1-\ell^{-2-z}},$$

and define the finite Euler factors

$$\eta_{\lambda;h}(w) = \sum_{e\mid h} e^{-\gamma+\delta-w} \prod_{\ell\mid h/e} c_{\ell}(\alpha-\beta+\gamma-\delta+w, \gamma-\delta, \alpha-\beta+\gamma-\delta),$$

$$\eta_{\lambda;h,k}(w) = \eta_{\lambda;h}(w)\eta_{\lambda^*;k}(w).$$

This is $\eta_{\alpha,\beta,\gamma,\delta,h,k}(0,0,w)$ in [1, equation (23)]; only this specialization is needed here. For a smooth weight $F(x,y,r)$, write

$$\widehat{F}_3(x,y,w) = \int_0^{\infty} F(x,y,r) r^{w-1}\,\mathrm{d}r$$

and

$$\widetilde{F}_\lambda(x,y;h,k,d)=\frac{1}{2\pi i}\int_{(1+\epsilon_0)}\widehat{F}_3(x,y,w)\zeta(w)\zeta(1+\alpha-\beta+\gamma-\delta+w)\\ \cdot d^{-w}\eta_{\lambda;h,k}(w)\,\mathrm{d}w, \tag{2.26}$$

where $\epsilon_0>0$ is fixed and small.

The weighted formula applies with u_h, v_k supported on dyadic intervals $[A,2A]$, $[B,2B]$, with F supported on $r\asymp H$, and with the derivative bounds

$$\partial_x^i\partial_y^j\partial_r^k F(x,y,r)\ll T^\varepsilon(1+x)^{-i}(1+y)^{-j}(1+r)^{-k},$$
$$\mathcal{W}^{(j)}(x)\ll_{j,J} T^\varepsilon(1+x)^{-j}(1+x/X^2)^{-J}$$

for fixed derivative orders and every fixed $J>0$. An overall size factor in F is extracted before using these bounds. The same formula applies after summing smooth dyadic partitions. For these weights, consider

$$\mathcal{S}_\lambda(F,\mathcal{W})=\sum_{hm_1m_2-kn_1n_2=r>0}\frac{u_hv_k}{m_1^\alpha m_2^\beta n_1^\gamma n_2^\delta}F(hm_1m_2,kn_1n_2,r)\mathcal{W}(m_1m_2n_1n_2).$$

The zero-frequency calculation in [1, Theorem 4.1 and its proof] produces four main terms,

$$\mathcal{Q}_{\alpha,\beta,\gamma,\delta}+\mathcal{Q}_{\beta,\alpha,\gamma,\delta}+\mathcal{Q}_{\alpha,\beta,\delta,\gamma}+\mathcal{Q}_{\beta,\alpha,\delta,\gamma}, \tag{2.27}$$

where the dependence on $F,\mathcal{W},u,v$ is suppressed and

$$\mathcal{Q}_\lambda=\frac{\zeta(1+\alpha-\beta)\zeta(1+\gamma-\delta)}{\zeta(2+\alpha-\beta+\gamma-\delta)}\sum_{d\ge1}\sum_{(h_0,k_0)=1}\frac{u_{dh_0}v_{dk_0}\,d}{(dh_0)^{1-\beta}(dk_0)^{1-\delta}}\\ \cdot\int_0^\infty\mathcal{W}\left(\frac{x^2}{d^2h_0k_0}\right)\widetilde{F}_\lambda(x,x;h_0,k_0,d)x^{-\beta-\delta}\,\mathrm{d}x. \tag{2.28}$$

The formula for the negative sign is obtained from the same zero-frequency expression (2.21).

There are two distinct errors in this passage. The nonzero frequencies are controlled by (2.24), before the Mellin inversions are undone. The contour shifts used to remove the ordering weights contribute an additional error. With $X=(MN)^{1/2}$, the bound on page 25 of [1] is

$$E_{\mathrm{rec}}\ll_\varepsilon T^\varepsilon(ABX^2)^{1/4}H\big(A^{1/2}+B^{1/2}\big). \tag{2.29}$$

This bound is for an unnormalized block; a scalar normalization of its weight multiplies the error by the same scalar. We will check this error separately rather than include it implicitly in the quadratic-divisor estimate.

The connection between (2.28) and the factors Z_λ is the local identity [1, equation (9), pp. 11–12]. In our notation it gives, for $(h,k)=1$,

$$h^\beta k^\delta\eta_{\lambda;h,k}(\beta+\delta)=\mathcal{B}_{(\alpha,-\delta,\gamma,-\beta)}(h)\mathcal{B}_{(\gamma,-\beta,\alpha,-\delta)}(k). \tag{2.30}$$

It is enough to verify this at prime powers; the identity for general h,k follows by multiplicativity.

For later use, put

$$C_\lambda(z)=(\alpha+z,-\delta-z,\gamma+z,-\beta-z). \tag{2.31}$$

The finite factor $\eta_{\lambda;h,k}(w)$ depends on the shifts only through their differences. Applying (2.30) with all four shifts increased by z therefore gives

$$h^{\beta+z}k^{\delta+z}\eta_{\lambda;h,k}(\beta+\delta+2z)=\mathcal{B}_{C_\lambda(z)}(h)\mathcal{B}_{C_\lambda(z)^*}(k). \tag{2.32}$$

This identity will identify the entire contour integrand, not just its residue at zero.

Part II. Proof of the asymmetric moment formula.

Proof of Lemma 4. We first work on a compact separated set as in (2.12). Write $f_h = f(\log h/L)$ and $g_k = g(\log k/L)$. All implied constants may depend on this compact set, on the profiles, and on Φ. The profiles and the weight remain fixed as $T \to \infty$.

Insert (2.14) into (2.6) and expand the two divisor functions. The direct contribution has the form

$$\sum_{h,k,m_1,m_2,n_1,n_2\geq 1} \frac{f_h g_k}{\sqrt{hkm_1m_2n_1n_2}\, m_1^{\alpha} m_2^{\beta} n_1^{\gamma} n_2^{\delta}} \int_{\mathbb{R}} \Phi(t/T) \left(\frac{hm_1m_2}{kn_1n_2}\right)^{-it} V_\lambda(m_1m_2n_1n_2, t)\, \mathrm{d}t. \tag{2.33}$$

The dual contribution is treated in the same way. We separate the diagonal $hm_1m_2 = kn_1n_2$ and first estimate the error in the off-diagonal calculation.

Choose small positive parameters u and ρ. By (2.16), the part with $m_1m_2n_1n_2 > T^{2+u}$ is negligible. Repeated integration by parts in t makes the terms with

$$T\left|\log \frac{hm_1m_2}{kn_1n_2}\right| > T^{\rho}$$

negligible as well. Here and below, negligible means $O(T^{-N})$ for any prescribed fixed N, after increasing the number of integrations by parts or the decay exponent.

On a remaining dyadic block, put

$$Q = AB, \qquad R = \max\{A, B\}, \qquad X = (MN)^{1/2}, \qquad \Lambda = T^{\rho}.$$

The two products are comparable, so $AM \asymp BN \asymp \sqrt{Q}\, X$. If their difference has size H, the preceding restrictions become

$$X \ll T^{1+u/2}, \qquad H \ll \Lambda \frac{\sqrt{Q}\, X}{T}. \tag{2.34}$$

These are the two restrictions that will convert the divisor-sum estimates into a moment estimate.

The coefficient normalization on the block is $P^{-1/2} = (\sqrt{Q}\, X)^{-1}$, and the time integral has length $O(T)$. Consequently the conversion factor for an unnormalized block error is

$$\frac{T}{\sqrt{Q}\, X}. \tag{2.35}$$

All ratios between the exact square-root factors and their dyadic sizes are included in the smooth weight.

We also record why the elementary weight replacements do not cost a new leading error. For fixed h, m_1, m_2 and the difference r, the number of factorizations of $hm_1m_2 \mp r$ as kn_1n_2 is $O_\varepsilon(T^{\varepsilon})$. There are $O(AMH) = O(\sqrt{Q}\, XH)$ choices before this factorization. Thus the absolute coefficient mass of a near-diagonal block in (2.33) is $O(T^{\varepsilon}H)$. The shift-dependent powers are bounded on the truncated ranges.

In particular, replacing the gamma weight by (2.18) is harmless. Likewise, on the ranges (2.34),

$$\left(1 \pm \frac{r}{kn_1n_2}\right)^{-it} = \exp\left(\mp \frac{itr}{kn_1n_2}\right) + O(\Lambda^2/T).$$

Multiplying these errors by the coefficient mass and integrating in t gives a total of $O(T^{(a+b)/2+\varepsilon})$, after choosing u, ρ sufficiently small and summing the dyadic blocks. This is smaller than the first term in (2.11).

We now verify that the refined quadratic-divisor bounds apply. Because the two profiles have compact support inside their open intervals, there are $a_0, b_0 > 0$ such that every nonzero coefficient satisfies

$$T^{a_0} \leq h \leq T^{a}, \qquad T^{b_0} \leq k \leq T^{b}.$$

Thus $Q \gg T^{a_0+b_0}$. Choosing $\rho + u/2$ sufficiently small in (2.34) gives $H \ll Q^{1/2+\eta}$. Moreover, a nonempty off-diagonal block has $H \gg 1$, and hence

$$MN = X^2 \gg \frac{T^{2-2\rho}}{Q}.$$

Since $Q \ll T^{a+b}$ and $a + b < 1/2$, this implies $MN \gg Q^{\delta_0}$ for some fixed $\delta_0 > 0$. The inequality $a + b < 1/2$ follows directly from $4(a + b) \le 3a + 5b < 2$.

Use (2.25) to order the factors within each side. The rapid decay of ω permits the restrictions $M_1 \ll M_2 P^\eta$ and $N_1 \ll N_2 P^\eta$, at a negligible cost. Interchanging the two sides when necessary gives $BN_1 \le AM_1$. This interchange leaves Q and R unchanged. All conditions in (2.20) are now satisfied.

The estimates (2.22) and (2.23) can be combined without replacing A and B by the larger one. Indeed,

$$A^{3/8}B^{1/8} \le Q^{1/8}R^{1/4}, \qquad Q^{1/4} \le Q^{1/8}R^{1/4}, \qquad P = QX^2.$$

Consequently the discrepancy between the divisor sum and its zero-frequency term is

$$E \ll_\varepsilon T^\varepsilon \left(H^{1/2}Q^{5/4}X^{1/2} + H^{3/4}Q^{1/2}R^{1/4}X^{3/4} + H^2 \right). \tag{2.36}$$

Multiplying by (2.35) and using the bound for H in (2.34) gives

$$E_{\mathrm{block}} \ll_\varepsilon T^\varepsilon \left(\Lambda^{1/2}T^{1/2}Q + \Lambda^{3/4}T^{1/4}Q^{3/8}R^{1/4}X^{1/2} + \Lambda^2 \frac{\sqrt{Q}\,X}{T} \right). \tag{2.37}$$

Since $Q \ll T^{a+b}$, $R \ll T^b$ and $X \ll T^{1+u/2}$, the three terms have respective bounds

$$T^{1/2+a+b+\rho/2+\varepsilon}, \qquad T^{3/4+(3a+5b)/8+3\rho/4+u/4+\varepsilon}, \qquad T^{(a+b)/2+2\rho+u/2+\varepsilon}.$$

The third is smaller than the first once the auxiliary parameters are sufficiently small.

The main-term recovery error (2.29) must also be included. With the same normalization, its contribution is at most

$$\begin{aligned} \frac{T}{\sqrt{Q}\,X} E_{\mathrm{rec}} &\ll_\varepsilon T^\varepsilon \Lambda Q^{1/4}X^{1/2}(A^{1/2} + B^{1/2}) \\ &\ll_\varepsilon T^{1/2+(a+3b)/4+\rho+u/4+\varepsilon}. \end{aligned}$$

Because $(a + 3b)/4 \le a + b$, this is absorbed by the first term of (2.11).

There are only logarithmically many dyadic blocks. For any prescribed final $\varepsilon > 0$, choose all auxiliary small-power parameters sufficiently small, including those needed for the finitely many weight derivatives in the Mellin separation. Absorbing the dyadic summation and these losses gives

$$E_{\mathrm{off}} \ll_\varepsilon T^{1/2+a+b+\varepsilon} + T^{3/4+(3a+5b)/8+\varepsilon}. \tag{2.38}$$

This is the step at which the separate lengths produce the condition $3a + 5b < 2$.

We next identify the main terms, retaining their full dependence on the shifts. On the diagonal, write $h = dh_0$, $k = dk_0$, with $(h_0, k_0) = 1$. The equation $hm = kn$ then gives $m = k_0 r$, $n = h_0 r$. The Euler factors defining Z imply

$$\sum_{r \ge 1} \frac{\tau_{\alpha,\beta}(k_0 r)\tau_{\gamma,\delta}(h_0 r)}{r^{1+2w}} = (h_0 k_0)^w Z_{\lambda + w\mathbf{1}}(h_0, k_0), \qquad \mathbf{1} = (1,1,1,1), \tag{2.39}$$

initially in a half-plane of absolute convergence. For example, a prime dividing h_0 contributes the shifted local factor in $\mathcal{B}_{\lambda+w\mathbf{1}}(h_0)$; primes dividing k_0 contribute the other factor, and all remaining primes contribute $\mathcal{A}_{\lambda+w\mathbf{1}}$.

Replacing V_λ by V^* on the diagonal costs only $O(T^\varepsilon)$. To see this, write the common product as $q = hm = kn$. On the truncated ranges the absolute diagonal coefficient sum is bounded by $\sum_{q\le T^C} d_3(q)^2/q = O(T^\varepsilon)$, where $d_3(q)$ denotes the number of ordered factorizations of q into three positive integers. The error $O(T^{-1+\varepsilon})$ in the weight is then integrated over an interval of length $O(T)$.

Insert (2.17) and use (2.39). The factor $(h_0k_0)^w$ cancels the corresponding factor from the Mellin weight. Thus the direct diagonal contribution is, up to the preceding error,

$$\int_{\mathbb{R}} \Phi(t/T)\frac{1}{2\pi i}\int_{(1)} \frac{G(w)}{w}\left(\frac{t}{2\pi}\right)^{2w} \mathcal{M}_{\lambda+w\mathbf{1}}(T;f,g)\,\mathrm{d}w\,\mathrm{d}t. \tag{2.40}$$

Move the inner contour to $\Re w = -1/8$. The possible zeta poles at $w = -(\alpha+\gamma)/2$ and the other three cross sums are canceled by G. The reciprocal zeta factor is regular in this strip, since its argument has real part greater than one for large T. Only the pole at $w = 0$ remains, and its residue is $\mathcal{M}_\lambda(T;f,g)$.

For completeness, a coarse bound suffices for the new contour. The local geometric series defining $\mathcal{B}$, together with polynomial vertical growth of the zeta factors, give on $\Re w = -1/8$

$$Z_{\lambda+w\mathbf{1}}(h_0,k_0) \ll_\varepsilon (1+|\Im w|)^C (h_0k_0)^{1/2+\varepsilon}.$$

Indeed the shifted divisor coefficients at a prime power grow at most like its $1/8 + O(1/L)$ power, times a divisor factor; the displayed $1/2+\varepsilon$ exponent is a convenient weaker bound. Summing with the coefficient $1/(dh_0k_0)$ gives

$$\mathcal{M}_{\lambda+w\mathbf{1}}(T;f,g) \ll_\varepsilon (1+|\Im w|)^C T^{(a+b)/2+\varepsilon}.$$

Since $(t/2\pi)^{2\Re w} \ll T^{-1/4}$ and G has Gaussian decay, the new contour contributes $O(T^{3/4+(a+b)/2+\varepsilon})$.

The dual diagonal is identical, with $\lambda^\dagger$ in place of λ and with the factor in (2.15). Consequently the two diagonal main terms are

$$\int_{\mathbb{R}} \Phi(t/T)\left(\mathcal{M}_\lambda(T;f,g) + \left(\frac{t}{2\pi}\right)^{-\alpha-\beta-\gamma-\delta}\mathcal{M}_{\lambda^\dagger}(T;f,g)\right)\mathrm{d}t.$$

Their contour errors are absorbed in (2.38), because

$$\frac{a+b}{2} \le \frac{3a+5b}{8}.$$

We explain the contribution corresponding to the (β,δ) swap. The other three follow by permuting α,β and γ,δ. After the elementary replacements already justified, apply (2.28) separately to the positive and negative off-diagonal shifts, with

$$\mathcal{W}(y) = V^*(y/t^2), \qquad F_t^\pm(x,y,r) = \frac{\vartheta(x/y)}{\sqrt{xy}}e^{\mp itr/y}\psi\left(\frac{r^2T^{2-2\rho}}{xy}\right).$$

Here ψ is smooth, equals one near zero, and vanishes outside a fixed compact interval. It implements the localization (2.34); changing the cutoff outside that range is negligible after integration in t. The fixed smooth function ϑ equals one near 1 and is supported where $x/y \asymp 1$. It is one on the retained terms and ensures that $tr/y \ll \Lambda$ throughout the support of the weight. Thus, after extracting its dyadic size, each fixed number of derivatives costs at

most a fixed power of Λ, as required for the Mellin separation above. Since $\vartheta(1) = 1$, it does not change the zero-frequency main term.

The two original signed sums need not be equal. However, their zero-frequency expressions evaluate the weights at $x = y$, so their weights there add to

$$F_t^+(x,x,r) + F_t^-(x,x,r) = \frac{2}{x}\cos(tr/x)\psi\left(\frac{r^2T^{2-2\rho}}{x^2}\right).$$

This is why a cosine occurs without any symmetry assumption on f, g. Taking the Mellin transform in r and putting $r = xv$ gives

$$2x^{w-1}\int_0^\infty \cos(tv)\psi(v^2T^{2-2\rho})v^{w-1}\,\mathrm{d}v.$$

As in [1, Section 2.4, p. 10], the smooth cutoff can be removed in this Mellin calculation. Its transform has residue one at zero; shifting the cutoff contour, or repeatedly integrating the tail by parts, makes the error arbitrarily small after insertion in the remaining integrals. The reason is that the oscillation at the cutoff scale is $T^\rho \to \infty$. The resulting expression is

$$2x^{w-1}\Gamma(w)\cos(\pi w/2)t^{-w}, \tag{2.41}$$

interpreted by meromorphic continuation from $0 < \Re w < 1$.

The remaining x-integral is evaluated directly from (2.17):

$$\begin{aligned}\int_0^\infty V^*\left(\frac{x^2}{d^2h_0k_0t^2}\right)x^{w-\beta-\delta-1}\,\mathrm{d}x \\ = \left(\frac{td\sqrt{h_0k_0}}{2\pi}\right)^{w-\beta-\delta}\frac{G((w-\beta-\delta)/2)}{w-\beta-\delta}.\end{aligned} \tag{2.42}$$

Use also

$$2(2\pi)^{-w}\Gamma(w)\cos(\pi w/2)\zeta(w) = \zeta(1-w).$$

Substitution in (2.28), followed by $w = \beta + \delta + 2z$, gives the following original crossed term, before integration against $\Phi(t/T)\,\mathrm{d}t$:

$$\left(\frac{t}{2\pi}\right)^{-\beta-\delta}\frac{1}{2\pi i}\int_{(c)}\frac{G(z)}{z}\mathcal{R}_\lambda(z)\,\mathrm{d}z, \tag{2.43}$$

where $c > 0$ is fixed and small. Moving the transformed contour to (c) crosses no pole, since the shift-dependent poles lie within $O(L^{-1})$ of zero. The arithmetic factor is

$$\begin{aligned}\mathcal{R}_\lambda(z) = \frac{\zeta(1+\alpha-\beta)\zeta(1+\gamma-\delta)}{\zeta(2+\alpha-\beta+\gamma-\delta)}\zeta(1+\alpha+\gamma+2z)\zeta(1-\beta-\delta-2z) \\ \cdot\sum_{d\ge1}\sum_{(h_0,k_0)=1}\frac{f_{dh_0}g_{dk_0}}{d\,h_0^{1-\beta-z}k_0^{1-\delta-z}}\eta_{\lambda;h_0,k_0}(\beta+\delta+2z).\end{aligned} \tag{2.44}$$

The factors of 2 in the cosine transform and the change of Mellin variable are accounted for by the functional equation and by $\mathrm{d}w = 2\,\mathrm{d}z$. In particular, there is no additional numerical factor in (2.43).

Now apply (2.32). The four zeta factors in (2.44), together with its denominator, are exactly $\mathcal{A}_{C_\lambda(z)}$. The powers $h_0^{\beta+z}k_0^{\delta+z}$ convert the finite Euler factor into the two $\mathcal{B}$ factors. Since $1/(dh_0k_0) = d/(hk)$, we obtain

$$\mathcal{R}_\lambda(z) = \mathcal{M}_{C_\lambda(z)}(T; f, g). \tag{2.45}$$

This is the precise arithmetic identification needed for the contour pairing.

The matching dual term uses $R_\lambda = (-\delta, -\gamma, -\beta, -\alpha)$, a permutation of $\lambda^\dagger$ within its two pairs. Its phase factor, after (2.15), is again $(t/2\pi)^{-\beta-\delta}$. Moreover, $C_{R_\lambda}(z)$ is $C_\lambda(-z)$ up to permutations within the two pairs. The factors Z and the auxiliary function G are invariant under these permutations. Changing z to $-z$ in the dual integral therefore combines the two terms into

$$\left(\frac{t}{2\pi}\right)^{-\beta-\delta} \frac{1}{2\pi i}\left(\int_{(c)} - \int_{(-c)}\right) \frac{G(z)}{z} \mathcal{M}_{C_\lambda(z)}(T; f, g)\,\mathrm{d}z. \tag{2.46}$$

The replacement of the dual gamma factor has a relative error $O(T^{-1})$. On either fixed contour, the same local-factor bounds as above give an integrated error $O(T^{(a+b)/2+\varepsilon})$, which is already admissible.

Besides $z = 0$, the only possible poles between these contours are

$$z = -\frac{\alpha+\gamma}{2}, \qquad z = -\frac{\beta+\delta}{2}.$$

They are canceled by G. The horizontal integrals tend to zero by its Gaussian decay. The residue theorem thus evaluates (2.46) as

$$\left(\frac{t}{2\pi}\right)^{-\beta-\delta} \mathcal{M}_{(\alpha,-\delta,\gamma,-\beta)}(T; f, g).$$

Permuting the shifts within the two pairs gives the remaining three one-swap terms in (2.8). Together with the two diagonal terms, these are exactly $\mathfrak{M}_\lambda(T; f, g)$. We have therefore proved (2.11) uniformly on compact separated scaled shift sets.

We finally justify the continuation used in the statement and remove the separation assumption. Put

$$x = (-\alpha, -\beta, \gamma, \delta).$$

The six summands are indexed by the six choices of two of these four coordinates for the first shift pair. Each individual summand has possible simple poles only on the hyperplanes $x_i = x_j$. At a generic point of such a hyperplane, pair the terms obtained by exchanging x_i and x_j between the two pairs. Their phase factors and all regular Euler factors agree on the hyperplane, whereas the polar factors are $\zeta(1 + x_i - x_j)$ and $\zeta(1 - x_i + x_j)$, with opposite residues. Thus the residues cancel.

To include intersections of hyperplanes, multiply the sum by $\prod_{i<j}(x_j - x_i)$. The product is holomorphic and vanishes on each of these hyperplanes by the preceding cancellation. It is therefore divisible by each linear factor, and hence by their product. This proves removability also at simultaneous collisions. It is the holomorphy mechanism noted after [1, Theorem 1.1]. In particular, the complete main term is holomorphic in all four shifts near the origin.

Fix now an arbitrary compact $K \subset \mathbb{C}^4$ of scaled shifts. Choose a polydisc centered at zero which contains K in its interior and whose four coordinate radii are pairwise distinct. On its distinguished boundary, the signed coordinates in (2.12) are uniformly separated, since the modulus of their difference is at least the difference of their radii. For large T, both the moment and the complete main term are holomorphic throughout this polydisc in the scaled variables. Applying the maximum-modulus principle successively in the four variables extends (2.11) from its distinguished boundary to K.

To see that the error is $o(T)$, put $\Delta = 2 - 3a - 5b > 0$. Then

$$\frac{3}{4} + \frac{3a+5b}{8} = 1 - \frac{\Delta}{8}, \qquad \frac{1}{2} + a + b = 1 - \frac{\Delta + b - a}{4} \le 1 - \frac{\Delta}{4}.$$

Taking, for example, $0 < \varepsilon < \Delta/16$ proves (2.10) uniformly on K.

Finally, the reality of f, g, Φ gives

$$I_\lambda(T; g, f) = \overline{I_{\overline{\lambda^*}}(T; f, g)}.$$

The complete main term has the same symmetry, after interchanging h, k in its coefficient sums. This proves the assertion with the profiles reversed. □

In particular, the formula applies locally uniformly to $I_{(0,0,z/L,0)}(T; g, f)$, which is the orientation needed for the U-numerator. Cauchy's integral formula also permits any fixed number of derivatives with respect to the scaled shifts. The next subsection uses this uniformity to combine the six main terms before specializing the shifts.

2.2.3 The U-numerator: from six swaps to the simplex kernel

To obtain the corresponding limits for the polynomial weights p, we first consider the generating moment

$$U_T(z) = \int_{\mathbb{R}} \Phi(t/T)\zeta(s)^3\overline{\chi(s)}A_f(\bar{s})A_g(s)e^{z/2}\zeta(\bar{s} + z/L)\, \mathrm{d}t.$$

The polynomial weights will be recovered by applying $p(\partial_z)$ at $z = 0$.

Our first step is to express $U_T(z)$ in terms of the asymmetric twisted fourth moment established in Section 2.2.2. By the functional equation, the product of zeta factors in $U_T(z)$ becomes

$$\zeta(s)^2\zeta(\bar{s})\zeta(\bar{s} + z/L),$$

with two factors on each side of the critical line. Thus, in the notation of the twisted fourth moment,

$$U_T(z) = e^{z/2}I_{(0,0,z/L,0)}(T; g, f). \tag{2.47}$$

Notice that the profiles appear in the order (g, f), since $A_g(s)$ and $A_f(\bar{s})$ occur on opposite sides. The asymmetric fourth-moment formula allows precisely this ordering.

We can therefore apply that formula to $U_T(z)$. It remains to extract the leading L^9 contribution from its six main terms and show that their sum gives the common kernel $K_z(a, b)$ introduced in Section 2.2.1. The resulting locally uniform limit in z will then allow us to recover the U-numerator formula by differentiation at $z = 0$.

We first work with separated scaled shifts $\sigma = (\sigma_1, \dots, \sigma_4)$, so that the actual shifts are given by $\lambda = \sigma/L$. This avoids the vanishing denominators that may occur in the individual main terms when shift parameters coincide. Although these terms may appear singular, their sum will admit a regular continuation to the shift configuration required for the U-numerator.

For convenience, we introduce

$$D(\sigma) = \prod_{i=1}^{2}\prod_{j=3}^{4}(\sigma_i + \sigma_j)$$

and

$$E_{r,s}(x) = \begin{cases} \dfrac{e^{-rx} - e^{-sx}}{s - r}, & r \neq s, \\ xe^{-rx}, & r = s. \end{cases}$$

The four factors in $D(\sigma)$ arise from the polar zeta factors associated with the two pairs of shifts. The function $E_{r,s}$ will appear naturally when we invert the remaining factors involving the amplifier profiles.

Let $\sigma^{(j)}$ and $e_j(\sigma)$, for $0 \leq j \leq 5$, denote the transformed shifts and the corresponding exponents from (2.8). Each of these six configurations contributes one main term.

To identify their leading contributions, we use the polar approximation

$$\frac{1}{L}\zeta\left(1+\frac{w}{L}\right) \longrightarrow \frac{1}{w}, \qquad L \to \infty,$$

for fixed $w \neq 0$. For the unswapped term, the resulting polar factors have the form

$$\frac{1}{D(\sigma)(u+v)(u+\sigma_3)(u+\sigma_4)(v+\sigma_1)(v+\sigma_2)}.$$

There are nine denominator factors in total, which accounts for the leading power L^9.

The variables u and v arise from Mellin inversion of the two amplifier profiles. For a profile h, we use the convention

$$\widetilde{h}(w) = \int_0^\infty h(x)e^{wx}\,\mathrm{d}x, \qquad h(\log n/L) = \frac{1}{2\pi i}\int_{(c)} \widetilde{h}(w)n^{-w/L}\,\mathrm{d}w.$$

Here u and v correspond to the profiles g and f, respectively. The factors involving these variables can be inverted using

$$\frac{1}{u+v} = \int_0^\infty e^{-t(u+v)}\,\mathrm{d}t$$

and

$$\mathcal{L}^{-1}\left(\frac{1}{(u+r)(u+s)}\right)(x) = E_{r,s}(x).$$

The first identity introduces a common integration variable t, while the second produces the two functions $E_{r,s}$ associated with the profiles.

Let A and B denote the logarithmic positions associated with the profiles f and g, respectively. Since we apply the twisted fourth-moment formula to $I_\lambda(T; g, f)$, the first amplifier has position B and the second has position A. Applying these identities to each of the six shifted main terms gives the following leading kernel:

$$\int_0^{\min(A,B)} \sum_{j=0}^{5} \frac{e^{e_j(\sigma)}}{D(\sigma^{(j)})} \times E_{\sigma_3^{(j)},\sigma_4^{(j)}}(B-t)E_{\sigma_1^{(j)},\sigma_2^{(j)}}(A-t)\,\mathrm{d}t. \tag{2.48}$$

Here the exponential factor $e^{e_j(\sigma)}$ arises from the corresponding power $(\tau/2\pi)^{e_j(\sigma)/L}$ in the twisted fourth-moment formula. Since $\tau \asymp T$ on the support of the smoothing function, this factor converges to $e^{e_j(\sigma)}$.

The expression in (2.48) gives the leading contribution for fixed logarithmic positions A and B. To recover the main term of the twisted fourth moment, we integrate this kernel against the two amplifier profiles, as follows from their Mellin inversion.

The integration against the smoothing function contributes $T\widehat{\Phi}(0)$, while the regular part of the Euler product gives the arithmetic constant a_3. Together with the previously obtained factor L^9 and the six swap contributions, this yields $a_3TL^9\widehat{\Phi}(0)$ times the profile average of (2.48), up to an error of $o(TL^9)$.

It remains to combine the six summands into a determinant representation, which will allow us to pass to the required coincident shifts and recover the simplex kernel.

To justify the limiting Mellin calculation, including its tails, fix a compact set K of scaled shifts such that $D(\sigma^{(j)}) \neq 0$ for $0 \leq j \leq 5$. Choose $R > 0$ bounding all components of the six transformed shift configurations on K, and take $c > R + 1$. On the vertical contours $u = c + i\xi$, $v = c + iv$, the five arguments involving u, v in the unswapped term are

$$\mathcal{W} = \{u+v, u+\sigma_3, u+\sigma_4, v+\sigma_1, v+\sigma_2\}.$$

For $w \in \mathcal{W}$, we have $\Re w > 1$, and absolute convergence gives, for $L \geq 1$,

$$|L^{-1}\zeta(1 + w/L)| \leq L^{-1}\zeta(1 + \Re w/L) \leq L^{-1} + (\Re w)^{-1} \leq 2.$$

The same bound applies to all six swap configurations.

The four remaining zeta factors depend only on the shifts. Since $D(\sigma^{(j)}) \neq 0$ on the compact set K, their scaled arguments stay uniformly away from zero, so these factors are uniformly bounded after division by L.

As in the proof of Lemma 2, cancellation of the first-order terms in the local Euler factors yields $1 + O(\ell^{-2+\delta})$ for some fixed $0 < \delta < 1$. This estimate is uniform on the Mellin contours for sufficiently large T, since the real parts of the scaled shifts are $O_K(L^{-1})$ and $|\ell^{-w}| = \ell^{-\Re w}$. Consequently, the regular Euler product is uniformly bounded. Since $f, g \in C_c^\infty$, their Mellin transforms satisfy

$$|\widetilde{g}(c + i\xi)| \ll_{g,c} (1 + |\xi|)^{-2}, \qquad |\widetilde{f}(c + i\upsilon)| \ll_{f,c} (1 + |\upsilon|)^{-2}.$$

Combining these estimates with the preceding bounds for the five Mellin-dependent zeta factors, the normalized integrand is bounded in absolute value by

$$\frac{C_{K,f,g}}{(1 + |\xi|)^2(1 + |\upsilon|)^2}, \tag{2.49}$$

uniformly for $\sigma \in K$ and all sufficiently large T. The right-hand side is integrable over $\mathbb{R}^2$.

On bounded segments of the Mellin contours, the nine normalized zeta factors converge uniformly to the reciprocals of their arguments. The regular Euler factor is holomorphic near the origin and takes the value a_3 there, as established by the nine-edge Euler-product calculation in Lemma 2. Since all its scaled arguments tend uniformly to zero, it also converges uniformly to a_3.

The majorant in (2.49) controls the remaining tails uniformly for $\sigma \in K$. Together with the uniform convergence on bounded contour segments, this gives uniform convergence of the normalized Mellin integrals on K. Restoring the main-term prefactor, the resulting error is $o(TL^9)$.

We now combine the six summands in (2.48) into a determinant. Put

$$x = (-\sigma_1, -\sigma_2, \sigma_3, \sigma_4), \qquad \rho(u, v) = (-1, v - 1, -u, 0),$$

and define the Vandermonde determinant

$$\Delta(x) = \prod_{0 \leq i < j \leq 3} (x_j - x_i).$$

Then the six-term expression admits the representation

$$\begin{aligned} \sum_{j=0}^{5} \frac{e^{e_j(\sigma)}}{D(\sigma^{(j)})} E_{\sigma_3^{(j)},\sigma_4^{(j)}}(B - t) E_{\sigma_1^{(j)},\sigma_2^{(j)}}(A - t) \\ = e^{x_0+x_1} \frac{\det\big(e^{x_i \rho_k(B-t,A-t)}\big)_{i,k=0}^{3}}{\Delta(x)}. \end{aligned} \tag{2.50}$$

To see the correspondence with the six swaps, expand the determinant along its first two columns. There are $\binom{4}{2} = 6$ choices of two rows. For the choice of rows $0, 1$, writing $u = B - t$ and $v = A - t$, the corresponding contribution is

$$\frac{e^{x_0+x_1}}{\Delta(x)} (e^{-x_0+x_1(v-1)} - e^{-x_1+x_0(v-1)})(e^{-x_2 u} - e^{-x_3 u}).$$

Since

$$\Delta(x) = (\sigma_1 - \sigma_2)(\sigma_4 - \sigma_3)D(\sigma),$$

this simplifies to

$$\frac{E_{\sigma_3,\sigma_4}(B-t)E_{\sigma_1,\sigma_2}(A-t)}{D(\sigma)},$$

which is precisely the unswapped term. The remaining five choices give the five swapped configurations, including their corresponding exponential factors.

The determinant in (2.50) is alternating in $x_0, \dots, x_3$. Consequently, its quotient by $\Delta(x)$ extends jointly holomorphically across $\Delta(x) = 0$. In particular, the apparent singularities of the combined six-term expression are removable.

We now evaluate this continuation at the required shift configuration

$$\sigma = (0, 0, z, 0), \qquad x = (0, 0, z, 0).$$

For this purpose, write $\rho = \rho(B-t, A-t)$. The confluent Hermite–Genocchi divided-difference identity gives

$$e^{z/2} \left.\frac{\det(e^{x_i\rho_j})_{i,j=0}^3}{\Delta(x)}\right|_{x=(0,0,z,0)} = \frac{\Delta(\rho)}{12}\mathbb{E}e^{z(\rho+\frac{1}{2}\mathbf{1})\cdot X}. \tag{2.51}$$

To verify the normalization, first suppose that $z \neq 0$ and the four components of ρ are distinct. In the confluent limit, the three rows corresponding to zero shifts become the rows with entries 1, ρ_j, and $\rho_j^2/2$, respectively. Moving the row corresponding to z to the bottom introduces a minus sign, which cancels the sign of the remaining Vandermonde factors. Their net contribution to the denominator is therefore z^3.

Thus

$$\left.\frac{\det(e^{x_i\rho_j})_{i,j=0}^3}{\Delta(x)}\right|_{x=(0,0,z,0)} = \frac{1}{z^3}\det\begin{pmatrix} 1 & 1 & 1 & 1 \\ \rho_0 & \rho_1 & \rho_2 & \rho_3 \\ \rho_0^2/2 & \rho_1^2/2 & \rho_2^2/2 & \rho_3^2/2 \\ e^{z\rho_0} & e^{z\rho_1} & e^{z\rho_2} & e^{z\rho_3} \end{pmatrix}.$$

Expanding along the last row gives

$$\left.\frac{\det(e^{x_i\rho_j})_{i,j=0}^3}{\Delta(x)}\right|_{x=(0,0,z,0)} = \frac{\Delta(\rho)}{2z^3}\sum_{j=0}^{3}\frac{e^{z\rho_j}}{\prod_{k\neq j}(\rho_j - \rho_k)}.$$

On the other hand, the divided-difference formula for the normalized simplex expectation yields

$$\mathbb{E}e^{z\rho\cdot X} = \frac{6}{z^3}\sum_{j=0}^{3}\frac{e^{z\rho_j}}{\prod_{k\neq j}(\rho_j - \rho_k)}.$$

Comparing these two expressions and using $\sum_{j=0}^3 X_j = 1$ gives (2.51). The identity extends to coincident components of ρ and to $z = 0$ by continuity. For the general divided-difference formula, see [3, formula (83)].

We return to the logarithmic positions $A = a$ and $B = b$. By the definitions of ρ, P and q,

$$\Delta(\rho(b-t, a-t)) = 12P(b-t, a-t),$$

and

$$\rho(b-t, a-t) + \frac{1}{2}\mathbf{1} = q(a, b, t).$$

Furthermore, $e^{x_0+x_1}=1$ at the required shift configuration. Thus, including the factor $e^{z/2}$ from (2.47), the limiting six-swap kernel becomes

$$\begin{aligned}&\int_0^{\min(a,b)} P(b-t,a-t)\mathbb{E}e^{zq(a,b,t)\cdot X}\,\mathrm{d}t\\&\qquad=K_z(a,b),\end{aligned}$$

exactly as defined in (2.2).

The holomorphic shift argument of Section 2.2.2 extends the locally uniform shifted moment asymptotic to the coincident configuration. Consequently,

$$\frac{U_T(z)}{a_3TL^9}\longrightarrow\widehat{\Phi}(0)\int_0^\infty\int_0^\infty f(a)g(b)K_z(a,b)\,\mathrm{d}a\,\mathrm{d}b,$$

locally uniformly in z.

Finally, Cauchy's integral formula allows us to apply $p(\partial_z)$ and then set $z=0$. By the definition of K_p in (2.3), this yields

$$k_U(p;f,g)=\int_0^\infty\int_0^\infty f(a)g(b)K_p(a,b)\,\mathrm{d}a\,\mathrm{d}b,$$

which proves the U-numerator formula in Proposition 3.

2.2.4 The V-numerator: the Durkan–Page formula

To prove the V-numerator formula, we first consider the generating moment

$$V_T(z)=\int_{\mathbb{R}}\Phi(t/T)\zeta(s)^3A_f(\bar{s})A_g(\bar{s})e^{z/2}\zeta(\bar{s}+z/L)\,\mathrm{d}t.$$

As in the U-case, the polynomial weights will be recovered by applying $p(\partial_z)$ at $z=0$.

We use the short-polynomial twisted fourth-moment formula of Durkan–Page [4, Theorem 16, $k=1$, and Section 5.2]. For scaled shifts $\sigma=(\sigma_1,\dots,\sigma_4)$, define

$$\mathcal{J}_\sigma^-(T;f,g)=\int_{\mathbb{R}}\Phi(t/T)\chi(s)\prod_{i=1}^2\zeta(s+\sigma_i/L)\prod_{j=3}^4\zeta(\bar{s}+\sigma_j/L)A_f(\bar{s})A_g(\bar{s})\,\mathrm{d}t.$$

This is the conjugate orientation of the moment $J_{1,1}$ in [4], with the shifts relabeled so that σ_1,σ_2 correspond to the s-factors. By the functional equation, $\chi(s)\zeta(\bar{s})=\zeta(s)$, so

$$V_T(z)=e^{z/2}\mathcal{J}^-_{(0,0,0,z)}(T;f,g).\tag{2.52}$$

Durkan–Page use amplifier coefficients of the form $P(\log(T^\eta/n)/\log T^\eta)$, which correspond to our profiles $f(a)=P(1-a/\eta)$. Their formula, stated for a squared amplifier, extends to two independent profiles by polarization:

$$A_f(w)A_g(w)=\frac{1}{4}(A_{f+g}(w)^2-A_{f-g}(w)^2).$$

Since both profiles have the same cutoff $\eta<1/4$, the polynomials $f+g$ and $f-g$ satisfy the same length restriction.

We now express the four residues in the notation of our profiles. Put

$$Q_\sigma(w)=(w+\sigma_1)(w+\sigma_2)(w-\sigma_3)(w-\sigma_4)$$

and define

$$F_\sigma(a,\omega)=\frac{1}{2\pi i}\oint \frac{e^{au}(u-\omega)}{Q_\sigma(u)}\,\mathrm{d}u.$$

Here and below, the contours are positively oriented circles enclosing all four roots $-\sigma_1,-\sigma_2,\sigma_3,\sigma_4$. For shifts in a fixed compact set, their radii can be chosen independently of the shifts.

The resulting residue kernel is

$$R_\sigma(a,b)=\frac{e^{-\sigma_3-\sigma_4}}{2\pi i}\oint \frac{e^{\omega}F_\sigma(a,\omega)F_\sigma(b,\omega)}{Q_\sigma(\omega)}\,\mathrm{d}\omega. \tag{2.53}$$

We first verify how this representation arises from the four-residue formula in [4]. For separated shifts, consider the outer residue at $\omega=-\sigma_1$. Its coefficient is

$$\begin{aligned} &e^{-\sigma_3-\sigma_4}\operatorname*{Res}_{\omega=-\sigma_1}\frac{e^{\omega}}{Q_\sigma(\omega)}\\ &\quad=\frac{e^{-\sigma_1-\sigma_3-\sigma_4}}{(\sigma_2-\sigma_1)(\sigma_1+\sigma_3)(\sigma_1+\sigma_4)}. \end{aligned}$$

At this pole, the factor $u+\sigma_1$ cancels in the inner contour, giving

$$F_\sigma(a,-\sigma_1)=\frac{1}{2\pi i}\oint \frac{e^{au}}{(u+\sigma_2)(u-\sigma_3)(u-\sigma_4)}\,\mathrm{d}u.$$

By inverse Laplace transformation,

$$F_\sigma(a,-\sigma_1)=\int_{\substack{x,y\ge 0\\ x+y\le a}} e^{-\sigma_2 x+\sigma_3 y+\sigma_4(a-x-y)}\,\mathrm{d}x\,\mathrm{d}y.$$

Integrating against $f(a)$ and substituting $v=a-x-y$, we obtain

$$\begin{aligned} &\int_0^\eta f(a)F_\sigma(a,-\sigma_1)\,\mathrm{d}a\\ &\quad=\int_{\substack{x,y,v\ge 0\\ x+y+v\le\eta}} f(x+y+v)e^{-\sigma_2 x+\sigma_3 y+\sigma_4 v}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}v. \end{aligned}$$

The corresponding expression for g has the same form.

After rescaling the simplex variables by η, these integrals agree with the two profile integrals in the first residue of the Durkan–Page formula for $k=1$, up to a Jacobian factor η^3 for each profile. The residues at $-\sigma_2,\sigma_3,\sigma_4$ reproduce the remaining three terms in the same way. Thus (2.53) combines all four residues into a single contour expression.

We next check the normalization. Writing $y=T^\eta$, the $k=1$ formula of Durkan–Page contains the factor $(\log y)^6$. The three shift differences in each residue contribute L^3 when the original shifts are written as $\alpha_i=\sigma_i/L$. Rescaling the two three-dimensional simplex integrals contributes η^{-6}. Hence

$$(\log y)^6L^3\eta^{-6}=(\eta L)^6L^3\eta^{-6}=L^9.$$

Moreover,

$$\left(\frac{T}{2\pi}\right)^{-(\alpha_3+\alpha_4)}=e^{-\sigma_3-\sigma_4}(1+O(L^{-1})),$$

uniformly for bounded scaled shifts. The arithmetic factor is a_3.

The error in [4, Theorem 16] is $O(TL^8)$ for fixed polynomial profiles and $\eta<1/4$. Polarization preserves this error bound. The convention at the single endpoint $n=T^\eta$, when

this is an integer, does not affect the leading term: the possible extra summand has size $O(T^{-\eta/2})$ on the critical line, while the other amplifier is $O(T^{\eta/2})$. Hence the difference between the two amplifier products is $O(1)$. Hölder's inequality and the shifted fourth-moment bound show that its contribution is $O(TL^4) = o(TL^9)$.

Consequently, for separated shifts,

$$\frac{\mathcal{J}_\sigma^-(T;f,g)}{a_3 TL^9} = \widehat{\Phi}(0)\int_0^\eta \int_0^\eta f(a)g(b)R_\sigma(a,b)\,\mathrm{d}a\,\mathrm{d}b + o(1). \tag{2.54}$$

We now extend this limit to coincident shifts. For shifts in a fixed compact set, the contours in (2.53) may be kept fixed. Their integrands are holomorphic in σ and uniformly bounded on the contours. Thus $R_\sigma(a,b)$ depends holomorphically on σ, including when roots of Q_σ coincide.

For the moment asymptotic, choose a polydisc containing the compact shift set, with distinct coordinate radii. On its distinguished boundary, the four roots remain uniformly separated. The error estimate from Durkan–Page is uniform there, and the difference between the normalized moment and its proposed main term is holomorphic in the shifts. The multidimensional Cauchy integral formula therefore extends the estimate to the interior. In particular, (2.54) holds locally uniformly in σ, including at coincident shifts.

We can now specialize to the shifts required for $V_T(z)$. Define the entire functions

$$\begin{aligned}
B_a(z) &= \int_0^a (a-u)e^{-zu}\,\mathrm{d}u,\\
C_a(z) &= -\frac{1}{2}\int_0^a (a-u)^2 e^{-zu}\,\mathrm{d}u,\\
H_m(z) &= \frac{1}{m!}\int_0^1 (1-u)^m e^{-zu}\,\mathrm{d}u, \qquad m=0,1,2.
\end{aligned}$$

For convenience, define

$$\begin{aligned}
J(a,b;z) = {} & B_a(z)B_b(z)H_2(z)\\
& + \big(B_a(z)C_b(z) + C_a(z)B_b(z)\big)H_1(z)\\
& + C_a(z)C_b(z)H_0(z).
\end{aligned} \tag{2.55}$$

The inverse Laplace identities

$$\frac{1}{2\pi i}\oint \frac{e^{au}}{u^2(u+z)}\,\mathrm{d}u = B_a(z), \qquad -\frac{1}{2\pi i}\oint \frac{e^{au}}{u^3(u+z)}\,\mathrm{d}u = C_a(z)$$

give

$$\frac{1}{2\pi i}\oint \frac{e^{au}(u-\omega)}{u^3(u+z)}\,\mathrm{d}u = B_a(z) + \omega C_a(z).$$

Similarly,

$$\frac{1}{2\pi i}\oint \frac{e^{\omega}}{\omega^{m+1}(\omega+z)}\,\mathrm{d}\omega = H_m(z), \qquad m=0,1,2.$$

It follows that

$$\begin{aligned}
J(a,b;z) = \frac{1}{2\pi i}\oint & \frac{e^{\omega}}{\omega^3(\omega+z)}\\
& \cdot \big(B_a(z)+\omega C_a(z)\big)\big(B_b(z)+\omega C_b(z)\big)\,\mathrm{d}\omega,
\end{aligned} \tag{2.56}$$

where the contour encloses 0 and $-z$. Indeed, expansion of the numerator produces exactly the three terms in (2.55).

At $\sigma = (0,0,0,z)$, we have

$$Q_\sigma(w) = w^3(w-z)$$

and hence

$$F_\sigma(a,\omega) = B_a(-z) + \omega C_a(-z).$$

Substitution into (2.53) and comparison with (2.56) yield

$$R_{(0,0,0,z)}(a,b) = e^{-z} J(a,b;-z).$$

Finally, combining (2.54) with (2.52), we obtain

$$\frac{V_T(z)}{a_3 T L^9} \longrightarrow \widehat{\Phi}(0) \int_0^\eta \int_0^\eta f(a)g(b)e^{-z/2} J(a,b;-z)\,\mathrm{d}a\,\mathrm{d}b, \tag{2.57}$$

locally uniformly near $z = 0$.

2.2.5 The reflection identity and completion of the numerator

The preceding calculation expresses the V-numerator in terms of the kernel J. We now show that this kernel agrees with the common numerator kernel K_z after reflecting its second argument.

Lemma 5 (Reflection)**.** *For $a, b > 0$ with $a + b < 1$, we have*

$$e^{-z/2} J(a,b;-z) = K_z(a, 1-b) \qquad (z \in \mathbb{C}). \tag{2.58}$$

Proof. Both sides are entire in z, so it suffices to establish the identity for $z \neq 0$.

Write

$$E(a,z) = e^{az} - 1 - az$$

and define

$$\begin{aligned} N(a,b,z) = {} & z(1-z/2)E(a,z)E(b,z) \\ & - \frac{z^3}{2}\left(a^2 E(b,z) + b^2 E(a,z)\right) \\ & + \frac{a^2 b^2 z^4}{4}(e^z - 1). \end{aligned}$$

We first evaluate $J(a,b;-z)$. From the definitions of B_a, C_a and H_m in the preceding subsection,

$$B_a(-z) = \frac{E(a,z)}{z^2}, \qquad C_a(-z) = -\frac{E(a,z) - a^2 z^2/2}{z^3},$$

and

$$H_m(-z) = \frac{e^z - \sum_{j=0}^m z^j/j!}{z^{m+1}}, \qquad m = 0,1,2.$$

Substituting these expressions into the definition of J and collecting terms gives

$$z^7 J(a,b;-z) = N(a,b,z). \tag{2.59}$$

We next evaluate $K_z(a, 1-b)$. Since $a + b < 1$, we have $\min(a, 1-b) = a$. For $0 \le t \le a$, the four translated simplex nodes are

$$0, \qquad a - t, \qquad b + t, \qquad 1.$$

By the divided-difference formula for the simplex expectation in the definition of K_z, we obtain

$$\begin{aligned}2z^3e^{z/2}K_z(a,1-b) = \int_0^a \big[&-(b-a+2t)(1-a+t)(1-b-t)\\ &+e^{z(a-t)}(b+t)(1-b-t)\\ &-e^{z(b+t)}(a-t)(1-a+t)\\ &+e^{z}(a-t)(b+t)(b-a+2t)\big]\,\mathrm{d}t.\end{aligned}$$

This expression remains valid when two simplex nodes coincide, by continuity.

Integrating the polynomial and exponential terms gives

$$2z^3e^{z/2}K_z(a,1-b) = \frac{2N(a,b,z)}{z^4}.$$

Consequently,

$$e^{z/2}K_z(a,1-b) = \frac{N(a,b,z)}{z^7}. \tag{2.60}$$

Comparing (2.59) and (2.60), we conclude that

$$e^{-z/2}J(a,b;-z) = K_z(a,1-b)$$

for $z \neq 0$. Since both sides are entire, the identity also holds at $z = 0$. □

We can now complete the proof of Proposition 3. Combining the locally uniform limit (2.57) with Lemma 5 yields

$$\frac{V_T(z)}{a_3TL^9} \longrightarrow \widehat{\Phi}(0)\int_0^\eta\int_0^\eta f(a)g(b)K_z(a,1-b)\,\mathrm{d}a\,\mathrm{d}b.$$

The convergence is locally uniform near $z = 0$. Hence Cauchy's integral formula allows us to apply $p(\partial_z)$ and then set $z = 0$. By the definition of K_p, this gives

$$k_V(p;f,g) = \int_0^\eta\int_0^\eta f(a)g(b)K_p(a,1-b)\,\mathrm{d}a\,\mathrm{d}b.$$

This proves the V-numerator formula. Together with the U-numerator limit established in Section 2.2.3, it completes the proof of Proposition 3.

The reflection identity concerns the limiting kernels only; the analytic V-moment estimate still requires the short-length condition $\eta < 1/4$.

2.3 Denominator

Unlike the numerator, which only involves correlations of individual tests with $\zeta(s)^3$, the denominator requires all pairwise products of the projected tests, including interactions between the two families U and V. By the projection identity (1.1), each Gram entry splits into an ordinary and a reflected product. Using the functional equation, these products reduce to two types of moments: balanced moments, with two amplifier factors on each side, and unbalanced moments, with three factors on one side and one on the other. The only exception is the reflected VV product, whose oscillatory contribution will be shown to be negligible.

We first establish the necessary approximation for the functional equation of centered derivatives, which introduces the polynomial reflection $p(x) \mapsto p(-x)$. We then evaluate the balanced and unbalanced moments separately. In both cases, we isolate the diagonal

contribution, control the off-diagonal terms, and apply the nine-edge diagonal limit from Lemma 2. The balanced case uses a smoothed approximate functional equation and produces an additional length constraint, whereas the unbalanced case uses a one-sided expansion. This leads to two explicit kernels, $B_{p,q}$ and $M_{p,q}$, from which we obtain all entries of the Gram matrix.

Throughout this section, the real profiles f, g, h, k are continuous and supported in $[0, A_i]$, for $i = 1, 2, 3, 4$, respectively, with

$$S_* := A_1 + A_2 + A_3 + A_4 < 1.$$

This condition holds for every pair of admissible tests from Section 1 and ensures sufficient control of the off-diagonal terms. The variables a, b, c, d below are integration variables, not support bounds. The polynomials p, q have degree at most four.

2.3.1 The functional-equation factor

Lemma 6. *For every fixed $R > 0$ and integer $j \geq 0$, we have*

$$\partial_z^j (R_T(t, z) - 1) \ll_{R,j} L^{-1}$$

uniformly for $T \leq t \leq 2T$ and $|z| \leq R$.

Proof. By Stirling's formula, the logarithmic derivative of χ satisfies

$$\frac{\chi'}{\chi}(s + w) = -\log \frac{t}{2\pi} + O(T^{-1})$$

uniformly for $T \leq t \leq 2T$ and $|w| \leq R/L$. Consequently,

$$\begin{aligned} \log R_T(t, z) &= z + \int_0^{z/L} \frac{\chi'}{\chi}(s + w)\, \mathrm{d}w \\ &= z - \frac{z}{L} \log \frac{t}{2\pi} + O_R\left(\frac{1}{TL}\right) \\ &= -\frac{z}{L} \log \frac{t}{2\pi T} + O_R\left(\frac{1}{TL}\right). \end{aligned}$$

Since $t/T \in [1, 2]$, the logarithm in the last expression is bounded. Thus

$$R_T(t, z) = 1 + O_R(L^{-1})$$

uniformly on every fixed disc. The bounds for the derivatives follow from Cauchy's integral formula, applied on a slightly larger fixed disc. □

Replacing $R_T(t, z)$ by 1 in the centered functional equation formally corresponds to the polynomial reflection $p \mapsto p^-$. This replacement is not an exact identity between centered derivatives. In the moment calculations below, we justify it by showing that the resulting error is negligible after integration.

2.3.2 The balanced kernel

For pairs of tests from the same family, the terms $H\overline{J}$ in (1.1) take the form

$$\begin{aligned} V_{p;f,g}\overline{V_{q;h,k}} &= C_p(s)C_q(\bar{s})\, A_f(s)A_g(s)A_h(\bar{s})A_k(\bar{s}), \\ U_{p;f,g}\overline{U_{q;h,k}} &= C_p(s)C_q(\bar{s})\, A_f(s)A_k(s)A_g(\bar{s})A_h(\bar{s}). \end{aligned}$$

Thus both are instances of the same balanced moment, up to a relabeling of the profiles.

For $a, b, c, d \geq 0$ with $a + b + c + d < 1$, define

$$\mathcal{R}(a, b; c, d) = \left\{ r \in \mathbb{R}_{\geq 0}^{2\times 2} : \begin{matrix} r_{11} + r_{12} \leq a, & r_{21} + r_{22} \leq b, \\ r_{11} + r_{21} \leq c, & r_{12} + r_{22} \leq d \end{matrix} \right\}.$$

Put

$$t_r = \sum_{i,j} r_{ij}, \qquad \Delta = \frac{a + b - c - d}{2}, \qquad w_r = \frac{1 - a - b - c - d + 2t_r}{2},$$

and define

$$B_{p,q}(a, b; c, d) = \int_{\mathcal{R}(a,b;c,d)} \int_{-w_r}^{w_r} p(\Delta - x) q(-\Delta - x) \, \mathrm{d}x \, \mathrm{d}r. \tag{2.61}$$

Proposition 7 (Balanced moment)**.** *The normalized integral*

$$\frac{1}{a_3 T L^9} \int_{\mathbb{R}} \Phi(t/T) C_p(s) C_q(\bar{s}) A_f(s) A_g(s) A_h(\bar{s}) A_k(\bar{s}) \, \mathrm{d}t$$

converges to

$$\widehat{\Phi}(0) \int f(a) g(b) h(c) k(d) B_{p,q}(a, b; c, d) \, \mathrm{d}a \, \mathrm{d}b \, \mathrm{d}c \, \mathrm{d}d.$$

Proof. We begin by introducing two complex shifts in order to handle the centered derivatives. For z, w in fixed complex discs, put

$$s_1 = s + \frac{z}{L}, \qquad s_2 = \bar{s} + \frac{w}{L},$$

and consider the shifted integral

$$I(z, w) := \int_{\mathbb{R}} \Phi(t/T) e^{(z+w)/2} \zeta(s_1) \zeta(s_2) \cdot A_f(s) A_g(s) A_h(\bar{s}) A_k(\bar{s}) \, \mathrm{d}t.$$

By the definition of the centered derivatives in Section 1, the integral in Proposition 7 is exactly

$$p(\partial_z) q(\partial_w) I(z, w)|_{z=w=0}.$$

It therefore suffices to analyze $I(z, w)$ uniformly for bounded z, w, and to apply the differential operators at the end of the proof.

To evaluate $I(z, w)$, we use a smoothed approximate functional equation for the product $\zeta(s_1)\zeta(s_2)$. This expresses the product as the sum of two rapidly convergent double Dirichlet series, one coming directly from the Dirichlet series and the other from the functional equation.

We first specify the smoothing. Write

$$\Gamma_{\mathbb{R}}(v) = \pi^{-v/2} \Gamma(v/2),$$

so that

$$\Gamma_{\mathbb{R}}(v) \zeta(v) = \Gamma_{\mathbb{R}}(1 - v) \zeta(1 - v).$$

For $\nu = 1, 2$, define

$$s_\nu^+ = s_\nu, \qquad s_\nu^- = 1 - s_\nu.$$

We choose the even smoothing function

$$G_{z,w}(u) = e^{u^2} \prod_{\nu=1}^{2} \left(1 - \frac{u^2}{s_\nu^2}\right) \left(1 - \frac{u^2}{(1 - s_\nu)^2}\right).$$

It satisfies $G_{z,w}(0) = 1$. The polynomial factors cancel the poles of the completed zeta functions encountered when shifting the Mellin contour.

For $x > 0$, define the two weights

$$W^{\pm}_{z,w}(x;t) = \frac{1}{2\pi i}\int_{(1)} G_{z,w}(u)\frac{\Gamma_{\mathbb{R}}(s_1^{\pm}+u)\Gamma_{\mathbb{R}}(s_2^{\pm}+u)}{\Gamma_{\mathbb{R}}(s_1^{\pm})\Gamma_{\mathbb{R}}(s_2^{\pm})}x^{-u}\frac{\mathrm{d}u}{u}.$$

Here $\int_{(1)}$ denotes integration along the vertical line $\Re u = 1$, oriented upwards.

The smoothed approximate functional equation now gives

$$\begin{aligned}\zeta(s_1)\zeta(s_2) = &\sum_{m,n\geq 1}\frac{W^{+}_{z,w}(mn;t)}{m^{s_1}n^{s_2}}\\ &+\chi(s_1)\chi(s_2)\sum_{m,n\geq 1}\frac{W^{-}_{z,w}(mn;t)}{m^{1-s_1}n^{1-s_2}}.\end{aligned}$$

With our choice of the pole-corrected smoothing function, this is an exact identity. Indeed, it follows by shifting the Mellin contour from $\Re u = 1$ to $\Re u = -1$, applying the functional equation of the completed zeta function, and using the evenness of $G_{z,w}$. Only the residue at $u = 0$ remains.

We call the first sum the direct term and the second sum the dual term.

For the subsequent moment calculation it is convenient to interchange m and n in the dual term. Since the weights depend on the summation variables only through their product mn, this gives

$$\begin{aligned}\zeta(s_1)\zeta(s_2) = \sum_{m,n\geq 1}\frac{(m/n)^{-it}}{\sqrt{mn}}\Big(&m^{-z/L}n^{-w/L}\,W^{+}_{z,w}(mn;t)\\ &+\chi(s_1)\chi(s_2)m^{w/L}n^{z/L}\,W^{-}_{z,w}(mn;t)\Big).\end{aligned}$$

In this form, both sums have the same oscillatory factor $(m/n)^{-it}$. Their difference lies in the shift-dependent coefficients and in the factor $\chi(s_1)\chi(s_2)$ appearing in the dual term.

We will insert this decomposition into $I(z,w)$, expand the four Dirichlet polynomials, and show that only the diagonal contributions survive in the leading term. The dependence on z, w will be kept throughout, allowing us to recover C_p and C_q by differentiation at the end.

Fix $\delta > 0$ such that $S_* + \delta < 1$. Write h_1, h_2 for the summation variables in $A_f(s)A_g(s)$ and k_1, k_2 for those in $A_h(\bar{s})A_k(\bar{s})$, and set

$$H = h_1h_2, \qquad K = k_1k_2, \qquad \mathcal{A}(h_1,h_2,k_1,k_2) = f\Big(\frac{\log h_1}{L}\Big)g\Big(\frac{\log h_2}{L}\Big)h\Big(\frac{\log k_1}{L}\Big)k\Big(\frac{\log k_2}{L}\Big).$$

The support assumptions imply $HK \leq T^{S_*}$. Inserting the preceding decomposition and expanding the four Dirichlet polynomials, we obtain terms with the common oscillatory factor

$$\Big(\frac{Hm}{Kn}\Big)^{-it}.$$

We first simplify the smoothing weights and show that the terms with $Hm \neq Kn$ are negligible.

Define

$$V_0(x) = \frac{1}{2\pi i}\int_{(1)} e^{u^2}x^{-u}\,\frac{\mathrm{d}u}{u} = \frac{1}{\sqrt{\pi}}\int_{\frac{1}{2}\log x}^{\infty} e^{-v^2}\,\mathrm{d}v.$$

For fixed bounded shifts z, w, Stirling's formula gives, uniformly for $T \le t \le 2T$ and u in a fixed vertical strip with $|\Im u| \le \log T$,

$$\frac{\Gamma_{\mathbb{R}}(s_1^{\pm}+u)\Gamma_{\mathbb{R}}(s_2^{\pm}+u)}{\Gamma_{\mathbb{R}}(s_1^{\pm})\Gamma_{\mathbb{R}}(s_2^{\pm})} = \left(\frac{t}{2\pi}\right)^{u}\left(1+O\left(\frac{(1+|u|)^{C}}{T}\right)\right),$$

where C is an absolute fixed constant. Moreover, the polynomial factor in $G_{z,w}(u)$ is $1+O((1+|u|)^{C}/T^{2})$ in this range. The Gaussian makes the complementary part of either vertical integral negligible. When the Mellin integrands defining $W_{z,w}^{\pm}$ and $V_0(2\pi x/t)$ are subtracted, their numerator vanishes at $u=0$. We may therefore move the contour of the difference to $\Re u = 0$ without crossing a pole. The preceding estimates, including the cancellation of the factor $1/u$ at the origin, yield

$$W_{z,w}^{\pm}(x;t) = V_0\left(\frac{2\pi x}{t}\right) + O(T^{-1}) \tag{2.62}$$

uniformly for $x>0$, $T\le t\le 2T$, and z, w in fixed complex discs. The same estimates give the corresponding bounds for any fixed number of t-derivatives, with an additional factor T^{-1} for each derivative.

For every fixed $B>1$, moving the defining Mellin contour to $\Re u = B$ gives

$$W_{z,w}^{\pm}(x;t), \quad V_0(2\pi x/t) \ll_B (t/x)^{B}.$$

Consequently, the portions of the two double sums with $mn > T^{1+\delta}$ are negligible. Indeed, the total absolute coefficient sum over the four amplifier variables is $O(T^{S_*/2})$, while

$$\sum_{mn>X}\frac{1}{\sqrt{mn}}\left(\frac{T}{mn}\right)^{B} \ll_B T^{B}X^{1/2-B}\log(2X).$$

The shift-dependent powers in the tails are at most $(mn)^{O(1/L)}$, and are absorbed by increasing B slightly. Taking $X = T^{1+\delta}$ and then B sufficiently large makes the integrated tail $o(TL^{9})$. On the truncated range, the total absolute coefficient sum satisfies

$$\sum_{\substack{h_1,h_2,k_1,k_2\\ mn\le T^{1+\delta}}}\frac{|\mathcal{A}(h_1,h_2,k_1,k_2)|}{\sqrt{mnHK}} \ll T^{(1+S_*+\delta)/2}(\log T)^{C}.$$

Thus the error $O(T^{-1})$ in (2.62), integrated over a time interval of length $O(T)$, contributes $O(T^{(1+S_*+\delta)/2}(\log T)^{C}) = o(TL^{9})$. All shift-dependent powers are bounded on the truncated range, and $\chi(s_1)\chi(s_2)$ is uniformly bounded.

It follows that, up to $o(TL^{9})$, the shifted integral is the sum over h_1, h_2, k_1, k_2 and $mn \le T^{1+\delta}$ of

$$\frac{\mathcal{A}(h_1,h_2,k_1,k_2)}{\sqrt{mnHK}}\int_{\mathbb{R}}\Phi(t/T)\left(\frac{Hm}{Kn}\right)^{-it}V_0\left(\frac{2\pi mn}{t}\right)e^{(z+w)/2} \cdot\left(m^{-z/L}n^{-w/L}+\chi(s_1)\chi(s_2)m^{w/L}n^{z/L}\right)\mathrm{d}t.$$

The amplitude after removal of the displayed oscillatory factor has j-th t-derivative $O_j(T^{-j})$, uniformly in all the truncated indices and in the fixed shift discs. If $Hm \ne Kn$ and the ratio $Hm/(Kn)$ is close to 1, integrality gives

$$\left|\log\frac{Hm}{Kn}\right| \gg \frac{1}{\max(Hm,Kn)} \gg (mnHK)^{-1/2}.$$

Away from 1 the logarithm is bounded below by a positive constant. Since $mnHK \le T^{1+\delta+S_*}$, in either case

$$T\left|\log\frac{Hm}{Kn}\right| \gg T^{(1-S_*-\delta)/2}.$$

Repeated integration by parts therefore bounds each off-diagonal time integral by

$$\ll_N T\left(1+T\left|\log\frac{Hm}{Kn}\right|\right)^{-N}.$$

Combining this with the preceding absolute coefficient bound, and choosing N sufficiently large, shows that all terms with $Hm \neq Kn$ contribute $o(TL^9)$.

We are left with the diagonal $Hm = Kn$. On this diagonal, Stirling's formula for the functional-equation factor gives

$$\chi(s_1)\chi(s_2) = \exp\Big(-\frac{z+w}{L}\log\frac{t}{2\pi}\Big)\,(1+O(T^{-1})) = e^{-(z+w)}(1+O(L^{-1}))$$

uniformly for $T \le t \le 2T$. The nine-edge diagonal limit of Lemma 2 shows that the diagonal sum, restricted to $mn \le T^{1+\delta}$ and normalized by L^9, is bounded. Hence replacing $\chi(s_1)\chi(s_2)$ by $e^{-(z+w)}$ changes the integral by $o(TL^9)$.

We next remove the remaining Gaussian smoothing. Put

$$Q = \frac{\log(mn)}{L}, \qquad t = T\tau, \qquad 1 \le \tau \le 2.$$

Then

$$V_0\Big(\frac{2\pi mn}{t}\Big) = \frac{1}{\sqrt{\pi}}\int_{\frac12\{L(Q-1)+\log(2\pi/\tau)\}}^{\infty} e^{-v^2}\,\mathrm{d}v.$$

For every $\varepsilon > 0$ this converges uniformly to 1 on $Q \le 1-\varepsilon$, and uniformly to 0 on $Q \ge 1+\varepsilon$, independently of $\tau \in [1,2]$. To justify passage to the limit in the diagonal sum, choose a continuous majorant of the strip $|Q-1| \le \varepsilon$, supported in $|Q-1| < 2\varepsilon$. Lemma 2 applies to this majorant: after normalization its limiting contribution tends to zero as $\varepsilon \downarrow 0$, because $Q = 1$ is a hyperplane of Lebesgue measure zero in the nine-edge coordinates. The restriction $Q \le 1+\delta$ makes the relevant region compact in combination with the four profile support bounds. Thus, uniformly for bounded z, w and $\tau \in [1,2]$, the smoothing may be replaced in the normalized diagonal sum by the sharp cutoff $Q \le 1$.

Apply Lemma 2 with the three indices on the s-side ordered as (m, h_1, h_2) and those on the $\bar{s}$-side as (n, k_1, k_2). Write $X = (x_{ij})_{1\le i,j\le 3}$ for the nine nonnegative edge variables, and put

$$a = \sum_j x_{2j}, \qquad b = \sum_j x_{3j}, \quad c = \sum_i x_{i2}, \quad d = \sum_i x_{i3},$$
$$R_0 = \sum_j x_{1j}, \quad C_0 = \sum_i x_{i1}.$$

Here R_0 and C_0 are the limiting logarithmic lengths of m and n, respectively. The diagonal cutoff becomes $R_0 + C_0 \le 1$. The direct term has limiting shifted weight $\exp(z(\frac12 - R_0) + w(\frac12 - C_0))$; the dual term has $\exp(z(C_0 - \frac12) + w(R_0 - \frac12))$. Since the weight is now independent of τ, integrating in time gives $T\widehat{\Phi}(0)$. Consequently,

$$\frac{I(z,w)}{a_3 TL^9} \longrightarrow \widehat{\Phi}(0)\int_{\substack{x_{ij}\ge 0\\ R_0+C_0\le 1}} f(a)g(b)h(c)k(d) \cdot \Big[e^{z(1/2-R_0)+w(1/2-C_0)} + e^{z(C_0-1/2)+w(R_0-1/2)}\Big]\prod_{i,j=1}^{3}\mathrm{d}x_{ij}.$$

It remains to identify the integral on the right with the proposed balanced kernel. Set

$$r_{11} = x_{22}, \quad r_{12} = x_{23}, \quad r_{21} = x_{32}, \quad r_{22} = x_{33}, \quad u = x_{11},$$

and write $t_r = \sum_{i,j=1}^{2} r_{ij}$. The nine-edge matrix is then

$$X = \begin{pmatrix} u & c - r_{11} - r_{21} & d - r_{12} - r_{22} \\ a - r_{11} - r_{12} & r_{11} & r_{12} \\ b - r_{21} - r_{22} & r_{21} & r_{22} \end{pmatrix}.$$

The nonnegativity of its entries is precisely $r \in \mathcal{R}(a, b; c, d)$ together with $u \geq 0$; the change from the nine entries of X to $(a, b, c, d, r_{11}, r_{12}, r_{21}, r_{22}, u)$ has Jacobian of absolute value one. Moreover,

$$R_0 = u + c + d - t_r, \qquad C_0 = u + a + b - t_r.$$

It follows that

$$R_0 + C_0 \leq 1 \quad \Longleftrightarrow \quad 0 \leq u \leq \frac{1 - a - b - c - d + 2t_r}{2} = w_r.$$

The support condition $a + b + c + d \leq S_* < 1$ ensures $w_r > 0$ on the relevant region.

For the direct term, make the substitution $x = u - w_r$. Then $-w_r \leq x \leq 0$, and with $\Delta = (a + b - c - d)/2$ we have

$$\frac{1}{2} - R_0 = \Delta - x, \qquad \frac{1}{2} - C_0 = -\Delta - x.$$

Thus its contribution is

$$\int_{\mathcal{R}(a,b;c,d)} \int_{-w_r}^{0} e^{z(\Delta - x) + w(-\Delta - x)} \, \mathrm{d}x \, \mathrm{d}r.$$

For the dual term, instead put $x = w_r - u$. Then $0 \leq x \leq w_r$, and

$$C_0 - \frac{1}{2} = \Delta - x, \qquad R_0 - \frac{1}{2} = -\Delta - x.$$

Its contribution is therefore

$$\int_{\mathcal{R}(a,b;c,d)} \int_{0}^{w_r} e^{z(\Delta - x) + w(-\Delta - x)} \, \mathrm{d}x \, \mathrm{d}r.$$

Adding the direct and dual contributions gives the uniform shifted limit

$$\frac{I(z, w)}{a_3 T L^9} \longrightarrow \widehat{\Phi}(0) \int_{\mathbb{R}^4_{\geq 0}} f(a) g(b) h(c) k(d) \int_{\mathcal{R}(a,b;c,d)} \int_{-w_r}^{w_r} e^{z(\Delta - x) + w(-\Delta - x)} \, \mathrm{d}x \, \mathrm{d}r \, \mathrm{d}a \, \mathrm{d}b \, \mathrm{d}c \, \mathrm{d}d.$$

All the estimates above are uniform on fixed complex discs in (z, w). Both sides are holomorphic in the shifts, so Cauchy's integral formula allows us to apply the fixed differential operators $p(\partial_z) q(\partial_w)$ and then set $z = w = 0$. Since

$$p(\partial_z) q(\partial_w) e^{z(\Delta - x) + w(-\Delta - x)} \Big|_{z=w=0} = p(\Delta - x) q(-\Delta - x),$$

we obtain exactly the kernel $B_{p,q}(a, b; c, d)$ in (2.61), and hence the asserted limit. This completes the proof. □

2.3.3 The unbalanced kernel

For pairs of tests from different families, the terms $H\overline{J}$ in (1.1) take the form

$$U_{p;f,g}\overline{V_{q;h,k}} = \chi(s)C_p(s)C_q(\bar{s})\,A_f(s)A_g(\bar{s})A_h(\bar{s})A_k(\bar{s}),$$
$$V_{p;f,g}\overline{U_{q;h,k}} = \chi(s)^{-1}C_p(s)C_q(\bar{s})\,A_f(s)A_g(s)A_h(\bar{s})A_k(s).$$

Unlike the balanced case, these products contain three amplifier factors on one side and only one on the other. Using the functional equation, we reduce them to unbalanced moments, which we evaluate next. For $a, b, c, d \geq 0$, put

$$H = a + b + c - d, \qquad \tau = \frac{1-H}{2}.$$

If $H \leq 0$, define $M_{p,q}(a, b, c; d) = 0$. For $H > 0$, let

$$\mathcal{S}(a,b,c;d) = \left\{ (r_1, r_2) \in \mathbb{R}^2_{\geq 0} : \begin{array}{r} r_1 \leq a, \\ r_2 \leq b, \\ 0 \leq r_3 := H - r_1 - r_2 \leq c \end{array} \right\}.$$

Then define

$$M_{p,q}(a,b,c;d) = \int_{\mathcal{S}(a,b,c;d)} \int_{\prod_{i=1}^3 [-r_i/2, r_i/2]} p(y_1+y_2+y_3-\tau) \cdot q(y_1+y_2+y_3+\tau)\, \mathrm{d}\mathbf{y}\, \mathrm{d}r_1\, \mathrm{d}r_2. \tag{2.63}$$

Proposition 8 (Unbalanced moment)**.** *The normalized integral*

$$\frac{1}{a_3 T L^9} \int_{\mathbb{R}} \Phi(t/T)\chi(s)^{-1}C_p(s)C_q(\bar{s})A_f(s)A_g(s)A_h(s)A_k(\bar{s})\, \mathrm{d}t$$

converges to

$$\widehat{\Phi}(0) \int f(a)g(b)h(c)k(d)M_{p,q}(a,b,c;d)\, \mathrm{d}a\, \mathrm{d}b\, \mathrm{d}c\, \mathrm{d}d.$$

Proof. We begin by introducing two complex shifts to handle the centered derivatives. For z, w in fixed complex discs, define

$$I(z,w) := \int_{\mathbb{R}} \Phi(t/T)\chi(s)^{-1}e^{(z+w)/2}\zeta\left(s+\frac{z}{L}\right)\zeta\left(\bar{s}+\frac{w}{L}\right) \cdot A_f(s)A_g(s)A_h(s)A_k(\bar{s})\, \mathrm{d}t.$$

By the definition of the centered derivatives, the integral in Proposition 8 is precisely

$$p(\partial_z)q(\partial_w)I(z,w)\big|_{z=w=0}.$$

As in the balanced case, we first establish a locally uniform asymptotic formula for $I(z, w)$ and apply the differential operators at the end.

The main difference from the balanced case is that the factor $\chi(s)^{-1}$ allows us to move the first zeta factor to the same side as the second. Indeed, the centered functional equation from Section 1 gives

$$\chi(s)^{-1}e^{z/2}\zeta\left(s+\frac{z}{L}\right) = R_T(t,z)e^{-z/2}\zeta\left(\bar{s}-\frac{z}{L}\right).$$

Consequently,

$$I(z,w) = \int_{\mathbb{R}} \Phi(t/T) R_T(t,z) e^{(w-z)/2} \zeta\left(\bar{s} - \frac{z}{L}\right) \zeta\left(\bar{s} + \frac{w}{L}\right) \cdot A_f(s) A_g(s) A_h(s) A_k(\bar{s})\, \mathrm{d}t.$$

We retain the factor $R_T(t,z)$ throughout the moment calculation. By Stirling's formula, uniformly for $T \le t \le 2T$ and z in a fixed disc,

$$R_T(t,z) = \left(\frac{t}{2\pi T}\right)^{-z/L} + O_R(T^{-1}).$$

The differentiated form of Stirling's formula also gives

$$\partial_t^j R_T(t,z) \ll_{j,R} T^{-j}$$

for every fixed integer $j \ge 0$. These estimates will allow us to integrate by parts in the off-diagonal terms. On the diagonal, Lemma 6 gives the simpler approximation

$$R_T(t,z) = 1 + O_R(L^{-1}).$$

Since both zeta factors now occur on the negative side, we can replace their product by a single, sufficiently long Dirichlet series. Put

$$w_1 = \bar{s} - \frac{z}{L}, \qquad w_2 = \bar{s} + \frac{w}{L},$$

and define

$$D_T(t;z,w) = \sum_{m,n \ge 1} \frac{e^{-mn/T^3}}{m^{w_1} n^{w_2}}.$$

We claim that, for every fixed $R > 0$ and sufficiently small $\varepsilon > 0$,

$$\zeta(w_1)\zeta(w_2) = D_T(t;z,w) + O_{\varepsilon,R}(T^{-1/2+\varepsilon}),$$

uniformly for $T \le t \le 2T$ and $|z|, |w| \le R$.

To establish this, Mellin inversion gives

$$D_T(t;z,w) = \frac{1}{2\pi i} \int_{(2)} \Gamma(u) T^{3u} \zeta(w_1+u)\zeta(w_2+u)\, \mathrm{d}u.$$

We truncate the integral at $|\Im u| \le T/4$ and shift the contour to

$$\Re u = -\frac{1}{2} - \varepsilon, \qquad 0 < \varepsilon < \frac{1}{4}.$$

The poles of the two zeta factors occur at $u = 1 - w_1$ and $u = 1 - w_2$. Their imaginary parts are $t + O_R(L^{-1})$, so both lie outside the rectangle used for the contour shift. Moreover, since $-\frac{1}{2} - \varepsilon > -1$, the only pole of $\Gamma(u)$ crossed during the shift is its simple pole at $u = 0$.

The residue at $u = 0$ is

$$\zeta(w_1)\zeta(w_2).$$

By Stirling's formula, the horizontal integrals and the discarded tails on the original line are exponentially small. It remains to estimate the integral on the new vertical line.

On this line, for $|\Im u| \le T/4$, the imaginary parts of $w_1 + u$ and $w_2 + u$ have absolute value comparable to T, while

$$\Re(w_j + u) = -\varepsilon + O_R(L^{-1}).$$

Applying the functional equation to each zeta factor and using the absolutely convergent Dirichlet series in the half-plane $\Re s > 1$, we obtain

$$|\zeta(w_1+u)\zeta(w_2+u)| \ll_{\varepsilon,R} T^{1+2\varepsilon}.$$

Since

$$|T^{3u}| = T^{-3/2-3\varepsilon}$$

and $\Gamma(u)$ is absolutely integrable along this vertical line, the remaining integral is $O_{\varepsilon,R}(T^{-1/2+\varepsilon})$. This proves the claimed replacement.

We next verify that its error is negligible in the moment integral. Since the four profiles are bounded and supported in $[0, A_1], \dots, [0, A_4]$, respectively,

$$\begin{aligned} |A_f(s)A_g(s)A_h(s)A_k(\bar{s})| &\ll \prod_{j=1}^{4} \sum_{n \le T^{A_j}} n^{-1/2} \\ &\ll T^{S_*/2}, \end{aligned}$$

where $S_* = A_1 + A_2 + A_3 + A_4 < 1$. The factor $R_T(t,z)e^{(w-z)/2}$ is uniformly bounded on fixed shift discs. Thus the integrated replacement error is

$$O_{\varepsilon,R}\left(T^{1/2+S_*/2+2\varepsilon}\right).$$

Choosing

$$0 < \varepsilon < \frac{1-S_*}{4}$$

makes the exponent strictly smaller than one. In particular, the error is $o(T)$, and hence $o(TL^9)$, uniformly on fixed shift discs.

We now insert the smoothed series into $I(z,w)$ and expand the four Dirichlet polynomials. Let $\ell_1, \ell_2, \ell_3, \ell_4$ denote their summation indices, and write

$$P = \ell_1\ell_2\ell_3, \qquad Q = \ell_4 mn.$$

For brevity, put

$$\begin{aligned} F_T(\ell_1,\ell_2,\ell_3,\ell_4) = f\left(\frac{\log \ell_1}{L}\right) g\left(\frac{\log \ell_2}{L}\right) \\ \cdot h\left(\frac{\log \ell_3}{L}\right) k\left(\frac{\log \ell_4}{L}\right). \end{aligned}$$

The resulting expression is

$$\begin{aligned} I(z,w) = e^{(w-z)/2} \sum_{\ell_1,\dots,\ell_4,m,n\ge 1} \frac{F_T(\ell_1,\dots,\ell_4)}{\sqrt{PQ}}\, m^{z/L} n^{-w/L} e^{-mn/T^3} \\ \cdot \int_{\mathbb{R}} \Phi(t/T) R_T(t,z) \exp\left(it \log \frac{Q}{P}\right) \mathrm{d}t + o(T). \end{aligned}$$

We separate the diagonal $P = Q$ from the off-diagonal $P \neq Q$. The diagonal equation is

$$\ell_1\ell_2\ell_3 = \ell_4 mn.$$

We first show that the off-diagonal contribution is negligible. Write

$$A_+ = A_1 + A_2 + A_3.$$

Since $\ell_j \le T^{A_j}$, we have $P \le T^{A_+}$. If $P \neq Q$, the integrality of P and Q implies

$$\left|\log \frac{Q}{P}\right| \gg T^{-A_+}.$$

Indeed, when P and Q are comparable, this follows from $|P-Q| \ge 1$; otherwise the logarithm is bounded away from zero.

Now Φ is smooth and compactly supported in $(1,2)$, and the derivative estimates for $R_T(t,z)$ imply

$$\frac{\mathrm{d}^j}{\mathrm{d}t^j}\left(\Phi(t/T)R_T(t,z)\right) \ll_{j,R,\Phi} T^{-j}.$$

Repeated integration by parts therefore gives, for every fixed integer $N \ge 1$,

$$\begin{aligned}\int_{\mathbb{R}} \Phi(t/T)R_T(t,z)\exp\left(it\log\frac{Q}{P}\right)\mathrm{d}t &\ll_{N,R,\Phi} T\left(T\left|\log\frac{Q}{P}\right|\right)^{-N}\\ &\ll_{N,R,\Phi} T^{1-N(1-A_+)}\end{aligned}$$

whenever $P \neq Q$.

It remains to sum this bound over the indices. The exponential smoothing ensures that the absolute sum of the coefficients grows at most polynomially in T. More precisely, for every fixed $\delta > 0$,

$$\sum_{m,n\ge 1} \frac{e^{-mn/T^3}|m^{z/L}n^{-w/L}|}{\sqrt{mn}} \ll_{R,\delta} T^{3/2+\delta},$$

uniformly on fixed shift discs. Combining this with the amplifier bound gives a total coefficient mass

$$O_{R,\delta}\left(T^{3/2+S_*/2+\delta}\right).$$

Hence the off-diagonal contribution is

$$O_{N,R,\delta}\left(T^{1+3/2+S_*/2+\delta-N(1-A_+)}\right).$$

Since $A_+ < 1$, we may choose N sufficiently large to make this $o(T)$. Thus only the diagonal contributes to the leading term.

On the diagonal $P = Q$, we have

$$mn = \frac{P}{\ell_4} \le P \le T^{A_+}.$$

Consequently,

$$e^{-mn/T^3} = 1 + O\left(T^{A_+-3}\right) = 1 + o(1)$$

uniformly over the diagonal summation indices.

Furthermore, Lemma 6 gives

$$\begin{aligned}\int_{\mathbb{R}} \Phi(t/T)R_T(t,z)\,\mathrm{d}t &= \int_{\mathbb{R}} \Phi(t/T)\,\mathrm{d}t + O_{R,\Phi}(T/L)\\ &= T\widehat{\Phi}(0) + O_{R,\Phi}(T/L).\end{aligned}$$

The absolute diagonal sum is $O(L^9)$, as follows from Lemma 2 and the compact support of the profiles. Since the shift-dependent factors $m^{z/L}$ and $n^{-w/L}$ are uniformly bounded on the diagonal, both replacements above produce an error $o(TL^9)$.

We therefore obtain

$$I(z,w) = T\widehat{\Phi}(0)e^{(w-z)/2} \sum_{\ell_1\ell_2\ell_3=\ell_4 mn} \frac{F_T(\ell_1,\dots,\ell_4)}{\ell_1\ell_2\ell_3}\, m^{z/L}n^{-w/L} + o(TL^9),$$

uniformly for z, w in fixed complex discs.

We now evaluate this diagonal sum using Lemma 2. Its three positive-side variables ℓ_1, ℓ_2, ℓ_3 correspond to the three row sums of the nine-edge matrix. The three negative-side variables m, n, ℓ_4 correspond to its column sums.

Fix row sums a, b, c and let d be the column sum corresponding to ℓ_4. Put

$$H = a + b + c - d.$$

We parameterize the nine nonnegative edge variables by the matrix

$$\begin{pmatrix} v_1 & r_1 - v_1 & a - r_1 \\ v_2 & r_2 - v_2 & b - r_2 \\ v_3 & r_3 - v_3 & c - r_3 \end{pmatrix},$$

where

$$r_1 + r_2 + r_3 = H, \qquad 0 \le r_1 \le a, \quad 0 \le r_2 \le b, \quad 0 \le r_3 \le c,$$

and

$$0 \le v_i \le r_i \qquad (i = 1, 2, 3).$$

The first two column sums are

$$v_1 + v_2 + v_3 \quad \text{and} \quad H - v_1 - v_2 - v_3,$$

while the third column sum is d, as required.

Eliminating

$$r_3 = H - r_1 - r_2$$

leaves precisely the region $\mathcal{S}(a, b, c; d)$ defined above. The change of variables from the nine edge coordinates to

$$(a, b, c, d, r_1, r_2, v_1, v_2, v_3)$$

has Jacobian of absolute value one.

The shift-dependent factors in the diagonal sum consequently become

$$m^{z/L}n^{-w/L} \longrightarrow \exp\left(z(v_1 + v_2 + v_3) - w(H - v_1 - v_2 - v_3)\right).$$

Including the centering factor $e^{(w-z)/2}$, the limiting shifted weight is

$$\exp\left(z\left(v_1 + v_2 + v_3 - \frac{1}{2}\right) + w\left(v_1 + v_2 + v_3 - H + \frac{1}{2}\right)\right).$$

For $H > 0$, Lemma 2 therefore yields

$$\lim_{T\to\infty} \frac{I(z,w)}{a_3 TL^9} = \widehat{\Phi}(0) \int_{\mathbb{R}^4_{\ge 0}} f(a)g(b)h(c)k(d) \cdot \int_{\mathcal{S}(a,b,c;d)} \int_{\prod_{i=1}^3 [0,r_i]} \exp\left(z\left(\sum_{i=1}^3 v_i - \frac{1}{2}\right)\right) \cdot \exp\left(w\left(\sum_{i=1}^3 v_i - H + \frac{1}{2}\right)\right) \mathrm{d}v\, \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}a\, \mathrm{d}b\, \mathrm{d}c\, \mathrm{d}d.$$

When $H < 0$, the corresponding region is empty; when $H = 0$, it has measure zero. Thus both cases give zero, in agreement with the definition of $M_{p,q}$.

We finally apply the differential operators $p(\partial_z)$ and $q(\partial_w)$.

All estimates above are uniform on fixed complex shift discs. Moreover, on the diagonal the logarithmic variables are confined to a fixed compact region, so the shifted diagonal weights are holomorphic and uniformly bounded on such discs. The convergence furnished by Lemma 2 is therefore locally uniform in z, w. Cauchy's integral formula justifies applying the differential operators to the limiting integral.

Writing

$$V = v_1 + v_2 + v_3,$$

we obtain the polynomial weight

$$p\left(V - \frac{1}{2}\right) q\left(V - H + \frac{1}{2}\right).$$

It follows that the limiting kernel before the final change of variables is

$$\int_{\mathcal{S}(a,b,c;d)} \int_{\prod_{i=1}^{3}[0,r_i]} p\left(\sum_{i=1}^{3} v_i - \frac{1}{2}\right) \cdot q\left(\sum_{i=1}^{3} v_i - H + \frac{1}{2}\right) \mathrm{d}v\, \mathrm{d}r_1\, \mathrm{d}r_2.$$

To obtain the symmetric form stated in the proposition, put

$$y_i = v_i - \frac{r_i}{2}, \qquad i = 1, 2, 3.$$

Then

$$-\frac{r_i}{2} \le y_i \le \frac{r_i}{2}, \qquad \sum_{i=1}^{3} v_i = \sum_{i=1}^{3} y_i + \frac{H}{2}.$$

Recalling that

$$\tau = \frac{1 - H}{2},$$

the two polynomial arguments become

$$\sum_{i=1}^{3} v_i - \frac{1}{2} = \sum_{i=1}^{3} y_i - \tau$$

and

$$\sum_{i=1}^{3} v_i - H + \frac{1}{2} = \sum_{i=1}^{3} y_i + \tau.$$

The change of variables has Jacobian one. Thus the limiting kernel is exactly

$$M_{p,q}(a,b,c;d) = \int_{\mathcal{S}(a,b,c;d)} \int_{\prod_{i=1}^{3}[-r_i/2, r_i/2]} p\left(y_1 + y_2 + y_3 - \tau\right) q\left(y_1 + y_2 + y_3 + \tau\right) \mathrm{d}\mathbf{y}\, \mathrm{d}r_1\, \mathrm{d}r_2,$$

which agrees with (2.63).

We conclude that

$$\lim_{T \to \infty} \frac{1}{a_3 T L^9} \int_{\mathbb{R}} \Phi(t/T) \chi(s)^{-1} C_p(s) C_q(\bar{s}) \cdot A_f(s) A_g(s) A_h(s) A_k(\bar{s})\, \mathrm{d}t = \widehat{\Phi}(0) \int_{\mathbb{R}^4_{\ge 0}} f(a) g(b) h(c) k(d) M_{p,q}(a,b,c;d)\, \mathrm{d}a\, \mathrm{d}b\, \mathrm{d}c\, \mathrm{d}d,$$

as claimed. □

2.3.4 Reflection errors and all projected Gram blocks

We first control the error arising from the centered functional equation. By the functional equation,

$$\chi(s)^{-1}C_q(s) - C_{q^-}(\bar{s}) = q(\partial_z)\left[(R_T(t,z)-1)e^{-z/2}\zeta(\bar{s}-z/L)\right]\Big|_{z=0}.$$

By Lemma 6 and Leibniz's rule, for every fixed polynomial q of degree at most four,

$$\left|\chi(s)^{-1}C_q(s) - C_{q^-}(\bar{s})\right| \ll_q L^{-1}\sum_{j=0}^{4}|C_{x^j}(\bar{s})|, \tag{2.64}$$

uniformly for $T \le t \le 2T$. We now show that the reflection error is negligible in the Gram matrix. Multiplying (2.64) by the two Dirichlet-polynomial factors of a test, say $A_f(s)A_g(\bar{s})$, and using

$$\left(\sum_{j=0}^{4} a_j\right)^2 \le 5\sum_{j=0}^{4} a_j^2,$$

we obtain

$$\begin{aligned}&\left|(\chi(s)^{-1}C_q(s) - C_{q^-}(\bar{s}))A_f(s)A_g(\bar{s})\right|^2\\ &\ll_q L^{-2}\sum_{j=0}^{4}|C_{x^j}(\bar{s})A_f(s)A_g(\bar{s})|^2.\end{aligned}$$

Choose a smooth nonnegative weight Ψ satisfying $\Psi \ge |\Phi|$. By Proposition 7, each term in the sum has weighted integral $O(TL^9)$. Consequently,

$$\begin{aligned}&\int_{\mathbb{R}} \Psi(t/T)\left|(\chi(s)^{-1}C_q(s) - C_{q^-}(\bar{s}))A_f(s)A_g(\bar{s})\right|^2 \mathrm{d}t\\ &\quad\ll L^{-2}\cdot TL^9 = O(TL^7).\end{aligned}$$

The other factor in the corresponding Gram product has squared weighted norm $O(TL^9)$, again by Proposition 7. Thus Cauchy–Schwarz gives

$$\begin{aligned}&\left|\int_{\mathbb{R}} \Phi(t/T)(\chi(s)^{-1}C_q(s) - C_{q^-}(\bar{s}))C_p(s)A_f(s)A_g(\bar{s})A_h(s)A_k(\bar{s})\,\mathrm{d}t\right|\\ &\quad\ll (TL^7)^{1/2}(TL^9)^{1/2}\\ &\quad= O(TL^8) = o(TL^9).\end{aligned}$$

The same argument applies to the other placements of the amplifier factors. Hence all reflection errors are negligible at the scale TL^9. Since the estimates use a nonnegative majorant of $|\Phi|$, no positivity assumption on Φ is needed.

Finally, we consider the reflected VV product. By the projection identity, this term initially takes the form

$$\chi(s)^{-3}C_p(s)C_q(s)A_f(s)A_g(s)A_h(s)A_k(s).$$

Applying the centered functional equation to $C_q(s)$, and using the reflection-error estimate established above, it suffices to show that

$$\int_{\mathbb{R}} \Phi(t/T)\chi(s)^{-2}C_p(s)C_{q^-}(\bar{s})A_f(s)A_g(s)A_h(s)A_k(s)\,\mathrm{d}t = o(TL^9).$$

We introduce bounded complex shifts for the two zeta factors and apply the one-sided expansion used in the proof of Proposition 8. After expanding the four Dirichlet polynomials, the resulting main terms contain the oscillatory factor

$$\chi(s)^{-1}\exp\left(it\log\frac{mn}{h_1h_2h_3h_4}\right),$$

where $m,n\geq 1$ and $h_i\leq T^{A_i}$.

The derivative of its phase is

$$\begin{aligned}\varphi'(t)&=\log\frac{t}{2\pi}+\log\frac{mn}{h_1h_2h_3h_4}+O(t^{-1})\\&\geq(1-S_*)L+O(1),\end{aligned}$$

since $h_1h_2h_3h_4\leq T^{S_*}$ and $mn\geq 1$. As $S_*<1$, we therefore have

$$\varphi'(t)\gg L$$

uniformly for $T\leq t\leq 2T$. In particular, the phase has no stationary point.

Writing $t=Tx$ and absorbing the shift-dependent powers of x into the smooth amplitude, repeated integration by parts gives, for every fixed integer $N\geq 1$,

$$\int_{\mathbb{R}}\Phi(t/T)\chi(s)^{-1}\exp\left(it\log\frac{mn}{h_1h_2h_3h_4}\right)(t/T)^{\tau}\,\mathrm{d}t\ll_{N,R}T^{1-N},$$

uniformly in the summation indices and for $|\tau|\leq R$, with R fixed. Here the uniform bounds for the derivatives of the reciprocal phase derivative ensure that the integrations by parts can be repeated.

The absolute sum of the coefficients in the one-sided expansion is bounded by $O(T^{6+S_*/2})$. Taking $N=9$, the total contribution of the oscillatory main terms is therefore

$$O(T^{6+S_*/2}T^{-8})=O(T^{-2+S_*/2}).$$

It remains to control the error in the one-sided replacement. The uniform replacement estimate used in the proof of Proposition 8 bounds its integrated contribution by

$$O\big(T^{1/2+2\varepsilon+S_*/2}\big).$$

Choosing $\varepsilon=(1-S_*)/8$ makes the exponent equal to $(3+S_*)/4<1$. Hence this contribution is also $o(TL^9)$.

All estimates are uniform on fixed shift discs, so the differential operators defining C_p and C_{q^-} may be applied afterwards. Consequently, the reflected VV contribution is $o(TL^9)$ and does not enter the limiting Gram matrix.

We can now assemble the Gram kernels for all pairs of projected tests. For two tests with polynomial and profile variables $(p;a,b)$ and $(q;c,d)$, respectively, define

$$\begin{aligned}G_{UU}(p,q;a,b,c,d)&=\frac{1}{2}\Big(B_{p,q}(a,d;b,c)+B_{p,q^-}(a,c;b,d)\Big),\\G_{UV}(p,q;a,b,c,d)&=\frac{1}{2}\Big(M_{q,p}(b,c,d;a)+M_{p,q^-}(a,c,d;b)\Big),\\G_{VV}(p,q;a,b,c,d)&=\frac{1}{2}B_{p,q}(a,b;c,d),\\G_{VU}(p,q;a,b,c,d)&=G_{UV}(q,p;c,d,a,b).\end{aligned}\tag{2.65}$$

To explain these formulas, consider first the UU block. The ordinary product is

$$U_{p;f,g}\overline{U_{q;h,k}} = C_p(s)C_q(\bar{s}) \cdot A_f(s)A_k(s)A_g(\bar{s})A_h(\bar{s}).$$

By Proposition 7, this contributes $B_{p,q}(a,d;b,c)$. For the reflected product, the centered functional equation replaces $\chi(s)^{-1}C_q(s)$ by $C_{q^-}(\bar{s})$, up to the negligible error established above. The resulting balanced moment contributes $B_{p,q^-}(a,c;b,d)$.

For the UV block, conjugating the ordinary product gives an unbalanced moment with profile ordering $(b,c,d;a)$ and polynomials (q,p). The reflected product, after applying the centered functional equation, gives the ordering $(a,c,d;b)$ with polynomials (p,q^-). These yield the two terms in G_{UV}.

For the VV block, the ordinary product is balanced, while the reflected contribution is $o(TL^9)$ by the preceding nonstationary phase argument. The VU formula follows by symmetry. In each case, the factor $1/2$ comes from the projection identity (1.1).

Finally, for admissible profiles f,g,h,k, the corresponding Gram coefficient is obtained by integrating the appropriate point kernel:

$$\int_{\mathbb{R}^4_{\geq 0}} f(a)g(b)h(c)k(d)\, G_{\tau\sigma}(p,q;a,b,c,d)\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c\,\mathrm{d}d,$$

where $\tau,\sigma \in \{U,V\}$ denote the two test families. This determines all entries of the denominator Gram matrix.

3 Assembly of the Lower Bound

3.1 Approximation by admissible profiles

The moment formulas established in Section 2 express the numerator and denominator of our test combinations as integrals of explicit kernels against admissible profiles. For the construction of a rational certificate, however, it is more convenient to evaluate these kernels directly at fixed points. We now show that such point evaluations can be approximated by genuine admissible profiles, so that the resulting finite-dimensional bounds remain applicable to the sixth moment.

Consider a finite collection of endpoint tests, indexed by $j = 1,\dots,m$. Each test is specified by a type U or V, a polynomial p_j, and two endpoints a_j, b_j. For a U-test, we require

$$0 < a_j \leq b_j, \qquad 3a_j + 5b_j < 2.$$

For a V-test, we fix a common cutoff $0 < \eta < 1/4$ and require

$$0 < a_j, b_j < \eta.$$

At this stage, these endpoint tests are only formal objects: they prescribe point evaluations of the kernels, rather than profiles satisfying the conditions of Section 1.2.

The corresponding numerator coefficients are

$$k_j = \begin{cases} K_{p_j}(a_j,b_j), & \text{for a } U\text{-test,} \\ K_{p_j}(a_j,1-b_j), & \text{for a } V\text{-test.} \end{cases}$$

Similarly, let G_{ij} denote the point evaluation of the appropriate projected Gram kernel obtained in Section 2.3 for the pair of tests i,j. Thus $k = (k_j)_{j=1}^m$ and $G = (G_{ij})_{i,j=1}^m$ are determined entirely by the chosen endpoints and polynomials.

For a fixed coefficient vector $c = (c_1, \dots, c_m) \in \mathbb{R}^m$, put

$$N = \sum_{j=1}^{m} c_j k_j = c^\top k, \qquad D = \sum_{i,j=1}^{m} c_i c_j G_{ij} = c^\top G c.$$

Whenever $D > 0$, the associated endpoint quotient is

$$R = \frac{9! N^2}{D}.$$

Our purpose is to show that this quotient is the limit of the corresponding quotients obtained from admissible test functions.

Proposition 9 (Approximation of endpoint tests)**.** *Let a finite collection of endpoint tests satisfy the conditions above, and let $c \in \mathbb{R}^m$ be fixed. There exists a sequence of collections of admissible U- and V-tests such that their numerator vectors $k^{(n)}$ and projected Gram matrices $G^{(n)}$ satisfy*

$$k_j^{(n)} \longrightarrow k_j, \qquad G_{ij}^{(n)} \longrightarrow G_{ij}$$

for every i, j.

Consequently, with

$$N_n = c^\top k^{(n)}, \qquad D_n = c^\top G^{(n)} c,$$

we have

$$N_n \longrightarrow N, \qquad D_n \longrightarrow D.$$

If $D > 0$, then $D_n > 0$ for all sufficiently large n, and

$$\frac{9! N_n^2}{D_n} \longrightarrow \frac{9! N^2}{D}. \tag{3.1}$$

Proof. We construct the approximating profiles separately for the two test families. The constructions differ because the U-family requires smooth compactly supported profiles, whereas the V-family requires polynomial profiles on a common interval.

Fix a nonnegative function $\psi \in C_c^\infty((0,1))$ satisfying

$$\int_{\mathbb{R}} \psi(x)\,\mathrm{d}x = 1.$$

Let $h_n > 0$ decrease to zero, and for an endpoint $a > 0$ define

$$f_{a,n}(x) = \frac{1}{h_n} \psi\left(\frac{a - x}{h_n}\right).$$

We choose $h_n < a$, which is possible simultaneously for all the finitely many U-endpoints once n is sufficiently large.

Since ψ is nonnegative and has integral one, the same is true of $f_{a,n}$. Moreover,

$$\operatorname{supp} f_{a,n} \subset (a - h_n, a) \subset (0, a).$$

In particular, $f_{a,n}$ is smooth and compactly supported strictly inside the required interval. The strict support conditions are therefore satisfied for every sufficiently large n, not merely in the limit.

For any bounded continuous function φ, a change of variables gives

$$\begin{aligned} \int_{\mathbb{R}} \varphi(x) f_{a,n}(x)\,\mathrm{d}x &= \int_0^1 \varphi(a - h_n u)\psi(u)\,\mathrm{d}u \\ &\longrightarrow \varphi(a). \end{aligned}$$

Thus, in the sense of weak convergence of probability measures,

$$f_{a,n}(x)\,\mathrm{d}x \Rightarrow \delta_a.$$

For a U-test with endpoints (a, b), we take $f_{a,n}$ and $f_{b,n}$ as its two profiles. The inequalities $0 < a \le b$ and $3a + 5b < 2$ ensure that these are admissible U-profiles. Their associated measures converge to δ_a and δ_b, respectively.

For the V-family, we must use a different construction, since the profiles are required to be polynomials on a common interval $[0, \eta]$, extended by zero outside that interval.

For the explicit certificate below, we shall use

$$\eta = \frac{19999}{80000} < \frac{1}{4}.$$

The construction itself works for any fixed $0 < \eta < 1/4$.

For each endpoint $0 < a < \eta$, set

$$\kappa_{a,n} = \left\lfloor \frac{na}{\eta} \right\rfloor$$

and define

$$\beta_{a,n}(x) = \frac{n+1}{\eta}\binom{n}{\kappa_{a,n}} \left(\frac{x}{\eta}\right)^{\kappa_{a,n}} \left(1 - \frac{x}{\eta}\right)^{n-\kappa_{a,n}}, \qquad 0 \le x \le \eta, \tag{3.2}$$

with $\beta_{a,n}(x) = 0$ outside $[0, \eta]$.

These are rescaled beta densities. In particular, they are nonnegative and, by the beta integral,

$$\int_0^\eta \beta_{a,n}(x)\,\mathrm{d}x = 1.$$

Their means and variances are

$$\mu_{a,n} = \frac{\eta(\kappa_{a,n}+1)}{n+2},$$

and

$$\begin{aligned}\sigma_{a,n}^2 &= \frac{\eta^2(\kappa_{a,n}+1)(n-\kappa_{a,n}+1)}{(n+2)^2(n+3)} \\ &\le \frac{\eta^2}{4(n+3)}.\end{aligned}$$

Since

$$\frac{\kappa_{a,n}}{n} \longrightarrow \frac{a}{\eta},$$

we obtain

$$\mu_{a,n} \longrightarrow a, \qquad \sigma_{a,n}^2 \longrightarrow 0.$$

It follows, for instance by Chebyshev's inequality, that

$$\beta_{a,n}(x)\,\mathrm{d}x \Rightarrow \delta_a.$$

It remains to verify the endpoint conditions. Because $0 < a < \eta$, we have

$$1 \le \kappa_{a,n} \le n - 1$$

for all sufficiently large n. Consequently, $\beta_{a,n}$ vanishes at both 0 and η, so that its extension by zero is continuous. On $[0, \eta]$ it is a polynomial, as required.

Unlike the U-profiles, the V-profiles are not confined to shrinking support intervals. Their support is contained in the fixed interval $[0, \eta]$, independently of the endpoint a. The condition $\eta < 1/4$ therefore ensures the required length restriction for every approximating V-test.

For a V-test with endpoints (a, b), we use $\beta_{a,n}$ and $\beta_{b,n}$. Since there are only finitely many endpoint tests, one common sufficiently large n makes every such profile admissible.

We emphasize that the polynomial degrees may increase with n. This causes no difficulty: for each fixed n, all profiles and polynomial degrees remain fixed when the moment formulas are applied as $T \to \infty$. Only after taking those moment limits do we let $n \to \infty$.

To pass from weak convergence of the profile measures to convergence of the moment coefficients, we need continuity of the corresponding kernels in their endpoint variables.

For the numerator kernel K_p, continuity follows from its compact integral representation in Section 2.2. On any compact subset of the admissible parameter domain, its integrand is continuous and uniformly bounded, so dominated convergence applies.

For the balanced kernel $B_{p,q}$, consider its integral representation from Section 2.3. Locally in the endpoint variables, all integration regions are contained in a single compact box. We may therefore write the integral over this fixed box, inserting the indicator function of the defining inequalities.

As the endpoints vary, these indicator functions converge pointwise except where one of the defining affine inequalities becomes an equality. Such exceptional sets are contained in finitely many affine hyperplanes and have Lebesgue measure zero. The remaining factors in the integrand depend continuously on the endpoints and are uniformly bounded on the fixed box. Dominated convergence therefore proves continuity of $B_{p,q}$.

The same reasoning applies to the unbalanced kernel $M_{p,q}$. Its integration region is again defined by affine inequalities depending continuously on the endpoints.

There is one additional boundary case to consider. Recall that $M_{p,q}$ is defined to be zero when

$$H = a + b + c - d \le 0.$$

For $H > 0$, its integration region in the first two variables is contained in

$$\{(r_1, r_2) \in \mathbb{R}^2_{\ge 0} : r_1 + r_2 \le H\}.$$

As H decreases to zero, this region shrinks to a set of measure zero. Since the integrand is locally bounded, the integral tends to zero. Thus $M_{p,q}$ is continuous also across $H = 0$.

Finally, all projected Gram kernels are finite linear combinations of these balanced and unbalanced kernels, with the polynomial reflections and permutations established in Section 2.3. They are therefore continuous in their endpoint variables as well.

For each fixed n, let $k^{(n)}$ and $G^{(n)}$ be the normalized limiting numerator vector and projected Gram matrix obtained by applying the moment formulas of Section 2 to the admissible profiles constructed above.

For example, the numerator coefficient of a U-test is

$$k_j^{(n)} = \int_{\mathbb{R}^2} f_{a_j,n}(x) f_{b_j,n}(y) K_{p_j}(x, y)\, \mathrm{d}x\, \mathrm{d}y,$$

whereas for a V-test it is

$$k_j^{(n)} = \int_{\mathbb{R}^2} \beta_{a_j,n}(x) \beta_{b_j,n}(y) K_{p_j}(x, 1 - y)\, \mathrm{d}x\, \mathrm{d}y.$$

The Gram entries $G_{ij}^{(n)}$ are obtained analogously by integrating the corresponding projected point kernels against the four profile measures belonging to tests i and j.

All these measures are probability measures supported in fixed compact intervals. Their weak convergence to the respective point masses implies weak convergence of their finite product measures. By the continuity of the kernels established above, we consequently obtain

$$k_j^{(n)} \longrightarrow k_j, \qquad G_{ij}^{(n)} \longrightarrow G_{ij}$$

for every pair of indices i, j.

Since the collection of tests is finite, these convergences hold simultaneously for all entries. For the fixed coefficient vector c, it follows that

$$N_n = \sum_{j=1}^{m} c_j k_j^{(n)} \longrightarrow \sum_{j=1}^{m} c_j k_j = N,$$
$$D_n = \sum_{i,j=1}^{m} c_i c_j G_{ij}^{(n)} \longrightarrow \sum_{i,j=1}^{m} c_i c_j G_{ij} = D.$$

If $D > 0$, the convergence $D_n \to D$ implies $D_n > 0$ for all sufficiently large n. Hence

$$\frac{9! N_n^2}{D_n} \longrightarrow \frac{9! N^2}{D},$$

as claimed.

Notice that we only require positivity of the quadratic form in the chosen direction c. No invertibility or positive definiteness of the entire Gram matrix is needed. □

3.2 From endpoint kernels to the sixth moment

We now use the moment formulas from Section 2 and the Cauchy–Schwarz inequality from Section 1.3 to convert this endpoint quotient into a lower bound for the sixth moment.

The moment formulas initially provide estimates with a smooth weight supported on an interval of the form $(Y, 2Y)$. To obtain a bound for the full moment $M_3(T)$, we shall combine finitely many such estimates over disjoint dyadic intervals.

Proposition 10 (Endpoint-to-moment transfer)**.** *Let a finite collection of endpoint tests and a real coefficient vector c satisfy the conditions of Proposition 9. Let k and G be the corresponding endpoint numerator vector and projected Gram matrix, and put*

$$N = c^\top k, \qquad D = c^\top G c.$$

If $D > 0$, then

$$\liminf_{T\to\infty} \frac{M_3(T)}{(a_3/9!)T(\log T)^9} \geq \frac{9! N^2}{D}. \tag{3.3}$$

Proof. Write

$$R = \frac{9! N^2}{D}.$$

If $N = 0$, then $R = 0$ and the assertion follows immediately from the nonnegativity of $M_3(T)$. We may therefore assume $R > 0$.

By Proposition 9, there are admissible profile collections with associated numerator and denominator coefficients N_n, D_n satisfying

$$N_n \longrightarrow N, \qquad D_n \longrightarrow D > 0.$$

In particular, $D_n > 0$ for all sufficiently large n, and the corresponding quotients

$$R_n = \frac{9! N_n^2}{D_n}$$

satisfy

$$R_n \longrightarrow R. \tag{3.4}$$

We first establish a weighted lower bound for each sufficiently large fixed n.

Let $\Phi \in C_c^\infty((1,2))$ be nonnegative, with $\widehat{\Phi}(0) = \int_{\mathbb{R}} \Phi(x)\,\mathrm{d}x > 0$. At scale Y, let $F_{c,n,Y}$ denote the projected linear combination from Section 1.3, formed using the admissible profiles of index n.

For the remainder of this step, n, the test polynomials, their coefficients, and Φ are all fixed while $Y \to \infty$.

The numerator moment formulas of Section 2 give

$$\frac{1}{a_3 Y(\log Y)^9} \int_{\mathbb{R}} \Phi(t/Y)\zeta(\tfrac{1}{2}+it)^3 \overline{F_{c,n,Y}(t)}\,\mathrm{d}t \longrightarrow \widehat{\Phi}(0) N_n. \tag{3.5}$$

Likewise, the denominator moment formulas, including all pairwise interactions between the projected tests, give

$$\frac{1}{a_3 Y(\log Y)^9} \int_{\mathbb{R}} \Phi(t/Y)|F_{c,n,Y}(t)|^2\,\mathrm{d}t \longrightarrow \widehat{\Phi}(0) D_n. \tag{3.6}$$

Since $\widehat{\Phi}(0) D_n > 0$, the denominator integral is positive for all sufficiently large Y. We may therefore apply the weighted Cauchy–Schwarz inequality:

$$\begin{aligned}
&\int_{\mathbb{R}} \Phi(t/Y)|\zeta(\tfrac{1}{2}+it)|^6\,\mathrm{d}t \\
&\qquad \geq \frac{\left|\int_{\mathbb{R}} \Phi(t/Y)\zeta(\tfrac{1}{2}+it)^3 \overline{F_{c,n,Y}(t)}\,\mathrm{d}t\right|^2}{\int_{\mathbb{R}} \Phi(t/Y)|F_{c,n,Y}(t)|^2\,\mathrm{d}t}.
\end{aligned}$$

Inserting the limits (3.5) and (3.6), we obtain

$$\begin{aligned}
\liminf_{Y\to\infty} \frac{\int_{\mathbb{R}} \Phi(t/Y)|\zeta(\tfrac{1}{2}+it)|^6\,\mathrm{d}t}{(a_3/9!)Y(\log Y)^9}
&\geq 9! \frac{\widehat{\Phi}(0)^2 N_n^2}{\widehat{\Phi}(0) D_n} \\
&= \widehat{\Phi}(0) R_n.
\end{aligned}$$

Thus every sufficiently large fixed profile index n gives a weighted sixth-moment lower bound with constant $\widehat{\Phi}(0)R_n$.

We next remove the smooth weight by combining estimates at finitely many dyadic scales.

Fix an integer $J \geq 1$, and for $0 \leq j < J$ put

$$Y_j = \frac{T}{2^{j+1}}.$$

Since Φ is supported in $(1,2)$, the function $\Phi(t/Y_j)$ is supported in

$$(Y_j, 2Y_j) = \left(\frac{T}{2^{j+1}}, \frac{T}{2^j}\right).$$

These intervals are pairwise disjoint and contained in $(0,T)$.

We now additionally require $0 \le \Phi \le 1$. It follows that

$$0 \le \sum_{j=0}^{J-1} \Phi(t/Y_j) \le \mathbf{1}_{[0,T]}(t).$$

Since $|\zeta(\frac{1}{2}+it)|^6$ is nonnegative, we may sum the weighted integrals to obtain

$$M_3(T) \ge \sum_{j=0}^{J-1} \int_{\mathbb{R}} \Phi(t/Y_j)|\zeta(\tfrac{1}{2}+it)|^6 \,\mathrm{d}t.$$

For each fixed j, we have $Y_j \to \infty$ as $T \to \infty$, and

$$\frac{Y_j}{T} = 2^{-j-1}, \qquad \frac{\log Y_j}{\log T} \longrightarrow 1.$$

Therefore the weighted lower bound, applied at scale Y_j, yields

$$\begin{aligned}\liminf_{T\to\infty} \frac{\int_{\mathbb{R}} \Phi(t/Y_j)|\zeta(\frac{1}{2}+it)|^6 \,\mathrm{d}t}{(a_3/9!)T(\log T)^9} \\ \ge 2^{-j-1}\widehat{\Phi}(0)R_n.\end{aligned}$$

Since J is fixed, we may combine these finitely many lower bounds. Consequently,

$$\begin{aligned}\liminf_{T\to\infty} \frac{M_3(T)}{(a_3/9!)T(\log T)^9} &\ge \widehat{\Phi}(0)R_n \sum_{j=0}^{J-1} 2^{-j-1} \\ &= \widehat{\Phi}(0)R_n(1-2^{-J}).\end{aligned}$$

It remains to show that the factors arising from profile approximation, smoothing, and the finite dyadic decomposition can be made arbitrarily close to one.

Let $0 < \rho < R$. By (3.4), we can first choose a sufficiently large, fixed profile index n such that

$$D_n > 0, \qquad R_n > \rho.$$

Next, choose a smooth weight satisfying

$$0 \le \Phi \le 1, \qquad \Phi \in C_c^\infty((1,2)),$$

whose integral $\widehat{\Phi}(0)$ is sufficiently close to one. Such weights exist, for example, by choosing Φ equal to one on $[1+\varepsilon, 2-\varepsilon]$ and letting ε become small.

Finally, choose the fixed integer J sufficiently large that

$$\widehat{\Phi}(0)R_n(1-2^{-J}) > \rho.$$

The preceding lower bound then gives

$$\liminf_{T\to\infty} \frac{M_3(T)}{(a_3/9!)T(\log T)^9} > \rho.$$

Since $\rho < R$ was arbitrary, letting $\rho \uparrow R$ proves

$$\liminf_{T\to\infty} \frac{M_3(T)}{(a_3/9!)T(\log T)^9} \ge R = \frac{9!N^2}{D}.$$

□

3.3 The explicit rational certificate

By Proposition 10, it remains to find a finite collection of endpoint tests and a rational coefficient vector c such that, with

$$N = c^\top k, \qquad D = c^\top G c,$$

we have

$$D > 0, \qquad \frac{9! N^2}{D} > 41.008.$$

On a subspace where the Gram matrix is positive definite, the quotient is maximized in the direction $G^{-1}k$. This motivates our choice of coefficients, but the proof uses only the explicit rational certificate below. Neither the number of tests nor the degree-four cutoff is claimed to be optimal.

3.3.1 Certificate data

The certificate consists of 250 rows

$$(\text{type}, A, B, r, c)$$

in the file `certificate_250.tsv` [5]. Each row specifies a family U or V, rational endpoints

$$a = \frac{A}{10^8}, \qquad b = \frac{B}{10^8},$$

a rational coefficient c, and a test polynomial

$$p_r(x) = T_r(2x), \qquad 0 \le r \le 4,$$

where T_r is the Chebyshev polynomial of the first kind.

There are 220 U-tests and 30 V-tests. For every U-test, the endpoints satisfy $0 < a \le b$, with

$$\min_{U\text{-tests}} (2 - 3a - 5b) = \frac{19997}{10^8} > 0.$$

For the V-tests, all endpoints are positive and

$$\max_{V\text{-tests}} \{a, b\} = 0.249975 < \eta = \frac{19999}{80000} < \frac{1}{4}.$$

Thus every test satisfies the admissibility conditions of Section 3.1.

3.3.2 Exact rational evaluation

We now explain why the endpoint kernels can be evaluated using exact rational arithmetic.

The integral representations from Section 2 involve polynomial integrands over regions defined by linear inequalities with rational coefficients. At the rational endpoints of our certificate, their evaluation reduces to finitely many exact polynomial integrations.

For the numerator kernel K_p, the simplex representation from Section 2.2 is evaluated using standard simplex moments. Since $\deg p \le 4$, only moments of order at most four are needed, followed by polynomial integration over a rational interval.

For the balanced kernel $B_{p,q}$, we slice the integration region according to the total edge mass

$$t = \sum_{i,j=1}^{2} r_{ij}.$$

The resulting slice density is piecewise polynomial. Splitting the interval at its finitely many rational breakpoints therefore reduces the evaluation to ordinary polynomial integrals.

For the unbalanced kernel $M_{p,q}$, we first integrate over the inner variables y_i. The remaining polynomial integral over a capped simplex is evaluated by inclusion–exclusion and standard simplex moment formulas.

Together with the projected Gram formulas from Section 2.3, these procedures determine every entry of k and G exactly. The exact evaluation routines are available in the repository [5].

3.3.3 Certified inequalities and conclusion

Let k and G be the endpoint values obtained from the certificate, and let c be its rational coefficient vector. Exact rational evaluation gives

$$N = c^\top k > 46159, \qquad 0 < D = c^\top G c < 18854000000000.$$

Consequently,

$$\begin{aligned} R := \frac{9!N^2}{D} &> \frac{9! \cdot 46159^2}{18854000000000} \\ &= \frac{1208080410327}{29459375000} \\ &> 41.008. \end{aligned}$$

This proves Theorem 1.

References

[1] S. Bettin, H. M. Bui, X. Li and M. Radziwiłł, *A quadratic divisor problem and moments of the Riemann zeta-function*, J. Eur. Math. Soc. **22** (2020), 3953–3980. doi:10.4171/JEMS/999. The equation, proposition and page numbers used here refer to arXiv:1609.02539v1.

[2] J. B. Conrey and A. Ghosh, *A conjecture for the sixth power moment of the Riemann zeta-function*, Int. Math. Res. Not. **1998**, no. 15, 775–780. doi:10.1155/S1073792898000476. arXiv:math/9807187.

[3] C. de Boor, *Math/CS 887: Approximation Theory*, class notes, Spring 2003, version of 21 November 2009, pp. 52–53, formula (83). `https://pages.cs.wisc.edu/~deboor/887/notes.pdf`.

[4] B. Durkan and T. Page, *Amplified moments of the Riemann zeta function*, arXiv:2606.27323v2 (2026).

[5] N. Fraccaro, *Sixth moment: Lean formalization and exact certificate*, repository snapshot `950bfaff8724d4adb1c37535634a139b70228df0`. `https://github.com/Niklas-Fraccaro/sixth-moment-lean/tree/950bfaff8724d4adb1c37535634a139b70228df0`. The endpoint table is `certificate_250.tsv`; the analytic limits and exact evaluations are in `SixthMoment/Numerator`, `SixthMoment/Denominator` and `SixthMoment/Certificate`.

[6] J. P. Keating and N. C. Snaith, *Random matrix theory and* $\zeta(1/2 + it)$, Comm. Math. Phys. **214** (2000), 57–89. doi:10.1007/s002200000261.