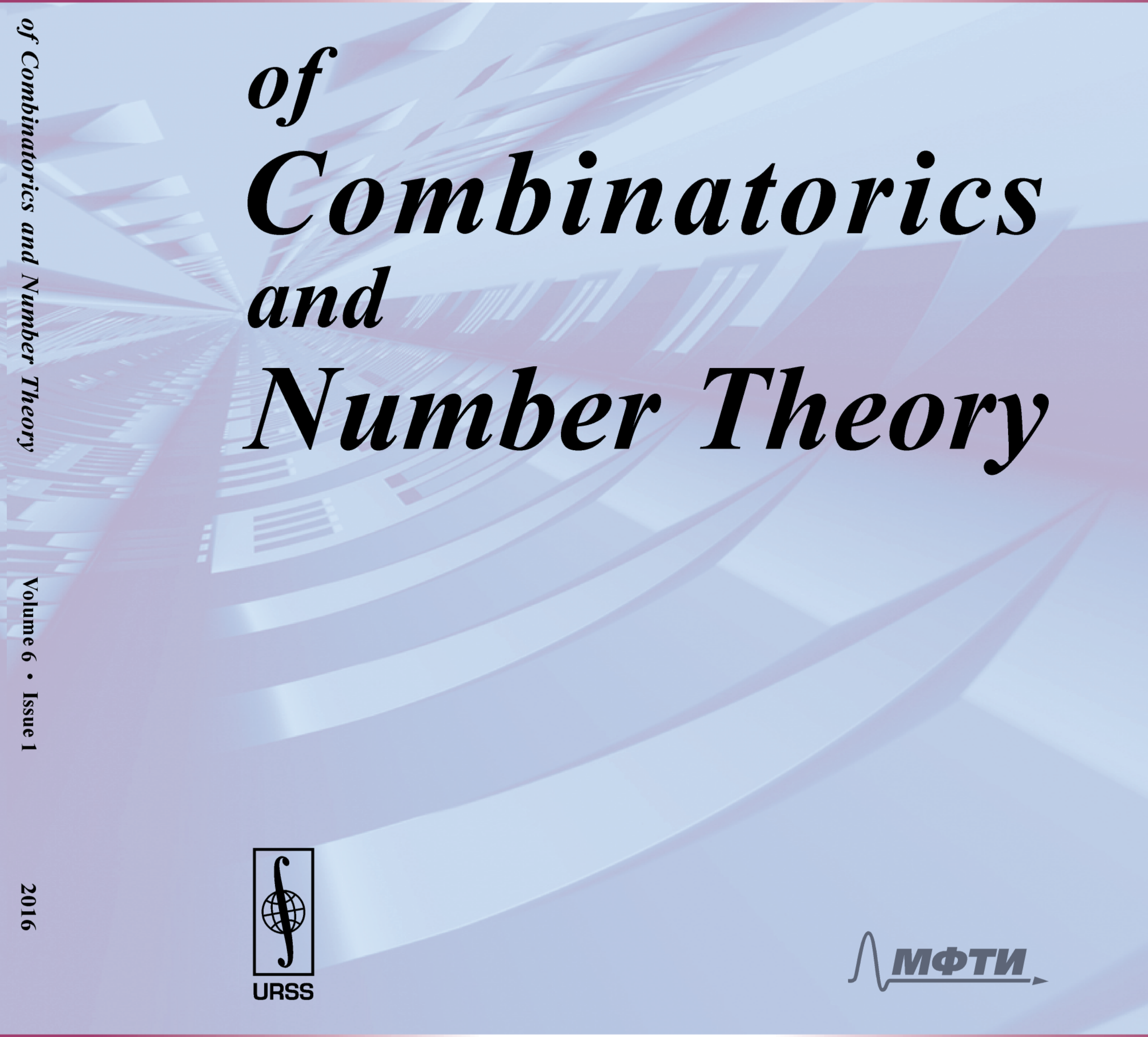
Reprint from

ISSN 2220-5438

# Moscow Journal of Combinatorics and Number Theory



URSS

МФТИ

Volume 6 • Issue 1

2016

Moscow Journal of Combinatorics and Number Theory Volume 6 • Issue 1 2016

URSS

Moscow Journal
of Combinatorics
and Number Theory
2016, vol. 6, iss. 1,
pp. 3–24 [3–24]

# On the correlations, Selberg integral and symmetry of sieve functions in short intervals, III

Giovanni Coppola (Avellino), Maurizio Laporta (Napoli)

**Abstract:** An arithmetic function $f$ is called a sieve function of range $Q$, if it is the convolution product of the constantly 1 function and $g$ such that $g(q) \ll_\varepsilon q^\varepsilon$, $\forall\, \varepsilon > 0$, for $q \le Q$, and $g(q) = 0$ for $q > Q$. Here we establish a new result on the autocorrelation of $f$ by using a famous theorem on bilinear forms of Kloosterman fractions by Duke, Friedlander and Iwaniec. In particular, for such correlations we obtain non-trivial asymptotic formulæ that are actually unreachable by the standard approach of the distribution of $f$ in the arithmetic progressions. Moreover, we apply our asymptotic formulæ to obtain new bounds for the so-called Selberg integral and symmetry integral of $f$, which are basic tools for the study of the distribution of $f$ in short intervals.

**Keywords:** correlations, Selberg integral, short intervals

**AMS Subject Classification:** 11N37, 11N25

**Received:** 14.04.2015

## 1. Introduction and statement of the results

A basic tool for the study of the distribution of an arithmetic function $f : \mathbb{N} \to \mathbb{C}$ in short intervals is the so-called *Selberg integral* of $f$, that is

$$J_f(N,h) \overset{def}{=} \int_N^{2N} \Big| \sum_{x<n\le x+h} f(n) - M_f(x,h) \Big|^2 \mathrm{d}x,$$

where $M_f(x,H)$ is the (short interval) *mean-value* of $f$ and $h, N \in \mathbb{N}$ are such that $h = o(N)$, as $N \to \infty$. Indeed, non-trivial bounds for $J_f(N,h)$ might yield results on the distribution of $f$ in *almost all* the short intervals $(x, x+h]$, i.e. for all $x \in [N, 2N] \cap \mathbb{N}$ with $o(N)$ exceptions. On the other side, the symmetry properties of $f$ in almost all the short intervals are linked to the *symmetry integral* of $f$,

$$I_f(N,h) \overset{def}{=} \int_N^{2N} \Big| \sum_{|n-x|\le h} \operatorname{sgn}(n-x) f(n) \Big|^2 \, \mathrm{d}x,$$

where the *sign* function is defined as $\operatorname{sgn}(0) \overset{def}{=} 0$, and $\operatorname{sgn}(t) \overset{def}{=} |t|/t$ if $t \neq 0$.

The aim of the present paper is to continue the study of $J_f(N,h)$ and $I_f(N,h)$, started in [2] and considered also in [3–5], for a *sieve function* $f$ of *range* $Q \ll N$, meaning that its *Eratosthenes transform* $g \overset{def}{=} f * \mu$ is supported in $[1, Q]$ and $g$ is *essentially bounded*, i.e. $g(q) \ll_\varepsilon q^\varepsilon$ $(\forall\, \varepsilon > 0)$. Here, we recall that $\ll$ is Vinogradov's notation, synonymous to Landau's $O$-notation. In particular, $\ll_\varepsilon$ means that the implicit constant might depend on an arbitrarily small $\varepsilon > 0$, which might change at each occurrence. Since by the Möbius inversion formula one has

$$f(n) = (g * \mathbf{1})(n) = \sum_{\substack{d|n \\ d\le Q}} g(d),$$

where $\mathbf{1}(n) = 1$ for all $n \in \mathbb{N}$, then $g$ is essentially bounded if and only if so is $f$. Moreover, since

$$\frac{1}{x}\sum_{n\le x} f(n) = \frac{1}{x}\sum_{d\le Q} g(d)\Big[\frac{x}{d}\Big] = \sum_{d\le Q}\frac{g(d)}{d} + O\Big(\frac{1}{x}\sum_{d\le Q}|g(d)|\Big),$$

where $[t]$ is the *integer part* of $t \in \mathbb{R}$ (hereafter, in sums over positive integers like $\sum_{a\le x} 1$ it is implicit that $a \ge 1$), we expect the (short interval) mean-value of $f$ to be independent of $x$, namely given by (see [5] for further comments)

$$M_f(h) \overset{def}{=} h \sum_{d\le Q} \frac{g(d)}{d}.$$

Before stating our results, let us introduce some auxiliary notation and convention. When $h = [N^\theta]$, with $\theta \in [0,1]$, we refer to $\theta$ as the *width* of the short interval $[x-h, x+h]$ or $(x, x+h]$. We adopt the convention that $\theta < \theta_0$ (resp. $\theta > \theta_0$)

means $\theta \le \theta_0 - \delta$ (resp. $\theta \ge \theta_0 + \delta$) for some absolute constant $\delta > 0$. Further, we say that $f$ has *level* $\lambda \in [0,1]$ if it is a sieve function of range $Q = [N^{\lambda}]$, and for $\lambda$ we adopt the same convention as for the width. Finally, given the arithmetic functions $\phi_1$ and $\phi_2$, we write $\phi_1(n) \lll \phi_2(n)$ to mean that $\phi_1(n) \ll_{\varepsilon} n^{\varepsilon}\phi_2(n)\ \forall\, \varepsilon > 0$ (as $n \to \infty$).

THEOREM 1. *Fix a small* $\delta > 0$. *If* $f : \mathbb{N} \to \mathbb{R}$ *has level* $\lambda \in (1/2, 1)$, *then*

$$J_f(N,h) \lll Nh + N^{\delta}Q^{2-\Delta}h^2 + N^{1-2\delta/3}h^2 + Qh^2,$$

$$I_f(N,h) \lll Nh + N^{\delta}Q^{2-\Delta}h^2 + N^{1-2\delta/3}h^2,$$

*as* $N \to \infty$, *where* $Q = [N^{\lambda}]$, $h = [N^{\theta}]$ *with* $\theta \in (0, 1/2)$, *and* $\Delta = 1/48$.

**Remark.** In [2] the above inequalities hold with $\Delta = 0$ (for a small $h$). In particular, such inequalities yield the non-trivial bound $N^{1-\varepsilon}h^2$ for both integrals $J_f(N,h)$ and $I_f(N,h)$ with $f$ of level $\lambda < (1+\theta)/2$ and for any width $\theta \in (0,1)$ (see Corollary 1.1 of [2]), whereas the previous theorem holds for $\lambda > 1/2$. By combining Theorem 1 above with the results given by Corollary 1.1 of [2] we immediately obtain the following non-trivial bounds that however improve on [2] estimates only in very short intervals, namely $\theta \in (0, 1/95)$.

COROLLARY 1. *Let* $\theta \in (0, 1/2)$ *be fixed. If* $f : \mathbb{N} \to \mathbb{R}$ *has level* $\lambda \in \big(0, \max\{(1+\theta)/2, 48/95\}\big)$, *then there exists* $\varepsilon_0 = \varepsilon_0(\theta, \lambda) > 0$ *such that*

$$J_f(N,h) \ll_{\varepsilon_0} N^{1-\varepsilon_0}h^2, \qquad I_f(N,h) \ll_{\varepsilon_0} N^{1-\varepsilon_0}h^2,$$

*as* $N \to \infty$, *where* $h = [N^{\theta}]$.

Note that $48/95 = 1/2 + 1/190 > (1+\theta)/2$ if and only if $\theta < 1/95$. Unlike [2], we derive Theorem 1 from a slight generalization concerning the *mixed* Selberg integral and the *mixed* symmetry integral of the sieve functions $f_1$ and $f_2$, namely

$$J_{f_1,f_2}(N,h) \overset{def}{=} \int_N^{2N} \prod_{c=1,2} \Big( \sum_{x<n\le x+h} f_c(n) - M_{f_c}(h) \Big)\, \mathrm{d}x,$$

$$I_{f_1,f_2}(N,h) \overset{def}{=} \int_N^{2N} \prod_{c=1,2} \Big( \sum_{|n-x|\le h} \mathrm{sgn}(n-x) f_c(n) \Big) \, \mathrm{d}x,$$

where as before we set

$$M_{f_c}(h) \overset{def}{=} h \sum_{d\le Q_c} \frac{g_c(d)}{d}, \quad (c=1,2),$$

provided that $g_c$ and $Q_c$ are the Eratosthenes transform and the range of $f_c$, respectively.

Theorem 2. *Fix a small $\delta > 0$. If for each $c = 1, 2$ the real sieve function $f_c$ has level $\lambda_c \in (1/2, 1)$ with $\lambda_1 \ge \lambda_2$, then*

$$J_{f_1,f_2}(N,h) \lll Nh + N^{\delta} Q_1^{53/48} Q_2^{7/8} h^2 + N^{1-2\delta/3} h^2 + Q_1 h^2,$$

$$I_{f_1,f_2}(N,h) \lll Nh + N^{\delta} Q_1^{53/48} Q_2^{7/8} h^2 + N^{1-2\delta/3} h^2,$$

*where $Q_c = [N^{\lambda_c}]$ and $h = [N^{\theta}]$ with $\theta \in (0, 1/2)$.*

It is plain that Theorem 1 follows immediately by taking $\lambda_1 = \lambda_2 = \lambda$, $Q_1 = Q_2$, and $f_1 = f_2 = f$ in Theorem 2. On the other side, since Theorem 1.1 and Corollary 1.1 of [2] can be easily extended to $J_{f_1,f_2}(N,h)$ and $I_{f_1,f_2}(N,h)$, then we can combine such a generalization with Theorem 2 to get the following immediate consequence.

Corollary 2. *Let $\theta \in (0, 1/2)$ be fixed and let $\lambda_1 \ge \lambda_2 > 0$ be such that $\lambda_1 + \lambda_2 < 1$ or $53\lambda_1 + 42\lambda_2 < 48$. If for each $c = 1, 2$ the real sieve function $f_c$ has level $\lambda_c$, then there exists $\varepsilon_0 = \varepsilon_0(\theta, \lambda_1, \lambda_2) > 0$ such that*

$$J_{f_1,f_2}(N,h) \ll_{\varepsilon_0} N^{1-\varepsilon_0} h^2, \qquad I_{f_1,f_2}(N,h) \ll_{\varepsilon_0} N^{1-\varepsilon_0} h^2,$$

*as $N \to \infty$, where $h = [N^{\theta}]$.*

Summarizing, we need to prove only Theorem 2 and this is accomplished in §4. To this end, we premise a short section on some further notation and basic formulæ, where we introduce the crucial auxiliary function $\mathcal{R}(a)$ in terms of the first Bernoulli periodic function. In §3 we give the necessary lemmata for Theorem 2.

The first lemma is a famous theorem of Duke, Friedlander and Iwaniec and it is the novelty of the present approach to estimating $\mathcal{R}(a)$. Such an estimate is the theme of the second lemma. The link to $J_{f_1,f_2}(N,h)$ and $I_{f_1,f_2}(N,h)$ is provided by the *correlations* of the sieve functions $f_1, f_2$ for which the third and last lemma gives a formula, with an error term taken under control by the new bound of $\mathcal{R}(a)$. We conclude the paper with a section of further comments and with an appendix including the proof of the Fourier expansion of the first Bernoulli periodic function on the rational numbers.

## 2. Some further notation and recurrent properties

As usual in number theory, $(m,n)$ denotes the greatest common divisor of integers $m$ and $n$. Although $(x,y)$ denotes also the pair with coordinates $x,y$ or the open interval with real endpoints $x,y$, the meaning will be evident from the context. For the same sake of brevity, we use to write $n \equiv a\ (q)$ instead of $n \equiv a \pmod q$. Moreover, we set $e(\alpha) \overset{def}{=} e^{2\pi i\alpha}\ \forall\,\alpha \in \mathbb{R}$ and $e_q(n) \overset{def}{=} e(n/q)\ \forall\,(n,q) \in \mathbb{Z}\times\mathbb{N}$. The *distance* of $\alpha \in \mathbb{R}$ from the integers is $\|\alpha\| \overset{def}{=} \min\limits_{n\in\mathbb{Z}} |\alpha - n|$ and $\{\alpha\} \overset{def}{=} \alpha - [\alpha]$ is its *fractional part*. For the main variable $N$ we set $L \overset{def}{=} \log N$.

Without further references, throughout the paper we will appeal to the well-known inequalities

$$\sum_{V_1<v\le V_2} e(v\alpha) \ll \min\left(V_2 - V_1, \frac{1}{\|\alpha\|}\right), \qquad \sum_{d|t} 1 \ll t^{\varepsilon}\ (\forall\, t \in \mathbb{N}, \forall\, \varepsilon > 0).$$

Let us recall that the *first Bernoulli periodic function* is defined as

$$\mathcal{B}_1(\alpha) \overset{def}{=} \begin{cases} \{\alpha\} - 1/2 & \text{if } \alpha \in \mathbb{R}\backslash\mathbb{Z}, \\ 0 & \text{otherwise,} \end{cases}$$

whose (finite) Fourier expansion on the rational numbers is given by (see the Appendix for the proof)

$$\mathcal{B}_1\left(\frac{n}{q}\right) = -\frac{1}{q}\sum_{j\le\frac{q}{2}} \cot\frac{\pi j}{q}\,\sin\frac{2\pi jn}{q} \qquad \forall\,(n,q) \in \mathbb{Z}\times\mathbb{N}\setminus\{1\}. \tag{2.1}$$

See that $\cot(\pi/2+k\pi)=0$, $\forall\, k\in\mathbb{Z}$. Such expansion is interpreted as $\mathcal{B}_1=0$ when $q=1$. It is easy to see that, for any $\alpha\in(0,+\infty)\setminus\mathbb{N}$, $d,q\in\mathbb{N}$ one has

$$\#\{m\in(\alpha,2\alpha]:\ m\equiv d\,(q)\}=\frac{[2\alpha]-[\alpha]}{q}+$$
$$+\begin{cases}\mathcal{B}_1\Big(\dfrac{[\alpha]-d}{q}\Big)-\mathcal{B}_1\Big(\dfrac{[2\alpha]-d}{q}\Big) & \text{if } q\not|\,[c\alpha]-\mathrm{d} \text{ for } \mathrm{c}=1,2,\\ O(1) & \text{otherwise.}\end{cases}\tag{2.2}$$

Given the functions $g_1,g_2$ supported in $[1,Q_1]$, $[1,Q_2]$, respectively, for all $a\in\mathbb{Z}\setminus\{0\}$ we set

$$\mathcal{R}(a)\overset{def}{=}\sum_{\ell|a}\sum_{q_1\sim\frac{Q_1}{\ell}}g_1(\ell q_1)\sum_{\substack{q_2\sim\frac{Q_2}{\ell}\\(q_1,q_2)=1}}g_2(\ell q_2)\sum_{c=1,2}(-1)^{c+1}\Big(\mathcal{B}_1\Big(\frac{[\alpha_c]+\overline{q}_1b}{q_2}\Big)+$$
$$+\mathcal{B}_1\Big(\frac{[\alpha_c]-\overline{q}_1b}{q_2}\Big)\Big),$$

where $x\sim X$ means that $x\in(X,2X]\cap\mathbb{N}$, the integer $\overline{q}_1\in[1,q_2]$ is defined by $\overline{q}_1q_1\equiv1\ (\mathrm{mod}\,q_2)$ when $(q_1,q_2)=1$, and we set $b\overset{def}{=}|a|/\ell$, $\alpha_c\overset{def}{=}cN/\ell q_1$. We explicitly remark that $\mathcal{R}(a)$ depends also on $g_1,g_2$ and $N$. Note that $Q_c\ge|a|$ ensures $(Q_c/\ell,2Q_c/\ell]\cap\mathbb{N}\ne\varnothing$ for every $\ell|a$. On the other side, $Q_1\ll|a|$ yields $\mathcal{R}(a)\lll|a|Q_2$, which in turn becomes $\mathcal{R}(a)\lll N^{1-\delta}$ when we assume also $Q_2=o(N^{1-\delta}/|a|)$ (the same property holds by interchanging the roles of $Q_1$ and $Q_2$).

## 3. Lemmata

The first lemma comes in a straightforward way from Theorem 2 of [6].

Lemma 1. *Let $N,Q_1,Q_2\in\mathbb{N}$ and $k\in\mathbb{Z}\setminus\{0\}$ such that $Q_1,Q_2\le N$ and $k\ll Q_1Q_2$, as $Q_1,Q_2\to\infty$. If $g_1,g_2:\mathbb{N}\to\mathbb{R}$ are essentially bounded and supported in $[Q_1,2Q_1]$ and $[Q_2,2Q_2]$, respectively, then*

$$\sum_{q_1\sim Q_1}g_1(q_1)\sum_{\substack{q_2\sim Q_2\\(q_1,q_2)=1}}g_2(q_2)e_{q_2}(k\,\overline{q}_1)\lll(Q_1Q_2)^{\frac{7}{8}}(Q_1+Q_2)^{\frac{11}{48}}.$$

The first part of the next lemma is basically Lemma B of [3], reformulated for a pair of essentially bounded functions with support in a bounded interval. Here we take this opportunity to give a more detailed proof. The second part is where we apply the previous lemma and it constitutes the novelty of the method, in comparison with [2] and [3].

LEMMA 2. *Fix a sufficiently small* $\delta > 0$. *Let* $N, Q_1, Q_2 \in \mathbb{N}$ *and* $a \in \mathbb{Z} \setminus \{0\}$ *such that* $Q_1 Q_2 \gg N^{1-2\delta/3}$, $Q_2 \ll Q_1 = o(N^{1-\delta})$ *and* $|a| = o(N)$, *as* $N \to \infty$. *If* $g_1, g_2 : \mathbb{N} \to \mathbb{R}$ *are essentially bounded and supported in* $[Q_1, 2Q_1]$ *and* $[Q_2, 2Q_2]$, *respectively, then for every* $\varepsilon > 0$ *one has*

*I)*
$$\mathcal{R}(a) = \frac{2}{\pi} \sum_{\ell | a} \sum_{q_1 \sim Q_1/\ell} g_1(\ell q_1) \sum_{\substack{q_2 \sim Q_2/\ell \\ (q_1, q_2) = 1}} g_2(\ell q_2) \sum_{j \le J} \frac{\Delta_j}{j} + O_\varepsilon(N^{1-\delta+\varepsilon}),$$

*where* $J = J(\ell, q_1, q_2, N, \delta) \overset{def}{=} [\ell q_1 q_2 N^{\delta-1}]$ *and*

$$\Delta_j = \Delta_j(\ell, q_1, q_2, a, N) \overset{def}{=} \Big( \sin \frac{2\pi[2N/\ell q_1]j}{q_2} - \sin \frac{2\pi[N/\ell q_1]j}{q_2} \Big) \cos \frac{2\pi\, \overline{q}_1 |a| j}{\ell q_2}.$$

*Also,*

*II)*
$$\mathcal{R}(a) \ll_\varepsilon N^\varepsilon (N^\delta Q_1^{53/48} Q_2^{7/8} + N^{1-\delta}).$$

PROOF. I) For every $\ell | a$, let us set $b \overset{def}{=} |a|/\ell \in \mathbb{N}$ and note that $J = [\ell q_1 q_2 N^{\delta-1}] = o(Q_2/\ell)$, while $M \overset{def}{=} Q_1 Q_2 N^{\delta-1} L^{-1} \ge 1$ from the hypothesis $Q_1 Q_2 \gg N^{1-2\delta/3}$. Moreover, since $\mathcal{B}_1 \ll 1$, the contribution to $\mathcal{R}(a)$ from all $\ell | a$ such that $\ell > M$ is trivially

$$\lll Q_1 Q_2 \sum_{\substack{\ell | a \\ \ell > M}} \frac{1}{\ell^2} \lll N^{1-\delta}.$$

Thus, let us consider the sum over $\ell \le M$ such that $\ell | a$. Together with $q_1 \sim Q_1/\ell$ and $q_2 \sim Q_2/\ell$, the condition $\ell \le M$ yields $J \to \infty$, as $N \to \infty$. By using formula (2.1) and the identity $\sin(x-w) - \sin(y-w) + \sin(x+w) - \sin(y+w) = 2(\sin x - \sin y)\cos w$, it is readily seen that

$$\mathcal{R}(a) = O_\varepsilon\Big(N^{1-\delta+\varepsilon}\Big) + 2 \sum_{\substack{\ell | a \\ \ell \le M}} \sum_{q_1 \sim Q_1/\ell} g_1(\ell q_1) \sum_{\substack{q_2 \sim Q_2/\ell \\ (q_1, q_2) = 1}} \frac{g_2(\ell q_2)}{q_2} \sum_{j \le q_2/2} \Delta_j \cot \frac{\pi j}{q_2}. \quad (3.1)$$

Let us split the last sum as

$$\sum_{j\le q_2/2} \Delta_j \cot\frac{\pi j}{q_2} = \sum_{j\le J} \Delta_j \cot\frac{\pi j}{q_2} + \sum_{J<j\le q_2/2} \Delta_j \cot\frac{\pi j}{q_2} = \mathcal{D}_1 + \mathcal{D}_2, \text{ say,}$$

and first evaluate the contribution to (3.1) from $\mathcal{D}_2$. To this end, note that

$$\begin{aligned}\Delta_j &= \frac{1}{2i}\cos\Big(\frac{2\pi\overline{q}_1 bj}{q_2}\Big)\sum_{c=1,2}(-1)^c\Big(e_{q_2}(j[\alpha_c]) - e_{q_2}(-j[\alpha_c])\Big) = \\ &= \frac{1}{4i}\Big(\mathcal{E}_{2,j} - \mathcal{E}_{1,j} + \overline{\mathcal{E}}_{1,j} - \overline{\mathcal{E}}_{2,j}\Big),\end{aligned}$$

where for $c = 1, 2$ we set $\alpha_c \overset{def}{=} cN/\ell q_1$ and

$$\mathcal{E}_{c,j} = \mathcal{E}_{c,j}(\ell, q_1, q_2, b, N) \overset{def}{=} e_{q_2}\big(j([\alpha_c] + \overline{q}_1 b)\big) + e_{q_2}\big(j([\alpha_c] - \overline{q}_1 b)\big).$$

Since $\cot(\pi/2) = 0$, by partial summation we can write

$$\begin{aligned}\mathcal{D}_2 &\ll \Big|\int_J^{q_2/2}\Big(\sum_{J<j\le v}\Delta_j\Big)\Big(\frac{\mathrm{d}}{\mathrm{d}v}\cot\frac{\pi v}{q_2}\Big)\mathrm{d}v\Big| \ll \\ &\ll q_2\int_J^{q_2/2}\frac{1}{v^2}\Big|\sum_{J<j\le v}\Big(\mathcal{E}_{2,j} - \mathcal{E}_{1,j} + \overline{\mathcal{E}}_{1,j} - \overline{\mathcal{E}}_{2,j}\Big)\Big|\mathrm{d}v.\end{aligned}$$

The contribution from $q_2 | q_1[\alpha_c] \pm b$ is trivially $\ll q_2\Big|\int_J^{q_2/2}\frac{\mathrm{d}v}{v}\Big| \lll q_2$, which in turn contributes to (3.1) as

$$\begin{aligned}&\lll \sum_{\ell|a}\sum_{q_1\sim Q_1/\ell}\#\{q_2\sim Q_2/\ell :\ (q_2, q_1) = 1 \text{ and } q_2|q_1[\alpha_c]\pm b\} \lll \\ &\lll \sum_{\ell|a}\Big(\sum_{\substack{q_1\sim Q_1/\ell \\ q_1\nmid b}}\ \sum_{\substack{(q_2,q_1)=1 \\ q_2|q_1[\alpha_c]\pm b}} 1 + \frac{Q_2}{\ell}\sum_{q_1|b} 1\Big) \lll (Q_1 + Q_2)\sum_{\ell|a}\frac{1}{\ell} \lll N^{1-\delta},\end{aligned}$$

where we have taken into account: $q_1[\alpha_c] \pm b \le (cN \pm |a|)/\ell \ll N/\ell$, $0 < |a| \ll N$ and $Q_2 \ll Q_1 = o(N^{1-\delta})$. On the other side, the contribution to $\mathcal{D}_2$ from $q_1[\alpha_c] \pm b \not\equiv 0 \pmod{q_2}$ amounts to

$$\ll \frac{q_2}{J} \sum_{c=1,2} \Big( \Big\| \frac{[\alpha_c] + \overline{q}_1 b}{q_2} \Big\|^{-1} + \Big\| \frac{[\alpha_c] - \overline{q}_1 b}{q_2} \Big\|^{-1} \Big).$$

Now, observe that $q_1[\alpha_c] \pm b \not\equiv 0 \pmod{q_2}$, with $(q_1, q_2) = 1$, yields $q_1[\alpha_c] \pm b \equiv rq_1 \pmod{q_2}$ for some $r$ such that $1 \le |r| \le q_2/2$, i.e.

$$\Big\| \frac{[\alpha_c] \pm \overline{q}_1 b}{q_2} \Big\| = \frac{|r|}{q_2}.$$

Therefore, since $J = [\ell q_1 q_2 N^{\delta - 1}]$, the contribution to (3.1) from $q_1[N/\ell q_1] + b \not\equiv 0 \pmod{q_2}$ through $\mathcal{D}_2$ is

$$\lll \frac{N^{1-\delta}}{Q_1 Q_2} \sum_{\ell | a} \ell \sum_{q_1 \sim \frac{Q_1}{\ell}} \sum_{\substack{q_2 \sim Q_2/\ell \\ (q_2, q_1) = 1 \\ q_1 [N/\ell q_1] + b \not\equiv 0\, (q_2)}} \Big\| \frac{[N/\ell q_1] + \overline{q}_1 b}{q_2} \Big\|^{-1} \lll$$

$$\lll \frac{N^{1-\delta}}{Q_1 Q_2} \sum_{\ell | a} \ell \sum_{q_2 \sim \frac{Q_2}{\ell}} q_2 \sum_{1 \le |r| \le \frac{q_2}{2}} \frac{1}{|r|} \sum_{\substack{q_1 \sim Q_1/\ell \\ (q_1, q_2) = 1 \\ q_1 [N/\ell q_1] + b \equiv r q_1\, (q_2)}} 1 \lll$$

$$\lll \frac{N^{1-\delta}}{Q_1 Q_2} \sum_{\ell | a} \ell \sum_{q_1 \sim \frac{Q_1}{\ell}} \sum_{1 \le |r| \le \frac{Q_2}{\ell}} \frac{1}{|r|} \sum_{\substack{q_2 \sim Q_2/\ell \\ q_2 | ([N/\ell q_1] - r) q_1 + b}} q_2 \lll$$

$$\lll \frac{N^{1-\delta} Q_2}{Q_1} \sum_{\ell | a} \frac{1}{\ell} \sum_{\substack{q_1 \sim \frac{Q_1}{\ell} \\ q_1 | b}} \sum_{\substack{1 \le |r| \le \frac{Q_2}{\ell} \\ ([N/\ell q_1] - r) q_1 + b = 0}} \frac{1}{|r|} +$$

$$+ \frac{N^{1-\delta}}{Q_1} \sum_{\ell | a} \frac{1}{\ell^{\varepsilon}} \sum_{q_1 \sim \frac{Q_1}{\ell}} \sum_{\substack{1 \le |r| \le \frac{Q_2}{\ell} \\ ([N/\ell q_1] - r) q_1 + b \ne 0}} \frac{1}{|r|} \lll$$

$$\lll \frac{N^{1-\delta} Q_2}{Q_1} + N^{1-\delta} \lll N^{1-\delta}.$$

The same bound holds in the other three cases $q_1[2N/\ell q_1] \pm b \not\equiv 0 \pmod{q_2}$, $q_1[N/\ell q_1] - b \not\equiv 0 \pmod{q_2}$.

Now, we turn our attention to $\mathcal{D}_1$. By using the expansion of the cotangent function in power series, for a fixed $K > 1$ we can write

$$\mathcal{D}_1 = \frac{q_2}{\pi} \sum_{j \le J} \frac{\Delta_j}{j} + \sum_{n=1}^{K-1} \frac{a_n}{q_2^n} \sum_{j \le J} j^n \Delta_j + O\Big(J\Big(\frac{J}{q_2}\Big)^K\Big),$$

where the coefficients $a_n \ll 1$ are given in terms of the Bernoulli numbers (see [10], Appendix B, exercise 11 and formula (B.20)). Note that the first sum gives the main term of the stated formula for $\mathcal{R}(a)$, whereas the $O$-term is $\ll q_2(Q_1/N^{1-\delta})^{K+1} \ll q_2$ from the hypothesis $Q_1 = o(N^{1-\delta})$. In order to see that also the sum over $n$ contributes to (3.1) as a remainder term, we first apply partial summation to write

$$\sum_{n=1}^{K-1} \frac{a_n}{q_2^n} \sum_{j \le J} j^n \Delta_j \ll \Big( \sum_{n=1}^{K-1} \frac{J^n}{q_2^n} \Big) \max_{v \le J} \Big| \sum_{j \le v} \Delta_j \Big|.$$

Then we observe that the argument previously used for $\mathcal{D}_2$ applies here, because it turns out that

$$\sum_{n=1}^{K-1} \frac{J^n}{q_2^n} \ll \sum_{n=1}^{K-1} \Big( \frac{Q_1}{N^{1-\delta}} \Big)^n \ll 1.$$

The formula I) is completely proved.

In order to prove the inequality stated in II), from what we have seen in the proof of I), it is plain that we may confine to consider only

$$\mathcal{R}_c(a) \overset{def}{=} \sum_{\substack{\ell | a \\ \ell \le M}} \sum_{q_1 \sim \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2) = 1}} g_2(\ell q_2) \sum_{j \le 4\frac{ML}{\ell}} \frac{\Sigma_j(c)}{j}, \quad (c = 1, 2),$$

where $M = Q_1 Q_2 N^{\delta - 1} L^{-1}$ is defined as above, and

$$\Sigma_j(c) \overset{def}{=} \sin \frac{2\pi [\alpha_c] j}{q_2} \cos \frac{2\pi \overline{q}_1 b j}{q_2}, \quad \text{with } \alpha_c = \frac{cN}{\ell q_1}, \ b = \frac{|a|}{\ell}.$$

Thus, we have to show that $\mathcal{R}_c(a) \lll N^{\delta} Q_1^{53/48} Q_2^{7/8} + N^{1-\delta}$, $\forall\, c = 1, 2$. To this end, we write

$$\begin{aligned}\Sigma_j(c) = \sin\frac{2\pi\alpha_c j}{q_2}\cos\frac{2\pi\overline{q}_1 bj}{q_2} - \\ &- \Big(1-\cos\frac{2\pi\{\alpha_c\}j}{q_2}\Big)\sin\frac{2\pi\alpha_c j}{q_2}\cos\frac{2\pi\overline{q}_1 bj}{q_2} - \\ &- \sin\frac{2\pi\{\alpha_c\}j}{q_2}\cos\frac{2\pi\alpha_c j}{q_2}\cos\frac{2\pi\overline{q}_1 bj}{q_2} = \\ &= \Sigma_j^{(0)}(c) - \Sigma_j^{(1)}(c) - \Sigma_j^{(2)}(c), \text{ say.}\end{aligned}$$

Accordingly we have $\mathcal{R}_c(a) = \mathcal{R}_c^{(0)}(a) - \mathcal{R}_c^{(1)}(a) - \mathcal{R}_c^{(2)}(a)$ with

$$\mathcal{R}_c^{(\nu)}(a) \overset{def}{=} \sum_{\substack{\ell \mid a \\ \ell \le M}} \sum_{q_1 \sim \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2)=1}} g_2(\ell q_2) \sum_{j \le 4\frac{ML}{\ell}} \frac{\Sigma_j^{(\nu)}(c)}{j}, \quad (\nu = 0, 1, 2).$$

By applying partial summation with respect to $q_1$ we see that

$$\begin{aligned}\sum_{q_1 \sim \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2)=1}} g_2(\ell q_2)\Sigma_j^{(0)}(c) = \sum_{q_1 \sim \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2)=1}} g_2(\ell q_2) \sin\frac{\pi c N j}{Q_1 q_2}\cos\frac{2\pi\overline{q}_1 bj}{q_2} + \\ + \frac{2\pi c N j}{\ell}\int_{Q_1/\ell}^{2Q_1/\ell} \sum_{\frac{Q_1}{\ell} < q_1 \le v} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2)=1}} \frac{g_2(\ell q_2)}{q_2}\cos\frac{2\pi\overline{q}_1 bj}{q_2}\cos\frac{2\pi c N j}{\ell q_2 v}\frac{\mathrm{d}v}{v^2}.\end{aligned}$$

Thus, from Lemma 1 it follows that $\mathcal{R}_c^{(0)}(a) \lll N^{\delta} Q_1^{53/48} Q_2^{7/8}$. Now let us prove $\mathcal{R}_c^{(1)}(a), \mathcal{R}_c^{(2)}(a) \lll N^{1-\delta}$. To this end, by using the expansion of the cosine function in power series, we fix an integer $K > 1$ and write

$$\begin{aligned}\sum_{j \le 4ML/\ell} \frac{\Sigma_j^{(1)}(c)}{j} &= \sum_{j \le 4ML/\ell} \Big(1 - \cos\frac{2\pi\{\alpha_c\}j}{q_2}\Big)\frac{\Sigma_j^{(0)}(c)}{j} = \\ &= \sum_{n=1}^{K-1} \frac{b_n}{q_2^n} \sum_{j \le 4ML/\ell} j^{n-1}\Sigma_j^{(0)}(c) + O\Big(\Big(\frac{ML}{\ell q_2}\Big)^K\Big) \ll\end{aligned}$$

$$\ll \frac{\ell}{ML} \max_{v\le 4ML/\ell} \Big| \sum_{j\le v} \Sigma_j^{(0)}(c) \Big| \sum_{n=1}^{K-1} \Big(\frac{4ML}{\ell q_2}\Big)^n \ll$$

$$\ll \frac{\ell N^{1-\delta}}{Q_1 Q_2} \max_{v\le \frac{4Q_1Q_2}{\ell N^{1-\delta}}} \Big| \sum_{j\le v} (\mathcal{E}'_{c,j} - \overline{\mathcal{E}}'_{c,j}) \Big|,$$

where we have set $\mathcal{E}'_{c,j} \overset{def}{=} e_{q_2}\big(j(\alpha_c + \overline{q}_1 b)\big) + e_{q_2}\big(j(\alpha_c - \overline{q}_1 b)\big)$, and we have used $b_n \ll 1$, $ML/(\ell q_2) = o(1)$ (the latter following straightforwardly from the hypothesis $Q_1 = o(N^{1-\delta})$). Now, it is easy to see that, according as $q_2 | \alpha_c \pm \overline{q}_1 b$ or not, the same arguments adopted in the proof of I) to treat the exponential sums lead to

$$\mathcal{R}_c^{(1)}(a) \lll \frac{N^{1-\delta}}{Q_1 Q_2} \sum_{\substack{\ell | a \\ \ell \le M}} \ell \sum_{q_1 \sim \frac{Q_1}{\ell}} \sum_{\substack{q_2 \sim \frac{Q_2}{\ell} \\ (q_1, q_2)=1}} \max_{v\le \frac{4Q_1Q_2}{\ell N^{1-\delta}}} \Big| \sum_{j\le v} (\mathcal{E}'_{c,j} - \overline{\mathcal{E}}'_{c,j}) \Big| \lll N^{1-\delta}.$$

In a completely similar way, we conclude also that $\mathcal{R}_c^{(2)}(a) \lll N^{1-\delta}$ after using the expansion of the sine function in power series and noticing that

$$4 \cos \frac{2\pi \alpha_c j}{q_2} \cos \frac{2\pi \overline{q}_1 b j}{q_2} = \mathcal{E}'_{c,j} + \overline{\mathcal{E}}'_{c,j}.$$

The lemma is completely proved. □

Now, we can state and prove the main lemma of the paper. It gives a fairly general asymptotic formula for the *correlation*

$$\mathcal{C}_{f_1,f_2}(a) \overset{def}{=} \sum_{n\sim N} f_1(n) f_2(n-a)$$

of real sieve functions $f_1, f_2$. In particular, it provides a strong level ($> 1/2$) for the autocorrelation $\mathcal{C}_f = \mathcal{C}_{f,f}$ of a real sieve function $f$. It is also worthwhile to remark that the next lemma applications to $J_{f_1,f_2}$ and $I_{f_1,f_2}$ (i.e., Theorem 2) improve the non-trivial bounds given in [2].

LEMMA 3. *Fix a sufficiently small* $\delta > 0$. *Let* $N, Q_1, Q_2 \in \mathbb{N}$ *such that* $Q_2 \to \infty$ *and* $Q_2 \ll Q_1 \ll N^{1-\delta}$, *as* $N \to \infty$. *If* $g_1, g_2 : \mathbb{N} \to \mathbb{R}$ *are essentially bounded and supported in* $[1, Q_1]$ *and* $[1, Q_2]$, *respectively, then for every* $\varepsilon > 0$ *and uniformly*

*$\forall\, a \in \mathbb{Z} \setminus \{0\}$, with $|a| = o(N)$, one has*

$$\frac{\mathcal{C}_{f_1,f_2}(a) + \mathcal{C}_{f_1,f_2}(-a)}{2} = N \sum_{\ell | a} \frac{1}{\ell} \sum\sum_{\substack{q_1 \le Q_1/\ell \\ q_2 \le Q_2/\ell \\ (q_1,q_2)=1}} \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} + $$

$$+ O_{\varepsilon}\big(N^{\delta+\varepsilon} Q_1^{53/48} Q_2^{7/8} + N^{1-2\delta/3+\varepsilon}\big),$$

*where $f_1 = g_1 * \mathbf{1}$, $f_2 = g_2 * \mathbf{1}$.*

Proof. First, we observe that

$$\mathcal{C}_{f_1,f_2}(a) = \sum_{n \sim N} f_1(n) \sum_{\substack{q_2 | n-a \\ q_2 \le Q_2}} g_2(q_2) = \sum_{q_2 \le Q_2} g_2(q_2) \sum_{\substack{n \sim N \\ n \equiv a\,(q_2)}} \sum_{\substack{q_1 | n \\ q_1 \le Q_1}} g_1(q_1) =$$

$$= \sum_{\ell | a} \sum_{q_2 \le Q_2} g_2(q_2) \sum_{\substack{q_1 \le Q_1 \\ (q_1,q_2)=\ell}} g_1(q_1) \sum_{\substack{n \sim N \\ n \equiv 0\,(q_1) \\ n \equiv a\,(q_2)}} 1 =$$

$$= \sum_{\ell | a} \sum_{q_1 \le \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \le \frac{Q_2}{\ell} \\ (q_1,q_2)=1}} g_2(\ell q_2) \sum_{\substack{m \sim \frac{N}{\ell q_1} \\ m \equiv \overline{q}_1 b\,(q_2)}} 1,$$

where we set $b = |a|/\ell$ as before. Then, plainly we can write

$$\frac{\mathcal{C}_{f_1,f_2}(a) + \mathcal{C}_{f_1,f_2}(-a)}{2} = \frac{1}{2} \sum_{\ell | a} \sum_{q_1 \le \frac{Q_1}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \le \frac{Q_2}{\ell} \\ (q_1,q_2)=1}} g_2(\ell q_2) \sum_{\substack{m \sim \frac{N}{\ell q_1} \\ m \equiv \pm \overline{q}_1 b\,(q_2)}} 1.$$

Now, let us set $Q_{c,k} \overset{def}{=} 2^{-k-1} Q_c$ for all $k = 0, \ldots, [\log_2 Q_c]$, $c = 1, 2$, and confine to the dyadic intervals $(Q_{c,k}/\ell, 2Q_{c,k}/\ell]$, where we define $\mathcal{R}_k(a)$ analogously to $\mathcal{R}(a)$. Thus, by the formula (2.2) we get

$$\frac{1}{2} \sum_{\ell | a} \sum_{q_1 \sim \frac{Q_{1,k}}{\ell}} g_1(\ell q_1) \sum_{\substack{q_2 \sim \frac{Q_{2,k}}{\ell} \\ (q_1,q_2)=1}} g_2(\ell q_2) \sum_{\substack{m \sim \frac{N}{\ell q_1} \\ m \equiv \pm \overline{q}_1 b\,(q_2)}} 1 = N \sum_{\ell | a} \frac{1}{\ell} \sum\sum_{\substack{q_c \sim \frac{Q_{c,k}}{\ell} \\ (q_1,q_2)=1}} \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} +$$

$$+ \frac{\mathcal{R}_k(a)}{2} + O_{\varepsilon}\Big(N^{\varepsilon} \sum_{\ell | a} \sum_{q_1 \sim \frac{Q_{1,k}}{\ell}} \sum_{\substack{q_2 \sim Q_{2,k}/\ell \\ q_2 | q_1 [\alpha_c] \pm b}} 1\Big).$$

Since $|a| = o(N)$ yields $q_1[cN/(\ell q_1)] \pm b \neq 0$, the latter $O$-term is $\lll Q_1 \lll N^{1-\delta}$. Moreover, we can assume that $Q_{1,k}Q_{2,k} \gg N^{1-2\delta/3}$, for otherwise trivially $\mathcal{R}_k(a) \lll \lll N^{1-2\delta/3}$. Hence, the conclusion follows from II) of Lemma 2. □

**Remark.** Since $f_1, f_2$ are essentially bounded, then for any $a > 0$ one has

$$\mathcal{C}_{f_1,f_2}(-a) = \sum_{n\sim N} f_1(n)f_2(n+a) = \sum_{N+a<n\le 2N+a} f_2(n)f_1(n-a) = \mathcal{C}_{f_2,f_1}(a) + O_\varepsilon(aN^\varepsilon).$$

In particular, if $f_1 = f_2$, this implies that

$$\frac{\mathcal{C}_{f_1,f_2}(a) + \mathcal{C}_{f_1,f_2}(-a)}{2} = \mathcal{C}_f(a) + O_\varepsilon\Big(N^\varepsilon|a|\Big).$$

Therefore, from the previous lemma we obtain the following formula for the value attained at $a = o(N)$ by the autocorrelation of a sieve function $f = g * \mathbf{1}$ of range $Q \ll N^{1-\delta}$:

$$\mathcal{C}_f(a) = N \sum_{\ell | a} \frac{1}{\ell} \sum_{\substack{q_1, q_2 \sim Q/\ell \\ (q_1,q_2)=1}}\sum \frac{g(\ell q_1)g(\ell q_2)}{q_1 q_2} + O_\varepsilon\Big(N^{\delta+\varepsilon}Q^{\frac{95}{48}} + N^{1-2\delta/3+\varepsilon}\Big),$$

that means level $\lambda = 1/2 + 1/190$ for autocorrelations of $f$.

## 4. Proof of Theorem 2

Let us consider the symmetry integral first, and write

$$\begin{aligned} I_{f_1,f_2}(N,h) &= \sum_{\substack{N-h<n,m\le 2N+h \\ 0\le|n-m|\le 2h}}\sum f_1(n)f_2(m) \int\limits_{\substack{N\le x\le 2N \\ |x-n|\le h \\ |x-m|\le h}} \operatorname{sgn}(x-n)\operatorname{sgn}(x-m)\mathrm{d}x = \\ &= \sum_{\substack{N+h<n,m\le 2N-h \\ 0\le|n-m|\le 2h}}\sum f_1(n)f_2(m) \int\limits_{\substack{|t|\le h \\ |t+(n-m)|\le h}} \operatorname{sgn}(t)\operatorname{sgn}(t+(n-m))\mathrm{d}t + O_\varepsilon(N^\varepsilon h^3) = \\ &= \sum_{N<n\le 2N} f_1(n) \sum_{0\le|a|\le 2h} f_2(n-a) \int\limits_{\substack{|t|\le h \\ |t+a|\le h}} \operatorname{sgn}(t)\operatorname{sgn}(t+a)\mathrm{d}t + O_\varepsilon(N^\varepsilon h^3). \end{aligned}$$

To simplify our exposition, the symbol (T) within some of the next formulæ will warn the reader of some *tails*, i.e. terms being $\lll h^3 \lll Nh$, that are discarded to abbreviate the formulæ themselves. Thus, the above equation becomes

$$I_{f_1,f_2}(N,h) \overset{(\mathrm{T})}{\sim} \sum_{0\le|a|\le 2h} W(a)\mathcal{C}_{f_1,f_2}(a), \qquad \text{with} \quad W(a) \overset{def}{=} \int\limits_{\substack{|t|\le h\\ |t-a|\le h}} \operatorname{sgn}(t)\operatorname{sgn}(t-a)\mathrm{d}t.$$

Since $f_1, f_2$ are essentially bounded, then

$$W(0)\mathcal{C}_{f_1,f_2}(0) = 2h\sum_{n\sim N} f_1(n)f_2(n) \lll Nh.$$

Moreover, note that $W$ is even and $W(a) \ll h$ uniformly for all $a$. Therefore, Lemma 3 implies that

$$\begin{aligned}\sum_{0<|a|\le 2h} W(a)\mathcal{C}_{f_1,f_2}(a) &= \sum_{0<a\le 2h} W(a)(\mathcal{C}_{f_1,f_2}(a) + \mathcal{C}_{f_1,f_2}(-a)) = \\ &= 2N\sum_{0<a\le 2h} W(a)\sum_{\ell|a}\frac{1}{\ell}\sum_{\substack{q_1\le Q_1/\ell\\ q_2\le Q_2/\ell\\ (q_1,q_2)=1}}\sum \frac{g_1(\ell q_1)g_2(\ell q_2)}{q_1q_2} + \\ &\quad + O_\varepsilon\big(N^{\delta+\varepsilon}Q_1^{53/48}Q_2^{7/8}h^2 + N^{1-2\delta/3+\varepsilon}h^2\big).\end{aligned}$$

Hence, the stated inequality for $I_{f_1,f_2}(N,h)$ follows from

$$\sum_{0<a\le 2h} W(a)\sum_{\ell|a}\frac{1}{\ell}\sum_{\substack{q_1\le Q_1/\ell\\ q_2\le Q_2/\ell\\ (q_1,q_2)=1}}\sum \frac{g_1(\ell q_1)g_2(\ell q_2)}{q_1q_2} \lll \sum_{\ell\le 2h}\frac{1}{\ell}\Big|\sum_{0<b\le 2h/\ell} W(\ell b)\Big| \lll h,$$

where we have applied the property (see [2], Lemma 2.4)

$$\sum_{\substack{0<a\le 2h\\ a\equiv 0\,(\ell)}} W(a) = 2\ell\Big\|\frac{h}{\ell}\Big\| \ll h, \quad \text{for } 1\le \ell\le 2h.$$

Now, let us turn our attention to the Selberg integral $J_{f_1,f_2}(N,h)$. First we observe that for any $c = 1, 2$

$$\int_N^{2N} \sum_{x<n\le x+h} f_c(n)\,\mathrm{d}x = \sum_{N<n\le 2N+h} f_c(n) \int\limits_{\substack{N\le x\le 2N\\ n-h\le x<n}} \mathrm{d}x =$$

$$= \sum_{N+h<n<2N-h} f_c(n) \int_{n-h}^{n} \mathrm{d}x + O_\varepsilon(N^\varepsilon h^2) =$$

$$= h\sum_{n\sim N} f_c(n) + O_\varepsilon(N^\varepsilon h^2).$$

Since

$$\sum_{n\sim N} f_c(n) = \sum_d g_c(d)\Big(\Big[\frac{2N}{d}\Big] - \Big[\frac{N}{d}\Big]\Big) = N\frac{M_{f_c}(h)}{h} + O_\varepsilon(N^\varepsilon Q_c),$$

then, by recalling that $Q_1 \ge Q_2$ and $M_{f_c}(h) \lll h$, we can write

$$J_{f_1,f_2}(N,h) = \int_N^{2N} \sum\sum_{x<n,m\le x+h} f_1(n)f_2(m)\mathrm{d}x - NM_{f_1}(h)M_{f_2}(h) + O_\varepsilon\big(N^\varepsilon(Q_1h^2+h^3)\big).$$

Now, by arguing as we have done before for $I_{f_1,f_2}(N,h)$, it is easy to see that

$$\int_N^{2N} \sum\sum_{x<n,m\le x+h} f_1(n)f_2(m)\mathrm{d}x = \sum_{0\le|a|\le h} (h-|a|)\sum_{n\sim N} f_1(n)f_2(n+a) + O_\varepsilon(N^\varepsilon h^3),$$

which yields

$$J_{f_1,f_2}(N,h) \overset{(\mathrm{T})}{\sim} \sum_{0<|a|\le h} (h-|a|)\mathcal{C}_{f_1,f_2}(-a) - NM_{f_1}(h)M_{f_2}(h) + O_\varepsilon\big(N^\varepsilon(Nh+Q_1h^2)\big).$$

Note that $h-|a|$ is an even function of the variable $a$. Thus, the previous calculations and Lemma 3 apply again here to get

$$\sum_{0<|a|\le h} (h-|a|)\mathcal{C}_{f_1,f_2}(-a) = \sum_{0<a\le h} (h-a)\big(\mathcal{C}_{f_1,f_2}(-a) + \mathcal{C}_{f_1,f_2}(a)\big) =$$

$$= 2N \sum_{\ell \le h} \frac{1}{\ell} \sum_{b \le h/\ell} (h - \ell b) \sum_{\substack{q_1 \le Q_1/\ell \\ q_2 \le Q_2/\ell \\ (q_1, q_2)=1}} \sum \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} +$$

$$+ O_\varepsilon\big(N^{\delta+\varepsilon} Q_1^{53/48} Q_2^{7/8} h^2 + N^{1-2\delta/3+\varepsilon} h^2\big).$$

By using the formula

$$\sum_{b \le h/\ell} (h - \ell b) = \frac{h^2}{2\ell} + O(h), \quad \forall\, \ell \le h,$$

we can write

$$2N \sum_{\ell \le h} \frac{1}{\ell} \sum_{b \le h/\ell} (h - \ell b) \sum_{\substack{q_1 \le Q_1/\ell \\ q_2 \le Q_2/\ell \\ (q_1, q_2)=1}} \sum \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} =$$

$$= Nh^2 \sum_{\ell \le h} \frac{1}{\ell^2} \sum_{\substack{q_1 \le Q_1/\ell \\ q_2 \le Q_2/\ell \\ (q_1, q_2)=1}} \sum \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} + O_\varepsilon(N^{1+\varepsilon} h) \overset{(T)}{\sim}$$

$$\overset{(T)}{\sim} Nh^2 \sum_{\ell=1}^{\infty} \frac{1}{\ell^2} \sum_{\substack{q_1 \le Q_1/\ell \\ q_2 \le Q_2/\ell \\ (q_1, q_2)=1}} \sum \frac{g_1(\ell q_1) g_2(\ell q_2)}{q_1 q_2} =$$

$$= Nh^2 \sum_{\ell=1}^{\infty} \sum_{\substack{q_1 \le Q_1 \\ q_2 \le Q_2 \\ (q_1, q_2)=\ell}} \sum \frac{g_1(q_1) g_2(q_2)}{q_1 q_2} =$$

$$= Nh^2 \sum_{\substack{q_1 \le Q_1 \\ q_2 \le Q_2}} \sum \frac{g_1(q_1)}{q_1} \frac{g_2(q_2)}{q_2} = N M_{f_1}(h) M_{f_2}(h).$$

Hence, we conclude that

$$\sum_{0<|a|\le h} (h-|a|) \mathcal{C}_{f_1,f_2}(-a) - N M_{f_1}(h) M_{f_2}(h) \ll_\varepsilon$$

$$\ll_\varepsilon N^{\delta+\varepsilon} Q_1^{53/48} Q_2^{7/8} h^2 + N^{1-2\delta/3+\varepsilon} h^2 + N^{1+\varepsilon} h.$$

Theorem 2 is completely proved. □

## 5. Further comments and remarks

1. Though analogous definitions and results can be easily formulated for complex valued sieve functions, here we stick to the real case for simplicity.
2. A famous example of sieve function is given by the truncated divisor sum $\Lambda_R$, exploited by Goldston in [8]. We refer the reader to [5] for an application of our recent study about the distribution of sieve functions to the case of $\Lambda_R$.
3. The key of the present approach is the treatment of the *error term* $\mathcal{R}_f(a)$ arising from the formula of the autocorrelation of a sieve function $f = g * \mathbf{1}$ written for any nonzero integer $a = o(N)$ as

$$\mathcal{C}_f(a) = \sum_{\ell|a} \sum \sum_{(d,q)=1} g(\ell q_1) g(\ell q_2) \frac{1}{q_2} \left( \left[ \frac{2N}{\ell q_1} \right] - \left[ \frac{N}{\ell q_1} \right] \right) + \mathcal{R}_f(a).$$

In [2] such an error term is defined by using the orthogonality of the additive characters. The estimate of $\mathcal{R}_f(a)$ leads to the gain $\Delta = 1/48$ given in Theorem 1 and, as noted above, it is essentially due to the non-trivial bound of the bilinear forms with Kloosterman fractions by Duke, Friedlander and Iwaniec (see Lemma 1). In this respect, the recent improvement obtained by Bettin and Chandee [1] would yield $\Delta = 1/20$. Moreover, we think that this result might lead to an improvement of ours in respect of the short length $h$ as well. We are going to show such a further achievement in a future paper.

4. In the literature the *level* of distribution of an arithmetic function $f$ in the arithmetic progressions is usually meant to be a positive real number $\lambda_{AP}(f)$ such that

$$\sum_{q \le Q} \max_{(a,q)=1} \Big| \sum_{\substack{n \le x \\ n \equiv a\,(q)}} f(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \le x \\ (n,q)=1}} f(n) \Big| \ll_\varepsilon x^{1-\varepsilon}$$

$$(\varphi(q) \overset{def}{=} |\{a \le q, (a,q) = 1\}|),$$

holds for $Q \le x^{\lambda_{AP}(f)-\delta}$ (where $\delta > 0$ is sufficiently small and depends on $\varepsilon > 0$). For example, the celebrated Bombieri-Vinogradov Theorem gives $\lambda_{AP}(\Lambda) = \frac{1}{2}$ for the von-Mangoldt function $\Lambda$.

It is a classical and standard argument (see [7] for example) to deduce from the previous inequality an asymptotic formula for the autocorrelations of a sieve function $f$ given by

$$f(n) = \sum_{\substack{q|n \\ q\le Q}} g(q), \quad \text{with} \quad Q \ll N^{\lambda_{AP}(f)-\delta}.$$

Somehow this justifies our definiton of *level* when we refer to a sieve function. Unfortunately, it seems to be very hard to reverse such a process, that is to say, information on the *level* of distribution of $f$ in the arithmetic progressions seems to be much stronger than knowledge about the autocorrelation of $f$. From this point of view, our Lemma 3 provides a substitute for what is still lacking from the study of distribution of $f$ in the arithmetic progressions.

## 6. Appendix: the first Bernoulli function on the rational numbers

Here we prove the formula (2.1), that equivalently we state as

$$\mathcal{B}_1\left(\frac{n}{q}\right) = \frac{i}{2q} \sum_{0<|j|\le q/2} \cot\frac{\pi j}{q} e_q(jn), \qquad \forall\,(n,q) \in \mathbb{Z}\times\mathbb{N}\setminus\{1\}.$$

To this end, it suffices to establish the following equality for the Fourier coefficient of $\mathcal{B}_1$:

$$c_{j,q} \overset{def}{=} \frac{1}{q} \sum_{0\le|r|\le q/2} \mathcal{B}_1\left(\frac{r}{q}\right) e_q(-jr) = \frac{i}{2q}\cot\frac{\pi j}{q}.$$

First, since $\mathcal{B}_1$ is odd, note that $c_{j,q} = \frac{i}{q}\Sigma$, where we set

$$\Sigma \overset{def}{=} -\frac{2}{q} \sum_{r\le[q/2]} r\sin\frac{2\pi jr}{q} + \sum_{r\le[q/2]} \sin\frac{2\pi jr}{q}.$$

Note that, for $R=[q/2]$ by Abel's lemma (see [10], Appendix A, exercise 3) one has

$$\sum_{r\le R} r\sin\frac{2\pi jr}{q} = R\sum_{r\le R}\sin\frac{2\pi jr}{q} - \sum_{r\le R-1}\Big(\sum_{\ell\le r}\sin\frac{2\pi j\ell}{q}\Big).$$

Therefore, by using the identities (see [9], formulæ n.1.342.1 and n.1.342.2), $\forall\, X \in \mathbb{N}, \forall\, \alpha \in \mathbb{R}\backslash\mathbb{Z}$,

$$\sum_{r\le X} \sin(2\pi\alpha r) = \sin^2(\pi\alpha X)\cot(\pi\alpha) + \frac{\sin(2\pi\alpha X)}{2},$$

$$\sum_{r\le X} \cos(2\pi\alpha r) = \frac{\sin(2\pi\alpha X)\cot(\pi\alpha)}{2} - \frac{1-\cos(2\pi\alpha X)}{2},$$

we can write

$$\begin{aligned}
\Sigma &= \Big(-\frac{2R}{q}+1\Big)\sum_{r\le R}\sin\frac{2\pi jr}{q} + \frac{2}{q}\sum_{r\le R-1}\Big(\sum_{\ell\le r}\sin\frac{2\pi j\ell}{q}\Big) = \\
&= \frac{2}{q}\cot\frac{\pi j}{q}\sum_{r\le R-1}\sin^2\frac{\pi jr}{q} + \Big(1-\frac{2R}{q}+\frac{1}{q}\Big)\sum_{r\le R}\sin\frac{2\pi jr}{q} - \frac{1}{q}\sin\frac{2\pi jR}{q} = \\
&= \frac{R}{q}\cot\frac{\pi j}{q} + \frac{1}{q}\Big(\cos\frac{2\pi jR}{q}-1\Big)\cot\frac{\pi j}{q} - \\
&\quad - \frac{1}{2q}\cot\frac{\pi j}{q}\Big(\sin\frac{2\pi jR}{q}\cot\frac{\pi j}{q} + \cos\frac{2\pi jR}{q} - 1\Big) + \\
&\quad + \frac{1+2\{q/2\}}{q}\Big(\sin^2\frac{\pi jR}{q}\cot\frac{\pi j}{q} + \frac{1}{2}\sin\frac{2\pi jR}{q}\Big) - \frac{1}{q}\sin\frac{2\pi jR}{q},
\end{aligned}$$

where

$$2\{q/2\} = q - 2[q/2] = q - 2R = \begin{cases} 1 & \text{if } q \text{ is odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Now, since from $2\pi jR = \pi jq - 2\pi j\{q/2\}$ it follows that

$$\begin{aligned}
\cos\frac{2\pi jR}{q} &= (-1)^j\cos\frac{2\pi j\{q/2\}}{q}, \\
\sin\frac{2\pi jR}{q} &= (-1)^{j+1}2\Big\{\frac{q}{2}\Big\}\sin\frac{\pi j}{q}, \\
\sin^2\frac{\pi jR}{q} &= \frac{1}{2} + \frac{(-1)^{j+1}}{2}\cos\frac{2\pi j\{q/2\}}{q},
\end{aligned}$$

then we get

$$\begin{aligned}\Sigma &= \frac{1}{2}\cot\frac{\pi j}{q} - 2\left\{\frac{q}{2}\right\}\frac{1}{2q}\cot\frac{\pi j}{q} + \\ &+ \frac{1}{2q}\left((-1)^j\cos\frac{\pi j 2\{q/2\}}{q} - 1\right)\cot\frac{\pi j}{q} + \\ &+ 2\left\{\frac{q}{2}\right\}\frac{(-1)^j}{2q}\sin\frac{\pi j}{q}\cot^2\frac{\pi j}{q} + \\ &+ \frac{1+2\{q/2\}}{q}\left(\frac{1}{2} - \frac{(-1)^j}{2}\cos\frac{\pi j 2\{q/2\}}{q}\right)\cot\frac{\pi j}{q} = \\ &= \frac{1}{2q}\cot\frac{\pi j}{q}\Bigg(q - 2\left\{\frac{q}{2}\right\} + 2\left\{\frac{q}{2}\right\}(-1)^j\cos\frac{\pi j}{q} + \\ &+ 2\left\{\frac{q}{2}\right\}\left(1 - (-1)^j\cos\frac{\pi j 2\{q/2\}}{q}\right)\Bigg) = \frac{1}{2}\cot\frac{\pi j}{q}.\end{aligned}$$

## 7. Acknowledgements

Research started while first author was a Fellow ”Ing. Giorgio Schirillo” of the Istituto Nazionale di Alta Matematica.

## Bibliography

1. ***S. Bettin, V. Chandee***, *Trilinear forms with Kloosterman fractions*, preprint, arXiv:1502.00769.
2. ***G. Coppola***, *On the Correlations, Selberg integral and symmetry of sieve functions in short intervals*, J. Combinatorics and Number Theory **2**:2 (2010), 91–105.
3. ***G. Coppola***, *On the Correlations, Selberg integral and symmetry of sieve functions in short intervals, II*, Int. J. Pure Appl. Math. **58**:3 (2010), 281–298.
4. ***G. Coppola, M. Laporta***, *Symmetry and short interval mean-squares*, preprint, arXiv:1312.5701.
5. ***G. Coppola, M. Laporta***, *Sieve functions in arithmetic bands*, preprint, arXiv:1503.07502.
6. ***W. Duke, J. Friedlander, H. Iwaniec***, *Bilinear forms with Kloosterman fractions*, Invent. Math. **128**:1 (1997), 23–43.
7. ***P. D. T. A. Elliott***, *On the Correlation of Multiplicative and the Sum of Additive Arithmetic Functions*, Mem. Amer. Math. Soc., **112**:538 (1994).
8. ***D. A. Goldston***, *On Bombieri and Davenport’s theorem concerning small gaps between primes*, Mathematika **39** (1992), 10–17.

9. ***I. S. Gradshteyn, I. M. Ryzhik***, *Tables of integrals*, *series*, *and products*, 5th edition, Academic Press, 1994.

10. ***H. L. Montgomery, R. C. Vaughan***, *Multiplicative Number Theory. I. Classical Theory*, Studies in Advanced Math. **97**, Cambridge University Press, 2007.

GIOVANNI COPPOLA

Università degli Studi di Napoli
Via Partenio 12, 83100,
Avellino(AV), Italy
giovanni.coppola@unina.it

MAURIZIO LAPORTA

Università degli Studi di Napoli
Dipartimento di Matematica e Appl.
Compl. Monte S. Angelo
Via Cinthia, 80126, Napoli, Italy
mlaporta@unina.it