

CROSS-ENERGY OF DIRICHLET PARTIAL SUMS. PART 7: MULTIPLICATIVE RESONANCE AND LARGE VALUES OF DIRICHLET POLYNOMIALS

PEDRO CACERES

ABSTRACT. In the cross-energy $C_2(n, s) = |X_n(s)|^2 - |X_{n-1}(s)|^2 - n^{-2\sigma}$ of the partial sums $X_n(s) = \sum_{k \leq n} k^{-s}$ of the Riemann zeta function, the subtracted term $n^{-2\sigma}$ is the diagonal energy. Read as a budget for zeros off the critical line it gives the classical zero-density theorems of Bohr–Landau and Carlson, and a mollifier empties it; what remains is the largest eigenvalue of a cross-energy matrix of the peaks of a short Dirichlet polynomial, the object of Montgomery’s large value conjecture. We study this obstruction numerically. For the peaks of $\sum n^{-it}$ the top eigenvector concentrates on smooth numbers, and restricting to rough supports lowers its growth without stopping it for $N \leq 800$. We carry out two of Guth’s polynomial-time checks of the conjecture and a direct search for worst cases with $|b_n| \leq 1$ up to $N = 8000$, $T = N^2$: at $N = 1000$ the coefficients found are close to completely multiplicative, and at $\sigma = 0.7$ random completely multiplicative (Steinhaus) coefficients give even more large values; at $\sigma = 0.75$ the search gives more. With 24 to 64 realizations per size, their number at $\sigma = 0.7$ exceeds $N^{2-2\sigma}$ by a mean factor from 1.85 at $N = 250$ to 9.2 at $N = 8000$, with a relative standard deviation below 10%; the fitted growth $N^{0.46}$ is pre-asymptotic. Harper’s high-moment asymptotics, which concern averages over the random function, suggest heuristically a log-normal regime that is consistent with Montgomery’s conjecture but dominates until $\log N$ is of order 10^2 ; their transfer to averages over t for one realization is not proved. Numerical tests of the conjecture should therefore be measured against the random multiplicative model, not against $N^{2-2\sigma}$. We do not prove the Riemann hypothesis or Montgomery’s conjecture.

1. INTRODUCTION

1.1. Context. Let $X_n(s) = \sum_{k \leq n} k^{-s}$, $s = \sigma + it$, and let

$$C_2(n, s) = |X_n(s)|^2 - |X_{n-1}(s)|^2 - n^{-2\sigma} = 2 \operatorname{Re} (X_{n-1}(s) n^{-\bar{s}})$$

be the cross-energy of consecutive partial sums, the basic object of this series [CE1]. Its limit at a zero ρ is $1/|\rho|^2$ on the critical line, and the sum of these limits is the slope C_B of the energy of the primes $I(X) = \int_1^X (\psi(x) - x)^2 x^{-2} dx$ under the Riemann hypothesis (RH) and simple zeros [CE5, Sect. 2]. Part 6 [CE6] studies that energy at second order: under RH, simple zeros and the spacing condition (H*) of [CE6] it is extraordinarily rigid, with fluctuations of standard deviation 0.0035 (a numerical value under the linear independence hypothesis) around $C_B \log X + K$, $K = 2.51248\dots$. This paper turns to the other side of the same quantity. Off the line, the diagonal term $n^{-2\sigma}$ that C_2 subtracts is the energy that mean value theorems control, and counting zeros with $\beta > \sigma$ becomes a problem about the points where a short Dirichlet polynomial $D(t) = \sum_{N < n \leq 2N} b_n n^{-it}$ is simultaneously large. Montgomery’s large value conjecture [Mont] would give the density hypothesis; Guth [Guth] has observed that there is no meaningful numerical evidence for it and has proposed quantities that can be checked in polynomial time.

Date: October 2026. Part 7 of the series *Cross-energy of Dirichlet partial sums* (seventeen parts; the plan of the series is in Part 1 [CE1]).

2020 Mathematics Subject Classification. Primary 11M26; Secondary 11M06, 11N64, 11K65, 11Y35.

Key words and phrases. Riemann zeta function, cross-energy, zero density, large values of Dirichlet polynomials, Montgomery’s conjecture, Guth’s checks, random multiplicative functions, Steinhaus model, log-normal tails.

1.2. Results. The results, with their status, are the following.

- (1) *From energy to large values* (Sect. 2): the classical route from mean values to zero density (Bohr–Landau, Carlson) read as a budget of the diagonal energy that C_2 subtracts, recalled from [CE10, Prop. 3.9] and not proved again; the remaining obstruction is the largest eigenvalue of a cross-energy matrix of peaks (Lemma 1, elementary). The content is classical; the formulation is ours. Numerically, the top eigenvector for the peaks of $\sum n^{-it}$ concentrates on smooth numbers, and on rough half-supports the worst case found is smaller but still grows for $N \leq 800$.
- (2) *Guth’s checks* (Sect. 3, numerical): fluctuations of the difference-set density (Check 1) and flattened Schatten norms (Check 2), and a direct search for worst cases with $|b_n| \leq 1$ (Check 3) for $N \leq 8000$, $T = N^2$. The coefficients found by the search are close to completely multiplicative.
- (3) *Calibration against random multiplicative functions* (Sect. 4; numerical, with a heuristic model): the diagonal part of the fourth moment of $\sum f(n)n^{it}$ is the multiplicative energy of $[N, 2N]$, and so is its expectation over f (Proposition 3, elementary). At $\sigma = 0.7$ Steinhaus random completely multiplicative coefficients give more large values than the search, at $\sigma = 0.75$ fewer; with 64, 32 and 24 realizations for $N \leq 2000$, $N = 4000$ and $N = 8000$ we give means, medians, standard deviations, confidence intervals, quantiles and a split-sample check of the seed dependence (Table 4). Harper’s high-moment asymptotics, used as a heuristic and labelled as such, bound the excess over $N^{2-2\sigma}$ by a power of N with exponent below 1 at computable sizes that disappears asymptotically; the agreement with the data is qualitative and holds at $\sigma = 0.7$ only (Remark 4).

1.3. What is classical and what is new. The zero-density theorems of Bohr–Landau and Carlson [T, Ch. 9], the Halász–Montgomery method and Montgomery’s conjecture [Mont], Bourgain’s counterexample [Bou], the resonance method [Sou, BS], the theorem of Guth and Maynard [GM] and Harper’s moment asymptotics [Hp, Hp2] are known and cited as such. New are: the reading of the mollifier as emptying the diagonal energy of C_2 (a formulation); the experiments on the cross-energy matrix of peaks and on smooth and rough supports; the implementation of Guth’s checks, including the Farey structure of the spikes in Check 1; the multiplicative structure of the worst cases found by the search; and the statistical calibration against the Steinhaus model. All of these are numerical observations or heuristics, not theorems. We have not found earlier numerical tests of Montgomery’s conjecture.

1.4. Organization. Sect. 2 explains how the diagonal energy leads to large values. Sect. 3 carries out Guth’s checks and the search for worst cases. Sect. 4 calibrates them against random multiplicative functions. Sect. 5 discusses what a numerical test of the conjecture should look like and lists open problems.

2. FROM ENERGY TO LARGE VALUES

Part 6 [CE6] describes the energy of the primes under RH. Here we ask what energy can say about zeros off the line. The answer is classical; we give it in the language of the cross-energy, briefly, because it leads to the large value problem of Sects. 3–4. Recall that $C_2(n, s)$ subtracts the *diagonal energy* $n^{-2\sigma}$ from the energy increment of the partial sums; averaged over t , the cross terms cancel and only the diagonal survives, and $\sum_n n^{-2\sigma}$ is finite exactly when $\sigma > \frac{1}{2}$. Let $N(\sigma; T, 2T)$ be the number of zeros with $\beta > \sigma$ and $T < \gamma \leq 2T$. We use the approximation $\zeta(s) = \sum_{n \leq x} n^{-s} - \frac{x^{1-s}}{1-s} + O(x^{-\sigma})$ for $|t| \leq \pi x$ [T, Thm. 4.11], the mean value theorem of Montgomery

and Vaughan [MV74],

$$(1) \quad \int_T^{2T} \left| \sum_{n \leq M} a_n n^{-it} \right|^2 dt = \sum_{n \leq M} |a_n|^2 (T + O(n)),$$

and Littlewood's lemma [T, Sect. 9.9]. With $a_n = n^{-\sigma_0}$ and $M = 2T$, (1) and the approximation give the classical mean value $\frac{1}{T} \int_T^{2T} |\zeta(\sigma_0 + it)|^2 dt = \zeta(2\sigma_0) + O(T^{1-2\sigma_0})$ for fixed $\frac{1}{2} < \sigma_0 < 1$ [T, Thm. 7.2]: the mean energy is the diagonal energy (Figure 1a). Littlewood's lemma on $[\sigma_0, 2] \times [T, 2T]$ and Jensen's inequality turn this into the theorem of Bohr and Landau in energy form, proved in [CE10, Prop. 3.9]: for $\frac{1}{2} < \sigma < 1$,

$$N(\sigma; T, 2T) \leq \frac{\log \zeta(\sigma + \frac{1}{2})}{2\pi(\sigma - \frac{1}{2})} T + o(T).$$

We do not repeat the proof. Status: classical [T, Sect. 9.15], proved in [CE10]. Each zero with $\beta > \sigma_0$ consumes $2\pi(\beta - \sigma_0)$ of the budget $\frac{T}{2} \log \zeta(2\sigma_0)$, and with $\sigma_0 = \frac{1}{2}(\sigma + \frac{1}{2})$ the fraction of the zeros in $(T, 2T]$ with $\beta > \sigma$ is at most $(1 + o(1)) \log \zeta(\sigma + \frac{1}{2}) / ((\sigma - \frac{1}{2}) \log T)$: almost all zeros are within $o(1)$ of the line (Figure 1b). At $\sigma_0 = \frac{1}{2}$ the diagonal energy $\sum_{n \leq 2T} 1/n \approx \log T$ is of the same order as the number of zeros per unit height, and the method loses exactly at the critical line, the abscissa at which the diagonal energy subtracted by C_2 ceases to be finite.

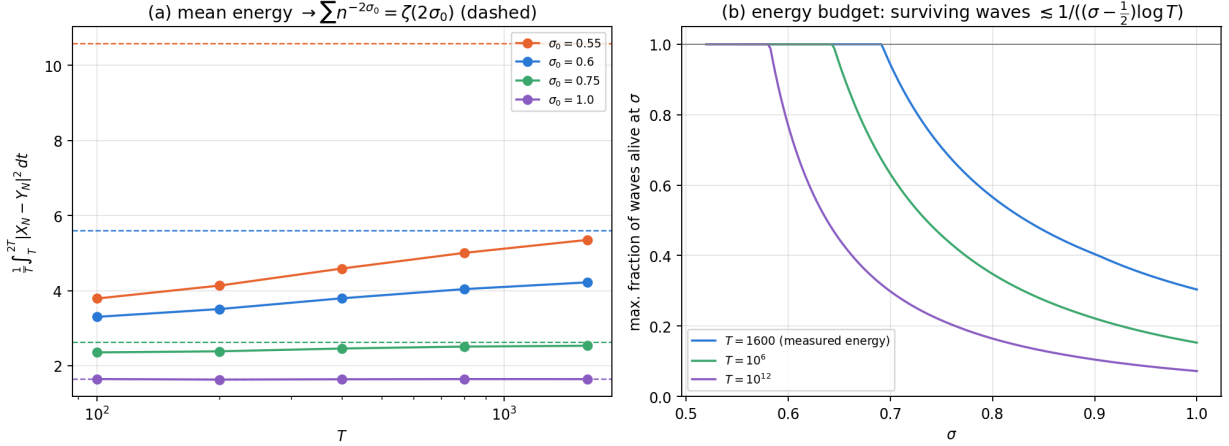


FIGURE 1. (a) The mean energy $\frac{1}{T} \int_T^{2T} |X_N - Y_N|^2 dt$, $Y_N(s) = N^{1-s}/(1-s)$, at $\sigma_0 = 0.55, 0.6, 0.75, 1$, against the diagonal energy $\zeta(2\sigma_0)$ (dashed). (b) The Bohr–Landau bound [CE10, Prop. 3.9] for the fraction of the zeros with real part $> \sigma$, with σ_0 optimized, for $T = 1600$ (measured mean energy), 10^6 and 10^{12} . Reproduced from [LV].

The Bohr–Landau bound is proportional to T because the diagonal energy does not decrease. A mollifier removes it: with $M_X(s) = \sum_{d \leq X} \mu(d) d^{-s}$ the coefficients of $\zeta M_X - 1$ are $c_n = \sum_{d|n, d \leq X} \mu(d)$, which vanish for $1 < n \leq X$. The *emptied* diagonal gives, with $X = T^{2\sigma_0-1}$, $\int_T^{2T} |\zeta M_X - 1|^2(\sigma_0 + it) dt \ll T^{1-(2\sigma_0-1)^2+\varepsilon}$, and Littlewood's lemma applied to $1 - (\zeta M_X - 1)^2$ gives Carlson's theorem $N(\sigma; T, 2T) \ll T^{4\sigma(1-\sigma)+\varepsilon}$ [T, Sect. 9.16]. An explicit zero-density estimate built on the same scheme is due to Kadiri, Lumley and Ng [KLN]: Littlewood's lemma applied to $\zeta M_X(2 - \zeta M_X) = 1 - (\zeta M_X - 1)^2$, with an explicit bound for the mean square of the mollified function. The best known exponents are much smaller: $\frac{12}{5}(1-\sigma)$ near $\sigma = \frac{1}{2}$ after Ingham and Huxley, and $\frac{30}{13}(1-\sigma)$ after Guth and Maynard [GM]; the density hypothesis is $2(1-\sigma)$ (Figure 2).

Status: classical; the reading of the mollifier as emptying the diagonal energy is ours. The two figures of this section are numerical and are reproduced from [LV]; the scripts that generated them are not among the ancillary files.

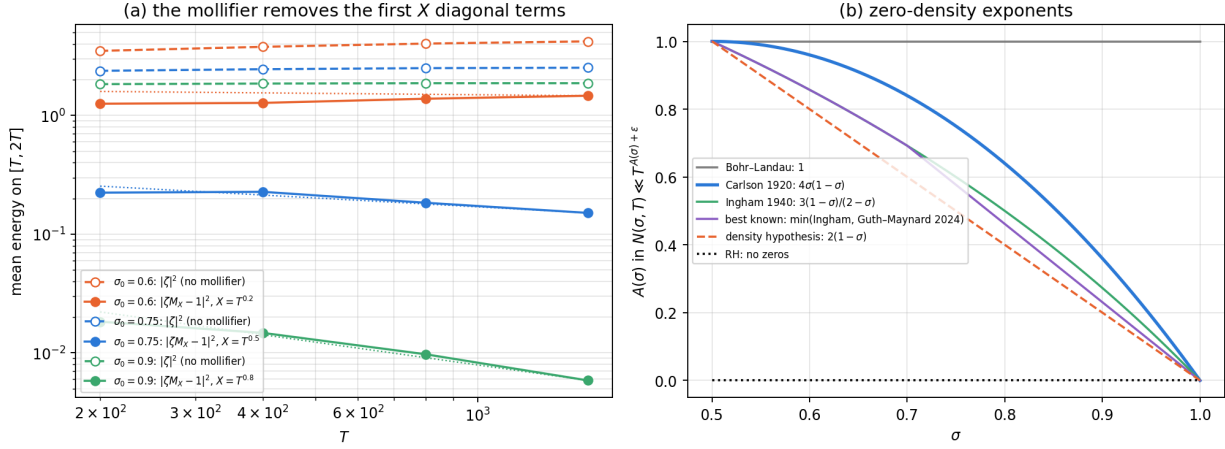


FIGURE 2. (a) Mean of $|\zeta|^2$ (no mollifier) and of $|\zeta M_X - 1|^2$ with $X = T^{2\sigma_0-1}$ on $[T, 2T]$, at $\sigma_0 = 0.6, 0.75, 0.9$; the mollifier removes the first X diagonal terms, and the dotted lines show the decay $T^{-(2\sigma_0-1)^2}$. (b) Zero-density exponents $A(\sigma)$ in $N(\sigma, T) \ll T^{A(\sigma)+\varepsilon}$: Bohr–Landau, Carlson, Ingham, the best known (Ingham, Guth–Maynard) and the density hypothesis $2(1 - \sigma)$. Reproduced from [LV].

2.1. Simultaneous peaks and the cross-energy matrix. The loss is the term $O(n)$ in (1): for $n > T$ the cross terms no longer cancel on average. Every zero with $\beta > \sigma$ forces the mollified polynomial to be large near γ , so counting zeros becomes counting the points where a short Dirichlet polynomial is simultaneously large. Let $D(t) = \sum_{N < n \leq 2N} b_n n^{-it}$, let $W = \{t_1, \dots, t_R\}$ be 1-separated, and let $v_r = (n^{-it_r})_{N < n \leq 2N}$. The *cross-energy matrix* of the peaks is

$$G_{rs} = \langle v_r, v_s \rangle = \sum_{N < n \leq 2N} n^{i(t_s - t_r)}, \quad G_{rr} = N, \quad \text{tr } G = RN.$$

Its diagonal is the diagonal energy and its off-diagonal entries are the cross-energies between peaks, as the energy matrix $M(U)$ of [CE5, Sect. 2] has the C_2 limits on its diagonal and the interference of the zeros off it.

Lemma 1 (duality). *For every b , $\sum_r |D(t_r)|^2 \leq \lambda_{\max}(G) \|b\|^2$. Hence if $|D(t_r)| \geq Y$ for all r , then $R \leq \lambda_{\max}(G) \|b\|^2 / Y^2$. Moreover $\lambda_{\max}(G) \leq (\text{tr } G^k)^{1/k}$ for every $k \geq 1$.*

Proof. With A the $R \times N$ matrix with rows v_r , $\sum_r |D(t_r)|^2 = \|Ab\|^2 \leq \lambda_{\max}(A^*A) \|b\|^2$, and $\lambda_{\max}(A^*A) = \lambda_{\max}(AA^*) = \lambda_{\max}(G)$. The last statement holds because G is positive semidefinite. $\square$

Status: elementary. Gershgorin's theorem with $\sum_{N < n \leq 2N} n^{iu} \ll N/|u| + T^{1/2} \log T$ for $1 \leq |u| \leq T$ gives $\lambda_{\max}(G) \ll (N + RT^{1/2}) \log T$, the Halász–Montgomery inequality [Mont]. Gershgorin uses only the size of each row; the traces $\text{tr } G^k$ see the coherence of the phases among many peaks, and the method of Guth and Maynard [GM] is a fine analysis of that structure.

Montgomery's large value conjecture [Mont], in the form of [Guth, Conj. 3.3]: if $|b_n| \leq 1$, $\sigma > \frac{1}{2}$, $N \leq T \leq N^{O(1)}$ and $|D(t)| > N^\sigma$ on a 1-separated set $W \subset [0, T]$, then $|W| \leq N^{2-2\sigma+o(1)}$. Montgomery showed that it implies the density hypothesis $N(\sigma, T) \ll T^{2(1-\sigma)+\varepsilon}$. The ℓ^2 form

$\lambda_{\max}(G) \ll T^\varepsilon(N+R)$ for arbitrary sets of peaks is false; Bourgain's counterexample [Bou] uses coefficients that are not flat.

Remark 2 (peaks of $\sum n^{-it}$). Take $T = N^2$, $R = N$, and for W either R random integers in $[N, T)$ or the R largest values of $|\sum_{N < n \leq 2N} n^{-it}|$ on the integers of $[N, T)$, chosen greedily with a minimal gap of 1 or 10. For random points $\lambda_{\max}(G)/(N+R)$ stays between 1.8 and 2.0 for $100 \leq N \leq 1600$; for peak points it grows, from 4.8 to 27 (gap 1) and from 2.5 to 17 (gap 10). At $N = R = 400$, gap 10, both matrices have trace RN , but the peak points separate an eigenvalue $13.7N$ from the rest of the spectrum, which lies below $6.5N$, while for random points all eigenvalues are below $3.9N$. The top eigenvector b of A^*A is not flat (the mean of $|b_n|/\max |b_n|$ is 0.18) and its largest coordinates sit on smooth numbers ($n = 630, 420, 720, 540, 504, 480, 560, 672$; the median largest prime factor of the 20 largest coordinates is 7, against 61 for all n in the range). This is the resonance mechanism of Soundararajan and of Bondarenko and Seip [Sou, BS], found here as the top eigenvector of the cross-energy matrix. It does not contradict the conjecture, which requires $|b_n| \leq 1$; Sect. 3 tests that bounded form directly. Status: numerical (peaks_matrix.py; Figure 3).

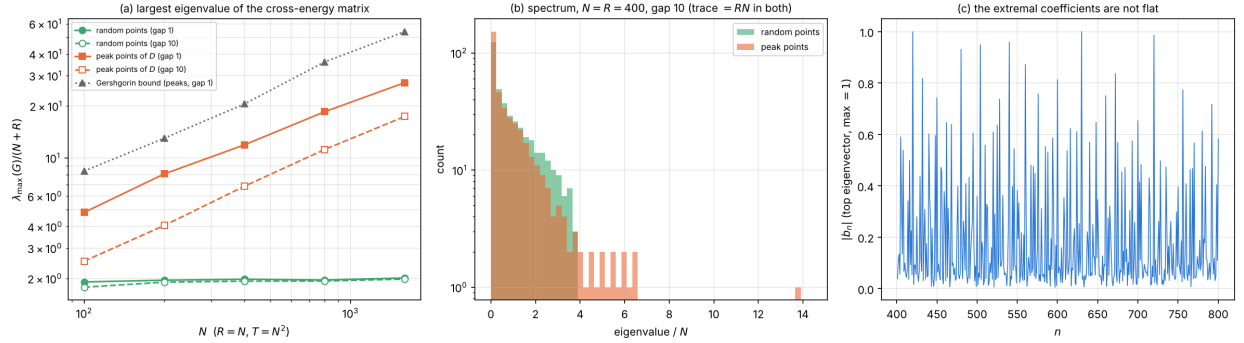


FIGURE 3. The cross-energy matrix of peaks, $T = N^2$, $R = N$. (a) $\lambda_{\max}(G)/(N+R)$ for random points and for peak points of $\sum n^{-it}$, with the Gershgorin bound. (b) Spectrum for $N = R = 400$, gap 10: same trace, but the peak points separate a large eigenvalue. (c) The top eigenvector is not flat and concentrates on smooth n .

Smooth, random and rough supports. The top eigenvector of Remark 2 lives on smooth numbers, and open problem 2 of Sect. 5 asks what happens on supports without small prime factors. A first experiment uses half-supports. Order the integers of $(N, 2N]$ by their largest prime factor $P(n)$, and take as support the $M = N/2$ integers with the smallest $P(n)$ (the smooth half), M random ones, or the M with the largest $P(n)$ (the rough half). Take $T = N^2$ and $R = M$ peaks on the integers of $[T, 2T)$ with minimal gap 10. Starting from $b = 1$ on the support, we alternate five times between taking the R largest well-separated values of $|D_b|$ and replacing b by the top eigenvector of A^*A . No step lowers $\lambda_{\max}$, so this is a local adversarial search. The values of $\lambda_{\max}(G)/(M+R)$ are:

N	100	200	400	800
smooth half	5.28	7.77	12.08	18.21
random half	3.39	5.18	7.59	11.53
rough half	3.38	4.45	5.64	7.48
R random points	1.70	1.80	1.96	1.90

Smooth numbers amplify the resonance. At $N = 800$ the smooth half gives 2.4 times the value of the rough half, and it grows faster: the fitted growth is $N^{0.60}$ or $(\log N)^{3.3}$ for the smooth half, against $N^{0.38}$ or $(\log N)^{2.1}$ for the rough half. Removing the smooth numbers does not stop the

growth, and for $N \leq 800$ a factor $(\log N)^{2.1}$, which is compatible with Montgomery’s conjecture, cannot be told from a power $N^{0.38}$. With fixed coefficients on all of $(N, 2N]$ and $R = N$ peaks of $|D_b|$ itself, the Rayleigh quotient $\sum_r |D_b(t_r)|^2 / (\|b\|^2(N + R))$ at $N = 100, 200, 400$ is 2.09, 2.79, 3.74 for $b = 1$; 2.24, 2.85, 3.96 for the mollifier coefficients $c_n = \sum_{d|n, d \leq \sqrt{N}} \mu(d)$; 2.04, 2.44, 2.84 for random signs; and 3.11, 4.54, 6.62 for the indicator of the smooth half. At these sizes the mollifier coefficients do not lower the energy at their own peaks, and random signs lower it most. Status: numerical (`smooth_rough.py`); the search is local, not exhaustive. The rows for the smooth and the rough half agree within 1.5% with the author’s working notes. There the random half, a different draw, gave 3.58, 4.15, 8.28, 9.81, and the smooth indicator, defined differently, gave 2.37, 5.24, 8.60.

3. GUTH’S CHECKS AND A SEARCH FOR WORST CASES

Guth [Guth] observes that there is no meaningful numerical evidence for Montgomery’s conjecture: exhaustive searches cost time exponential in N . He proposes quantities that can be computed in polynomial time [Guth, Sects. 6 and 7.1]. We compute two of them and add a direct search for worst cases with $|b_n| \leq 1$.

Setting. Let $D(t) = \sum_{n=N}^{2N} b_n e^{it\phi_n}$ with t on the integers of $[0, T)$. We compare the Dirichlet frequencies $\Phi_{\text{Dir}} = \{\log n\}$; Guth’s “almost counterexample” $\Phi_{\text{AC}} = \{\sqrt{n/N}\}$, which shares many of the properties of $\{\log n\}$ used in current proofs but for which the conjecture should fail; the affinely rescaled set Φ_{ACs} with the span $\log 2$ of Φ_{Dir} (*span-matched*; without it, the wider near-diagonal band of Φ_{AC} , whose span is $\sqrt{2} - 1$, inflates several quantities for trivial reasons); and $N + 1$ random points of $[0, \log 2]$. All values of D on the integer grid are computed exactly to rounding by a blocked matrix product; for $N = 8000$ and $T = N^2$ one evaluation covers $6.4 \cdot 10^7$ points.

3.1. Check 1: fluctuations of the difference-set density. Let $\rho_{1/T}(\xi)$ be the number of differences $\phi_m - \phi_n$, $m \neq n$, in bins of width $1/T$, let ρ_{smooth} be its moving average over 2001 bins and $\rho_{\text{fluc}} = \rho_{1/T} - \rho_{\text{smooth}}$. The statistic is $\int |\rho_{\text{fluc}}|$ over $0.02 < |\xi| < 0.9L$ (L the span), which excludes the trivial spike at 0 and the edges, divided by the same quantity for a Poisson process with mean ρ_{smooth} . Guth predicts that for Φ_{Dir} the fluctuation concentrates in sparse spikes at $\log(p/q)$, and that for Φ_{AC} it looks like noise.

$T = N^{3/2}$, $N =$	500	1000	2000	4000
Dirichlet	0.661	0.459	0.420	0.391
almost counterexample	0.919	0.620	0.678	0.747
almost counterexample, span-matched	0.822	0.711	0.764	0.816
random	0.992	0.994	1.002	0.989
spikes ($z > 8$), Dirichlet	12	44	140	376

TABLE 1. $\int |\rho_{\text{fluc}}|$ relative to Poisson noise; $z = \rho_{\text{fluc}}/\rho_{\text{smooth}}^{1/2}$, spikes counted on both sides of 0.

At $T = N^{3/2}$ the outcome matches Guth’s picture (Table 1, Figure 4a). For Φ_{Dir} the spikes sit at $\log(p/q)$ with small p, q , their number grows with N , and the remaining fluctuation decreases relative to Poisson noise; for Φ_{AC} the fluctuation moves towards Poisson noise from $N = 1000$ on, and the ratio is still increasing at $N = 4000$. Each spike has a fine structure not mentioned in [Guth]: since $|mq - np| \geq 1$ unless $m/n = p/q$, no difference lies within about $1/(2qN)$ of $\log(p/q)$ (for $p/q = \frac{3}{2}$, $|\xi - \log \frac{3}{2}| \geq 1/(4N)$), so every spike is surrounded by an *exclusion zone*, whose edges carry *satellite peaks* formed by the nearest Farey neighbours, $|mq - np| = 1, 2, \dots$ (Figure 4b). At $T = N^2$ the spike set $\{\log(p/q) : q \lesssim T/N\}$ is dense at scale $1/T$; the Dirichlet ratio stays at 0.912–0.915 for $300 \leq N \leq 2000$, and the description “smooth plus sparse spikes” no

longer applies. Status: numerical (`diffsets.py`, `fig_diffsets.py`). An earlier implementation [LV] gave the same values for $N \geq 2000$ and smaller ratios at $N = 500$ (where the smoothing window covers about a quarter of the range); the qualitative picture is the same.

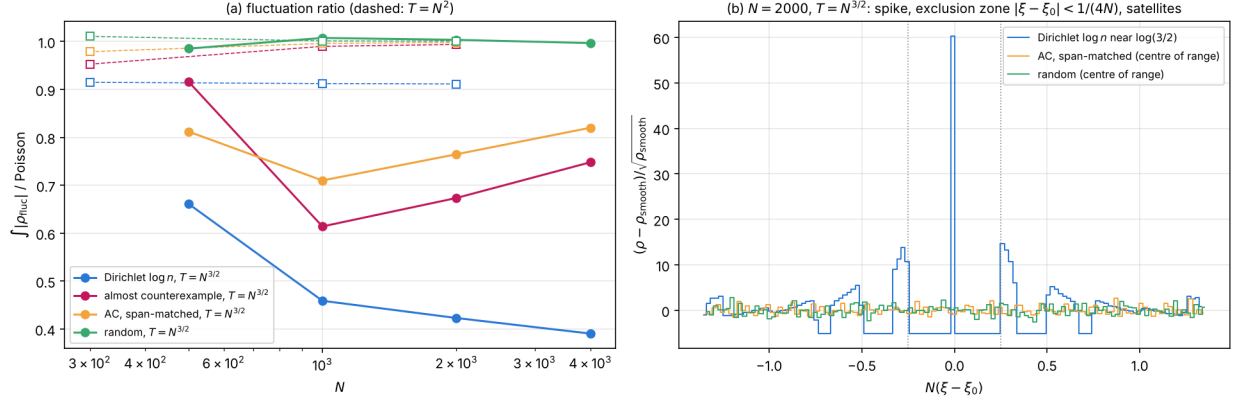


FIGURE 4. (a) Fluctuation ratio of Table 1; dashed: $T = N^2$. (b) Normalized fluctuation near $\xi_0 = \log \frac{3}{2}$ for $N = 2000$, $T = N^{3/2}$: a spike, its Farey exclusion zone (dotted lines at $\pm 1/(4N)$) and satellites for Dirichlet frequencies, against noise for the span-matched almost counterexample and for random frequencies.

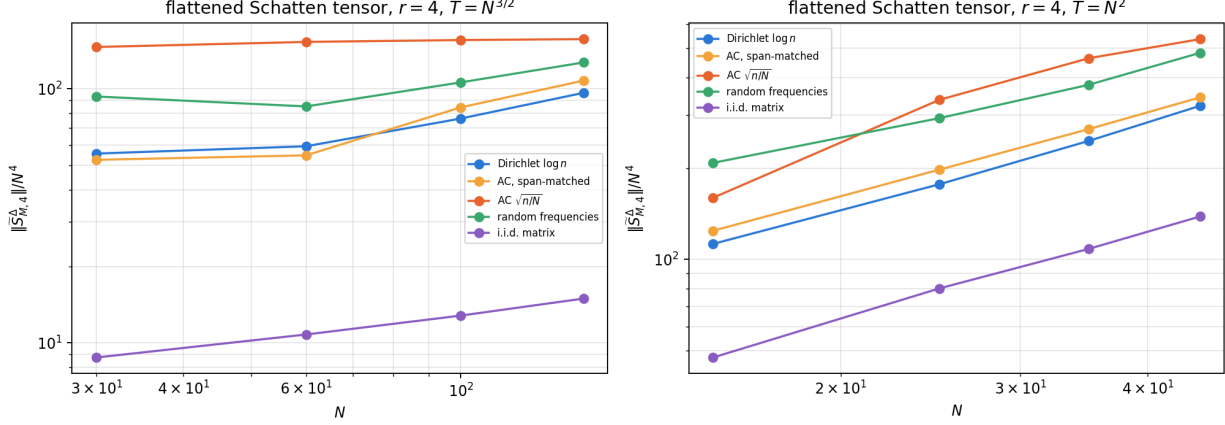
3.2. Check 2: flattened Schatten tensors. Guth also considers the multilinear forms

$$S_{M,r}(x_1, \dots, x_r) = \sum x_1(t_1) \cdots x_r(t_r) G_{t_1 t_2} \cdots G_{t_r t_1}$$

built from the cross-energy matrix $G_{tt'} = \sum_n e^{i(t-t')\phi_n}$ on the full grid, with the diagonal removed, and the operator norms of their flattenings. In [LV] we computed these norms by Lanczos iteration for $r \leq 4$ and $N \leq 150$. The raw almost counterexample has a much larger norm than the Dirichlet case, as Guth expects, but most of the gap comes from the narrower span: after span matching the Dirichlet norm is smaller by only about 10% at $N = 100$ – 150 for $T = N^{3/2}$. At these sizes the norms are dominated by the band $|t - t'| \lesssim 1/\text{span}$, where $|G_{tt'}| \approx N$ for every frequency set, and the flattenings do not yet show the expected separation (Table 2, Figure 5); for $T = N^2$ the gap after span matching is about 6% at $N = 45$. Finer decompositions of S or much larger N would be needed to see it. Status: numerical [LV]; the Lanczos products were checked against dense computations, and the scripts of this check are not among the ancillary files.

$T = N^{3/2}$, $N =$	Dirichlet	AC	AC (span-matched)	random freq.	i.i.d.
30	55.6	146.1	52.5	93.3	8.8
60	59.4	152.9	54.7	85.2	10.8
100	76.4	155.6	84.5	106.0	12.8
150	96.4	156.8	107.6	126.9	14.9

TABLE 2. Flattened Schatten norms $\|\tilde{S}_{M,4}^\Delta\|/N^4$ (flattening $12|34$, diagonal removed) for the Dirichlet frequencies, the almost counterexample $\{\sqrt{n/N}\}$ (AC), its span-matched rescaling, random frequencies and i.i.d. phases [LV].

FIGURE 5. Flattened Schatten norms, $r = 4$, for the frequency sets of Table 2 [LV].

3.3. Check 3: worst cases with $|b_n| \leq 1$. Search. For a target number R of points we alternate two steps: W is set to the R largest values of $|D_b|$, and $b \leftarrow \text{sgn}(A_W^* D_b|_W)$, where A_W is the matrix of phases on W and $\text{sgn}(z) = z/|z|$. The second step maximizes the linearization of $\sum_{t \in W} |D_b(t)|^2$ over $|b_n| \leq 1$, so the objective never decreases; by Lemma 1 this is a power iteration on the cross-energy matrix restricted to the unit polydisc. We use $R = N$, 20 iterations from $b = 1$, and report the number of 10-separated t with $|D(t)| > N^\sigma$; this count is insensitive to the width of the peaks, since nearly every cluster contains exactly one 10-separated point. We also evaluate, without search, $b_n = 1$, a random completely multiplicative unimodular f ($f(p)$ independent and uniform on the circle, the Steinhaus model), and independent random phases.

$T = N^2, \sigma = 0.7, N =$	250	500	1000	2000	3000	4000	8000
Dirichlet, search	37	72	155	314	508	696	1488
span-matched AC, search	28	53	80	71	67	45	13
random frequencies, search	31	58	84	79	59	56	18
Dirichlet, $b_n = 1$	–	5	14	21	–	37	65
Dirichlet, random multiplicative f	51	110	233	485	–	1003	2023
Dirichlet, random phases	3	1	0	0	–	0	0
$N^{2-2\sigma}$	28	42	63	96	122	145	220

TABLE 3. Number of 10-separated $t \in [0, N^2)$ with $|D(t)| > (N+1)^{0.7}$; the row $N^{2-2\sigma}$ is $(N+1)^{0.6}$ rounded. For random multiplicative f the entries are means over 64 realizations for $N \leq 2000$, 32 for $N = 4000$ and 24 for $N = 8000$ (Table 4); earlier runs of [LV], with 8 realizations for $N \leq 2000$ and 2 for $N \geq 4000$, gave 53, 111, 231, 480, 975, 2050.

Results. At $T = N^{3/2}$ the three frequency sets behave alike: for $R = N$ the ratio $R/N^{2-2\sigma_{\text{eff}}}$, where $N^{\sigma_{\text{eff}}} = \min_W |D|$, stays roughly constant for Dirichlet frequencies (4.6, 6.1, 5.8, 6.1 for $N = 500, \dots, 4000$). At $T = N^2$ Dirichlet frequencies admit markedly more large values than the other sets (Table 3, Figure 6a). At $\sigma = 0.7$ the ratio of the Dirichlet count to $N^{2-2\sigma}$ grows from 1.3 at $N = 250$ to 4.8 at $N = 4000$ and 6.8 at $N = 8000$; its local logarithmic slope stays between 0.36 and 0.59 (0.50 between 4000 and 8000). At $\sigma = 0.75$ the ratio stays bounded, around 2 (151 points against $N^{1/2} = 89$ at $N = 8000$). For the span-matched almost counterexample and for

random frequencies the search finds fewer points than $N^{2-2\sigma}$ from $N = 2000$ on, and only 13 and 18 at $N = 8000$.

Structure of the search output. The coefficients found by the search are close to completely multiplicative. For $N = 1000$ the averages $|\text{mean}_k b_{kp} \overline{b_{kq}}|$ lie between 0.92 and 0.98 for every ratio p/q tested, against 0.09 for random pairs, and the cocycle relations hold to within 0.04 rad (Figure 6c). Moreover, at $\sigma = 0.7$, a *random* completely multiplicative f , with no optimization, produces even more large values than the search, while random phases without multiplicative structure produce essentially none. At $\sigma = 0.75$ the order is reversed: the search finds 62 points at $N = 1000$ and 151 at $N = 8000$, against means 21.0 and 52.9 (and 95% quantiles 25 and 60) for the random multiplicative model (Table 4). Status: numerical (`core.py`, `big8k.py`, `mult8k.py`; earlier runs `clusbig.py`, `lvs.py`, `mult.py`). Rerunning `big8k.py` at $N = 1000$ and `mult8k.py` at $N = 500$ and 1000 reproduces the corresponding entries of Table 3, and the entries for the random multiplicative model are from `mult_stats.py` (Sect. 4.2); the other entries (in particular $N = 250$, $N = 3000$ and $N \geq 2000$), the ratios at $T = N^{3/2}$ and the multiplicativity statistics of Figure 6c are from the runs of [LV].

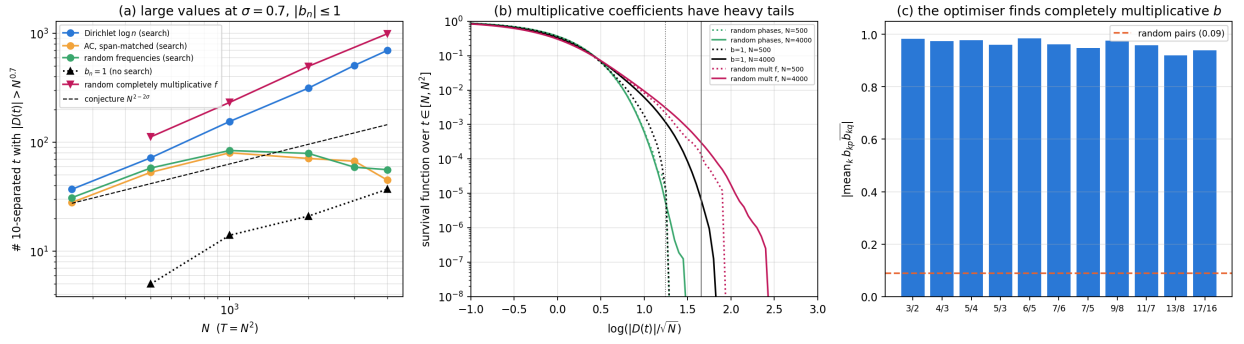


FIGURE 6. (a) Counts of Table 3 for $N \leq 4000$, from the runs of [LV] (for the random multiplicative model these are the earlier means, not those of Table 4). (b) Survival function of $\log(|D(t)|/\sqrt{N})$ over $t \in [N, N^2]$; the threshold $(\sigma - \frac{1}{2}) \log N$ for $\sigma = 0.7$ is marked. (c) Multiplicativity of the output of the search, $N = 1000$.

4. CALIBRATION: RANDOM MULTIPLICATIVE FUNCTIONS

Why should multiplicative coefficients produce many large values, and is the excess over $N^{2-2\sigma}$ a counterexample in the making? The theory of random multiplicative functions suggests answers to both questions; we separate what is proved from what is heuristic.

Proposition 3 (the diagonal of the fourth moment is the multiplicative energy). *Let f be completely multiplicative with $|f(n)| = 1$, and $D_f(t) = \sum_{N \leq n_i \leq 2N} f(n_i) n_i^{it}$. Then*

$$|D_f(t)|^4 = \sum_m \left| \sum_{\substack{n_1 n_2 = m \\ N \leq n_i \leq 2N}} 1 \right|^2 + \sum_{\substack{n_1 n_2 \neq n_3 n_4 \\ N \leq n_i \leq 2N}} f(n_1 n_2) \overline{f(n_3 n_4)} \left(\frac{n_1 n_2}{n_3 n_4} \right)^{it},$$

so the diagonal part of the fourth moment is the multiplicative energy $\mathcal{M}(N) = \#\{n_1 n_2 = n_3 n_4, N \leq n_i \leq 2N\}$, the same for every such f and for $f = 1$. If f is a Steinhaus random multiplicative function, then $\mathbb{E}_f |D_f(t)|^4 = \mathcal{M}(N)$ for every t .

Proof. Expand $|D_f|^4 = D_f^2 \overline{D_f}^2$ and use $f(n_1) f(n_2) = f(n_1 n_2)$ and $|f| = 1$. For the Steinhaus model, $\mathbb{E}_f f(m) \overline{f(m')} = 0$ for $m \neq m'$, so the off-diagonal terms have expectation 0. $\square$

Status: proved (elementary). For independent random phases the expected fourth moment is instead $2(N+1)^2 - (N+1)$. Throughout, the polynomial has $N+1$ terms, and thresholds and normalizers are computed with $N+1$ in place of N (so $N^{2-2\sigma}$ stands for $(N+1)^{2-2\sigma}$). At $T = N^2$ the off-diagonal terms are numerically negligible in the mean over t :

N	250	500	1000	2000	4000	8000
$\mathcal{M}(N)/(N+1)^2$ (exact count)	2.895	3.075	3.266	3.458	3.647	3.837
mean of $ D_f ^4/(N+1)^2$, random multiplicative f	2.88	3.07	3.26	3.46	3.65	3.84
same, random phases	1.98	1.99	1.99	2.00	2.00	2.00

The multiplicative energy grows like $\log N$ (about $1.39 + 0.27 \log N$ in this range), whereas a Gaussian model gives the constant 2. Status: numerical (`mult_energy.py` for the exact counts; `calib.py` for the moments, runs of [LV]).

4.1. Harper's high moments: a heuristic for the t -process. *What Harper's theorem says, and what it does not.* For a Steinhaus random multiplicative function f (the values $f(p)$ independent and uniform on the unit circle) and $1 \leq q \leq c \log x / \log \log x$, Harper [Hp2] proved

$$(2) \quad \mathbb{E}_f \left| \sum_{n \leq x} f(n) \right|^{2q} = e^{-q^2 \log q - q^2 \log \log(2q) + O(q^2)} x^q (\log x)^{(q-1)^2}.$$

This is an average over the random function f at a fixed length x . Our counts are averages over $t \in [0, T)$ for one fixed realization of f ,

$$\frac{1}{T} \sum_{0 \leq t < T} \left| \sum_{N \leq n \leq 2N} f(n) n^{it} \right|^{2q},$$

and the two are not automatically equivalent. One fact connects them. For each fixed t the function $n \mapsto f(n) n^{it}$ is again a Steinhaus random multiplicative function, because $f(p) p^{it}$ is uniform on the circle and independent over p ; hence, by Fubini, the expectation over f of the t -average equals $\mathbb{E}_f |D_f(0)|^{2q}$ exactly, and the expected number of integers $t \in [0, T)$ with $|D_f(t)| > N^\sigma$ (all of them, not the 10-separated ones that we count) is $T \mathbb{P}_f(|D_f(0)| > N^\sigma)$. What is not known is that the t -average of a single realization concentrates near this expectation; moreover Harper's theorem concerns $\sum_{n \leq x}$, not the dyadic range $[N, 2N]$, for which the asymptotic of $\mathbb{E}_f |D_f(0)|^{2q}$ is not proved. So what follows is a heuristic, not a consequence of [Hp2].

The log-normal regime. The factor $(\log x)^{(q-1)^2}$ in (2) is the signature of tails of log-normal type: for large q it grows like $e^{q^2 \log \log x}$, the $2q$ -th moment of e^Y with Y Gaussian of variance $\frac{1}{2} \log \log x$, the variance $\frac{1}{2} \sum_{p \leq x} 1/p$ of the logarithm of a short random Euler product. (For $q < 1$ there is better than square-root cancellation [Hp]; here $q > 1$.) Proposition 3 gives the dyadic analogue of the case $q = 2$ of (2) for the expectation over f , and the table above confirms numerically that the t -average of one realization is close to it.

The predicted count. Take (2) as a model for the moments of D_f over $t \in [0, T)$, and write $L = \log N$, $\ell = \log \log N$. By Markov's inequality the number of t with $|D_f(t)| > N^\sigma$ is at most $T N^{q(1-2\sigma)} L^{(q-1)^2}$ times the factor $e^{-q^2 \log q - q^2 \log \log(2q) + O(q^2)}$, whose $O(q^2)$ part is not explicit. Ignoring that factor, with $T = N^2$ the logarithm of the bound divided by $N^{2-2\sigma}$ is $L - (2\sigma - 1)(q - 1)L + (q - 1)^2 \ell$, which is minimal at $q - 1 = (\sigma - \frac{1}{2})L/\ell$. This gives the following.

Heuristic prediction, not a consequence of Harper's theorem for the sampled t -process:

$$(3) \quad \log \frac{\#\{t : |D_f(t)| > N^\sigma\}}{N^{2-2\sigma}} \lesssim L - \frac{(\sigma - \frac{1}{2})^2 L^2}{\ell}.$$

Equivalently, the Legendre transform of the moments (2) without the factor $e^{O(q^2 \log q)}$ corresponds to the tail $\mathbb{P}(\log |D_f/\sqrt{N}| > y) \approx e^{-2y - y^2/\ell}$, and at $y = (\sigma - \frac{1}{2})L$ with $T = N^2$ this gives exactly the right side of (3); this tail is not that of a centred Gaussian law (it corresponds to a Gaussian

with mean $-\ell$, variance $\ell/2$ and total mass e^ℓ), and (3) is a Markov upper bound, not a prediction of the count. The linear term dominates as long as $(\sigma - \frac{1}{2})^2 L/\ell < 1$, that is, for $\log N$ up to roughly $\ell/(\sigma - \frac{1}{2})^2$, which is of order 10^2 for $\sigma = 0.7$. The dropped factor shifts this crossover by an amount we cannot quantify, since its $O(q^2)$ term is of the same order as the terms kept, so we use (3) only qualitatively. Asymptotically L^2/ℓ beats L and the model is consistent with Montgomery's conjecture; at computable sizes, where the tail is governed by moments of order $q \approx 1 + (\sigma - \frac{1}{2})L/\ell$ (between 1.65 and 1.82 for $250 \leq N \leq 8000$ at $\sigma = 0.7$), it allows an excess over $N^{2-2\sigma}$ that grows like a power of N with exponent below 1. Status: heuristic; it uses a proved theorem about a related random object, outside the range in which that theorem applies.

4.2. Statistics of the random multiplicative model. In an earlier version of these experiments the largest sizes rested on two realizations. We repeated the count for the Steinhaus model with $R = 64$ independent realizations for $N \leq 2000$, $R = 32$ for $N = 4000$ and $R = 24$ for $N = 8000$ (seeds $1000, \dots, 999 + R$; at $N = 8000$ each realization evaluates D_f at $6.4 \cdot 10^7$ points). Table 4 gives, for the number of 10-separated $t \in [0, N^2)$ with $|D_f(t)| > (N+1)^\sigma$, the mean, median and standard deviation over the realizations, a 95% confidence interval for the mean (Student t), the empirical 5% and 95% quantiles, and the means over the two halves of the seeds.

N	R	mean	median	s.d.	95% CI of the mean	$q_{0.05}$	$q_{0.95}$	halves
$\sigma = 0.7$								
250	64	51.0	51.0	4.1	[50.0, 52.1]	44.3	57.0	51.3 / 50.8
500	64	109.7	109.0	7.3	[107.9, 111.6]	100.0	120.0	111.5 / 108.0
1000	64	232.8	234.0	11.1	[230.1, 235.6]	213.2	249.8	234.7 / 231.0
2000	64	484.7	484.0	14.2	[481.2, 488.3]	462.3	506.0	486.8 / 482.7
4000	32	1003.0	1006.0	22.5	[994.8, 1011.1]	961.2	1034.7	1008.1 / 997.8
8000	24	2022.7	2018.0	27.6	[2011.1, 2034.4]	1987.3	2061.4	2024.2 / 2021.2
$\sigma = 0.75$								
250	64	10.1	10.0	2.2	[9.5, 10.7]	7.0	14.0	10.6 / 9.6
500	64	14.5	15.0	2.4	[13.9, 15.1]	11.0	18.0	14.4 / 14.5
1000	64	21.0	21.5	3.4	[20.1, 21.8]	15.0	25.0	20.7 / 21.2
2000	64	30.4	30.5	4.3	[29.3, 31.5]	24.0	37.0	29.6 / 31.2
4000	32	40.6	41.0	6.3	[38.3, 42.8]	30.6	50.5	40.1 / 41.1
8000	24	52.9	53.0	5.6	[50.5, 55.2]	43.6	60.0	52.2 / 53.6

TABLE 4. Number of 10-separated $t \in [0, N^2)$ with $|D_f(t)| > (N+1)^{0.7}$ for Steinhaus random completely multiplicative f , over R realizations; $N^{2-2\sigma} = 28, 42, 63, 96, 145, 220$. The last column gives the means over the first and the second half of the seeds. Lower block: the same at $\sigma = 0.75$, against $N^{1/2} = 16, 22, 32, 45, 63, 89$.

Comparison with the data. At $\sigma = 0.7$ the mean ratios to $N^{2-2\sigma}$ are 1.85, 2.63, 3.69, 5.07, 6.92, 9.21 for $N = 250, 500, 1000, 2000, 4000, 8000$. The relative standard deviation across realizations decreases from 8% at $N = 250$ to 1.4% at $N = 8000$, the two halves of the seeds agree within their sampling error ($|z| \leq 2$, the largest $z = 1.97$ at $N = 500$), and the confidence intervals are narrow compared with the growth, so the excess is a property of the model and not of particular seeds. A least-squares fit of the logarithm of the mean ratio against $\log N$ gives the exponent 0.46 ([0.458, 0.468] by a bootstrap over realizations; realization i uses the same seed at every N , so the values $f(p)$ for $p \leq 2N$ are shared across sizes and the counts are mildly correlated across N , by 0.16–0.33, which the bootstrap ignores), with local exponents 0.51, 0.49, 0.46, 0.45, 0.41 between consecutive sizes; the local exponents decrease slowly with N . The exponent is positive and below 1, as the heuristic (3) allows at these sizes; the agreement is only qualitative, since the local slope

of the right side of (3) is about 0.8 in this range, against measured 0.41–0.51, consistently with (3) being an upper bound. It is a pre-asymptotic description of $250 \leq N \leq 8000$, not an estimate of a limiting exponent: by (3) the excess should eventually stop growing, at sizes far beyond computation. At $\sigma = 0.75$ the heuristic fails: the right side of (3) has local slope about 0.6, but the counts stay below $N^{1/2}$, with mean ratios 0.64, 0.65, 0.66, 0.68, 0.64, 0.59 and no clear trend (least-squares exponent -0.015 , bootstrap interval $[-0.031, 0.001]$). The threshold $(N+1)^{0.75}$ lies at $\log |D_f/\sqrt{N}| \approx L/4$, half way to the trivial maximum $L/2$, where the tail of the finite sum is cut off and the moment model, which ignores the factor $e^{-q^2 \log q}$ (with $q \approx 2$ here), is least reliable. The bulk of the distribution barely moves: the variance of $\log |D_f/\sqrt{N}|$ is 0.47 at $N = 250$ and 0.49 at $N = 8000$, against $\pi^2/24 = 0.411$ for a complex Gaussian (random phases give 0.40–0.41); the excess lives in the tail, which becomes visibly heavier with N (Figure 8c). Status: numerical (`mult_stats.py` for Table 4 and Figure 7; `calib.py`, `lognorm.py`, `figcal.py` and runs of [LV] for Figure 8).

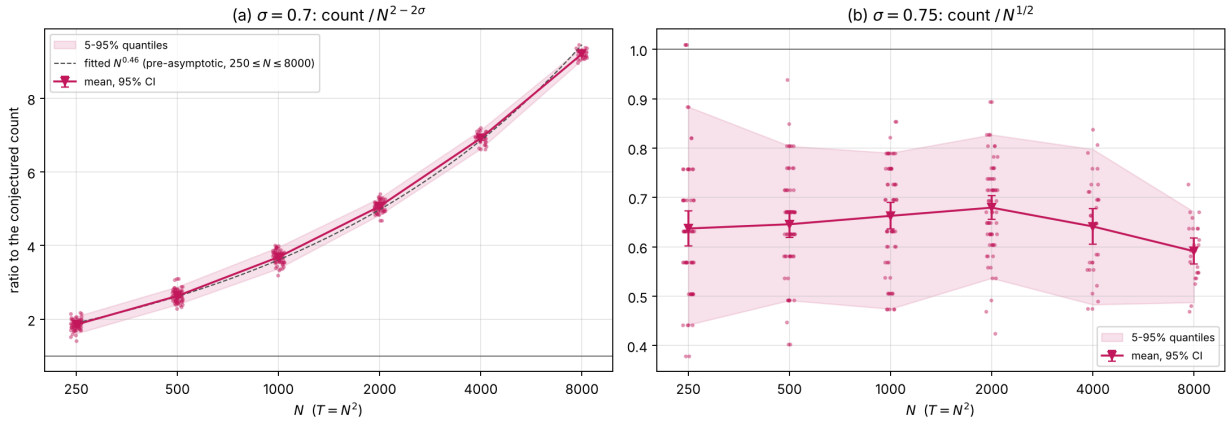


FIGURE 7. Counts of 10-separated $t \in [0, N^2)$ with $|D_f(t)| > (N+1)^\sigma$ for Steinhaus random multiplicative f , divided by $N^{2-2\sigma}$. (a) $\sigma = 0.7$: individual realizations (dots), mean with 95% confidence interval, and the band between the 5% and 95% quantiles; dashed: the fitted power. The fitted exponent describes $250 \leq N \leq 8000$ only and is pre-asymptotic. (b) $\sigma = 0.75$, against $N^{1/2}$.

Remark 4 (calibration). At $T = N^2$, the coefficients $|b_n| \leq 1$ produced by a heuristic ascent for large values of $\sum_{N \leq n \leq 2N} b_n n^{it}$ are, numerically, close to completely multiplicative, and Steinhaus random completely multiplicative coefficients exceed $N^{2-2\sigma}$ at $\sigma = 0.7$ by a mean factor that grows from 1.85 to 9.2 over $250 \leq N \leq 8000$, with small dispersion across realizations. Harper’s moment asymptotics explain this qualitatively, through the heuristic (3), by a log-normal regime that is consistent with Montgomery’s conjecture but dominates until $\log N$ is of order 10^2 . Numerical tests of the conjecture must therefore be measured against the random multiplicative model, not against $N^{2-2\sigma}$. At $\sigma = 0.75$ the comparison goes the other way: the search exceeds the model, and the model does not exceed $N^{1/2}$. Status: numerical, with a heuristic model; for one realization we have not proved that the t -average of $|D_f|^{2q}$ over $[0, N^2)$ follows (2) for any q , only that the diagonal part of the case $q = 2$ is $\mathcal{M}(N)$.

Related work and limitations. The maxima of randomized models of ζ on short intervals are governed by branching random walks [ABH], and the Gaussian behaviour of $\log |\zeta|$ extends to large deviations of size $\asymp \log \log T$ [Rad]. These works do not address Guth’s checks, and we have not found earlier numerical tests of Montgomery’s conjecture. The sizes here are modest ($N \leq 8000$).

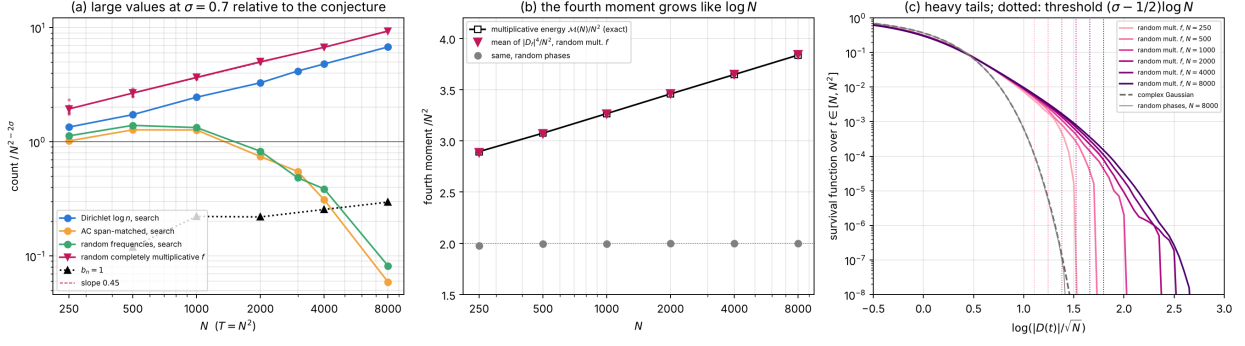


FIGURE 8. Calibration against random multiplicative functions, $T = N^2$, from the runs of [LV] (two realizations at $N = 4000$ and $N = 8000$; Table 4 and Figure 7 replace panel (a) for the random model). (a) Number of 10-separated t with $|D(t)| > N^{0.7}$, divided by $N^{0.6}$: searches for three frequency sets, $b_n = 1$, and random completely multiplicative f (small dots: individual realizations); the slope 0.45 is pre-asymptotic. (b) The fourth moment of D_f equals the multiplicative energy of $[N, 2N]$ and grows like $\log N$; random phases give 2. (c) Survival function of $\log(|D_f|/\sqrt{N})$ for random multiplicative f , $250 \leq N \leq 8000$, against the complex Gaussian law; dotted: the thresholds $(\sigma - \frac{1}{2})\log N$ for $\sigma = 0.7$.

for the large value counts); the search gives lower bounds for the worst case, not certificates; and the calibration transfers a theorem about averages over f to averages over t for one f , which Fubini's theorem reduces, for the expectation, to the dyadic moments $\mathbb{E}_f |D_f(0)|^{2q}$ (themselves known only for $q = 2$), and which the agreement of the fourth moment and the small dispersion across realizations support for single realizations, but do not prove.

5. DISCUSSION AND OPEN PROBLEMS

5.1. Two faces of the Euler product. The series follows one quantity, the energy of partial sums with the diagonal removed. Integrated over $\log x$, the energy of the primes is extraordinarily rigid: under RH, simple zeros and (H^*) its fluctuations around the line $C_B \log X + K$ have standard deviation 0.0035 (under LI), and in the random models of the primes and of μ examined in [CE6] the median energy is 8 to 150 times larger. The rigidity comes from the discrete spectrum of the zeros and from the weights $1/(\gamma - \gamma')$ that damp high frequencies. Evaluated on short ranges $N \leq n \leq 2N$, the same multiplicative structure produces the opposite effect: the moments of $\sum f(n)n^{it}$ grow like $(\log N)^{(q-1)^2}$, the tails are heavier than Gaussian, and, numerically at $\sigma = 0.7$, the number of large values exceeds the conjectured count by a factor that keeps growing at every computable size. The Euler product protects the energy and amplifies the peaks. An energy argument for counting zeros off the line has to treat this multiplicative structure explicitly; the mollified coefficients $c_n = \sum_{d|n, d \leq X} \mu(d)$, which vanish on $1 < n \leq X$, are the natural class in which to test it (open problem 2).

5.2. What a numerical test of Montgomery's conjecture should look like. A numerical exponent fitted to counts of large values at $N \leq 10^4$ is not evidence for or against the conjecture: the random multiplicative model, which is consistent with it, produces growing excesses at these sizes. A meaningful test compares the counts for a given frequency set and coefficient class with the same counts for the random multiplicative model, with its sampling distribution. In that comparison the coefficients found by our search are *not* more extreme than random multiplicative functions at $\sigma = 0.7$: they give fewer large values, below the 5% quantile of the model at every

N tested (Tables 3 and 4). At $\sigma = 0.75$ they are: the search gives about three times the model's mean, above its 95% quantile, and about twice $N^{1/2}$. This is the more interesting comparison, and it calls for longer searches and larger N at $\sigma \geq 0.75$. The same caution applies to every numerical tendency measured at sizes where $\log \log N$ still matters.

5.3. What is not achieved. Nothing in this paper bears on the truth of RH or of Montgomery's conjecture. Sect. 2 is classical apart from its formulation and the experiments. Sects. 3–4 give numerical evidence and a heuristic calibration, not theorems: the search is local and gives lower bounds for the worst case, the sizes are modest, and the use of Harper's theorem is heuristic.

5.4. Open problems.

- (1) Prove that the t -moments of $\sum_{N \leq n \leq 2N} f(n)n^{it}$ over $[0, N^2)$ follow Harper's asymptotics for $1 < q \leq 2$ for almost every realization f , and turn (3) into a statement about counts.
- (2) Repeat Check 3 with coefficients supported on integers without small prime factors, the class relevant for mollified polynomials; Sect. 2 suggests that the cross-energy $\text{tr } G^k$ of peaks is much smaller there. The half-supports of Sect. 2 give a first answer for $\lambda_{\max}$: on the rough half the worst case is about 2.4 times smaller at $N = 800$, but it still grows.
- (3) Extend the statistics of Table 4 to $N = 1.6 \cdot 10^4$ and $3.2 \cdot 10^4$, where each realization costs about 8 and 64 times more (the blocked evaluation costs about N^3), to see whether the local exponent continues to decrease.

DATA AND CODE

The ancillary Python scripts (NumPy, SciPy, Matplotlib) are in the folder `anc/largevalues`, whose file `README.txt` lists, for each script, the section, table or figure it reproduces and its output: `core.py` and `helpers.py` (evaluation of D on the integer grid, frequency sets, counts), `peaks_matrix.py` and `smooth_rough.py` (Sect. 2), `diffsets.py` and `fig_diffsets.py` (Check 1), `big8k.py` and `mult8k.py` (Check 3), `mult_energy.py`, `calib.py`, `lognorm.py`, `figcal.py`, `mult_stats.py` and `fig_stats.py` (Sect. 4). The computations are in double precision; D is evaluated exactly to rounding by blocked matrix products.

Provenance. This paper contains the material on large values (Sects. 7–9) of an earlier version of Part 6, titled “Fluctuations and large values of the prime energy”, and, together with [CE6], supersedes Parts II–IV of the preprint [LV]. The statistics of Sect. 4.2 are new; Check 2 and part of Check 3 (as stated there) are taken from [LV].

Use of AI tools. The computations and drafting of this paper, and of the preprint [LV] that it absorbs, were carried out with the assistance of an AI system (Claude, Anthropic), under the author's direction. The author checked the results and takes responsibility for them.

REFERENCES

- [ABH] L.-P. Arguin, D. Belius, A. J. Harper, Maxima of a randomized Riemann zeta function, and branching random walks, *Ann. Appl. Probab.* 27 (2017), 178–215.
- [Bou] J. Bourgain, Remarks on Montgomery's conjectures on Dirichlet sums, in *Geometric Aspects of Functional Analysis (1989–90)*, Lecture Notes in Math. 1469, Springer, 1991, 153–165.
- [BS] A. Bondarenko, K. Seip, Extreme values of the Riemann zeta function and its argument, *Math. Ann.* 372 (2018), 999–1015.
- [GM] L. Guth, J. Maynard, New large value estimates for Dirichlet polynomials, *Ann. of Math.* 203 (2026), no. 2; arXiv:2405.20552.
- [Guth] L. Guth, Large value estimates in number theory, harmonic analysis, and computer science, preprint, arXiv:2503.07410, 2025.
- [Hp] A. J. Harper, Moments of random multiplicative functions, I: low moments, better than squareroot cancellation, and critical multiplicative chaos, *Forum Math. Pi* 8 (2020), e1.
- [Hp2] A. J. Harper, Moments of random multiplicative functions, II: high moments, *Algebra Number Theory* 13 (2019), 2277–2321.

- [KLN] H. Kadiri, A. Lumley, N. Ng, Explicit zero density for the Riemann zeta function, *J. Math. Anal. Appl.* 465 (2018), no. 1, 22–46.
- [LV] P. Caceres, The cross-energy of Dirichlet partial sums: from the energy law of the primes to large values of Dirichlet polynomials, preprint, 2026.
- [Mont] H. L. Montgomery, *Topics in Multiplicative Number Theory*, Lecture Notes in Math. 227, Springer, 1971.
- [MV74] H. L. Montgomery, R. C. Vaughan, Hilbert’s inequality, *J. London Math. Soc.* (2) 8 (1974), 73–82.
- [Rad] M. Radziwiłł, Large deviations in Selberg’s central limit theorem, preprint, arXiv:1108.5092, 2011.
- [Sou] K. Soundararajan, Extreme values of zeta and L -functions, *Math. Ann.* 342 (2008), 467–486.
- [T] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., revised by D. R. Heath-Brown, Oxford University Press, 1986.
- [CE1] P. Caceres, Cross-energy of Dirichlet partial sums. Part 1: The C-transform and the cross-energy, preprint, 2026.
- [CE5] P. Caceres, Cross-energy of Dirichlet partial sums. Part 5: The energy of the primes, preprint, 2026.
- [CE6] P. Caceres, Cross-energy of Dirichlet partial sums. Part 6: Second-order asymptotics and limiting fluctuations of the prime energy, preprint, 2026.
- [CE10] P. Caceres, Cross-energy of Dirichlet partial sums. Part 10: Jensen’s formula, the energy slope and Weil’s functional, preprint, 2026.

UNIVERSIDAD EUROPEA DE VALENCIA, VALENCIA, SPAIN

Email address: `pedrojesus.caceres@universidadeuropea.es`