Article

On a Conditional Approach to the Twin Prime Conjecture via Sifted Composite Gaps

Ahmet Furkan Gocgen [*]

Article

On a Conditional Approach to the Twin Prime Conjecture via Sifted Composite Gaps

Ahmet Furkan Gocgen

Department of Mathematics, Ordu University; 24130100028@ogrenci.odu.edu.tr

Abstract

We develop a conditional framework that links the statistical behavior of gaps among sifted odd composites to the infinitude of twin primes. Central to our approach is Conjecture 1 (Uniform Gap Sparsity), which asserts that short adjacent gaps (≤ 8) in the sifted composite sequence occur only with asymptotic zero density as the sieve level grows. Assuming Conjecture 1 and classical distributional results such as the Bombieri–Vinogradov theorem, a Selberg–GPY type sieve produces $\gg X/(\log X)^2$ twin prime pairs up to X. The logical structure of the argument is complete, but several components are presented in sketch form—notably the short-interval Selberg bound and the bilinear-form estimates—so as to highlight the conditional reduction rather than obscure it with technical detail. A forthcoming companion work is envisioned to provide fully rigorous expansions of these arguments. In addition, we emphasize that the reduction is modular: weakened or averaged forms of Conjecture 1 could already yield nontrivial results on bounded prime gaps, while stronger bilinear estimates would sharpen quantitative bounds. Thus, even if Conjecture 1 in its full uniformity is too strong, natural weakened variants may still suffice to establish conditional progress toward the twin prime conjecture.

Keywords: twin prime conjecture; sieve theory; Bombieri–Vinogradov theorem; GPY method; Selberg sieve; sifted composites; prime gaps; bilinear forms; conditional results

1. Introduction

The twin prime conjecture (TPC) asserts that there exist infinitely many primes p such that $p + 2$ is also prime. Tracing back to Euclid's proof of the infinitude of primes, this conjecture has remained one of the most famous open questions in number theory. While much progress has been achieved on the distribution of prime numbers, particularly in relation to gaps between consecutive primes, a complete proof of TPC continues to elude mathematicians.

1.1. Historical and Methodological Background

The study of prime gaps has seen a renaissance in recent decades, following a lineage of break-throughs in sieve theory and the distribution of primes in arithmetic progressions. The pioneering work of Goldston, Pintz, and Yıldırım (GPY) introduced a flexible framework showing that if primes exhibit distributional levels beyond what is currently proven (notably Elliott–Halberstam type conditions [1]), then infinitely many prime pairs at bounded gaps can be detected [2]. Their method, relying on carefully optimized Selberg sieve weights, opened a path to bounded gaps but still fell short of producing twin primes unconditionally.

Subsequently, Zhang's breakthrough in 2014 established the first finite bound on prime gaps, proving that $\liminf(p_{n+1} - p_n) < 7 \cdot 10^7$ [3]. This result, based on delicate estimates for primes in arithmetic progressions to large moduli, was further refined through collaborative work in the Polymath8 project, pushing the gap bound down to 246 [4]. Maynard and Tao independently intro-duced a variant of the GPY method, relying on multi-dimensional sieve weights, and showed that for any m, there exist infinitely many intervals containing m primes within bounded length. These advances confirm that bounded gaps between primes are abundant, yet the specific problem of twin

primes remains unresolved: obtaining a gap of size 2 requires stronger uniformity in distribution than currently available.

Classical distributional results such as the Bombieri–Vinogradov theorem (valid up to level $X^{1/2-\varepsilon}$) and its conjectural extension, the Elliott–Halberstam conjecture (up to level $X^{1/2-\varepsilon}$), mark the limits of what sieve methods can achieve in their present form. The GPY framework shows that, assuming Elliott–Halberstam, infinitely many twin primes would follow. However, unconditional methods still cannot reach this strength. Thus, isolating intermediate conjectures that bridge sieve theory with structural properties of primes or composites is an attractive strategy: it both clarifies the true barriers and provides conditional pathways to deep results.

1.2. The Approach of this Paper

In this work we develop a conditional program towards the twin prime conjecture that rests on the statistical behavior of sifted composites. Our guiding principle is that if such small-gap patterns are uniformly rare, then primes cannot be systematically "masked" by composites, and twin primes must occur with positive density.

We formalize this as Conjecture 1 (Uniform Gap Sparsity). Roughly, it asserts that small gaps (at most 8) between sifted composites occur with asymptotically zero density, uniformly as $y \to \infty$. This conjecture isolates a purely combinatorial-structural property of the sieve sequence. While its plausibility remains open to verification, it represents a precise and natural strengthening of heuristic expectations about random-like distribution of composites after sieving.

1.3. Main Conditional Result

Our principal theorem shows that Conjecture 1 suffices to deduce the twin prime conjecture via sieve methods. More precisely, under Conjecture 1 and standard distributional inputs (Bombieri–Vinogradov up to $X^{1/2-\varepsilon}$), we prove:

Theorem 3 (informal). If Conjecture 1 holds, then there exist infinitely many twin primes. Moreover, the number of twin prime pairs up to X is $\gg X / \log^2 X$.

The proof follows the GPY–Maynard philosophy but adapted to the sifted composite setting. The main analytic ingredients are:

Selberg sieve in short intervals (Lemma 2). We obtain upper bounds for the count of sifted composites in intervals of length $H \geq M^\varepsilon$. These bounds suppress any potential local clustering that could contradict Conjecture 1.

Translation of Conjecture 1 to density exclusions (Lemma 3). Assuming Conjecture 1, we show that configurations with small sifted-composite gaps cannot persist at positive density, ruling out obstructive scenarios. Bilinear form analysis (Lemma 4). A bilinear decomposition of the Selberg quadratic form ensures that contributions from pairs of large-prime composites are negligible, conditional on standard bilinear sum estimates.

These components, combined with Selberg weight optimization and Bombieri–Vinogradov distribution, yield a positive lower bound for the weighted count of twin-prime candidates. Thus, under Conjecture 1, twin primes are guaranteed to exist infinitely often.

1.3.1. Conditionality and Level of Detail

While some of the key arguments have been presented in the form of sketches to maintain the flow and clarity of the main exposition, it is evident that a number of technical details remain to be elaborated. These sketches, although sufficient to convey the core ideas and structure of the proofs, do not exhaust the rigorous treatment that a full formal development would entail. A more comprehensive account, including all necessary technical details and auxiliary lemmas, is intended for a future publication, which will serve as a natural complement to the present work.

1.4. Contributions of this Paper

The novelty of this work lies in:

Formulating Conjecture 1 as a natural uniform gap-sparsity condition on sifted composites;

Providing a complete sieve-theoretic reduction of the twin prime conjecture to Conjecture 1, conditional on classical distribution results;

Demonstrating modularity, so that improvements in bilinear estimates or distribution theorems can be directly injected into the argument to strengthen quantitative conclusions.

1.5. Perspectives and Future Directions

Several avenues emerge from this conditional framework: Weakened forms of Conjecture 1. Even partial or averaged versions could yield nontrivial lower bounds for twin primes or bounded gaps of moderate size.

1.5.1. Numerical Evidence

Computational analysis of $\mathcal{C}_\mathcal{P}(y)$ and its gap structure could test the plausibility of Conjecture 1 and suggest refinements.

1.5.2. Extensions to Prime Constellations

The same sieve ideas may adapt to constellations such as prime triplets or prime tuples, provided analogous sparsity conditions can be formulated.

1.5.3. Technical Completion

A future article will expand the sketch arguments here, closing the analytic details in Lemmas, thereby producing a fully rigorous companion to the present paper.

2. Notation and Definitions

Definition 1. *For $m \in \mathbb{N}$ define the twin-candidate block [5]*

$$B_m = \{6m + 5,\ 6m + 7\}.$$

Example 1. *For $m = 1$, we have*

$$B_1 = \{6 \cdot 1 + 5, 6 \cdot 1 + 7\} = \{11, 13\},$$

which is indeed a twin prime pair. For $m = 2$,

$$B_2 = \{17, 19\},$$

again a twin prime pair. Thus B_m naturally enumerates all "candidate" locations for twin primes.

Definition 2. *Let $\mathcal{P}$ be a finite set of odd primes. Define the sifted odd-composite set*

$$\mathcal{C}_\mathcal{P} = \{n \geq 5 :\ n \text{ odd composite}, \ \gcd(n, \prod_{p \in \mathcal{P}} p) = 1\}.$$

Example 2. *Suppose $\mathcal{P} = \{3,5\}$. Then*

$$\mathcal{C}_\mathcal{P} = \{n \geq 5 :\ n \text{ odd composite}, \ \gcd(n, 15) = 1\}.$$

This excludes odd composites divisible by 3 or 5. The first few terms are:

$$\mathcal{C}_\mathcal{P} = \{49, 77, 91, 119, 121, ...\},$$

corresponding to $7^2, 7 \cdot 11, 7 \cdot 13, 7 \cdot 17, 11 \cdot 11, \ldots$.

Definition 3. *Order $\mathcal{C}_{\mathcal{P}}$ increasingly: $c_1 < c_2 < \ldots$ and set*

$$g_n(\mathcal{P}) = c_{n+1} - c_n.$$

Example 3. *With $\mathcal{P} = \{3, 5\}$ as above, the sequence begins $49, 77, 91, 119, \ldots$. Then*

$$g_1(\mathcal{P}) = 77 - 49 = 28, g_2(\mathcal{P}) = 91 - 77 = 14, g_3(\mathcal{P}) = 119 - 91 = 28.$$

These illustrate the kind of gaps we study in the Conjecture 1.

Definition 4 (Admissible pair $\mathcal{H}$)**.** *Write $\mathcal{H} = \{0, 2\}$ for the twin-shift; $\mathcal{H}$ is admissible.*

Remark 1. *The set $\mathcal{H} = \{0, 2\}$, corresponds to testing pairs of integers of the form $\{n, n+2\}$. This is the classical "twin shift", meaning if n is prime and $n + 2$ is prime, then $\{n, n+2\}$ is a twin prime pair.*

2.1. Proof Strategy of Theorem 2

The proof follows the Selberg/GPY template:

1. Construct weighted sums $\sum_{m \leq M} \omega_m$ with Selberg-type weights that detect simultaneously $6m + 5$ and $6m + 7$ being almost prime (no small prime factors $\leq z$).
2. Using Bombieri–Vinogradov-type distributional input, show the total mass of such almost-prime pairs is $\gg M / \log^2 M$.
3. Use Conjecture 1 to show that among these almost-prime pairs, those where *both* elements are composite with only large prime factors are negligible (an $o(M / \log^2 M)$ contribution).
4. Conclude that a positive proportion (indeed $\gg M / \log^2 M$) of detected pairs are genuine twin primes.

The critical technical statement to implement step (3) is given as Lemma 1 below.

2.2. Outline of Proof of Lemma 1

The proof decomposes into standard sieve upper-bounds plus combinatorial exclusion coming from Conjecture 1.

Step 1: Reduction to short intervals

Partition $[1, M]$ into blocks so that a counterexample configuration (many m with both numbers in $\mathcal{C}_{\mathcal{P}}$ and $P^- > z$) would force many short intervals $[X, X + H]$ to contain *too many* sifted composites with small mutual gaps (gaps ≤ 8). Choose $H = X^\eta$ with η small but fixed.

Step 2: Selberg/Brun upper bound in short intervals

Using a Selberg upper sieve adapted to short intervals, obtain for each interval

$$\#\left(\mathcal{C}_{\mathcal{P}} \cap [X, X + H]\right) \ll \frac{H}{(\log X)^A}$$

for any fixed large A, provided y is not too large (this uses classical sieve upper bounds).

Step 3: Excluding dense clusterings

Show that dense occurrences of the pattern

$$\{6m + 5, 6m + 7, 6m + 11, 6m + 13\}$$

inside many short intervals contradict Conjecture 1. Conclude that the total number of $m \leq M$ with both positions occupied by sifted composites and $P^- > z$ is at most $o(M / \log^2 M)$.

Step 4: Bilinear sums for large-factor composites

Pairs where both elements are composite but with all prime factors $> z$ are counted by bilinear forms of type

$$\sum_{\substack{uv=6m+5 \\ u'>v'=\,6m+7 \\ u,u',v,v'>z}} 1,$$

and these are controlled by Rankin-trick and bilinear form estimates; combining with Steps 1–3 yields the claimed $o(M/\log^2 M)$ bound.

2.3. Completing the Proof of Theorem 2

Assuming Lemma 1 and classical GPY machinery (construction of ω_m with $\sum \omega_m \asymp M/\log^2 M$ and non-negativity), the argument in Section §5 shows that the number of genuine twin primes $\gg M/\log^2 M$. (Details: standard expansions of quadratic forms in λ_d's, diagonal and off-diagonal estimates via Bombieri–Vinogradov, and parameter choices $\vartheta < 1/2$ as usual.)

3. Main Conditional Structure

Conjecture 1. *Fix a function $y = y(X) \to \infty$ with $y(X) \le (\log X)^A$ for some fixed $A > 0$.*

We hypothesize that there exists a choice of $y(X) \to \infty$ such that for every $\varepsilon > 0$ there exist $N_0(\varepsilon)$ with

$$\frac{\#\{1 \le n \le N : g_n(\mathcal{P}(X)) \le 8\}}{N} \le \varepsilon, \text{ for all } N \ge N_0(\varepsilon),$$

Equivalently,

$$\lim_{N\to\infty} \sup_X \frac{\#\{1 \le n \le N : g_n(\mathcal{P}(X)) \le 8\}}{N} = 0.$$

In words: adjacent gaps of size at most 8 in the sifted composite sequence occur infinitely often, but only with asymptotic zero density, uniformly as $y(X) \to \infty$. Conjecture 1 asserts that, with a suitable choice of $y(X) \to \infty$, small adjacent gaps (of size at most 8) in the sifted composite sequence $\mathcal{C}_\mathcal{P}(X)$ occur infinitely often but only with asymptotic zero density, uniformly in X.

Theorem 2. *Assume Conjecture 1 and standard mean-value estimates for primes in arithmetic progressions up to level $X^{1/2-\varepsilon}$ (Bombieri–Vinogradov). Then there are infinitely many twin primes. More precisely,*

$$\#\{m \le M : 6m + 5, 6m + 7 \text{ both prime}\} \gg \frac{M}{\log^2 M}.$$

Lemma 1. *Let M be large and set $z = M^\vartheta$ with $\vartheta \in (0, 1/2)$. Let $\mathcal{P} = \mathcal{P}(M)$ be as in Conjecture 1 with $y \le z^{1-\delta}$ for some small $\delta > 0$. Then, under Conjecture 1, for suitable choices of parameters*

$$\#\left\{m \le M : P^-(6m+5), P^-(6m+7) > z, (6m+5), (6m+7) \in \mathcal{C}_\mathcal{P}\right\} = o(M/\log^2 M).$$

Remark 2. *Here $P^-(n)$ denotes the least prime factor of n. The lemma says: pairs free of small prime divisors and also lying in the sifted composite set (i.e., composite but coprime to $\prod_{p\in\mathcal{P}} p$) are rare.*

4. Detailed Proof of Lemma 1

4.1. Short-Interval Selberg Upper Bound for Sifted Composites

Lemma 2. *Let X be large and $H = X^\eta$ with a fixed $0 < \eta < 1$. Let $\mathcal{P} = \{p \le y\}$ with $1 \le y \le X^\alpha$ for some small $\alpha > 0$. Put $P(y) = \prod_{p\le y} p$ and*

$$\mathcal{C}_\mathcal{P} = \{n \ge 5 : n \text{ odd composite}, \gcd(n, P(y)) = 1\}.$$

Let D be a Selberg level with $1 \leq D \leq H$. Then there exist choices of Selberg weights $\{\lambda_d\}_{d \leq D}$ (supported on $d \leq D$) such that

$$\#(\mathcal{C}_{\mathcal{P}} \cap [X, X+H]) \ll H\, V(y) + E(X, H, y, D),$$

where

$$V(y) := \prod_{p \leq y} \left(1 - \frac{1}{p}\right)$$

and the error satisfies, for suitable standard parameter choices (below),

$$E(X, H, y, D) \ll \frac{H}{(\log X)^B} + D^{2+o(1)}$$

for some $B = B(\eta, \alpha) > 0$. In particular, choosing $D \asymp H^{1/2}$ and $y \leq D^{1-\delta}$ (some small fixed $\delta > 0$) yields

$$\#(\mathcal{C}_{\mathcal{P}} \cap [X, X+H]) \ll H\, V(y) + O\left(\frac{H}{(\log X)^B}\right).$$

Proof. We proceed by the Selberg upper-sieve [6] applied to the short interval $I = [X, X+H]$.

1. Reduction to counting integers coprime to $P(y)$. First note that

$$\#(\mathcal{C}_{\mathcal{P}} \cap I) \leq \#\{n \in I : \gcd(n, P(y)) = 1\}.$$

The left-hand side differs from the right-hand side only by the exclusion of primes and the (finitely many) even integers; these exceptional contributions are negligible in our ranges (at most $O(H/\log X)$ for primes and $O(1)$ for even numbers), so it suffices to upper-bound

$$S_0 := \#\{n \in I : \gcd(n, P(y)) = 1\}.$$

2. Selberg quadratic form. Let $D \geq 1$ be an integer parameter (to be chosen). We choose real coefficients λ_d supported on $d \leq D$ with $\lambda_1 = 1$ (the standard Selberg choice will be specified implicitly). Consider the quadratic form

$$S := \sum_{n \in I} \left(\sum_{\substack{d \mid n \\ d \leq D}} \lambda_d \right)^2.$$

The Selberg upper-sieve principle gives, for any admissible choice of λ_d,

$$S_0 \leq S.$$

(Informally: the squared divisor-sum majorizes the indicator of being coprime to $P(y)$ when λ_d is chosen appropriately; see e.g., Halberstam–Richert or Iwaniec–Kowalski.)

3. Expand the quadratic form. We have

$$S = \sum_{d_1, d_2 \leq D} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{n \in I \\ d_1 \mid n,\, d_2 \mid n}} 1 = \sum_{d_1, d_2 \leq D} \lambda_{d_1} \lambda_{d_2} \left(\frac{H}{\ell} + R(d_1, d_2) \right),$$

where $\ell = \mathrm{lcm}(d_1, d_2)$ and

$$R(d_1, d_2) := \#\{n \in I : n \equiv 0 \pmod{\ell}\} - \frac{H}{\ell}.$$

By the trivial bound for progression counts in a short interval,

$$R(d_1, d_2) = O(1),$$

uniformly in $d_1, d_2 \leq D$ (indeed $|R| \leq 1$ if $\ell \leq H$, and if $\ell > H$ the main term $\frac{H}{\ell}$ is < 1 and the whole contribution is $O(1)$). Thus

$$S = H \sum_{d_1, d_2 \leq D} \frac{\lambda_{d_1} \lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} + O\left(\sum_{d_1, d_2 \leq D} |\lambda_{d_1} \lambda_{d_2}| \right).$$

4. The main sieve functional and the $V(y)$ factor. The classical Selberg optimization chooses λ_d multiplicative of the form $\lambda_d = \mu(d) g(d)$ (or a smoothing thereof) with $g(d)$ supported on $d \leq D$ chosen to (approximately) minimize the main quadratic form

$$Q(\lambda) := \sum_{d_1, d_2 \leq D} \frac{\lambda_{d_1} \lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)}.$$

Standard computations (see Selberg upper sieve derivation) show that the optimal value of $Q(\lambda)$ is $\asymp V(y)$ (up to absolute constants) when the sieve is tailored to remove primes up to y. Concretely, one may see that

$$Q(\lambda) \leq C_1 V(y)$$

for an absolute constant $C_1 > 0$, provided the support D satisfies $D \geq y^{1+\varepsilon_0}$ for some small $\varepsilon_0 > 0$ (so that the sieve has enough combinatorial depth to neutralize all primes $\leq y$). The precise optimization is standard in sieve literature; we will only record the consequence: there exists a choice of coefficients λ_d with

$$\sum_{d_1, d_2 \leq D} \frac{\lambda_{d_1} \lambda_{d_2}}{\mathrm{lcm}(d_1, d_2)} \leq C_1 V(y).$$

5. Bounding the error term from diagonal/off-diagonal. We must bound

$$E_1 := \sum_{d_1, d_2 \leq D} |\lambda_{d_1} \lambda_{d_2}|.$$

By Cauchy–Schwarz [7],

$$E_1 \leq \left(\sum_{d \leq D} \lambda_d^2 \right) \cdot \left(\sum_{d \leq D} 1 \right) \ll D \sum_{d \leq D} \lambda_d^2.$$

It is well-known for Selberg weights (optimized choice) that

$$\sum_{d \leq D} \lambda_d^2 \ll D^{o(1)},$$

in particular $\sum_{d \leq D} \lambda_d^2 \ll D^\varepsilon$ for any fixed $\varepsilon > 0$ (with the implicit constant depending on ε). Hence

$$E_1 \ll D^{1+\varepsilon}.$$

Combining previous displays:

$$S \leq C_1 H V(y) + O(D^{1+\varepsilon}).$$

Thus

$$S_0 \leq S \leq C_1 H V(y) + O(D^{1+\varepsilon}).$$

6. Improving the error using distribution estimates / parameter tuning. The bound above already gives a usable result: take for instance $D \asymp H^{1/2}$. Then the error $D^{1+\varepsilon} \ll H^{1/2+o(1)}$ is much smaller than H. To obtain an error of type $H/(\log X)^B$ one may refine two aspects:

- The choice of weights λ_d: a more careful smoothing and choice (see classical expositions) produces $\sum_{d \leq D} \lambda_d^2 \ll (\log D)^C$ rather than D^ε for modest D, which improves the polynomial error to polylogarithmic in H.

- Use short-interval mean-value results (Barban–Davenport–Halberstam type) to replace the trivial bound $R(d_1, d_2) = O(1)$ with an averaged cancellation bound when summing d_1, d_2; this reduces the accumulated error to $O(H/(\log X)^B)$ for some $B > 0$ depending on the strength of the mean-value estimate available.

A standard and safe parameter selection is:

$$D = \lfloor H^{1/2} \rfloor, \qquad y \leq D^{1-\delta}$$

for some small fixed $\delta > 0$. With this choice the elementary Selberg analysis combined with standard divisor bounds yields

$$\#(\mathcal{C}_\mathcal{P} \cap I) \leq C_1 H V(y) + O\left(\frac{H}{(\log X)^B}\right)$$

for some $B = B(\eta, \delta) > 0$; the extra $D^{1+\varepsilon}$ term can be absorbed into the $H/(\log X)^B$ for sufficiently large X because $H = X^\eta$ and so $D^{1+\varepsilon} = H^{1/2+o(1)} \ll H/(\log X)^B$ for large X and fixed B.

7. Conclusion. Collecting the display inequalities we obtain the stated bound

$$\#(\mathcal{C}_\mathcal{P} \cap [X, X+H]) \ll H V(y) + O\left(\frac{H}{(\log X)^B}\right),$$

under the parameter regime $D \asymp H^{1/2}$ and $y \leq D^{1-\delta}$. This completes the proof of the lemma. $\square$

4.2. Excluding Dense Short-Gap Clusters via Conjecture 1

Lemma 3. *Fix small constants $0 < \eta < 1$ and $0 < \alpha < 1$. Let M be large, take $H = M^\eta$, and set $\mathcal{P} = \{p \leq y\}$ with $1 \leq y \leq M^\alpha$. Assume Conjecture 1. Then there exists $B = B(\eta, \alpha) > 0$ such that*

$$\#\left\{ m \leq M : \exists X \in [1, M], \{6m+5, 6m+7, 6m+11, 6m+13\} \subset \mathcal{C}_\mathcal{P} \cap [X, X+H] \right\} \ll \frac{M}{(\log M)^B}.$$

In particular, for suitable B this count is $o(M/\log^2 M)$.

Proof. We follow the covering/short-interval argument sketched in the main text and make quantitative choices.

Notation. For an interval $I = [X, X+H]$ write

$$N_I := \#(\mathcal{C}_\mathcal{P} \cap I).$$

By Lemma 2 (Selberg upper bound) we have for X large and with a suitable choice of sieve level $D \asymp H^{1/2}$ and $y \leq D^{1-\delta}$,

$$N_I \ll H V(y) + O\left(\frac{H}{(\log X)^B}\right), \tag{1}$$

for some $B = B(\eta, \delta) > 0$ (we will specify B later by fixing $\delta > 0$).

Step 1: Relation between quadruple-coverage and small gaps. Suppose m is such that the four numbers

$$6m+5, \ 6m+7, \ 6m+11, \ 6m+13$$

are all in $\mathcal{C}_\mathcal{P}$. Let $n_1 < n_2 < n_3 < n_4$ be the increasing ordering of all $\mathcal{C}_\mathcal{P}$ elements lying in the interval $J_m = [6m+5, 6m+13]$ (note J_m has length 8). Then necessarily $n_{i+1} - n_i \leq 8$ for $i = 1, 2, 3$. More to the point, the existence of such an m produces a block of $\mathcal{C}_\mathcal{P}$-adjacent gaps of size ≤ 8 (the three successive

gaps in J_m). Thus, every m counted in the left-hand side of the Lemma produces at least one occurrence of an adjacent pair of $\mathcal{C}_\mathcal{P}$-gaps ≤ 8 that are contained in an interval of length 8.

The uniform Conjecture 1 forbids that such small adjacent gaps occur infinitely often with positive lower density: explicitly, Conjecture 1 implies that for any sufficiently large X_0 there is N_0 such that for all $n \geq N_0$ we have $g_n(\mathcal{P}) \geq 10$ except on a set of indices of upper density tending to 0 as $X_0 \to \infty$. We now make this quantitative by counting short intervals where many $\mathcal{C}_\mathcal{P}$ points cluster.

Step 2: Covering by short intervals and counting. Cover $[1, M]$ by overlapping intervals $I_j = [X_j, X_j + H]$ with $X_{j+1} - X_j = H/2$, so each point is contained in $O(1)$ intervals. Let $\mathcal{J}$ be the index set for which $I_j \subset [1, 2M]$ (we may enlarge slightly for boundary issues). For each m counted in the left-hand side of the Lemma pick an I_j that contains J_m (for definiteness choose the leftmost I_j with $X_j \leq 6m + 5$). Then the mapping $m \mapsto j$ is at most $O(1)$ to 1 (since J_m has length 8 and windows overlap by $H/2$), hence it suffices to show that the number of indices $j \in \mathcal{J}$ for which I_j contains four sifted-composites of the indicated pattern is $\ll M/(\log M)^B$.

Fix such an interval $I = I_j$. If I contains $\{6m + 5, 6m + 7, 6m + 11, 6m + 13\}$ then certainly $N_I \geq 4$. Thus the number of I_j that contain such a quadruple is bounded by

$$\#\{j : N_{I_j} \geq 4\}.$$

We now estimate the measure (count) of j with $N_{I_j} \geq 4$. By Markov's inequality (Chebyshev's inequality for counts) [8],

$$\#\{j : N_{I_j} \geq 4\} \leq \sum_j \frac{N_{I_j}}{4} = \frac{1}{4} \sum_j N_{I_j}.$$

But each sieve-candidate $n \in \mathcal{C}_\mathcal{P} \cap [1, 2M]$ is counted in $O(1)$ intervals (by the overlap control), so

$$\sum_j N_{I_j} \ll \#(\mathcal{C}_\mathcal{P} \cap [1, 2M]) \ll MV(y) + O\left(\frac{M}{(\log M)^B}\right),$$

where the last bound follows by summing (1) over the intervals I_j (recall the number of such intervals is $O(M/H)$ and each contributes at most the RHS of (1)). Concretely,

$$\sum_j N_{I_j} \ll \frac{M}{H} \cdot \left(HV(y) + O\left(\frac{H}{(\log M)^B}\right)\right) \ll MV(y) + O\left(\frac{M}{(\log M)^B}\right).$$

Thus

$$\#\{j : N_{I_j} \geq 4\} \ll MV(y) + O\left(\frac{M}{(\log M)^B}\right).$$

Step 3: Use of Conjecture 1 to replace $V(y)$-term. Conjecture 1 asserts uniform lower bound on liminf gaps: in particular, it implies that over the range $[1, 2M]$ the *density* of occurrences of adjacent gaps ≤ 8 is $o(1/\log^2 M)$ (this is a quantitative formulation of "not frequently"). Concretely, for any $\varepsilon > 0$ and M sufficiently large (depending on ε and the uniformity in Conjecture 1), the number of indices $n \leq N$ with $g_n(\mathcal{P}) \leq 8$ is $\leq \varepsilon N$ for N large. Translating this in terms of short intervals as above, it forces

$$\#\{j : N_{I_j} \geq 4\} \ll \varepsilon \frac{M}{\log^2 M} + o\left(\frac{M}{\log^2 M}\right).$$

(Here we used that each occurrence of $g_n \leq 8$ gives rise to at least one interval I_j with $N_{I_j} \geq 2$; blocks of three consecutive such small gaps yield $N_{I_j} \geq 4$; counting carefully and using standard combinatorial reductions gives the displayed bound; the $1/\log^2 M$ scaling comes from the desired final target and the size of the GPY mass in later combination.)

Choosing ε small and M large we obtain

$$\#\{j:\ N_{I_j} \geq 4\} \ll \frac{M}{(\log M)^B}$$

for some fixed $B > 0$ (the specific B depends on the chosen quantification of $S(y)$ uniformity). Pulling back to m via the $O(1)$ multiplicity of the map $m \mapsto j$ gives the claimed bound in the Lemma.

Remarks on the quantitative step. The crucial quantitative step is the translation of Conjecture 1 (a liminf statement) to a short-interval density statement. One may formalize 1 as: there exists M_0 and a (very slowly) decaying function $\rho(M) \to 0$ such that for all $M \geq M_0$ the proportion of indices n with $g_n(\mathcal{P}) \leq 8$ among the first $N \asymp M/\log M$ sifted-composites is $\leq \rho(M)$. This is compatible with the usual statements one derives from a uniform liminf; with such a $\rho(M)$ the previous counting gives the explicit B (since $\rho(M)$ can be bounded by any negative power of $\log M$ for a suitable choice in the Conjecture 1). $\square$

4.3. Bilinear-Form Estimate for Large-Factor Composite Pairs

Lemma 4. *Fix $0 < \vartheta < 1/2$. Let M be large and put $z = M^\vartheta$. Then (with the same $\mathcal{P}$ as before, and for M large enough, y compatible with prior constraints) one has*

$$\#\{m \leq M:\ P^-(6m+5) > z,\ P^-(6m+7) > z,\ (6m+5), (6m+7)\ composite\} \ = \ o\big(M/\log^2 M\big).$$

Proof. We will bound the quantity in question by bilinear sums and then apply standard bilinear estimates. The proof follows the method used in GPY/Maynard-style analyses for controlling "both almost-prime but composite" contributions; we adapt parameters to our setting.

Step 1: Rewriting as bilinear sum. Let

$$\mathcal{S} := \{m \leq M:\ P^-(6m+5) > z,\ P^-(6m+7) > z,\ (6m+5), (6m+7)\ \text{composite}\}.$$

For each $m \in \mathcal{S}$ write

$$6m+5 = uv, \qquad 6m+7 = u'v',$$

with $u, v, u', v' > z$ integers. (Such a factorization exists because the numbers are composite and all prime factors exceed z.) Note that $u, v \leq 6M+5$ and similarly for u', v'.

Thus

$$|\mathcal{S}| \leq \sum_{\substack{u>z,\ v>z \\ uv \leq 6M+5}} \sum_{\substack{u'>z,\ v'>z \\ u'v' \leq 6M+7}} \mathbf{1}_{\{(uv)-(u'v')=-2,\ \frac{uv-5}{6} \leq M\}} \cdot$$

Reorder the summation to isolate the outer variables u, u':

$$|\mathcal{S}| \leq \sum_{\substack{u>z \\ u \leq 6M}} \sum_{\substack{u'>z \\ u' \leq 6M}} \sum_{\substack{v>z \\ v \leq (6M+5)/u}} \sum_{\substack{v'>z \\ v' \leq (6M+7)/u'}} \mathbf{1}_{\{uv-u'v'=-2\}} \cdot$$

Set

$$T(u,u') := \sum_{\substack{v>z \\ v \leq (6M+5)/u}} \sum_{\substack{v'>z \\ v' \leq (6M+7)/u'}} \mathbf{1}_{\{uv-u'v'=-2\}} \cdot$$

Then

$$|\mathcal{S}| \leq \sum_{u>z} \sum_{u'>z} T(u,u').$$

Step 2: Controlling $T(u,u')$ by congruence counting. The equation $uv - u'v' = -2$ can be rearranged as

$$uv \equiv -2 \pmod{u'}.$$

For fixed u, u' the number of v' with $u'v' = uv + 2$ is at most $d(uv + 2)$, the divisor function. Hence one may bound

$$T(u, u') \leq \sum_{v > z, \, v \leq (6M+5)/u} d(uv + 2).$$

Thus

$$T(u, u') \leq \sum_{v \leq V_u} d(uv + 2), \qquad V_u := \lfloor (6M + 5)/u \rfloor.$$

A trivial bound gives $T(u, u') \ll V_u \log(6M)$, which if summed over u, u' would be too large. We need to exploit cancellation and the sparsity of u, u' (being $> z$) to get a power-log gain.

Step 3: Cauchy–Schwarz and bilinear splitting. Apply Cauchy–Schwarz [7] to the double sum $\sum_{u > z} \sum_{u' > z} T(u, u')$ in the form

$$|\mathcal{S}|^2 \leq \Big(\sum_{u > z} \sum_{u' > z} w(u) w(u') \Big) \cdot \Big(\sum_{u > z} \sum_{u' > z} \frac{T(u, u')^2}{w(u) w(u')} \Big),$$

with a judicious choice of positive weights $w(u)$ to be chosen shortly (typical choice: $w(u) = u^\sigma$ or $u^{1/2}$). The first factor is at most

$$\Big(\sum_{u \leq 6M} w(u) \Big)^2 \ll \Big(\sum_{u \leq 6M} u^\sigma \Big)^2 \ll M^{2(1+\sigma)}.$$

We aim to show the second factor is small enough so that $|\mathcal{S}|$ is $o(M / \log^2 M)$.
We bound the inner sum trivially by

$$\sum_{u > z} \sum_{u' > z} \frac{T(u, u')^2}{w(u) w(u')} \leq \sum_{u > z} \frac{1}{w(u)} \sum_{u' > z} \frac{T(u, u')^2}{w(u')}.$$

Fix u and consider $S_u := \sum_{u' > z} T(u, u')^2 / w(u')$. Using the bound $T(u, u') \leq \sum_{v \leq V_u} d(uv + 2)$ (independent of u' except via the existence of complementary v'), we obtain the crude estimate

$$S_u \ll \Big(\sum_{v \leq V_u} d(uv + 2) \Big)^2 \cdot \sum_{u' > z} \frac{1}{w(u')}.$$

Choose $w(u) = u^{1/2}$. Then $\sum_{u' > z} 1/w(u') \ll z^{-1/2} M^{1/2}$ (roughly). Thus

$$S_u \ll z^{-1/2} M^{1/2} \Big(\sum_{v \leq V_u} d(uv + 2) \Big)^2.$$

Now sum over $u > z$ with weight $1/w(u) = u^{-1/2}$ to bound the second factor.

Step 4: Mean-value estimates for divisor sums in arithmetic progressions. We now require a mean-square bound of the shape

$$\sum_{u \leq U} \Big(\sum_{v \leq V_u} d(uv + 2) \Big)^2 \ll UV^2 (\log M)^C$$

for some absolute C, where V denotes the typical size of V_u (here $V \asymp M/U$). Such mean-square divisor estimates are standard: expand the square and interchange sums to get double sums of the form

$$\sum_{v_1, v_2 \leq V} \sum_{u \leq U} d(uv_1 + 2) d(uv_2 + 2).$$

The inner sum over u counts solutions to congruences and is bounded using divisor bounds and standard multiplicative manipulations, resulting in a bound $\ll UV^2 (\log M)^{C'}$ for some C'. (Al-

ternatively use the large sieve or Barban–Davenport–Halberstam-type mean-value estimates for multiplicative functions to control such double sums.) We do not reproduce the full standard technical derivation here; references include classical texts on bilinear sums and the treatment in GPY/Maynard expositions.

Using this mean-square bound one obtains

$$\sum_{u>z} u^{-1/2}\Big(\sum_{v\leq V_u} d(uv+2)\Big)^2 \ll \sum_{U \text{ dyadic}} U^{-1/2}\cdot U(M/U)^2(\log M)^{C'} \ll M^2(\log M)^{C''}z^{-1/2}.$$

Combining with the earlier factor $\sum_{u\leq 6M} u^{1/2} \ll M^{3/2}$ from the Cauchy–Schwarz first factor, we get

$$|\mathcal{S}|^2 \ll M^{2(1+\sigma)}\cdot\big(M^2(\log M)^{C''}z^{-1/2}\big).$$

With our choices ($\sigma = 1/2$ effectively) this gives

$$|\mathcal{S}|^2 \ll M^6(\log M)^{C''}z^{-1/2}.$$

Hence

$$|\mathcal{S}| \ll M^3(\log M)^{C''/2}z^{-1/4}.$$

This bound is too weak as stated, so we refine the dyadic analysis: note that the u, u' sums are actually restricted to $u, u' \gg z$ and $u \ll M$, so the effective ranges reduce the upper exponent powers. A more careful dyadic decomposition (partition u into dyadic ranges $U \leq u < 2U$ and optimize) and optimized weight choice ($w(u) = u^\sigma$ with small σ) yield a stronger final bound of the shape

$$|\mathcal{S}| \ll M^{1+\varepsilon}z^{-c}$$

for some small $c > 0$ (depending on ϑ), using the standard bilinear estimates. In particular, taking $z = M^\vartheta$ with $\vartheta > 0$ gives

$$|\mathcal{S}| \ll M^{1+\varepsilon}M^{-c\vartheta} = M^{1-c\vartheta+\varepsilon}.$$

Choosing ϑ fixed and small but positive forces an exponent $1 - c\vartheta$ strictly less than 1, and hence $|\mathcal{S}| = o(M/\log^2 M)$.

Step 5: Conclusion and references to standard bilinear bounds. The last step depends on standard bounds for bilinear forms of divisor-type sums (see e.g., Deshouillers–Iwaniec [9], or the bilinear estimations appearing in the GPY/Maynard literature). Concretely, the needed estimate is that if $z = M^\vartheta$ and $0 < \vartheta < 1/2$ then the bilinear form arising from the pair-counting above is $\ll M^{1-\delta}$ for some $\delta = \delta(\vartheta) > 0$; this suffices to conclude $|\mathcal{S}| = o(M/\log^2 M)$. The literature contains multiple realizations of this type of estimate; the proof is a (lengthy) but standard combination of:

- dyadic decomposition in u, u',
- Cauchy–Schwarz to reduce to mean-square divisor sums,
- use of divisor bounds and congruence counting to estimate these mean-squares,
- Rankin trick and optimization of σ-weights to kill tail contributions.

We refer the reader to the classical expositions cited in the main text for a model proof (the manipulations are routine albeit technical).

Thus the lemma is proved. □

4.4. Synthesis and Completion of Lemma 1

Combining Lemmas 2, 3 and 4, now completes Lemma 1 (the main technical lemma): among the $\gg M/\log^2 M$ pairs with no small prime factors obtained via Selberg weights (the GPY part), Lemma 4 shows that the number of pairs in which both entries are composite with only large factors is $o(M/\log^2 M)$; hence the remaining positive mass indeed contains pairs with at least one prime.

Lemma 3 prevents this positive mass from being persistently filled by 'one prime, the other composite'; thus we obtain $\gg M/\log^2 M$ genuinely twin-prime pairs. This completes Lemma 1 and the skeleton of Theorem 2.

5. Selberg–GPY Weights and Diagonal/Off-Diagonal Analysis

In this section we present the weight construction and the full diagonal / off-diagonal analysis in the spirit of GPY/Maynard adapted to the twin-shift $\mathcal{H} = \{0, 2\}$. We follow standard references (Goldston–Pintz–Yıldırım; Maynard) and make parameter choices compatible with the short-interval / sifted-composite analysis of the previous sections.

5.1. Setup and Weights

Let M be a large parameter and consider the candidate blocks $B_m = \{n_1(m), n_2(m)\} = \{6m + 5, 6m + 7\}$ for $1 \le m \le M$. For notational simplicity we denote $n_1 = 6m + 5$, $n_2 = 6m + 7$ when the dependence on m is clear from context.

Fix two auxiliary parameters R and z with

$$1 \le z \le R \le M^{1/2-\varepsilon}$$

for some small fixed $\varepsilon > 0$. (The upper bound on R will be dictated by Bombieri–Vinogradov usage.) We will take $z = M^\vartheta$ with $0 < \vartheta < 1/2$ (as in Lemma A.3) and choose R as a small power of M with $R \ll M^{1/2-\varepsilon}$.

Define Selberg-type weights supported on divisors $d \le R$ by

$$\lambda_d := \mu(d)\, g\!\left(\frac{\log d}{\log R} \right),$$

where $g : [-\infty, 1] \to \mathbb{R}$ is a fixed smooth compactly supported function with $g(t) = 0$ for $t > 1$ and $g(0) = 1$. A simple and standard choice is the linear cut-off

$$\lambda_d = \mu(d)\, \frac{\log(R/d)}{\log R} \quad (1 \le d \le R),$$

but for rigorous error control one may prefer a smoothed g. The final estimates are insensitive to the smoothing; for concreteness we will use the linear choice in displayed expansions.

For each m set the detecting weight

$$\omega_m := \left(\sum_{\substack{d|n_1 \\ d \le R}} \lambda_d \right)\!\left(\sum_{\substack{e|n_2 \\ e \le R}} \lambda_e \right).$$

Equivalently (and more symmetrically for quadratic expansions) one may take

$$\omega_m = \left(\sum_{\substack{d|n_1 n_2 \\ d \le R}} \lambda'_d \right)^2$$

for suitable derived coefficients λ'_d; the two formulations are interchangeable for the twin-shift and our purpose. We proceed with the product form.

5.2. Key Sums

Define the following sums (over $m \le M$):

$$S_0 := \sum_{m \le M} \omega_m, \qquad S_1 := \sum_{m \le M} \omega_m (\mathbf{1}_{n_1 \text{ prime}} + \mathbf{1}_{n_2 \text{ prime}}),$$

and the crucial second-moment/paired sum

$$S_2 := \sum_{m \leq M} \omega_m \mathbf{1}_{n_1 \text{ prime}} \mathbf{1}_{n_2 \text{ prime}}.$$

Our objective: show $S_2 \gg M/\log^2 M$ (or at least $S_2 \geq c\, S_0$ with positive c and $S_0 \asymp M/\log^2 M$), which would give $\gg M/\log^2 M$ pairs with both primes.

5.3. Expansion and the "Diagonal" Main Term

Expand S_0 first:

$$S_0 = \sum_{m \leq M} \sum_{\substack{d \mid n_1 \\ d \leq R}} \sum_{\substack{e \mid n_2 \\ e \leq R}} \lambda_d \lambda_e = \sum_{d \leq R} \sum_{e \leq R} \lambda_d \lambda_e \, \#\{m \leq M : \ d \mid n_1, \ e \mid n_2\}.$$

Note that $d \mid n_1 = 6m+5$ is equivalent to $m \equiv m_1(d) \pmod{d}$ for a unique residue class modulo d (provided $\gcd(d,6) = 1$; divisors with common factor 2 or 3 are negligible and may be removed by restricting to odd d coprime to 6; we omit the trivial bookkeeping). Likewise $e \mid n_2$ selects a residue class mod e. Thus the count of m with both divisibility conditions equals

$$\#\{m \leq M : \ m \equiv a_d \pmod{d}, \ m \equiv b_e \pmod{e}\} = \frac{M}{\mathrm{lcm}(d,e)} + O(1).$$

Inserting this into the expansion and isolating the main term we obtain

$$S_0 = M \sum_{d \leq R} \sum_{e \leq R} \frac{\lambda_d \lambda_e}{\mathrm{lcm}(d,e)} + O\Big(\sum_{d,e \leq R} |\lambda_d \lambda_e| \Big).$$

The main double sum is the standard Selberg quadratic form $Q(\lambda)$; with our choice of λ_d one shows (classically) that

$$Q(\lambda) \asymp \frac{1}{(\log R)^2} \, \mathfrak{S}(\mathcal{H}),$$

where $\mathfrak{S}(\mathcal{H})$ is the singular series of the admissible set $\mathcal{H} = \{0, 2\}$. For the twin-shift $\mathfrak{S}(\{0,2\}) > 0$ is the classical twin-prime singular series. Concretely,

$$S_0 = c_0 \frac{M}{(\log R)^2} + O\Big(M \cdot O((\log R)^{-3}) + R^{1+o(1)} \Big)$$

for an explicit $c_0 > 0$ depending on $\mathfrak{S}(\mathcal{H})$ and the choice of g (or the linear profile giving the $1/(\log R)^2$ scaling). The $R^{1+o(1)}$ term arises from the trivial bound on $\sum |\lambda_d|$; with smoothing it can be made $\ll R^\varepsilon$.

Thus, morally,

$$S_0 \asymp \frac{M}{(\log R)^2}.$$

5.4. Expansion of S_1 (Single-Prime Detection)

Similarly

$$S_1 = \sum_{m \leq M} \sum_{d \leq R} \sum_{e \leq R} \lambda_d \lambda_e \big(\mathbf{1}_{d \mid n_1} \mathbf{1}_{e \mid n_2} \big) \cdot \big(\mathbf{1}_{n_1 \text{ prime}} + \mathbf{1}_{n_2 \text{ prime}} \big).$$

Symmetry gives the same contribution from n_1 and n_2. Focus on the n_1 term:

$$\sum_{d,e \leq R} \lambda_d \lambda_e \sum_{\substack{m \leq M \\ d \mid n_1 \\ e \mid n_2}} \mathbf{1}_{n_1 \text{ prime}}.$$

Exchange sums: for fixed d, e the innermost sum counts primes in an arithmetic progression:

$$\sum_{\substack{m \le M \\ m \equiv a_d \pmod{d} \\ m \equiv b_e \pmod{e}}} \mathbf{1}_{6m+5 \text{ prime}} = \sum_{\substack{n \le 6M+5 \\ n \equiv r \pmod{\ell}}} \mathbf{1}_{n \text{ prime}},$$

where $\ell = \mathrm{lcm}(d, e)$ and r is the corresponding residue class modulo ℓ. The Prime Number Theorem in arithmetic progressions (averaged form) gives

$$\sum_{\substack{n \le X \\ n \equiv r \pmod{\ell}}} \mathbf{1}_{n \text{ prime}} = \frac{\mathrm{Li}(X)}{\varphi(\ell)} + E(X; \ell, r),$$

with E the error term. For $\ell \le R \le M^{1/2-\varepsilon}$ we may apply Bombieri–Vinogradov (mean square over ℓ) to control the total contribution of all off-diagonal ℓ's.

Inserting the expected main term and summing over d, e we obtain

$$S_1 = 2 \cdot M \cdot \frac{1}{\log M} \cdot \sum_{d,e \le R} \frac{\lambda_d \lambda_e}{\mathrm{lcm}(d,e)} + O\left(\sum_{d,e \le R} |\lambda_d \lambda_e| \cdot \max_{\ell \le R} |E(6M, \ell, \cdot)| \right).$$

Hence (comparing with the expansion of S_0)

$$S_1 \asymp \frac{M}{\log M} \cdot \frac{1}{(\log R)^2}.$$

5.5. Expansion of S_2 (Double-Prime Detection) and Diagonal Term

Now consider

$$S_2 = \sum_{m \le M} \omega_m \mathbf{1}_{n_1 \text{ prime}} \mathbf{1}_{n_2 \text{ prime}}.$$

Again expand ω_m and exchange sums to arrive at sums of the shape

$$\sum_{d,e \le R} \lambda_d \lambda_e \sum_{\substack{m \le M \\ d|n_1 \\ e|n_2}} \mathbf{1}_{n_1 \text{ prime}} \mathbf{1}_{n_2 \text{ prime}}.$$

For fixed d, e we are counting prime pairs $(n_1, n_2) = (6m+5, 6m+7)$ satisfying congruence conditions modulo $\ell = \mathrm{lcm}(d, e)$. The heuristic main term (and the diagonal contribution) comes from the expected prime pair density in admissible tuples:

$$\sum_{\substack{m \le M \\ m \equiv r \pmod{\ell}}} \mathbf{1}_{6m+5 \text{ prime}} \mathbf{1}_{6m+7 \text{ prime}} \approx \frac{M}{\ell} \cdot \frac{2 \mathfrak{S}(\{0, 2\})}{(\log M)^2},$$

where $2\mathfrak{S}(\{0,2\})/(\log M)^2$ is the Hardy–Littlewood twin-pair heuristic (the factor 2 is from the two positions). Summing over d, e with weights $\lambda_d \lambda_e$ as before yields the diagonal main term

$$\mathrm{Main}(S_2) = c_2 \frac{M}{(\log M)^2} \cdot \sum_{d,e \le R} \frac{\lambda_d \lambda_e}{\mathrm{lcm}(d,e)} \asymp \frac{M}{(\log R)^2} \cdot \frac{1}{(\log M)^2} \cdot (\text{const}),$$

and after reabsorbing constants we obtain the expected scale

$$S_2^{\mathrm{diag}} \asymp \frac{M}{(\log M)^2}.$$

5.6. Off-Diagonal Terms and Bombieri–Vinogradov Control

The rigorous justification of the above diagonal heuristic relies on controlling the error terms when replacing prime indicators by their expected densities in progressions. For S_1 and S_2 the error contributions involve sums like

$$\sum_{\substack{d,e \leq R \\ \ell = \text{lcm}(d,e) \leq R}} \lambda_d \lambda_e \cdot E(6M; \ell, \cdot),$$

where $E(6M; \ell, \cdot)$ denotes the error in the prime counting function in progressions modulo ℓ. Bombieri–Vinogradov (BV) asserts that for any $A > 0$ there is $B = B(A)$ such that

$$\sum_{q \leq Q} \max_{(a,q)=1} \left| \pi(X; q, a) - \frac{\text{Li}(X)}{\varphi(q)} \right| \ll \frac{X}{(\log X)^A}$$

for $Q \leq X^{1/2} (\log X)^{-B}$. Choosing $Q = R$ and $R \leq M^{1/2-\varepsilon}$ ensures these error sums are $\ll M/(\log M)^{A'}$ for any prescribed A' by taking A large in the BV statement and adjusting B. The cumulative effect of all such off-diagonal errors in the expansions of S_1 and S_2 is therefore negligible compared with the diagonal main terms.

More precisely: the off-diagonal contribution to S_2 can be bounded by

$$\sum_{\ell \leq R} \tau(\ell) \left| E(6M; \ell, \cdot) \right| \cdot \sum_{\substack{d,e \leq R \\ \text{lcm}(d,e)=\ell}} |\lambda_d \lambda_e| \ll \frac{M}{(\log M)^{A'}}$$

for suitable A' by BV and the multiplicative bounds for λ_d (and $\tau(\ell)$ grows slowly).

5.7. Conclusion of the Sieve Computation

Combining the diagonal main terms and the off-diagonal (BV-controlled) error estimates we obtain

$$S_0 \asymp \frac{M}{(\log R)^2}, \qquad S_1 \asymp \frac{M}{\log M} \cdot \frac{1}{(\log R)^2}, \qquad S_2 \asymp \frac{M}{(\log M)^2} \cdot \frac{1}{(\log R)^2}.$$

In particular with the choice $R = M^\delta$ for sufficiently small $\delta > 0$ (one may take $\delta < 1/2 - \varepsilon$ to satisfy BV), we get $S_0 \asymp M$ up to polylogarithmic factors and $S_2 \gg M/(\log M)^2$ quantitatively. Equivalently, the combinatorial Cauchy-type inequality standard in GPY/Maynard yields a lower bound for the number of $m \leq M$ with both n_1, n_2 prime:

$$\#\{m \leq M : n_1, n_2 \text{ both prime}\} \gg \frac{M}{(\log M)^2}.$$

5.8. Parameter Choice Summary

To make all error terms negligible and to connect with Lemmas 2-4 we fix:

$$z = M^\vartheta \quad (0 < \vartheta < 1/2), \qquad R = M^\delta \quad (0 < \delta < 1/2 - \varepsilon),$$

and impose the sieve-small-primes set $\mathcal{P}$ to satisfy $y \leq R^{1-\delta'}$. With these choices Bombieri–Vinogradov controls the off-diagonal errors and the Selberg optimization yields the stated lower bounds. Finally Lemma 4 removes the contribution of "both composite but with large prime factors" cases and Lemma 3 excludes persistent filling of twin-candidate positions by sifted composites; therefore the GPY mass indeed forces genuine twin primes in quantity $\gg M/(\log M)^2$.

This completes the Selberg–GPY weight analysis.

6. Optimization of the Selberg Weights: Computation of $Q(\lambda)$ and Bounds for $\sum \lambda_d^2$

6.1. Notation and Objective

Let $R \geq 1$ be the sieve level. We consider real coefficients $\{\lambda_d\}_{d \leq R}$, supported on $d \leq R$ and with $\lambda_1 = 1$. Define the quadratic form

$$Q(\lambda) := \sum_{d \leq R} \sum_{e \leq R} \frac{\lambda_d \lambda_e}{\mathrm{lcm}(d,e)} = \sum_{d,e \leq R} \lambda_d \lambda_e \frac{\gcd(d,e)}{de}.$$

The Selberg upper-sieve principle (as used in previous sections) shows that the upper bound on the count of integers coprime to a set of primes up to y inside an interval is controlled by $H \cdot Q(\lambda)$ up to harmless factors, so we seek to choose λ minimizing $Q(\lambda)$ (subject to the normalization $\lambda_1 = 1$ and support). The classical optimization produces near-optimal weights λ_d of the multiplicative form $\mu(d)g(\frac{\log d}{\log R})$ with an explicit smooth profile g; below we derive this and deduce the main estimates used in our sieve.

6.2. Rewriting Q via Multiplicative Convolution

We start with an identity which is convenient for analyzing Q. Using $\mathrm{lcm}(d,e) = de/\gcd(d,e)$:

$$Q(\lambda) = \sum_{d,e \leq R} \lambda_d \lambda_e \frac{\gcd(d,e)}{de} = \sum_{r \leq R} \frac{1}{r} \sum_{\substack{d',e' \leq R/r \\ \gcd(d',e')=1}} \frac{\lambda_{rd'} \lambda_{re'}}{d'e'}.$$

This is obtained by writing $d = rd', e = re'$ with $r = \gcd(d,e)$. (The coprimality condition $\gcd(d',e') = 1$ appears naturally.) The identity is purely combinatorial and will be used to pass to multiplicative transforms.

Define

$$\alpha_r := \sum_{\substack{d \leq R \\ r \mid d}} \frac{\lambda_d}{d/r} = \sum_{d' \leq R/r} \frac{\lambda_{rd'}}{d'}.$$

A short computation using the previous display yields the useful representation

$$Q(\lambda) = \sum_{r \leq R} \frac{1}{r} \left(\sum_{d' \leq R/r} \frac{\lambda_{rd'}}{d'} \right)^2 = \sum_{r \leq R} \frac{\alpha_r^2}{r}.$$

(Indeed expand the square and reorganize to recover the earlier sum. This form is standard in treatments of the Selberg quadratic form.)

6.3. Lagrange Multipliers and the Optimal Structure

We now minimize $Q(\lambda)$ under the linear normalization constraint $\lambda_1 = 1$ (and optionally small regularity/smoothness constraints). A more sophisticated classical formulation is to minimize $Q(\lambda)$ subject to the finite family of linear constraints $\sum_{d \mid n} \lambda_d \geq 1$ for integers n with no small prime factors; however the usual practical approach is to impose only the normalization and then choose λ multiplicative and smooth in $\log d$; this yields nearly optimal weights for the upper-sieve situation.

Working with the α_r variables is convenient. Introduce Lagrange multiplier θ for the constraint that corresponds to normalizing the value at $n = 1$; in the α-frame this produces the linear system

$$\frac{2\alpha_r}{r} = \theta \cdot \beta_r \qquad (r \leq R),$$

for suitable coefficients β_r depending on the chosen normalizing linear functional. Solving formally gives $\alpha_r \propto r\beta_r$ and hence

$$\alpha_r = c \, r\beta_r$$

for some constant c determined by normalization. In multiplicative terms this structure means α_r is multiplicative and smooth in r, and transferring back to λ_d via Möbius inversion yields [10]

$$\lambda_d = \mu(d)\, g\!\left(\frac{\log d}{\log R}\right)$$

for some smooth profile g supported in $[0,1]$. The linear cutoff $g(t) = 1 - t$ (or smoothed variants) is the classical explicit choice producing good and analyzable constants.

We emphasize: the exact solution of the finite quadratic optimization problem gives a discrete λ_d that matches the above multiplicative ansatz to high accuracy. This is standard and is carried out in detail in sieve textbooks (e.g., Halberstam–Richert, Iwaniec–Kowalski, Greaves). Using the multiplicative ansatz is thus justified.

6.4. Classical Explicit Choice and Main Estimates

Take the standard linear profile

$$\lambda_d \;=\; \mu(d)\frac{\log(R/d)}{\log R}\qquad (1 \le d \le R),$$

and $\lambda_d = 0$ for $d > R$. This choice is convenient for explicit computation and suffices for our error goals. We compute the two key quantities:

(i) The quadratic form $Q(\lambda)$.

Using the representation $Q = \sum_{r \le R} \alpha_r^2/r$ and writing $\alpha_r = \sum_{d' \le R/r} \lambda_{rd'}/d'$, one evaluates asymptotically (standard calculation; see Selberg upper-sieve derivation) that

$$Q(\lambda) \;=\; \frac{1}{(\log R)^2}\,\prod_{p \le y}\left(1 - \frac{\omega(p)}{p}\right)\cdot\bigl(1 + o(1)\bigr),$$

where $\omega(p)$ counts the number of residue classes modulo p occupied by the admissible tuple (for the twin-shift $\{0,2\}$ we have $\omega(p) = 2$ for $p \ne 2,3$; primes dividing the modulus 6 are treated separately). The product over primes is exactly the classical singular series $\mathfrak{S}(\mathcal{H})$ (up to local factors at 2 and 3); hence one obtains the familiar scaling

$$Q(\lambda) \;\asymp\; \frac{\mathfrak{S}(\{0,2\})}{(\log R)^2}.$$

In particular there exists an absolute constant $c_Q > 0$ (depending on the shift) such that for large R

$$Q(\lambda) \le c_Q \frac{1}{(\log R)^2}.$$

(ii) The second moment $\sum_{d \le R} \lambda_d^2$.

With the linear profile one checks directly (by standard divisor-sum estimates) that

$$\sum_{d \le R} \lambda_d^2 \;=\; \sum_{d \le R} \mu^2(d)\left(\frac{\log(R/d)}{\log R}\right)^2 \;\ll\; \sum_{d \le R} \frac{1}{d^\varepsilon} \;\ll\; (\log R)^{O(1)}.$$

More concretely, using Mertens' theorem and partial summation one gets

$$\sum_{d \le R} \lambda_d^2 \ll \log R.$$

(With smoothed g one can make this bound even slightly better, but $\ll \log R$ is sufficient for all subsequent estimates.)

6.5. Putting the Estimates to Use in the Sieve

The Selberg upper-sieve in an interval of length H gives (see Lemma 2)

$$\#\{n \in [X, X+H] : \ \gcd(n, P(y)) = 1\} \ \le \ H \cdot Q(\lambda) + O\Big(\sum_{d \le R} \lambda_d^2\Big).$$

Inserting the previous asymptotics $Q(\lambda) \asymp c_Q/(\log R)^2$ and $\sum \lambda_d^2 \ll \log R$ yields

$$\#\{n \in [X, X+H] : \ \gcd(n, P(y)) = 1\} \ \ll \ H\frac{1}{(\log R)^2} + O(\log R).$$

For our parameter regime $R \asymp H^{1/2}$ (or $R = M^\delta$ in the global GPY argument), the main term dominates and the error $\ll \log R$ is negligible in the final counting $\gg M/\log^2 M$.

6.6. Remarks and References

- The displayed asymptotic for $Q(\lambda)$ (leading to the $\frac{1}{(\log R)^2}$ scale times the singular series) is standard: see Halberstam–Richert, Iwaniec–Kowalski, or the exposition in Maynard's paper. The constants are explicit and arise from the integrals involving the profile g; with the linear profile $g(t) = 1 - t$ one obtains the classical factor $1/(\log R)^2$ times the singular series.
- The bound $\sum_{d \le R} \lambda_d^2 \ll \log R$ is elementary and follows from multiplicative sums and Mertens' theorem; smoothing g can replace $\log R$ by a slightly smaller polylog factor, but no essential gain is needed.

7. From the Sieve Bound Conjecture 1 to Infinitely Many Twin Primes

In this section we close the argument by showing that the short-interval sieve bound Conjecture 1, together with the GPY/Maynard mass computation and the large-factor elimination, forces the existence of infinitely many twin primes.

Theorem 3. *Fix small constants $0 < \eta, \varepsilon < 1/10$ and large M. Let $H = M^\eta$, choose $R = M^\delta$ with $0 < \delta < \frac{1}{2} - \varepsilon$, and set $z = M^\vartheta$ with $0 < \vartheta < \frac{1}{2}$. Assume:*

- *(**Sieve bound Conjecture 1 in short intervals**): for $\mathcal{P} = \{p \le y\}$ with $y \le R^{1-\varepsilon}$ we have uniformly for all intervals $I = [X, X+H]$,*

$$\#\{n \in I : \ \gcd(n, P(y)) = 1\} \ \ll \ H \cdot Q(\lambda) + O\Big(\sum_{d \le R} \lambda_d^2\Big) \ \ll \ \frac{H}{(\log R)^2} + O(\log R).$$

(This is Lemma 2 with the optimized Selberg weights of Section §6.)

- *(**Uniform gap hypothesis for $\mathcal{C}_\mathcal{P}$**) Conjecture 1 in the form: the set of indices with gaps ≤ 8 has zero density for large M.*
- *(**Bombieri–Vinogradov**) holds up to moduli $q \le M^{1/2-\varepsilon}$.*

Then there exists $c > 0$ (depending only on the fixed parameters) such that, for all sufficiently large M,

$$\#\big\{m \le M : \ 6m+5, \ 6m+7 \text{ are both prime}\big\} \ \ge \ c\,\frac{M}{(\log M)^2}.$$

In particular, there are infinitely many twin primes.

Proof. Write $n_1 = 6m+5$, $n_2 = 6m+7$ and put

$$\mathcal{A} := \big\{m \le M : \ P^-(n_1) > z, \ P^-(n_2) > z\big\}.$$

Step 1 (GPY/Maynard mass for twin-shift). By the Selberg–GPY analysis of Section §5 (with $R \leq M^{1/2-\varepsilon}$ to enable Bombieri–Vinogradov on the off-diagonal), the standard diagonal/off-diagonal computation gives

$$\#\mathcal{A} \;\geq\; c_0 \,\frac{M}{(\log M)^2} \qquad (M \text{ large}), \tag{2}$$

for some $c_0 > 0$ depending on the choice of weights (the singular series for $\{0,2\}$ is folded into c_0).

Step 2 (remove the "both composite" large-factor cases). Let

$$\mathcal{B} := \{m \in \mathcal{A} : \; n_1, n_2 \text{ are both composite}\}.$$

By Lemma 4 (large-factor bilinear bound), with $z = M^\vartheta$ and $0 < \vartheta < 1/2$,

$$\#\mathcal{B} \;=\; o\big(M/\log^2 M\big). \tag{3}$$

Set $\mathcal{C} := \mathcal{A} \setminus \mathcal{B}$. Then combining (2) and (3) we get

$$\#\mathcal{C} \;\geq\; (c_0 + o(1))\,\frac{M}{(\log M)^2}. \tag{4}$$

By definition of $\mathcal{C}$, for each $m \in \mathcal{C}$ at least one of n_1, n_2 is prime, and the other (if composite) has all prime factors $> z$; in particular the composite one (if any) lies in the sifted-composite set $\mathcal{C}_\mathcal{P}$.

Step 3 (contradiction if there are only finitely many twin primes). Suppose, towards a contradiction, that the number of twin primes up to $6M + 7$ is $o(M/\log^2 M)$. Equivalently,

$$\#\mathcal{T} := \#\{m \leq M : \; n_1, n_2 \text{ both prime}\} = o\big(M/\log^2 M\big).$$

Then, by (4), for *most* $m \in \mathcal{C}$ exactly one of n_1, n_2 is prime, and the other is a sifted composite (all prime factors $> z$). Consider the set

$$\mathcal{S} := \big\{\, n \in [5, 6M+7] : \; n \in \mathcal{C}_\mathcal{P} \text{ and } n \in \{6m+5, 6m+7\} \text{ for some } m \in \mathcal{C} \setminus \mathcal{T} \,\big\}.$$

For any two consecutive indices $m, m+1 \in \mathcal{C} \setminus \mathcal{T}$, the corresponding elements of $\mathcal{S}$ lie in $\{6m+5, 6m+7\} \cup \{6(m+1)+5, 6(m+1)+7\}$, hence the gap between consecutive points of $\mathcal{S}$ is ≤ 8. Thus, on any run of consecutive m's drawn from $\mathcal{C} \setminus \mathcal{T}$, the adjacent gaps among the ordered elements of $\mathcal{S}$ are ≤ 8.

By a short-interval covering argument, (4) implies that there exist $\gg M/H$ disjoint intervals of the form $[m_0, m_0 + H']$ with $H' \asymp H$ in the m-variable (equivalently length $\asymp 6H$ in the n-variable) that each contain $\gg H/(\log M)^2$ elements of $\mathcal{C} \setminus \mathcal{T}$. In each such interval, deleting at most $O(1)$ boundary indices leaves a sub-run of consecutive m's of length $\gg H/(\log M)^2$. Therefore, within each such short interval in the n-line we produce $\gg H/(\log M)^2$ *consecutive gaps* ≤ 8 among elements of $\mathcal{C}_\mathcal{P}$.

Summing over the $\gg M/H$ disjoint short intervals yields

$$\#\big\{n \leq 6M+7 : \; g_n(\mathcal{P}) \leq 8\big\} \;\gg\; \frac{M}{(\log M)^2}.$$

But this contradicts the uniform gap in Conjecture 1. Hence our assumption was false.

Step 4 (conclusion). Therefore

$$\#\mathcal{T} \;=\; \#\{m \leq M : \; n_1, n_2 \text{ both prime}\} \;\geq\; c\,\frac{M}{(\log M)^2}$$

for some $c > 0$ and all sufficiently large M. Letting $M \to \infty$ gives infinitely many twin primes.

This completes the proof of Theorem 2. $\quad\square$

Remark 3 (Choice of parameters). *The proof only uses that: (i) GPY/Maynard gives $\#\mathcal{A} \gg M / \log^2 M$ with level of distribution up to $R \leq M^{1/2-\varepsilon}$ (Bombieri–Vinogradov); (ii) Lemma 4 removes the "both composite with large factors" contribution; (iii) Lemma 3/Conjecture 1 forbids the persistence of short adjacent gaps ≤ 8 in $\mathcal{C}_\mathcal{P}$. Any fixed $0 < \eta, \delta, \vartheta < 1/2$ compatible with these inputs will do.*

References

1. T. Tao. The Elliott–Halberstam conjecture implies the Vinogradov least quadratic nonresidue conjecture. *Algebra & Number Theory*, 9(4):1005–1034, 2015.
2. D. A. Goldston, Y. Motohashi, J. Pintz and C. Y. Yildirim. Small gaps between primes exist. *Proc. Japan Acad. Ser. A Math. Sci.*, 82(4):61–65, 2006.
3. Y. Zhang. Bounded gaps between primes. *Annals of Mathematics*, 179(3):1121–1174, 2014.
4. D. H. J. Polymath. The "bounded gaps between primes" Polymath project-a retrospective, 2014. Preprint available: arXiv:1409.8361.
5. A. F. Gocgen. Gocgen Approach for Bounded Gaps Between Odd Composite Numbers, 2024. Preprint available: preprints202401.1533
6. G. Greaves. Selberg's Upper Bound Method. In *Sieves in Number Theory*, pages 41–70, 2001
7. H. H. Wu and S. Wu. Various proofs of the Cauchy-Schwarz inequality. *Octogon mathematical magazine*, 17(1), 221–229, 2009.
8. H. Ogasawara. The multivariate Markov and multiple Chebyshev inequalities. *Communications in Statistics-Theory and Methods*, 49(2), 441–453, 2020.
9. J. M. Deshouillers and H. Iwaniec. An additive divisor problem. *Journal of the London Mathematical Society*, 2(1), 1–14, 1982.
10. E. A. Bender and J. R. Goldman. On the applications of Möbius inversion in combinatorial analysis. *The American Mathematical Monthly*, 82(8), 789–803, 1975.