1 Introduction

In this thesis we shall look at two different problems concerning the distribution of prime numbers. Two famous conjectures regarding the primes are:

- The twin prime conjecture, which asserts that there are infinitely many primes p for which $p+2$ is also a prime.
- The conjecture that the interval $[x, x+\sqrt{x}]$ always contains at least one prime whenever x is sufficiently large ($x \geq 117$ might be sufficient).

Although it appears to be impossible to turn these conjectures into theorems with present methods, it is possible to get quantitative results through sieve methods that show us how "far" we are from a solution. The above conjectures have been attacked in different ways; here we shall give a new upper bound for the number of pairs of twin primes below x of the form

$$(1)\ \pi_2(x) < k(1+\varepsilon)\frac{x}{\log^2 x} \text{ for every } x > x_0(\varepsilon)$$

(the constant k we give is new) and a result concerning the second conjecture of the form "The interval $\left[x, x+x^{\frac{1}{2}+\varepsilon}\right]$ always contains a number with a prime factor $> x^{24/25}$ for $x > x_0(\varepsilon)$".

The first nontrivial result regarding twin primes was given by Brun in 1919. He devised a rather complicated sieve method which gives (1) with a large k. A weaker, yet startling corollary of this is that whether or not there are infinitely many pairs of twin primes, the sum $\sum\left(\frac{1}{p}+\frac{1}{p+2}\right)$ over all pairs $(p, p+2)$ of twin primes converges. Selberg introduced a sieve method which is simpler, and for many purposes more efficient. As an application, (1) holds with $k = 16C_2$ where C_2, called the twin prime constant, is defined by

$$C_2 = \prod_{p \geq 3} \left(1 - \frac{1}{(p-1)^2}\right) = 0.6601618158...$$

It has been conjectured (for example, by Hardy and Littlewood) that $2C_2 \frac{x}{\log^2 x}$ is an asymptotically correct formula for the number of pairs of twin primes below x. (The constant C_2 has been calculated with high accuracy; see [W1].) It is plausible that if one can give a positive lower bound for $\pi_2(x)$ (and thus prove the twin prime conjecture), then one can also give a good upper bound. Attempts to give a best possible upper bound (as in this thesis) therefore carry valuable indications as to whether the twin prime conjecture is within reach. Alas, it doesn't seem to be at the moment.

However, a combination of Selberg's method with the Bombieri-Vinogradov Theorem is enough to replace 16 by 8 in the upper bound. I shall return to this theorem and other essential tools in Chapter 2. Compared to the reduction by a factor 2 here, only small improvements have been made since this was discovered in the 1960's. Pan [P] replaced 8 by 7.9280 on using the switching principle - a method involving rewriting one or more of the terms in the Buchstab identity. (Alt. spelling: Bukhshtab.) Chen [C] used a more advanced version and obtained 7.8342. Iwaniec [I2] found a new form of the error term in the linear sieve which allowed further improvements, even without using the switching principle. Fouvry and Iwaniec [F3] gave $\frac{68}{9} \approx 7.5556$, Fouvry [F1] gave $\frac{128}{17} \approx 7.5294$ and Bombieri, Friedlander and Iwaniec [B4] gave 7. The superior last paper has literally been a "starting point" for the further improvements which have been done (also the one we present here).

Fouvry and Grupp [F2] added the switching principle to the calculations and obtained 6.908 - more accurate calculations give 6.9075. Wu [W2] applied a more advanced version of the switching principle and obtained 6.8354.

We shall apply a method of Heath-Brown in order to get an improvement. The method involves replacing the characteristic function for prime numbers by a function more manageable with the available sieve machinery. The method has previously been used by Heath-Brown [H3] in his work on the problem about primes in short intervals. (Ironically, we shall not use this method in the treatment of that problem.) One disadvantage with this method is that, although it is clear that we can improve on Bombieri *et al.*'s estimate 7, it is not highly compatible with the switching principle. It turns out that we can combine it with Fouvry and Grupp's approach, but not with Wu's. It could therefore be regarded as pure luck that the constant we end up with in the end is 6.8325, which is only slightly better than Wu's result.

Our treatment of the twin prime problem contains a new result on the existence of subsets of a set of positive real numbers with sums in certain intervals. This is necessary in order to be able to use Heath-Brown's method.

There have been several different approaches to the second conjecture. The most common is possibly the one already mentioned - to give a largest possible constant t such that the interval $\left(x, x + x^{\frac{1}{2}+\varepsilon}\right]$ always contains an integer with a prime factor larger than x^t for $x > x_0(\varepsilon)$. Results of this type date back to Ramachandra, who gave $t = \frac{15}{26} = 0.5769...$ in [R1] (although it appears that $\frac{1}{\sqrt{3}} - \varepsilon = 0.5773... - \varepsilon$ is permissible), and later $\frac{5}{8} = 0.625$ in

[R2]. (In fact, he did this with $x^{\frac{1}{2}+\varepsilon}$ replaced by $x^{\frac{1}{2}}$.) Jutila [J] gave $t = \frac{2}{3} - \varepsilon = 0.666... - \varepsilon$. Balog improved this to 0.730... in [B1] and to 0.772... in [B2]. Balog, Harman and Pintz [B3] gave $t = 0.82$ and noted that 0.824 is attainable with refined calculations. Heath-Brown [H4] noted that $t = \frac{5}{6} - \varepsilon = 0.8333... - \varepsilon$ can be obtained through a combination of Jutila's method and his own generalized Vaughan identity [H3], and gave $t = \frac{11}{12} - \varepsilon = 0.91666... - \varepsilon$ on applying the method mentioned before. Later, Heath-Brown and Jia [H5] obtained $t = \frac{17}{18} - \varepsilon = 0.9444... - \varepsilon$ on applying certain mean-value theorems, and Harman [H2] obtained $t = \frac{19}{20} = 0.95$ on using a "rôle reversal trick". The part of our approach which is new, is an investigation of the four-dimensional terms in the iterated Buchstab identity. We also use Harman's rôle reversal trick, and we put some effort into making the calculations as accurate as possible. Thus we find that $t = \frac{24}{25} = 0.96$ is admissible. Without the rôle reversal trick I obtained $t = \frac{58}{61} = 0.9508...$, and this seemed to be very close to being optimal. Here we encounter some difficulties which make it hard to say just how far we are from a "best possible" result. But at least I think that, for example, $\frac{25}{26} \approx 0.9615$ is currently out of reach.

The calculation of the constants we obtain in both these problems involve lots of numerical integration. The improvement over the best previous result is very small for the twin prime problem, and the margin with which our result regarding almost-primes in intervals holds is also very small. It is therefore extremely important to be able to do the integration with high accuracy. It is essential that the method of integration is simple and efficient at the same time. We have chosen a special case of the Newton-Cotes Quadrature Formulas, namely:

$$\int_{x_0}^{x_0+6\triangle x} y(x)dx \approx \frac{\triangle x}{140}\left(41y_0 + 216y_1 + 27y_2 + 272y_3 + 27y_4 + 216y_5 + 41y_6\right)$$

.

where $y_i = y(x_0 + i\triangle x)$. See [K1], p. 771-772.

2 Preliminaries

In this Chapter, I will go through the most important background results, without labelling each result as a "lemma".

First, the Prime Number Theorem: Let $\pi(x)$ denote the number of prime numbers $\leq x$. Then,

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1$$

Define

$$\Lambda(n) = \begin{cases} \log p \text{ if } n \text{ is a power of a prime } p \\ 0 \text{ otherwise} \end{cases}$$

(called the von Mangoldt function), and

$$\psi(x; q, a) = \sum_{n\leq x, n\equiv a (\operatorname{mod} q)} \Lambda(n)$$

The Prime Number Theorem may be restated as

$$\psi(x; 1, 0) \sim x$$

Let $\mathcal{A}$ be a finite set of positive integers. The expression $S(\mathcal{A}, z)$ stands for the number of elements in $\mathcal{A}$ with no prime factor $< z$. $\mathcal{A}_d$ stands for the set of integers u for which $du \in \mathcal{A}$. The number of elements in $\mathcal{A}$ could be written as $|\mathcal{A}|$ or $S(\mathcal{A}, 1)$ but is normally written as X.

Throughout the text, p always denotes a prime number, e.g., $\prod_{p<m} \frac{p}{p-1}$ is the product over all primes p less than m of $\frac{p}{p-1}$. As usual, ε always denotes a "small" positive number, and can normally be replaced by $o(1)$.

I have chosen to treat the problem about primes in small intervals first, as the methods involved are not quite as deep as for the twin prime problem. An essential tool here is the Buchstab identity:

$$S(\mathcal{A},z) = S(\mathcal{A},w) - \sum_{w \le p < z} S(\mathcal{A}_p, p) \qquad (0 < w < z)$$

The proof is trivial: The first term on the RHS counts the elements with no prime factor $< w$, and the second term counts the terms whose smallest prime factor is in $[w,z)$. We can iterate this identity, i.e., we can apply it to the last term and possibly repeat the procedure. This leads to

$$S(\mathcal{A},z) = S(\mathcal{A},w) - \sum_{w \le p_1 < z} S(\mathcal{A}_{p_1}, w) + \sum_{w \le p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2}, w) - \ldots$$
$$\pm \sum_{w \le p_k < \ldots < p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 \ldots p_k}, p_k)$$

One could also have introduced different w-values.

In Chapter 3 we shall see that some of the terms involved here can be estimated asymptotically. The idea is then to deduce from Buchstab's identity a valid inequality $S(\mathcal{A},z) > \sum_1 \pm \sum_2 \pm \ldots$ where all the terms on the RHS can be estimated asymptotically and have a positive sum.

When dealing with twin primes, we resort to deeper results from sieve theory. I shall not attempt to derive the established results concerning the *linear sieve*. For any integer d, choose $\omega(d)$ so that $\frac{\omega(d)}{d}$ approximates $\frac{|\mathcal{A}_d|}{X}$, and define $R_d = |\mathcal{A}_d| - \frac{\omega(d)}{d}X$, $W(z) = \prod_{p<z} \left(1 - \frac{\omega(p)}{p}\right)$ and $P(z) = \prod_{p<z} p$ - this is standard sieve terminology. Following Iwaniec, we deal with a linear

sieve if the following inequalities hold:

$$\frac{W(w)}{W(z)} \leq \frac{\log z}{\log w}\left(1+\frac{K}{\log w}\right);$$

$$\sum_{w \leq p < z}\sum_{\alpha \geq 2}\frac{\omega(p^\alpha)}{p^\alpha} \leq \frac{L}{\log 3w}$$

which are to hold for all $z > w \geq 2$ with some constants $K, L > 1$. The "standard" way of treating the twin prime problem is to take $\mathcal{A}$ to be the set $\{p+2 \mid p < x\}$, which gives $\omega(p) = \frac{p}{p-1}$ for odd primes p, $\omega(2) = 0$ and $\omega(d) = \prod_{p|d} \omega(p)$ for square-free integers d. This gives:

$$W(z) = \prod_{3 \leq p < z} \frac{p-2}{p-1} \sim 2C_2 \prod_{p<z} \frac{p-1}{p} \sim \frac{2C_2}{e^\gamma \log z}$$

where C_2 was defined on page 2; see Chapter 3 of [H1] for details. For the parameters D and z, we put $s = \frac{\log D}{\log z}$, and we then have

$$S(\mathcal{A}, z) \leq XW(z)\{F(s) + A(s,K,D)\} + R(\mathcal{A}, D)$$

$$S(\mathcal{A}, z) \geq XW(z)\{f(s) - A(s,K,D)\} - R(\mathcal{A}, D)$$

where the functions $F(s)$ and $f(s)$ (due to Jurkat and Richert) are continuous and satisfy

$$\begin{cases} F(s) = \frac{2e^\gamma}{s},\ 0 < s \leq 3 \\ f(s) = 0,\ 0 < s \leq 2 \\ \frac{d}{ds}(sF(s)) = f(s-1),\ s > 3 \\ \frac{d}{ds}(sf(s)) = F(s-1),\ s > 2 \end{cases}$$

Here, the first error term $A(s,K,D)$ tends to zero as s or D approaches infinity and K remains constant. $R(\mathcal{A}, D)$ is defined to be $\sum_{d<D,\ d|P(z)} |R_d|$. We are now almost ready to

obtain the constant 8 previously announced as the upper bound for $\frac{\pi_2(x)}{C_2 \frac{x}{\log^2 x}}$. What we need is that the error term doesn't become too big if $D < x^{\frac{1}{2}-\varepsilon}$ say, and taking z to be any number between $D^{\frac{1}{3}+\varepsilon}$ and $D^{\frac{1}{2}-\varepsilon}$ then gives:

$$(2)\ \pi_2(x) < S(\mathcal{A}, z) + O(z) < (1+o(1))\{W(z)F(s)\}X$$

$$= (1+o(1))\left\{\frac{8C_2}{\log x}\right\}\frac{x}{\log x} = (1+o(1))8C_2\frac{x}{\log^2 x}$$

The result that takes care of the error term here is known as the *Bombieri-Vinogradov Theorem*, and says that

$$\sum_{q \leq Q} \max_{(a,q)=1} \left| \psi(x;q,a) - \frac{x}{\phi(q)} \right| < K\frac{x}{\log^A x}$$

for any $A > 0$ with $Q = \frac{\sqrt{x}}{\log^B x}$, where B and K depend on A alone.

A function related to F and f is the Buchstab (or Bukhshtab) function $\omega(s)$ which is continuous and satisfies

$$\begin{cases} \omega(s) = \frac{1}{s} \text{ for } 0 < s \leq 2 \\ \frac{d}{ds}\left(s\omega(s)\right) = \omega(s-1) \text{ for } s > 2 \end{cases}$$

If $s > 1$, then the number of integers $\leq x$ with no prime factor $\leq x^{\frac{1}{s}}$ is asymptotically $s\omega(s)\frac{x}{\log x}$; this is *equivalent* to the Prime Number Theorem.

Iwaniec [I2] found a new form of the error term in the linear sieve. First, we need to define *well factorable functions*. Let Q be any positive integer, and let λ be an arithmetic function defined for the integers from 1 to Q. Suppose it has the property that for any Q_1, $Q_2 \geq 1$, $Q_1Q_2 = Q$ (Q_1 and Q_2 are not necessarily integers) there exist two functions λ_1, λ_2 supported in $[1, [Q_1]]$ and $[1, [Q_2]]$ respectively such that $|\lambda_1| \leq 1$, $|\lambda_2| \leq 1$ and $\lambda = \lambda_1 * \lambda_2$, i.e.,

$$\sum_{q=1}^{Q} \frac{\lambda(q)}{q^s} = \left(\sum_{q_1=1}^{[Q_1]} \frac{\lambda_1(q_1)}{q_1^s} \right) \left(\sum_{q_2=1}^{[Q_2]} \frac{\lambda_2(q_2)}{q_2^s} \right).$$

Then we say that λ is well factorable of level Q. In the paper just referred to, an example of a well factorable function can be found.

When dealing with convolutions (e.g., $\lambda = \lambda_1 * \lambda_2$), we often refer to the components as "factors" (λ_1 is thus a factor in λ in the example).

Iwaniec proved that for $0 < \delta < \frac{1}{8}$ and $s = \frac{\log Q}{\log z}$ we have

$$S(\mathcal{A}, z) \leq XW(z) \left\{ F(s) + O\left(\delta + \delta^{-8} e^K (\log Q)^{-\frac{1}{3}} \right) \right\} + \sum_{l < \exp(8\delta^{-3})} \sum_{q | P(z)} \lambda_l(q) R_q$$

$$S(\mathcal{A}, z) \geq XW(z) \left\{ f(s) - O\left(\delta + \delta^{-8} e^K (\log Q)^{-\frac{1}{3}} \right) \right\} - \sum_{l < \exp(8\delta^{-3})} \sum_{q | P(z)} \lambda_l(q) R_q$$

where each λ_l is well factorable of level Q; note that the parameter D is typically replaced by Q when we deal with well factorable functions. With the absolute sign being replaced by multiplication by a well factorable function, it is possible to improve on the Bombieri-Vinogradov Theorem. Indeed, Bombieri *et al.* [B4] proved that for $a \neq 0$, $\varepsilon > 0$ and $Q = x^{\frac{4}{7}-\varepsilon}$, we have

$$\sum_{(q,a)=1} \lambda(q) \left(\psi(x; q, a) - \frac{x}{\phi(q)} \right) \ll x (\log x)^{-A}$$

for any well factorable function λ of level Q and any $A > 0$. The constant implied in $\ll$ depends at most on ε, a and A. Whereas it is easy to see that this is sufficient to allow us to replace 8 by 7 in (2), our task will be to see how much improvement we can pull out of Bombieri *et al.*'s methods on using more "manageable" functions than ψ. This is "legal" as long as they are strictly bigger than ψ (or, equivalently, some other function which has a "direct" connection to $\pi(n)$). Although we won't worry about the technical parts of Bombieri *et al.*'s method, let me mention that it is based on the dispersion method, Fourier analysis and Kloosterman sums, rather than the "large sieve inequality" on which the Bombieri-Vinogradov Theorem is based.

When working with problems of this kind, it is often convenient to avoid having to deal with numbers with too small factors. This is done using the *fundamental lemma* of sieve theory: Let $D \geq 2$, $z = D^{1/s}$ with $s \geq 3$. There exist two sequences $\{\lambda_d^+\}_{d \leq D}$ and $\{\lambda_d^-\}_{d \leq D}$ such that

$$\begin{cases} \left|\lambda_d^{\pm}\right| \leq 1 \\ \left(\lambda^- * 1\right)(n) = 1 = \left(\lambda^+ * 1\right)(n) \text{ if } n \text{ has no prime factor } < z \\ \left(\lambda^- * 1\right)(n) \leq 0 \leq \left(\lambda^+ * 1\right)(n) \text{ otherwise} \\ \sum_{d \leq D} \frac{\lambda_d^{\pm}}{d} = \prod_{p<z} \left(1 - \frac{1}{p}\right)\left(1 + O\left(\exp(-s \log s)\right)\right) \end{cases}$$

In the end, we will also need to apply the *switching principle.* We start by deriving an inequality of Pan, used by Fouvry and Grupp, with an improvement noted by Wu. Suppose $0 < z_1 < z$. The first two iterations of the Buchstab identity give

$$S(\mathcal{A}, z) = S(\mathcal{A}, z_1) - \sum_{z_1 \leq p_1 < z} S(\mathcal{A}_{p_1}, p_1)$$

$$S(\mathcal{A}, z) = S(\mathcal{A}, z_1) - \sum_{z_1 \leq p < z} S(\mathcal{A}_p, z_1) + \sum_{z_1 \leq p_3 < p_1 < z} S(\mathcal{A}_{p_1 p_3}, p_3)$$

Adding together yields

$$2S(\mathcal{A}, z) = 2S(\mathcal{A}, z_1) - \left\{ \sum_{z_1 \leq p < z} S(\mathcal{A}_p, z_1) \right\} + \left\{ \sum_{z_1 \leq p_3 < p_1 < z} S(\mathcal{A}_{p_1 p_3}, p_3) - \sum_{z_1 \leq p_1 < z} S(\mathcal{A}_{p_1}, p_1) \right\}$$

$= 2S(\mathcal{A}, z_1) - \Omega_1 + M$, say. We have

$$M = \sum_{z_1 \le p_3 < p_1 < z} S(\mathcal{A}_{p_1 p_3}, p_3) - \sum_{z_1 \le p_3 < z} S(\mathcal{A}_{p_3}, p_3)$$

$$= \left(\sum_{z_1 \le p_3 < p_1 < z} S(\mathcal{A}_{p_1 p_3}, p_1) + \sum_{z_1 \le p_3 \le p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 p_3}, p_2) \right)$$

$$- \left(\sum_{z_1 \le p_3 < z} S(\mathcal{A}_{p_3}, z) + \sum_{z_1 \le p_3 \le p_1 < z} S(\mathcal{A}_{p_1 p_3}, p_1) \right)$$

$$\le \sum_{z_1 \le p_3 \le p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 p_3}, p_2)$$

$$= \sum_{z_1 \le p_3 < p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 p_3}, p_2) + \sum_{z_1 \le p_3 < p_1 < z} S(\mathcal{A}_{p_1 p_3^2}, p_3)$$

On replacing the last term by the trivial upper bound $O\left(\frac{x}{z_1}\right)$, we get (with $\mathcal{A}$ being as above)

$$\pi_2(x) \le S(\mathcal{A}, z) + O(z) \le S(\mathcal{A}, z_1) - \frac{1}{2}\Omega_1 + \frac{1}{2}\Omega_2 + O(z) + O\left(\frac{x}{z_1}\right)$$

where $\Omega_2 = \sum_{z_1 \le p_3 < p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 p_3}, p_2)$. The trick here is to let $\mathcal{B}$ be the sequence of numbers less than x of the form $p_1 p_2 p_3 m - 2$ where $z_1 \le p_3 < p_2 < p_1 < z$ and every prime factor in m is bigger than p_2. We then have

$$\Omega_2 = S(\mathcal{B}, x^{\frac{1}{2}}) + O(x^{\frac{1}{2}})$$

and in this way we can achieve a better upper bound.

3 Plan for 'almost-primes in intervals'

I shall now outline how one shows that there is always a number in the interval $\left(x, x+x^{\frac{1}{2}+\varepsilon}\right]$ with a prime factor $> x^{1-\theta}$ for $\theta = \frac{1}{25}$. Let K be a large fixed number, and let $P = x^{\theta/K}$. Rather than sifting the integers in the interval $\mathcal{I}(x) = \left(x, x+x^{\frac{1}{2}+\varepsilon}\right]$ directly, we shall look for primes in $\mathcal{A}$, the set of integers n for which $nm \in \mathcal{I}(x)$ for some m for which the coefficient for m^{-s} in $\left(\sum_{P<p\le 2P} p^{-s}\right)^K$ is positive (counted with their multiplicity). This is perhaps a cumbersome way of saying that m is the product of K prime numbers in the interval $(P, 2P]$, but an essential part of the calculations makes use of Dirichlet polynomials. Note that $P^K \simeq x^\theta$. We will also make heavy use of the Buchstab identity.

Let V be a subregion of $[0,1]^k$. Then we have the following results of Heath-Brown (see, for example, [H4]):

Theorem 1 *We can give an asymptotic formula for*

$$\sum_{\left(\frac{\log p_1}{\log x},\ldots,\frac{\log p_k}{\log x}\right)\subseteq V} S\left(\mathcal{A}_{p_1\ldots p_k}, q\right)$$

if for every point $(\alpha_1,\ldots,\alpha_k) \in V$ *we have* $\sum_{i\in M} \alpha_i \in \left[\frac{1}{2}-\theta, \frac{1}{2}\right]$ *for some subset* M *of* $\{1,2,\ldots,k\}$.

Theorem 2 *We can give an asymptotic formula for*

$$\sum_{\left(\frac{\log p_1}{\log x},\ldots,\frac{\log p_k}{\log x}\right)\subseteq V} S\left(\mathcal{A}_{p_1\ldots p_k}, x^\theta\right)$$

if for every point $(\alpha_1,\ldots,\alpha_k) \in V$ *we have* $\sum_{i\in M} \alpha_i \le \frac{1}{2}$ *and* $\sum_{i\notin M} \alpha_i \le \frac{1}{4}$ *for some subset* M *of* $\{1,2,\ldots,k\}$.

What we need is a lower bound for $S(\mathcal{A}, x^{\frac{1-\theta}{2}})$, and the Buchstab identity gives

$$S(\mathcal{A}, x^{\frac{1-\theta}{2}}) = S(\mathcal{A}, x^{\frac{1}{2}-\theta}) - \sum_{x^{\frac{1}{2}-\theta} \leq p < x^{\frac{1-\theta}{2}}} S(\mathcal{A}_p, p)$$

where the sum on the RHS can be estimated by Theorem 1. Consider what is left:

$$S(\mathcal{A}, x^{\frac{1}{2}-\theta}) = S(\mathcal{A}, x^{\theta}) - \sum_{x^{\theta} \leq p < x^{\frac{1}{2}-\theta}} S(\mathcal{A}_p, p)$$

(the first term on the RHS is OK by Theorem 2 but not the second)

$$(3) \quad = S(\mathcal{A}, x^{\theta}) - \sum_{x^{\theta} \leq p < x^{\frac{1}{2}-\theta}} S(\mathcal{A}_p, x^{\theta}) + \sum_{x^{\theta} \leq p_2 < p_1 < x^{\frac{1}{2}-\theta}} S(\mathcal{A}_{p_1p_2}, p_2).$$

The second term is now OK by Theorem 2. The third term must be further examined.

The first thing one should note here is that $\sum_{x^{\theta} \leq p_2 < p_1 < x^{\frac{1}{2}-\theta},\ p_1p_2^2 > x^{1-\theta}} S(\mathcal{A}_{p_1p_2}, p_2)$ is zero. More generally, when dealing with $\sum_{\left(\frac{\log p_1}{\log x}, \ldots, \frac{\log p_k}{\log x}\right) \subseteq V} S\left(\mathcal{A}_{p_1 \ldots p_k}, p_k\right)$, we may assume that $\alpha_1 + \ldots + \alpha_{k-1} + 2\alpha_k \leq 1 - \theta$.

At this stage we could stop, and say that

$$S(\mathcal{A}, x^{\frac{1}{2}-\theta}) \geq S(\mathcal{A}, x^{\theta}) - \sum_{x^{\theta} \leq p < x^{\frac{1}{2}-\theta}} S(\mathcal{A}_p, x^{\theta}) + \sum_{x^{\theta} \leq p_2 < p_1 < x^{\frac{1}{2}-\theta},\ p_1p_2 \in \left[x^{\frac{1}{2}-\theta}, x^{\frac{1}{2}}\right]} S(\mathcal{A}_{p_1p_2}, p_2)$$

where all terms on the RHS can be estimated asymptotically. However, very often it pays off to *expand* (i.e., apply Buchstab's identity) rather than *discard* terms. In the case $\theta \approx 0$ it is known that, essentially, all terms of the form

$$\sum_{p_1 \ldots p_{2k-1} p_{2k}^3 > x^{1-\theta}} S(\mathcal{A}_{p_1 \ldots p_{2k}}, p_{2k})$$

should be discarded (when only using the Buchstab identity). See [I1] for details. But for larger values of θ, the possibility that Theorem 1 applies to the new terms makes expanding profitable in a larger range. In particular, Monte Carlo methods suggest that for $\theta = \frac{1}{25}$ this *always* pays off, as long as it is possible (see below), except possibly in some regions that are so small that the difference in whether we do it one way or the other hardly matters. (This is the conclusion of some testing I did on a computer.)

Generally, one still has to be careful when deciding whether or not to expand. Consider the expansion of $\sum_{\left(\frac{\log p_1}{\log x},\dots,\frac{\log p_k}{\log x}\right)\subseteq V} S\left(\mathcal{A}_{p_1\dots p_k}, p_k\right)$, assuming that Theorem 1 does not apply to any subregion of V. We assume that k is even, so that the sign is positive. For each term in the sum we have

$$S\left(\mathcal{A}_{p_1\dots p_k}, p_k\right) = S\left(\mathcal{A}_{p_1\dots p_k}, x^\theta\right) - \sum_{x^\theta \le p_{k+1} < p_k} S\left(\mathcal{A}_{p_1\dots p_{k+1}}, p_{k+1}\right)$$

$$= S\left(\mathcal{A}_{p_1\dots p_k}, x^\theta\right) - \sum_{x^\theta \le p_{k+1} < p_k} S\left(\mathcal{A}_{p_1\dots p_{k+1}}, x^\theta\right) + \sum_{x^\theta \le p_{k+2} < p_{k+1} < p_k} S\left(\mathcal{A}_{p_1\dots p_{k+2}}, p_{k+2}\right).$$

It is essential that the first two terms on the RHS can be estimated asymptotically. This is because the first one is bigger than the LHS, and the second term is negative. Therefore, we must make sure that for every point $(\alpha_1, \dots, \alpha_k)$ in V and every $\alpha_{k+1} \in \{0\} \cup [\theta, \alpha_k]$, with $\alpha_1 + \dots + \alpha_k + 2\alpha_{k+1} \le 1 - \theta$, there is a subset M of $\{1, 2, \dots, k+1\}$ such that either $\sum_{i\in M} \alpha_i \in \left[\frac{1}{2} - \theta, \frac{1}{2}\right]$, or $\sum_{i\in M} \alpha_i \le \frac{1}{2}$ and $\sum_{i\notin M} \alpha_i \le \frac{1}{4}$. It would not make any difference if we took the trouble to check if Theorem 1 applies to any of the terms $S\left(\mathcal{A}_{p_1\dots p_{k+1}}, p_{k+1}\right)$ after the first iteration, as it would necessarily apply to any term in its expansion.

When this procedure is repeated, one eventually ends up with an inequality

$$S(\mathcal{A}, x^{\frac{1-\theta}{2}}) \ge \sum\nolimits_1 \pm \sum\nolimits_2 \pm \dots$$

where each sum $\sum_i$ can be estimated asymptotically by Theorems 1 and 2. The RHS is equal to

$$S(\mathcal{A}, x^{\frac{1-\theta}{2}}) - \sum\nolimits_1' - \sum\nolimits_2' - ...$$

where the $\sum_i'$-s are sums over the regions that we have discarded. Typically, these "bad" regions consist mostly of disjoint convex regions that are relatively easy to describe (at least in 4 dimensions; see Chapter 4). We shall use this observation to ease the calculations in the following way.

Let $y = x^{\frac{1}{2}+\varepsilon}$, so that we have $\mathcal{I}(x) = (x, x+y]$. Each sum $\sum_i$ over a "good" region is asymptotically $\frac{y}{x}\sum_i''$ where $\sum_i''$ is a sum over the same region, but in which $\mathcal{A}$ has been replaced by $\mathcal{B}$. $\mathcal{B}$ is defined like $\mathcal{A}$, with $\mathcal{I}(x)$ replaced by $(x, 2x]$. (I will soon return to why this is true.) If we define each $\sum_i'''$ by $\sum_i'$ in a corresponding way, we get

$$S(\mathcal{A}, x^{\frac{1-\theta}{2}}) \geq \frac{y}{x}\left(S(\mathcal{B}, x^{\frac{1-\theta}{2}}) - \sum\nolimits_1''' - \sum\nolimits_2''' - ...\right)$$

and the RHS can be estimated asymptotically with the Prime Number Theorem (and some numerical integration).

In order to treat the last term $\sum\limits_{x^\theta \leq p_2 < p_1 < x^{\frac{1}{2}-\theta}} S(\mathcal{A}_{p_1p_2}, p_2)$ in (3), we divide the corresponding region V into six subregions according to how one proceeds. First, note that in this case we have

$$V = \left\{(\alpha_1, \alpha_2) \mid \theta \leq \alpha_2 \leq \alpha_1 \leq \frac{1}{2} - \theta, \alpha_1 + 2\alpha_2 \leq 1 - \theta\right\}$$

by previous remarks. We now write V as a disjoint union of six sets, save that they may in pairs have common boundaries. Thus $V = A \cup B \cup C \cup D \cup E \cup F$ where

$$A = \left\{(\alpha_1, \alpha_2) \mid \alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_1 - \alpha_2 \leq \theta, \alpha_2 \leq \frac{1}{4}\right\}$$

.

$$B = \left\{(\alpha_1, \alpha_2) \mid \alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_1 \leq \frac{1}{2} - \theta, \alpha_2 \leq \frac{1}{8}\right\}$$
$$C = \left\{(\alpha_1, \alpha_2) \mid \theta \leq \alpha_2 \leq \alpha_1, \alpha_1 + \alpha_2 \leq \frac{1}{2} - \theta\right\}$$
$$D = \left\{(\alpha_1, \alpha_2) \mid \theta \leq \alpha_2 \leq \alpha_1, \frac{1}{2} - \theta \leq \alpha_1 + \alpha_2 \leq \frac{1}{2}\right\}$$
$$E = \left\{(\alpha_1, \alpha_2) \mid \frac{1}{4} \leq \alpha_2 \leq \alpha_1, \alpha_1 + 2\alpha_2 \leq 1 - \theta\right\}$$
$$F = \left\{(\alpha_1, \alpha_2) \mid \frac{1}{8} \leq \alpha_2 \leq \frac{1}{4}, \alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_2 + \theta \leq \alpha_1 \leq \frac{1}{2} - \theta\right\}$$

In the interior of $A \cup B \cup C$, Theorem 1 doesn't apply to $\sum S(\mathcal{A}_{p_1p_2}, p_2)$, but in this region we can always expand this to

$$\sum S(\mathcal{A}_{p_1p_2}, x^\theta) - \sum_{x^\theta \leq p_3 < p_2} S(\mathcal{A}_{p_1p_2p_3}, x^\theta) + \sum_{x^\theta \leq p_4 < p_3 < p_2} S(\mathcal{A}_{p_1p_2p_3p_4}, p_4)$$

The first two terms can be estimated because: In A, we have

$$\alpha_3 \leq \frac{1 - \theta - \alpha_1 - \alpha_2}{2}$$
$$\Rightarrow \alpha_1 + \alpha_3 \leq \frac{1 - \theta + \alpha_1 - \alpha_2}{2} \leq \frac{1}{2}$$

and $\alpha_2 \leq \frac{1}{4}$. In B we have $\alpha_1 \leq \frac{1}{2}$ and $\alpha_2 + \alpha_3 \leq 2\alpha_2 \leq \frac{1}{4}$, and in C we have $\alpha_1 + \alpha_2 \leq \frac{1}{2}$ and $\alpha_3 \leq \alpha_2 \leq \frac{1}{4}$. Obviously Theorem 1 does apply in D, as was pointed out before, so we do not have to do anything more with that. In contrast, the whole of E must be discarded, because $\alpha_1 \geq \alpha_2 \geq \frac{1}{4}$.

It is in F that we will apply the so-called rôle reversal trick. First, note that for every point (α_1, α_2) in F there is an α_3 such that neither Theorem applies: If $\alpha_1 + 3\alpha_2 \leq 1 - \theta$, then we may choose $\alpha_3 = \alpha_2$. Then $\alpha_1 + \alpha_3 \geq \frac{1}{2}$ and

$$\alpha_2 + \alpha_3 = 2\alpha_2 = (\alpha_1 + 3\alpha_2) - (\alpha_1 + \alpha_2) \leq (1 - \theta) - \frac{1}{2} = \frac{1}{2} - \theta$$

which shows that Theorem 1 does not apply. We also have $\alpha_2+\alpha_3=2\alpha_2\geq\frac{1}{4}$ which shows that Theorem 2 doesn't apply either. If $\alpha_1+3\alpha_2\geq 1-\theta$ then choose $\alpha_3=\frac{1-\theta-\alpha_1-\alpha_2}{2}$; this gives the same conclusions as we have

$$\alpha_1+\alpha_3=\frac{1-\theta+\alpha_1-\alpha_2}{2}\geq\frac{1}{2}$$

$$\alpha_2+\alpha_3=\frac{1-\theta-\alpha_1+\alpha_2}{2}\leq\frac{1}{2}-\theta$$

$$\begin{aligned}\alpha_2+\alpha_3=\frac{1-\theta-\alpha_1+\alpha_2}{2}&=\frac{1-\theta+(\alpha_1+3\alpha_2)-2(\alpha_1+\alpha_2)}{2}\\ &\geq\frac{1-\theta+(1-\theta)-2\left(\frac{3}{4}-\theta\right)}{2}=\frac{1}{4}\end{aligned}$$

This means that we cannot expand in the usual way, because we would encounter terms

$$-\sum S\left(\mathcal{A}_{p_1p_2p_3},x^\theta\right)$$

which cannot be estimated. But we have that

$$\begin{aligned}-\sum S\left(\mathcal{A}_{p_1p_2p_3},p_3\right)&=-\sum_{p|t\Rightarrow p\geq p_3}S\left(\mathcal{A}_{p_2p_3t},\sqrt{\frac{x^{1-\theta}}{p_2p_3t}}\right)\\ &=-\sum_{p|t\Rightarrow p\geq p_3}S\left(\mathcal{A}_{p_2p_3t},x^\theta\right)+\sum_{p|t\Rightarrow p\geq p_3}S\left(\mathcal{A}_{p_2p_3tp_4},p_4\right)\end{aligned}$$

where Theorem 2 applies to the first term on the RHS. Indeed, letting

$$\beta=\frac{\log t}{\log x}=1-\theta-\alpha_1-\alpha_2-\alpha_3$$

gives $\alpha_3+\beta=1-\theta-\alpha_1-\alpha_2\leq\frac{1}{2}-\theta$ and $\alpha_2\leq\frac{1}{4}$. For the last term, one estimates the part where Theorem 1 applies and discards the rest. With the large range for p_4 (which is $\left[x^\theta,\sqrt{p_1}\right]$), one expects that it is better to expand as normally if the value of p_3 allows it. However, in some parts of F, it is still better to just discard the 2-factor term. It seems like the best procedure is to let a computer check lots of small subregions and determine which choice is better in each one. The problem is that it appears to be very complicated to classify the regions in which there is no sum of a subset of

$$\{\alpha_2, \alpha_3, 1-\theta-\alpha_1-\alpha_2-\alpha_3, \alpha_4\}$$

lying in $\left[\frac{1}{2}-\theta, \frac{1}{2}\right]$. On letting a computer include a whole "box" in the integration if there is only one point inside having this property, one does get a valid upper bound for the contribution, but one loses accuracy. There are a few other regions where we use this procedure because of the great difficulties this avoids. In particular, our treatment of the "bad" regions in 6 or more dimensions will be fairly brief, because the numerical contribution is small. I will return to this matter in Chapters 4 and 5.

Before turning to Chapter 4, the classification of the 4-dimensional discarded terms, I will show how Theorem 1 and 2 are proven. Let

$$I(\xi, T) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{\xi^s}{s} ds$$

where ξ, T and c are positive real numbers. Then we have

$$I(\xi, +\infty) = \lim_{T \to +\infty} I(\xi, T) = \begin{cases} 0,\ 0 < \xi < 1 \\ \frac{1}{2},\ \xi = 1 \\ 1,\ \xi > 1 \end{cases}$$

This leads to *Perron's formula*:

$$\sum_{n \le x}{}' a_n = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} A(s) \frac{x^s}{s} ds$$

where $A(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ is absolutely convergent for $\operatorname{Re} s = c$. $\sum'$ means that if x is an integer, then $\frac{a_x}{2}$ comes instead of a_x into the sum. Moreover, we have the estimate

$$|I(\xi, T) - I(\xi, +\infty)| < \begin{cases} \xi^c \min\left(1, \frac{1}{T|\log \xi|}\right) \text{ if } \xi \neq 1 \\ \frac{c}{T} \text{ if } \xi = 1 \end{cases}$$

Let $y = x^{\frac{1}{2}+\varepsilon}$ (as above). Then $\sum_{\left(\frac{\log p_1}{\log x}, \dots, \frac{\log p_k}{\log x}\right) \subseteq V} S\left(\mathcal{A}_{p_1 \dots p_k}, q\right)$ can be written as

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(\sum \frac{1}{p_1^s}\right) \dots \left(\sum \frac{1}{p_k^s}\right) \left(\sum_{P<p\leq 2P} \frac{1}{p^s}\right)^K \left(\sum_{r|n \Rightarrow r>q} \frac{1}{n^s}\right) \frac{(x+y)^s - x^s}{s} ds$$

where the summations are still over the range

$$\left(\frac{\log p_1}{\log x}, \dots, \frac{\log p_k}{\log x}\right) \subseteq V$$

There are various ways in which one can estimate this. Harman uses Van der Corput-estimates; Heath-Brown has a more elementary approach. In either case, one gets an estimate which has a main term equal to

$$\frac{y}{x} \sum_{\left(\frac{\log p_1}{\log x}, \dots, \frac{\log p_k}{\log x}\right) \subseteq V} S\left(\mathcal{B}_{p_1 \dots p_k}, q\right)$$

as already mentioned, and an error term

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \left(\sum \frac{1}{p_1^s}\right) \dots \left(\sum \frac{1}{p_k^s}\right) \left(\sum_{P<p\leq 2P} \frac{1}{p^s}\right)^{K-1} \left(\sum_{r|n \Rightarrow r>q} \frac{1}{n^s}\right) ds$$

where $T = x^{\frac{1-\varepsilon}{2}}$. From here one only needs to show that $\frac{\text{error term}}{\text{main term}} \to 0$, as the main term can be estimated asymptotically with the Prime Number Theorem. In particular, it is sufficient to show that the error term is $O(T \log^J x)$ for some constant J. Now, under the hypothesis of Theorem 1 this can be written as

$$\frac{1}{2\pi i}\int F(s)G(s)ds$$

with F and G being Dirichlet polynomials supported on integers below $\sqrt{x}$. Thus we can apply Hölder's inequality to obtain

$$\frac{1}{2\pi i}\int F(s)G(s)ds \ll \left(\int |F(s)|^2\, ds\right)^{\frac{1}{2}} \left(\int |G(s)|^2\, ds\right)^{\frac{1}{2}} \ll T^{\frac{1}{2}+\frac{1}{2}} \log^{J_1} x = T\log^{J_1} x$$

Under the hypothesis of Theorem 2 we have $q = x^\theta$, which means that the integrand in the error term can be written as $F(s)G(s)Z(s)$ where $F(s)$ is supported below $\sqrt{x}$, $G(s)$ is supported below $x^{\frac{1}{4}}$, and $Z(s)$ is sufficiently "similar" to $\zeta(s)$ that we can use the fourth power moment estimate for this function if the support is above $x^{\frac{1}{4}}$ and obtain

$$\frac{1}{2\pi i}\int F(s)G(s)Z(s)ds \ll \left(\int |F(s)|^2\, ds\right)^{\frac{1}{2}} \left(\int |G(s)|^4\, ds\right)^{\frac{1}{4}} \left(\int |Z(s)|^4\, ds\right)^{\frac{1}{4}}$$

$$\ll T^{\frac{1}{2}+\frac{1}{4}+\frac{1}{4}} \log^{J_2} x = T\log^{J_2} x$$

If the support for Z is below $x^{\frac{1}{4}}$ then one takes G and Z together and proceeds as in the first case. A fuller explanation is given in [H4].

Can we get more out of Hölder's inequality, i.e., add to the list of theorems one can use? The answer is that we can't do more that is relevant in our discussion (as our treatment of terms with 6 or more factors is very simplified), but that it is actually possible to write 1 as the sum of reciprocals of even numbers in other ways than starting with $\frac{1}{2} + \frac{1}{2}$ and splitting up each term. The following result covers possible analogues of both Theorem 1 and Theorem 2.

Proposition 3 *Suppose that* $\{n_i\}_{1\le i\le k}$ *is a finite sequence of positive integers such that* $n_1 \le ... \le n_k$ *and* $\sum_{1\le i\le k} \frac{1}{n_i} = 2$*, and that for any proper subset* S *of* $\{1, 2, ..., k\}$ *with at least two elements, the reciprocal of* $\sum_{i\in S} \frac{1}{n_i}$ *is not an integer. Then one of these holds:*

- $k = 2$ and $n_1 = n_2 = 1$
- $k = 8$ and $n_1 = 2$, $n_2 = n_3 = 3$, $n_4 = n_5 = n_6 = n_7 = 5$, $n_8 = 30$
- $k \ge 9$ and $n_k \ge 28$

The proof is rather cumbersome, but the lighter version with the last two possibilities replaced by "$n_k \ge 20$" is not difficult to see, and I will give the proof here. Note that it is essential that we can prove that $n_k \ge \frac{1}{2\theta} = 12.5$ if $n_1 = n_2 = 1$ is excluded, in order that we can conclude that Theorem 1 and Theorem 2 are in a sense best possible.

By a *prime power*, I will mean a positive power of a prime number (i.e., 1 is excluded).

Case 1: Suppose q is a prime power and that no multiples of q, possibly except q itself, are in the sequence. To avoid a multiple of q in the denominator of $\sum \frac{1}{n_i}$, the number r of occurrences of q must satisfy $(q, r) > 1$. But if $r \ge 1$, this would conflict the property of the sequence, so q itself can not be in the sequence either.

Case 2: Suppose q is a prime power and that no multiples of q, possibly except q and/or $2q$, are in the sequence. If the number of occurrences of $2q$ is zero, then we are back in Case 1. If the number of occurrences of $2q$ is ≥ 2, then we have $\frac{1}{2q} + \frac{1}{2q} = \frac{1}{q}$, so the condition of the sequence is violated. So we may assume that there is exactly one occurrence of $2q$. Then, to avoid a multiple of q in the denominator of $\sum \frac{1}{n_i}$, there must be r occurrences of q

where $(2r+1, 2q) = (2r+1, q) > 1$. But the reciprocal of

$$\frac{(2r+1,q)-1}{2} \times \frac{1}{q} + \frac{1}{2q} = \frac{(2r+1,q)}{2q}$$

would then be an integer. So neither q nor $2q$ is permissible.

So suppose there is such a sequence other than $\{1, 1\}$. We are assuming that the largest integer in the sequence is at most 19, so that Case 1 rules out 11, 13, 16, 17 and 19. Having ruled out 16, Case 1 also rules out 8. The only remaining multiples of 4 are 4 and 12. To avoid a multiple of 4 in the denominator of $\sum \frac{1}{n_i}$, we can't have one without the other, but $\frac{1}{4} + \frac{1}{12} = \frac{1}{3}$, so we can rule them out too. Case 2 rules out 7, 9, 14 and 18. The remaining possible integers are: 2, 3, 5, 6, 10, 15. If there are any multiples of 5 in the sequence, then 15 must be one of them, by Case 2. But since $\frac{1}{10} + \frac{1}{15} = \frac{1}{6}$, this rules out 10. There can not be as many as two occurrences of 15; indeed, there would then have to be another multiple of 5 in the sequence to avoid a multiple of 5 in the denominator of $\sum \frac{1}{n_i}$, and $\frac{3}{15} = \frac{1}{5}$, $\frac{1}{5} + \frac{2}{15} = \frac{1}{3}$. We conclude that if there are any multiples of 5 in the sequence, then there are exactly three occurrences of 5 and one of 15, as $\frac{3}{5} + \frac{1}{15} = \frac{2}{3}$ (the only way to avoid a multiple of 5 in the denominator). This means that 3 and multiples of 5 are mutually exclusive possibilities, and no matter which one occurs (if any), the contribution to $\sum \frac{1}{n_i}$ is at most $\frac{2}{3}$. The other possible integers in the sequence, 2 and 6, can not come in pairs, so $\sum \frac{1}{n_i} \leq \frac{2}{3} + \frac{1}{2} + \frac{1}{6} < 2$, a contradiction. This concludes the proof of the weaker version of Proposition 3. $\square$

4 Classification of the 4-dimensional regions

Recall that in the regions

$$A = \left\{ (\alpha_1, \alpha_2) \mid \alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_1 - \alpha_2 \leq \theta, \alpha_2 \leq \frac{1}{4} \right\}$$

$$B = \left\{ (\alpha_1, \alpha_2) \mid \alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_1 \leq \frac{1}{2} - \theta, \alpha_2 \leq \frac{1}{8} \right\}$$

$$C = \left\{ (\alpha_1, \alpha_2) \mid \theta \leq \alpha_2 \leq \alpha_1, \alpha_1 + \alpha_2 \leq \frac{1}{2} - \theta \right\}$$

we may apply Buchstab's identity twice, for $p_i = x^{\alpha_i}$, to obtain

$$S\left(\mathcal{A}_{p_1 p_2}, p_2\right) = S\left(\mathcal{A}_{p_1 p_2}, x^\theta\right) - \sum_{x^\theta \leq p_3 < p_2} S\left(\mathcal{A}_{p_1 p_2 p_3}, p_3\right)$$

$$= S\left(\mathcal{A}_{p_1 p_2}, x^\theta\right) - \sum_{x^\theta \leq p_3 < p_2} S\left(\mathcal{A}_{p_1 p_2 p_3}, x^\theta\right) + \sum_{x^\theta \leq p_4 < p_3 < p_2} S\left(\mathcal{A}_{p_1 p_2 p_3 p_4}, p_4\right)$$

where the first term on the RHS can be estimated by Theorem 2, and the second can be estimated by Theorem 2 (but not entirely by Theorem 1). Here we shall give the subregions of

$$\{(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \mid (\alpha_1, \alpha_2) \in A \cup B \cup C,\ \theta \leq \alpha_4 \leq \alpha_3 \leq \alpha_2\}$$

for which the corresponding part of the last term has to be discarded. Admittedly, the rôle reversal trick is not considered here because of the difficulties involved. In particular, one would not expect an improvement on using the same simple procedure for trying this out as we do in region F here, because of the lost accuracy in the boundaries of the regions. Indeed, the higher the number of dimensions is, the greater the loss of accuracy is because of the limitations of the acceptable computer running time.

For the corresponding 6-dimensional regions, the contribution is sufficiently small that it is sensible to give an upper bound in essentially the same way as for region F.

From now on, each region X in the (α_1, α_2) −plane is replaced by

$$\{(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \mid (\alpha_1, \alpha_2) \in X,\, \theta \le \alpha_4 \le \alpha_3 \le \alpha_2\}$$

Although we are not interested in the parts where either $\alpha_1 + \alpha_2 + 2\alpha_3 > 1 - \theta$ or $\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 > 1 - \theta$, we shall check in each case whether this calls for extra conditions, so that the calculations in the end can be as simple as possible.

Consider the expansion

$$\sum S(\mathcal{A}_{p_1p_2p_3p_4}, p_4) = \sum S(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta) - \sum_{x^\theta \le p_5 < p_4} S(\mathcal{A}_{p_1p_2p_3p_4p_5}, x^\theta) + \sum_{x^\theta \le p_6 < p_5 < p_4} S(\mathcal{A}_{p_1 \dots p_6}, p_6)$$

Our task is to determine, as far as it is feasible, best possible bounds for the subregions in which the first two terms on the RHS can not be estimated because Theorem 1 and 2 do not apply in the *interiors*. (I will only list *new* conditions in each region.) Thus, the corresponding terms in the complementary regions *can* be estimated. In the difficult cases we encounter, where we do not take the trouble to find the best possible conditions, it is of course essential that we can still estimate everything in the complements asymptotically.

Region A: We have seen that $\alpha_1 + \alpha_3 \le \frac{1}{2}$, and we must have $\alpha_1 + \alpha_3 \le \frac{1}{2} - \theta$ to avoid Theorem 1. This implies that $\alpha_1 + \alpha_2 + 2\alpha_3 \le 2\left(\frac{1}{2} - \theta\right) < 1 - \theta$.

If $\alpha_2 + \alpha_3 + \alpha_4 \ge \frac{1}{2}$, then we can not even estimate $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta\right)$: If Theorem 1 were applicable, then there would be a sum of a subset of $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ in $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$. But if that subset had only one element, the sum would be at most $\alpha_1 \leq \frac{1}{4} + \theta$. If it had two elements, the sum would either be $\alpha_1 + \alpha_2 \geq \frac{1}{2}$, or at most $\alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta$. If it had at least three elements, the sum would be at least $\alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{2}$. If Theorem 2 were applicable, we could for example determine the subset with the largest sum $\leq \frac{1}{2}$, and the sum of the complement would be $\leq \frac{1}{4}$. But here, the subset with the largest sum $\leq \frac{1}{2}$ is clearly $\{\alpha_1, \alpha_3\}$, and the sum of the complement is

$$\alpha_2 + \alpha_4 = (\alpha_2 + \alpha_3 + \alpha_4) - \alpha_3 \geq \frac{1}{2} - \left(\frac{1}{2} - \theta - \alpha_1\right) = \alpha_1 + \theta \geq \frac{1}{4} + \theta > \frac{1}{4}$$

We must also impose the condition

$$\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$$

in this case.

Suppose that $\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2}$. To avoid Theorem 1, we must impose

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$$

This implies

$$\alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta + (\alpha_1 - \alpha_2) \leq \frac{1}{2}$$

so we must further impose $\alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$. This implies that

$$\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 \leq 2\,(\alpha_1 + \alpha_3 + \alpha_4) \leq 2\left(\frac{1}{2} - \theta\right) < 1 - \theta$$

so α_5 may be as large as α_4. To avoid Theorem 2 for this value of α_5, we must have $\alpha_1 + \alpha_3 + 2\alpha_4 = \alpha_1 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$. But this implies that

$$\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2} - (\alpha_1 - \alpha_2) \geq \frac{1}{2} - \theta$$

so to avoid Theorem 1, we must impose $\alpha_2 + \alpha_3 + 2\alpha_4 = \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$. Under this restriction, there is clearly no way of having the sum of one subset less than $\frac{1}{2}$ and the sum of the complement less than $\frac{1}{4}$, as

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \alpha_1 + (\alpha_2 + \alpha_3 + 2\alpha_4) \geq \frac{1}{4} + \frac{1}{2}$$

The nonexistence of subsets with sums in $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$ is equally easy to establish on considering sums of 2, 3 and 4 terms separately. For smaller values of α_5, we would still have to require $\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$.

We thus have the two regions

$$A_1 : \alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{2}, \alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$$

$$A_2 : \alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$$

Actually, it is equally easy to determine the corresponding regions in 6 dimensions, 8 dimensions, etc. for $(\alpha_1, \alpha_2) \in A$. Here, I will only mention this as a "curiosity" without the derivation, because we shall rely on a computer search for "bad" regions in 6 or more dimensions, and the numerical contribution from those coming from expansions of A is very small. Thus for $2k$ dimensions, $k \geq 3$, one gets

$$\left\{ \begin{array}{c} \alpha_1 + \alpha_3 + \alpha_4 + \ldots + \alpha_{2k-1} \leq \frac{1}{2} - \theta \\ \alpha_2 + \alpha_3 + \ldots + \alpha_{2i-1} + 2\alpha_{2i} \leq \frac{1}{2} \text{ for } i = 2, 3, \ldots, k-1 \\ \alpha_2 + \alpha_3 + \ldots + \alpha_{2k} \geq \frac{1}{2} \end{array} \right\}$$

.

and

$$\left\{\begin{array}{c} \alpha_1+\alpha_3+\alpha_4+\ldots+\alpha_{2k}\le\frac{1}{2}-\theta \\ \alpha_2+\alpha_3+\ldots+\alpha_{2i-1}+2\alpha_{2i}\le\frac{1}{2}\text{ for } i=2,3,\ldots,k-1 \\ \alpha_2+\alpha_3+\ldots+\alpha_{2k-1}+2\alpha_{2k}\ge\frac{1}{2} \end{array}\right\}$$

Region B: Here we have $\alpha_1+\alpha_2+3\alpha_3\le\alpha_1+4\alpha_2\le1-\theta$, so that $\theta\le\alpha_4\le\alpha_3\le\alpha_2$ implies $\max\left(\alpha_1+\alpha_2+2\alpha_3,\alpha_1+\alpha_2+\alpha_3+2\alpha_4\right)\le1-\theta$. There are three convex subregions in which neither of $\alpha_1+\alpha_3$ and $\alpha_1+\alpha_4$ lies in $\left[\frac{1}{2}-\theta,\frac{1}{2}\right]$:

If $\alpha_1+\alpha_3\le\frac{1}{2}-\theta$, then

$$\alpha_1+\alpha_2+\alpha_3+3\alpha_4\le(\alpha_1+\alpha_4)+4\alpha_2\le\left(\frac{1}{2}-\theta\right)+\frac{1}{2}=1-\theta$$

so that α_5 may be as large as α_4. To avoid Theorem 2, we must then have $\alpha_2+2\alpha_4\ge\frac{1}{4}$. This implies

$$\alpha_1+2\alpha_4=(\alpha_1+\alpha_2)-2\alpha_2+(\alpha_2+2\alpha_4)\ge\frac{1}{2}-\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

whereas

$$\alpha_2+\alpha_3+2\alpha_4\le4\alpha_2+3\left((\alpha_1+\alpha_3)-(\alpha_1+\alpha_2)\right)\le4\alpha_2-3\theta\le\frac{1}{2}-3\theta<\frac{1}{2}-\theta$$

It is easy to check that neither Theorem applies under these conditions, which would also be required for smaller values of α_5. That settles this case.

If $\alpha_1+\alpha_3\ge\frac{1}{2}$, $\alpha_1+\alpha_4\le\frac{1}{2}-\theta$ then we still have $\alpha_1+\alpha_2+\alpha_3+3\alpha_4\le1-\theta$. Since α_5 once again may be as large as α_4, the analysis is very similar to the previous case. Two necessary requirements are $\alpha_1+2\alpha_4\ge\frac{1}{2}$, $\alpha_2+\alpha_3+\alpha_4\ge\frac{1}{4}$. This is also sufficient, because

$$\alpha_2+\alpha_3+2\alpha_4\le4\alpha_2+2\left((\alpha_1+\alpha_4)-(\alpha_1+\alpha_2)\right)\le4\alpha_2-2\theta\le\frac{1}{2}-2\theta<\frac{1}{2}-\theta$$

Finally, consider the case $\alpha_1+\alpha_4 \geq \frac{1}{2}$. Clearly we must have $\alpha_2+\alpha_3+2\alpha_4 \geq \frac{1}{4}$ (otherwise Theorem 2 will apply for any α_5), and we shall see that it is sufficient. Suppose firstly that $\alpha_2+\alpha_3+\alpha_4 \geq \frac{1}{4}$. Since

$$\alpha_2+\alpha_3+\alpha_4 \leq 3\alpha_2 \leq \frac{3}{8} < \frac{1}{2}-\theta$$

not even $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta\right)$ can be estimated. Suppose secondly that $\alpha_2+\alpha_3+\alpha_4 \leq \frac{1}{4}$. Then

$$\alpha_1+\alpha_2+\alpha_3+3\alpha_4 \leq \alpha_1+\frac{5}{3}(\alpha_2+\alpha_3+\alpha_4) \leq \frac{1}{2}-\theta+\frac{5}{12} < 1-\theta$$

which means that we can take $\alpha_5=\alpha_4$, in which case neither Theorem applies.

We thus have these regions:

$$B_1 : \alpha_1+\alpha_3 \leq \frac{1}{2}-\theta, \alpha_2+2\alpha_4 \geq \frac{1}{4}$$

$$B_2 : \alpha_1+\alpha_4 \leq \frac{1}{2}-\theta, \alpha_1+\alpha_3 \geq \frac{1}{2}, \alpha_1+2\alpha_4 \geq \frac{1}{2}, \alpha_2+\alpha_3+\alpha_4 \geq \frac{1}{4}$$

$$B_3 : \alpha_1+\alpha_4 \geq \frac{1}{2}, \alpha_2+\alpha_3+2\alpha_4 \geq \frac{1}{4}$$

Region C: This is the difficult case. First, we remove some obvious subregions in which Theorem 1 applies to the sum of three or four coordinates. It never applies to the sum of two of them. The remaining region consists of some convex disjoint subregions which we treat separately. We have:

$$C0 : \alpha_1+\alpha_2+\alpha_3+\alpha_4 \leq \frac{1}{2}-\theta$$

$$C1 : \alpha_1+\alpha_2+\alpha_3 \leq \frac{1}{2}-\theta, \alpha_1+\alpha_2+\alpha_3+\alpha_4 \geq \frac{1}{2}$$

$$C2 : \alpha_1+\alpha_2+\alpha_4 \leq \frac{1}{2}-\theta, \alpha_1+\alpha_2+\alpha_3 \geq \frac{1}{2}$$

$$C3 : \alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_1 + \alpha_2 + \alpha_4 \geq \frac{1}{2}$$

$$C4 : \alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2}$$

$$C5 : \alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{2}$$

Of course, in the whole of C we have

$$\alpha_2 \leq \frac{\alpha_1 + \alpha_2}{2} \leq \frac{1}{4} - \frac{\theta}{2}$$

and

$$\alpha_1 + \alpha_2 + 2\alpha_3 \leq 2(\alpha_1 + \alpha_2) < 1 - \theta$$

and in $C \setminus C5$ we also have

$$\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq (\alpha_1 + \alpha_2) + (\alpha_2 + \alpha_3 + \alpha_4) < 1 - \theta;$$

in $C5$ this is an extra condition.

Trivially, $C0$ doesn't give any terms to which Theorem 2 cannot be applied twice, but we include it for completeness. In $C1$ we can always take $\alpha_5 = \alpha_4$ since

$$\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 \leq 2(\alpha_1 + \alpha_3 + \alpha_4) < 1 - \theta;$$

this also holds in $C2$ and $C3$. To avoid Theorem 2 (with $\alpha_1 + \alpha_2 + \alpha_3 < \frac{1}{2}$, $\alpha_4 + \alpha_5 \leq \frac{1}{4}$) we must at least have $\alpha_4 \geq \frac{1}{8}$, and it is easy to see that for $\alpha_5 = \alpha_4$, this also avoids Theorem 1: The sum of any three terms is at most $\frac{1}{2} - \theta$, and the sum of any four terms is at least $\frac{1}{2}$.

In $C2$ we must at least have $\alpha_3 + \alpha_4 \geq \frac{1}{4}$ to avoid $\alpha_1 + \alpha_2 + \alpha_4 \leq \frac{1}{2} - \theta$, $\alpha_3 + \alpha_5 \leq \frac{1}{4}$ for $\alpha_5 = \alpha_4$. If the sum of a subset is at most $\frac{1}{4}$ then the subset is either $\{\alpha_4, \alpha_5\}$ or a singleton set; in either case the sum of the complement is at least $\frac{1}{2}$. It is easy to see that Theorem 1 doesn't apply either, if $\alpha_5 = \alpha_4$. The reason is the same as in $C1$, except that the sum of any three terms is either $\alpha_1 + \alpha_2 + \alpha_3 \geq \frac{1}{2}$ or at most $\alpha_1 + \alpha_2 + \alpha_4 \leq \frac{1}{2} - \theta$.

Let us also finish $C5$ before turning to the more difficult regions $C3$ and $C4$. In $C5$ we have

$$\alpha_3 + \alpha_4 \geq \frac{1}{2} - \alpha_2 \geq \frac{1}{4} + \frac{\theta}{2} > \frac{1}{4}$$

so we can not even estimate $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^{\theta}\right)$. All we have to do here is to impose $\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$.

So the regions we have found so far are:

$$C1_1 : \alpha_4 \geq \frac{1}{8}$$

$$C2_1 : \alpha_3 + \alpha_4 \geq \frac{1}{4}$$

$$C5_1 : \alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$$

Region $C3$: We have already seen that α_5 can be as large as α_4, and to avoid Theorem 2 with $\alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$, $\alpha_2 + \alpha_5 \leq \frac{1}{4}$ for some α_5, we must at least impose $\alpha_2 + \alpha_4 \geq \frac{1}{4}$.

If the stronger condition $\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$ is satisfied, then we also have

$$\alpha_3 + 2\alpha_4 = (\alpha_2 + \alpha_3 + 2\alpha_4) - \alpha_2 \geq \frac{1}{2} - \left(\frac{1}{4} - \frac{\theta}{2}\right) > \frac{1}{4}$$

Suppose $\alpha_5 = \alpha_4$. The largest sum of a subset $\leq \frac{1}{4}$ is either α_1, α_2, $\alpha_3 + \alpha_4$ or $2\alpha_4$. But in each case, the sum of the complement is $\geq \frac{1}{2}$, so Theorem 2 doesn't apply. A subset whose sum lies in $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$ would have to have exactly 3 elements (by the extra requirement), which is impossible by the definition of $C3$, so Theorem 1 doesn't apply either.

Suppose $\alpha_2 + \alpha_3 + 2\alpha_4 \le \frac{1}{2}$. Then we must have $\alpha_1 \ge \frac{1}{4}$ and $\alpha_3 + 2\alpha_4 \ge \frac{1}{4}$ to avoid Theorem 2. If $\alpha_2 + \alpha_3 + 2\alpha_4 \le \frac{1}{2} - \theta$, then we may choose $\alpha_5 = \alpha_4$, and the argument is similar to that in the previous case. Otherwise, we may choose $\alpha_5 = \frac{1}{2} - \theta - \alpha_2 - \alpha_3 - \alpha_4$ to avoid Theorem 1 for sums not involving α_1. Note that we then still have

$$\alpha_5 = \left(\frac{1}{2} - \theta - \alpha_3 - \alpha_4\right) - \alpha_2 \ge \alpha_1 - \alpha_2 = 2\alpha_1 - (\alpha_1 + \alpha_2) \ge \frac{1}{2} - \left(\frac{1}{2} - \theta\right) = \theta$$

The only way that Theorem 2 can apply is if $\alpha_2 + \alpha_5 \le \frac{1}{4}$ or $\alpha_3 + \alpha_4 + \alpha_5 \le \frac{1}{4}$. A third candidate for a maximal sum below $\frac{1}{4}$ would be $\alpha_3 + \alpha_4$, but Theorem 2 would then require $\alpha_1 + \alpha_2 + \alpha_5 \le \frac{1}{2}$, which would imply

$$\alpha_2 + \alpha_5 \le \frac{1}{2} - \alpha_1 \le \frac{1}{4}$$

However,

$$\alpha_2 + \alpha_5 = \frac{1}{2} - \theta - (\alpha_3 + \alpha_4) \ge \frac{1}{2} - \theta - \left(\frac{1}{2} - \theta - \alpha_1\right) = \alpha_1 \ge \frac{1}{4}$$

and

$$\alpha_3 + \alpha_4 + \alpha_5 = \frac{1}{2} - \theta - \alpha_2 = \frac{1}{2} - \theta + \alpha_1 - (\alpha_1 + \alpha_2) \ge \alpha_1 \ge \frac{1}{4}$$

Theorem 1 doesn't apply either, since

$$\alpha_1 + \alpha_3 + \alpha_4 + \alpha_5 = \alpha_1 + (\alpha_3 + \alpha_4 + \alpha_5) \ge \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

This gives the two regions

$$C3_1 : \alpha_2 + \alpha_3 + 2\alpha_4 \ge \frac{1}{2}$$

$$C3_2 : \alpha_1 \ge \frac{1}{4}, \alpha_2 + \alpha_4 \ge \frac{1}{4}, \alpha_3 + 2\alpha_4 \ge \frac{1}{4}, \alpha_2 + \alpha_3 + 2\alpha_4 \le \frac{1}{2}$$

In $C4$ we shall consider three different cases separately. The critical question is which ones of α_1 and $\alpha_3 + \alpha_4$ are at least $\frac{1}{4}$. Of course, at least one is, as their sum is $\geq \frac{1}{2}$. That gives us the *easy* case ($\alpha_1 \geq \frac{1}{4}$, $\alpha_3 + \alpha_4 \geq \frac{1}{4}$), the *mediocre* case ($\alpha_1 \leq \frac{1}{4}$) and the *hard* case ($\alpha_3 + \alpha_4 \leq \frac{1}{4}$).

The "easy" case is very similar to $C5$: We can not even estimate $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta\right)$, so the whole region has to be discarded.

In the "mediocre" case ($\alpha_1 \leq \frac{1}{4}$) we are going to do things a little differently. There are several inequalities which have to be satisfied in order that we *don't* have to discard, so in the calculations we will first integrate over the whole region and then remove the contribution from the non-discarded subregion. Note that the only way that Theorem 2 can apply is for $\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$, $\alpha_1 \leq \frac{1}{4}$. Our task is therefore to decide, for any point $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ here, whether there exists

$$\alpha_5 \in \left[\theta, \min\left(\alpha_4, \frac{1 - \theta - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4}{2}\right)\right]$$

such that

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}, \; \alpha_1 + \alpha_5 \geq \frac{1}{4}$$

(this implies $\alpha_5 \geq \theta$, since $\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$) and no sum of a subset lies in $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$. A sum in $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$ must have the summands α_1 (since the two largest sums without it are $\alpha_2+\alpha_3+\alpha_4 \leq \frac{1}{2}-\theta$ and $\alpha_2+\alpha_3+\alpha_4+\alpha_5 \geq \frac{1}{2}$), α_5 (since Theorem 1 doesn't apply without it) and one other element (since $\alpha_1+\alpha_5 \leq \alpha_1+\alpha_2 \leq \frac{1}{2}-\theta$ and $\alpha_1+\alpha_3+\alpha_4+\alpha_5 > \alpha_1+\alpha_3+\alpha_4 \geq \frac{1}{2}$), so the only possibilities are

$$\alpha_1 + \alpha_2 + \alpha_5, \; \alpha_1 + \alpha_3 + \alpha_5, \; \alpha_1 + \alpha_4 + \alpha_5$$

Suppose $\alpha_2 - \alpha_3 \geq \theta$. Then we can take

$$\alpha_5 = \frac{1}{2} - \theta - \alpha_1 - \alpha_3$$

since this gives

$$\alpha_5 < \frac{1}{2} - (\alpha_1 + \alpha_3) \leq \alpha_4$$

and

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 = 1 - 2\theta - \alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 \leq 1 - 2\theta < 1 - \theta$$

As for the requirements made to avoid Theorem 2, we have

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \frac{1}{2} - \theta - \alpha_1 + \alpha_2 + \alpha_4 \geq \frac{1}{2} - \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2}$$

and

$$\alpha_1 + \alpha_5 = \frac{1}{2} - \theta - \alpha_3 \geq \frac{1}{2} - \alpha_2 \geq \frac{1}{4} + \frac{\theta}{2} > \frac{1}{4}$$

and the Theorem 1-related requirements are trivially satisfied: $\alpha_1 + \alpha_3 + \alpha_5 = \frac{1}{2} - \theta$ and

$$\alpha_1 + \alpha_2 + \alpha_5 = \frac{1}{2} - \theta + (\alpha_2 - \alpha_3) \geq \frac{1}{2}$$

Suppose, independently of "$\alpha_2 - \alpha_3 \geq \theta$", that $\alpha_3 - \alpha_4 \geq \theta$. Then we can choose

$$\alpha_5 = \frac{1}{2} - \alpha_1 - \alpha_3$$

since this implies $\alpha_5 \leq \alpha_4$ (trivially) and

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 = 1 - \alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 \leq 1 - \theta - \alpha_1 + \alpha_2 \leq 1 - \theta$$

Then we have

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \frac{1}{2} - \alpha_1 + \alpha_2 + \alpha_4 \geq \frac{1}{2}$$

and

$$\alpha_1 + \alpha_5 = \frac{1}{2} - \alpha_3 \geq \frac{1}{4} + \frac{\theta}{2} > \frac{1}{4}$$

which avoids Theorem 2, and we have

$$\alpha_1 + \alpha_3 + \alpha_5 = \frac{1}{2}$$

and

$$\alpha_1 + \alpha_4 + \alpha_5 = \frac{1}{2} - (\alpha_3 - \alpha_4) \leq \frac{1}{2} - \theta$$

which avoids Theorem 1. So, either of $\alpha_2 - \alpha_3 \geq \theta$ and $\alpha_3 - \alpha_4 \geq \theta$ implies that we are dealing with a term which is to be discarded. Let us assume that $\alpha_2 - \alpha_3 \leq \theta$ and $\alpha_3 - \alpha_4 \leq \theta$ in the remaining analysis. The Theorem 1-requirements then reduce to:

$$\alpha_1 + \alpha_2 + \alpha_5 \leq \frac{1}{2} - \theta \text{ OR } \alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

The first one of these gives the inequality system

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

$$\alpha_1 + \alpha_5 \geq \frac{1}{4}$$

$$\alpha_1 + \alpha_2 + \alpha_5 \leq \frac{1}{2} - \theta$$

(as the last inequality implies $\alpha_5 \leq \alpha_4$ (trivially) and $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 \leq 2\left(\frac{1}{2} - \theta\right) < 1 - \theta$). On rewriting this as

$$\alpha_5 \geq \frac{1}{2} - \alpha_2 - \alpha_3 - \alpha_4$$

$$\alpha_5 \geq \frac{1}{4} - \alpha_1$$

$$\alpha_5 \leq \frac{1}{2} - \theta - \alpha_1 - \alpha_2$$

we see that the *existence* of an α_5 which satisfies this is equivalent to the inequality system

$$\frac{1}{2} - \theta - \alpha_1 - \alpha_2 \geq \frac{1}{2} - \alpha_2 - \alpha_3 - \alpha_4 \Leftrightarrow \alpha_1 + \theta \leq \alpha_3 + \alpha_4$$

$$\frac{1}{2} - \theta - \alpha_1 - \alpha_2 \geq \frac{1}{4} - \alpha_1 \Leftrightarrow \alpha_2 \leq \frac{1}{4} - \theta$$

of which the last inequality is always satisfied:

$$\alpha_2 = \alpha_1 + (\alpha_2 + \alpha_3 + \alpha_4) - (\alpha_1 + \alpha_3 + \alpha_4) \leq \frac{1}{4} + \left(\frac{1}{2} - \theta\right) - \frac{1}{2} = \frac{1}{4} - \theta$$

This leaves $\alpha_1 + \theta \leq \alpha_3 + \alpha_4$ in this case.

Having settled that, we consider the alternative case $\alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$, which gives the inequality system

$$\alpha_5 \leq \alpha_4$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 \leq 1 - \theta$$

$$\alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

(the last inequality implies $\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$ and $\alpha_1 + \alpha_5 \geq \frac{1}{4}$). In the same way as above, the existence of an α_5 is equivalent to the system

$$\alpha_1 + 2\alpha_4 \geq \frac{1}{2}$$

$$\frac{1}{2} - \alpha_1 - \alpha_4 \leq \frac{1 - \theta - \alpha_1 - \alpha_2 - \alpha_3 - \alpha_4}{2} \Leftrightarrow \alpha_1 + \alpha_4 \geq \alpha_2 + \alpha_3 + \theta$$

But if $\alpha_1 + 2\alpha_4 \geq \frac{1}{2}$ is satisfied, then we have

$$\alpha_1 + \alpha_4 \geq \frac{1}{2} - \alpha_4 = \left(\frac{1}{2} - \theta - \alpha_4\right) + \theta \geq (\alpha_2 + \alpha_3) + \theta$$

so we only need to impose $\alpha_1 + 2\alpha_4 \geq \frac{1}{2}$ in this case.

The region we have to discard is therefore the *complement* of the region

$$C4(\text{mediocre})_{\text{1-COMPLEMENT}} : \alpha_2 - \alpha_3 \leq \theta, \alpha_3 - \alpha_4 \leq \theta, \alpha_1 + \theta \geq \alpha_3 + \alpha_4, \alpha_1 + 2\alpha_4 \leq \frac{1}{2}$$

In the "hard" case ($\alpha_3 + \alpha_4 \leq \frac{1}{4}$), at least the analysis is somewhat simplified by the inequality

$$\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 \leq (\alpha_1 + \alpha_2) + 2(\alpha_3 + \alpha_4) \leq \frac{1}{2} - \theta + \frac{1}{2} = 1 - \theta$$

which means that the maximum value of α_5 is always α_4. This means that we must always have

$$\alpha_3 + 2\alpha_4 \geq \frac{1}{4}$$

in order to have a chance to avoid Theorem 2. We will assume this from now on.

If there is a subset M of $\{1, 2, 3, 4, 5\}$ such that either

$$\sum_{m \in M} \alpha_m \in \left[\frac{1}{2} - \theta, \frac{1}{2}\right]$$

or

$$\sum_{m \in M} \alpha_m \leq \frac{1}{2}, \ \sum_{m \notin M} \alpha_m \leq \frac{1}{4}$$

then we may assume that M is one of the following:

$$\{1, 2\}$$

$$\{1, 2, 5\}$$

$$\{1, 3, 5\}$$

$$\{1, 4, 5\}$$

$$\{2, 3, 4, 5\}$$

Since $\alpha_1 + \alpha_2 \leq \frac{1}{2} - \theta$, the possibilities $\{1\}$, $\{1,3\}$ and $\{1,4\}$ are superfluous.

Once again, we shall consider a few special cases separately according to difficulty. These are:

$$I : \alpha_2 + \alpha_3 \leq \frac{1}{4}$$

$$II : \alpha_2 + \alpha_4 \leq \frac{1}{4} \leq \alpha_2 + \alpha_3$$

$$III : \alpha_2 + \alpha_4 \geq \frac{1}{4}$$

Note that in I and II, we have

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \leq 2\left(\alpha_2 + \alpha_4\right) \leq \frac{1}{2}$$

In Region I, the list of possible choices for M can be reduced to

$$\{1,2\} \rightarrow \text{impose } \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{4}$$

$$\{1,4,5\} \rightarrow \text{impose } \alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

$$\{2,3,4,5\} \rightarrow \ \text{impose } \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \leq \frac{1}{2} - \theta$$

This gives the inequality system

$$\left.\begin{array}{r} \theta \leq \\ \frac{1}{4} - \alpha_3 - \alpha_4 \leq \\ \frac{1}{2} - \alpha_1 - \alpha_4 \leq \end{array}\right\} \alpha_5 \left\{\begin{array}{l} \leq \alpha_4 \\ \leq \frac{1}{2} - \theta - \alpha_2 - \alpha_3 - \alpha_4 \end{array}\right.$$

which gives six inequalities for the first 4 coordinates. Two of them are the trivial

$$\alpha_4 \geq \theta,\ \alpha_3 + 2\alpha_4 \geq \frac{1}{4}$$

and we also have

$$\alpha_2 \leq \frac{1}{2} - \theta - \alpha_1 \leq \frac{1}{4} - \theta$$

and

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{2}{3}(\alpha_2 + \alpha_3) \leq \frac{1}{6} < \frac{1}{2} - 2\theta$$

which leaves us with the two extra inequalities

$$\alpha_1 + 2\alpha_4 \geq \frac{1}{2}, \, \alpha_1 \geq \alpha_2 + \alpha_3 + \theta$$

in Region I.

Region II: The list of possible choices for M can be reduced to

$$\{1,2\} \to \text{impose } \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{4}$$

$$\{1,3,5\} \to \text{impose } \alpha_1 + \alpha_3 + \alpha_5 \geq \frac{1}{2}$$

$$\{1,4,5\} \to \text{impose } \alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2} \text{ OR } \alpha_1 + \alpha_4 + \alpha_5 \leq \frac{1}{2} - \theta$$

$$\{2,3,4,5\} \to \text{ impose } \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \leq \frac{1}{2} - \theta$$

So we have two possible constraints for $\alpha_1 + \alpha_4 + \alpha_5$.

The first possibility can be treated almost exactly like Region I which we just settled; the only difference is that it is not completely trivial that

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - 2\theta$$

However, one of the extra conditions we ended up with, namely $\alpha_1 \geq \alpha_2 + \alpha_3 + \theta$, implies

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \alpha_1 - \theta + \alpha_4 \leq \alpha_1 + \alpha_2 - \theta \leq \frac{1}{2} - 2\theta$$

So the extra inequalities we end up with are once again

$$\alpha_1 + 2\alpha_4 \geq \frac{1}{2},\ \alpha_1 \geq \alpha_2 + \alpha_3 + \theta$$

The other possibility yields

$$\left.\begin{array}{r} \theta \leq \\ \\ \frac{1}{4} - \alpha_3 - \alpha_4 \leq \\ \\ \frac{1}{2} - \alpha_1 - \alpha_3 \leq \end{array}\right\} \alpha_5 \left\{\begin{array}{l} \leq \alpha_4 \\ \\ \leq \frac{1}{2} - \theta - \alpha_2 - \alpha_3 - \alpha_4 \\ \\ \leq \frac{1}{2} - \theta - \alpha_1 - \alpha_4 \end{array}\right.$$

which gives 9 inequalities for the first four coordinates. However, all those based on $\alpha_5 \leq \alpha_4$ are trivial. The others are:

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - 2\theta$$

$$\alpha_2 \leq \frac{1}{4} - \theta \text{ (trivial, see Region } I)$$

$$\alpha_1 \geq \alpha_2 + \alpha_4 + \theta$$

$$\alpha_1 + \alpha_4 \leq \frac{1}{2} - 2\theta$$

$$\alpha_1 \leq \alpha_3 + \frac{1}{4} - \theta \text{ (follows from } \alpha_1 + \alpha_2 \leq \frac{1}{2} - \theta \leq \alpha_2 + \alpha_3 + \frac{1}{4} - \theta)$$

$$\alpha_3 - \alpha_4 \geq \theta$$

The inequality $\alpha_3 - \alpha_4 \geq \theta$ implies

$$\alpha_1 + \alpha_4 \leq \alpha_1 + \alpha_3 - \theta \leq \frac{1}{2} - 2\theta$$

and we also have that $\alpha_1 \geq \alpha_2 + \alpha_4 + \theta$ implies

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \alpha_1 - \theta + \alpha_3 \leq \alpha_1 + \alpha_2 - \theta \leq \frac{1}{2} - 2\theta$$

That leaves us with the extra conditions

$$\alpha_1 \geq \alpha_2 + \alpha_4 + \theta,\ \alpha_3 - \alpha_4 \geq \theta$$

in this case. However, the regions defined by the two sets of conditions we have sifted out in Region II need not be distinct. The complete requirement for Region II is therefore

$$\left\{ \begin{array}{c} \alpha_1 + 2\alpha_4 \geq \frac{1}{2} \\ \\ \alpha_1 \geq \alpha_2 + \alpha_3 + \theta \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_1 \geq \alpha_2 + \alpha_4 + \theta \\ \\ \alpha_3 - \alpha_4 \geq \theta \end{array} \right\}$$

Region III is the most difficult case, because not only are there three different ways to avoid Theorem 1 for sums involving α_1, we must also combine these with the two possibilities

$$\text{(a) } \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \leq \frac{1}{2} - \theta$$

and

$$\text{(b) } \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

With $\alpha_2 + \alpha_4$ being rather big, the requirements for avoiding Theorem 2 are reduced to:

$$\alpha_1 + \alpha_2 + \alpha_5 \geq \frac{1}{2},\ \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{4}$$

and the three ways to avoid Theorem 1 for sums involving α_1 are:

$$\text{(c) } \alpha_1 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

$$\text{(d) } \alpha_1 + \alpha_4 + \alpha_5 \leq \frac{1}{2} - \theta,\ \alpha_1 + \alpha_3 + \alpha_5 \geq \frac{1}{2}$$

$$\text{(e) } \alpha_1 + \alpha_3 + \alpha_5 \leq \frac{1}{2} - \theta$$

We will consider the cases involving requirement (a) first, since this resembles the situation in Region II. Indeed, the combinations (a)+(c) and (a)+(d) give exactly the same inequalities as in Region II, and the new element

$$\alpha_2 + \alpha_4 \geq \frac{1}{4}$$

doesn't make any difference in the calculations. That leaves us with the combination (a)+(e). Here we have

$$\left.\begin{array}{r} \theta \leq \\ \frac{1}{4} - \alpha_3 - \alpha_4 \leq \\ \frac{1}{2} - \alpha_1 - \alpha_2 \leq \end{array}\right\} \alpha_5 \left\{\begin{array}{l} \leq \alpha_4 \\ \leq \frac{1}{2} - \theta - \alpha_2 - \alpha_3 - \alpha_4 \\ \leq \frac{1}{2} - \theta - \alpha_1 - \alpha_3 \end{array}\right.$$

Out of the nine inequalities we get from this, the ones coming from $\alpha_5 \leq \alpha_4$ are once again trivial. The other six are

$$\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - 2\theta$$

$$\alpha_2 \leq \frac{1}{4} - \theta \text{ (trivial)}$$

$$\alpha_1 \geq \alpha_3 + \alpha_4 + \theta$$

$$\alpha_1 + \alpha_3 \leq \frac{1}{2} - 2\theta$$

$$\alpha_1 \leq \alpha_4 + \frac{1}{4} - \theta \text{ (follows from } \alpha_1 + \alpha_2 \leq \frac{1}{2} - \theta \leq \alpha_2 + \alpha_4 + \frac{1}{4} - \theta)$$

$$\alpha_2 - \alpha_3 \geq \theta$$

In the same way as in the second case of Region II, we do the reductions

$$\alpha_1 \geq \alpha_3 + \alpha_4 + \theta \Rightarrow \alpha_2 + \alpha_3 + \alpha_4 \leq \alpha_1 + \alpha_2 - \theta \leq \frac{1}{2} - 2\theta$$

and

$$\alpha_2 - \alpha_3 \geq \theta \Rightarrow \alpha_1 + \alpha_3 \leq \alpha_1 + \alpha_2 - \theta \leq \frac{1}{2} - 2\theta$$

So, for the case where (a) is assumed, we get the region

$$\left\{ \begin{array}{c} \alpha_1 + 2\alpha_4 \geq \frac{1}{2} \\ \\ \alpha_1 \geq \alpha_2 + \alpha_3 + \theta \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_1 \geq \alpha_2 + \alpha_4 + \theta \\ \\ \alpha_3 - \alpha_4 \geq \theta \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_1 \geq \alpha_3 + \alpha_4 + \theta \\ \\ \alpha_2 - \alpha_3 \geq \theta \end{array} \right\}$$

Condition (b) is actually a simple one, because it trivially implies both our Theorem 2-avoiding requirements. It also trivially implies that $\alpha_5 \geq \theta$, so all we need to do is to combine

$$\frac{1}{2} - \alpha_2 - \alpha_3 - \alpha_4 \leq \alpha_5 \leq \alpha_4$$

with each of (c), (d) and (e). Combination with (c) gives the inequalities

$$\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$$

$$\alpha_1 + 2\alpha_4 \geq \frac{1}{2}$$

which can't be reduced. Combination with (d) gives

$$\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$$

$$\alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2} \text{ (trivial)}$$

$$\alpha_1 + \theta \leq \alpha_2 + \alpha_3$$

$$\alpha_3 - \alpha_4 \geq \theta$$

which can't be further reduced (after throwing out the trivial one). Finally, combination with (e) gives

$$\alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$$

$$\alpha_1 + \theta \le \alpha_2 + \alpha_4$$

Here, however, we can do the reduction

$$\alpha_1 + \theta \le \alpha_2 + \alpha_4 \Rightarrow \alpha_2 + \alpha_3 + 2\alpha_4 \ge \alpha_1 + \alpha_3 + \alpha_4 + \theta > \frac{1}{2}$$

and the complete set of requirements for region III is thus

$$\left\{ \begin{array}{c} \alpha_1 + 2\alpha_4 \ge \frac{1}{2} \\ \alpha_1 \ge \alpha_2 + \alpha_3 + \theta \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_1 \ge \alpha_2 + \alpha_4 + \theta \\ \alpha_3 - \alpha_4 \ge \theta \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_1 \ge \alpha_3 + \alpha_4 + \theta \\ \alpha_2 - \alpha_3 \ge \theta \end{array} \right\}$$

$$\text{or } \left\{ \begin{array}{c} \alpha_2 + \alpha_3 + 2\alpha_4 \ge \frac{1}{2} \\ \alpha_1 + 2\alpha_4 \ge \frac{1}{2} \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \alpha_2 + \alpha_3 + 2\alpha_4 \ge \frac{1}{2} \\ \alpha_1 + \theta \le \alpha_2 + \alpha_3 \\ \alpha_3 - \alpha_4 \ge \theta \end{array} \right\} \text{ or } \{\alpha_1 + \theta \le \alpha_2 + \alpha_4\}$$

The complicated shape of the set of requirements in Region $C4$(hard) (with lots of AND's and OR's) makes it pretty hard to work out a representation suitable for numerical integration with high accuracy, although it would probably be possible to write a computer program that could do this. Instead, we shall find an upper bound for the contribution in the following way:

- Integrate over a convex region which covers $C4$(hard) - we will use:

  $$\left\{ \begin{array}{c} (\alpha_1, ..., \alpha_4) \mid \theta \le \alpha_4 \le ... \le \alpha_1, \alpha_1 + \alpha_2 \le \frac{1}{2} - \theta, \\ \alpha_1 + \alpha_3 + \alpha_4 \ge \frac{1}{2}, \alpha_3 + \alpha_4 \le \frac{1}{4} \le \alpha_3 + 2\alpha_4 \end{array} \right\}$$

  Note that the condition on $\alpha_2 + \alpha_3 + \alpha_4$ in $C4$ was dropped, since the other constraints yield

  $$\alpha_2 + \alpha_3 + \alpha_4 = (\alpha_1 + \alpha_2) + 2(\alpha_3 + \alpha_4) - (\alpha_1 + \alpha_3 + \alpha_4) \le \left(\frac{1}{2} - \theta\right) + \frac{2}{4} - \frac{1}{2} = \frac{1}{2} - \theta$$

- With a computer search, remove (a lower bound for) the contribution from "boxes" of the form

$$\alpha_i \in [\alpha_i', \alpha_i' + \Delta\alpha] \text{ for } i = 1, 2, 3, 4$$

which are entirely inside the covering region, and entirely outside $C4$(hard).

The wisdom of using this relatively simple procedure is confirmed by the fact that the contribution from $C4$(hard) is actually pretty small; see Chapter 5.

In $C4$(mediocre) the procedure is similar:

- Integrate over

$$\left\{(\alpha_1, ..., \alpha_4) \mid \theta \leq \alpha_4 \leq ... \leq \alpha_1, \alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2}, \alpha_1 \leq \frac{1}{4}\right\}$$

Note that

$$\alpha_1 + \alpha_2 = 2\alpha_1 - (\alpha_1 + \alpha_3 + \alpha_4) + (\alpha_2 + \alpha_3 + \alpha_4) \leq \frac{2}{4} - \frac{1}{2} + \left(\frac{1}{2} - \theta\right) = \frac{1}{2} - \theta$$

- Remove the contribution from "boxes" entirely inside $C4(\text{mediocre})_{\text{1-COMPLEMENT}}$

The contribution from $C4$(mediocre) is even smaller than that from $C4$(hard).

We also need to do an analysis of region F. If α_3 is sufficiently small, then we can find 4-dimensional regions pretty much as we just did for A, B and C. The requirement is that either $\alpha_2 + \alpha_3 \leq \frac{1}{4}$ or $\alpha_1 + \alpha_3 \leq \frac{1}{2}$ holds; we therefore consider $\alpha_1 - \alpha_2 \geq \frac{1}{4}$ and $\alpha_1 - \alpha_2 \leq \frac{1}{4}$ separately.

Suppose firstly that $\alpha_1 - \alpha_2 \geq \frac{1}{4}$, so that $\alpha_2 + \alpha_3 \leq \frac{1}{4}$ is required. Then we have

$$\alpha_1 + \alpha_2 + 3\alpha_3 \leq \alpha_1 + 2\left(\alpha_2 + \alpha_3\right) \leq \frac{1}{2} - \theta + \frac{1}{2} = 1 - \theta$$

so α_4 may be as large as α_3. Note that $\alpha_2+\alpha_3+\alpha_4 \le \frac{3}{2}(\alpha_2+\alpha_3) \le \frac{3}{8} < \frac{1}{2}-\theta$, so that the situation is exactly like in B, and we get three regions again:

$$F_1 : \alpha_1-\alpha_2 \ge \frac{1}{4}, \alpha_2+\alpha_3 \le \frac{1}{4}, \alpha_1+\alpha_3 \le \frac{1}{2}-\theta, \alpha_2+2\alpha_4 \ge \frac{1}{4}$$

$$F_2 : \alpha_2+\alpha_3 \le \frac{1}{4}, \alpha_1+\alpha_4 \le \frac{1}{2}-\theta, \alpha_1+\alpha_3 \ge \frac{1}{2}, \alpha_1+2\alpha_4 \ge \frac{1}{2}, \alpha_2+\alpha_3+\alpha_4 \ge \frac{1}{4}$$

$$F_3 : \alpha_2+\alpha_3 \le \frac{1}{4}, \alpha_1+\alpha_4 \ge \frac{1}{2}, \alpha_2+\alpha_3+2\alpha_4 \ge \frac{1}{4}$$

The condition $\alpha_1-\alpha_2 \ge \frac{1}{4}$ was dropped in F_2 and F_3, since we have

$$\alpha_1-\alpha_2 = (\alpha_1+\alpha_3)-(\alpha_2+\alpha_3) \ge \frac{1}{2}-\frac{1}{4} = \frac{1}{4}$$

in those regions.

Suppose secondly that $\alpha_1-\alpha_2 \le \frac{1}{4}$, so that $\alpha_1+\alpha_3 \le \frac{1}{2}$ is required. To avoid Theorem 1 we must have $\alpha_1+\alpha_3 \le \frac{1}{2}-\theta$, which implies that $\alpha_3 < \alpha_2$ and that

$$\alpha_1+\alpha_2+2\alpha_3 < 2(\alpha_1+\alpha_3) < 1-\theta$$

Also,

$$\alpha_3 \le \frac{1}{2}-\theta-\alpha_1 = \frac{1}{2}-\theta-\frac{1}{2}((\alpha_1+\alpha_2)+(\alpha_1-\alpha_2)) \le \frac{1}{2}-\theta-\frac{1}{2}\left(\frac{1}{2}+\theta\right) = \frac{1}{4}-\frac{3\theta}{2}$$

It should be noted that in the cases where $\alpha_2+\alpha_3$ is required to be $\ge \frac{1}{4}$, the condition $\alpha_1-\alpha_2 \le \frac{1}{4}$ (or even $\le \frac{1}{4}-\theta$, which occurs in one case) is superfluous, since we then have

$$\alpha_1-\alpha_2 \le (\alpha_1+\alpha_3)-(\alpha_2+\alpha_3) \le \left(\frac{1}{2}-\theta\right)-\frac{1}{4} = \frac{1}{4}-\theta < \frac{1}{4}$$

If $\alpha_2+\alpha_3+\alpha_4 \ge \frac{1}{2}$, then we have indeed

$$\alpha_2+\alpha_4 \ge \frac{1}{4}+\frac{3\theta}{2} > \frac{1}{4}$$

so we can not even estimate $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta\right)$. We must, however, impose

$$\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$$

here, and we get

$$F_4 : \alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{2}, \alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$$

For the remaining regions we have $\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$ (the interval $\left[\frac{1}{2} - \theta, \frac{1}{2}\right]$ was avoided as usual) so that

$$\alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq (\alpha_1 + \alpha_3) + (\alpha_2 + \alpha_3 + \alpha_4) \leq 2\left(\frac{1}{2} - \theta\right) < 1 - \theta$$

If $\alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2}$, $\alpha_2 + \alpha_4 \leq \frac{1}{4}$ then Theorem 2 applies. We therefore have three cases left to consider:

$$(i)\ \ \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2},\ \alpha_2 + \alpha_4 \geq \frac{1}{4}$$

$$(ii)\ \ \alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2},\ \alpha_2 + \alpha_4 \geq \frac{1}{4}$$

$$(iii)\ \ \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2},\ \alpha_2 + \alpha_4 \leq \frac{1}{4}$$

where $\alpha_1 - \alpha_2 \leq \frac{1}{4}$, $\alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta$, $\alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$ is assumed is each case.

The case (i) is easy: We still can not even estimate $\sum S\left(\mathcal{A}_{p_1p_2p_3p_4}, x^\theta\right)$. This gives us

$$F_5 : \alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_4 \geq \frac{1}{4}, \alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2}, \alpha_2 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$$

In case (ii), we require $\alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta$ to avoid Theorem 1. Then we have

$$\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 \leq 2\left(\alpha_1 + \alpha_3 + \alpha_4\right) < 1 - \theta$$

so that α_5 may be as large as α_4. This gives at least the extra condition $\alpha_1 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$ to avoid Theorem 2. We need an α_5 for which

$$\alpha_1 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2}$$

$$\alpha_2 + \alpha_5 \geq \frac{1}{4}$$

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \notin \left[\frac{1}{2} - \theta, \frac{1}{2}\right]$$

If $\alpha_2 + \alpha_3 + 2\alpha_4 \leq \frac{1}{2} - \theta$, then $\alpha_5 = \alpha_4$ will do. Otherwise, choose

$$\alpha_5 = \frac{1}{2} - \theta - \alpha_2 - \alpha_3 - \alpha_4 < \alpha_4$$

We then have

$$\alpha_5 \geq \frac{1}{2} - \alpha_1 - \alpha_3 - \alpha_4 \geq \theta \text{ (because } \alpha_1 - \alpha_2 \geq \theta \text{ in } F)$$

$$\alpha_1 + \alpha_3 + \alpha_4 + \alpha_5 \geq \frac{1}{2} \text{ (same reason)}$$

$$\alpha_2 + \alpha_5 = \frac{1}{2} - \theta - \alpha_3 - \alpha_4 \geq \alpha_1 \geq \frac{1}{4} + \frac{\theta}{2} > \frac{1}{4}$$

which concludes this case, and we have

$$F_6 : \alpha_1 + \alpha_3 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_4 \geq \frac{1}{4}, \alpha_1 + \alpha_3 + 2\alpha_4 \geq \frac{1}{2}$$

The remaining case is (iii), where $\alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2}$ and $\alpha_2 + \alpha_4 \leq \frac{1}{4}$. Here, we have

$$\alpha_1 + \alpha_2 + \alpha_3 + 3\alpha_4 \leq (\alpha_1 + \alpha_3) + 2(\alpha_2 + \alpha_4) \leq \left(\frac{1}{2} - \theta\right) + \frac{1}{2} = 1 - \theta$$

so that α_5 may be as large as α_4. If $\alpha_1 + 2\alpha_4 \geq \frac{1}{2}$, then $\alpha_2 + 2\alpha_4 \geq \frac{1}{4}$ and neither Theorem applies for $\alpha_5 = \alpha_4$. In particular,

$$\alpha_2 + \alpha_3 + 2\alpha_4 = 2(\alpha_2 + \alpha_4) + ((\alpha_1 + \alpha_3) - (\alpha_1 + \alpha_2)) \leq \frac{1}{2} - \theta$$

This gives us

$$F_7 : \alpha_1 - \alpha_2 \le \frac{1}{4}, \alpha_1 + \alpha_3 \le \frac{1}{2} - \theta, \alpha_2 + \alpha_4 \le \frac{1}{4}, \alpha_1 + 2\alpha_4 \ge \frac{1}{2}$$

If $\alpha_1 + 2\alpha_4 \le \frac{1}{2}$, then we must have $\alpha_2 + \alpha_3 \ge \frac{1}{4}$ to avoid Theorem 2. This is the most difficult case, so let us start with a few observations. We have

$$\alpha_2 - \alpha_3 = ((\alpha_1 + \alpha_2) - (\alpha_1 + \alpha_3)) \ge \frac{1}{2} - \left(\frac{1}{2} - \theta\right) = \theta$$

which implies

$$\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 \le 2\,(\alpha_2 + \alpha_4) - (\alpha_2 - \alpha_3) \le \frac{1}{2} - \theta$$

and

$$\alpha_1 = (\alpha_1 + \alpha_3 + \alpha_4) + (\alpha_2 - \alpha_3) - (\alpha_2 + \alpha_4) \ge \frac{1}{2} + \theta - \frac{1}{4} = \frac{1}{4} + \theta$$

of which the latter implies

$$\alpha_3 = (\alpha_1 + \alpha_3) - \alpha_1 \le \left(\frac{1}{2} - \theta\right) - \left(\frac{1}{4} + \theta\right) = \frac{1}{4} - 2\theta$$

which in turn implies

$$\alpha_2 + \alpha_3 + \alpha_4 = (\alpha_2 + \alpha_4) + \alpha_3 \le \frac{1}{4} + \left(\frac{1}{4} - 2\theta\right) = \frac{1}{2} - 2\theta$$

For sums involving α_5, we require

$$\alpha_1 + \alpha_3 + \alpha_5 \ge \frac{1}{2}$$

to avoid Theorem 2, since $\alpha_2 + \alpha_4 \le \frac{1}{4}$;

$$\alpha_1 + \alpha_4 + \alpha_5 \le \frac{1}{2} - \theta$$

to avoid Theorem 1, since $\alpha_1 + 2\alpha_4 \leq \frac{1}{2}$; and

$$\alpha_2 + \alpha_4 + \alpha_5 \geq \frac{1}{4}$$

to avoid Theorem 2, since $\alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta$. It is easy to check that this suffices.

This gives us the inequality system

$$\left.\begin{array}{r} \theta \leq \\ \frac{1}{2} - \alpha_1 - \alpha_3 \leq \\ \frac{1}{4} - \alpha_2 - \alpha_4 \leq \end{array}\right\} \alpha_5 \left\{\begin{array}{l} \leq \alpha_4 \\ \leq \frac{1}{2} - \theta - \alpha_1 - \alpha_4 \end{array}\right.$$

which leads to

$$\alpha_4 \geq \theta \text{ - OK}$$

$$\alpha_1 + \alpha_3 + \alpha_4 \geq \frac{1}{2} \text{ - true by def.}$$

$$\alpha_2 + 2\alpha_4 \geq \frac{1}{4} \text{ - new condition}$$

$$\alpha_1 + \alpha_4 \leq \frac{1}{2} - 2\theta \text{ - new condition}$$

$$\alpha_3 - \alpha_4 \geq \theta \text{ - new condition}$$

$$\alpha_1 - \alpha_2 \leq \frac{1}{4} - \theta \text{ - derived earlier}$$

If we remove superfluous conditions, we thus get the following regions:

$$F_1 : \alpha_1 - \alpha_2 \geq \frac{1}{4}, \alpha_2 + \alpha_3 \leq \frac{1}{4}, \alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta, \alpha_2 + 2\alpha_4 \geq \frac{1}{4}$$

$$F_2 : \alpha_2 + \alpha_3 \leq \frac{1}{4}, \alpha_1 + \alpha_4 \leq \frac{1}{2} - \theta, \alpha_1 + \alpha_3 \geq \frac{1}{2}, \alpha_1 + 2\alpha_4 \geq \frac{1}{2}, \alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{4}$$

$$F_3 : \alpha_2 + \alpha_3 \leq \frac{1}{4}, \alpha_1 + \alpha_4 \geq \frac{1}{2}, \alpha_2 + \alpha_3 + 2\alpha_4 \geq \frac{1}{4}$$

$$F_4 : \alpha_1+\alpha_3 \le \frac{1}{2}-\theta, \alpha_2+\alpha_3+\alpha_4 \ge \frac{1}{2}, \alpha_1+\alpha_2+\alpha_3+2\alpha_4 \le 1-\theta$$

$$F_5 : \alpha_1+\alpha_3 \le \frac{1}{2}-\theta, \alpha_2+\alpha_4 \ge \frac{1}{4}, \alpha_1+\alpha_3+\alpha_4 \ge \frac{1}{2}, \alpha_2+\alpha_3+\alpha_4 \le \frac{1}{2}-\theta$$

$$F_6 : \alpha_1+\alpha_3+\alpha_4 \le \frac{1}{2}-\theta, \alpha_2+\alpha_4 \ge \frac{1}{4}, \alpha_1+\alpha_3+2\alpha_4 \ge \frac{1}{2}$$

$$F_7 : \alpha_1-\alpha_2 \le \frac{1}{4}, \alpha_1+\alpha_3 \le \frac{1}{2}-\theta, \alpha_2+\alpha_4 \le \frac{1}{4}, \alpha_1+2\alpha_4 \ge \frac{1}{2}$$

$$F_8 : \alpha_1+\alpha_3 \le \frac{1}{2}-\theta, \alpha_2+\alpha_3 \ge \frac{1}{4}, \alpha_2+\alpha_4 \le \frac{1}{4}, \alpha_1+2\alpha_4 \le \frac{1}{2},$$

$$\alpha_2+2\alpha_4 \ge \frac{1}{4}, \alpha_1+\alpha_4 \le \frac{1}{2}-2\theta, \alpha_1+\alpha_3+\alpha_4 \ge \frac{1}{2}, \alpha_3-\alpha_4 \ge \theta$$

When α_3 is bigger than both $\frac{1}{2}-\alpha_1$ and $\frac{1}{4}-\alpha_2$, and we use the rôle reversal trick, we simply let the computer check if the sum of any subset of $\{\alpha_2, \alpha_3, 1-\theta-\alpha_1-\alpha_2-\alpha_3, \alpha_4\}$ lies in $\left[\frac{1}{2}-\theta, \frac{1}{2}\right]$. Here, the range for α_4 is $\left[\theta, \frac{\alpha_1}{2}\right]$. The candidates are: $\alpha_2+\alpha_4$, $\alpha_2+\alpha_3+\alpha_4$, $1-\theta-\alpha_1-\alpha_2+\alpha_4$, $1-\theta-\alpha_1-\alpha_3+\alpha_4$, $1-\theta-\alpha_1-\alpha_2-\alpha_3+\alpha_4$. It is easy to check that the other ones are outside.

Our treatment of Region F is similar to that of the complicated parts of $C4$. The contribution from the 2-dimensional form of F is easy to evaluate with high accuracy. Then we go through the 4-dimensional "boxes" for each "square"

$$[\alpha_1', \alpha_1'+\Delta\alpha] \times [\alpha_2', \alpha_2'+\Delta\alpha]$$

and compare a *lower* bound for the corresponding 2-dimensional contribution with an *upper* bound for the corresponding 4-dimensional contribution. For the purpose of evaluating the latter, we check if there is (or might be) a point in the interior lying in one of the F_i's $(i = 1, ..., 8)$, and/or a point where $\alpha_3 \ge \max\left(\frac{1}{2}-\alpha_1, \frac{1}{4}-\alpha_2\right)$.

To be able to perform numerical integration with high accuracy in $A \cup B \cup C$ except for some subregions of $C4$, we must rewrite the inequality systems such that the range for each α_i is determined by the α_j's with $j < i$ (for example). There is a straightforward algorithm for doing this transformation, which means that I don't have to go through all the calculations; I'll just give the results. The algorithm goes like this: The system of inequalities that determines a region R, together with the trivial requirements $\theta \leq \alpha_4 \leq \alpha_3 \leq \alpha_2 \leq \alpha_1 \leq \frac{1}{2} - \theta$, can be written as

$$f_1(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \geq 0$$

$$f_2(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \geq 0$$

$$(...)$$

$$f_k(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \geq 0$$

Combining the f_i's with a negative coefficient for α_4 with those with a positive coefficient for α_4 while keeping the remaining ones gives a new system

$$g_1(\alpha_1, \alpha_2, \alpha_3) \geq 0$$

$$(...)$$

$$g_l(\alpha_1, \alpha_2, \alpha_3) \geq 0$$

and in the obvious way one also gets

$$h_1(\alpha_1, \alpha_2) \geq 0$$

$$(...)$$

$$h_m(\alpha_1, \alpha_2) \geq 0$$

and in the end one gets the range for α_1 - a convex subset of the real line, i.e., an interval. Then, the next step is to divide this range into smaller intervals according to which h_i's are "dominating", i.e., which pairs of inequalities $h_i(\alpha_1, \alpha_2) \geq 0$, $h_j(\alpha_1, \alpha_2) \geq 0$ with negative coefficient for α_2 in h_i and positive coefficient in h_j imply all the other ones. Then, for a given range S for (α_1, α_2), one does the same with the requirements for α_3. Unless we end up with just two g_i's (a lower and upper bound for α_3), we have to split up S accordingly following the same procedure as above, and consider each subregion separately. Of course, the final step (including α_4) is done in a similar way and the whole procedure is terminating. Fortunately, in some regions the splitting is fairly simple.

The results are presented in a form valid for $0 < \theta < \frac{1}{24}$. In the calculations (Chapter 5), I will assume that $\theta = \frac{1}{25}$. When I give the constraints that a region is defined by, these are always meant to be in addition to (not instead of) constraints already given for any region containing it.

Recall that the two-dimensional Region A is defined by: $\alpha_1 + \alpha_2 \geq \frac{1}{2}, \alpha_1 - \alpha_2 \leq \theta, \alpha_2 \leq \frac{1}{4}$. Region A_1 (defined by $\alpha_1 + \alpha_3 \leq \frac{1}{2} - \theta, \alpha_2 + \alpha_3 + \alpha_4 \geq \frac{1}{2}, \alpha_1 + \alpha_2 + \alpha_3 + 2\alpha_4 \leq 1 - \theta$) becomes

$$\left\{ \begin{array}{c} \frac{1}{4} \leq \alpha_1 \leq \frac{1}{4} + \frac{\theta}{2} \\ \frac{1}{2} - \alpha_1 \leq \alpha_2 \leq \frac{1}{4} \\ \frac{1}{4} - \frac{\alpha_2}{2} \leq \alpha_3 \leq \frac{1-\theta-\alpha_1-\alpha_2}{3} \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \alpha_3 \end{array} \right\} \cup \left\{ \begin{array}{c} \frac{1}{4} \leq \alpha_1 \leq \frac{1}{4} + \frac{\theta}{2} \\ \frac{1}{2} - \alpha_1 \leq \alpha_2 \leq \frac{1}{4} \\ \frac{1-\theta-\alpha_1-\alpha_2}{3} \leq \alpha_3 \leq \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \frac{1-\theta-\alpha_1-\alpha_2-\alpha_3}{2} \end{array} \right\}$$

.

$$
\cup \left\{ \begin{array}{c} \frac{1}{4} + \frac{\theta}{2} \le \alpha_1 \le \frac{1}{4} + \theta \\ \alpha_1 - \theta \le \alpha_2 \le \frac{1}{4} \\ \frac{1}{4} - \frac{\alpha_2}{2} \le \alpha_3 \le \frac{1-\theta-\alpha_1-\alpha_2}{3} \\ \frac{1}{2} - \alpha_2 - \alpha_3 \le \alpha_4 \le \alpha_3 \end{array} \right\} \cup \left\{ \begin{array}{c} \frac{1}{4} + \frac{\theta}{2} \le \alpha_1 \le \frac{1}{4} + \theta \\ \alpha_1 - \theta \le \alpha_2 \le \frac{1}{4} \\ \frac{1-\theta-\alpha_1-\alpha_2}{3} \le \alpha_3 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \le \alpha_4 \le \frac{1-\theta-\alpha_1-\alpha_2-\alpha_3}{2} \end{array} \right\}
$$

and region A_2 (defined by $\alpha_1 + \alpha_3 + \alpha_4 \le \frac{1}{2} - \theta, \alpha_2 + \alpha_3 + 2\alpha_4 \ge \frac{1}{2}$) becomes

$$
\left\{ \begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{1}{4} + \frac{\theta}{2} \\ \frac{1}{2} - \alpha_1 \le \alpha_2 \le \frac{1}{4} \\ \frac{1}{6} - \frac{\alpha_2}{3} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \alpha_3 \end{array} \right\} \cup \left\{ \begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{1}{4} + \frac{\theta}{2} \\ \frac{1}{2} - \alpha_1 \le \alpha_2 \le \frac{1}{4} \\ \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le \frac{1}{2} - 2\theta - 2\alpha_1 + \alpha_2 \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \frac{1}{2} - \theta - \alpha_1 - \alpha_3 \end{array} \right\}
$$

$$
\cup \left\{ \begin{array}{c} \frac{1}{4} + \frac{\theta}{2} \le \alpha_1 \le \frac{1}{4} + \theta \\ \alpha_1 - \theta \le \alpha_2 \le \frac{1}{4} \\ \frac{1}{6} - \frac{\alpha_2}{3} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \alpha_3 \end{array} \right\} \cup \left\{ \begin{array}{c} \frac{1}{4} + \frac{\theta}{2} \le \alpha_1 \le \frac{1}{4} + \theta \\ \alpha_1 - \theta \le \alpha_2 \le \frac{1}{4} \\ \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le \frac{1}{2} - 2\theta - 2\alpha_1 + \alpha_2 \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \frac{1}{2} - \theta - \alpha_1 - \alpha_3 \end{array} \right\}
$$

The two-dimensional Region B is defined by $\alpha_1 + \alpha_2 \ge \frac{1}{2}, \alpha_1 \le \frac{1}{2} - \theta, \alpha_2 \le \frac{1}{8}$. Region B_1 is defined by $\alpha_1 + \alpha_3 \le \frac{1}{2} - \theta, \alpha_2 + 2\alpha_4 \ge \frac{1}{4}$, which becomes

$$
\left\{ \begin{array}{c} \frac{3}{8} \le \alpha_1 \le \frac{5-8\theta}{12} \\ \frac{1}{2} - \alpha_1 \le \alpha_2 \le \frac{1}{8} \\ \frac{1}{8} - \frac{\alpha_2}{2} \le \alpha_3 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{8} - \frac{\alpha_2}{2} \le \alpha_4 \le \alpha_3 \end{array} \right\} \cup \left\{ \begin{array}{c} \frac{5-8\theta}{12} \le \alpha_1 \le \frac{7}{16} - \theta \\ 2\alpha_1 - \frac{3}{4} + 2\theta \le \alpha_2 \le \frac{1}{8} \\ \frac{1}{8} - \frac{\alpha_2}{2} \le \alpha_3 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{8} - \frac{\alpha_2}{2} \le \alpha_4 \le \alpha_3 \end{array} \right\}
$$

Region B_2 is defined by $\alpha_1 + \alpha_4 \le \frac{1}{2} - \theta, \alpha_1 + \alpha_3 \ge \frac{1}{2}, \alpha_1 + 2\alpha_4 \ge \frac{1}{2}, \alpha_2 + \alpha_3 + \alpha_4 \ge \frac{1}{4}$, which

.

becomes

$$\left\{\begin{array}{c}\frac{3}{8}\leq\alpha_1\leq\frac{2}{5}\\ \frac{1}{2}-\alpha_1\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{2}{5}\leq\alpha_1\leq\frac{5}{12}-\frac{\theta}{3}\\ \frac{1}{2}-\alpha_1\leq\alpha_2\leq\frac{\alpha_1}{4}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{2}{5}\leq\alpha_1\leq\frac{3-4\theta}{7}\\ \frac{\alpha_1}{4}\leq\alpha_2\leq\frac{3\alpha_1-1}{2}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\frac{\alpha_1}{2}-\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{2}{5}\leq\alpha_1\leq\frac{3-4\theta}{7}\\ \frac{\alpha_1}{4}\leq\alpha_2\leq\frac{3\alpha_1-1}{2}\\ \frac{\alpha_1}{2}-\alpha_2\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{2}{5}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ \frac{3\alpha_1-1}{2}\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{5}{12}-\frac{\theta}{3}\leq\alpha_1\leq\frac{3-4\theta}{7}\\ \frac{\alpha_1+\theta-\frac{1}{4}}{2}\leq\alpha_2\leq2\alpha_1-\frac{3}{4}+\theta\\ \alpha_1-\alpha_2-\frac{1}{4}+\theta\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{5}{12}-\frac{\theta}{3}\leq\alpha_1\leq\frac{3-4\theta}{7}\\ 2\alpha_1-\frac{3}{4}+\theta\leq\alpha_2\leq\frac{\alpha_1}{4}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{3-4\theta}{7}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ \frac{\alpha_1+\theta-\frac{1}{4}}{2}\leq\alpha_2\leq\frac{\alpha_1}{4}\\ \alpha_1-\alpha_2-\frac{1}{4}+\theta\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{3-4\theta}{7}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ \frac{\alpha_1}{4}\leq\alpha_2\leq2\alpha_1-\frac{3}{4}+\theta\\ \alpha_1-\alpha_2-\frac{1}{4}+\theta\leq\alpha_3\leq\frac{\alpha_1}{2}-\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{3-4\theta}{7}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ \frac{\alpha_1}{4}\leq\alpha_2\leq2\alpha_1-\frac{3}{4}+\theta\\ \frac{\alpha_1}{2}-\alpha_2\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{3-4\theta}{7}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ 2\alpha_1-\frac{3}{4}+\theta\leq\alpha_2\leq\frac{3\alpha_1-1}{2}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\frac{\alpha_1}{2}-\alpha_2\\ \frac{1}{4}-\alpha_2-\alpha_3\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}\cup\left\{\begin{array}{c}\frac{3-4\theta}{7}\leq\alpha_1\leq\frac{1}{2}-2\theta\\ 2\alpha_1-\frac{3}{4}+\theta\leq\alpha_2\leq\frac{3\alpha_1-1}{2}\\ \frac{\alpha_1}{2}-\alpha_2\leq\alpha_3\leq\alpha_2\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_4\leq\frac{1}{2}-\theta-\alpha_1\end{array}\right\}.$$

and Region B_3 is defined by $\alpha_1+\alpha_4\geq\frac{1}{2},\alpha_2+\alpha_3+2\alpha_4\geq\frac{1}{4}$, which becomes

$$\left\{\begin{array}{c}\frac{3}{8}\leq\alpha_1\leq\frac{7}{16}\\ \frac{1}{2}-\alpha_1\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{2}-\alpha_1\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{7}{16}\leq\alpha_1\leq\frac{1}{2}-\theta\\ \frac{1}{16}\leq\alpha_2\leq\alpha_1-\frac{3}{8}\\ \frac{1}{12}-\frac{\alpha_2}{3}\leq\alpha_3\leq\alpha_2\\ \frac{\frac{1}{4}-\alpha_2-\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{7}{16}\leq\alpha_1\leq\frac{11}{24}\\ \alpha_1-\frac{3}{8}\leq\alpha_2\leq3\alpha_1-\frac{5}{4}\\ \frac{1}{12}-\frac{\alpha_2}{3}\leq\alpha_3\leq2\alpha_1-\alpha_2-\frac{3}{4}\\ \frac{\frac{1}{4}-\alpha_2-\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{7}{16}\leq\alpha_1\leq\frac{11}{24}\\ \alpha_1-\frac{3}{8}\leq\alpha_2\leq3\alpha_1-\frac{5}{4}\\ 2\alpha_1-\alpha_2-\frac{3}{4}\leq\alpha_3\leq\alpha_2\\ \frac{1}{2}-\alpha_1\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{7}{16}\leq\alpha_1\leq\frac{11}{24}\\ 3\alpha_1-\frac{5}{4}\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{2}-\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{2}-\alpha_1\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{11}{24}\leq\alpha_1\leq\frac{1}{2}-\theta\\ \alpha_1-\frac{3}{8}\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{12}-\frac{\alpha_2}{3}\leq\alpha_3\leq2\alpha_1-\alpha_2-\frac{3}{4}\\ \frac{\frac{1}{4}-\alpha_2-\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{11}{24}\leq\alpha_1\leq\frac{1}{2}-\theta\\ \alpha_1-\frac{3}{8}\leq\alpha_2\leq\frac{1}{8}\\ 2\alpha_1-\alpha_2-\frac{3}{4}\leq\alpha_3\leq\alpha_2\\ \frac{1}{2}-\alpha_1\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

What is left is region C, defined by $\theta\leq\alpha_2\leq\alpha_1,\alpha_1+\alpha_2\leq\frac{1}{2}-\theta$. Starting up with $C1_1$

defined by $\alpha_1+\alpha_2+\alpha_3 \le \frac{1}{2}-\theta, \alpha_4 \ge \frac{1}{8}$ we get

$$\left\{\begin{array}{c} \frac{1}{8} \le \alpha_1 \le \frac{1-2\theta}{6} \\ \frac{1}{8} \le \alpha_2 \le \alpha_1 \\ \frac{1}{8} \le \alpha_3 \le \alpha_2 \\ \frac{1}{8} \le \alpha_4 \le \alpha_3 \end{array}\right\} \cup \left\{\begin{array}{c} \frac{1-2\theta}{6} \le \alpha_1 \le \frac{1}{4}-\theta \\ \frac{1}{8} \le \alpha_2 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{8} \le \alpha_3 \le \alpha_2 \\ \frac{1}{8} \le \alpha_4 \le \alpha_3 \end{array}\right\}$$

$$\cup \left\{\begin{array}{c} \frac{1-2\theta}{6} \le \alpha_1 \le \frac{3-8\theta}{16} \\ \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_2 \le \alpha_1 \\ \frac{1}{8} \le \alpha_3 \le \frac{1}{2}-\theta-\alpha_1-\alpha_2 \\ \frac{1}{8} \le \alpha_4 \le \alpha_3 \end{array}\right\} \cup \left\{\begin{array}{c} \frac{3-8\theta}{16} \le \alpha_1 \le \frac{1}{4}-\theta \\ \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_2 \le \frac{3}{8}-\theta-\alpha_1 \\ \frac{1}{8} \le \alpha_3 \le \frac{1}{2}-\theta-\alpha_1-\alpha_2 \\ \frac{1}{8} \le \alpha_4 \le \alpha_3 \end{array}\right\}$$

and with $C2_1$, defined by $\alpha_1+\alpha_2+\alpha_4 \le \frac{1}{2}-\theta, \alpha_1+\alpha_2+\alpha_3 \ge \frac{1}{2}$, $\alpha_3+\alpha_4 \ge \frac{1}{4}$ we get

$$\left\{\begin{array}{c} \frac{1}{6} \le \alpha_1 \le \frac{3-4\theta}{16} \\ \frac{1}{4}-\frac{\alpha_1}{2} \le \alpha_2 \le \alpha_1 \\ \frac{1}{2}-\alpha_1-\alpha_2 \le \alpha_3 \le \alpha_2 \\ \frac{1}{4}-\alpha_3 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_2 \end{array}\right\} \cup \left\{\begin{array}{c} \frac{3-4\theta}{16} \le \alpha_1 \le \frac{1}{4}-\theta \\ \frac{1}{4}-\frac{\alpha_1}{2} \le \alpha_2 \le \frac{3-4\theta}{8}-\alpha_1 \\ \frac{1}{2}-\alpha_1-\alpha_2 \le \alpha_3 \le \alpha_2 \\ \frac{1}{4}-\alpha_3 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_2 \end{array}\right\}$$

$$\cup \left\{\begin{array}{c} \frac{3-4\theta}{16} \le \alpha_1 \le \frac{1}{4}-\theta \\ \frac{3-4\theta}{8}-\alpha_1 \le \alpha_2 \le \alpha_1 \\ \alpha_1+\alpha_2-\frac{1}{4}+\theta \le \alpha_3 \le \alpha_2 \\ \frac{1}{4}-\alpha_3 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_2 \end{array}\right\}$$

In $C3$, defined by $\alpha_1+\alpha_3+\alpha_4 \le \frac{1}{2}-\theta, \alpha_1+\alpha_2+\alpha_4 \ge \frac{1}{2}$, the splitting gives 15 regions for each of the two convex subregions. We will therefore write them down in a shorter form, which is still easy to read. Thus in $C3_1$, defined by $\alpha_2+\alpha_3+2\alpha_4 \geq \frac{1}{2}$, we get:

$$\text{Each of } \left\{ \begin{array}{c} \frac{1+2\theta}{6} \leq \alpha_1 \leq \frac{1}{5} \\ \frac{\frac{1}{2}+\theta-\alpha_1}{2} \leq \alpha_2 \leq \alpha_1 \end{array} \right\}, \left\{ \begin{array}{c} \frac{1}{5} \leq \alpha_1 \leq \frac{1-2\theta}{4} \\ \frac{\frac{1}{2}+\theta-\alpha_1}{2} \leq \alpha_2 \leq \frac{1-3\alpha_1}{2} \end{array} \right\}$$

$$\text{combined with each of } \left\{ \begin{array}{c} \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_3 \leq \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_4 \leq \alpha_3 \end{array} \right\},$$

$$\left\{ \begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \leq \alpha_3 \leq \alpha_2-\theta \\ \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_4 \leq \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array} \right\};$$

$$\text{each of } \left\{ \begin{array}{c} \frac{1}{5} \leq \alpha_1 \leq \frac{3-2\theta}{14} \\ \frac{1-3\alpha_1}{2} \leq \alpha_2 \leq \alpha_1 \end{array} \right\}, \left\{ \begin{array}{c} \frac{3-2\theta}{14} \leq \alpha_1 \leq \frac{1-2\theta}{4} \\ \frac{1-3\alpha_1}{2} \leq \alpha_2 \leq \frac{\frac{3}{2}-\theta-5\alpha_1}{2} \end{array} \right\}$$

$$\text{combined with each of } \left\{ \begin{array}{c} \frac{1}{6}-\frac{\alpha_2}{3} \leq \alpha_3 \leq 2\alpha_1+\alpha_2-\frac{1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \leq \alpha_4 \leq \alpha_3 \end{array} \right\},$$

$$\left\{ \begin{array}{c} 2\alpha_1+\alpha_2-\frac{1}{2} \leq \alpha_3 \leq \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_4 \leq \alpha_3 \end{array} \right\}, \left\{ \begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \leq \alpha_3 \leq \alpha_2-\theta \\ \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_4 \leq \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array} \right\};$$

$$\left\{ \begin{array}{c} \frac{3-2\theta}{14} \leq \alpha_1 \leq \frac{1-2\theta}{4} \\ \frac{\frac{3}{2}-\theta-5\alpha_1}{2} \leq \alpha_2 \leq \alpha_1 \end{array} \right\} \text{ combined with each of } \left\{ \begin{array}{c} \frac{1}{6}-\frac{\alpha_2}{3} \leq \alpha_3 \leq \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \leq \alpha_4 \leq \alpha_3 \end{array} \right\},$$

$$\left\{ \begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \leq \alpha_3 \leq 2\alpha_1+\alpha_2-\frac{1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \leq \alpha_4 \leq \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array} \right\}, \left\{ \begin{array}{c} 2\alpha_1+\alpha_2-\frac{1}{2} \leq \alpha_3 \leq \alpha_2-\theta \\ \frac{1}{2}-\alpha_1-\alpha_2 \leq \alpha_4 \leq \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array} \right\};$$

$$\text{and } \left\{ \begin{array}{c} \frac{1-2\theta}{4} \leq \alpha_1 \leq \frac{3}{10}-\theta \\ \frac{3\alpha_1-\frac{1}{2}+3\theta}{2} \leq \alpha_2 \leq \frac{1}{2}-\theta-\alpha_1 \end{array} \right\} \text{ combined with each of}$$

.

$$\left\{\begin{array}{c} \frac{1}{6}-\frac{\alpha_2}{3} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \alpha_3 \end{array}\right\}, \left\{\begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le \frac{1}{2}-2\theta-2\alpha_1+\alpha_2 \\ \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2} \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array}\right\}$$

Similarly, in $C3_2$, defined by $\alpha_1 \ge \frac{1}{4}, \alpha_2+\alpha_4 \ge \frac{1}{4}, \alpha_3+2\alpha_4 \ge \frac{1}{4}, \alpha_2+\alpha_3+2\alpha_4 \le \frac{1}{2}$, we write the solution as

$$\left\{\begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{1}{3}-\theta \\ \frac{\alpha_1+\theta}{2} \le \alpha_2 \le \frac{1}{6} \end{array}\right\} \text{ combined with each of}$$

$$\left\{\begin{array}{c} \frac{1}{4}-\alpha_2 \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{4}-\alpha_2 \le \alpha_4 \le \alpha_3 \end{array}\right\}, \left\{\begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le \frac{1}{4}-\theta-\alpha_1+\alpha_2 \\ \frac{1}{4}-\alpha_2 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array}\right\};$$

$$\left\{\begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{1}{3}-\theta \\ \frac{1}{6} \le \alpha_2 \le \frac{1-\theta-\alpha_1}{4} \end{array}\right\} \text{ combined with each of } \left\{\begin{array}{c} \frac{1}{12} \le \alpha_3 \le 2\alpha_2-\frac{1}{4} \\ \frac{1}{8}-\frac{\alpha_3}{2} \le \alpha_4 \le \alpha_3 \end{array}\right\},$$

$$\left\{\begin{array}{c} 2\alpha_2-\frac{1}{4} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{4}-\alpha_2 \le \alpha_4 \le \alpha_3 \end{array}\right\}, \left\{\begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le \frac{1}{4}-\theta-\alpha_1+\alpha_2 \\ \frac{1}{4}-\alpha_2 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array}\right\};$$

$$\text{each of } \left\{\begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{3}{10}-\theta \\ \frac{1-\theta-\alpha_1}{4} \le \alpha_2 \le \frac{3\alpha_1-\frac{1}{2}+3\theta}{2} \end{array}\right\}, \left\{\begin{array}{c} \frac{3}{10}-\theta \le \alpha_1 \le \frac{1}{3}-\theta \\ \frac{1-\theta-\alpha_1}{4} \le \alpha_2 \le \frac{1}{2}-\theta-\alpha_1 \end{array}\right\}$$

$$\text{combined with each of } \left\{\begin{array}{c} \frac{1}{12} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_1}{2} \\ \frac{1}{8}-\frac{\alpha_3}{2} \le \alpha_4 \le \alpha_3 \end{array}\right\},$$

$$\left\{\begin{array}{c} \frac{\frac{1}{2}-\theta-\alpha_1}{2} \le \alpha_3 \le 2\alpha_2-\frac{1}{4} \\ \frac{1}{8}-\frac{\alpha_3}{2} \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array}\right\}, \left\{\begin{array}{c} 2\alpha_2-\frac{1}{4} \le \alpha_3 \le \frac{1}{4}-\theta-\alpha_1+\alpha_2 \\ \frac{1}{4}-\alpha_2 \le \alpha_4 \le \frac{1}{2}-\theta-\alpha_1-\alpha_3 \end{array}\right\};$$

$$\left\{\begin{array}{c} \frac{1}{4} \le \alpha_1 \le \frac{3}{10}-\theta \\ \frac{3\alpha_1-\frac{1}{2}+3\theta}{2} \le \alpha_2 \le \frac{1}{2}-\theta-\alpha_1 \end{array}\right\} \text{ combined with each of}$$

.

$$\left\{\begin{array}{c}\frac{1}{12} \le \alpha_3 \le \frac{1}{6} - \frac{\alpha_2}{3} \\ \frac{1}{8} - \frac{\alpha_3}{2} \le \alpha_4 \le \alpha_3\end{array}\right\}, \left\{\begin{array}{c}\frac{1}{6} - \frac{\alpha_2}{3} \le \alpha_3 \le \frac{1}{2} - 2\theta - 2\alpha_1 + \alpha_2 \\ \frac{1}{8} - \frac{\alpha_3}{2} \le \alpha_4 \le \frac{\frac{1}{2}-\alpha_2-\alpha_3}{2}\end{array}\right\},$$

$$\left\{\begin{array}{c}\frac{1}{2} - 2\theta - 2\alpha_1 + \alpha_2 \le \alpha_3 \le 2\alpha_2 - \frac{1}{4} \\ \frac{1}{8} - \frac{\alpha_3}{2} \le \alpha_4 \le \frac{1}{2} - \theta - \alpha_1 - \alpha_3\end{array}\right\}, \left\{\begin{array}{c}2\alpha_2 - \frac{1}{4} \le \alpha_3 \le \frac{1}{4} - \theta - \alpha_1 + \alpha_2 \\ \frac{1}{4} - \alpha_2 \le \alpha_4 \le \frac{1}{2} - \theta - \alpha_1 - \alpha_3\end{array}\right\}$$

In the difficult Region $C4$, defined by $\alpha_2 + \alpha_3 + \alpha_4 \le \frac{1}{2} - \theta, \alpha_1 + \alpha_3 + \alpha_4 \ge \frac{1}{2}$, we first consider the easy case, where we have $\alpha_1 \ge \frac{1}{4}, \alpha_3 + \alpha_4 \ge \frac{1}{4}$ which gives

$$\left\{\begin{array}{c}\frac{1}{4} \le \alpha_1 \le \frac{1-2\theta}{3} \\ \frac{1}{8} \le \alpha_2 \le \frac{1-2\theta}{6} \\ \frac{1}{8} \le \alpha_3 \le \alpha_2 \\ \frac{1}{4} - \alpha_3 \le \alpha_4 \le \alpha_3\end{array}\right\} \cup \left\{\begin{array}{c}\frac{1}{4} \le \alpha_1 \le \frac{1-2\theta}{3} \\ \frac{1-2\theta}{6} \le \alpha_2 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{8} \le \alpha_3 \le \frac{\frac{1}{2}-\theta-\alpha_2}{2} \\ \frac{1}{4} - \alpha_3 \le \alpha_4 \le \alpha_3\end{array}\right\}$$

$$\cup \left\{\begin{array}{c}\frac{1}{4} \le \alpha_1 \le \frac{1-2\theta}{3} \\ \frac{1-2\theta}{6} \le \alpha_2 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{\frac{1}{2}-\theta-\alpha_2}{2} \le \alpha_3 \le \alpha_2 \\ \frac{1}{4} - \alpha_3 \le \alpha_4 \le \frac{1}{2} - \theta - \alpha_2 - \alpha_3\end{array}\right\} \cup \left\{\begin{array}{c}\frac{1-2\theta}{3} \le \alpha_1 \le \frac{3}{8} - \theta \\ \frac{1}{8} \le \alpha_2 \le \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{8} \le \alpha_3 \le \alpha_2 \\ \frac{1}{4} - \alpha_3 \le \alpha_4 \le \alpha_3\end{array}\right\}$$

In $C4$(mediocre/hard) the procedure will be to treat the *whole* region as if it were to be discarded - save that we may impose $\alpha_3 + 2\alpha_4 \ge \frac{1}{4}$ in $C4$(hard) - and then do a computer search for subregions that we may remove again. Thus $C4$(mediocre), defined by $\alpha_1 \le \frac{1}{4}$,

.

becomes

$$\left\{\begin{array}{c}\frac{1}{6}+\frac{2\theta}{3}\le\alpha_1\le\frac{1}{4}\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_2\le\frac{1}{6}-\frac{\theta}{3}\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_3\le\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{6}+\frac{2\theta}{3}\le\alpha_1\le\frac{1}{4}\\ \frac{1}{6}-\frac{\theta}{3}\le\alpha_2\le\alpha_1-\theta\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_3\le\frac{\frac{1}{2}-\theta-\alpha_2}{2}\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{1}{6}+\frac{2\theta}{3}\le\alpha_1\le\frac{1}{4}\\ \frac{1}{6}-\frac{\theta}{3}\le\alpha_2\le\alpha_1-\theta\\ \frac{\frac{1}{2}-\theta-\alpha_2}{2}\le\alpha_3\le\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\frac{1}{2}-\theta-\alpha_2-\alpha_3\end{array}\right\}$$

and $C4$(hard), defined by $\alpha_3+\alpha_4\le\frac{1}{4}\le\alpha_3+2\alpha_4$ becomes

$$\left\{\begin{array}{c}\frac{1}{4}\le\alpha_1\le\frac{5}{16}\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_2\le\frac{1}{8}\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_3\le\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{4}\le\alpha_1\le\frac{5}{16}\\ \frac{1}{8}\le\alpha_2\le\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{4}-\frac{\alpha_1}{2}\le\alpha_3\le\frac{1}{8}\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{4}\le\alpha_1\le\frac{1}{4}+\theta\\ \frac{1}{8}\le\alpha_2\le\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{8}\le\alpha_3\le\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\frac{1}{4}-\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{1}{4}+\theta\le\alpha_1\le\frac{5}{16}\\ \frac{1}{8}\le\alpha_2\le\frac{3}{4}-2\alpha_1\\ \frac{1}{8}\le\alpha_3\le\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\frac{1}{4}-\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{4}+\theta\le\alpha_1\le\frac{5}{16}\\ \frac{3}{4}-2\alpha_1\le\alpha_2\le\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{8}\le\alpha_3\le\frac{3}{4}-2\alpha_1\\ \frac{1}{2}-\alpha_1-\alpha_3\le\alpha_4\le\frac{1}{4}-\alpha_3\end{array}\right\}$$

.

$$\cup\left\{\begin{array}{c}\frac{1}{4}+\theta\leq\alpha_1\leq\frac{5}{16}\\ \frac{3}{4}-2\alpha_1\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{3}{4}-2\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\frac{1}{4}-\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{5}{16}\leq\alpha_1\leq\frac{1}{3}\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_2\leq\frac{3}{4}-2\alpha_1\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_3\leq\alpha_2\\ \frac{1}{2}-\alpha_1-\alpha_3\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{5}{16}\leq\alpha_1\leq\frac{1}{3}\\ \frac{3}{4}-2\alpha_1\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{4}-\frac{\alpha_1}{2}\leq\alpha_3\leq\frac{3}{4}-2\alpha_1\\ \frac{1}{2}-\alpha_1-\alpha_3\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{5}{16}\leq\alpha_1\leq\frac{1}{3}\\ \frac{3}{4}-2\alpha_1\leq\alpha_2\leq\frac{1}{8}\\ \frac{3}{4}-2\alpha_1\leq\alpha_3\leq\alpha_2\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{5}{16}\leq\alpha_1\leq\frac{1}{3}\\ \frac{1}{8}\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{3}{4}-2\alpha_1\leq\alpha_3\leq\frac{1}{8}\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{5}{16}\leq\alpha_1\leq\frac{3}{8}-\theta\\ \frac{1}{8}\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{8}\leq\alpha_3\leq\alpha_2\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\frac{1}{4}-\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{3}\leq\alpha_1\leq\frac{3}{8}-\theta\\ \frac{1}{12}\leq\alpha_2\leq\frac{1}{8}\\ \frac{1}{12}\leq\alpha_3\leq\alpha_2\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}\cup\left\{\begin{array}{c}\frac{1}{3}\leq\alpha_1\leq\frac{3}{8}-\theta\\ \frac{1}{8}\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{12}\leq\alpha_3\leq\frac{1}{8}\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

$$\cup\left\{\begin{array}{c}\frac{3}{8}-\theta\leq\alpha_1\leq\frac{5}{12}-\theta\\ \frac{1}{12}\leq\alpha_2\leq\frac{1}{2}-\theta-\alpha_1\\ \frac{1}{12}\leq\alpha_3\leq\alpha_2\\ \frac{1}{8}-\frac{\alpha_3}{2}\leq\alpha_4\leq\alpha_3\end{array}\right\}$$

Finally, there is $C5_1$, defined by $\alpha_2+\alpha_3+\alpha_4\geq\frac{1}{2}, \alpha_1+\alpha_2+\alpha_3+2\alpha_4\leq 1-\theta$, which

.

becomes

$$\left\{\begin{array}{c}\frac{1}{6} \leq \alpha_1 \leq \frac{1-\theta}{5} \\ \frac{1}{6} \leq \alpha_2 \leq \alpha_1 \\ \frac{1}{4} - \frac{\alpha_2}{2} \leq \alpha_3 \leq \alpha_2 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \alpha_3\end{array}\right\} \cup \left\{\begin{array}{c}\frac{1-\theta}{5} \leq \alpha_1 \leq \frac{1}{3} - \theta \\ \frac{1}{6} \leq \alpha_2 \leq \frac{1-\theta-\alpha_1}{4} \\ \frac{1}{4} - \frac{\alpha_2}{2} \leq \alpha_3 \leq \alpha_2 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \alpha_3\end{array}\right\}$$

$$\cup \left\{\begin{array}{c}\frac{1-\theta}{5} \leq \alpha_1 \leq \frac{1}{4} - \frac{\theta}{2} \\ \frac{1-\theta-\alpha_1}{4} \leq \alpha_2 \leq \alpha_1 \\ \frac{1}{4} - \frac{\alpha_2}{2} \leq \alpha_3 \leq \frac{1-\theta-\alpha_1-\alpha_2}{3} \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \alpha_3\end{array}\right\} \cup \left\{\begin{array}{c}\frac{1-\theta}{5} \leq \alpha_1 \leq \frac{1}{4} - \frac{\theta}{2} \\ \frac{1-\theta-\alpha_1}{4} \leq \alpha_2 \leq \alpha_1 \\ \frac{1-\theta-\alpha_1-\alpha_2}{3} \leq \alpha_3 \leq \alpha_2 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \frac{1-\theta-\alpha_1-\alpha_2-\alpha_3}{2}\end{array}\right\}$$

$$\cup \left\{\begin{array}{c}\frac{1}{4} - \frac{\theta}{2} \leq \alpha_1 \leq \frac{1}{3} - \theta \\ \frac{1-\theta-\alpha_1}{4} \leq \alpha_2 \leq \frac{1}{2} - \theta - \alpha_1 \\ \frac{1}{4} - \frac{\alpha_2}{2} \leq \alpha_3 \leq \frac{1-\theta-\alpha_1-\alpha_2}{3} \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \alpha_3\end{array}\right\} \cup \left\{\begin{array}{c}\frac{1}{4} - \frac{\theta}{2} \leq \alpha_1 \leq \frac{1}{3} - \theta \\ \frac{1-\theta-\alpha_1}{4} \leq \alpha_2 \leq \frac{1}{2} - \theta - \alpha_1 \\ \frac{1-\theta-\alpha_1-\alpha_2}{3} \leq \alpha_3 \leq \alpha_2 \\ \frac{1}{2} - \alpha_2 - \alpha_3 \leq \alpha_4 \leq \frac{1-\theta-\alpha_1-\alpha_2-\alpha_3}{2}\end{array}\right\}$$

5 Calculations (almost-primes in intervals)

Almost all the background we need has been discussed in previous chapters. Here is a very brief summary. Let $\mathcal{A}$ denote the set of integers n for which $mn \in \left(x, x + x^{\frac{1}{2}+\varepsilon}\right]$ for some m being the product of K primes between P and $2P$, where $P^K \simeq x^\theta$. To prove that the interval $\left(x, x + x^{\frac{1}{2}+\varepsilon}\right]$ contains numbers with a prime factor $\geq x^{1-\theta}$, it suffices to prove that $\mathcal{A}$ contains prime numbers, which is true if we can prove that

$$S\left(\mathcal{A}, x^{\frac{1-\theta}{2}}\right) > 0$$

As explained in Chapter 3, this is true if the following holds:

$$\begin{aligned}
&S\left(\mathcal{B}, x^{\frac{1-\theta}{2}}\right) - \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}\right)\in E} S\left(\mathcal{B}_{p_1p_2}, p_2\right) \\
&- \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}\right)\in F} \min\left(S\left(\mathcal{B}_{p_1p_2}, p_2\right), \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}, \frac{\log p_3}{\log x}, \frac{\log p_4}{\log x}\right)\in F} \begin{matrix} S\left(\mathcal{B}_{p_1p_2p_3p_4}, p_4\right) \\ \text{or} \\ S\left(\mathcal{B}_{p_2p_3tp_4}, p_4\right)\end{matrix} \right) \\
&- \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}, \frac{\log p_3}{\log x}, \frac{\log p_4}{\log x}\right)\in A_1\cup\ldots\cup C5_1} S\left(\mathcal{B}_{p_1p_2p_3p_4}, p_4\right) \\
&- \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}, \frac{\log p_3}{\log x}, \frac{\log p_4}{\log x}, \frac{\log p_5}{\log x}, \frac{\log p_6}{\log x}\right)\in X_6} S\left(\mathcal{B}_{p_1p_2p_3p_4p_5p_6}, p_6\right) \\
&- \sum_{\left(\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}, \frac{\log p_3}{\log x}, \frac{\log p_4}{\log x}, \frac{\log p_5}{\log x}, \frac{\log p_6}{\log x}, \frac{\log p_7}{\log x}, \frac{\log p_8}{\log x}\right)\in X_8} S\left(\mathcal{B}_{p_1p_2p_3p_4p_5p_6p_7p_8}, p_8\right) > 0
\end{aligned}$$

where X_6 and X_8 are 6- and 8-dimensional analogues of $A_1 \cup \ldots \cup F_8$ whose contribution will be evaluated later. There are some parts of $A \cup B \cup C \cup F$ (in 4 dimensions) which have neither been discarded nor covered by Theorem 1, and they give rise to 6-dimensional regions. X_6 thus consists of the "bad" parts of those. X_8 is defined correspondingly in 8 dimensions, except that everything that is not covered by Theorem 1 is, for practical reasons, considered a "bad" region.

The asymptotic formula for

$$\sum_{\left(\frac{\log p_1}{\log x},\ldots,\frac{\log p_k}{\log x}\right)\in V} S\left(\mathcal{B}_{p_1\ldots p_k},q\right)$$

is

$$Y \int\limits_{(\alpha_1,..,\alpha_k)\in V} \cdots \int \frac{d\alpha_1}{\alpha_1}\ldots\frac{d\alpha_k}{\alpha_k}\frac{\omega\left(\frac{1-\theta-\alpha_1-\ldots-\alpha_k}{\beta}\right)}{\beta}$$

where $\beta = \frac{\log q}{\log x}$ and $Y = \frac{x\sum\frac{\gamma(m)}{m}}{\log x}$; the γ's come from

$$\left(\sum_{P<p\leq 2P} p^{-s}\right)^K = \sum \frac{\gamma(m)}{m^s}$$

Dividing by Y everywhere simplifies the notation.

Thus the contribution from $S\left(\mathcal{B}, x^{\frac{1-\theta}{2}}\right)$ is $\frac{2}{1-\theta}\omega\left(\frac{1-\theta}{\left(\frac{1-\theta}{2}\right)}\right) = \frac{1}{1-\theta}$, which in our case is $\frac{25}{24} = 1.0416666666...$ The contribution from region E can be written in a simple way, as we here have $\frac{1-\theta-\alpha_1-\alpha_2}{\alpha_2} \in [1,2]$. With $\omega(t) = \frac{1}{t}$ in this interval we get

$$\iint\limits_{(\alpha_1,\alpha_2)\in E} \frac{d\alpha_1 d\alpha_2}{\alpha_1\alpha_2\left(1-\theta-\alpha_1-\alpha_2\right)}$$

which is 0.2323012939977473. This was found on using the previously mentioned Newton-Cotes Quadrature Formula (special case)

$$\int\limits_{x_0}^{x_0+6\triangle x} y(x)dx \approx \frac{\triangle x}{140}\left(41y_0+216y_1+27y_2+272y_3+27y_4+216y_5+41y_6\right)$$

and observing how quickly this converges with increasing number of checkpoints.

In the other regions we typically encounter larger values of $\frac{1-\theta-\alpha_1-...-\alpha_k}{\beta}$; for those I have adopted the method for calculating $\omega(t)$ described in [M1]. To sharpen the accuracy, the program automatically splits any suggested integration where the values would fall on both sides of 2 into two integrals with 2 as an endpoint. This is because $\omega(t)$ has a cusp at $t = 2$, as it is $\frac{1}{t}$ in $[1, 2]$ and $\frac{1+\log(t-1)}{t}$ in $[2, 3]$.

There are some cases where we integrate over boxes located by a computer (this is done for the *complements* of $C4$(mediocre/hard), for Region F and for the 6- and 8-dimensional terms). In these cases we use the exact (and basic) integration formulas

$$\int \cdots \int_{\alpha_i \in [u_i, v_i] \forall i} \prod_i \frac{d\alpha_i}{\alpha_i} = \prod_i \log \frac{v_i}{u_i}$$

with possible modifications if some of the intervals intersect, i.e.,

$$\int \cdots \int_{u \leq \alpha_k \leq ... \alpha_1 \leq v} \prod_i \frac{d\alpha_i}{\alpha_i} = \frac{1}{k!} \left(\log \frac{v_i}{u_i} \right)^k$$

and we use a simple upper (or lower) bound for ω which holds in the whole box.

To get a best possible result in X_6 and X_8 I also used the result

$$\int \cdots \int_{x_1, ..., x_n \in [0,1], \sum x_i \in [k-1, k]} dx_1 ... dx_n = \frac{A(n, k)}{n!}$$

where $A(n, k)$ denotes the Euler number; see [K2]. Basically I looked up 6-dimensional "boxes" of points $(\alpha_1, ..., \alpha_6)$ in which there was a positive part in which, for every $\beta \leq \min\left(\alpha_4, \frac{1-\theta-\alpha_1-...-\alpha_4}{2}\right)$, $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \beta$ could be written as $S_1 + S_2$ where each term would go into one of the S_i's and $S_i \leq \frac{1}{2^i}$ for each i. Then the program would compare possible reductions of the volume of the box and pick the best one. If it were possible to do the same with $\alpha_1 + ... + \alpha_6 + \beta$, the factor $\frac{\omega\left(\frac{1-\theta-\alpha_1-...-\alpha_6}{\alpha_6}\right)}{\alpha_6}$ could be replaced by $\left(\max \omega\left(\frac{1-\theta-\alpha_1-...-\alpha_8}{\alpha_8}\right)\right)\left(\frac{1}{\alpha_6}+\frac{\log\frac{\alpha_6}{\theta}-1}{\theta}\right)$; the latter factor being equal to

$$\int_{\theta}^{\alpha_6}\frac{d\alpha_7}{\alpha_7}\int_{\theta}^{\alpha_7}\frac{d\alpha_8}{\alpha_8^2}$$

Clearly these are crude methods which are only acceptable because the contributions from these regions are so small anyway. I also used upper/lower bounds for the Buchstab function in all the "box-search cases". For example, we have $\omega(x) \leq \frac{1}{c}$ for every $x \geq c$, where c satisfies $c \log c = 1$, so $c = 1.76322...$

As a rule of thumb, decimals that remained unchanged when the number of checkpoints in each dimension was doubled were accepted, although the change in the next decimal was considered. The results were:

$S\left(\mathcal{B}, x^{\frac{1-\theta}{2}}\right) \rightarrowtail$1.041666666666666....(+)

Region $E \rightarrowtail$0.2323012939977473 (-)

Region $A \rightarrowtail$0.0070119 (-)

Region $B \rightarrowtail$0.0638410 (-)

Region $C1 \rightarrowtail$0.0012679298 (-)

Region $C2 \rightarrowtail$0.002119988 (-)

Region $C3 \rightarrowtail$0.0246993 (-)

Region $C4$(easy)$\rightarrowtail$0.0154415 (-)

Region $C4$(mediocre)(total)$\rightarrowtail$0.00628690 (-)

Region $C4$(hard)(total - upper bound)$\rightarrowtail$0.05044928 (-)

Region $C4$(mediocre/hard)(removed - lower bound)

$\rightarrowtail 0.001166912 + 0.0210369316 = 0.0222038436$ (+)

Region $C5 \rightarrowtail$0.0210566 (-)

Region F(2-dim. estimate minus lower bound for removed part)

$\rightarrowtail 0.790357746958687954 - 0.17088863515 < 0.61946911181$ (-)

Regions X_6, X_8(upper bound)$\rightarrowtail$0.0107548463 (-)

Total sum $> 0.00917 > \frac{1}{110}$.

On switching back from $\mathcal{B}$ to $\mathcal{A}$, we multiply by $\frac{x^{\frac{1}{2}+\varepsilon}}{x}$ and obtain:

Theorem 4 *If* $\theta \geq \frac{1}{25}$, *then* $S\left(\mathcal{A}, x^{\frac{1-\theta}{2}}\right) > \frac{1}{110} \frac{x^{\frac{1}{2}+\varepsilon} \sum \frac{\gamma(m)}{m}}{\log x}$ *for every positive* ε *and all sufficiently large* x.

6 The results of Bombieri *et al.*

This Chapter is devoted to a summary of the relevant results of [B4].

Let ε be a "small" positive number as usual, and let $x \sim X$ stand for $X < x \leq 2X$. Let $\{\alpha_m\}_{m\sim M}$, $\{\beta_n\}_{n\sim N}$, $\{\gamma_q\}_{q\sim Q}$ and $\{\delta_r\}_{r\sim R}$ be four sequences. We define

$$\|\alpha\| = \left(\sum_m |\alpha_m|^2\right)^{\frac{1}{2}}$$

and the meaning of $\|\beta\|$ is corresponding. We shall assume that the sequences satisfy some or all of the following conditions:

(A_1) $M = x^{1-\nu}$, $N = x^{\nu}$, with $\varepsilon \leq \nu \leq 1-\varepsilon$.

(A_2) $\{\beta_n\}_{n\sim N}$ is well distributed in arithmetic progressions to small moduli, that is, for any $d \geq 1$, $k \geq 1$, $l \neq 0$, $(k,l) = 1$ we have

$$\sum_{n\equiv l (\operatorname{mod} k),(n,d)=1} \beta_n - \frac{1}{\phi(k)} \sum_{(n,dk)=1} \beta_n \ll \|\beta\| N^{\frac{1}{2}} \tau(d)^B (\log 2N)^{-A}$$

with some $B > 0$ and any $A > 0$, the constant implied in $\ll$ depending on A alone.

(A_3)$|\gamma_q| \leq \tau(q)^B$, $|\delta_r| \leq \tau(r)^B$, $QR < x$.

(A_4) $\beta_n = 0$ if n has a prime factor $\leq N_0$.

(A_5) $N^{1-\varepsilon} \sum_n |\beta_n|^4 \ll \left(\sum_n |\beta_n|^2\right)^2$

(A_6) $\alpha_m = \begin{cases} 1 \text{ if } (m, P(z_0)) = 1 \text{ with } z_0 \leq \exp\left(\frac{\log x}{\log\log x}\right) \\ 0 \text{ otherwise} \end{cases}$

Bombieri *et al.* then give estimates of the sum

$$\mathcal{D}(M,N,Q,R) = \sum_{q\sim Q, r\sim R, (qr,a)=1} \gamma_q \delta_r \left(\sum\sum_{m\sim M, n\sim N, mn\equiv a (\operatorname{mod} qr)} \alpha_m \beta_n - \frac{1}{\phi(qr)} \sum\sum_{m\sim M, n\sim N, (mn,qr)=1} \alpha_m \beta_n \right)$$

under various conditions on M, N, Q and R.

Lemma 5 *If* $(A_1)-(A_4)$ *are satisfied, then we have*

$$\mathcal{D}(M,N,Q,R) \ll_A \|\alpha\| \;\; \|\beta\| \, x^{\frac{1}{2}} (\log x)^{-A}$$

with any $A>0$ *provided*

$$x^{\varepsilon}R < N < x^{-\varepsilon} \min\left\{x^{\frac{1}{2}}Q^{-\frac{1}{2}}, x^2Q^{-5}R^{-1}, xQ^{-2}R^{-\frac{1}{2}}\right\}$$

Lemma 6 *If* $(A_1)-(A_5)$ *are satisfied, then we have*

$$\mathcal{D}(M,N,Q,R) \ll_A \|\alpha\| \;\; \|\beta\| \, x^{\frac{1}{2}} (\log x)^{-A}$$

with any $A>0$ *provided*

$$x^{\varepsilon}R < N < x^{-\varepsilon} \min\left\{\left(\frac{xR}{Q^2}\right)^{\frac{1}{2}}, \left(\frac{x}{Q}\right)^{\frac{2}{5}}, \left(\frac{x^2}{Q^3}\right)^{\frac{1}{4}}\right\}$$

Lemma 7 *If* (A_1), (A_3) *and* (A_6) *are satisfied, then we have*

$$\mathcal{D}(M,N,Q,R) \ll \|\beta\| \, x^{\frac{1}{2}-\varepsilon} M^{\frac{1}{2}}$$

provided

$$M > x^{\varepsilon} \max\left\{Q, x^{-1}QR^4, Q^{\frac{1}{2}}R, x^{-2}Q^3R^4\right\}$$

Lemma 5, Lemma 6 and Lemma 7 are Theorem 1, Theorem 2 and Theorem 5* of [B4]. On applying Heath-Brown's combinatorial lemma

$$(4)\;\; \Lambda(n) = \sum_{j=1}^{J} (-1)^{j-1} \binom{J}{j} \sum_{m_1,\dots,m_j \le x^{1/J}} \mu(m_1)\dots\mu(m_j) \sum_{n_1\dots n_j m_1 \dots m_j = n} \log n_1$$

(Lemma 5 of [B4]) they deduce

Corollary 8

$$\sum_{(q,a)=1} \lambda(q) \left(\psi\,(x;q,a) - \frac{x}{\phi(q)} \right) \ll x\,(\log x)^{-A}$$

for any well factorable function $\lambda(q)$ of level Q and any $A > 0$ where $a \neq 0$, $Q = x^{\frac{4}{7}-\varepsilon}$ and the constant implied in $\ll$ depends at most on ε, a and A.

This is Theorem 10 of [B4]. The linear sieve then gives the constant 7 as previously mentioned (see Chapter 2). To see that the conditions of Lemma 5-7 allow $Q = x^{\frac{4}{7}-\varepsilon}$ requires analysis. In the next chapter I will derive the complete set of conditions that follow on assuming $Q \geq x^{\frac{4}{7}}$, from which this can be deduced.

7 The method for twin primes

When analyzing the conditions in Bombieri *et al.*'s theorems, a central problem is: What is the range for N (or M) for a given value of QR? If we assume that $QR = x^{\frac{4}{7}-\varepsilon}$, then we can apply at least one of Lemma 5, 6 and 7 if either

$$N \in \left[x^{\frac{1}{7}}, x^{\frac{3}{7}}\right]$$

or condition (A_6) is satisfied with $M > x^{\frac{8}{21}}$, as we shall see soon. When (4) is split into pieces in which the summation range for each term has the form $[K, (1+\varepsilon)K]$, there is always a subset of the terms whose product is in a satisfactory range, when $J = 7$. On the other hand, when $QR > x^{\frac{4}{7}}$, a "hole" around $x^{\frac{2}{7}}$ is formed that prevents a direct use of prime numbers in the analysis (at least with the representations of these that we know). We shall treat the case $QR > x^{\frac{4}{7}}$ by introducing a function $h(n)$ which is never smaller than $l(n)$, the characteristic function for primes, and which is a sum of functions each having a factor in the required (restricted) ranges. By the "Fundamental Lemma" (mentioned in Chapter 2) we then have

$$\pi_2(x) = \sum_{n \le x} l(n)l(n+2) \le \sum_{n \le x} h(n) \left(\lambda^+ * 1\right)(n+2) = \sum_{d \le x+2} \lambda^+(d) \sum_{n \le x, n \equiv -2 (\mathrm{mod}\, d)} h(n)$$

and the RHS can be estimated when we know $\sum h(n)$.

In the analysis of the conditions of Lemma 5-7, we will do the following simplifications. First of all, we will ignore all $x^{\pm\varepsilon}$-factors. Clearly these factors don't affect the value of the constant we end up with, and if we (as we will do here) end up with a (seemingly) irrational number, we can just round off upwards at a suitable point and forget about the

ε-s altogether. For the same reason, we may also read an inequality "$x < y$" as "$x \leq y$" if $y - x$ can be arbitrarily small (as in our treatment of almost-primes in intervals). Finally, it is simpler to work directly with the exponents $\frac{\log N}{\log x}$, $\frac{\log R}{\log x}$ etc. We have already introduced $N = x^{\nu}$, and we also put $R = x^{\rho}$. Since the product QR is so important, we put $Q = x^{\vartheta - \rho}$ so that QR has a simple form. The conditions of Lemma 5 then become

$$\rho \leq \nu \leq \min\left\{\frac{1-\vartheta+\rho}{2}, 2-5\vartheta+4\rho, 1-2\vartheta+\frac{3\rho}{2}\right\}$$

Suppose ϑ is fixed. Since both the LHS and the RHS are increasing functions of ρ, the range for ν for which there exists a satisfactory ρ is equal to the range for ρ for which there exists a satisfactory ν, which implies: We may replace "$\leq \nu \leq$" by "$\leq$", solve for ρ, and this is the range for ν. This yields

$$\left\{\begin{array}{c} \rho \leq \frac{1-\vartheta+\rho}{2} \\ \rho \leq 2-5\vartheta+4\rho \\ \rho \leq 1-2\vartheta+\frac{3\rho}{2} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{c} \rho \leq 1-\vartheta \\ \rho \geq \frac{5\vartheta-2}{3} \\ \rho \geq 4\vartheta-2 \end{array}\right\}$$

which reduces to

$$4\vartheta - 2 \leq \rho \leq 1 - \vartheta$$

as we are assuming that $\vartheta > \frac{4}{7}$. In the same way, the conditions of Lemma 6 yield

$$\left\{\begin{array}{c} \rho \leq \frac{1}{2}-\vartheta+\frac{3\rho}{2} \\ \rho \leq \frac{2}{5}(1-\vartheta+\rho) \\ \rho \leq \frac{1}{2}-\frac{3\vartheta}{4}+\frac{3\rho}{4} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{c} \rho \geq 2\vartheta-1 \\ \rho \leq \frac{2}{3}(1-\vartheta) \\ \rho \leq 2-3\vartheta \end{array}\right\}$$

which reduces to

$$2\vartheta - 1 \leq \rho \leq 2 - 3\vartheta$$

for the same reason. Note that both these inequalities imply $\vartheta \le \frac{3}{5}$; therefore we cannot achieve anything for $\vartheta > \frac{3}{5}$.

Lemma 7 refers specifically to $\{\alpha_m\}$, so let us take $M = x^{\mu}$. Here we have a more direct approach to the determination of the range. The conditions are

$$(i)\ \mu \ge \vartheta - \rho$$

$$(ii)\ \mu \ge -1 + \vartheta + 3\rho$$

$$(iii)\ \mu \ge \frac{\vartheta + \rho}{2}$$

$$(iv)\ \mu \ge -2 + 3\vartheta + \rho$$

The value of ρ that gives the largest range for μ under the conditions (i) and (iii) is $\rho = \frac{\vartheta}{3}$, for which the condition is $\mu \ge \frac{2\vartheta}{3}$. Since $\vartheta < \frac{3}{4}$, the conditions (ii) and (iv) are also satisfied if $\rho = \frac{\vartheta}{3}$ and $\mu \ge \frac{2\vartheta}{3}$.

We summarize what we have found so far as follows. For a fixed ϑ in the range $\left[\frac{4}{7}, \frac{3}{5}\right]$, we have an analogue of the Bombieri-Vinogradov Theorem for the function h with exponent ϑ, if it is a sum of functions each having a factor such that if its size is x^{ν}, then either

$$\nu \in [2\vartheta - 1, 2 - 3\vartheta] \cup [4\vartheta - 2, 1 - \vartheta]$$

or the factor satisfies (A_6) and $\nu \ge \frac{2\vartheta}{3}$.

Suppose we know the sizes of the "prime" factors of a function. Can we simplify the search for a factor in the interesting ranges? Since the boundary $\frac{2\vartheta}{3}$ is inside the interval $[4\vartheta - 2, 1 - \vartheta]$, the most important case is the one in which each (A_6)-type factor has support below $x^{4\vartheta-2}$. Then we can use the following more general result with $A = 2\vartheta - 1$, $B = 2 - 3\vartheta$, $C = 4\vartheta - 2$ and $D = 1 - \vartheta$. We will see later that the range indicated is in a sense best possible.

Lemma 9 *Suppose* $0 \le A \le B < C \le D \le \frac{1}{2}$ *with* $2A \le D$. *Let* $\{x_i\}$ *be a finite set of positive numbers smaller than* C *with* $\sum x_i = 1$, *and let* M *be the sum of a subset* Ξ. *Then there is also a sum of a subset in* $[A, B] \cup [C, D]$ *if either (a)* $A + C \le M \le 2B$ *or (b) for every integer* $n \ge 2$, M *belongs to either*

$$[0, 1 - A - nC],$$

$$[1 - nB, A]$$

or

$$(5)\quad \left(\left[\frac{nC-1}{n-1}, D - C\right] \cup \left[C - B, \frac{nD-1}{n-1}\right]\right)$$
$$\cap \left(\left[\frac{1 + (n-1)A - nD}{n}, B - A\right] \cup [1 + (n-1)A - nD, A]\right).$$

Proof: We shall assume throughout that the conditions on M are satisfied, and that there is no sum of a subset in the required range, and arrive at contradictions. Thus we may assume without loss of generality that at most one element in Ξ and at most one element in its complement are smaller than A, and that the other x_i's lie in (B, C). Indeed, if x_i and x_j are two elements in Ξ smaller than A, then their sum is smaller than D by the conditions of the Lemma, and the statement that there *is* a sum of a subset in $[A, B] \cup [C, D]$ (the desired contradiction) is even stronger with $\{x_i, x_j\}$ replaced by $x_i + x_j$. We call this to "combine" x_i and x_j. The same argument applies for the complement of Ξ. Now, if (a) is satisfied, then the requirement $M \geq A + C$ implies that at least two of the elements in Ξ lie in (B, C), which conflicts the other requirement $M \leq 2B$: contradiction. Assume that (b) is satisfied henceforth. On taking n to be larger than $\frac{1-A-B}{C}$, one can see easily that $M < A$. Under our assumptions, M is thus equal to one of the x_i's. To derive a contradiction we will select n to be the number of x_i's between B and C. This number can't be ≤ 1, as it would imply

$$\sum x_i < M + A + C < 2A + C < D + D \leq 1$$

hence the restriction "$n \geq 2$" which avoids division by zero in (5). If $M \leq 1 - A - nC$, then we get the impossible inequality

$$\sum x_i < M + A + nC \leq 1$$

Similarly, if $M \geq 1 - nB$, then we get

$$\sum x_i > M + nB \geq 1$$

It remains to derive a contradiction from the hypothesis "M belongs to (5)". If there isn't an x_i below A other than the one in Ξ, then it suffices to consider

$$M \in \left[\frac{nC-1}{n-1}, D - C\right] \cup \left[C - B, \frac{nD-1}{n-1}\right]$$

Let x_1 and x_2 be the smallest and the largest x_i in (B, C) respectively. Thus if

$$\frac{nC-1}{n-1} \leq M \leq D - C$$

then we have

$$C \leq \frac{(n-1)M+1}{n} = M + \frac{1-M}{n} \leq M + x_2 < M + C \leq D$$

so that the sum $M + x_2$ lies in $[C, D]$: contradiction. Similarly, if

$$C - B \leq M \leq \frac{nD - 1}{n - 1}$$

then we have

$$C \leq M + B < M + x_1 \leq M + \frac{1 - M}{n} = \frac{(n-1)M + 1}{n} \leq D.$$

It remains to investigate the case in which there is an x_i in the complement of Ξ below A, which can not be combined with any of the x_i's above B. In this case we shall see that it suffices to consider

$$M \in \left(\left[\frac{1 + (n-1)A - nD}{n}, B - A \right] \cup [1 + (n-1)A - nD, A] \right)$$

Let x_1 be the element in the complement of Ξ below A, and let x_2 be the smallest x_i above B. If

$$M \geq 1 + (n-1)A - nD$$

then we have

$$x_1 + x_2 \leq x_1 + \frac{1 - M - x_1}{n} = \frac{1 - M + (n-1)x_1}{n} < \frac{1 - M + (n-1)A}{n} \leq D$$

which means that x_1 and x_2 can be combined, contrary to our assumption. Finally, if

$$\frac{1 + (n-1)A - nD}{n} \leq M \leq B - A$$

then clearly we have

$$M + x_1 < M + A \leq B$$

so to avoid

$$M + x_1 \in [A, B]$$

we must have $x_1 < A - M$. But this gives, as above,

$$x_1 + x_2 \leq \frac{1 - M + (n-1)x_1}{n} < \frac{1 - M + (n-1)(A - M)}{n} = \frac{1 + (n-1)A}{n} - M \leq D.$$

This concludes the proof of Lemma 9. □

Let us see what this gives in our case where $A = 2\vartheta - 1$, $B = 2 - 3\vartheta$, $C = 4\vartheta - 2$, $D = 1 - \vartheta$. The range

$$[A + C, 2B]$$

in (a) becomes

$$[6\vartheta - 3, 4 - 6\vartheta]$$

which is non-empty if $\vartheta \leq \frac{7}{12}$. We only need to consider sums that are less than $\frac{1}{2}$, and since this interval is symmetric about $\frac{1}{2}$, we can replace it by

$$\left[6\vartheta - 3, \frac{1}{2}\right].$$

When "translating" the range (b), we must find the union of $[0, 1 - A - nC] \cup [1 - nB, A]$ and (5) for each $n \geq 2$, and then find their intersection.

The case $n = 2$: Both $[1 - nB, A] = [6\vartheta - 3, 2\vartheta - 1]$ and (5) are empty, since

$$\left[\frac{1 + (n-1)A - nD}{n}, B - A\right] \cup [1 + (n-1)A - nD, A] = [2\vartheta - 1, 3 - 5\vartheta] \cup [4\vartheta - 2, 2\vartheta - 1] = \emptyset$$

and this leaves us with $[0, 1 - A - nC] = [0, 6 - 10\vartheta]$.

The case $n = 3$: We have

$$[0, 1 - A - nC] \cup [1 - nB, A] = [0, 8 - 14\vartheta] \cup [9\vartheta - 5, 2\vartheta - 1] = \emptyset$$

so we are left with (5), which is the intersection of

$$\left[\frac{nC-1}{n-1}, D - C\right] \cup \left[C - B, \frac{nD-1}{n-1}\right] = \left[\frac{12\vartheta - 7}{2}, 3 - 5\vartheta\right] \cup \left[7\vartheta - 4, \frac{2 - 3\vartheta}{2}\right]$$

with

$$\left[\frac{1 + (n-1)A - nD}{n}, B - A\right] \cup [1 + (n-1)A - nD, A] = \left[\frac{7\vartheta - 4}{3}, 3 - 5\vartheta\right] \cup [7\vartheta - 4, 2\vartheta - 1]$$

This is easy to simplify, since

$$\frac{7\vartheta - 4}{3} = \frac{2}{3}\left(\frac{12\vartheta - 7}{2}\right) + \frac{1}{3}(3 - 5\vartheta)$$

and

$$3 - 5\vartheta < \frac{2 - 3\vartheta}{2} < 2\vartheta - 1$$

The first of these two observations implies that either $\left[\frac{12\vartheta-7}{2}, 3 - 5\vartheta\right]$ and $\left[\frac{7\vartheta-4}{3}, 3 - 5\vartheta\right]$ are both empty, or $\frac{12\vartheta-7}{2} \leq \frac{7\vartheta-4}{3}$; in either case we can write the region as

$$\left(\left[\frac{7\vartheta - 4}{3}, 3 - 5\vartheta\right] \cup \left[7\vartheta - 4, \frac{2 - 3\vartheta}{2}\right]\right) \cap \left(\left[\frac{7\vartheta - 4}{3}, 3 - 5\vartheta\right] \cup [7\vartheta - 4, 2\vartheta - 1]\right)$$

With the second observation, it is easy to see that the intersection is equal to

$$\left[\frac{7\vartheta - 4}{3}, 3 - 5\vartheta\right] \cup \left[7\vartheta - 4, \frac{2 - 3\vartheta}{2}\right]$$

The case $n = 4$: The interval $[0, 1 - A - nC] = [0, 10 - 18\vartheta]$ is empty, and we shall see that (5) is contained in $[1 - nB, A] = [12\vartheta - 7, 2\vartheta - 1]$. Indeed, the smallest number in

$$\left[\frac{nC-1}{n-1}, D - C\right] \cup \left[C - B, \frac{nD-1}{n-1}\right]$$

(if any) is at least $\min\left(\frac{nC-1}{n-1}, C-B\right) = \min\left(\frac{16\vartheta}{3}-3, 7\vartheta-4\right) = 7\vartheta-4$ which is bigger than $12\vartheta-7$, and the largest number in

$$\left[\frac{1+(n-1)A-nD}{n}, B-A\right] \cup [1+(n-1)A-nD, A]$$

(if any) is at most $\max(B-A, A) = \max(3-5\vartheta, 2\vartheta-1) = 2\vartheta-1$. So the interval we end up with is

$$[12\vartheta-7, 2\vartheta-1]$$

The case $n \geq 5$: The interval $[1-nB, A]$ covers $[0, A]$, so this doesn't give any further restriction.

The intersection: What remains is to observe that

$$\begin{aligned}
&[0, 6-10\vartheta] \cap \left(\left[\frac{7\vartheta-4}{3}, 3-5\vartheta\right] \cup \left[7\vartheta-4, \frac{2-3\vartheta}{2}\right]\right) \cap [12\vartheta-7, 2\vartheta-1] \\
&\quad = \left[\max\left(0, \frac{7\vartheta-4}{3}, 12\vartheta-7\right), \min(6-10\vartheta, 3-5\vartheta, 2\vartheta-1)\right] \\
&\quad \cup \left[\max(0, 7\vartheta-4, 12\vartheta-7), \min\left(6-10\vartheta, \frac{2-3\vartheta}{2}, 2\vartheta-1\right)\right] \\
&= \left[\max\left(\frac{7\vartheta-4}{3}, 12\vartheta-7\right), 3-5\vartheta\right] \cup \left[7\vartheta-4, \min\left(6-10\vartheta, \frac{2-3\vartheta}{2}\right)\right] \\
&\quad = \left[\max\left(\frac{7\vartheta-4}{3}, 12\vartheta-7\right), 3-5\vartheta\right] \cup \left[7\vartheta-4, \frac{2-3\vartheta}{2}\right]
\end{aligned}$$

- the last simplification follows from

$$\frac{2-3\vartheta}{2} = \frac{1}{2}(7\vartheta-4) + \frac{1}{2}(6-10\vartheta)$$

Lemma 9 doesn't settle the general case, but the region we have ended up with here is optimal, as the interested reader can check with the following examples:

$$x_1 = 2\vartheta-1, x_2 = x_3 = 2-3\vartheta, x_4 = 4\vartheta-2 \quad \left(\text{for } \frac{4}{7} < \vartheta < \frac{3}{5}\right)$$

$$x_1 = 3 - 5\vartheta, x_2 = 7\vartheta - 4, x_3 = x_4 = 2 - 3\vartheta, x_5 = 4\vartheta - 2 \quad \left(\text{for } \frac{7}{12} < \vartheta < \frac{10}{17}\right)$$

$$x_1 = \frac{2 - 3\vartheta}{2}, x_2 = x_3 = x_4 = \frac{\vartheta}{2} \quad \left(\text{for } \frac{4}{7} < \vartheta < \frac{10}{17}\right)$$

$$x_1 = \frac{7\vartheta - 4}{3}, x_2 = ... = x_8 = \frac{1 - \vartheta}{3} \quad \left(\text{for } \frac{4}{7} < \vartheta < \frac{17}{29}\right)$$

$$x_1 = 12\vartheta - 7, x_2 = ... = x_5 = 2 - 3\vartheta \quad \left(\text{for } \frac{17}{29} < \vartheta < \frac{10}{17}\right)$$

If $\vartheta > \frac{7}{12}$, then there is a problematic "hole" between $3 - 5\vartheta$ and $7\vartheta - 4$ here. We shall therefore assume that

$$\frac{4}{7} < \vartheta \leq \frac{7}{12}$$

from now on. In this range, we have $\frac{7\vartheta - 4}{3} > 12\vartheta - 7$. We may therefore summarize our findings as follows:

Corollary 10 *Suppose $\frac{4}{7} < \vartheta \leq \frac{7}{12}$, and let $\{x_i\}$ be a finite set of positive numbers smaller than $4\vartheta - 2$ with $\sum x_i = 1$. If there is a sum of a subset in*

$$\left[\frac{7\vartheta - 4}{3}, \frac{2 - 3\vartheta}{2}\right] \cup \left[6\vartheta - 3, \frac{1}{2}\right]$$

then there is also a sum of a subset in

$$[2\vartheta - 1, 2 - 3\vartheta] \cup [4\vartheta - 2, 1 - \vartheta]$$

To avoid having to mention ϑ in every other sentence, I will assume that ϑ is fixed in the rest of this Chapter.

Definition 1 *A function $f(n)$ $(n \leq x)$ has a* convenient factorization *if it is possible to express it as*

$$\sum_i a_i \sum_{s_j \in S(i,j) \forall j, s_1 \ldots s_{k(i)} = n} \prod_{j=1}^{k(i)} g_{ij}(s_j)$$

where the a_i's are constants, the $k(i)$'s and s_j's are integers, $S(i,j) = \left[x^{\gamma(i,j)}, (1+\varepsilon)x^{\gamma(i,j)}\right]$ for all i, j; and for every i we either have (a) There is a set U such that

$$\sum_{j\in U} \gamma(i,j) \in [2\vartheta - 1, 2 - 3\vartheta] \cup [4\vartheta - 2, 1 - \vartheta]$$

or (b) There is a j such that g_{ij} satisfies A_6 and $\gamma(i,j) \geq \frac{2\vartheta}{3}$. (As indicated in Chapter 2, the g's and their convolutions may be referred to as "factors".)

With Corollary 10 we may extend the range under (a) with

$$\left[\frac{7\vartheta - 4}{3}, \frac{2 - 3\vartheta}{2}\right] \cup \left[6\vartheta - 3, \frac{1}{2}\right]$$

provided that we avoid functions with just two "large" factors, neither satisfying A_6. Our object is to find a function $h(n)$ which satisfies the following criteria:

- $h(n)$ has a convenient factorization
- $h(n) \geq 0$, and $h(p) \geq 1$ for any prime p
- $\sum h(n)$ is as small as possible

Let

$$I_1 = \left[\frac{2 - 3\vartheta}{2}, 2\vartheta - 1\right]$$

$$I_2 = [2 - 3\vartheta, 4\vartheta - 2]$$

$$I_3 = [1 - \vartheta, 6\vartheta - 3]$$

$$I_4 = [4 - 6\vartheta, \vartheta]$$

$$I_5 = [3 - 4\vartheta, 3\vartheta - 1]$$

$$I_6 = \left[2 - 2\vartheta, \frac{3\vartheta}{2}\right]$$

$$I_7 = [1 - \varepsilon, 1]$$

Let β_i denote the characteristic function for prime numbers p with $\frac{\log p}{\log x}$ in I_i for $i = 1, 2, 3, 4, 5, 6, 7,$ and let $\chi(n)$ denote the characteristic function for integers whose prime factors are all bigger than $L = x^{\frac{2-3\vartheta}{2}-\varepsilon}$. We shall see what convolutions of the β_i's and χ (later referred to as the "building stones" of h) give functions with convenient factorizations, and later find a linear combination of these which is a suitable candidate for $h(n)$.

One may assume that $h(n) = 0$ if n has a factor m satisfying

$$\frac{\log m}{\log x} \in [2\vartheta - 1, 2 - 3\vartheta] \cup [4\vartheta - 2, 1 - \vartheta]$$

for otherwise, it would be trivial to add or subtract a convolution of characteristic functions for primes in regions near the prime factors of n - and this has a convenient factorization by definition.

If n has a factor m satisfying

$$\frac{\log m}{\log x} \in \left[\frac{7\vartheta - 4}{3}, \frac{2 - 3\vartheta}{2} - \varepsilon\right]$$

then it follows from the shape of the "building stones" of h that $h(n) = 0$.

Finally, if n has a factor m satisfying

$$\frac{\log m}{\log x} \in \left[6\vartheta - 3, \frac{1}{2}\right]$$

then the argument of Lemma 9(a) shows that one may assume that $h(n) = 0$ (as in the first case), except when n has only two prime factors. But in the construction of h, we will make sure that $h(n) = 0$ whenever n has 2 or 3 prime factors (for reasons to be explained later).

The following result is a simple generalization of a result by Heath-Brown (private communication).

Lemma 11 *The function $\chi(n)$ has a convenient factorization, and so does any convolution $(\chi * g)(n)$ where $g(n)$ is equal to one of $\chi(n), \beta_1(n), ..., \beta_4(n)$ or a convolution of these, as long as there is always a piece corresponding to a factor χ supported above $x^{\frac{1}{3}}$ in any splitting.*

Proof: With $z = \exp\left(\sqrt{\log x}\right)$ we have

$$\sum \frac{\chi(n)}{n^s} = \prod_{p>L} (1-p^{-s})^{-1} = \left(\prod_{p>z} (1-p^{-s})^{-1}\right)\left(\prod_{z<p\le L} (1-p^{-s})\right)$$
$$= \left(\sum \frac{Z(n)}{n^s}\right) \exp\left(-\sum_{z<p\le L}\left(\frac{1}{p^s} + \frac{1}{2p^{2s}} + \frac{1}{3p^{3s}} + ...\right)\right)$$
$$= \left(\sum \frac{Z(n)}{n^s}\right)\left(\exp\left(-\sum_{z<p\le L}\frac{1}{p^s}\right)\right)\exp\left(-\sum_{z<p\le L}\left(\frac{1}{2p^{2s}} + \frac{1}{3p^{3s}} + ...\right)\right)$$
$$= Q_1 Q_2 Q_3$$

The function $Z(n)$ in Q_1 satisfies A_6. Q_2 is a linear combination of expressions $\left(\sum \frac{P(n)}{n^s}\right)^i$ where P has support below L, and only those with $i < \sqrt{\log x}$ contribute to the terms with $n < x$. And clearly no Dirichlet coefficient of

$$\frac{1}{i!}\left(\sum \frac{P(n)}{n^s}\right)^i$$

is bigger than 1 in absolute value. Q_3 is negligible because the number of positive integers below x with a squared prime factor bigger than z is $O\left(x \exp\left(-k\sqrt{\log x}\right)\right)$.

If no factor $\chi(n)$ produced a convenient factorization, there would be a piece

$$Z * P_1 * ... * P_j$$

where the support for Z has the form

$$\left[x^{\zeta}, (1+\varepsilon)x^{\zeta}\right]$$

with $\zeta < 4\vartheta - 2$ (since Z satisfies A_6), and the support for each P_i has the form

$$\left[x^{K_i}, (1+\varepsilon)x^{K_i}\right]$$

with $K_i < \frac{7\vartheta-4}{3}$ (since they are smaller than $\frac{2-3\vartheta}{2}$ and may not belong to $\left[\frac{7\vartheta-4}{3}, \frac{2-3\vartheta}{2}\right]$). The sum of any two K_i's must also be $< \frac{7\vartheta-4}{3}$, since $2 \times \frac{7\vartheta-4}{3} < \frac{2-3\vartheta}{2}$, and trivially this extends to the sum of all of them. And since

$$\frac{7\vartheta - 4}{3} + (4\vartheta - 2) < 1 - \vartheta$$

we would also have $\zeta + \sum K_i < 4\vartheta - 2 \leq \frac{1}{3}$. This concludes the proof. □

We now have 17 functions we can use:

$$g_1(n) = \chi(n)$$

$$g_2(n) = (\chi * \chi)(n)$$

$$g_3(n) = (\chi * \chi * \chi)(n)$$

$$g_4(n) = \left(\beta_1 * \chi\right)(n)$$

$$g_5(n) = \left(\beta_1 * \chi * \chi\right)(n)$$

$$g_6(n) = (\beta_1 * \beta_1 * \chi)(n)$$

$$g_7(n) = (\beta_1 * \beta_1 * \chi * \chi)(n)$$

$$g_8(n) = (\beta_2 * \chi)(n)$$

$$g_9(n) = (\beta_2 * \chi * \chi)(n)$$

$$g_{10}(n) = (\beta_1 * \beta_1 * \beta_1 * \chi)(n)$$

$$g_{11}(n) = (\beta_1 * \beta_2 * \chi)(n)$$

$$g_{12}(n) = (\beta_3 * \chi)(n)$$

$$g_{13}(n) = (\beta_1 * \beta_1 * \beta_1 * \beta_1 * \chi)(n)$$

$$g_{14}(n) = (\beta_1 * \beta_1 * \beta_2 * \chi)(n)$$

$$g_{15}(n) = (\beta_1 * \beta_3 * \chi)(n)$$

$$g_{16}(n) = (\beta_2 * \beta_2 * \chi)(n)$$

$$g_{17}(n) = (\beta_4 * \chi)(n)$$

It is trivial to see what values these functions produce for square-free values of n, but I am including a table for the record. As I mentioned in the proof of Lemma 12, we need not worry about numbers with a squared prime factor. There are 15 types of factorizations of n that have to be considered, corresponding to the 15 partitions of the number 7.

By a *representation* for n I mean the intervals I_i that $\frac{\log p_j}{\log x}$ belong to for the prime factors of n, so that for example (1114) represents the n's for which

$$n = p_1 p_2 p_3 p_4$$

with

$$\left\{\frac{\log p_1}{\log x}, \frac{\log p_2}{\log x}, \frac{\log p_3}{\log x}\right\} \in I_1;\ \frac{\log p_4}{\log x} \in I_4$$

This is the table: The horizontal entries are the representations, the vertical entries are the functions, and the table values are the function values for the integers having those representations. Unfortunately, there is not enough room for the definitions of the functions g_i; this may make checking slightly inconvenient.

.

	7	16	25	34	115	124	133	223	1114	1123	1222	11113	11122	111112	1111111
g_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
g_2	2	4	4	4	8	8	8	8	16	16	16	32	32	64	128
g_3	3	9	9	9	27	27	27	27	81	81	81	243	243	729	2187
g_4	0	1	0	0	2	1	1	0	3	2	1	4	3	5	7
g_5	0	2	0	0	8	4	4	0	24	16	8	64	48	160	448
g_6	0	0	0	0	2	0	0	0	6	2	0	12	6	20	42
g_7	0	0	0	0	4	0	0	0	24	8	0	96	48	320	1344
g_8	0	0	1	0	0	1	0	2	0	1	3	0	2	1	0
g_9	0	0	2	0	0	4	0	8	0	8	24	0	32	32	0
g_{10}	0	0	0	0	0	0	0	0	6	0	0	24	6	60	210
g_{11}	0	0	0	0	0	1	0	0	0	2	3	0	6	5	0
g_{12}	0	0	0	1	0	0	2	1	0	1	0	1	0	0	0
g_{13}	0	0	0	0	0	0	0	0	0	0	0	24	0	120	840
g_{14}	0	0	0	0	0	0	0	0	0	2	0	0	12	120	0
g_{15}	0	0	0	0	0	0	2	0	0	2	0	4	0	0	0
g_{16}	0	0	0	0	0	0	0	2	0	0	6	0	2	0	0
g_{17}	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0

To find the optimal combination of these functions is not as difficult as it might appear with the number of different cases we have here - if we make a few reasonable assumptions.

As we shall see in Chapter 8, the contribution to

$$\sum_{m|n, m>1 \Rightarrow \frac{\log m}{\log x} \in I_1 \cup \ldots \cup I_7} n$$

from numbers with few prime factors is rather much bigger than that from numbers with more prime factors. We will therefore assume that our function $h(n)$, a linear combination of the g_i's, satisfies

$$(6)\ h(n) = \begin{cases} 1 \text{ if } n \text{ is a prime} \\ 0 \text{ if } n \text{ has 2 or 3 prime factors} \end{cases}$$

This can for example be achieved by selecting

$$h(n) = 3g_1(n) - \frac{3}{2}g_2(n) + \frac{1}{3}g_3(n)$$

which conveniently also produces nonnegative values for integers with at least four prime factors.

Let $n_{t_1 \ldots t_m}$ denote any integer with representation $(t_1 \ldots t_m)$. We then have, for *any* choice of $h(n)$,

$$6h(n_7) - 2h(n_{16}) - 6h(n_{25}) - 3h(n_{34}) + 3h(n_{124}) + 3h(n_{223}) - h(n_{1222}) = 0$$

which gives

$$h(n_{1222}) = 6$$

provided h satisfies (6). This can be checked using the table. Some other identities that I found are:

$$(7)\ 6h(n_7)-12h(n_{16})-5h(n_{34})+6h(n_{115})+6h(n_{124})+6h(n_{133})-h(n_{223})$$

$$-h(n_{1114})-6h(n_{1123})+h(n_{11122})=0$$

$$(8)\ 30h(n_7)-30h(n_{16})-6h(n_{25})-25h(n_{34})+15h(n_{124})+20h(n_{133})$$

$$+10h(n_{1114})-10h(n_{1123})-5h(n_{11113})+h(n_{111112})=0$$

$$(9)\ 90h(n_7)-105h(n_{34})-126h(n_{115})+70h(n_{133})$$

$$+105h(n_{1114})-35h(n_{11113})+h(n_{1111111})=0$$

If we, for short, replace $h(n_{1114})$, $h(n_{1123})$, $h(n_{11113})$, $h(n_{11122})$, $h(n_{111112})$ and $h(n_{1111111})$ by A, B, C, D, E and F respectively, then we get from (7)-(9) together with (6) the identities

$$(10)\ A+6B-D=6$$

$$(11)\ \ -10A+10B+5C-E=30$$

$$(12)\ \ -105A+35C-F=90$$

We require all of A, B, C, D, E, F to be nonnegative, and they should be as small as possible - particularly A and B (as they correspond to only 4 factors). (10) implies

$$A+6B\geq 6$$

Numbers with the representations (1114) and (1123) give approximately the same contribution (see Chapter 8), so it seems like the optimal choice is $A=0$, $B=1$. If we put $C=4$, $D=0$, $E=0$ and $F=50$, then (10)-(12) all hold. The contribution from numbers with representation (1111111) is extremely small, which means that the "high" value of F is not a problem. So the question is whether C could have been smaller. But then B would have to be bigger ((11) gives $2B + C \geq 6$), and the contribution from numbers with representation (11123) is more than twice as big as the contribution from numbers with representation (11113) (the quotient is indeed much bigger), so this wouldn't work.

Finally, a combination which does produce the desired values is

$$h(n) = 3g_1(n) - \frac{3}{2}g_2(n) + \frac{1}{3}g_3(n) - g_{10}(n) - \frac{1}{3}g_{13}(n) - \frac{5}{2}g_{14}(n)$$

8 Calculations (twin primes)

Although our function $h(n)$ only produces positive values for integers with representations (7), (1123), (1222), (11113) and (1111111), we shall see what contribution other numbers *would* have given to $\sum h(n)$ to justify some assumptions we made in the construction of $h(n)$. Since we can easily construct functions which take care of integers with a factor m such that $\frac{\log m}{\log x}$ is not in any of the intervals $I_1, ..., I_6$, we may assume that *all* nontrivial factors m (not only the prime factors) satisfy

$$\frac{\log m}{\log x} \in I_1 \cup ... \cup I_6$$

By the Prime Number Theorem, the number of integers below x of the form $p_1...p_k$ for a given range for each p_i, $1 \leq i \leq k-1$ is asymptotically

$$\sum_{p_1,...,p_{k-1}} \frac{\frac{x}{p_1...p_{k-1}}}{\log \frac{x}{p_1...p_{k-1}}} = \frac{x}{\log x} \sum \frac{1}{p_1} \sum \frac{1}{p_2} ... \sum \frac{1}{p_{k-1}} \frac{\log x}{\log \frac{x}{p_1...p_{k-1}}}$$

$$\sim \frac{x}{\log x} \int \frac{dl_1}{l_1 \log l_1} \int \frac{dl_2}{l_2 \log l_2} ... \int \frac{dl_{k-1}}{l_{k-1} \log l_{k-1}} \frac{\log x}{\log \frac{x}{l_1...l_{k-1}}}$$

which after the substitutions $l_i = x^{x_i}$, $1 \leq i \leq k-1$ becomes

$$\frac{x}{\log x} \int \frac{dx_1}{x_1} \int \frac{dx_2}{x_2} ... \int \frac{dx_{k-1}}{x_{k-1}} \frac{1}{1 - x_1 - ... - x_{k-1}}$$

Thus, with the notation we used in Chapter 7, we have

$$\sum_{n \leq x} h(n) \sim \frac{x}{\log x} \sum_{(t_1...t_i)} \left(h(n_{t_1...t_i}) \int_{t_1...t_i} \right)$$

where

$$\int_7 = 1$$

$$\int_{16} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx}{x}\frac{1}{1-x} = \log\frac{3\vartheta(2\vartheta-1)}{2(1-\vartheta)(2-3\vartheta)}$$

$$\int_{25} = \int_{2-3\vartheta}^{4\vartheta-2} \frac{dx}{x}\frac{1}{1-x} = \log\frac{2(2\vartheta-1)(3\vartheta-1)}{(2-3\vartheta)(3-4\vartheta)}$$

$$\int_{34} = \int_{1-\vartheta}^{6\vartheta-3} \frac{dx}{x}\frac{1}{1-x} = \log\frac{3\vartheta(2\vartheta-1)}{2(1-\vartheta)(2-3\vartheta)} = \int_{16}$$

$$\int_{115} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{x_1}^{2\vartheta-1} \frac{dx_2}{x_2}\frac{1}{1-x_1-x_2}$$

$$\int_{124} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{1-\vartheta-x_1}^{4\vartheta-2} \frac{dx_2}{x_2}\frac{1}{1-x_1-x_2}$$

$$\int_{133} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{1-\vartheta}^{\vartheta-x_1} \frac{dx_2}{x_2}\frac{1}{1-x_1-x_2}$$

$$\int_{223} = \int_{2-3\vartheta}^{\frac{\vartheta}{2}} \frac{dx_1}{x_1} \int_{x_1}^{\vartheta-x_1} \frac{dx_2}{x_2}\frac{1}{1-x_1-x_2}$$

$$\int_{1114} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{x_1}^{2\vartheta-1} \frac{dx_2}{x_2} \int_{\max(x_2,1-\vartheta-x_1-x_2)}^{2\vartheta-1} \frac{dx_3}{x_3}\frac{1}{1-x_1-x_2-x_3}$$

$$\int_{1123} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{x_1}^{2\vartheta-1} \frac{dx_2}{x_2} \int_{1-\vartheta-x_1}^{\vartheta-x_1-x_2} \frac{dx_3}{x_3}\frac{1}{1-x_1-x_2-x_3}$$

$$\int_{1222} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{1-\vartheta-x_1}^{\frac{1-x_1}{3}} \frac{dx_2}{x_2} \int_{x_2}^{\frac{1-x_1-x_2}{2}} \frac{dx_3}{x_3}\frac{1}{1-x_1-x_2-x_3}$$

$$\int_{11113} = \int_{\frac{2-3\vartheta}{2}}^{\frac{\vartheta}{4}} \frac{dx_1}{x_1} \int_{x_1}^{\frac{\vartheta-x_1}{3}} \frac{dx_2}{x_2} \int_{\max(x_2,1-\vartheta-x_1-x_2)}^{\frac{\vartheta-x_1-x_2}{2}} \frac{dx_3}{x_3} \int_{x_3}^{\vartheta-x_1-x_2-x_3} \frac{dx_4}{x_4}\frac{1}{1-x_1-x_2-x_3-x_4}$$

.

$$\int_{11122} = \int_{\frac{2-3\vartheta}{2}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{x_1}^{\frac{2\vartheta-1+x_1}{2}} \frac{dx_2}{x_2} \int_{\max(x_2,1-\vartheta-x_1-x_2)}^{2\vartheta-1+x_1-x_2} \frac{dx_3}{x_3} \int_{1-\vartheta-x_1}^{\frac{1-x_1-x_2-x_3}{2}} \frac{dx_4}{x_4} \frac{1}{1-x_1-x_2-x_3-x_4}$$

$$\int_{111112} = \int_{\frac{2-3\vartheta}{2}}^{\frac{\vartheta}{4}} \frac{dx_1}{x_1} \int_{\max\left(x_1,\frac{3-4\vartheta-3x_1}{2}\right)}^{\frac{\vartheta}{4}} \frac{dx_2}{x_2} \int_{\max(x_2,1-\vartheta-x_1-x_2)}^{\frac{\vartheta-x_2}{3}} \frac{dx_3}{x_3} \int_{x_3}^{\frac{\vartheta-x_2-x_3}{2}} \frac{dx_4}{x_4} \to$$

$$\to \int_{x_4}^{\vartheta-x_2-x_3-x_4} \frac{dx_5}{x_5} \frac{1}{1-x_1-...-x_5}$$

$$\int_{1111111} = \int_{\frac{2-3\vartheta}{2}}^{\frac{1}{7}} \frac{dx_1}{x_1} \int_{\max\left(x_1,\frac{4-5\vartheta}{4}-x_1\right)}^{\frac{1-x_1}{6}} \frac{dx_2}{x_2} \int_{\max(x_2,1-\vartheta-x_1-x_2)}^{\frac{1-x_1-x_2}{5}} \frac{dx_3}{x_3} \int_{x_3}^{\frac{1-x_1-x_2-x_3}{4}} \frac{dx_4}{x_4} \to$$

$$\to \int_{x_4}^{\frac{1-x_1-x_2-x_3-x_4}{3}} \frac{dx_5}{x_5} \int_{x_5}^{\frac{1-x_1-x_2-x_3-x_4-x_5}{2}} \frac{dx_6}{x_6} \frac{1}{1-x_1-...-x_6}$$

All the "nontrivial" limits of integration here can be derived from

$$\sum_{i\in S, x_i\in I_{j(i)}} x_i \in I_{\sum_{i\in S} j(i)}$$

which is to hold for any set S of indices.

To get as high accuracy as possible when computing these using (once again) the Newton-Cotes Quadrature formula, I will split the integrals so that there are no max-expressions - the same as I did in the end of Chapter 4 (but this time with the integrals written explicitly). We have

$$\int_{1114} = \int_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} \frac{dx_1}{x_1} \int_{x_1}^{\frac{1-\vartheta-x_1}{2}} \frac{dx_2}{x_2} \int_{1-\vartheta-x_1-x_2}^{2\vartheta-1} \frac{dx_3}{x_3} \frac{1}{1-x_1-x_2-x_3}$$

$$+ \int_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} \frac{dx_1}{x_1} \int_{\frac{1-\vartheta-x_1}{2}}^{2\vartheta-1} \frac{dx_2}{x_2} \int_{x_2}^{2\vartheta-1} \frac{dx_3}{x_3} \frac{1}{1-x_1-x_2-x_3} + \int_{\frac{1-\vartheta}{3}}^{2\vartheta-1} \frac{dx_1}{x_1} \int_{x_1}^{2\vartheta-1} \frac{dx_2}{x_2} \int_{x_2}^{2\vartheta-1} \frac{dx_3}{x_3} \frac{1}{1-x_1-x_2-x_3}$$

Let

$$\int^{(5,1)} = \frac{1}{x_1x_2x_3} \int\limits_{x_3}^{\vartheta-x_1-x_2-x_3} \frac{dx_4}{x_4} \frac{1}{1-x_1-x_2-x_3-x_4}$$

$$\int^{(5,2)} = \frac{1}{x_1x_2x_3} \int\limits_{1-\vartheta-x_1}^{\frac{1-x_1-x_2-x_3}{2}} \frac{dx_4}{x_4} \frac{1}{1-x_1-x_2-x_3-x_4}$$

$$\int^{(6)} = \frac{1}{x_1x_2x_3} \int\limits_{x_3}^{\frac{\vartheta-x_2-x_3}{2}} \frac{dx_4}{x_4} \int\limits_{x_4}^{\vartheta-x_2-x_3-x_4} \frac{dx_5}{x_5} \frac{1}{1-x_1-...-x_5}$$

$$\int^{(7)} = \frac{1}{x_1x_2x_3} \int\limits_{x_3}^{\frac{1-x_1-x_2-x_3}{4}} \frac{dx_4}{x_4} \int\limits_{x_4}^{\frac{1-x_1-x_2-x_3-x_4}{3}} \frac{dx_5}{x_5} \int\limits_{x_5}^{\frac{1-x_1-x_2-x_3-x_4-x_5}{2}} \frac{dx_6}{x_6} \frac{1}{1-x_1-...-x_6}$$

Then we have

$$\int_{11113} = \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{x_1}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int\limits_{1-\vartheta-x_1-x_2}^{\frac{\vartheta-x_1-x_2}{2}} dx_3 \int^{(5,1)} + \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{\frac{1-\vartheta-x_1}{2}}^{\frac{\vartheta-x_1}{3}} dx_2 \int\limits_{x_2}^{\frac{\vartheta-x_1-x_2}{2}} dx_3 \int^{(5,1)}$$

$$+ \int\limits_{\frac{1-\vartheta}{3}}^{\frac{\vartheta}{4}} dx_1 \int\limits_{x_1}^{\frac{\vartheta-x_1}{3}} dx_2 \int\limits_{x_2}^{\frac{\vartheta-x_1-x_2}{2}} dx_3 \int^{(5,1)}$$

$$\int_{11122} = \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{x_1}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int\limits_{1-\vartheta-x_1-x_2}^{2\vartheta-1+x_1-x_2} dx_3 \int^{(5,2)} + \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{\frac{1-\vartheta-x_1}{2}}^{\frac{2\vartheta-1+x_1}{2}} dx_2 \int\limits_{x_2}^{2\vartheta-1+x_1-x_2} dx_3 \int^{(5,2)}$$

$$+ \int\limits_{\frac{1-\vartheta}{3}}^{2\vartheta-1} dx_1 \int\limits_{x_1}^{\frac{2\vartheta-1+x_1}{2}} dx_2 \int\limits_{x_2}^{2\vartheta-1+x_1-x_2} dx_3 \int^{(5,2)}$$

$$\int_{111112} = \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{3-4\vartheta}{5}} dx_1 \int\limits_{\frac{3-4\vartheta-3x_1}{2}}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int\limits_{1-\vartheta-x_1-x_2}^{\frac{\vartheta-x_2}{3}} dx_3 \int^{(6)} + \int\limits_{\frac{3-4\vartheta}{5}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{x_1}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int\limits_{1-\vartheta-x_1-x_2}^{\frac{\vartheta-x_2}{3}} dx_3 \int^{(6)}$$

$$+ \int\limits_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int\limits_{\frac{1-\vartheta-x_1}{2}}^{\frac{\vartheta}{4}} dx_2 \int\limits_{x_2}^{\frac{\vartheta-x_2}{3}} dx_3 \int^{(6)} + \int\limits_{\frac{1-\vartheta}{3}}^{\frac{\vartheta}{4}} dx_1 \int\limits_{x_1}^{\frac{\vartheta}{4}} dx_2 \int\limits_{x_2}^{\frac{\vartheta-x_2}{3}} dx_3 \int^{(6)}$$

$$\int_{1111111} = \int_{\frac{2-3\vartheta}{2}}^{\frac{4-5\vartheta}{8}} dx_1 \int_{\frac{4-5\vartheta}{4}-x_1}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int_{1-\vartheta-x_1-x_2}^{\frac{1-x_1-x_2}{5}} dx_3 \int^{(7)} + \int_{\frac{4-5\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int_{x_1}^{\frac{1-\vartheta-x_1}{2}} dx_2 \int_{1-\vartheta-x_1-x_2}^{\frac{1-x_1-x_2}{5}} dx_3 \int^{(7)}$$

$$+ \int_{\frac{2-3\vartheta}{2}}^{\frac{1-\vartheta}{3}} dx_1 \int_{\frac{1-\vartheta-x_1}{2}}^{\frac{1-x_1}{6}} dx_2 \int_{x_2}^{\frac{1-x_1-x_2}{5}} dx_3 \int^{(7)} + \int_{\frac{1-\vartheta}{3}}^{\frac{1}{7}} dx_1 \int_{x_1}^{\frac{1-x_1}{6}} dx_2 \int_{x_2}^{\frac{1-x_1-x_2}{5}} dx_3 \int^{(7)}$$

The function $W(z)$ is the same for this sequence as for the sequence of primes $+\ 2$, because there are no small prime factors. Therefore, if one doesn't use the switching principle, the "new" constant one ends up with is

$$(13)\ \frac{4}{\vartheta}\left(1+\int_{1123} +6\int_{1222} +4\int_{11113} +50\int_{1111111}\right)$$

which has its minimum value at $\vartheta = 0.5785507$. (It has the form

$$\frac{4}{\vartheta}\left(1+O\left(\left(\vartheta-\frac{4}{7}\right)^3\right)\right)$$

so we don't need to do any calculations to know that the minimum value is below 7.) Here we have, with 7 decimals accuracy (although the integration formula allows a much higher accuracy)

$$\int_{16} = \int_{34} = 2.019172 \times 10^{-1}$$

$$\int_{25} = 2.429480 \times 10^{-1}$$

$$\int_{115} = 2.098307 \times 10^{-2}$$

$$\int_{124} = 3.883109 \times 10^{-2}$$

$$\int_{133} = 1.199154 \times 10^{-2}$$

$$\int_{223} = 1.800771 \times 10^{-2}$$

$$\int_{1114} = 1.238156 \times 10^{-3}$$

$$\int_{1123} = 1.044879 \times 10^{-3}$$

$$\int_{1222} = 5.041910 \times 10^{-4}$$

$$\int_{11113} = 1.236947 \times 10^{-5}$$

$$\int_{11122} = 7.453181 \times 10^{-6}$$

$$\int_{111112} = 3.490074 \times 10^{-8}$$

$$\int_{1111111} = 4.037450 \times 10^{-11}$$

The value of (13) is ≈ 6.9423095, which is not even as good as Fouvry and Grupp's result.

We shall now apply the switching principle, following Pan's/Fouvry and Grupp's approach with the improved form of the last term, so we let $\mathcal{A}$ be the weighted sequence given by $h(n)$ with 2 added to each term, and consider

$$(14)\ \pi_2(x) \leq S(\mathcal{A}, z) + O(z) \leq S(\mathcal{A}, z_1) - \frac{1}{2}\Omega_1 + \frac{1}{2}\Omega_2 + O(z) + O\left(\frac{x}{z_1}\right)$$

where

$$\Omega_1 = \sum_{z_1 \leq p < z} S(\mathcal{A}_p, z_1),\ \Omega_2 = \sum_{z_1 \leq p_3 < p_2 < p_1 < z} S(\mathcal{A}_{p_1 p_2 p_3}, p_2)$$

We have $z_1 = x^{\xi_1}$, $z = x^{\xi}$ where the values of the constants will be approximately $\xi_1 = 0.150$, $\xi = 0.244$ and $\vartheta = 0.578$. In the calculations, we shall assume that the following inequalities hold:

$$\tfrac{1}{7} < \xi_1 < \xi < \tfrac{1}{4}$$

$$2\xi_1 + \xi < \vartheta$$

$$2\xi_1 + 3\xi > 1$$

$$5\xi_1 + \xi < 1$$

Some of these are observations, that ease the calculations in retrospect, made from the optimal choice.

The problem about using the switching principle is that we simply can't switch the sequence given by $h(n)$. This particularly affects our evaluation of Ω_2; the evaluation of the other terms is pretty much as in [F2].

The first step in evaluating Ω_2 is to write $\mathcal{A}$ as {primes} + {other terms}. The primes can then be treated as in [F2], and we are left with "other terms". As a next step one would attempt a direct estimate of Ω_2 with the linear sieve and get the bound

$$\frac{4\,|\mathcal{B}|}{\log x}\frac{C_2}{2e^{\gamma}} \iiint\limits_{\xi_1 \le t_3 \le t_2 \le t_1 \le \xi} \frac{dt_1}{t_1}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3} F\left(\frac{\varphi - t_1 - t_2 - t_3}{t_2}\right)$$
$$= \frac{4\,|\mathcal{B}|}{\log x} C_2 \iiint\limits_{\xi_1 \le t_3 \le t_2 \le t_1 \le \xi} \frac{dt_1}{t_1}\frac{dt_2}{t_2}\frac{dt_3}{t_3}\frac{1}{\varphi - t_1 - t_2 - t_3}$$

where φ is ϑ or $\frac{4}{7}$ according to whether the sequence $\mathcal{B}$ comes from a function with a convenient factorization or not. (Here I used that $F(s) = \frac{2e^{\gamma}}{s}$ for $0 < s \le 3$.) In fact, based on our observations in Chapter 7 (p. 89-91), it is easy to see that $(\beta_1 * \beta_2 * \beta_2 * \beta_2)(n)$ must be treated separately with sieving limit $x^{\frac{4}{7}}$, whereas the rest can be approximated from above with the function

$$\frac{1}{6}g_{13}(n) + \frac{1}{2}g_{14}(n)$$

which, for a square-free number n produces the value 1 if n has representation (1123), 4 if it has representation (11113), 6 if it has representation (11122), 30 if it has representation (111112) and 140 if it has representation (1111111). As usual, we annihilate numbers with factors in the ranges that we have discussed before.

But the direct linear sieve estimate only works if

$$t_1 + t_2 + t_3 < \varphi$$

which is not true in the whole range of integration. In particular, $3\xi > \vartheta$. One way to go about this problem is to note the inequality

$$(15) \qquad \sum_{z_1 \leq p_3 < p_2 < p_1 < z} S\left(\mathcal{A}_{p_1 p_2 p_3}, p_2\right) \leq \sum_{z_1 \leq p_2 < p_1 < z} \frac{m\left(p_1, p_2\right) - 2}{m\left(p_1, p_2\right) - 1} S\left(\mathcal{A}_{p_1 p_2}, p_2\right)$$

where $m = m\left(p_1, p_2\right)$ is the maximum number of prime factors in the interval $[z_1, z)$ for integers near x, divisible by $p_1 p_2$ and with no prime factors below z_1. (Of course, the linear sieve gives the upper bound

$$\frac{4\left|\mathcal{B}\right|}{\log x} C_2 \iint_{\xi_1 \leq t_2 \leq t_1 \leq \xi} \frac{dt_1}{t_1} \frac{dt_2}{t_2} \frac{1}{\varphi - t_1 - t_2}$$

for $\sum S\left(\mathcal{A}_{p_1 p_2}, p_2\right)$, as $5\xi_1 > \vartheta$.) Indeed, for a fixed term in $\mathcal{A}$, $\sum S\left(\mathcal{A}_{p_1 p_2 p_3}, p_2\right)$ counts the number of prime factors in this region except two of them, and $\sum S\left(\mathcal{A}_{p_1 p_2}, p_2\right)$ counts the number of prime factors except one, assuming there is no prime factor below z_1. To determine m when p_1, p_2, z_1 and z are given is not too hard:

Lemma 12 *Suppose* $0 < A < B \leq 1$. *Let* m *be the maximum number of* x_i*'s that can be* $< B$ *if* $\{x_i\}$ *is a set of numbers* $\geq A$ *with* $\sum x_i = 1$. *If* $\left[\frac{1}{A}\right] > \left[\frac{1}{B}\right]$, *then* $m = \left[\frac{1}{A}\right]$, *otherwise* $m = \left[\frac{1-B}{A}\right]$.

Proof: $\left[\frac{1}{A}\right]$ is clearly an upper bound for m, and if $\left[\frac{1}{A}\right] > \left[\frac{1}{B}\right]$, then we can take

$$x_1 = ... = x_m = \frac{1}{m}$$

with $m = \left[\frac{1}{A}\right]$. On the other hand, if we have $\left[\frac{1}{A}\right] = \left[\frac{1}{B}\right]$, then it is not possible to have $x_i < B$ for each term. Indeed, this would lead to

$$\left[\frac{1}{A}\right](\max x_i) < 1, \ \left(\left[\frac{1}{A}\right] + 1\right)(\min x_i) > 1$$

Therefore, in this case there is one x_i which is $\geq B$. This implies that

$$m \leq \left[\frac{1-B}{A}\right]$$

and since we can take

$$x_1 = ... = x_m = A, \, x_{m+1} = 1 - Am \geq B$$

for this value, the proof is complete. □

Of course, whether we use $\leq$ or $<$ in the limits does not affect the numerical value in the end. So for our application, we will plug in $A = \frac{t_2}{1-t_1-t_2}$, $B = \frac{\xi}{1-t_1-t_2}$ (where t_i is related to p_i in the usual way), add 2 to the resulting m (since we have calculated the number of possible factors in addition to p_1 and p_2) and obtain the following result:

Corollary 13 *If* $\left[\frac{1-t_1-t_2}{t_2}\right] > \left[\frac{1-t_1-t_2}{\xi}\right]$, *then* $m = 2 + \left[\frac{1-t_1-t_2}{t_2}\right] = 1 + \left[\frac{1-t_1}{t_2}\right]$, *otherwise* $m = 1 + \left[\frac{1-t_1-\xi}{t_2}\right]$.

With our values of ξ_1 and ξ we have generally

$$\left[\frac{1-t_1-t_2}{\xi}\right] = 2$$

because $\xi < \frac{1}{4}$ and $2\xi_1 + 3\xi > 1$, and the conditions can be simplified as follows:

- If $t_1 + 4t_2 > 1$, then $\left[\frac{1-t_1-t_2}{t_2}\right] = 2$, and $m = 1 + \left[\frac{1-t_1-\xi}{t_2}\right] = 3$
- If $t_1 + 4t_2 < 1 < t_1 + 5t_2$, then $\left[\frac{1-t_1-t_2}{t_2}\right] > 2$, and $m = 1 + \left[\frac{1-t_1}{t_2}\right] = 5$
- If $t_1 + 5t_2 < 1$, then $m = 1 + \left[\frac{1-t_1}{t_2}\right] = 6$

We may now find the exact region in which it pays off to use (15). We have

$$\int_{\xi_1}^{t_2} \frac{dt_3}{t_3} \frac{1}{\varphi - t_1 - t_2 - t_3} = \frac{1}{\varphi - t_1 - t_2} \log \frac{t_2 (\varphi - t_1 - t_2 - \xi_1)}{\xi_1 (\varphi - t_1 - 2t_2)}$$

which has to be compared with

$$\frac{1}{\varphi - t_1 - t_2} \frac{m-2}{m-1}$$

where m depends on t_1 and t_2 as lined out above, under the restriction $t_1 + 2t_2 < \varphi$. A possible rewriting leads to comparison of

$$\frac{t_2 (\varphi - t_1 - t_2 - \xi_1)}{\xi_1 (\varphi - t_1 - 2t_2)}$$

with

$$\exp \frac{m-2}{m-1}$$

and it turns out that in the region

$$\{(t_1, t_2) \mid \xi_1 \le t_2 \le t_1 \le \xi\}$$

we have:

$$t_1 + 4t_2 > 1,\, t_1 + 2t_2 < \varphi \Rightarrow \frac{t_2 (\varphi - t_1 - t_2 - \xi_1)}{\xi_1 (\varphi - t_1 - 2t_2)} > \exp \frac{3}{4}$$

$$t_1 + 5t_2 < 1 \Rightarrow t_1 + 2t_2 < \varphi,\, \frac{t_2 (\varphi - t_1 - t_2 - \xi_1)}{\xi_1 (\varphi - t_1 - 2t_2)} < \exp \frac{3}{4}$$

for either value of φ. So if $t_1 + 4t_2 > 1$, then the factor $\psi(t_1, t_2)$ in

$$\int \frac{dt_1}{t_1} \int \frac{dt_2}{t_2} \frac{\psi(t_1, t_2)}{1 - t_1 - t_2}$$

is $\frac{m-2}{m-1} = \frac{1}{2}$, otherwise it is

$$\log \min \left(\exp \frac{3}{4}, \frac{t_2 (\varphi - t_1 - t_2 - \xi_1)}{\xi_1 (\varphi - t_1 - 2t_2)} \right)$$

It only remains to write down the integrals and give the numerical results. Since

$$\frac{\vartheta}{5} < \xi_1 < \frac{\vartheta}{3}$$

we can use the formula

$$F(s) = \frac{2e^\gamma}{s} \left(1 + \int_2^{s-1} \frac{\log(t-1)}{t} dt \right)$$

for $3 \le s \le 5$ to get an upper bound for $S(\mathcal{A}, z_1)$, namely

$$\frac{4}{\vartheta} \left(1 + \int_2^{\frac{\vartheta}{\xi_1} - 1} \frac{\log(t-1)}{t} dt \right) \left(1 + \int_{1123} + 6 \int_{1222} + 4 \int_{11113} + 50 \int_{1111111} \right) \frac{C_2 x}{\log^2 x}$$

If $2 \le s \le 4$, then we have

$$f(s) = 2e^\gamma \frac{\log(s-1)}{s}$$

and since $\xi_1 > \frac{\vartheta}{5}$, $2\xi_1 + \xi < \vartheta$, we conclude that the lower bound for $\Omega_1 = \sum_{z_1 \le p < z} S(\mathcal{A}_p, z_1)$ is

$$4 \left(\int_{\xi_1}^{\xi} \frac{\log \left(\frac{\vartheta - t}{\xi_1} - 1 \right)}{t (\vartheta - t)} dt \right) \left(1 + \int_{1123} + 6 \int_{1222} + 4 \int_{11113} + 50 \int_{1111111} \right) \frac{C_2 x}{\log^2 x}$$

Let

$$r_1(\xi_1, \varphi) = \frac{\varphi + \xi_1 \left(3 \exp \frac{3}{4} - 1 \right) - \sqrt{\varphi^2 + \xi_1^2 \left(3 \exp \frac{3}{4} - 1 \right)^2 - 2\varphi\xi_1 \left(1 + \exp \frac{3}{4} \right)}}{4}$$

$$r_2(\xi_1,\varphi,t_1)=\frac{\varphi-t_1+\xi_1\left(2\exp\frac{3}{4}-1\right)-\sqrt{(\varphi-t_1)^2+\xi_1^2\left(2\exp\frac{3}{4}-1\right)^2-2\xi_1(\varphi-t_1)}}{2}$$

$$J(\vartheta)=\int_{1123}+4\int_{11113}+6\int_{11122}+30\int_{111112}+140\int_{1111111}$$

We shall now give the upper bound for $\Omega_2=\sum_{z_1\le p_3<p_2<p_1<z} S(\mathcal{A}_{p_1p_2p_3},p_2)$. For the prime number part, Fouvry and Grupp derived the bound

$$\frac{8}{1+\xi_1}\left(\iiint\limits_{\xi_1\le t_3\le t_2\le t_1\le\xi}\omega\left(\frac{1-t_1-t_2-t_3}{t_2}\right)\frac{dt_1}{t_1}\frac{dt_2}{t_2^2}\frac{dt_3}{t_3}\right)\frac{C_2x}{\log^2 x}$$

(with the rôles of t_2 and t_3 switched) under the conditions $\frac{1}{7}<\xi_1<\frac{1}{5}$, which follow from our assumptions. To maintain the accuracy when using numerical integration, we shall split this up according to the integer part of

$$\frac{1-t_1-t_2-t_3}{t_2}$$

We have

$$\omega(s)=\begin{cases}\frac{1}{s},\ 1<s\le 2\\ \frac{1+\log(s-1)}{s},\ 2<s\le 3\\ \frac{1}{s}\left(1+\log(s-1)+\int\limits_2^{s-1}\frac{\log(u-1)}{u}du\right),\ 3<s\le 4\end{cases}$$

Since we also have $5\xi_1+\xi<1$, the contribution from the prime number terms can be written as

$$\Bigg\{\int\limits_{\xi_1}^{\xi}\frac{dt_1}{t_1}\int\limits_{\xi_1}^{t_1}\frac{dt_2}{t_2}\int\limits_{\xi_1}^{t_2}\frac{dt_3}{t_3}\frac{1}{1-t_1-t_2-t_3}+\int\limits_{\xi_1}^{\frac{1}{5}}\frac{dt_1}{t_1}\int\limits_{\xi_1}^{t_1}\frac{dt_2}{t_2}\int\limits_{\xi_1}^{t_2}\frac{dt_3}{t_3}\frac{\log\left(\frac{1-t_1-t_3}{t_2}-2\right)}{1-t_1-t_2-t_3}$$

$$+\int\limits_{\frac{1}{5}}^{\xi}\frac{dt_1}{t_1}\int\limits_{\xi_1}^{\frac{1-t_1}{4}}\frac{dt_2}{t_2}\int\limits_{\xi_1}^{t_2}\frac{dt_3}{t_3}\frac{\log\left(\frac{1-t_1-t_3}{t_2}-2\right)}{1-t_1-t_2-t_3}+\int\limits_{\frac{1}{5}}^{\xi}\frac{dt_1}{t_1}\int\limits_{\frac{1-t_1}{4}}^{\frac{1-\xi_1-t_1}{3}}\frac{dt_2}{t_2}\int\limits_{\xi_1}^{1-t_1-3t_2}\frac{dt_3}{t_3}\frac{\log\left(\frac{1-t_1-t_3}{t_2}-2\right)}{1-t_1-t_2-t_3}$$

$$+\int_{\xi_1}^{\frac{1}{6}}\frac{dt_1}{t_1}\int_{\xi_1}^{t_1}\frac{dt_2}{t_2}\int_{\xi_1}^{t_2}\frac{dt_3}{t_3}\frac{\int_2^{\frac{1-t_1-t_3}{t_2}-2}\frac{\log(u-1)}{u}du}{1-t_1-t_2-t_3}+\int_{\frac{1}{6}}^{\xi}\frac{dt_1}{t_1}\int_{\xi_1}^{\frac{1-t_1}{5}}\frac{dt_2}{t_2}\int_{\xi_1}^{t_2}\frac{dt_3}{t_3}\frac{\int_2^{\frac{1-t_1-t_3}{t_2}-2}\frac{\log(u-1)}{u}du}{1-t_1-t_2-t_3}$$

$$\left.+\int_{\frac{1}{6}}^{\xi}\frac{dt_1}{t_1}\int_{\frac{1-t_1}{5}}^{\frac{1-\xi_1-t_1}{4}}\frac{dt_2}{t_2}\int_{\xi_1}^{1-t_1-4t_2}\frac{dt_3}{t_3}\frac{\int_2^{\frac{1-t_1-t_3}{t_2}-2}\frac{\log(u-1)}{u}du}{1-t_1-t_2-t_3}\right\}\frac{8}{1+\xi_1}\frac{C_2x}{\log^2 x}$$

The contribution from "other terms", which we have just been discussing, is

$$4\left\{\left(\int_{\xi_1}^{r_1\left(\xi_1,\frac{4}{7}\right)}\frac{dt_1}{t_1}\int_{\xi_1}^{t_1}\frac{dt_2}{t_2}\frac{\log\frac{t_2\left(\frac{4}{7}-t_1-t_2-\xi_1\right)}{\xi_1\left(\frac{4}{7}-t_1-2t_2\right)}}{\frac{4}{7}-t_1-t_2}\right)\left(6\int_{1222}\right)\right.$$

$$+\left(\int_{r_1\left(\xi_1,\frac{4}{7}\right)}^{\xi}\frac{dt_1}{t_1}\int_{\xi_1}^{r_2\left(\xi_1,\frac{4}{7},t_1\right)}\frac{dt_2}{t_2}\frac{\log\frac{t_2\left(\frac{4}{7}-t_1-t_2-\xi_1\right)}{\xi_1\left(\frac{4}{7}-t_1-2t_2\right)}}{\frac{4}{7}-t_1-t_2}\right)\left(6\int_{1222}\right)$$

$$+\left(\int_{r_1\left(\xi_1,\frac{4}{7}\right)}^{\frac{1}{5}}\frac{dt_1}{t_1}\int_{r_2\left(\xi_1,\frac{4}{7},t_1\right)}^{t_1}\frac{dt_2}{t_2}\frac{\frac{3}{4}}{\frac{4}{7}-t_1-t_2}\right)\left(6\int_{1222}\right)$$

$$+\left(\int_{\frac{1}{5}}^{\xi}\frac{dt_1}{t_1}\int_{r_2\left(\xi_1,\frac{4}{7},t_1\right)}^{\frac{1-t_1}{4}}\frac{dt_2}{t_2}\frac{\frac{3}{4}}{\frac{4}{7}-t_1-t_2}\right)\left(6\int_{1222}\right)$$

$$+\left(\int_{\xi_1}^{r_1(\xi_1,\vartheta)}\frac{dt_1}{t_1}\int_{\xi_1}^{t_1}\frac{dt_2}{t_2}\frac{\log\frac{t_2(\vartheta-t_1-t_2-\xi_1)}{\xi_1(\vartheta-t_1-2t_2)}}{\vartheta-t_1-t_2}\right)J(\vartheta)$$

$$+\left(\int_{r_1(\xi_1,\vartheta)}^{\xi}\frac{dt_1}{t_1}\int_{\xi_1}^{r_2(\xi_1,\vartheta,t_1)}\frac{dt_2}{t_2}\frac{\log\frac{t_2(\vartheta-t_1-t_2-\xi_1)}{\xi_1(\vartheta-t_1-2t_2)}}{\vartheta-t_1-t_2}\right)J(\vartheta)$$

$$+\left(\int_{r_1(\xi_1,\vartheta)}^{\frac{1}{5}}\frac{dt_1}{t_1}\int_{r_2(\xi_1,\vartheta,t_1)}^{t_1}\frac{dt_2}{t_2}\frac{\frac{3}{4}}{\vartheta-t_1-t_2}\right)J(\vartheta)$$

$$+\left(\int_{\frac{1}{5}}^{\xi}\frac{dt_1}{t_1}\int_{r_2(\xi_1,\vartheta,t_1)}^{\frac{1-t_1}{4}}\frac{dt_2}{t_2}\frac{\frac{3}{4}}{\vartheta-t_1-t_2}\right)J(\vartheta)$$

$$+\left(\int_{\frac{1}{5}}^{\xi}\frac{dt_1}{t_1}\int_{\frac{1-t_1}{4}}^{t_1}\frac{dt_2}{t_2}\left(\frac{\frac{1}{2}\times 6\int_{1222}}{\frac{4}{7}-t_1-t_2}+\frac{\frac{1}{2}J(\vartheta)}{\vartheta-t_1-t_2}\right)\right)\Bigg\}\frac{C_2 x}{\log^2 x}$$

The optimal choice of parameters appears to be, with 7 decimals accuracy,

$$\vartheta = 0.5782797$$

$$\xi_1 = 0.1503298$$

$$\xi = 0.2441184$$

which gives

$$S(\mathcal{A}, z_1) < 7.72661\ 031718\frac{C_2 x}{\log^2 x}$$

$$\Omega_1 > 2.18284\ 213424\frac{C_2 x}{\log^2 x}$$

$$\Omega_2 < 0.39444\ 307702\frac{C_2 x}{\log^2 x}$$

and by (14), our "new constant" is 6.83241 07886. High accuracy is preserved everywhere in the calculations and allows this many decimals.

Theorem 14 *For sufficiently large x, we have*

$$\pi_2(x) < 6.8325 C_2\frac{x}{\log^2 x}$$

This is a modest improvement of Wu's 6.8354.

9 References

[B1] A. Balog, *Numbers with a large prime factor.* STUDIA SCI. MATH. HUNGAR., 15(1980), 139-146.

[B2] A. Balog, *Numbers with a large prime factor II.* TOPICS IN CLASSICAL NUMBER THEORY, Coll. Math. Soc. János Bolyai 34, Amsterdam: Elsevier North-Holland, 1984, 49-67.

[B3] A. Balog, G. Harman and J. Pintz, *Numbers with a large prime factor IV.* J. LONDON MATH. SOC. (2), 28(1983), 218-226.

[B4] E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli.* ACTA MATH., 156(1986), 203-251.

[C] J. R. Chen, *On the Goldbach's problem and the sieve methods.* SCI. SINICA, 21(1978), 701-739.

[F1] E. Fouvry, *Autour du théorème de Bombieri-Vinogradov.* ACTA MATH., 152(1984), 219-244.

[F2] E. Fouvry and F. Grupp, *On the switching principle in sieve theory.* J. REINE U. ANGEW. MATH., 370(1986), 101-126.

[F3] E. Fouvry and H. Iwaniec, *Primes in arithmetic progressions.* ACTA ARITH., 42(1983), 197-218.

[H1] H. Halberstam and H.-E. Richert, *Sieve methods.* Academic Press, 1974.

[H2] G. Harman, *The greatest prime factor of the integers in an interval.* (Unpublished manuscript)

[H3] D. R. Heath-Brown, *Prime numbers in short intervals and a generalized Vaughan identity.* CAN. J. MATH., 34(1982), 1365-1377.

[H4] D. R. Heath-Brown, *The largest prime factor of the integers in an interval.* SCI. SINICA, 39(1996), 449-476.

[H5] D. R. Heath-Brown and C. H. Jia, *The largest prime factor of the integers in an interval, II.* J. REINE U. ANGEW. MATH., 498(1998), 35-59.

[I1] H. Iwaniec, *Rosser's sieve.* ACTA ARITH., 36(1980), 171-202.

[I2] H. Iwaniec, *A new form of the error term in the linear sieve.* ACTA ARITH., 37(1980), 307-320.

[J] M. Jutila, *On numbers with a large prime factor.* J. INDIAN MATH. SOC. (N. S.), 37(1973), 43-53(1974).

[K1] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers.* McGraw-Hill Book Company, 1968.

[K2] L. L. Kupper and I. A. Salama, *A geometric interpretation for the Eulerian numbers.* AMER. MATH. MONTHLY, 93(1986), 51-52.

[M1] G. Marsaglia, J. C. W. Marsaglia and A. Zaman, *Numerical solution of some classical differential-difference equations.* MATH. COMP. 53(1989), 191-201.

[M2] H. L. Montgomery, *Topics in Analytic Number Theory.* Lecture notes in mathematics (227), Springer-Verlag, 1971, (Chapter 6).

[P] C. D. Pan, *A new mean value theorem and its applications.* Recent progress in analytic number theory, 1, London 1981, 275-287.

[R1] K. Ramachandra, *A note on numbers with a large prime factor.* J. London Math. Soc. (2), 1(1969), 303-306.

[R2] K. Ramachandra, *A note on numbers with a large prime factor. II.* J. Indian Math. Soc. (N. S.), 34(1970), 39-48(1971).

[W1] J.-W. Wrench Jr., *Evaluation of Artin's constant and the twin prime constant.* Math. Comp., 15(1961), 396-398.

[W2] J. Wu, *Sur la suite des nombres premiers jumeaux.* Acta Arith., 55(1990), 365-394.