



# *The Mertens conjecture revisited*

Herman te Riele, CWI, Amsterdam

(joint work with Tadej Kotnik, Univ. of Ljubljana, Slovenia)

*7th Algorithmic Number Theory Symposium*

*Technische Universität Berlin, 23 – 28 July 2006*

# Introduction

The Möbius function  $\mu(n)$  is defined as  $\mu(1) = 1$ ,  $\mu(n) = (-1)^k$  if  $n$  is the product of  $k$  different primes, and  $\mu(n) = 0$  otherwise.

Then

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

is the difference between the number of squarefree positive integers  $n \leq x$  with an **even** number of prime factors and those with an **odd** number of prime factors.

The **Mertens conjecture states that  $|M(x)|/\sqrt{x} < 1$  for all  $x > 1$ .**

This – but also the weaker assumption  $|M(x)|/\sqrt{x} < C$  for all  $x > 1$  and some  $C > 1$  – would imply the truth of the Riemann hypothesis (RH).

In fact, it is known that **RH  $\iff \forall \epsilon > 0, \lim_{x \rightarrow \infty} M(x)/x^{\frac{1}{2} + \epsilon} = 0$ .**

## Introduction, 2

The Mertens conjecture was shown to be false by Odlyzko and Te Riele in 1985 with help of the lattice basis reduction ( $L^3$ ) algorithm of A.K. Lenstra, H.W. Lenstra, Jr., and L. Lovász (1982) for finding short vectors in lattices.

They proved the **existence** of some  $x$  for which  $M(x)/\sqrt{x} > 1.06$ , and of some other  $x$  for which  $M(x)/\sqrt{x} < -1.009$ .

In 1987, Pintz gave an **effective** disproof of the Mertens conjecture in the sense that he proved that  $|M(x)|/\sqrt{x} > 1$  for some  $x \leq \exp(3.21 \times 10^{64})$ .

Nowadays, it is generally believed that the function  $M(x)/\sqrt{x}$  is **unbounded**, both in the positive and in the negative direction.

Kotnik and Van de Lune, e.g., have conjectured that  $M(x)/\sqrt{x} = \Omega_{\pm}(\sqrt{\log \log \log x})$ .

## 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  $\gamma_1 = 14.1347\dots$  and

$$\gamma_j < \gamma_{j+1}, j = 1, 2, \dots$$

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  $\rho_j$  ordered according to

**non-increasing** values of  $\alpha_j$ , and denote them by  $\rho_j^* = \frac{1}{2} + i\gamma_j^*$

with the corresponding quantities  $\psi_j^*, \alpha_j^*, j = 1, 2, \dots$ .

For example, the first five  $\rho_j^*$ 's coincide with the first five  $\rho_j$ 's, but

$$\rho_6^* = \rho_7, \rho_7^* = \rho_{10}, \text{ and } \rho_8^* = \rho_6$$

(with  $\alpha_6^* = \alpha_7 = 0.0163\dots, \alpha_7^* = \alpha_{10} = 0.0141\dots$  and

$$\alpha_8^* = \alpha_6 = 0.0137\dots).$$

## The first ten $\gamma_j$ 's

$j$	$\gamma_j$	$\psi_j$	$\alpha_j$	$\gamma_j^*$
1	14.1347	1.6933	0.0891	14.1347
2	21.0220	1.3264	0.0418	21.0220
3	25.0109	1.8851	0.0291	25.0109
4	30.4249	1.0169	0.0252	30.4249
5	32.9351	2.1297	0.0220	32.9351
6	37.5862	1.2636	0.0137	40.9187
7	40.9187	1.3540	0.0164	49.7738
8	43.3271	2.2052	0.0126	37.5862
9	48.0052	0.7096	0.0133	48.0052
10	49.7738	2.0372	0.0142	43.3271

## *Direct approach*

Systematic computations of  $M(x)$  for all  $x \in [1, X]$  by Mertens and many others have not led to a disproof of the Mertens conjecture. For  $X = 10^{14}$ , Kotnik and Van de Lune found the largest **positive** value of  $M(x)/\sqrt{x}$  to be 0.571 for  $x = 7\,766\,842\,813$  and the largest **negative** value to be  $-0.525$  for  $x = 71\,578\,936\,427\,177$ .

## Another approach

Another approach is based on the following theorem of Titchmarsh:

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

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

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

On the Riemann hypothesis, we have  $\rho = \frac{1}{2} + i\gamma$ , giving:

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

Hence, as  $n$  increases, the sum in (2) will eventually converge to  $M(x)/\sqrt{x}$ , with error on the order of magnitude of  $1/\sqrt{x}$ . However, very little is known about the rate of this convergence, as the coefficients  $|\rho_j \zeta'(\rho_j)|^{-1}$  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 by Kotnik and Van de Lune: several thousands of terms generally suffice to bring the error below 1%, but for much larger  $x$  this approach is not feasible.

## Ingham's tric

The tric of Ingham was to consider, instead of (2), a **weighthed average of the function**  $M(x)/\sqrt{x}$ . In that case the terms of the sum in (2) are multiplied by a function of bounded support, and the series in (1) is transformed into a finite sum. Two such cases will appear in what follows.

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

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

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

## Ingham's tric, 2

Then we have the following

**Theorem 2** *Let*

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

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 real  $y_0$**  we have

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

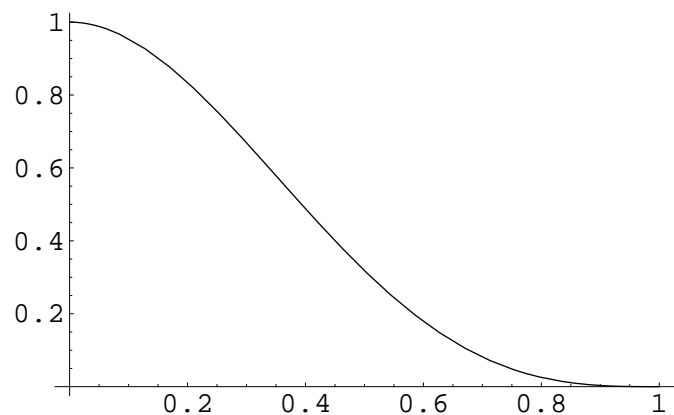
and **any value  $h(y, T)$  is approximated arbitrarily closely, and infinitely often, by  $M(x)/\sqrt{x}$ .**

Notice that also **negative** values of  $y_0$  are allowed.



# Graph of $(1 - t) \cos(\pi t) + \pi^{-1} \sin(\pi t)$

```
Plot[(1 - t) * Cos[Pi * t] + Sin[Pi * t] / Pi, {t, 0, 1}]
```



# *An inhomogeneous Diophantine approximation problem*

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, the sum of the *coefficients* of  $\cos(\gamma y - \psi_{\gamma})$  in Theorem 2 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  $2\pi$** , then we could make  $h(y, T)$  arbitrarily large. This would contradict, by Theorem 2, the conjecture of Mertens or any weaker form given above.

## An inhomogeneous ..., 2

If the  $\gamma$ 's were linearly independent over the rationals, then by Kronecker's theorem 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  $m_\gamma$ . 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  $\gamma$ 's there should exist any linear dependencies over the rationals.

# *The lattice basis reduction algorithm*

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 (A.K.), Lenstra (H.W., Jr.) and Lovász for finding short vectors in lattices.**

With this algorithm, the inhomogeneous Diophantine approximation problem could be solved for a much larger number of terms than before the time that  $L^3$  was known. The **“prize” to pay** was that any value of  $y$  that would come out was quite large. Therefore, the first 2000  $\gamma$ 's were computed with an accuracy of about 100 decimal digits.

The best lower and upper bounds found in 1985 for  $\overline{m}$  and  $\underline{m}$  were **1.06** and **-1.009**, respectively.

## How is $L^3$ applied?

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

$$(\gamma_j^* y - \psi_j^*) \bmod 2\pi, \quad 1 \leq j \leq n, \quad (3)$$

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

# The input matrix for $L^3$

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

where  $\nu$  is an integer satisfying  $2n \leq \nu \leq 4n$ .

The  $L^3$  algorithm produces a reduced basis  $\underline{v}'_1, \underline{v}'_2, \dots, \underline{v}'_{n+2}$  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  $1 \leq j \leq n$ :

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

and the  $(n+2)$ -nd coordinate is  $z$ , for some integers  $z, m_1, m_2, \dots, m_n$ . 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  $\alpha_j^*$  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 approximately be equal to:

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

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

# Example

$$n = 4, \nu = 8, L = \begin{bmatrix} -129 & 1 & 480 & 0 & 0 & 0 \\ -69 & 1 & 0 & 328 & 0 & 0 \\ -82 & 1 & 0 & 0 & 274 & 0 \\ -41 & 1 & 0 & 0 & 0 & 255 \\ 65536 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L' = \begin{bmatrix} 1 & -51 & -55 & -66 & -213 & -65 \\ 1 & -51 & -55 & -262 & 61 & -5 \\ 1 & -51 & 219 & -66 & 7 & -18 \\ 1 & 204 & -55 & -66 & -12 & 23 \\ 0 & 0 & 0 & 0 & 0 & 65536 \\ 1 & -51 & -55 & -66 & -267 & 64 \end{bmatrix}$$

$$z = 64, y = z/1024 = 0.0625$$

## Example, cont.

norms of vectors of  $L$ :

2.236, 255, 274, 328, 480, 65536.227

product= $1.6 \times 10^{15}$

norms of vectors of  $L'$ :

2.236, 228.079, 245.073, 293.373, 347.235, 65536.070

product= $8.3 \times 10^{14}$

$j$	$\alpha_j^*$	$\cos(\gamma_j^* y - \psi_j^*)$	$\alpha_j^* \cos(\gamma_j^* y - \psi_j^*)$
1	0.0891	0.6896	0.0614
2	0.0418	0.9999	0.0418
3	0.0291	0.9486	0.0276
4	0.0252	0.6336	0.0160
<b>sum:</b>	<b>0.1852</b>		<b>0.1468</b>

## Application of $L^3$

We have applied the  $L^3$  algorithm with the matrix (4) as input, for all the combinations  $(\nu, n)$  in the range  $\nu = 8, 9, \dots, 400$ ,  $n = \lceil \nu/4 \rceil, \lceil \nu/4 \rceil + 1, \dots, \lceil \nu/2 \rceil$ . To this end we used the function *qflll* from the PARI/GP package. For a given  $\nu$ , the precision by which the computations were carried out was chosen to be  $\log_{10}(2^{2\nu})$  decimal digits. For each combination of  $\nu$  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 above with  $y$  in the neighborhood of  $z/1024$ , and  $T = \gamma_{10000}$ .

The  $\gamma_j$ 's were computed to an accuracy of about 250 decimal digits using the Mathematica package, and, as a check, using the PARI/GP package. The computing time was about 600 CPU hours on the SGI Altix 3700 Aster system of the Academic Computing Centre Amsterdam (SARA).

# Scatter plot of the large positive values of $h$

Figure 1 gives for each  $\nu = 8, 9, \dots, 400$  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  $\nu$ , 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  $\nu$  in the neighborhood of 800.

*Large positive values of  $h$ ,  
hence of  $M(x)/\sqrt{x}$*

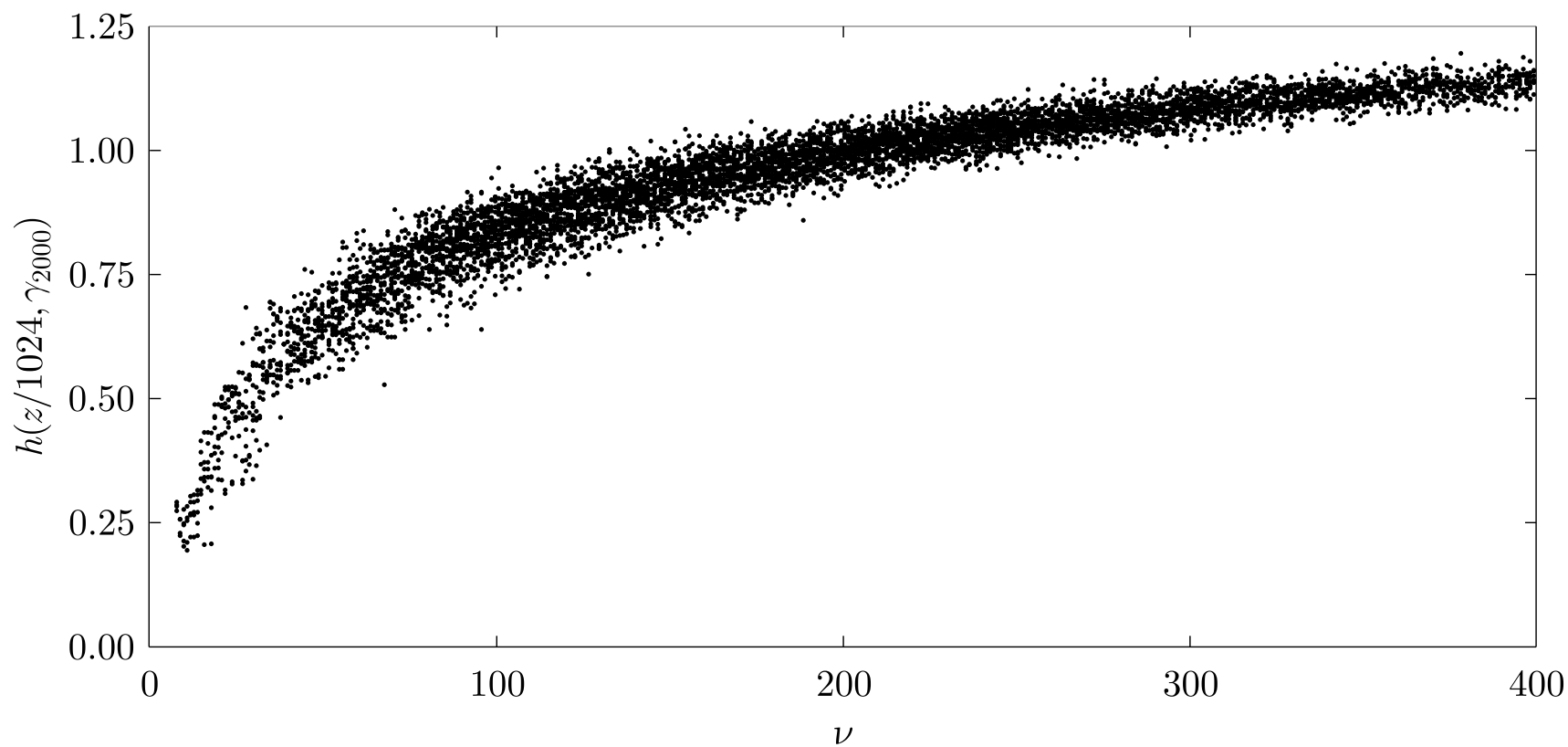


FIGURE 1

# Champions

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

$y = -233029271\ 5134531215\ 0140181996\ 7723401020\ 4456785091\ \backslash$   
 $6681557518\ 6743434036\ 9240230890\ 8933261706\ 9029233958\ \backslash$   
 $2730162362.807965$  ( $\log_{10} |y| = 108.3 \dots$ )

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

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

$y = -1608\ 7349754400\ 0919817483\ 9640165505\ 4685212472\ \backslash$   
 $2284778177\ 5539303027\ 5350690810\ 7957194829\ 6433602695\ \backslash$   
 $1442102295\ 3212754000.679958$  ( $\log_{10} |y| = 113.2 \dots$ )

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

# *Behaviour of $M(e^y)/e^{y/2}$ near the cham-* *pions*

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).

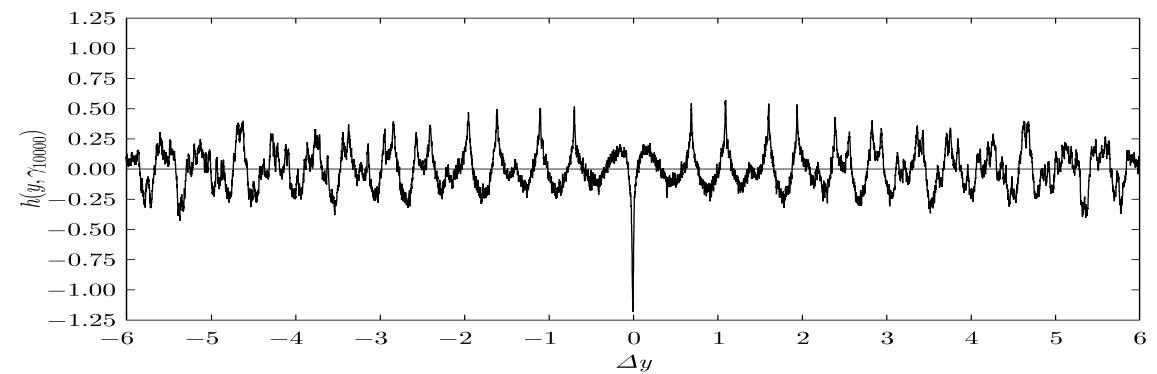
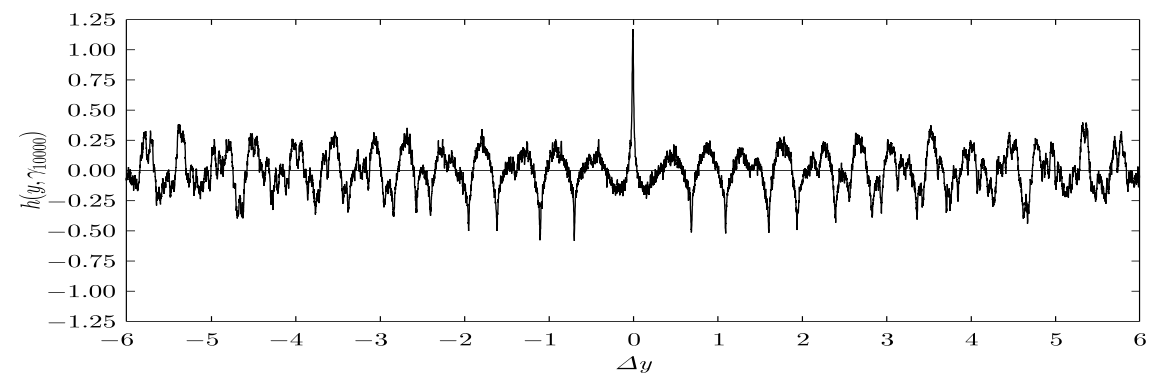
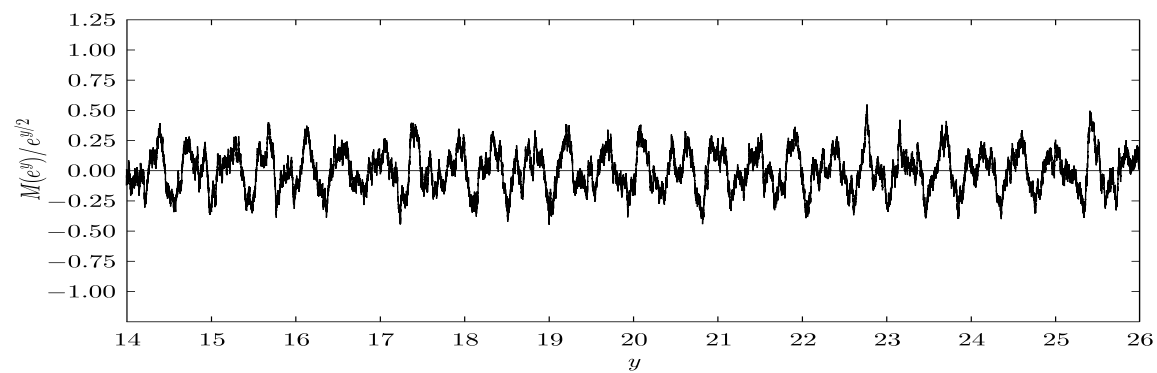


FIGURE 2

## Effective result of Pintz

**Theorem 3 (Pintz, 1987)** *Let*

$$h_P(y, T) := 2 \sum_{0 < \gamma < T} e^{-1.5 \times 10^{-6} \gamma^2} \frac{\cos(\gamma y - \psi_\gamma)}{|\rho \zeta'(\rho)|}. \quad (4)$$

*If there exists a  $y \in [e^7, e^{5 \times 10^4}]$  with  $|h_P(y, T)| > 1 + e^{-40}$  for  $T = 1.4 \times 10^4$ , then  $|M(x)|/\sqrt{x} > 1$  for some  $x < e^{y + \sqrt{y}}$ .*

For the number  $y = y_0 \approx 3.2097 \times 10^{64}$  as given by Odlyzko and Te Riele, 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  $x < \exp(3.21 \times 10^{64})$ .**

## *Improvement of Pintz's result*

We have computed  $h_P(y, T)$  for many *smaller* values of  $y$ , resulting from our application of the  $L^3$  algorithm above, in order to attempt to 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.**

# $\Omega_{\pm}$ -behaviour of $M(x)/\sqrt{x}$

In Figure 3, we extend the study of Kotnik and van de Lune with our results obtained here. 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, found by Kotnik and Van de Lune and the solid circles give the values of  $h_1(y, \gamma_{10^4})$  found by the  $L^3$  algorithm.

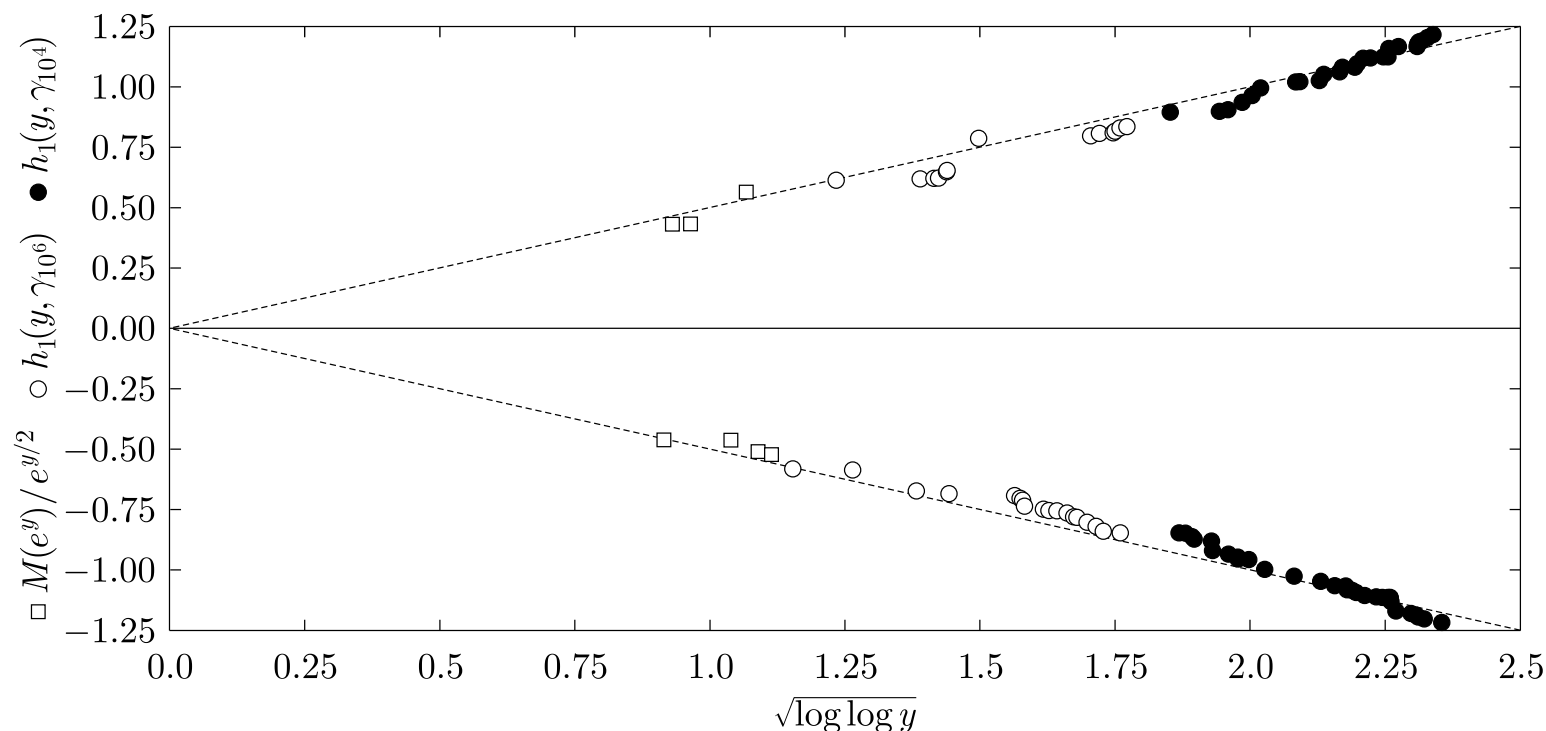


FIGURE 3

We observe that 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.

# Conclusions

- known lower bound 1.06 for  $\limsup M(x)/\sqrt{x}$  raised to **1.218**  
known upper bound  $-1.009$  for  $\liminf M(x)/\sqrt{x}$  lowered to  **$-1.229$**
- explicit upper bound  $\exp(3.21 \times 10^{64})$  of Pintz on the smallest number for which the Mertens conjecture is false, reduced to  **$\exp(1.59 \times 10^{40})$**
- numerical evidence for  $M(x)/\sqrt{x} = \Omega_{\pm}(\sqrt{\log \log \log x})$