hexagon:2611.01021v1

# The Selberg integral at polynomial lengths and pair correlation beyond one

Wren Aquila

25 November 2026

**Abstract**

For every sufficiently large $T$, Montgomery's normalized form factor is at least 0.500001 throughout $1 \le \alpha \le 1.0000009$, assuming the Riemann hypothesis for $\zeta$. On a second, wider interval a smaller positive increment also improves the line $3/2-\alpha$. The arithmetic result is unconditional. For $1 \le h \le X^{0.001}$, the Selberg integral on $[X, 2X]$ exceeds $Xh\{\frac{1}{2}\log(X/h^3) + a\log X\}$ with an absolute $a > 0$ for all sufficiently large $X$. We obtain the increment from a second family of Ramanujan sums, supported beyond the square root. The argument also gives the form-factor increment unconditionally for complex differences of zeta zeros.

## 1 Statements and prior work

Write $e(t) = \exp(2\pi it)$, $L = \log X$, $\psi(t) = \sum_{n\le t}\Lambda(n)$ and $\vartheta(t) = \sum_{p\le t}\log p$. Define

$$J(X,h) = \int_X^{2X} |\psi(t+h) - \psi(t) - h|^2 dt.$$

Let $\rho = \beta + i\gamma$ run over the nontrivial zeros of $\zeta$, with multiplicity. With both sums over $0 < \gamma, \gamma' \le T$, put

$$F_T(\alpha) = \frac{2\pi}{T\log T}\sum_{\rho,\rho'} T^{i\alpha(\gamma-\gamma')}\frac{4}{4+(\gamma-\gamma')^2}, \quad F_{T,\mathrm{c}}(\alpha) = \frac{2\pi}{T\log T}\sum_{\rho,\rho'} T^{\alpha(\rho-\rho')}\frac{4}{4-(\rho-\rho')^2}.$$

Use the following two pairs of constants:

$$(\theta_1, a_1) = \left(.000039, \frac{74}{100000}\frac{2605072188928623487}{10^{21}}\right), \qquad (\theta_2, a_2) = (.001, 1.8\cdot 10^{-15}).$$

**Theorem 1.** *For $i = 1, 2$, uniformly for $1 \le h \le X^{\theta_i}$,*

$$J(X,h) \ge Xh\{\tfrac{1}{2}\log(X/h^3) + a_i\log X - o(\log X)\}.$$

*For $1 \le h \le X^{1/3-\epsilon}$ and any fixed $\epsilon > 0$, the same bound holds with $a_i = 0$. All these estimates are unconditional and hold with $\vartheta$ in place of $\psi$.*

**Theorem 2.** *Assume the Riemann hypothesis for $\zeta$. For $i = 1, 2$, uniformly for $1 \le \alpha \le 1/(1-\theta_i)$,*

$$F_T(\alpha) \ge 3/2 - \alpha + a_i\alpha - o(1).$$

*For $1 \le \alpha \le 3/2-\epsilon$ the same inequality holds with $a_i = 0$. Without a zero hypothesis the corresponding bounds hold for $F_{T,\mathrm{c}}$.*

In particular the two normalized form factors have lower bound 0.500001 on $[1, 1.0000009]$ under their respective hypotheses, for sufficiently large $T$. For $1 \le h \le X^{.0000006}$ the unconditional lower bound for $J(X,h)$ is $0.500001Xh\log X$ for sufficiently large $X$. The strict inequality $a_1 > 1.92775 \cdot 10^{-6}$ gives slack at both right endpoints. The statements at a fixed point with $\alpha > 1$ use an arithmetic estimate beyond the known endpoint asymptotic $F_T(1) = 1 + o(1)$.

Goldston–Gonek–Özlük–Snyder (GGÖS) obtained $3/2 - \alpha - \epsilon$ on $[1, 3/2 - \epsilon]$ under GRH for Dirichlet $L$-functions. Walker's Theorem 1.10 under the progression conjecture AP(27/53) gives the larger $127/53 - 100\alpha/53$ on $1 \le \alpha < 95/94$, which exceeds every bound of Theorem 2 on that interval and meets the line $3/2 - \alpha$ at $\alpha = 95/94$. The reductions in GGÖS, Lemma 1, and Walker, Lemma 4.1, use only RH for $\zeta$. Walker asks for an improvement of the slope bound under GRH for Dirichlet $L$-functions. Baluyot, Goldston, Suriajaya and Turnage-Butterbaugh (BGSTB) proved nonnegativity of $F_{T,\mathrm{c}}$ and evaluated it on $0 \le \alpha \le 1$. [5, 10, 1] Alpöge has announced a proof of the Riemann hypothesis for $\zeta$ and of GRH for Dirichlet $L$-functions [11]. If confirmed, the GGÖS line and the Goldston–Yıldırım variance bound below would follow from those authors' theorems, and the two form factors would coincide. The positive increments in Theorems 1 and 2 improve those respective bounds.

Goldston obtained the unconditional half coefficient at polylogarithmic lengths. Juhas, Theorem 2, extended this to $h \le \exp(c(\log X)^{3/5}(\log\log X)^{-1/5})$. His Theorem 11 gives, at polynomial lengths and for infinitely many $X$, a lower bound of order $Xh\log X/\varphi(r_0)$ for the corresponding variance of primes in an arithmetic progression to some modulus $r_0 \le h$. Goldston–Yıldırım give the coefficient $\frac{1}{2}\log(X/h^3)$ to $X^{1/3-\epsilon}$ under GRH for Dirichlet $L$-functions. Our restriction to large Ramanujan denominators adapts the minor-arc method used by Liu, Perelli, Hooley and Harper–Soundararajan for the variance of primes in arithmetic progressions. Juhas uses signed shift sums and Vaughan's mean value theorem at smaller lengths. Here completion of the shift sum lets the high-conductor test reach polynomial lengths. Polymath supplies the Type I and II dispersion estimates, Walker the off-diagonal divisor estimate, and Hooley's factorization the mixed diagonal. [3, 4, 8, 6, 7, 9, 10]

## 2 The ordinary Selberg integral

Fix $H = h$, $1 \le H \le X^{1/3-100\eta}$, with $\eta > 0$ as small as needed. Use the full band $U = HX^{2\eta} < r \le R = \sqrt{X/H}X^{-10\eta}$,

$$B(n) = \sum_{U<r\le R} \frac{\mu(r)c_r(n)}{\varphi(r)} = \sum_{d\mid n} \lambda_d, \qquad A = \sum_{U<r\le R} \frac{\mu^2(r)}{\varphi(r)} = \log(R/U) + O(1).$$

Here $c_r$ is the Ramanujan sum. The divisor coefficients are supported on squarefree $d \le R$ and satisfy $|\lambda_d| \ll L^3$, by the identity $c_r(n) = \sum_{d\mid(r,n)} d\mu(r/d)$. The last harmonic estimate follows from the Dirichlet series $\zeta(1+s)\prod_p(1-p^{-1-s})(1+p^{-s}/(p-1))$, whose second factor has value one at zero and is absolutely convergent for $\Re s > -1/2$. Let $0 \le \omega, W \le 1$ be smooth weights supported on $[1,2]$ and $[0,1]$, equal to one outside endpoint collars of width $O(1/L)$ *inside* those intervals. Their fixed derivatives cost log powers. Write $D(t) = \sum_n B(n)W((n-t)/h)$ and use the measure $\omega(t/X)dt$. The covariance with the prime count, divided by $h$, is

$$\int_0^1 \sum_p (\log p)\omega((p-hv)/X) \sum_j B(p+j)W(v+j/H)\,dv. \tag{2.1}$$

Its zero shift contributes $XA\int\omega\int W + O_D(XL^{-D})$. For the principal character on $r$ the coefficient at shift $j$ is $\mu^2(r)c_r(j)/\varphi(r)^2$. At nonzero shifts sum $\sum_{0<|j|\le H+1}|c_r(j)| \ll H\tau(r)$, and then use $r > U$. This is power-small compared with $X$.

For the nonprincipal part expand each divisor indicator on units modulo $d = fz$ in primitive characters $\chi$ of conductor $f > 1$. If $(j, d) > 1$ the original congruence has no macroscopic prime. Remove the condition $(j, z) = 1$ by inclusion over $v_1 \mid z$. For $H_1 = H/v_1$ the shift sum is $U_\chi = \sum_m \overline{\chi}(m) W(v + m/H_1)$ up to a unit scalar. Values $v_1 > 2H$ contribute zero. The prime sum $S_\chi$ has a smooth weight at length $O(X)$. It suffices on each dyad $f \asymp F$ to bound $F^{-1} \sum^*_{f,\chi} |S_\chi U_\chi|$ to arbitrary logarithmic saving from $X$. The outer cost is $\sum_{z \le R} \tau(z) L^{O(1)}/\varphi(z) \ll L^{O(1)}$. For $F \le L^E$ use Siegel–Walfisz and Pólya–Vinogradov. For $L^E < F \le X^{1/3-\eta}$, Vaughan's mean value theorem for $S_\chi$ and Pólya–Vinogradov on the shifts give

$$L^{O(1)}(X/F + X^{5/6} + \sqrt{X}F)\sqrt{F}.$$

Choose $E$ sufficiently large for the desired logarithmic saving.

For $X^{1/3-\eta} < F \le R$ use Vaughan's identity with both parameters $Z = HX^{3\eta}$. If a free variable has marked partner of length at most $Z$, Pólya–Vinogradov in the free variable and the large sieve in the shift cost $ZF^{3/2}\sqrt{H}X^{o(1)}$. This saves a power of $X$ since $R \le \sqrt{X/H}X^{-10\eta}$ and $H \le X^{1/3-100\eta}$. In the other blocks, with $M \ge N$, $Y = MN \ll X$, both factors have length at most $X^{1-2\eta}/H$. Indeed the free pieces have marked length at most $Z^2$, and in the last convolution both variables exceed $Z$, up to dyadic constants. Complete the primitive shift by Poisson summation. Its prefactor is $H_1/\sqrt{F}$ up to a unit scalar and a bounded dyadic factor, and its dual length is $F/H_1$ up to an arbitrarily small power. For each sign of the dual variable $\ell$, the Fourier weight depends on $f$ only through $u = H_1|\ell|/f$. Write it as a Mellin integral on $\Re s = \varepsilon' > 0$. Its transform has $L^1$ norm bounded by a log power, uniformly in $v$, by integration by parts applied to $\widehat{W}(\pm u)e(\pm vu)$. The modulus contributes only the scalar $f^s$, and the factor $|\ell|^{-s}$ is independent of $f$. Attach the dual sum to the shorter factor $N$. Cauchy and two multiplicative large sieves, with divisor bounds for the convolution, give the squared bound

$$X^{o(1)}(YH_1F^2 + YH_1M + YFN + Y^2/F), \tag{2.2}$$

where arbitrarily small fixed powers from truncation and the Mellin line can be absorbed in the strict exponent margins. Indeed the product before expansion is $(H_1^2/F^3)(F^2+M)M(F^2+NF/H_1)NF/H_1$. The four terms save a fixed power relative to $X^2$. Removing proper prime powers in the character estimate costs at most $X^{1/2}F\sqrt{H}L^{O(1)}$ by Cauchy and the shift large sieve. This proves that (2.1) has only the stated diagonal to arbitrary logarithmic accuracy.

For the norm, write $m = n+j$, $t = n-hv$. Counting progressions with the two divisor coefficients costs $O((1+H)R^2L^6)$ after division by $h$. Periodic orthogonality leaves coefficient $\mu^2(r)c_r(j)/\varphi(r)^2$ at equal denominators. Only $j = 0$ contributes to logarithmic accuracy. Thus $\|D\|^2 = XhA \int \omega \int W^2 + O_D(XhL^{-D})$. The test has zero periodic mean. The floor error when pairing with the constant $h$ is $O(h^2RL^3) = o(Xh)$. The covariance with the centered prime count is $XhA \int \omega \int W + O_D(XhL^{-D})$ after that power-saving correction. Cauchy now gives the base bound of Theorem 1 for $\vartheta$, on letting $\eta \downarrow 0$. The window of higher prime powers has squared norm $\ll h(1+h)\sqrt{X}L^3$: each power belongs to a set of starts of length $h$, and any one window contains at most $1+h$ integers. The same bound follows for $\psi$ by the triangle inequality.

# 3 Transfer to zeta zeros

## 3.1 Testing a prime Dirichlet polynomial

Let $X = T^{\alpha},\ H = X/T,\ L = \log X$. Then $1 \le H \le X^{1/3-c_\epsilon}$. Set $U = HX^{2\eta}$, $R = \sqrt{X/H}\, X^{-10\eta}$, and put

$$B(n) = \sum_{U<r\le R} \frac{\mu(r)c_r(n)}{\varphi(r)}, \qquad A = \sum_{U<r\le R} \frac{\mu^2(r)}{\varphi(r)} = \log(R/U) + O(1),$$

where $0 < \eta$ is sufficiently small relative to $\epsilon$. We use the $q = 1$ band, with all denominators present. We recall the estimate on which the proof depends. Suppose $b_0$ is smooth and compactly supported in a fixed subinterval of $(0,\infty)$, and $V$ is Schwartz on the line. Then

$$\sum_p \log p\, b_0(p/X) \sum_j B(p+j)V(j/H) = AV(0)\sum_p \log p\, b_0(p/X) + O_D(XL^{-D}). \tag{3.2}$$

One may require $p + j > 0$ here since the Schwartz tail costs an arbitrary negative power; $B(0)$ has the trivial polynomial bound. Constants are controlled by finitely many smooth seminorms of $b_0, V$ for each fixed $D, \eta, \epsilon$. Similarly

$$\sum_n b_0(n/X)B(n)\sum_j B(n+j)V(j/H) = XAV(0)\int b_0(u)\,du + O_D(XL^{-D}). \tag{3.3}$$

For the first identity use the argument of §2 with $b_0(p/X)V(j/H)$ in place of $\omega((p-hv)/X)W(v + j/H)$, omitting the $v$-integral. Pólya–Vinogradov with partial summation applies to both tails. Poisson gives the same prefactor $H_1/\sqrt{f}$ and dual length $fX^{o(1)}/H_1$ with $H_1 = H/v_1$. Schwartz decay permits the truncations and the Mellin separation in the proof of the window estimate. At zero shift the principal term equals $A$, while the remaining principal terms cost $O(X^{1-c_\eta})$ since $U > HX^{2\eta}/2$ and $\sum_{0<|j|\le HX^{\eta'}}(j,r) \ll HX^{\eta'}\tau(r)$. Schwartz tails allow any $\eta' > 0$ small enough. For the nonprincipal part, primitive conductor $f \sim F$ and imprimitive divisor $v$ give $U_{\psi,v} = \sum_m \bar\psi(m)V(mv/H)$. If $v$ exceeds $H$ by a small power, arbitrary Schwartz decay absorbs the contribution. In the intermediate case $H/v < 1$ the same completion and large-sieve bounds hold after an arbitrarily small power loss. For (3.3) the smooth progression count uses $O(HX^{\eta'}R^2L^{O(1)})$ total floor error. Orthogonality at the two denominators and the same nonzero-shift estimate prove the assertion.

Both identities permit $b_0(p/X)V(j/H)$ to be replaced by a kernel $K(p/X, j/H)$, smooth and supported in a fixed compact interval in its first argument, uniformly Schwartz with all derivatives in the second. The corresponding diagonal is $K(p/X, 0)$. For detail, expand periodically in the first argument on a fixed interval containing the support in its interior, then multiply by a fixed cutoff equal to one on that support. The Fourier coefficients as functions of the second argument have any prescribed Schwartz seminorm $O_m((1+|k|)^{-m})$, by integration by parts. The prime endpoints and partial summations for the $k$-th frequency cost only a polynomial in $1 + |k|$; choose $m$ larger. This also applies to (3.3) with $n$ in place of $p$.

Take a fixed nonnegative $g \in C_c^\infty((0,1))$, bounded by one, with mass $g_0 = \int g$, and define $\widehat{g}(u) = \int g(v)e^{-2\pi iuv}\,dv$. Let $a$ be one of the real weights

$$w_1(v) = v^{1/2}\mathbf{1}_{0<v\le 1}, \qquad w_2(v) = v^{-3/2}\mathbf{1}_{v>1}.$$

Take a real $b \in C_c^\infty((0,1))$ in the first case, or $b \in C_c^\infty((1,\infty))$ in the second. Define

$$S_a(t) = \sum_n \Lambda(n)a(n/X)n^{-it}, \qquad D_b(t) = \sum_m B(m)b(m/X)m^{-it}.$$

We claim

$$\int g(t/T)S_a(t)\overline{D_b(t)}\,dt = TXAg_0\int a(u)b(u)\,du + O_D(TXL^{-D}),$$
$$\int g(t/T)|D_b(t)|^2\,dt = TXAg_0\int b(u)^2\,du + O_D(TXL^{-D}). \tag{3.4}$$

Constants may depend on the fixed smooth functions. To see the first identity integrate the phase at $m = n + j$. The kernel is $T\widehat{g}((T/(2\pi))\log(n/m))$. Outside $n$ comparable to $X$, with $m/X$ in the compact support of $b$, it costs an arbitrary negative power of $X$. Indeed $\sum_n \Lambda(n)|a(n/X)| \ll XL$ for either weight and $\sum_m |b(m/X)B(m)| \ll X^{1+o(1)}$, by the divisor expansion of $B$. At $n \asymp X$, restrict to $|j| \le HX^{\eta'}$ by the same decay, allowing power-small error. On this range replace $b((n + j)/X)$ by $b(n/X)$ and $(T/(2\pi))\log(n/(n + j))$ by $-Tj/(2\pi n)$. Their respective pointwise errors are $O(X^{\eta'}/T)$ and $O(X^{2\eta'}/T)$. Even after summing all $j$ the relative error before the scale $TX$ is $O((1 + H)X^{3\eta'+o(1)}/T)$, power-small. The distance of $\operatorname{supp} b$ from the endpoint 1 leaves $ab$ smooth. The proper prime powers with $n \asymp X$ cost $O(T\sqrt{X}(1 + H)X^{\eta'+o(1)})$, also power-small. Now apply (3.2) with

$$K(\beta, u) = a(\beta)b(\beta)\widehat{g}(-u/(2\pi\beta)).$$

The PNT on its smooth compact support supplies the first main term in (3.4). For the second identity use (3.3) and $b(\beta)^2$ instead. The same linearization errors remain negligible because $B(n) \ll \tau(n)L^3$ on the positive integers.

Consequently Cauchy gives

$$\int g(t/T)|S_a(t)|^2dt \ge (1 - o(1))TXAg_0\,\frac{(\int ab)^2}{\int b^2}. \tag{3.5}$$

For both weights $\int_0^\infty a(u)^2du = 1/2$. Approximating $a$ in $L^2$ by these compact interior $b$'s makes the ratio tend to $1/2$. These are fixed-function approximations before $T \to \infty$.

### 3.2 Zeros on the critical line

Write $F(X, T) = (T\log T/2\pi)F_T(\log X/\log T)$. For $B_0 > 2$, let $\Psi$ be the nonnegative GGÖS cutoff approximating the indicator of $[0, 1]$ with endpoint collars of width $1/\log^{B_0} T$. It is one on the support of our fixed $g$ for sufficiently large $T$. By GGÖS, Lemma 1, equivalently Walker, Lemma 4.1, on RH for $\zeta$ alone,

$$F(X, T) = \frac{I_1}{2\pi X^2} + \frac{X^2 I_2}{2\pi} + O(T\log^{2-B_0} T) + O_\delta(X^{1+\delta}/T), \tag{3.6}$$

where

$$I_1 = \int \Psi(t/T)\left|\sum_{n\le X}\Lambda(n)n^{1/2-it} - \int_1^X u^{1/2-it}\,du\right|^2 dt,$$
$$I_2 = \int \Psi(t/T)\left|\sum_{n>X}\Lambda(n)n^{-3/2-it} - \int_X^\infty u^{-3/2-it}\,du\right|^2 dt.$$

Changing the signs of all Mellin phases in one square has no effect. On the support of $g(t/T)$ one has $|t| \gg_g T$, so the continuous integral in either square after scaling to $S_{w_1}$ or $S_{w_2}$ is $O_g(X/T)$ by direct integration. Its squared $g$-norm is $O_g(X^2/T) = o(TX)$ here. Since $\Psi \ge g$, (3.5) and the triangle inequality thus yield

$$I_1 \ge (1/2 - o(1))TX^2A, \qquad I_2 \ge (1/2 - o(1))TX^{-2}A$$

after allowing $g_0 \to 1$ and improving the $L^2$ approximation on each side. In (3.6) choose a small fixed $\delta > 0$. Its errors are $o(T \log T)$ for the stated $\alpha$. The two terms add, and

$$\frac{A}{\log T} = \frac{\log(\sqrt{T}/(X/T))}{\log T} - 12\eta\alpha + o(1) = \tfrac{3}{2} - \alpha - 12\eta\alpha + o(1).$$

Taking the smooth approximation losses and $\eta$ arbitrarily small proves the base estimate of Theorem 2.

### 3.3 Complex differences of zeros

Let $\rho = 1/2+\delta+i\gamma$ run over the nontrivial zeros. Define the statistic of Baluyot–Goldston–Suriajaya–Turnage-Butterbaugh (BGSTB) by

$$F_{\mathrm{c}}(X,T) = \sum_{0<\gamma,\gamma'\le T} X^{\rho-\rho'} \frac{4}{4-(\rho-\rho')^2}.$$

Unconditionally, in the same range,

$$\frac{2\pi F_{\mathrm{c}}(T^\alpha, T)}{T\log T} \ge \tfrac{3}{2} - \alpha - o(1). \tag{3.7}$$

On ordinary RH the statistic agrees with $F$, giving a second proof of the base estimate of Theorem 2. The two unconditional ingredients of BGSTB that we use are their Lemmas 1 and 3. Write

$$M_T(t) = \sum_{0<\gamma\le T} \frac{2X^{\delta+i(\gamma-t)}}{1+((t-\gamma)+i\delta)^2}.$$

Lemma 3, using the symmetry $\rho \mapsto 1-\bar\rho$, gives exactly $2\pi F_{\mathrm{c}}(X,T) = \int_{\mathbb{R}} |M_T(t)|^2 dt$. Let $M_\infty$ denote the same sum over all zeros, including negative ordinates. On the support of $g(t/T)$, both endpoints $0, T$ have distance $\gg_g T$. By the local zero-count bound $N(u+1)-N(u) \ll \log(|u|+2)$ and $|\delta| \le 1/2$, we have the uniform pointwise bound

$$|M_\infty(t) - M_T(t)| \ll_g X^{1/2}\log(2T)/T.$$

The squared norm of this error over the interior is $O_g(X\log^2 T/T) = o(T)$. In particular no equality with a sharply truncated integral on the entire interval $[0,T]$ is required here.

Lemma 1 of BGSTB, which is the unconditional explicit formula on the line $3/2+it$, gives on the same support

$$M_\infty(t) = -X^{-1/2}S_{w_1+w_2}(t) + O_g\Big(X^{-1}\log T + X^{1/2}/T^2 + X^{-5/2}/T\Big).$$

This error also has squared norm $o(T)$. The mean-square lower bound (3.5) applies to $w_1+w_2$: choose a single $b$ with compact support away from $0, 1, \infty$ approximating the sum in $L^2$. Its support may have two components. In (3.4) terms across the fixed gap about 1 are power-small after $t$-integration, and the same proof uses the two interior cutoffs. The new squared mass is $\int (w_1+w_2)^2 = 1$. Thus

$$2\pi F_{\mathrm{c}}(X,T) \ge \int g(t/T)|M_T(t)|^2 dt \ge (1-o(1))TA.$$

Here the small losses in $g, b, \eta$ are taken as before. This proves (3.7). BGSTB's Theorem 1 already evaluates their complex-difference statistic unconditionally through $|\alpha| = 1$. Away from RH, (3.7) refers to the whole complex differences of the zeros.

## 4 Conductors in the displaced prime sums

For any squarefree denominator family $\mathcal{R}$, put $B_{\mathcal{R}}(n) = \sum_{r\in\mathcal{R}} \mu(r)c_r(n)/\varphi(r)$ and $A_{\mathcal{R}} = \sum_{r\in\mathcal{R}} 1/\varphi(r)$. For the two disjoint families write $B_i = B_{\mathcal{R}_i}$ and $A_i = A_{\mathcal{R}_i}$. The first is a full band below $\sqrt{X}$ and the second lies just above $\sqrt{X}$. In this section $U$ denotes any lower bound for the denominators under consideration. We use unit stride $s = 1$ throughout.

**Small conductors, including every exception.** For a prime $p \asymp X$ the prime is a unit at every test denominator. Expand $c_r(p + js)$ on the multiplicative characters of the units modulo $r$. For a primitive $\chi$ mod $f$, $r = fm'$, squarefreeness gives the following coefficient of $\chi(p)$ after multiplication by $\mu(r)/\varphi(r)$:

$$\frac{\mu(f)f\,\overline{\chi}(-js)}{\varphi(f)^2}\,\frac{c_{m'}(js)}{\varphi(m')^2}. \tag{4.7}$$

Here $\chi \bmod 1 \equiv 1$, including at zero. At a prime of $f$ the unnormalized local sum over units is $p_0\overline{\chi}_{p_0}(-js)$, since $c_{p_0}(n) = p_0\mathbf{1}_{p_0|n} - 1$. At a prime of $m'$ the unit sum is $-c_{p_0}(js)$. This proves (4.7). The character projection in this expansion agrees exactly with projecting the divisor indicators modulo $d \mid r$: each such function on the units depends only on the class at $d$. Nonunit target classes at $d$ give the zero function on units.

The zero shift in (4.7) contributes $A$. For the other shifts use $(s, m') = 1$ and $\sum_{0<|j|\le J} |c_{m'}(js)| \ll J\tau(m')$. Here $m' > U/f$ for each $f \le K$. Summing up to $\varphi(f)$ characters and all $f \le K$, the absolute total of the coefficients summed over nonzero shifts is

$$\ll X^{o(1)}J \sum_{f\le K} \frac{f}{\varphi(f)} \sum_{m'>U/f} \frac{\tau(m')}{\varphi(m')^2} \ll X^{o(1)}JK^2/U. \tag{4.8}$$

Thus small conductors cost a power after the trivial prime estimate, uniformly in their character sums, even in the presence of an exceptional zero.

**Unrestricted divisors below the square root.** Write $R_0 = X^{1/2-b}$ and $K = X^k$. For each fixed nonzero shift retain the primitive conductors $f > K$ at $d \le R_0$ and sum absolute discrepancies over $d$. For a balanced Vaughan block $MN \asymp X$, both lengths at least $X^{.32}$, write $d = fz$, $f \asymp F$. The multiplicative large sieve on both factors costs

$$X^{o(1)}(F\sqrt{X} + \sqrt{XM} + \sqrt{XN} + X/F)/\varphi(z).$$

The sum on $z$ costs a log power. Free long blocks use elementary progression counting on the free variable as below. For the small $b, k$ used here these estimates save $X^{\min(b,k)-o(1)}$ per shift.

**Factorization for the larger divisors.** For squarefree $d$ and a compatible reduced class $a$ set

$$u_K(n, a; d) = \mathbf{1}_{n\equiv a(d)} - \varphi(d)^{-1} \sum_{\substack{\chi \bmod d\\ \operatorname{cond}\chi\le K}} \overline{\chi}(a)\chi(n).$$

This conductor subtraction is Drappeau's [2]. Factor the large divisor as $d = qr$ with every prime of size at most the core roughness threshold on the inner factor $r$. Denote that threshold by $X^a$. In the formulas for $u_K$ the letter $a$ in a residue slot denotes the prescribed class. The factorizations and the allowed parameters are specified in §§6–7. They have $K < X^a$ and $NX^{-3e-\delta} \ll r \ll NX^{-3e}$,

where $N$ is the shorter bilinear length and $e$ the completion parameter. For a sequence $g$, split the discrepancy into

$$E_1 = \sum_{n\equiv a(qr)} g(n) - \frac{1}{\varphi(q)} \sum_{\substack{n\equiv a(r)\\(n,q)=1}} g(n),$$

$$E_2 = \frac{1}{\varphi(q)} \sum_{(n,q)=1} g(n) u_K(n,a;r).$$

Their sum is $\sum g(n)u_K(n,a;qr)$. Indeed every character of conductor $\le K$ at $qr$ is induced from $r$. The average over the units at $q$ passes those characters exactly.

**Change of the outer residue.** The following accounting strengthens the error in the derivation of (5.18) in Polymath for the chosen splits, using $x = X$. Take $P$ to be the product of all primes in the moduli after deleting the primes dividing the shift. Average over a single global reduced class $b \bmod P$ and use the triangle inequality as just before their (5.18). It suffices to bound the sum of absolute differences between the two conditions $mn \equiv b_i \bmod q$, $i = 1, 2$, in the common class $mn \equiv a \bmod r$. Insert complex signs $c_{q,r}$ bounded by one, setting them zero outside our chosen splits. Cauchy in $r, m$ gives their (5.19): the original sum is bounded by $x^{o(1)}(RM\Sigma)^{1/2}$. In the four-term expansion of $\Sigma$, the terms below are compared with $MN^2/R$.

- At zero translate $\ell = (n_2 - n_1)/r = 0$, the displayed bound on page 47 is $x^{1+o(1)}$. As $R/N \ll x^{-3e}$ it saves a power. The smooth majorant of the $m$-interval may be fixed. Completion (Lemma 4.9(ii)) discards nonzero Fourier frequencies beyond the $x^e$ cutoff with arbitrary power saving, by repeated integration by parts.

- For $\ell \neq 0$, the zero-frequency main term from $(q_1, q_2) = 1$ is independent of $b_1, b_2$. All four copies have the same $c_{q_1,r}\overline{c_{q_2,r}}$, so it cancels. If $q_0 = (q_1, q_2) > 1$, it is at least $x^a$. The bound immediately preceding (5.26) is

  $$x^{o(1)} \left( MN + \frac{MN^2}{Rx^a} \right).$$

  For example the congruence in $\ell$ allows at most $1 + O(N/(Rq_0))$ choices per $n$, and the outer sum contributes $\sum_{q_0 \ge x^a} q_0^{-1}(1 + N/(Rq_0))$. Individual divisor bounds $x^{o(1)}$ suffice.

- For the nonzero frequencies, (5.27)–(5.28) remove the masks by absolute values *outside* the $n$-sum. The resulting nonnegative sums are bounded by the full sums in the proof of Polymath Theorem 5.8(i),(iv), restricted to the one-fold densely divisible moduli and their chosen factorizations as in Section 6. Here $P$ is the product of the allowed primes and the shifted residue is fixed compatibly at every prime of $P$. Its (5.31), summed as on page 50, gives $x^{-e+o(1)}MN^2/R$, including the $x^e$ in (5.26). The two branch conditions are $54\varpi+15\delta+5\sigma < 1$ and $68\varpi+14\delta < 1$. For $N$ below the transition $x^{1/2-2\varpi-c_0}$ use (i), above it use (iv), taking the same sufficiently small fixed $c_0 > 0$. The input $\beta$ in Theorem 5.8 need only be a coefficient sequence. The preliminary Siegel–Walfisz use on pages 42 and 45 concerns other pieces of the published proof.

Taking square roots after the nonzero-frequency estimate and the two zero-frequency bounds gives a relative saving of order $X^{-\min(e/2,a/2)+o(1)}$ for the outer change. The terminal exponents used for each application are given below.

**The inner modulus.** Expand $E_2$ over primitive characters of conductor $f \mid r,\ f > K$. For fixed $q$ write $r = fz$, use that $(f, z) = 1$ and insert $\mathbf{1}_{(\cdot, zq)=1}$ into each sequence. For $f \asymp F$, discard the mask after absolute values. The two classical multiplicative large sieves give the bound, apart from the exterior $1/(\varphi(q)\varphi(z))$,

$$x^{o(1)}F^{-1}\{(F^2+M)(F^2+N)MN\}^{1/2} \ll x^{o(1)}(F\sqrt{x}+\sqrt{xM}+\sqrt{xN}+x/F).$$

Here $F \le 2R \ll \sqrt{x}\,x^{-3e}$, $F \gg K/2$, and the longer sequence has length $\ll x^{.68}$. All four terms save a fixed power. The sums of $1/\varphi(z)$ and $1/\varphi(q)$ cost log powers. This estimates the inner discrepancy. To restrict $mn$ to an interval, use truncated Perron on each endpoint taken at a half-integer. Height $x^{10}$ makes the error $O(x^{-7})$ per term since $mn \asymp x$; the integration costs $O(\log x)$. The inserted phases $m^{it}, n^{it}$ do not affect divisor bounds or the uniformity in Theorem 5.8. The real part $1/\log x$ has bounded cost.

**Vaughan blocks and proper powers.** Take both Vaughan cutoffs $Z = X^{1/3}$. The pieces on the macroscopic support are $\mu_{\le Z} * \log$, $-\mu_{\le Z} * \Lambda_{\le Z} * 1$ and $(\mu_{>Z} * 1) * \Lambda_{>Z}$. In the last the two variables exceed $Z$; in the first two the marked partner of the free variable is at most $Z^2$. If the marked partner has length $\ll X^{n_0}$ with $n_0 < 1/3$, counting the free variable in a reduced progression against its unit mean costs $O(\tau(d)L^{O(1)})$ per marked value and modulus. The projected nonprincipal characters on the free variable cost $X^{o(1)}K^{3/2}/\varphi(d)$ after Pólya–Vinogradov and inclusion over primes dividing $d/f$. If the free variable itself has length below $X^{n_0}$ and its marked partner is at most $Z^2$, their product misses the support. Proper powers in the shifted divisor count cost $X^{1/2+o(1)}$ per translation: each nonzero shifted integer has $X^{o(1)}$ divisors. At a zero argument the cost is $X^{1/2+\alpha_1+v+o(1)}$. Their small-character projection costs $KX^{1/2+o(1)}$ after summation over $d = fz$.

# 5 The two diagonals

Write $L = \log X$. For a family $\mathcal{R}$ of squarefree integers let

$$B_{\mathcal{R}}(n) = \sum_{r\in\mathcal{R}} \frac{\mu(r)c_r(n)}{\varphi(r)}, \qquad A_{\mathcal{R}} = \sum_{r\in\mathcal{R}} \frac{1}{\varphi(r)}.$$

Here $c_r$ is the Ramanujan sum. Write $\mathbb{E}_{\mathrm{per}}$ for exact averaging over the common period of functions of this type. Distinct denominators are orthogonal under this average, so $\mathbb{E}_{\mathrm{per}}B_{\mathcal{R}}^2 = A_{\mathcal{R}}$. We compare the periodic average with the actual mean on the integers at scale $X$.

## 5.1 A mixed moment with one unrestricted band

**Lemma 3.** *Let $q \le X^C$ and fix $\xi > 0$. Let $B$ be the sum above over every squarefree $r \le Z$ coprime to $q$, or the difference of two such sums, $Z \le X$. Let $D(n) = \sum_{d\mid n} v_d$, where $|v_d| \le L^C$, $d$ is squarefree and $d \le Y \le X^{1-\xi}$. For a smooth $w$ of compact support in $(0,\infty)$, with fixed derivatives bounded by log powers, we have*

$$\sum_n w(n/X)B(n)D(n) = X\int w\,\mathbb{E}_{\mathrm{per}}(BD) + O\Big(Xe^{-c_\xi\sqrt{L}}\Big). \tag{5.1}$$

*The same holds for $X^{1-o(1)}$ in place of the physical scale while $q, Z, Y$ obey the original size bounds, provided $Z \le X^{O(1)}$ and the displayed power gap for $Y$ persists.*

*Proof.* Consider the full sum through $Z$ with divisor coefficients

$$b_h = \frac{\mu(h)h}{\varphi(h)}\mathbf{1}_{(h,q)=1}\sum_{\substack{v\le Z/h\\(v,hq)=1}}\frac{\mu^2(v)}{\varphi(v)}.$$

The Dirichlet series of $BD$ is $\zeta(s)F(s)$, where $F(s)=\sum_{h,d} b_h v_d/[h,d]^s$ is entire. The residue $F(1)$ is the exact periodic mean. Write $\phi_s(u)=\prod_{p\mid u}(p^s-1)$ on squarefree $u$. Expanding $(h,d)^s = \sum_{u\mid(h,d)}\phi_s(u)$ gives

$$F(s)=\sum_{u\le\min(Z,Y)}\frac{\mu^2(u)\phi_s(u)}{u^{2s}}\,\Gamma_b(u,s)\Gamma_v(u,s),\qquad \Gamma_v(u,s)=\sum_{d\le Y/u,\ (d,u)=1}\frac{v_{ud}}{d^s}. \tag{5.2}$$

Define $\Gamma_b$ by the same formula with $Z,b$. This is the mixed form of Hooley's factorization, as in Walker (3.2)–(3.5).

Use $\sigma=1-\lambda,\ \lambda=c/\sqrt{L}$ for a sufficiently small fixed $c>0$. On this line,

$$\left|\sum_{\substack{d\le V\\(d,h)=1}}\frac{\mu(d)d}{\varphi(d)d^{\sigma+it}}\right|\ll(1+|t|)L^{C_1}\qquad(V\le X^{O(1)},\ h\le X^{O(1)}). \tag{5.3}$$

Indeed the classical zero-free region gives $\sum_{m\le v}\mu(m)\ll v\exp(-c_0\sqrt{\log(2v)})$. Partial summation bounds $\sum_{m\le V}\mu(m)m^{-\sigma-it}$ by $O(1+|t|)$: in the integral $\lambda\log v\le(c_0/2)\sqrt{\log v}$. To impose $d/\varphi(d)$ and coprimality, convolve $\mu$ with the coefficients of an absolutely convergent correction. Its local factor at $p\nmid h$ is $(1-p^{1-s}/(p-1))/(1-p^{-s})$, and at $p\mid h$ it is $(1-p^{-s})^{-1}$. The absolute norm on the line is $O(L^{C_1})$, uniformly in $h$: for $p\le L^2$, $p^\lambda=O(1)$, while $h$ has $O(L)$ larger prime divisors. This proves (5.3).

Expanding $b_{ud}$ in (5.2) gives the sum in (5.3) with $h=uvq$ and $V=Z/(uv)$, integrated against the discrete weights $\mu^2(v)/\varphi(v)$ and a prefactor of magnitude at most $u/\varphi(u)$. Thus $|\Gamma_b|\ll(1+|t|)L^{C_2}$. Direct summation gives $|\Gamma_v|\ll L^{C_2}(Y/u)^\lambda$, absorbing $1/\lambda$; also $|\phi_s(u)|\ll u^\sigma L^{C_2}$. The remaining $u$-sum is harmonic, since $\sigma+\lambda=1$. We obtain

$$|F(\sigma+it)|\ll(1+|t|)L^{C_3}Y^\lambda.$$

Shift the Mellin integral for the smooth sum to this line. Zeta has a polynomial bound there, including a factor $O(1/\lambda)$ near the real axis. Rapid decay of the Mellin transform against fixed powers of $t$ costs log powers. Thus the error is $O(XL^{C_4}(Y/X)^\lambda)$, as asserted. Intervals reaching 0 can instead be divided into dyadic pieces, or truncated with a smooth logarithmic collar when that is the chosen test. □

## 5.2 The diagonal for the larger denominators

Let $\mathcal{C}$ be any family of squarefree $X^a$-rough integers in $[X^{1/2+\alpha_0},X^{1/2+\alpha_1}]$ with $0<\alpha_0<\alpha_1$, and fix $0<v<a$ with $2\alpha_1+2v<a$. Write $S$ for the full Ramanujan sum through $X^v$ with coefficients $\mu(u)/\varphi(u)$ and $A_S=\sum_{u\le X^v}\mu^2(u)/\varphi(u)$. Take arbitrary complex $|w_c|\le1$ and put

$$B_w(n)=S(n)\sum_{c\in\mathcal{C}}w_c\frac{\mu(c)c_c(n)}{\varphi(c)}.$$

Here $c_c(n)$ denotes the Ramanujan sum at denominator $c$.

**Lemma 4.** *For every fixed smooth compactly supported macro weight $W$, for some $\xi > 0$ depending only on the fixed power gaps,*

$$\sum_n W(n/X)|B_w(n)|^2 = X \int W\, A_S \left( \sum_{c\in\mathcal{C}} \frac{|w_c|^2}{\varphi(c)} - \left| \sum_{c\in\mathcal{C}} \frac{\mu(c) w_c}{\varphi(c)} \right|^2 \right) + O(X^{1-\xi}).$$

In the divisor expansion each coefficient is $O(L^3)$, $L = \log X$, as follows by summing $d/\varphi(d)$ times $1/\varphi(r/d)$ over denominators containing $d$. In its square let $d_1, d_2$ be divisors of small-times-core denominators. If one medium prime is omitted from either core, or if the two cores share a prime, then $[d_1, d_2] \le X^{1+2\alpha_1+2v-a}$. This argument applies to each representation of the divisor coefficient; a divisor containing a full core identifies that core uniquely, since no different core can contain it. The same conclusion for a large lcm also follows directly from the product upper bound independent of representation. Replacing the smooth count by its periodic mean for all such small lcms costs $O(X^{1-\xi})$, since each lcm $l$ admits at most $\tau(l)^2$ pairs of divisors and the floor error for each pair is bounded. Every other lcm contains two disjoint full cores and exceeds $X^{1+2\alpha_0}$, so no observed integer is divisible by it. At a full core $c$ the terminal divisor coefficient is $w_c\mu(c)c/\varphi(c)$. The small factors are prime to every core. The missed periodic mean is therefore

$$A_S \sum_{\substack{c,c'\in\mathcal{C}\\(c,c')=1}} \frac{\mu(c) w_c\, \mu(c')\overline{w_{c'}}}{\varphi(c)\varphi(c')}.$$

Dropping the coprimality condition costs $O(X^{-a+o(1)})$, including the $O(L)$ factor $A_S$, since $\sum_{p>X^a}(\sum_{p|c} 1/\varphi(c))^2 \ll L^{O(1)} \sum_{p>X^a} p^{-2}$. The full periodic second moment is $A_S \sum |w_c|^2/\varphi(c)$ by orthogonality. Subtraction proves the lemma. It also proves the cross covariance by polarization.

# 6 A finite lower sum and a per-shift estimate

## 6.1 The reciprocal mass of the larger denominators

Use $\sigma = .167,\ \varpi = .00075,\ \delta = .008,\ \nu = .001$, and put

$$a = .003, \quad \alpha_0 = .00001, \quad \alpha_1 = .00073, \quad v = .00074.$$

Take all squarefree $X^\delta$-densely divisible cores $c$ in $[X^{1/2+\alpha_0}, X^{1/2+\alpha_1}]$ with prime factors greater than $X^a$, weight one, and independently every squarefree small part $u \le X^v$. The split is unique. For a sorted list of prime log exponents, dense divisibility follows if each exponent is at most $\delta$ plus the sum of the preceding exponents. Indeed the consecutive subset-sum gaps stay at most $\delta$ by induction. We shall prove

$$\liminf_{X\to\infty} \sum_{c\in\mathcal{C}} \frac{1}{\varphi(c)} \ge 2605072188928623487/10^{21}. \tag{6.2}$$

With $a_* = a_1$ from Section 1, the extra covariance coefficient is at least $(a_* - o(1))L$, since $\sum_{u\le X^v} \mu^2(u)/\varphi(u) = vL + O(1)$.

Use exponent units $10^{-5}$, bin width $d = 25$, and integer mass scale $Z = 10^{21}$. State $(s, h)$ stores a lower reciprocal weight in units of $1/Z$ for prefixes with $h \le 45$ factors whose exponents sum to between $sd$ and $(s + h)d$. Start with weight $Z$ at $(0, 0)$. At bin $[kd, (k + 1)d)$, starting at 300 units, allow zero through three prime factors. The zero choice copies every entry. A nonzero choice requires

$(k+1)d \le 800+sd$. For $j$ distinct factors the limiting reciprocal mass is $\log(1+1/k)^j/j!$ by Mertens, since repeated prime values have negligible mass. From an old integer weight $w$, add

$$w' = \left\lfloor \frac{w}{(k+1)^j j!} \right\rfloor \quad \text{to the entry } (u,h') = (s+jk, h+j)$$

of the new table, retaining $h' \le 45$ and $(u+h')d + (k+1)d \le 50001$. The inequality $\log(1+1/k) \ge 1/(k+1)$ makes this a lower sum. Take each such bin with lower endpoint less than $50073/2$. All transitions use the previous bin table.

Whenever a nonzero transition additionally obeys $2ud + 800 \ge 50073$, insert a final prime. For a true prefix with exponent sum $t$ its allowed exponent interval is $[50001-t, 50073-t]$. It is above all previous primes, by the retention condition, and its top is at most $800+t$, by the additional condition. Thus the whole core is dense. The reciprocal mass of the final prime is at least $72/(50073-ud)+o(1)$ uniformly over the prefixes. Use uniform Mertens on this compact interval of positive exponents and $\log(B/A) \ge (B-A)/B$. The unique largest prime and the unique prefix multiset prevent repetition across transitions. Ties at endpoints may be omitted. Add $\lfloor 72w'/(50073-ud) \rfloor$ to an initially zero output. All rounding is down. The integer recurrence, reproduced in the appendix, gives the integer in (6.2), with 11,420,845 accepted transitions of positive integer weight out of 17,086,369 nonzero retained transitions of positive integer weight. All bin constants and counts are fixed before sending $X$ to infinity.

### 6.2 A dispersion saving with specified exponents

Take $e = 10^{-5}$ in the role of Polymath's $\varepsilon$. Put $c_0 = 4e$ and split at shorter bilinear length $N = X^{1/2-2\varpi-c_0}$. Keep all ordinary $X^{o(1)}$ factors until the end. In the bilinear blocks $N \gg X^{.333}$ up to dyadic constants and $N \ll \sqrt{X}$. The Vaughan cutoff is $X^{1/3}$. Free blocks with short marked factor below $X^{.333}$ cost $X^{.333+.502+o(1)}$ in progression counting, as in the free-block estimate of §4. A free factor shorter than this with its marked complement at most $X^{2/3}$ cannot reach the macroscopic product range.

Take the full squarefree base band $X^{u_0} < r \le X^{1/2-b}$. Each divisor of an extra denominator omitting a core prime is at most $X^{.5+\alpha_1+v-a} = X^{.49847}$. Thus for $0 < b < 10^{-5}$ every divisor above the base cutoff contains its full core and already exceeds $X^{.50001}$. Its upper bound $X^{.50147} < X^{.5+2\varpi}$ meets precisely Polymath (5.13). Write the extra divisor as $cd_s$ with $d_s \le X^v$. Dense divisibility of $c$ at $NX^{-3e}/d_s$ gives an inner part $r = d_s r_1$ with $NX^{-3e-\delta} \ll r \ll NX^{-3e}$. Every prime of the outer part $q$ exceeds $X^a$. The entire product is one-fold densely divisible since its small primes can be prepended. For each nonzero translate remove its prime divisors from the index set before fixing its compatible unit class.

Put $n = \log N/\log X$, allowing $o(1)$ in dyadic endpoints. After dividing the square bound by its squared target before saving, the three Type I terms on Polymath p. 56 have exponents at most

$$\tfrac{1}{4} + 9\varpi + 5\delta/2 + 9e - 5n/6, \quad 8\varpi + 3\delta + 11e - n, \quad -c_0 + e.$$

At $n \ge .333$ the first two are less than $-2e$, and the third is $-3e$. On the Type II side (5.36) and the three estimates directly below it give exponents at most

$$1 + 12\varpi + 7\delta/2 + (25/2)e - (5/2)n, \quad 8\varpi + 3\delta + 11e - n, \quad -2e$$

for $n \ge .5 - 2\varpi - c_0 + o(1)$; the first two are below $-2e$. These are the terminal smallness requirements on $e$ in the (i),(iv) exponential-sum arguments. Conditions (5.12)–(5.15) hold; in particular $rq^2 \ll X^{1+4\varpi+3e+\delta-.333+o(1)} \ll X^{.69}$. Proposition 5.12 uses one-fold dense divisibility and Proposition 5.10 the squarefree conditions. Their incomplete exponential-sum inequalities are uniform up to

$X^{o(1)}$. The completion tails in (5.22)–(5.26) save an arbitrary power for fixed $e$ by increasing the derivative order of the smooth cutoff. Thus (5.31) gives $X^{-e+o(1)}$ as displayed directly after that equation, including the exterior $X^e$ from (5.26). The zero translate costs a relative $X^{-3e+o(1)}$. For the nontrivial zero-frequency gcd the outer primes exceed $X^a$, giving relative error $X^{-a+o(1)} + O(X^{-3e+o(1)})$. The coprime main terms cancel. The outer residue change after Cauchy therefore costs at most $X^{1-e/2+o(1)}$.

Subtract conductors up to $K = X^k$ on the inner factor. The bilinear large sieve there has relative errors $O(X^{o(1)}(X^{-3e} + X^{-.15} + X^{-k}))$, since $F \le r \ll \sqrt{X}X^{-3e}$, $F > K/2$, and both bilinear lengths exceed $X^{.333+o(1)}$. The free-block Pólya–Vinogradov error with $K \le X^{10^{-5}}$ has a larger margin. These bounds, including Perron for the product endpoints, give relative saving $X^{-\min(k,e/2)+o(1)}$ per translate on the extra divisors. For all base divisors the usual two large sieves give $X^{-\min(k,b)+o(1)}$. Both estimates admit the squarefree divisor coefficients, of size $O(L^3)$, by taking absolute values in the moduli. The low primitive conductors on a nonzero shift keep the exact Ramanujan coefficient (4.7).

### 6.3 A bound using a per-shift estimate on both bands

Fix $0 \le \theta \le 3 \cdot 10^{-6}$, then an arbitrarily small $0 < \eta < 10^{-7}$. Set

$$k = b = \theta + \eta, \qquad u_0 = 3\theta + 4\eta, \qquad J = X^{\theta+\eta/4}.$$

For $H \le X^\theta$ the Schwartz shift tails past $J$ cost an arbitrary power. The high conductors through $J$ save a fixed power by the preceding estimates. The low conductors cost relative $X^{o(1)}JK^2/X^{u_0}$, also power-small. The dense-core diagonal, Lemma 4, applies as $2\alpha_1 + 2v = .00294 < a$. Its subtracted square is nonnegative up to the power-small shared-prime term. The full-base mixed diagonal follows from the Hooley factorization in Lemma 3, with periodic mean zero for the disjoint bands. At nonzero shifts Walker's Proposition 2.1 costs $X^{47/74+(53/74)(.50147)+o(1)}$ per shift, with a saving exceeding .005 before summation. The integer mean and proper-power errors have larger margins.

The total prime covariance coefficient divided by $L$ is at least $1/2 + a_* - 4\theta - 5\eta - o(1)$. The norm coefficient is at most the actual covariance coefficient up to $o(1)$, by the three diagonals. The Cauchy tests in Sections 2 and 3 now apply. Take the fixed smooth approximation errors arbitrarily small, then let $\eta$ tend to zero after the large-scale limit. This gives $F_T(\alpha), F_{T,\mathrm{c}}(\alpha) \ge 4 - (7/2 - a_*)\alpha - o(1)$ under their respective hypotheses: take $\theta = 1 - 1/\alpha$ and $L/\log T = \alpha$. Uniformity follows by covering the compact $\theta$-range by finitely many intervals of arbitrarily small fixed length and using the estimate at each right endpoint. For the shorter interval $1 \le \alpha \le 1.0000002$ one can simply use $\theta = 2 \cdot 10^{-7}$ throughout. Since $a_* - 8 \cdot 10^{-7} > 10^{-6}$, the stated numerical coefficients have fixed room. □

## 7 Combining the signed band with the extra denominators

Set $H = h$ or $X/T$. For the full band take all squarefree denominators strictly above $U = HX^{2\eta}$ through $R = \sqrt{X/H}\, X^{-10\eta}$, with fixed $\eta > 0$ small. Then

$$A_0 = \tfrac{1}{2}\log(X/H^3) - 12\eta \log X + O(1).$$

The signed-shift covariance and norm in Sections 2 and 3 apply to this band for the sliding window and for the two-variable kernel in the zero statistic. They apply throughout the present ranges since $H < X^{1/3-100\eta}$.

Only the restricted extra test needs a per-translation prime estimate. Suppose its denominators are $cu$, with $c$ squarefree and $X^a$-rough in $[X^{1/2+\alpha_0}, X^{1/2+\alpha_1}]$, and $u \le X^v$ squarefree; $c$ is $X^\delta$-densely divisible. Choose $0 < v < a,\ 2(\alpha_1 + v) < a,\ \alpha_1 + v < 2\varpi$. The split is unique. Every divisor

omitting a core prime has size at most $X^{1/2-a+\alpha_1+v}$. The others exceed $X^{1/2+\alpha_0}$ and admit exactly the inner-outer factorization of the factorization in §6. In particular there is no lift of moduli near $\sqrt{X}$. For each nonzero shift expand the divisor test and remove all conductors through $K = X^k$. On the smaller divisors the two ordinary large sieves, and free Vaughan terms, save $X^{\min(k,a-\alpha_1-v)+o(1)}$ relative to the trivial exponent 1, where here and below $o(1)$ allows either sign. On the larger divisors the dispersion proof gives a relative error

$$X^{-\min(e/2,\, a/2,\, k)+o(1)}. \tag{7.1}$$

Indeed its nonzero-frequency relative squared error is $X^{-e+o(1)}$, the shared outer prime costs $X^{-a+o(1)}$, and the zero translate of the bilinear difference costs $X^{-3e+o(1)}$. The inner-factor character estimate has relative error $X^{-\min(3e,k,.15)+o(1)}$. All small primitive conductors at this modulus are induced from the inner factor, since $k < a$. For the remaining part of the original prime covariance sum, formula (4.7) has complementary denominator $m' \geq X^{1/2+\alpha_0}/f$ for $f \leq K$. Thus its absolute error over $0 < |j| \leq J$ is at most

$$X^{1+o(1)} JK^2 / X^{1/2+\alpha_0}.$$

This estimate uses the large size of each *extra* Ramanujan denominator, independently of $U$. At $j = 0$ the full extra weight on an ordinary macroscopic prime is its exact harmonic mass $A_1$.

### 7.1 Numerical margins and a second range of shapes

Here are the parameters for the two estimates. All decimal entries are rational numbers.

| $i$ | $\delta$ | $\varpi$ | $e$ | $k$ | $a$ | $\alpha_0$ | $\alpha_1$ | $v$ |
|---|---|---|---|---|---|---|---|---|
| 1 | .008 | .00075 | .00008 | .0001 | .003 | .00001 | .00073 | .00074 |
| 2 | .00211 | .00002 | .00202 | .0012 | .00207 | .000001 | .000021 | .000017 |

Take $\sigma = .167$, the free/short threshold $n_0 = .33332 < 1/3$, and split the two dispersion cases at $n = 1/2 - 2\varpi - 4e$ for $n = \log N / \log X$, $N$ the shorter bilinear length. Every free variable shorter than $X^{n_0}$ with marked partner at most $X^{2/3}$ misses the macroscopic support. A marked partner shorter than that cutoff has the direct progression count already in the dispersion proof. In both rows $54\varpi + 15\delta + 5\sigma < 1$, $68\varpi + 14\delta < 1$ and $1 + 4\varpi + 3e + \delta - n_0 < .70$, verifying the product-size condition (5.15). In particular $Q < N$. The Type I first and second exponents, maximized at $n_0$, are

$$1/4 + 9\varpi + 5\delta/2 + 9e - 5n_0/6, \qquad 8\varpi + 3\delta + 11e - n_0.$$

They are strictly below $-2e$ in each row, and the third is $-3e$. At $n \geq 1/2 - 2\varpi - 4e$, the Type II exponents $1 + 12\varpi + 7\delta/2 + (25/2)e - (5/2)n$, $8\varpi + 3\delta + 11e - n$ are strictly below $-2e$; the third is at most $-2e$. These are exactly the three terms in the squared terminal estimates on pp. 53 and 56 of Polymath. Its fixed $\epsilon$ at completion is our $e$, which also gives $R_{\text{inner}} \ll NX^{-3e}$ and permits a fixed sufficiently high order of integration by parts. The individual conditions used in the reduction to (7.1) are set out in the per-translation proof in §6, which uses the same exponents before substituting different numbers. Both the omitted-core saving and (7.1) exceed $\theta_i$ in their respective rows. Use $J = HX^\tau$ for $\tau > 0$ smaller than these strict margins and $\eta$.

For row 1 the integer lower sum of §6 already gives $A_1/L \geq a_1 - o(1)$. For row 2 first choose a prime with log exponent in $(.00207, .00211)$, of reciprocal mass at least .018 for large $X$. Put $s = \delta + \sum t_i$, where $t_i$ are the chosen prime exponents. Until $s \geq (.5 + \alpha_1 + \delta)/2$, choose the next exponent between $.62s$ and $.99s$. These choices have reciprocal mass at least .46 each by Mertens. Their exponents increase strictly, including the second choice, since $.62(.00211 + .00207) > .00211$

and $.62(1+.62) > .99$. There are at most nine of these choices as $.00418(1.62)^9 > (.5+\alpha_1+\delta)/2$. At the stopping time insert a final prime with exponent in

$$[.5+\alpha_0+\delta-s,\ .5+\alpha_1+\delta-s].$$

Its lower endpoint exceeds $a$, since $s \le 1.99(.5+\alpha_1+\delta)/2$. Its upper endpoint is at most $s$ and at most $(.5+\alpha_1+\delta)/2$. Its reciprocal mass is at least .000075 after omitting previously chosen primes, by the interval width .000020 and uniform Mertens. At every insertion the new exponent is at most $\delta$ plus the sum of earlier exponents. Induction on subset-sum gaps proves one-fold dense divisibility. Each squarefree core occurs at most eleven times, since only the final prime needs designation and the preceding list is increasing. The harmonic mass of the complete extra denominators, divided by $L$, is therefore at least

$$.000017 \cdot .018 \cdot (.46)^9 \cdot .000075/11 - o(1) \ > \ a_2$$

for sufficiently large $X$.

### 7.2 The combined norm and transfer

The extra test in the window has zero periodic mean. Its pairing with the constant $h$ is therefore $O(h^2 X^{1/2+\alpha_1+v} L^{O(1)})$ by the divisor expansion. This is $o(Xh)$ here. The mixed diagonal at zero shift is $O(X\exp(-c\sqrt{\log X}))$, by the mixed Hooley factorization in Lemma 3. The rank-one subtraction in Lemma 4 applies to the extra square since $2\alpha_1 + 2v < a$. The base diagonal has its periodic mean $A_0$ to a power saving by direct progression counting since $R^2 \ll X^{1-20\eta}$. At nonzero shifts, Walker's Proposition 2.1 on the full divisor sum has error $X^{47/74+(53/74)(.5+\alpha_1+v)+o(1)}$ per shift, with room for $J$ in each case. The periodic off-diagonal contributes $X^{1+o(1)}J/U$ after summing, which saves a power. The covariance coefficient is $A_0+A_1$ and the test norm coefficient is at most $A_0+A_1+o(L)$. Cauchy, followed by the two zero transfers of §3, gives Theorems 1 and 2 after the fixed approximation errors and $\eta$ tend to zero after the asymptotic limit. All errors are uniform in each displayed range. The slopes use $H = X/T,\ \log X = \alpha \log T$. The numerical .500001 statements have strict slack at their endpoints since $a_1 > 1.92775 \cdot 10^{-6}$. □

## A An integer certificate for the denominator mass

The following Python code implements the recurrence in Section 6. All operations used in the lower sum are on integers. The returned triple gives the numerator in units of $10^{-21}$, the number of retained positive-weight transitions, and the number of those admitting a last prime.

```
from math import factorial

def core_certificate():
    step, density, lower, upper = 25, 800, 50001, 50073
    states = {(0, 0): 10**21}
    mass = retained = accepted = 0
    for left in range(300, upper // 2, step):
        k = left // step
        updated = states.copy()
        for (s, count), prev in states.items():
            if left + step > density + s * step:
                continue
            for j in range(1, min(3, 45 - count) + 1):
                u, h = s + j * k, count + j
```

```
                if (u + h) * step + left + step > lower:
                    break
                w = prev // ((k + 1)**j * factorial(j))
                if w == 0:
                    break
                retained += 1
                updated[u, h] = updated.get((u, h), 0) + w
                if 2 * u * step + density >= upper:
                    accepted += 1
                    mass += 72 * w // (upper - u * step)
        states = updated
    return mass, retained, accepted

assert core_certificate() == (2605072188928623487, 17086369, 11420845)
```

# References


[1] S. Baluyot, D. A. Goldston, A. I. Suriajaya and C. L. Turnage-Butterbaugh, An unconditional Montgomery theorem for pair correlation of zeros of the Riemann zeta-function.

[2] S. Drappeau, Sums of Kloosterman sums in arithmetic progressions, and the error term in the dispersion method.

[3] D. A. Goldston, The second moment for prime numbers, Quart. J. Math. Oxford (2) **35** (1984), 153–163.

[4] D. A. Goldston, A lower bound for the second moment of primes in short intervals, Expo. Math. **13** (1995), 366–376.

[5] D. A. Goldston, S. M. Gonek, A. E. Özlük and C. Snyder, On the pair correlation of zeros of the Riemann zeta-function.

[6] D. A. Goldston and C. Y. Yıldırım, On the second moment for primes in an arithmetic progression.

[7] A. J. Harper and K. Soundararajan, Lower bounds for the variance of sequences in arithmetic progressions: primes and divisor functions.

[8] A. L. Juhas, Lower bound estimates related to primes in short intervals, doctoral thesis, Universität Münster, 2016.

[9] D. H. J. Polymath, New equidistribution estimates of Zhang type.

[10] A. Walker, Correlations of sieve weights and distributions of zeros.

[11] L. Alpöge, announcement concerning the Riemann hypothesis and Dirichlet $L$-functions, 20 November 2026.