Lecture Notes in Computer Science 4076:156–167, 2006.

The Mertens conjecture revisited

Tadej Kotnik[1] and Herman te Riele[2]

[1] Faculty of Electrical Engineering, University of Ljubljana, Slovenia,
tadej.kotnik@fe.uni-lj.si
[2] CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands,
herman@cwi.nl

Abstract. Let $M(x) = \sum_{1 \le n \le x} \mu(n)$ where $\mu(n)$ is the Möbius function. The Mertens conjecture that $|M(x)|/\sqrt{x} < 1$ for all $x > 1$ was disproved in 1985 by Odlyzko and te Riele [13]. In the present paper, the known lower bound 1.06 for $\limsup M(x)/\sqrt{x}$ is raised to 1.218, and the known upper bound -1.009 for $\liminf M(x)/\sqrt{x}$ is lowered to -1.229. In addition, the explicit upper bound of Pintz [14] on the smallest number for which the Mertens conjecture is false, is reduced from $\exp(3.21 \times 10^{64})$ to $\exp(1.59 \times 10^{40})$. Finally, new numerical evidence is presented for the conjecture that $M(x)/\sqrt{x} = \Omega_{\pm}(\sqrt{\log\log\log x})$.

1 Introduction

The Möbius function $\mu(n)$ is defined as follows

$$\mu(n) := \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is divisible by the square of a prime number,} \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes.} \end{cases}$$

Taking the sum of the values of $\mu(n)$ for all $1 \le n \le x$, we obtain the function

$$M(x) := \sum_{1 \le n \le x} \mu(n),$$

which is the difference between the number of squarefree positive integers $n \le x$ with an *even* number of prime factors and those with an *odd* number of prime factors. The *Mertens conjecture* [11] states that

$$|M(x)|/\sqrt{x} < 1 \quad \text{for } x > 1.$$

This, but also the weaker assumption

$$|M(x)|/\sqrt{x} < C \quad \text{for } x > 1 \text{ and some } C > 1 \tag{1}$$

would imply the truth of the Riemann hypothesis. The Mertens conjecture was shown to be false by Odlyzko and te Riele in 1985 [13]. They proved the existence of some x for which $M(x)/\sqrt{x} > 1.06$, and some other x for which $M(x)/\sqrt{x} <$

-1.009. In 1987, Pintz [14] gave an *effective* disproof of the Mertens conjecture by showing that $|M(x)|/\sqrt{x} > 1$ for some $x \le \exp(3.21 \times 10^{64})$. Nowadays, it is generally believed that the function $M(x)/\sqrt{x}$ is *unbounded*, both in the positive and in the negative direction. In [8], for example, it is conjectured that

$$M(x)/\sqrt{x} = \Omega_{\pm}(\sqrt{\log\log\log x}). \tag{2}$$

In this paper, we improve the above results by showing that there exists an x for which $M(x)/\sqrt{x} > 1.218$ and an x for which $M(x)/\sqrt{x} < -1.229$ (Section 2), and that there exists an $x < \exp(1.59 \times 10^{40})$ for which the Mertens conjecture is false (Section 3). In addition, we provide new numerical evidence to support (2) (Section 4).

Notation The complex zeros of the Riemann zeta function are denoted by $\rho_j = \frac{1}{2} + i\gamma_j$ (we work in the range where the Riemann hypothesis is known to be true) with $\gamma_1 = 14.1347...$ and $\gamma_j < \gamma_{j+1}$, $j = 1, 2, \ldots$. Furthermore, we write $\psi_j = \arg \rho_j \zeta'(\rho_j)$ and $\alpha_j = |\rho_j \zeta'(\rho_j)|^{-1}$. We also consider the zeros ρ_j ordered according to *non-increasing* values of α_j, and denote them by $\rho_j^* = \frac{1}{2} + i\gamma_j^*$ with the corresponding quantities ψ_j^*, α_j^*, $j = 1, 2, \ldots$. For example, the first five ρ_j^*'s coincide with the first five ρ_j's, but $\rho_6^* = \rho_7$, $\rho_7^* = \rho_{10}$, and $\rho_8^* = \rho_6$ (with $\alpha_6^* = \alpha_7 = 0.0163\ldots, \alpha_7^* = \alpha_{10} = 0.0141\ldots$ and $\alpha_8^* = \alpha_6 = 0.0137\ldots$).

2 Improvement of the upper and lower bounds for $M(x)/\sqrt{x}$

2.1 Background

We describe the approach which led to the disproof of the Mertens conjecture and which is the basis of the experiments which we have carried out to extend the results of Odlyzko and te Riele concerning the function $M(x)/\sqrt{x}$.

For large x, computational results on $M(x)$ are generally based on the following result due to Titchmarsh [18, Theorem 14.27].

Theorem 1. *If all the zeros of the Riemann zeta-function are simple, then there is an increasing sequence $\{T_n\}$ such that*

$$M(x) = \lim_{n\to\infty} \sum_{|\gamma|<T_n} \frac{x^\rho}{\rho\zeta'(\rho)} - R(x) + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}(2\pi/x)^{2n}}{(2n)!n\,\zeta(2n+1)} \tag{3}$$

where $R(x) = 2 - \frac{\mu(x)}{2}$ if x is an integer, and $R(x) = 2$ otherwise.

On the Riemann hypothesis, we have $\rho = \frac{1}{2} + i\gamma$, so that (3) can be rewritten as

$$\frac{M(x)}{\sqrt{x}} = 2 \lim_{n\to\infty} \sum_{0<\gamma<T_n} \frac{\cos(\gamma \log x - \psi_\gamma)}{|\rho\zeta'(\rho)|} + O(x^{-1/2}) \tag{4}$$

where we have also taken into account that in Theorem 1, $R(x) = O(1)$ and the second series is $O(x^{-2})$. Hence, as T increases, the sum in (4) will eventually converge to $M(x)/\sqrt{x}$, with the remaining error on the order of magnitude of $1/\sqrt{x}$. However, very little is known about the rate of this convergence, as the coefficients $|\rho\zeta'(\rho)|^{-1}$ do not form a monotonically decreasing sequence, but instead behave quite irregularly. For some values of x up to 10^{14}, this rate of convergence has been studied computationally, and several thousands of terms generally suffice to bring the error below 1% [8], but for much larger x this approach is not feasible.

If instead of an isolated value of $M(x)/\sqrt{x}$, a weighted average of this function in some x-range is of interest, the problem becomes somewhat more tractable. Namely, in these cases the terms of the above sum are multiplied by a function of bounded support, and the series in (3) is transformed into a finite sum. Two such cases will appear in Theorems 2 and 3 below

We write $x = e^y$, $-\infty < y < \infty$, and define

$$m(y) := M(x)x^{-1/2} = M(e^y)e^{-y/2},$$

and

$$\overline{m} := \limsup_{y\to\infty} m(y), \quad \underline{m} := \liminf_{y\to\infty} m(y).$$

Then we have the following [5–7, 13]

Theorem 2. *Let*

$$h(y,T) := 2\sum_{0<\gamma<T}\left[(1-\frac{\gamma}{T})\cos(\pi\frac{\gamma}{T}) + \pi^{-1}\sin(\pi\frac{\gamma}{T})\right]\frac{\cos(\gamma y - \psi_\gamma)}{|\rho\,\zeta'(\rho)|} \tag{5}$$

with

$$\psi_\gamma = \arg\rho\,\zeta'(\rho),$$

where $\rho = \beta + i\gamma$ are the complex zeros of the Riemann zeta function which satisfy $\beta = \frac{1}{2}$ and which are simple. Then for any y_0,

$$\underline{m} \le h(y_0,T) \le \overline{m}$$

and any value $h(y,T)$ is approximated arbitrarily closely, and infinitely often, by $M(x)/\sqrt{x}$.

Since

$$(1-t)\cos(\pi t) + \pi^{-1}\sin(\pi t) > 0 \text{ for } 0 < t < 1$$

and since it is known that $\sum_\rho |\rho\,\zeta'(\rho)|^{-1}$ diverges [18, Section 14.27], the sum of the *coefficients* of $\cos(\gamma y - \psi_\gamma)$ in (5) can be made arbitrarily large by choosing T large enough. Consequently, if we could find a value of y such that *all* of the $\gamma y - \psi_\gamma$ are close to integer multiples of 2π, then we could make $h(y,T)$ arbitrarily large. This would contradict, by Theorem 2, any conjecture of the form (1). If the γ's were linearly independent over the rationals, then by Kronecker's theorem (see, e.g., [4, Theorem 442]) there would indeed exist, for any $\epsilon > 0$, integer values of y satisfying

$$|\gamma y - \psi_\gamma - 2\pi m_\gamma| < \epsilon$$

for all $\gamma \in (0, T)$ and certain integers m_γ. This would show that $h(y, T)$, and hence $M(x)/\sqrt{x}$, can be made arbitrarily large. On the same assumptions, a similar argument can be given to imply that $h(y, T)$, and hence $M(x)/\sqrt{x}$, can be made arbitrarily large on the negative side. No good reason is known why among the γ's there should exist any linear dependencies over the rationals (see, e.g., [1]).

The approach which actually led to a disproof of the Mertens conjecture was based on the now well-known lattice basis reduction (L^3–) algorithm of Lenstra, Lenstra and Lovász [9] for finding short vectors in lattices. With this algorithm, the above mentioned inhomogeneous Diophantine approximation problem could be solved for a much larger number of terms in (5) than before. Any value of y that would come out was likely to be quite large, viz., of the order of 10^{70} in size. Therefore, it was necessary to compute the first 2000 γ's with an accuray of about 75 decimal digits (actually, 100 decimal digits were used). The best lower and upper bounds found for $\overline{m}$ and $\underline{m}$ were 1.06 and -1.009, respectively.

2.2 Computation of new lower and upper bounds for $M(x)/\sqrt{x}$

In order to find a y such that each of the numbers

$$\eta_j := (\gamma_j^* y - \psi_j^*) \bmod 2\pi, \ 1 \le j \le n, \tag{6}$$

is small, Odlyzko transformed this problem into a problem about short vectors in lattices, as described in [13]. The lattice L used is generated by the columns $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_{n+2}$ of the following $(n+2) \times (n+2)$ matrix (here $[x]$ means the greatest integer $\le x$):

$$\begin{array}{cccccc}
-[\sqrt{\alpha_1^*}\psi_1^* 2^\nu] & [\sqrt{\alpha_1^*}\gamma_1^* 2^{\nu-10}] & [2\pi\sqrt{\alpha_1^*}2^\nu] & 0 & \ldots & 0 \\
-[\sqrt{\alpha_2^*}\psi_2^* 2^\nu] & [\sqrt{\alpha_2^*}\gamma_2^* 2^{\nu-10}] & 0 & [2\pi\sqrt{\alpha_2^*}2^\nu] & \ldots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
-[\sqrt{\alpha_n^*}\psi_n^* 2^\nu] & [\sqrt{\alpha_n^*}\gamma_n^* 2^{\nu-10}] & 0 & 0 & \ldots & [2\pi\sqrt{\alpha_n^*}2^\nu] \\
2^\nu n^4 & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 & \ldots & 0
\end{array} \tag{7}$$

where ν is an integer satisfying $2n \le \nu \le 4n$. The L^3 algorithm produces a reduced basis $\underline{v}'_1, \underline{v}'_2, \ldots, \underline{v}'_{n+2}$ for the lattice L, where each new basis vector is a linear combination of the $n+2$ given basis vectors. Now the $(n+1)$-st coordinate of $\underline{v}_1$, which has value $2^\nu n^4$, is very large compared to all the other entries of the original basis. Since the reduced basis is a basis for the lattice L, it should contain precisely one vector $\underline{w}$ which has a nonzero coordinate in the $(n+1)$-st position and that coordinate should be $\pm 2^\nu n^4$. Without loss of generality this may be taken to be $2^\nu n^4$. Given the original lattice basis, the j-th coordinate of this vector $\underline{w}$ equals, for $1 \le j \le n$:

$$z\left[\sqrt{\alpha_j^*}\gamma_j^* 2^{\nu-10}\right] - \left[\sqrt{\alpha_j^*}\psi_j^* 2^\nu\right] - m_j\left[2\pi\sqrt{\alpha_j^*}2^\nu\right]$$

and the $(n+2)$-nd coordinate is z, for some integers $z, m_1, m_2, \ldots, m_n$. If the length of $\underline{w}$ is small, all of the

$$z\sqrt{\alpha_j^*}\gamma_j^* 2^{\nu-10} - \sqrt{\alpha_j^*}\psi_j^* 2^\nu - m_j 2\pi\sqrt{\alpha_j^*}2^\nu$$

will be small, i.e., all of the

$$\beta_j = \sqrt{\alpha_j^*}(y\gamma_j^* - \psi_j^* - 2\pi m_j)$$

will be very small, where $y = z/1024$. The reason for the presence of the numbers α_j^* in the lattice basis is that we want to make the sum

$$\sum_{j=1}^{n} \alpha_j^* \cos(\gamma_j^* y - \psi_j^* - 2\pi m_j)$$

large. If the cos-arguments are all close to zero, this sum will be approximately

$$\sum_{j=1}^{n} \alpha_j^* - \frac{1}{2}\sum_{j=1}^{n}[\sqrt{\alpha_j^*}(\gamma_j^* y - \psi_j^* - 2\pi m_j)]^2,$$

and therefore we want the second sum to be small. This corresponds to minimizing the euclidean norm of the vector $(\beta_1, \beta_2, \ldots, \beta_n)$ which is what the L^3 algorithm attempts to do.

In order to obtain values of y for which $h(y,T)$ will be negative, similar lattices can be used with only one change, namely that ψ_j^* is replaced by $\psi_j^* + \pi$, so that the cosine-arguments mod 2π will be close to π and the cosine-values close to -1.

We have applied the L^3 algorithm with the matrix (7) as input, for all the combinations (ν, n) in the range $\nu = 8, 9, \ldots, 400$, $n = [\nu/4], [\nu/4]+1, \ldots, [\nu/2]$. To this end we used the function *qflll* from the PARI/GP package [15]. For a given ν, the precision by which the computations were carried out was chosen to be $\log_{10}(2^{2\nu})$ decimal digits. For each combination of ν and n a number $z = z(\nu, n)$ was generated as described above and we computed the local maximum of $h(y,T)$ as defined in (5) with y in the neighborhood of $z/1024$, and $T = \gamma_{10000}$. The γ_j's were computed to an accuracy of about 250 decimal digits using the Mathematica package [10], and, as a check, using the PARI/GP package. Figure 1 gives for each $\nu = 8, 9, \ldots, 400$ and for each value of $z(\nu, n)$ which was found by the

L^3 algorithm, a scatter plot of the positive values of $h(z(\nu,n)/1024,\gamma_{2000})$. For increasing values of ν, the corresponding h-values are increasing on average, but at a rate that seems to decrease. For the negative values of h the pattern is very similar. Reaching 1.3 and -1.3 would likely require a value of ν in the neighborhood of 800.

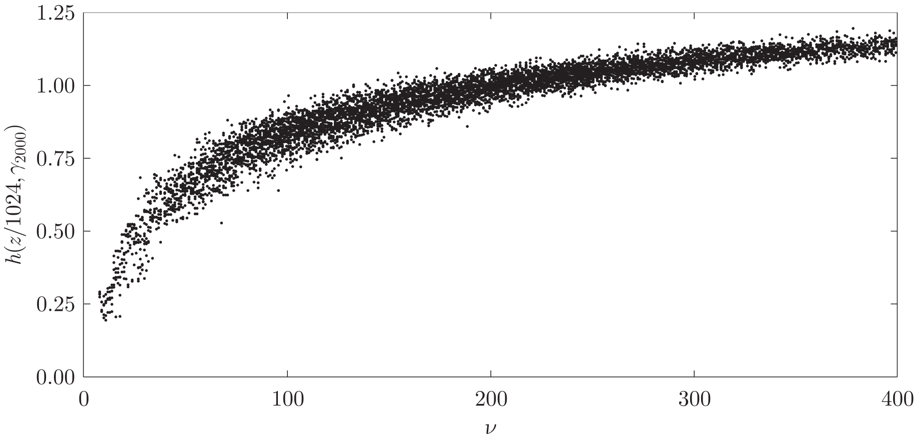

FIGURE 1

For the most promising values of h obtained, we computed the local maximum resp. minimum of $h(y,\gamma_{10000})$ in the neighborhood of $y = z/1024$. On the positive side, our champion (found with $\nu = 379, n = 98$) is

$$y = -233029271\ 5134531215\ 0140181996\ 7723401020\ 4456785091\backslash\\ 6681557518\ 6743434036\ 9240230890\ 8933261706\ 9029233958\ 2730162362.807965$$

with

$$h(y,\gamma_{10000}) = 1.218429$$

and on the negative side, our champion (found with $\nu = 396, n = 102$) is

$$y = -1608\ 7349754400\ 0919817483\ 9640165505\ 4685212472\ 2284778177\backslash\\ 5539303027\ 5350690810\ 7957194829\ 6433602695\ 1442102295\ 3212754000.679958$$

with

$$h(y,\gamma_{10000}) = -1.229385.$$

Figure 2 compares the typical behaviour of $M(e^y)/e^{y/2}$ (top) with the behaviour of $h(y,\gamma_{10000})$ around the 1.218–spike (middle) and around the -1.229–spike (bottom). Notice the four large negative spikes to the left and to the right of the champion positive spike, and the four large positive spikes to the left and to the right of the champion negative spike. This suggests that a very large spike

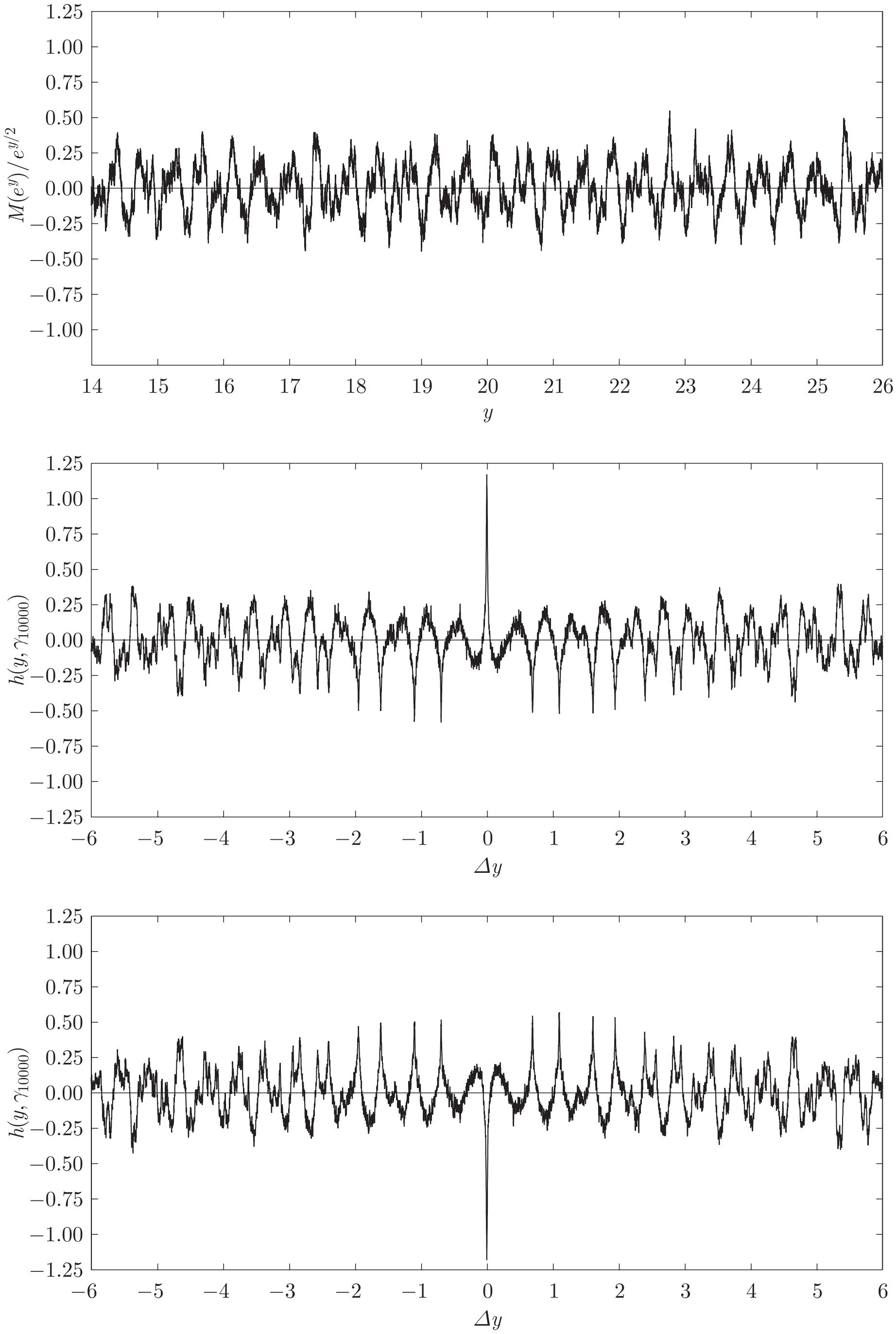

FIGURE 2

in one direction is usually accompanied by several large spikes in the opposite direction. Notice also that the bottom graph, when inverted with respect to the horizontal axis, very much resembles the middle graph. This is explained by the fact that the two functions plotted there are sums of cosines of which the first 98 main terms are aligned in $\Delta y = 0$.

Part of the L^3–output was also used to reduce the upper bound for which the Mertens conjecture is known to be false (Section 3) and for the computations concerning the growth rate of $M(x)/\sqrt{x}$ (Section 4).

3 Reduction of the smallest known x for which $|M(x)/\sqrt{x}| > 1$

Two years after Odlyzko and te Riele disproved the Mertens conjecture, Pintz [14] published a theorem which gives an *explicit* upper bound for the smallest x for which the Mertens conjecture is false:

Theorem 3. *If there exists a* $y \in [e^7, e^{5\times 10^4}]$ *with*

$$h_P(y,T) := 2 \sum_{0<\gamma<T} e^{-1.5\times 10^{-6}\gamma^2} \frac{\cos(\gamma y - \psi_\gamma)}{|\rho\zeta'(\rho)|} > 1 + e^{-40} \qquad (8)$$

for $T = 1.4 \times 10^4$, *then* $|M(x)|/\sqrt{x} > 1$ *for some* $x < e^{v+\sqrt{v}}$.

For the number $y = y_0 \approx 3.2097 \times 10^{64}$ as given by Odlyzko and te Riele in their Table 3 (line $i = 21$) in [13], on request of Pintz, te Riele computed $h_P(y_0, T)$ and found the value -1.00223, which implies, by Pintz's Theorem, that the Mertens conjecture is false for some $x < \exp(3.21 \times 10^{64})$.

We have computed $h_P(y,T)$ for many *smaller* values of y, resulting from our application of the L^3 algorithm in Section 2.2, in order to attempt to further reduce the upper bound for the smallest x for which the Mertens conjecture is false. The smallest y for which we found a value of $|h_P(y,T)| > 1 + e^{-40}$ is:

$$y = 1\,5853191167\,3595000428\,9014722171\,6268116204.984802$$

with $h_P(y,T) = -1.00819$. This shows that there exists an

$$x < \exp(1.59 \times 10^{40})$$

for which the Mertens conjecture is false. It is very likely that there is still substantial room for improvement of this result. For example, in [8] it is suggested that the first violation of the Mertens conjecture should occur not too far from $x \approx \exp(5.15 \times 10^{23})$.

4 Estimation of the order of magnitude of $M(x)/\sqrt{x}$

4.1 Existing results and conjectures

The strongest unconditional results on the order of magnitude of $M(x)$ are of the general form $M(x) = o(x)$. Thus Walfisz [19] proved that

$$M(x) = O\left(x \exp\left(-A\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\right)\right) \quad \text{for some } A > 0$$

and Ford [2] has recently shown that we may take $A = 0.2098$. A proof of the Riemann hypothesis would strengthen this to

$$M(x) = O(x^{1/2+\varepsilon}) \quad \text{for every } \varepsilon > 0.$$

It is also known that

$$M(x) = \Omega_\pm(x^{1/2})$$

and from the disproof of the Mertens conjecture by Odlyzko and te Riele it follows that the multiplicative constant is larger than 1 into both the positive and the negative direction. Nonetheless, the question whether $M(x)/\sqrt{x}$ is unbounded remains open, although many experts suppose that this is the case, and some arguments in favor of this have been presented in Section 2.1. During the last decades, several conjectures on the order of magnitude of $M(x)/\sqrt{x}$ have been set forth. Good and Churchhouse [3], as well as Lévy in a comment to Saffari [17], have conjectured that

$$\limsup_{x\to\infty} \frac{|M(x)|}{\sqrt{x\log\log x}} = C \tag{9}$$

with $C = \frac{\sqrt{12}}{\pi} = 1.1026\ldots$ according to Good and Churchhouse, whereas $C = \frac{6\sqrt{2}}{\pi^2} = 0.8597\ldots$ according to Lévy. Conjectures of the type (9) seem questionable, however, both on theoretical grounds and because of very poor agreement with experimental observations (see e.g. [13, p. 140] and [8, pp. 479–480]). More recently, Ng [12], partly building on unpublished work by Gonek, conjectured that

$$\limsup_{x\to\infty} \frac{|M(x)|}{\sqrt{x}(\log\log\log x)^{5/4}} = B \tag{10}$$

for some $B > 0$, while Kotnik and van de Lune [8] observed experimentally that estimates of the largest positive and negative values of $M(x)/\sqrt{x}$ in the range $10^4 \le x \le 10^{10^{10}}$ are quite close to $\frac{1}{2}\sqrt{\log\log\log x}$ and $-\frac{1}{2}\sqrt{\log\log\log x}$, respectively. If this would also hold asymptotically, it would contradict both (9) and (10). In a somewhat more conservative spirit, Kotnik and van de Lune finally conjectured that

$$M(x)/\sqrt{x} = \Omega_\pm(\sqrt{\log\log\log x}),$$

which is weaker than both (9) and (10), as these correspond to $M(x)/\sqrt{x} = \Omega(\sqrt{\log\log x})$ and $M(x)/\sqrt{x} = \Omega((\log\log\log x)^{5/4})$, respectively.

4.2 New results

Theorem 1 suggests that sums of the type

$$h_1(y,T) := 2 \sum_{0<\gamma<T} \frac{\cos(\gamma y - \psi_\gamma)}{|\rho\zeta'(\rho)|}$$

should, as T increases, eventually converge to the respective values of $M(e^y)/\sqrt{e^y}$. As mentioned in Section 2.1, the essential problem of the estimation of $M(e^y)/\sqrt{e^y}$ by means of such sums is that their convergence can only be estimated empirically. Observing that the sums obtained with the first ten thousand ζ-zeros ($T = \gamma_{10^4} = 9877.782...$) and those obtained with the first million ζ-zeros ($T = \gamma_{10^6} = 600269.677...$) consistently differ by less than 1% (see e.g. Tables 3 and 4 in [8], as well as Figure 4 in the same source), Kotnik and van de Lune used these sums in estimating the largest positive and negative values of $M(x)/\sqrt{x}$ in the range $10^4 \leq x \leq 10^{10^{10}}$ (see previous subsection), and thereby came to the observations mentioned in the preceding subsection. In terms of y, the range they investigated is $9.210... \leq y \leq 23025850929.940...$, and the estimates were obtained by sampling the sums in y-increments of 0.0005, and then searching for the local extrema if an exceptionally large value was found. While this approach practically eliminated the possibility of missing an exceptionally large positive or negative value, due to uniform y-increments a significant extension of this approach (e.g., to the smallest values of y for which $|M(e^y)|/\sqrt{e^y}$ is known to exceed 1; see Section 3) would be impossible.

As the L^3 algorithm can be used to generate isolated y-values for which either $M(e^y)/\sqrt{e^y}$ or $-M(e^y)/\sqrt{e^y}$ is likely to be very large, this offered a possibility to considerably extend the estimation of the order of magnitude of $M(e^y)/\sqrt{e^y}$ by means of the sums $h_1(y,T)$.

In Figure 3, we extend the study of Kotnik and van de Lune [8] with the estimates $h_1(y,\gamma_{10^4})$ obtained in this manner. The hollow squares and circles give the increasingly large values of $M(e^y)/\sqrt{e^y}$ and $h_1(y,\gamma_{10^6})$, respectively, obtained in [8], and the solid circles give the values of $h_1(y,\gamma_{10^4})$ found by the L^3 algorithm. We observe that also with y up to $\approx 10^{110}$ (i.e., $\sqrt{\log\log y}$ up to ≈ 2.35), the estimates of the largest positive and negative values of $M(x)/\sqrt{x}$ are quite close to $\frac{1}{2}\sqrt{\log\log\log x}$ and $-\frac{1}{2}\sqrt{\log\log\log x}$, respectively. Nevertheless, at the very largest y-values the positive and negative estimates appear to be systematically somewhat above the first and somewhat below the second of these two functions, respectively. This suggests that a further extension, perhaps to $y \approx 10^{500}$, could provide some additional insight into these observations.

5 Conclusions and discussion

We have presented improvements of $\limsup M(x)/\sqrt{x}$ and $\liminf M(x)/\sqrt{x}$ with respect to the results obtained by Odlyzko and te Riele in 1985, and an improvement of the upper bound of the smallest x for which $|M(x)| > \sqrt{x}$ with respect to the result obtained by Pintz in 1987. The approach that led to our results

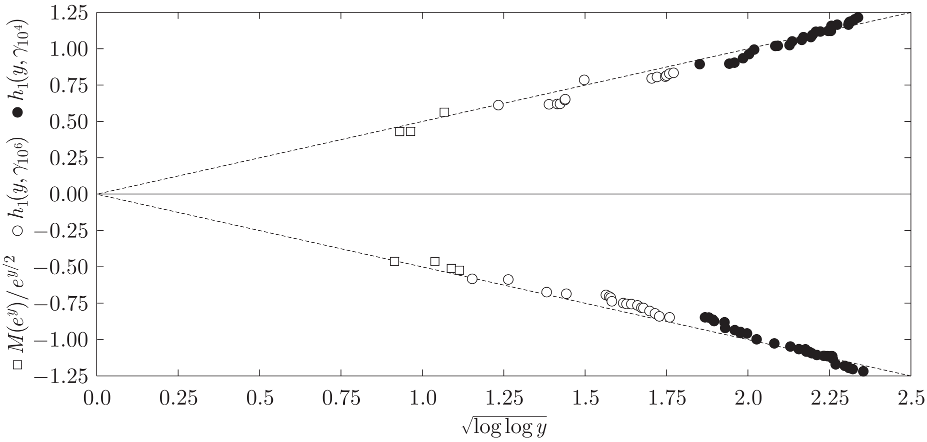

FIGURE 3

was based on systematic application of the L^3 algorithm to a relatively extensive set of combinations (ν, n), and also on considerable increase in computing power with respect to what was available to the researchers twenty years ago.

There are several directions in which our results could be improved further. Concerning the lim sup and lim inf results, the methods described in Section 2 could be extended in an elementary manner to larger values of ν and n, and as we state in that section, values of ν close to 800 would probably lead to lim sup and lim inf values close to 1.3 and -1.3, respectively. In contrast, a further improvement of the upper bound of the first violation of the Mertens conjecture would very likely require a strengthening of Theorem 3, in the sense of reducing the value of the constant 1.5×10^{-6} in the exponent occurring in (8). Namely, there are values of y smaller than the upper bound presented in Section 3 for which the sum in (8) only falls short of exceeding 1 by a small amount. An example is

$$y = 35499\,1618091406\,4844654619\,4090311725.687444$$

for which

$$h_P(y, 1.4 \times 10^4) = 0.991549.$$

Finally, as we discussed in Section 4, the largest positive and negative estimates of $M(x)/\sqrt{x}$ agree reasonably well with $\frac{1}{2}\sqrt{\log\log\log x}$ and $-\frac{1}{2}\sqrt{\log\log\log x}$, respectively, throughout the investigated range, but the estimates close to the very end of this range slightly, yet consistently exceed these two functions. An extension of the study presented here should clarify whether this behavior continues, and perhaps becomes more pronounced at even larger x.

Acknowledgements. We would like to thank Jan van de Lune for his valuable suggestions. The work of one of us (TK) has been performed under the Project HPC-Europa (RII3–CT–2003–506079), with the support of the European Community – Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme.

References

1. P.T. Bateman, J.W. Brown, R.S. Hall, K.E. Kloss and R.M. Stemmler. Linear relations connecting the imaginary parts of the zeros of the zeta function. In: A.O.L. Atkin and B.J. Birch (editors). *Computers in Number Theory*. New York, 1971, 11–19.
2. K. Ford. Vinogradov's integral and bounds for the Riemann zeta function. *Proc. Lond. Math. Soc.*, 85:565–633, 2002.
3. I.J. Good and R.F. Churchhouse. The Riemann hypothesis and pseudorandom features of the Möbius sequence. *Math. Comp.*, 22:857–861, 1968.
4. G.H. Hardy and E.M. Wright. *An introduction to the theory of numbers*. Oxford at the Clarendon Press, Fourth Edition, 1975.
5. A.E. Ingham. On two conjectures in the theory of numbers. *Amer. J. Math.*, 64:313–319, 1942.
6. W.B. Jurkat. On the Mertens conjecture and related general Ω-theorems. In H. Diamond, editor, *Analytic Number Theory*, pages 147–158, Providence, 1973.
7. W. Jurkat and A. Peyerimhoff. A constructive approach to Kronecker approximations and its application to the Mertens conjecture. *J. reine angew. Math.*, 286/287:322–340, 1976.
8. T. Kotnik and J. van de Lune. On the order of the Mertens function. *Exp. Math.*, 13:473–481, 2004.
9. A.K. Lenstra, H.W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.
10. Mathematica, version 5.0, Wolfram Research, Inc., Champaign, IL, 2003. `http://www.wolfram.com/`
11. F. Mertens. Über eine zahlentheoretische Funktion. *Sitzungsberichte Akad. Wiss. Wien IIa*, 106:761–830, 1897.
12. N. Ng. The distribution of the summatory function of the Möbius function. *Proc. Lond. Math. Soc.*, 89:361–389, 2004.
13. A.M. Odlyzko and H.J.J. te Riele. Disproof of the Mertens conjecture. *J. reine angew. Math.*, 357:138–160, 1985.
14. J. Pintz. An effective disproof of the Mertens conjecture. *Astérisque*, 147–148:325–333, 1987.
15. PARI/GP, version `2.2.11 (alpha)`, Bordeaux, 2005. `http://pari.math.u-bordeaux.fr/`
16. H.J.J. te Riele. Computations concerning the conjecture of Mertens. *J. reine angew. Math.*, 311/312:356–360, 1979.
17. B. Saffari. Sur la fausseté de la conjecture de Mertens. Avec une observation par Paul Lévy. *C. R. Acad. Sci.*, 271(A):1097–1101, 1970.
18. E.C. Titchmarsh. *The Theory of the Riemann Zeta-Function*. Oxford, 1951.
19. A. Walfisz. *Weylsche Exponentialsummen in der neueren Zahlentheorie*. VEB Deutscher Verlag, 1963.