Acta Math. Hungar.
103 (3) (2004), 243–254.

ON VARIANTS OF THE LARGER SIEVE

E. S. CROOT III (Atlanta) and C. ELSHOLTZ (Egham)

Abstract. Gallagher's larger sieve is a powerful tool, when dealing with sequences of integers that avoid many residue classes. We present and discuss various variants of Gallagher's larger sieve.

1. Introduction

The large sieve has its origins in the work of Linnik and Rényi. It was developed to deal with sequences that avoid a positive proportion of residue classes. It was later strengthened and simplified by Roth, Bombieri, Davenport, Halberstam, Montgomery, Selberg, Gallagher and many others. For a survey see Montgomery [15] and Bombieri [1].

It is known that Montgomery's large sieve [15] is a useful method when sifting sequences that avoid many residue classes modulo primes. But for sequences that avoid on average more than half of the residue classes it is preferable to use Gallagher's *larger* sieve [8]. For example, if a sequence $\mathcal{A} \subset [1, N]$ avoids $\omega(p) = \frac{p-1}{2}$ residue classes modulo the primes $2 < p \leqq \sqrt{N}$, then Gallagher's larger sieve gives the upper bound $\mathcal{A}(N) \ll \sqrt{N}$. Moreover the squares are the standard example to show that here both, the large and the larger sieve achieve the correct order of magnitude (see Section 3).

Recently there emerged quite a few new applications of Gallagher's larger sieve (see e.g. Dujella [4], Elsholtz [5], [6], Gyarmati [10], Hegyvári and Sárközy [12]) so that it seemed worthwhile looking for variants of this sieve having some advantages over Gallagher's version. (We would like to point out that Gallagher had various contributions to the large sieve. In addition to the *larger* sieve that we consider here, he gave a simple approach to the large sieve [7] and developed a sieve that allows to sift modulo powers of primes, [9].)

Key words and phrases: large sieve method, Gallagher's larger sieve.
2000 Mathematics Subject Classification: 11M35.

Let us state Montgomery's large sieve and Gallagher's larger sieve first. We then state and prove our new variants of it. In the final section we discuss the advantages or disadvantages of these variants.

THEOREM 1 (Montgomery [15]). *Let $\mathcal{P}$ denote the set of primes. Let $\mathcal{A} \subset [1, N]$ denote a set of integers which lies outside $\omega(p)$ residue classes modulo the prime p. Here $\omega : \mathcal{P} \to \mathbb{N}$ with $0 \leqq \omega(p) \leqq p-1$. Then the following bound on the counting function $\mathcal{A}(N)$ holds:*

$$\mathcal{A}(N) \leqq \frac{N+Q^2-1}{L}, \quad \textit{where} \quad L = \sum_{q \leqq Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p-\omega(p)}.$$

One usually takes $Q = \sqrt{N}$.

THEOREM 2 (Gallagher's larger sieve, [8]). *Let $\mathcal{S}$ denote a set of primes or powers of primes, such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Then the following bound holds, provided the denominator is positive:*

$$|\mathcal{A}| \leqq \frac{-\log N + \sum_{q \in \mathcal{S}} \Lambda(q)}{-\log N + \sum_{q \in \mathcal{S}} \dfrac{\Lambda(q)}{\nu(q)}}.$$

Here Λ denotes the von Mangoldt function defined by

$$\Lambda(q) = \begin{cases} \log p, & \text{if } \ q = p^r \text{ with prime } p, \\ 0, & \text{otherwise.} \end{cases}$$

2. New variants of the larger sieve

THEOREM 3 (Variant 1). *Let $\mathcal{S} \subset [2, Q]$ denote a set of primes or powers of primes such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Then,*

$$|\mathcal{A}| \leqq \max\left(Q, \frac{23N \exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{q}\right)}{\exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{\nu(q)}\right)}\right).$$

For most applications one chooses $\mathcal{S}$ to be the set of all primes in some interval $[2, Q]$, where Q is less than the upper bound one is going to prove

for $|\mathcal{A}|$. Then $Q \sim \sum_{q\in\mathcal{S}} \Lambda(q) \sim C \exp\Big(\sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{q}\Big)$ for some positive constant C.

THEOREM 4 (Variant 2). *Let $\mathcal{S}$ denote a set of primes or powers of primes such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Let $G = \max_{q\in\mathcal{S}} \nu(q)$. Then the following inequality holds, provided the denominator is positive:*

$$|\mathcal{A}| \leqq \frac{-G\log N + \sum_{q\in\mathcal{S}} \nu(q)\Lambda(q)}{-G\log N + \sum_{q\in\mathcal{S}} \Lambda(q)}.$$

The following two variants look odd for a sieve bound on $|\mathcal{A}|$ but they may still be useful.

THEOREM 5 (Variant 3). *Let $\mathcal{S}$ denote a set of primes or powers of primes such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Suppose that $|\mathcal{A}| > \sum_{q\in\mathcal{S}} \Lambda(q)$. Then,*

$$1 + \log N \geqq \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)}.$$

THEOREM 6 (Variant 4). *Let $\mathcal{S}$ denote a set of primes or powers of primes such that $\mathcal{A} \subset [1, N]$ lies in at most $\nu(q)$ residue classes modulo q, for each $q \in \mathcal{S}$. Then the following inequality holds, provided the denominator is positive:*

$$|\mathcal{A}| \leqq \frac{-\log N + \sum_{q\in\mathcal{S}} \Lambda(q)}{-\log N + \frac{1}{|\mathcal{A}|^2} \sum_{q\in\mathcal{S}} \Lambda(q) \sum_{c=0}^{q-1} |a \in \mathcal{A} :\ a \equiv c \bmod q|^2}.$$

PROOF OF VARIANT 1. For any integer q let

$$B(q) := \Big|\big\{(a, a') \in \mathcal{A}^2 :\ a \neq a',\ q \mid a - a'\big\}\Big|$$

and

$$C(q) := \sum_{c=0}^{q-1} \Big|\{a \in \mathcal{A} :\ a \equiv c \bmod q\}\Big|^2.$$

Note that $B(q) = C(q) - |\mathcal{A}|$.

For any pair of distinct integers $a, a' \in \mathcal{A}$, we have that

$$\sum_{q|a-a'} \Lambda(q) = \sum_{p^r|a-a'} \log p = \log|a - a'| < \log N,$$

where $\sum_{p^r|a-a'}$ is taken over all primes and powers of primes that divide $a-a'$. So,

$$(1) \qquad \sum_{a\in\mathcal{A}} \sum_{\substack{a'\in\mathcal{A},\\ a'\neq a}} \sum_{q|a-a'} \Lambda(q) = \sum_{q\leqq N} B(q)\Lambda(q) < \left(|\mathcal{A}|^2 - |\mathcal{A}|\right) \log N.$$

$C(q)$ attains its minimum value when all the elements of $\mathcal{A}$ are as evenly distributed as possible amongst the $\nu(q)$ progressions modulo q which $\mathcal{A}$ occupies. Thus,

$$C(q) \geqq \nu(q) \left(\frac{|\mathcal{A}|}{\nu(q)}\right)^2 = \frac{|\mathcal{A}|^2}{\nu(q)}.$$

This gives

$$(2) \qquad \sum_{q\in\mathcal{S}} B(q)\Lambda(q) \geqq |\mathcal{A}|^2 \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)} - |\mathcal{A}| \sum_{q\in\mathcal{S}} \Lambda(q).$$

For $q \notin \mathcal{S}$ we do not assume that the set $\mathcal{A}$ avoids any residue class modulo q. Thus, for each q the smallest that $B(q)$ can be occurs if all the elements of $\mathcal{A}$ are equally distributed amongst the q progressions modulo q; so we use for $q < |\mathcal{A}|$ with $q \notin \mathcal{S}$,

$$B(q) = C(q) - |\mathcal{A}| > \frac{|\mathcal{A}|^2}{q} - |\mathcal{A}|.$$

Thus,

$$\sum_{q\leqq|\mathcal{A}|,\ q\notin\mathcal{S}} B(q)\Lambda(q) \geqq |\mathcal{A}|^2 \sum_{q\leqq|\mathcal{A}|,\ q\notin\mathcal{S}} \frac{\Lambda(q)}{q} - |\mathcal{A}| \sum_{q\leqq|\mathcal{A}|,\ q\notin\mathcal{S}} \Lambda(q)$$

$$= |\mathcal{A}|^2 \left(\sum_{q\leqq|\mathcal{A}|} \frac{\Lambda(q)}{q} - \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{q} \right) - |\mathcal{A}| \sum_{q\leqq|\mathcal{A}|,\ q\notin\mathcal{S}} \Lambda(q)$$

$$> |\mathcal{A}|^2 \left(\log|\mathcal{A}| - 2.06 - \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{q} \right) - |\mathcal{A}| \sum_{q\leqq|\mathcal{A}|,\ q\notin\mathcal{S}} \Lambda(q).$$

Here we have used formula (3.21) of Rosser and Schoenfeld [18] which implies that $\sum_{q\leqq|\mathcal{A}|} \frac{\Lambda(q)}{q} > \sum_{p\leqq|\mathcal{A}|} \frac{\Lambda(p)}{p} > \log|\mathcal{A}| + E - \frac{1}{2\log|\mathcal{A}|} > \log|\mathcal{A}| - 2.06$

for $|\mathcal{A}| \geqq 2$, where $E = -1.33258\ldots$. For larger $|\mathcal{A}|$ the constant 2.06 can be replaced by a better constant.

Combining this with (1) and (2), we get

$$\left(|\mathcal{A}|^2 - |\mathcal{A}|\right) \log N > \sum_{q \leqq |\mathcal{A}|} B(q)\Lambda(q)$$

$$= \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} B(q)\Lambda(q) + \sum_{q \leqq |\mathcal{A}|,\ q \notin \mathcal{S}} B(q)\Lambda(q)$$

$$\geqq |\mathcal{A}|^2 \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} \frac{\Lambda(q)}{\nu(q)} - |\mathcal{A}| \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} \Lambda(q)$$

$$+ |\mathcal{A}|^2 \left(-2.06 + \log|\mathcal{A}| - \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} \frac{\Lambda(q)}{q} \right) - |\mathcal{A}| \sum_{q \leqq |\mathcal{A}|,\ q \notin \mathcal{S}} \Lambda(q).$$

Dividing through by $|\mathcal{A}|^2$, gives

$$\left(1 - \frac{1}{|\mathcal{A}|}\right) \log N + \frac{1}{|\mathcal{A}|} \sum_{q \leqq |\mathcal{A}|} \Lambda(q) \geqq -2.06 + \log|\mathcal{A}|$$

$$- \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} \frac{\Lambda(q)}{q} + \sum_{q \leqq |\mathcal{A}|,\ q \in \mathcal{S}} \frac{\Lambda(q)}{\nu(q)};$$

simplifying and rearranging terms gives

$$\log|\mathcal{A}| \leqq 3.1 + \log N + \sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{q} - \sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{\nu(q)}.$$

Here we used Theorem 12 of Rosser and Schoenfeld [18], which implies that for $|\mathcal{A}| na 0$ we have $\frac{1}{|\mathcal{A}|} \sum_{q \leqq |\mathcal{A}|} \Lambda(q) < 1.039$.

It then follows that

$$|\mathcal{A}| \leqq \frac{23N \exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{q}\right)}{\exp\left(\sum_{q \in \mathcal{S}} \frac{\Lambda(q)}{\nu(q)}\right)}. \qquad \square$$

PROOF OF GALLAGHER'S LARGER SIEVE AND OF VARIANT 4. From (1) and (2) we can also easily arrive at Gallagher's original version:

$$\big(|\mathcal{A}|^2 - |\mathcal{A}|\big)\log N > \sum_{q\in\mathcal{S}} B(q)\Lambda(q) \geqq |\mathcal{A}|^2 \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)} - |\mathcal{A}| \sum_{q\in\mathcal{S}} \Lambda(q).$$

Dividing through by $|\mathcal{A}|$ proves

$$|\mathcal{A}| \leqq \frac{-\log N + \sum_{q\in\mathcal{S}} \Lambda(q)}{-\log N + \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)}}.$$

If we work with $C(q)$ instead of $C(q) \geqq \frac{|\mathcal{A}|}{\nu(q)}$, then the same line of argument proves Variant 4. □

PROOF OF VARIANT 2. The proof closely follows Gallagher. Let $A(h,q)$ denote the number of elements of the set $\mathcal{A}$ with $a \equiv h \bmod q$. Then by the Cauchy–Schwarz inequality

$$|\mathcal{A}|^2 = \left(\sum_{h=1}^{q} A(h,q)\right)^2 \leqq \nu(q) \sum_{h=1}^{q} \big(A(h,q)\big)^2,$$

since $A(h,q) = 0$ for all but $\nu(q)$ values of h. Summing over $\mathcal{S}$ we find that

$$|\mathcal{A}|^2 \sum_{q\in\mathcal{S}} \Lambda(q) \leqq \sum_{q\in\mathcal{S}} \Lambda(q)\nu(q) \sum_{\substack{a,a'\in\mathcal{A} \text{ with} \\ a\equiv a' \bmod q}} 1 = \sum_{|d|\leqq N} \sum_{a-a'=d} \sum_{q|d,\, q\in\mathcal{S}} \Lambda(q)\nu(q)$$

$$\leqq |\mathcal{A}| \sum_{q\in\mathcal{S}} \Lambda(q)\nu(q) + G\big(|\mathcal{A}|^2 - |\mathcal{A}|\big)\log N,$$

since for $d \neq 0$ one has that $\sum_{q|d} \Lambda(q) = \log|d| \leqq \log N$. This implies that

$$|\mathcal{A}|^2 \left(-G\log N + \sum_{q\in\mathcal{S}} \Lambda(q)\right) \leqq |\mathcal{A}| \left(-G\log N + \sum_{q\in\mathcal{S}} \Lambda(q)\nu(q)\right),$$

which proves the theorem. □

PROOF OF VARIANT 3. This is an immediate corollary of Gallagher's original version: Suppose that

$$\sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)} > 1 + \log N.$$

Then

$$|\mathcal{A}| \leqq \frac{-\log N + \sum_{q\in\mathcal{S}} \Lambda(q)}{-\log N + \sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)}} \leqq \frac{\sum_{q\in\mathcal{S}} \Lambda(q)}{1},$$

which contradicts the assumption $|\mathcal{A}| > \sum_{q\in\mathcal{S}} \Lambda(q)$. The same holds if one takes

$$\frac{1}{|\mathcal{A}|^2} \sum_{p\leqq Q} (\log p) \sum_{c=0}^{p-1} \big|\{a \in \mathcal{A} : a \equiv c \bmod p\}\big|^2$$

instead of $\sum_{p\leqq Q} \frac{\log p}{\nu(p)}$, by the intermediate steps of the proof of Gallagher's sieve. □

3. A standard example

The squares are the standard example of a sequence for which both Montgomery's and Gallagher's sieve, are sharp. As we could not find the details of this important example anywhere in the literature we decided to include them here.

Let $\mathcal{A} = \{n^2 : n \in \mathbb{N}\}$ be the sequence of the squares. These lie modulo odd primes in $\nu(p) = \frac{p+1}{2}$ residue classes. For $p = 2$ we have $\nu(p) = 2$. Gallagher's larger sieve (Theorem 2) and Variant 1 give the correct upper bound $\mathcal{A}(N) = O\big(N^{1/2}\big)$ which of course is best possible, apart from the O-constant. To see this (in the case of Theorem 2), one simply chooses $Q = cN^{1/2}$ with a sufficiently large constant c and obtains:

$$|\mathcal{A}| \leqq \frac{-\log N + \sum_{p\leqq Q} \log p}{-\log N + \sum_{p\leqq Q} \frac{\log p}{\nu(p)}} \ll \frac{Q}{-\log N + \frac{\log 2}{2} + \sum_{2<p\leqq Q} \frac{\log p}{(p+1)/2}}$$

$$\ll \frac{Q}{-\log N + 2\log Q + O(1)} \ll Q = O\big(N^{1/2}\big).$$

The O-constant can be improved by recalling that Gallagher's larger sieve can also sift modulo powers of primes, and observing that $a^2 \not\equiv kp \bmod p^2$, for $k = 1, \ldots, p-1$.

Montgomery's sieve also gives the correct order of magnitude. Here $\omega(2) = 0$ and $\omega(p) = \frac{p-1}{2}$ for odd primes. Since f defined by

$$f(q) = \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}$$

is a multiplicative function,

$$L = \sum_{q \leqq Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)}$$

can be evaluated by Wirsing's theorem, see for example Theorem II.4.1 of [20]. Here $\sum_{p \leqq x} f(p) \log p = \big(1 + o(1)\big)\, x$ holds and therefore $L \sim CQ$ for a suitable positive constant C. Again, with $Q = cN^{1/2}$ it follows that $\mathcal{A}(N) = O\big(N^{1/2}\big)$. As in the application of Gallagher's larger sieve, the fact that the sequence avoids modulo squares of odd primes $\frac{p(p-1)}{2} + p - 1$ classes can be used for a slight improvement of the O-constant. But one would have to appeal to an extension of Montgomery's sieve (see for example Johnsen [13], Gallagher [9], Selberg [21], Motohashi [16], and Ramaré and Ruzsa [17]). Here an application of formula (4) of [9] would suffice. In this applications it would be important to sieve as before modulo p the $\omega(p)$ classes, and only to use the $p - 1$ additional classes modulo p^2. An application of this sieve, where one sifts modulo p^2 *any* $\frac{p(p-1)}{2} + p - 1$ classes (ignoring the information modulo p) would lead to a much weaker result, only.

4. Discussion

We now discuss the advantages of the variants of the larger sieve, presented above. In many standard applications the original version and these variants are of the same strength. Let us consider a sieve problem with $\omega(p) = p - \nu(p) \approx cp$, where c is a constant with $0 < c < 1$. Of course, $\omega(p)$ and $\nu(p)$ are integers, so that usually one would have for example $\omega(p) = \frac{p+1}{2}$, if $c = \frac{1}{2}$, but let us put for simplicity $\omega(p) = p - \nu(p) \geqq cp$. Wirsing's theorem allows to estimate the denominator of Montgomery's sieve:

$$L = \sum_{q \leqq Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p - \omega(p)} \gg Q(\log Q)^{\frac{c}{1-c} - 1}.$$

Choosing $Q = \sqrt{N}$ gives

$$|\mathcal{A}| \ll \sqrt{N} (\log N)^{\frac{1-2c}{1-c}}.$$

Using Gallagher's larger sieve, we find that for some sufficiently large constant C the choice $\mathcal{S} = \{p : p \leqq Q = CN^{1-c}\}$ gives

$$|\mathcal{A}| \ll N^{1-c}.$$

This shows that for $\omega(p) < \frac{p}{2}$ one should use Montgomery's sieve. For $\omega(p) > \frac{p}{2}$ Gallagher's sieve is the preferable choice. For $c = \frac{1}{2}$ both sieves are of about the same strength as was studied before with the example of the squares. However, the upper bounds above suggest an asymmetry. If $\omega(p)$ oscillates around $\frac{p}{2}$ this oscillation can be used for a saving, using Gallagher's larger sieve.

If for example, for a constant $0 < \alpha < \frac{1}{2}$,

$$\omega(p) = \begin{cases} \dfrac{p}{2} + \dfrac{p}{2(\log p)^{\alpha}} & \text{if } p \equiv 1 \bmod 4 \\ \dfrac{p}{2} - \dfrac{p}{2(\log p)^{\alpha}} & \text{if } p \equiv 3 \bmod 4, \end{cases}$$

then one can work out that Montgomery's sieve will still give $|\mathcal{A}| \ll N^{1/2}$, whereas Gallagher's larger sieve will give an upper bound of

$$|\mathcal{A}| = O\left(\frac{N^{1/2}}{\exp\left(c_\alpha (\log N)^{1-2\alpha}\right)}\right),$$

for a suitable positive constant c_α.

Similarly, it is possible to construct an example, where the arithmetic mean of $\frac{\omega(p)}{p}$ is well below $\frac{1}{2}$, but where it is preferable to use Gallagher's sieve. Let

$$\omega(p) = \begin{cases} \dfrac{6}{7}p & \text{if } p \equiv 1 \bmod 4 \\ 0 & \text{if } p \equiv 3 \bmod 4. \end{cases}$$

Here the arithmetic mean is $\frac{3}{7}$, but using Gallagher's sieve one finds an upper bound of $O\left(N^{2/7}\right)$.

Even though these examples may appear to be artificial, let us remark that such a kind of asymmetry between Gallagher's and Montgomery's sieve was successfully used in Elsholtz [5].

For the case of cubes, $\mathcal{A} = \{n^3 : n \in \mathbb{N}\}$, Gallagher's sieve is not optimal. Here one has that

$$\nu(p) = \begin{cases} 2 & \text{if } p = 2,3 \\ p & \text{if } p \equiv -1 \bmod 6 \\ \dfrac{p+2}{3} & \text{if } p \equiv \quad 1 \bmod 6. \end{cases}$$

In an application of Montgomery's sieve such an oscillation influences the power of the logarithmic factor only, giving here an upper bound estimate of $\mathcal{A}(N) = O\big(N^{1/2}(\log N)^{1/2}\big)$. For an application of Gallagher's sieve in this case the optimal choice of the set $\mathcal{S}$ is not obvious. One considers whether it is better that the set $\mathcal{S}$ contains the primes $p \equiv -1 \bmod 6$ or not. If $\mathcal{S}$ contains all primes up to some level Q, then we find that

$$|\mathcal{A}| \ll \frac{-\log N + Q + O(1)}{-\log N + \frac{\log Q}{2} + \frac{3\log Q}{2} + O(1)} \ll \frac{Q}{-\log N + 2\log Q + O(1)} \ll N^{\frac{1}{2}},$$

with $Q = CN^{\frac{1}{2}}$. If $\mathcal{S}$ only contains the primes $p \equiv 1 \bmod 6$ up to some level Q, then we have

$$|\mathcal{A}| \ll \frac{-\log N + \frac{Q}{2} + O(1)}{-\log N + \frac{3\log Q}{2} + O(1)} \ll N^{\frac{2}{3}},$$

with $Q = CN^{\frac{2}{3}}$. Here, the fact that primes without any sifting effect were omitted even weakens the result.

The situation is different in the case of Variant 1. Here we choose $\mathcal{S} = \big\{p \leqq Q : p \equiv 1 \bmod 6\big\}$.

$$|\mathcal{A}| \ll \frac{N \exp\Big(\sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{q}\Big)}{\exp\Big(\sum_{q\in\mathcal{S}} \frac{\Lambda(q)}{\nu(q)}\Big)} = \frac{N \exp\big(\frac{1}{2}\log Q\big)}{\exp\big(\frac{3}{2}\log Q\big)} = \frac{NQ^{\frac{1}{2}}}{Q^{\frac{3}{2}}} = \frac{N}{Q}.$$

In making the optimal choice of Q we have to respect $\sum_{q\in\mathcal{S}} \Lambda(p) < |\mathcal{A}|$ (or $Q \ll |\mathcal{A}|$) so that

$$|\mathcal{A}| \ll N^{\frac{1}{2}},$$

with $Q = CN^{\frac{1}{2}}$, Interestingly enough, the other choice $\mathcal{S} = \{p : p \leqq Q\}$ leads to the very same result. It is an advantage of Variant 1 that an additional prime with $\nu(p) = p$ does not influence the result since here a factor

of $\exp\left(\frac{\Lambda(q)}{q}\right)$ occurs in the numerator and denominator, so that these factors cancel each other. So, in the case of Variant 1 the choice of $\mathcal{S}$ is easier here. It is only necessary to choose the optimal Q.

If $\nu(p) \approx p^\alpha$ (with $0 < \alpha < 1$), then Montgomery's sieve gives $\mathcal{A}(N) \ll N^{\frac{\alpha}{2}}$, whereas Gallagher's sieve with $\mathcal{S} = \left\{p : p \leqq Q = C(\log N)^{\frac{1}{1-\alpha}}\right\}$ shows that $\mathcal{A}(N) \ll (\log N)^{\frac{\alpha}{1-\alpha}}$. Variant 1 cannot handle this case. Moreover, if $|\mathcal{A}|$ is very small, then $|\mathcal{A}|^2 - |\mathcal{A}| < |\mathcal{A}|^2$ in equality (1) weakens the result.

Variant 2 cannot handle the above cases with $\omega(p) \approx cp$ since one needs that $G \leqq \frac{Q}{\log N}$. But it can deal with problems that use small values $\nu(p)$. For $\nu(p) \approx p^\alpha$ we also find with $Q = C(\log N)^{\frac{1}{1-\alpha}}$ that $\mathcal{A}(N) \ll (\log N)^{\frac{\alpha}{1-\alpha}}$.

Even though Gallagher's original version might be stronger than Variant 2, for many applications Variant 2 will completely suffice. In some applications it may be easier to have control over $\sum_p \nu(p) \log p$ rather than $\sum_p \frac{\log p}{\nu(p)}$. Moreover, the term $\sum_p \nu(p) \log p$ is more familiar in applications of the small sieve.

References

[1] E. Bombieri, Le grand crible dans la théorie analytique des nombres, second edition, *Astérisque,* **18** (1987).

[2] E. Croot and C. Elsholtz, On thin sets of primes expressible as sumsets, submitted.

[3] H. Davenport, *Multiplicative Number Theory,* third edition, revised by H. L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag (New York, 2000).

[4] A. Dujella, On the size of Diophantine m-tuples, *Math. Proc. Cambridge Philos. Soc.,* **132** (2002), 23–33.

[5] C. Elsholtz, The inverse Goldbach problem, *Mathematika,* **48** (2001), 151–158.

[6] C. Elsholtz, The distribution of sequences in residue classes, *Proc. Amer. Math. Soc.,* **130** (2002), 2247–2250.

[7] P. X. Gallagher, The large sieve, *Mathematika,* **14** (1967), 14–20.

[8] P. X. Gallagher, A larger sieve, *Acta Arith.,* **18** (1971), 77–81.

[9] P. X. Gallagher, Sieving by prime powers, *Acta Arith.,* **24** (1973/74), 491–497.

[10] K. Gyarmati, On a problem of Diophantus, *Acta Arith.,* **97** (2001), 53–65.

[11] H. Halberstam, and H.-E. Richert, *Sieve Methods,* London Mathematical Society Monographs, No. 4, Academic Press (London, 1974).

[12] N. Hegyvári and A. Sárközy, On Hilbert cubes in certain sets, *Ramanujan J.,* **3** (1999), 303–314.

[13] J. Johnsen, On the large sieve method in GF $[q, x]$, *Mathematika,* **18** (1971), 172–184.

[14] H. L. Montgomery, *Topics in Multiplicative Number Theory,* Lecture Notes in Mathematics, Vol. 227, Springer (Berlin-New York, 1971).

[15] H. L. Montgomery, The analytic principle of the large sieve, *Bull. Amer. Math. Soc.,* **84** (1978), 547–567.

[16] Y. Motohashi, *Lectures on Sieve Methods and Prime Number Theory,* Tata Institute Lecture Notes, 72. Springer (Berlin, 1983).

[17] O. Ramaré and I. Ruzsa, Additive properties of dense subsets of sifted sequences, *J. Théor. Nombres Bordeaux,* **13** (2001), 559–581.

[18] J. B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.,* **6** (1962), 64–94.

[19] A. Sárközy, A note on the arithmetic form of the large sieve, *Studia Sci. Math. Hungar.,* **27** (1992), 83–95.

[20] W. Schwarz and J. Spilker, *Arithmetical Functions,* Cambridge University Press, London Mathematical Society Lecture Notes 184 (Cambridge, 1994).

[21] A. Selberg, Remarks on multiplicative functions, in: *Number Theory Day* (Proc. Conf., Rockefeller Univ., New York, 1976), pp. 232–241. Lecture Notes in Math., Vol. 626, Springer (Berlin, 1977).

(Received October 24, 2002; revised November 5, 2003)

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF MATHEMATICS
125 SKILES
ATLANTA, GA 30332
U.S.A.
E-MAIL: ECROOT@MATH.GATECH.EDU

DEPARTMENT OF MATHEMATICS
ROYAL HOLLOWAY UNIVERSITY OF LONDON
EGHAM, SURREY TW20 0EX, UK
E-MAIL: CHRISTIAN.ELSHOLTZ@RHUL.AC.UK