Multidimensional Selberg theorem and fluctuations of the zeta zeros via Malliavin calculus

Ciprian A. Tudor

Multidimensional Selberg theorem and fluctuations of the zeta zeros via Malliavin calculus

Ciprian A. Tudor [1]
, [1] Laboratoire Paul Painlevé, Université de Lille 1
F-59655 Villeneuve d'Ascq, France.
tudor@math.univ-lille1.fr

January 11, 2016

Abstract

We give new contributions on the distribution of the zeros of the Riemann zeta function by using the techniques of the Malliavin calculus. In particular, we obtain the error bound in the multidimensional Selberg' s central limit theorem concerning the zeta zeros on the critical line and we discuss some consequences concerning the asymptotic behavior of the mesoscopic fluctuations of the zeta zeros.

2010 AMS Classification Numbers: 60F05, 60H05, 11M06.

Key words: Riemann zeta function, Selberg theorem, Stein's method, Malliavin calculus, central limit theorem, fluctuations of zeta zeros, Wasserstein distance.

1 On the Riemann zeta function and Selberg's theorem

The Riemann zeta function is usually defined, for $\mathrm{Re}\, s > 1$, as

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} \tag{1}$$

and for $\mathrm{Re}\, s \leq 1$, as an analytic continuation of (1). The Riemann zeta function is strongly related to the prime numbers theory via the Euler product formula

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

for $\mathrm{Re}\, s > 1$, where p ranges over primes. The distribution of the zeta zeros is one of the outstanding problems in mathematics. We know that $\zeta(-2n) = 0$ for every $n \geq 1$. The points $s = -2n$ are called the *trivial zeros* of the zeta function. It is known that the possible *non-trivial zeros* of the zeta-function could only lie inside the *critical strip* $0 < \mathrm{Re}\, s < 1$. They are of great interest since their distribution leads to many important results in prime numbers theory.

We also know that the numbers of zeta zeros is infinite and they lie symmetrical about the real axis and about the vertical line $\mathrm{Re}\, s = \frac{1}{2}$. *The Riemann hypothesis* posits that all the non-trivial zeros lies on the *critical line* $\mathrm{Re}\, s = \frac{1}{2}$.

A probabilistic way to analyze the zeta zeros is to look at the values of $\log \zeta(s)$ on the critical line $s = \frac{1}{2} + \mathbf{i}t$ and to consider t as a random variable uniformly distributed that takes huge values. That is, one considers $t \sim \mathcal{U}[T, 2T]$ with T close to infinity. By $\mathcal{U}[a, b]$ we will denote throughout this work the uniform distribution over the interval $[a, b], a < b$.

Selberg' s theorem (see [10], [11], [12] or the surveys [4], [14], [16]) gives the asymptotic distribution of $\log \zeta(s)$ on the critical line $\mathrm{Re}\, s = \frac{1}{2}$. Selberg theorem says that, if t is a random variable uniformly distributed over the interval $[T, 2T]$, then the sequence

$$\frac{\log \zeta\left(\frac{1}{2} + \mathbf{i}t\right)}{\sqrt{\frac{1}{2} \log \log T}} \tag{2}$$

converges in distribution to a complex-valued standard normal random variable $X_1 + \mathbf{i}X_2$ with $X_1, X_2 \sim N(0, 1)$ being independent random variables. There are several versions of this theorem. In particular, the result (2) holds if $t \sim U[0, T]$ of, more generally, if $t \sim U[aT, bT]$ with $b > a \geq 0$.

Selberg's theorem basically says that the zeta zeros does not affect too much the behavior of ζ on the critical line. Actually, the primes do most of the work. The very small normalization of order $\sqrt{\log \log T}$ is usually interpreted as a *repulsion of zeros* (see [14], [1], [2]).

Selberg theorem is actually equivalent to the convergence of the real and imaginary parts of $\log \zeta$ on the critical line, i.e. (recall that, if s is a complex number, then $\log s = \log |s| + \mathbf{i} \arg s$)

$$\frac{\log \left|\zeta\left(\frac{1}{2} + \mathbf{i}t\right)\right|}{\sqrt{\frac{1}{2} \log \log T}} \to^{(d)}_{T \to \infty} X_1 \sim N(0, 1) \tag{3}$$

and

$$\frac{\arg \log \zeta\left(\frac{1}{2} + \mathbf{i}t\right)}{\sqrt{\frac{1}{2} \log \log T}} \to^{(d)}_{T \to \infty} X_2 \sim N(0, 1) \tag{4}$$

where " $\to^{(d)}$ " stands for the convergence in distribution. The idea of the proof of (3) and (4) is (see [10], [11], [12], [14], [15]) to approximate $\log \zeta\left(\frac{1}{2} + \mathbf{i}t\right)$ by a Dirichlet series and to look to the behavior of this Dirichlet series, which can be easier handled. The Dirichlet approximation will be of the form

$$\sum_{p \leq T^{\varepsilon}} \frac{1}{p^{\frac{1}{2} + \mathbf{i}t}} = \sum_{p \leq T^{\varepsilon}} \frac{\cos(t \log p)}{\sqrt{p}} + \mathbf{i} \sum_{p \leq T^{\varepsilon}} \frac{\sin(t \log p)}{\sqrt{p}}. \tag{5}$$

with $t \sim \mathcal{U}[aT, bT]$, $b > a \geq 0$ and ε small enough. We will work throughout with $\varepsilon = 1$.

One of the main issues in our work is to study the speed of the convergence in (3) and (4). Some information concerning the rate of convergence to the normal distribution in the Selberg's central limit theorem can be found in the more recent Selberg's work [13] or in [16], [17]. Actually, it follows from [13] (see also Appendix A in [17] for a detailed proof) that (for ε sutably chosen in (5)) the Kolmogorov distance between the sequence (2) and the standard normal distribution is less than $C\frac{\log\log\log T}{\sqrt{\log\log T}}$ (throughout, by C we will denote a generic strictly positive constant that may change from one line to another). It is actually shown in [13], [17] that the Kolmogorov distance between the Dirichlet series (5) and the Gaussian law $N(0,1)$ is less $C\frac{1}{\sqrt{\log\log T}}$ and it is deduced that the distance between (2) and $N(0,1)$ is less than $C\frac{\log\log\log T}{\sqrt{\log\log T}}$. It seems that there are not results concerning other metrics. Therefore, in a first step, we will investigate the rate of convergence in the (one-dimensional) Selberg's theorem in terms of the Wasserstein distance.

We use recent techniques based on Malliavin calculus combined with Stein method (see [5]) in order to obtain our error bounds. Let us describe our new contributions. First, concerning the one-dimensional Selberg's theorem: we prove that the distance (under several metrics, such as the Kolmogorov, total variation, Wasserstein or Fortet -Mourier metrics) between the Dirichlet approximation (5) and the standard normal law if less than $C\frac{1}{\log\log T}$. This improves the known result for the Kolmogorov distance. We also prove that the rate of convergence of $\log\zeta(\frac{1}{2} + \mathbf{i}UT)$ to $N(0,1)$ is, under the Wasserstein metric, less than $C\frac{1}{\sqrt{\log\log T}}$ and this improves the result in [13], [17] (which states, recall, that the Kolmogorov distance between (2) and $N(0,1)$ is less than $C\frac{\log\log\log T}{\sqrt{\log\log T}}$). In our work, U denotes a standard uniform random variable i.e. $U \sim \mathcal{U}[0,1]$.

We also study the multidimensional context. The multidimensional extension of the Selberg's theorem has been proved more recently. First, in the paper [3], the authors showed that for any $0 < \lambda_1 < \ldots \lambda_d$, the random vector

$$X_T := \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(\log\zeta\left(\frac{1}{2} + \mathbf{i}P_i\right)\right)_{i=1,..,d} \tag{6}$$

with $P_i = Ue^{(\log T)^{\lambda_i}}$ $(i = 1, \ldots, d)$, converges in distribution, as $T \to \infty$ to $(\lambda_1 Y_1, \ldots, \lambda_d Y_d)$ where $Y_1, \ldots, Y_d$ are independent standard complex Gaussian random variables. There is no correlation between the components of the limit vector because the evaluation points P_i are rather distant one from each other. When these points are less distanced, then non-trivial correlations appear in the limit. The result is due to [1]. In this reference, the author showed that for $P_i = TU + f_T^{(i)}$ with $f_T^{(i)} - f_T^{(j)}$ not too big (the exact meaning is given later), then the random vector (6) converges in law to a d-dimensional complex Gaussian vector with dependent components.

We will regard these results from the Malliavin calculus point of view and we will give the associated error bounds. We will treat the case when $P_i = TU + iT$ (here the space between points is pretty big and the limit is a Gaussian vector with independent components) and $P_i = TU + f_T^{(i)}$ with $f_T^{(i)} - f_T^{(j)}$ small if $i \neq j$ (non-trivial correlations appear). We will see that the order of the speed of convergence to the limit distribution is not affected by the distance between the evalution points.

All these results lead to several consequences for the fluctuation of the number of zeros of the zeta function on the critical line. More precisely, we prove that the number of zeta zeros on the critical line $\mathrm{Re}\, s = \frac{1}{2}$ between some random heights satisfies a central limit theorem and we obtain the associated error bound. Our results extend the findings in [1], [2] or [3].

Our paper is organized as follows. In Section 2 we analyze the speed of convergence in the classical Selberg's theorem under several metrics via the Stein' s method combined with Malliavin calculus. In particular, we obtain explicit formulas for the Malliavin operators applied to the random variables in the left-hand side of (3), (4). In Section 3 we make the same study in the multidimensional settings, while in Section 4 we apply our findings to prove new results concerning the number of zeros of the Riemann zeta function. In the Appendix we included some elements from the Malliavin calculus and from the prime numbers theory needed in our work.

Throughout the paper we fix H a real and separable Hilbert space and $(W(h), h \in H)$ an isonomal Gaussian process (as introduced in Section 5.2) on the probability space $(\Omega, \mathcal{F}, P)$.

2 Rate of convergence in the Selberg theorem via Malliavin calculus

We here study the error bound corresponding to the weak convergences (3) and (4). The error bound will be obtained in two steps: first we measure the distance between the series (5) and the standard normal distribution and then we will use an old result in [12].

2.1 Rate of convergence for the Dirichlet series

Consider the family $(X_T)_{T>0}$ given by

$$X_T = \sum_{p \leq T} \left[\frac{\cos(TU \log p)}{\sqrt{p}} - \mathbf{E} \frac{\cos(TU \log p)}{\sqrt{p}} \right] \tag{7}$$

where the sum is taken over the primes p and U is $\mathcal{U}[0,1]$ distributed. In the sequel, we will assume

$$U = e^{-\frac{1}{2}(W(f)^2 + W(g)^2)} \tag{8}$$

with $f, g \in H$, $\|f\| = \|g\| = 1$ and $\langle f, g \rangle = 0$ (all the scalar products and norms in the paper will be considered in H if no further precision is made). In (8), W stands for a Gaussian isonormal process as described in the Appendix, Section 5.2. This implies that $W(f)$ and $W(g)$ are independent standard normal random variables.

The sequence X_T corresponds to the real part of $\log\zeta\left(\frac{1}{2}+\mathbf{i}TU\right)$. A similar analysis can be done for the imaginary part. We briefly describe the main steps.

We will measure the distance between the sequence

$$\frac{1}{\sqrt{\frac{1}{2}\log\log T}}X_T$$

and the standard normal distribution. For every $p\leq T$, denote by

$$X_{T,p}=\frac{\cos(TU\log p)}{\sqrt{p}}-\mathbf{E}\frac{\cos(TU\log p)}{\sqrt{p}} \tag{9}$$

so $X_T=\sum_{p\leq T}X_{T,p}$.

We need to compute $\langle DX_T, D(-L)^{-1}X_T\rangle$. This quantity is crucial when one uses the Stein method combined with the Malliavin calculus (see [5], see also (31) in Theorem 1). Let us do this computation. We can write

$$\begin{aligned}
&\langle DX_T, D(-L)^{-1}X_T\rangle\\
=&\sum_{p_1,p_2\leq T}\langle DX_{T,p_1}, D(-L)^{-1}X_{T,p_2}\rangle=\sum_{p_1,p_2\leq T}\langle DX_{T,p_2}, D(-L)^{-1}X_{T,p_1}\rangle\\
=&\frac{1}{2}\sum_{p_1,p_2\leq T}\left(\langle DX_{T,p_1}, D(-L)^{-1}X_{T,p_2}\rangle+\langle DX_{T,p_2}, D(-L)^{-1}X_{T,p_1}\rangle\right).
\end{aligned}\tag{10}$$

Using the series expansion of the cosinus function $\cos x=\sum_{k\geq 0}\frac{(-1)^k}{(2k)!}x^{2k}$ we get

$$X_{T,p}=\frac{1}{\sqrt{p}}\sum_{k\geq 0}\frac{(-1)^k}{(2k)!}(T\log p)^{2k}(U^{2k}-\mathbf{E}U^{2k}).$$

So

$$\begin{aligned}
&\langle DX_{T,p_1}, D(-L)^{-1}X_{T,p_2}\rangle+\langle DX_{T,p_2}, D(-L)^{-1}X_{T,p_1}\rangle\\
&=\sum_{k,l\geq 0}\frac{(-1)^{k+l}}{(2k)!(2l)!}T^{2k+2l}\frac{(\log p_1)^{2k}}{\sqrt{p_1}}\frac{(\log p_2)^{2l}}{\sqrt{p_2}}\\
&\left[\langle D(U^{2k}-\mathbf{E}U^{2k}), D(-L)^{-1}(U^{2l}-\mathbf{E}U^{2l})\rangle+\langle D(U^{2l}-\mathbf{E}U^{2l}), D(-L)^{-1}(U^{2k}-\mathbf{E}U^{2k})\rangle\right].
\end{aligned}$$

Let us use the notation, for every $k>0$

$$G_k:=U^k-\mathbf{E}U^k=e^{-\frac{k}{2}(W(f)^2+W(g)^2)}-\mathbf{E}e^{-\frac{k}{2}(W(f)^2+W(g)^2)}. \tag{11}$$

Relation (10) becomes

$$
\begin{aligned}
\langle DX_T, D(-L)^{-1}X_T\rangle &= \sum_{p_1,p_2\le T}\sum_{k,l\ge 0}\frac{(-1)^{k+l}}{(2k)!(2l)!}T^{2k+2l}\frac{(\log p_1)^{2k}}{\sqrt{p_1}}\frac{(\log p_2)^{2l}}{\sqrt{p_2}}\langle DG_k, D(-L)^{-1}G_l\rangle \\
&= \frac{1}{2}\sum_{p_1,p_2\le T}\sum_{k,l\ge 0}\frac{(-1)^{k+l}}{(2k)!(2l)!}T^{2k+2l}\frac{(\log p_1)^{2k}}{\sqrt{p_1}}\frac{(\log p_2)^{2l}}{\sqrt{p_2}} \\
&\quad \left[\langle DG_{2k}, D(-L)^{-1}G_{2l}\rangle + \langle DG_{2l}, D(-L)^{-1}G_{2k}\rangle\right].
\end{aligned}
\tag{12}
$$

where we used the symmetry of the sums. Consequently, it is necessary to calculate $\langle DG_{2k}, D(-L)^{-1}G_{2l}\rangle + \langle DG_{2l}, D(-L)^{-1}G_{2k}\rangle$. This will be done in the next lemma and it will be used several times in the paper.

Lemma 1 *Let G_k be given by (11). Then for every $k,l>0$*

$$
\langle DG_{2k}, D(-L)^{-1}G_{2l}\rangle + \langle DG_{2l}, D(-L)^{-1}G_{2k}\rangle = U^{2k}\frac{2k}{2l+1}\left(1-U^{2l}\right)+U^{2l}\frac{2l}{2k+1}\left(1-U^{2k}\right)
\tag{13}
$$

with U given by (8).

Proof: For every two smooth centered random variables F,G we have

$$
\begin{aligned}
&\langle DF, D(-L)^{-1}G\rangle + \langle DG, D(-L)^{-1}F\rangle \\
&= \left(\langle D(F+G), D(-L)^{-1}(F+G)\rangle - \langle DF, D(-L)^{-1}F\rangle - \langle DG, D(-L)^{-1}G\rangle\right).
\end{aligned}
\tag{14}
$$

We use the following formula proved in [8]: if $Y=f(N)-\mathbf{E}[f(N)]$ where $f\in C_b^1(\mathbb{R}^n;\mathbb{R})$ with bounded derivatives and $N=(N_1,...,N_n)$ is a Gaussian vector with zero mean and covariance matrix $K=(K_{i,j})_{i,j=1,..,n}$ then

$$
\langle D(-L)^{-1}(Y-\mathbf{E}[Y]), DY\rangle_H = \int_0^1 da\mathbf{E}'\left[\sum_{i,j=1}^{n}K_{i,j}\frac{\partial f}{\partial x_i}(N)\frac{\partial f}{\partial x_j}(aN+\sqrt{1-a^2}N')\right].
\tag{15}
$$

Here N' denotes an independent copy of N, the variables N and N' are defined on a product probability space $(\Omega\times\Omega', \mathcal{F}\otimes\mathcal{F}, P\times P')$ and $\mathbf{E}'$ denotes the expectation with respect to the probability measure P'.

In our case, for every $k\ge 1$,

$$
U^{2k}-\mathbf{E}U^{2k} = G_{2k} = h(W(f),W(g)) - \mathbf{E}h(W(f),W(g))
$$

with $h(x,y)=e^{-k(x^2+y^2)}$. Denote by

$$
G_{k,a} = e^{-k\left[(aW(f)+\sqrt{1-a^2}W'(f))^2+(aW(g)+\sqrt{1-a^2}W'(g))^2\right]}.
\tag{16}
$$

Then, by (15) we find

$$
\begin{aligned}
&\langle DG_{2k}, D(-L)^{-1}G_{2k}\rangle \\
= \quad & \int_0^1 da(2k)^2\mathbf{E}'\left[U^{2k}G_{k,a}W(f)(a(W(f)+\sqrt{1-a^2}W'(f))\right] \\
& + \int_0^1 da(2k)^2\mathbf{E}'\left[U^{2k}G_{k,a}W(g)(a(W(g)+\sqrt{1-a^2}W'(g))\right] \\
= \quad & \int_0^1 da4k^2aU^{2k}\left(W(f)^2+W(g)^2\right)\mathbf{E}'(G_{k,a}) \\
& + \int_0^1 da4k^2\sqrt{1-a^2}U^{2k}\left[W(f)\mathbf{E}'(W'(f)G_{k,a})+W(g)\mathbf{E}'(W'(g)G_{k,a})\right]. \qquad (17)
\end{aligned}
$$

Let us first calculate $\mathbf{E}'(G_{k,a})$ with $G_{k,a}$ given by (16). We have

$$
\mathbf{E}'(G_{k,a}) = g(aW(f))g(aW(g)) \qquad (18)
$$

where $g(c)=\mathbf{E}e^{-k(c+\sqrt{1-a^2}Z)^2}$ with Z a standard normal random variable. By standard calculations

$$
\begin{aligned}
g(c) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-k(c+\sqrt{1-a^2}x)^2}e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}e^{-kc^2}\int_{\mathbb{R}} dxe^{-k(1-a^2)x^2}e^{-\frac{x^2}{2}}e^{-2kc\sqrt{1-a^2}x} \\
&= \frac{1}{\sqrt{2\pi}}e^{-kc^2}e^{\frac{2k^2c^2(1-a^2)}{1+2k(1-a^2)}}\int_{\mathbb{R}} e^{-\frac{1}{2}(1+2k(1-a^2))y^2}dy = \frac{1}{\sqrt{1+2k(1-a^2)}}e^{-\frac{kc^2}{1+2k(1-a^2)}}.
\end{aligned}
$$

By (18), we obtain

$$
\mathbf{E}'(G_{k,a}) = \frac{1}{1+2k(1-a^2)}e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}. \qquad (19)
$$

We also need to compute $\mathbf{E}'(G_{k,a}W'(f)$. We have

$$
\mathbf{E}'(G_{k,a}W'(f)) = \mathbf{E}'\left[W'(f)e^{-k(aW(f)+\sqrt{1-a^2}W'(f))^2}\right]g(aW(g))
$$

where $g(c)$ has been computed just above. We will find that

$$
\mathbf{E}'\left[W'(f)e^{-k(aW(f)+\sqrt{1-a^2}W'(f))^2}\right] = m(aW(f))
$$

with $m(c)=\mathbf{E}'\left[Ze^{-k(c+\sqrt{1-a^2}Z)^2}\right]$. Moreover

$$
\begin{aligned}
m(c) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} xe^{-k(c+\sqrt{1-a^2}x)^2}e^{-\frac{x^2}{2}}dx \\
&= e^{-\frac{kc^2}{1+2k(1-a^2)}}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}(y-\frac{2kc\sqrt{1-a^2}}{1+2k(1-a^2)})e^{-\frac{1}{2}(1+2k(1-a^2))y^2}dy \\
&= -e^{-\frac{kc^2}{1+2k(1-a^2)}}\frac{2kc\sqrt{1-a^2}}{(1+2k(1-a^2))^{\frac{3}{2}}}.
\end{aligned}
$$

Hence

$$\mathbf{E}'(G_{k,a}W'(f)) = -e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}\frac{2kaW(f)\sqrt{1-a^2}}{(1+2k(1-a^2))^2} \tag{20}$$

and similarly,

$$\mathbf{E}'(G_{k,a}W'(g)) = -e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}\frac{2kaW(g)\sqrt{1-a^2}}{(1+2k(1-a^2))^2}. \tag{21}$$

Now, formula (17) becomes

$$\begin{aligned}
& \langle DG_{2k}, D(-L)^{-1}G_{2k}\rangle \\
= & \int_0^1 da4k^2a\frac{1}{1+2k(1-a^2)}U^{2k}(W(f)^2+W(g)^2)e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}} \\
& -\int_0^1 da4k^2(1-a^2)\frac{2ka}{(1+2k(1-a^2))^2}U^{2k}(W(f)^2+W(g)^2)e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}.
\end{aligned}$$

Let us denote by

$$S = W(f)^2+W(g)^2. \tag{22}$$

Since

$$a\frac{1}{1+2k(1-a^2)} - (1-a^2)\frac{2ka}{(1+2k(1-a^2))^2} = \frac{a}{(1+2k(1-a^2))^2} \tag{23}$$

we can write

$$\langle DG_{2k}, D(-L)^{-1}G_{2k}\rangle \quad = \quad 4k^2U^{2k}S\int_0^1 da\frac{a}{(1+2k(1-a^2))^2}e^{-\frac{ka^2S}{1+2k(1-a^2)}}.$$

Note that

$$\left(\frac{a^2}{(1+2k(1-a^2))^2}\right)' = \frac{2a(1+2k)}{(1+2k(1-a^2))^2}. \tag{24}$$

Thus, by the change of variables $\frac{a^2}{(1+2k(1-a^2))^2} = z$

$$\begin{aligned}
\langle DG_{2k}, D(-L)^{-1}G_{2k}\rangle &= 4k^2U^{2k}S\frac{1}{2(2k+1)}\int_0^1 dze^{-kSz} \\
&= 4k^2U^{2k}S\frac{1}{2(2k+1)}\int_0^1 dze^{-kSz} = 2kU^{2k}\frac{1}{2k+1}\left(1-e^{-kS}\right).
\end{aligned}$$

We can compute

$$\langle D(G_{2k}+G_{2l}), D(-L)^{-1}(G_{2k}+G_{2l})\rangle$$

by using again (15) with $h(x,y)=e^{-k(x^2+y^2)}+e^{-l(x^2+y^2)}$. We will have

$$
\begin{aligned}
&\langle D(G_{2k}+G_{2l}), D(-L)^{-1}(G_{2k}+G_{2l})\rangle\\
=\ & \int_0^1 daa(2kU^{2k}+2lU^{2l})(W(f)^2+W(g)^2)\mathbf{E}'(2kG_{k,a}+2lG_{l,a})\\
&+\int_0^1 da\sqrt{1-a^2}(2kU^{2k}+2lU^{2l})\\
&\left[W(f)\mathbf{E}'W'(f)(2kG_{k,a}+2lG_{l,a})+W(g)\mathbf{E}'W'(g)(2kG_{k,a}+2lG_{l,a})\right]
\end{aligned}
$$

and this implies

$$
\begin{aligned}
&\langle D(G_{2k}+G_{2l}), D(-L)^{-1}(G_{2k}+G_{2l})\rangle-\langle DG_{2k}, D(-L)^{-1}G_{2k}\rangle-\langle DG_{2l}, D(-L)^{-1}G_{2l}\rangle\\
=\ & \int_0^1 daa4kl(W(f)^2+W(g)^2)\left[U^{2k}\mathbf{E}'G_{l,a}+U^{2l}\mathbf{E}'(G_{k,a}\right]\\
&+\int_0^1 da\sqrt{1-a^2}4kl\left[U^{2k}W(f)\mathbf{E}'W'(f)G_{l,a}+U^{2l}W(f)\mathbf{E}'W'(f)G_{k,a}\right.\\
&\left.+U^{2k}W(g)\mathbf{E}'W'(g)G_{l,a}+U^{2l}W(g)\mathbf{E}'W'(g)G_{k,a}\right].
\end{aligned}
$$

Consequently, from relations (19), (20), and (21), we obtain

$$
\begin{aligned}
&\langle D(G_k+G_l), D(-L)^{-1}(G_k+G_l)\rangle-\langle DG_k, D(-L)^{-1}G_k\rangle-\langle DG_l, D(-L)^{-1}G_l\rangle\\
=\ & \int_0^1 daa4kl(W(f)^2+W(g)^2)\\
&\left[\frac{1}{1+k(1-a^2)}G_l e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}+\frac{1}{1+2l(1-a^2)}G_k e^{-\frac{la^2(W(f)^2+W(g)^2)}{1+2l(1-a^2)}}\right]\\
&-\int_0^1 da\sqrt{1-a^2}4kl(W(f)^2+W(g)^2)\\
&\left[G_l e^{-\frac{ka^2(W(f)^2+W(g)^2)}{1+2k(1-a^2)}}\frac{2ka\sqrt{1-a^2}}{(1+2k(1-a^2))^2}+G_k e^{-\frac{la^2(W(f)^2+W(g)^2)}{1+2l(1-a^2)}}\frac{2la\sqrt{1-a^2}}{(1+2l(1-a^2))^2}\right].
\end{aligned}
$$

To conclude (13), it suffices to use (14), (23) and (24). ■

We obtain the explicit form of the terms needed in the Stein-Malliavin bound (31).

Proposition 1 *For every $T>0$, with X_T given by (7), we have*

$$
\begin{aligned}
\langle DX_T, D(-L)^{-1}X_T\rangle &= \sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\\
&\times\left[\sin(TU\log p_2)\sin(TU\log p_1)-U\sin(TU\log p_2)\sin(T\log p_1)\right]. \qquad (25)
\end{aligned}
$$

Proof: By (12) and relation (13) in Lemma 1, we have

$$\begin{aligned}
\langle DX_T, D(-L)^{-1}X_T\rangle &= \frac{1}{2}\sum_{p_1,p_2\le T}\sum_{k,l\ge 0}\frac{(-1)^{k+l}}{(2k)!(2l)!}T^{2k+2l}\frac{(\log p_1)^{2k}}{\sqrt{p_1}}\frac{(\log p_2)^{2l}}{\sqrt{p_2}}\\
&\quad\times\left[2kU^{2k}\frac{1}{2l+1}\left(1-e^{-lS}\right)+2lU^{2l}\frac{1}{2k+1}\left(1-e^{-kS}\right)\right]\\
&= \sum_{p_1,p_2\le T}\sum_{k\ge 0;l\ge 1}\frac{(-1)^{k+l}}{(2k+1)!(2l-1)!}T^{2k+2l}\frac{(\log p_1)^{2k}}{\sqrt{p_1}}\frac{(\log p_2)^{2l}}{\sqrt{p_2}}U^{2l}\left(1-e^{-kS}\right).
\end{aligned}$$

So,

$$\begin{aligned}
\langle DX_T, D(-L)^{-1}X_T\rangle &= \sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\sum_{l\ge 1}\frac{(-1)^l}{(2l-1)!}(T\log p_2)^{2l}U^{2l}\\
&\quad\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p_1)^{2k}(1-U^{2k}). \qquad (26)
\end{aligned}$$

We first compute the sum over l. We have

$$\begin{aligned}
\sum_{l\ge 1}\frac{(-1)^l}{(2l-1)!}(T\log p_2)^{2l}G_l &= \sum_{l\ge 0}\frac{(-1)^{l+1}}{(2l+1)!}(T\log p_2)^{2l+2}U^{2l+2}\\
&= -TU\log p_2\sin(TU\log p_2). \qquad (27)
\end{aligned}$$

Concerning the sum over k

$$\begin{aligned}
&\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p_1)^{2k}(1-U^{2k}) = \sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p_1)^{2k}(1-U^{2k})\\
&= \frac{1}{T\log p_1}\left[\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p_1)^{2k+1}-U^{-1}\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)!}(T\log p_1)^{2k+1}U^{2k+1}\right]\\
&= \frac{1}{T\log p_1}\left(\sin(T\log p_1)-U^{-1}\sin(TU\log p_1)\right). \qquad (28)
\end{aligned}$$

By combining (26), (27) and (28), we get (25). ■

Using the same lines, we can treat the imaginary part of $\log\zeta(\frac{1}{2}+\mathbf{i}t)$, with $t=TU$.

Proposition 2 *Denote by*

$$Y_T=\sum_{p\le T}\left(\frac{\sin(TU\log p)}{\sqrt{p}}-\frac{\mathbf{E}\sin(TU\log p)}{\sqrt{p}}\right). \qquad (29)$$

Then

$$\langle DY_T, D(-L)^{-1}Y_T\rangle = \sum_{p_1,p_2\leq T} \frac{\log p_1}{\log p_2}\frac{1}{\sqrt{p_1p_2}}\left[(U-1)\cos(TU\log p_2) - U\cos(TU\log p_2)\cos(T\log p_1) + \cos(TU\log p_1)\cos(TU\log p_2)\right].$$

Proof: As in the proof of Proposition 1,

$$\begin{aligned}
&\langle DY_T, D(-L)^{-1}Y_T\rangle \\
= \;& \frac{1}{2}\sum_{p_1,p_2\leq T}\frac{1}{\sqrt{p_1p_2}}\sum_{k,l\geq 0}\frac{(-1)^{k+1++1l}}{(2k+1)!(2l+1)!}T^{2k+1+2l+1}\frac{(\log p_1)^{2k+1}}{\sqrt{p_1}}\frac{(\log p_2)^{2l+1}}{\sqrt{p_2}} \\
& \left[\langle DG_{2k+1}, D(-L)^{-1}G_{2l+1}\rangle - \langle DG_{2l+1}, D(-L)^{-1}G_{2k+1}\rangle\right]
\end{aligned} \tag{30}$$

with G_{2k+1} from (11) and by Lemma 1 (by replacing k, l by $k+\frac{1}{2}, l+\frac{1}{2}$ respectively),

$$\begin{aligned}
&\left[\langle DG_{2k+1}, D(-L)^{-1}G_{2l+1}\rangle - \langle DG_{2l+1}, D(-L)^{-1}G_{2k+1}\rangle\right] \\
= \;& (2k+1)U^{2k+1}\frac{1}{2l+2}\left(1-U^{2l+1}\right) + (2l+1)U^{2l+1}\frac{1}{2k+2}\left(1-U^{2k+1}\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\langle DY_T, D(-L)^{-1}Y_T\rangle \\
= \;& \sum_{p_1,p_2\leq T}\frac{1}{\sqrt{p_1p_2}}\sum_{k\geq 0}\frac{(-1)^k}{(2k+2)!}(T\log p_1)^{2k+1}(1-U^{2k+1})\sum_{l\geq 0}\frac{(-1)^l}{(2l)!}(TU\log p_2)^{2l+1}.
\end{aligned}$$

To conclude, it remains to notice that

$$\sum_{l\geq 0}\frac{(-1)^l}{(2l)!}(TU\log p_2)^{2l+1} = TU\log p_2\cos(TU\log p_2)$$

and

$$\sum_{k\geq 0}\frac{(-1)^k}{(2k+2)!}(T\log p_1)^{2k+1}(1-U^{2k+1}) = \frac{1}{T\log p_1}\left(1-\cos(T\log p_1) + \frac{1}{U}(\cos(T\log p_1 U) - 1)\right).$$

■

Let us take a moment to introduce certain notion on the distance between probability distributions and to recall some links between these topic and Malliavin calculus. Let X, Y be two random variables. The distance between the law of X and the law of Y is usually defined by ($\mathcal{L}(F)$ denotes the law of F)

$$d(\mathcal{L}(X), \mathcal{L}(Y)) := \sup_{h\in\mathcal{H}}|\mathbf{E}h(X) - \mathbf{E}h(Y)|$$

where $\mathcal{H}$ is a suitable class of functions. For example, if $\mathcal{H}$ is the set of indicator functions $1_{(-\infty,z]}, z \in \mathbb{R}$ we obtain the Kolmogorov distance (for simplicity, we will always write $d(X,Y)$ instead of $d(\mathcal{L}(X), \mathcal{L}(Y))$)

$$d_K(X,Y) := d_K(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{z\in\mathbb{R}} |P(X \leq z) - P(Y \leq z)|.$$

If $\mathcal{H}$ is the set of 1_B with B a Borel set, one has the total variation distance

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{B\in\mathcal{B}(\mathbb{R})} |P(X \in B) - P(Y \in B)|$$

while for $\mathcal{H} = \{h; \|h\|_L \leq 1\}$ ($\|\cdot\|_L$ is the Lipschitz norm) one has the Wasserstein distance denoted d_W. We will focus in our work on these metrics. We will use the generic notation $d(X,Y)$ when our claim concerns all the metrics introduced above.

Let us recall the Stein bound for the normal approximation in terms of the Malliavin operators. See Section 5 in [5].

Theorem 1 *If F is a random variable in $\mathbb{D}^{1,4}$ with $\mathbf{E}F = 0$ and N is a standard normal random variable, then*

$$d(F,N) \leq C\mathbf{E}\left|1 - \langle DF, D(-L)^{-1}F\rangle\right|. \tag{31}$$

We have the following result.

Theorem 2 *For every $T > 0$, let X_T, Y_T be given by (7), (29) respectively. Denote by*

$$F_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} X_T, \quad G_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} Y_T. \tag{32}$$

Then for T large enough,

$$d(F_T,N) \leq C\frac{1}{\log\log T} \text{ and } d(G_T,N) \leq C\frac{1}{\log\log T}. \tag{33}$$

Proof: Clearly $\mathbf{E}F_T = 0$ and from relation (25) in Proposition 1

$$\begin{aligned}
&\langle DX_T, D(-L)^{-1}X_T\rangle \\
= & -\sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \frac{\log p_2}{\log p_1} U\sin(TU\log p_2)\left(\sin(T\log p_1) - U^{-1}\sin(TU\log p_1)\right) \\
= & -\sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \frac{\log p_2}{\log p_1} U\sin(TU\log p_2)\sin(T\log p_1) \\
& +\sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \frac{\log p_2}{\log p_1} \sin(TU\log p_2)\sin(TU\log p_1)
\end{aligned}$$

and by separating the diagonal and non-diagonal parts in the second sum above, and by using $\sin^2(x) = \frac{1-\cos(2x)}{2}$,

$$\begin{aligned}
&\langle DX_T, D(-L)^{-1}X_T\rangle \\
= &\sum_{p\le T}\frac{\sin^2(TU\log p)}{p} + \sum_{p_1,p_2\le T;p_1\neq p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin(TU\log p_2)\sin(TU\log p_1) \\
&-\sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}U\sin(TU\log p_2)\sin(T\log p_1) \\
= &\sum_{p\le T}\frac{1}{2p} - \sum_{p\le T}\frac{\cos(2TU\log p)}{2p} + \sum_{p_1,p_2\le T;p_1\neq p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin(TU\log p_2)\sin(TU\log p_1) \\
&-\sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}U\sin(TU\log p_2)\sin(T\log p_1).
\end{aligned}$$

Therefore,

$$\mathbf{E}\left|1-\langle DF_T, D(-L)^{-1}F_T\rangle\right| \quad\le\quad A_{1,T}+A_{2,T}+A_{3,T}+A_{4,T}+A_{5,T}$$

where

$$A_{1,T} = \left|1-\frac{1}{\frac{1}{2}\log\log T}\sum_{p\le T}\frac{1}{2p}\right|, A_{2,T} = \frac{1}{\frac{1}{2}\log\log T}\left|\mathbf{E}\sum_{p\le T}\frac{2\cos(2TU\log p)}{2p}\right|,$$

$$A_{3,T} = \frac{1}{\frac{1}{2}\log\log T}\left|\mathbf{E}\sum_{p_1,p_2\le T^{\varepsilon};p_1\neq p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin(TU\log p_2)\sin(TU\log p_1)\right|$$

and

$$A_{4,T} = \frac{1}{\frac{1}{2}\log\log T}\left|\mathbf{E}\sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}U\sin(TU\log p_2)\sin(T\log p_1)\right|.$$

Since for every $a\in\mathbb{R}$, one has $\mathbf{E}\cos(aTU\log p) = \frac{1}{aT\log p}\sin(aT\log p)$, we have

$$|\mathbf{E}\cos(aTU\log p)| \le C\frac{1}{T\log p}. \tag{34}$$

The inequality (34) gives immediately, via (68)

$$|A_{2,T}| \le C\frac{1}{T\log\log T}\sum_{p\le T}\frac{1}{p\log p} \le C\frac{1}{T\log\log T}\sum_{p\le T}\frac{1}{p} \le C\frac{1}{T}.$$

To bound $A_{3,T}$, we use $\sin(x)\sin(y)=\frac{1}{2}(\cos(x-y)-\cos(x-y))$ and, if $p_1\neq p_2$,

$$|\mathbf{E}\cos(aTU(\log p_1 \pm \log p_2)| \leq C\frac{1}{T} \tag{35}$$

to get

$$\begin{aligned}
|A_{3,T}| &\leq C\frac{1}{T\log\log T}\sum_{p_1\leq T}\frac{1}{\sqrt{p_1}\log p_1}\sum_{p_2\leq T}\frac{\log p_2}{\sqrt{p_2}} \leq C\frac{1}{T\log\log T}\sum_{p_1\leq T}\frac{1}{\sqrt{p_1}}\log T\sum_{p_2\leq T}\frac{1}{\sqrt{p_2}}\\
&\leq C\frac{1}{\log\log T\log T}\frac{1}{\log\log T}
\end{aligned}$$

where we used the estimate (67). Finally, to deal with the summand $A_{4,T}$, we majorize $|\sin(T\log p_1)|$ by 1, we use

$$|\mathbf{E}U\sin(TU\log p_2)| \leq C\frac{1}{T\log p_2} \tag{36}$$

and we will have, from (36) and (67)

$$|A_{4,T}| \leq C\frac{1}{T\log\log T}\sum_{p_1\leq T}\frac{1}{\sqrt{p_1}\log p_1}\sum_{p_2\leq T}\frac{1}{\sqrt{p_2}} \leq C\frac{1}{\log\log T(\log T)^2}.$$

The speed of the convergence will be given by the dominant term $A_{1,T}$. Actually, from (68),

$$|A_{1,T}| \leq C\frac{1}{\log\log T}.$$

Obviously, similar arguments apply to the sequence G_T from (32). ■

Remark 1 *Our inequalities (33) improve the bounds obtained in [13] (see also Appendix A in [17]) where it was proved that for large T*

$$d_K(F_T,N)\leq C\frac{1}{\sqrt{\log\log T}} \text{ and } d_K(G_T,N)\leq C\frac{1}{\sqrt{\log\log T}}$$

where d_K is the Kolmogorov distance.

2.2 Rate of convergence in the Selberg theorem

The real part of $\log\zeta(s)$ on the critical, where ζ is the Riemann zeta function (1) , can be approximated by the family $X_T+\mathbf{E}X_T$ where X_T is given by (7). More precisely, if t is a random variable uniformly distributed on $[0,T]$, then $\log\left|\zeta(\frac{1}{2}+\mathbf{i}t)\right|$ is "close" (we explain below what that means) to $\sum_{p\leq T}\frac{\cos(TU\log p)}{\sqrt{p}}$ with $U\sim\mathcal{U}[0,1]$. Since we have estimated is the previous paragraph the distance between X_T and the standard normal law, we will be able to measure how far is $\log\left|\zeta(\frac{1}{2}+\mathbf{i}t)\right|$ from the standard normal distribution. Actually, we have (with $t = TU \sim \mathcal{U}[0,T]$)

$$\begin{aligned}
\frac{\log|\zeta\left(\frac{1}{2}+\mathbf{i}t\right)|}{\sqrt{\frac{1}{2}\log\log T}} &= \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{\cos(TU\log p)-\mathbf{E}\cos(TU\log p)}{\sqrt{p}} \\
&+ \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\left(\log|\zeta\left(\frac{1}{2}+\mathbf{i}t\right)|-\sum_{p\leq T}\frac{\cos(TU\log p)}{\sqrt{p}}\right) \\
&+ \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{\mathbf{E}\cos(TU\log p)}{\sqrt{p}}.
\end{aligned}$$

As mentioned above, the second summand in the right side above is known to be "small". The exact meaning is described in the following result which has been obtained in [12]. The reader may also consult the survey [4] for the detailed steps of the proof.

Lemma 2 *For every $k, j \geq 1$ integers, it holds*

$$\mathbf{E}\left|\log|\zeta\left(\frac{1}{2}+\mathbf{i}T(U+j)\right)|-\sum_{p\leq T}\frac{\cos(T\log p(U+j))}{\sqrt{p}}\right|^{2k} = \mathcal{O}(1). \tag{37}$$

We do not need to assume the Riemann hypothesis in order to have the result in Lemma 2. We will use the Wasserstein distance (introduced in this section) to measure how far is $\log\left|\zeta(\frac{1}{2}+\mathbf{i}t)\right|$ from the standard normal distribution. From the definition of this metric, one can see that

$$d_W(F,G) \leq \mathbf{E}|F-G| \leq \left(\mathbf{E}\,|F-G|^2\right)^{\frac{1}{2}} \tag{38}$$

if F, G are two random variables in $L^2(\Omega)$. Using the triangular inequality for the Wasserstein distance, we write, with $t = TU$ and U as in (8)

$$
\begin{aligned}
& d_W\left(\frac{\log|\zeta\left(\frac{1}{2}+\mathbf{i}t\right)|}{\sqrt{\frac{1}{2}\log\log T}}, N(0,1)\right) \\
\leq\ & d_W\left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{\cos(TU\log p)-\mathbf{E}\cos(TU\log p)}{\sqrt{p}}, N(0,1)\right) \\
+\ & d_W\left(\frac{\log|\zeta\left(\frac{1}{2}+\mathbf{i}t\right)|}{\sqrt{\frac{1}{2}\log\log T}}, \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{\cos(TU\log p)-\mathbf{E}\cos(TU\log p)}{\sqrt{p}}\right) \\
\leq\ & d_W\left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\frac{\cos(TU\log p)-\mathbf{E}\cos(TU\log p)}{\sqrt{p}}\right) \\
+\ & \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\mathbf{E}\left|\log|\zeta\left(\frac{1}{2}+\mathbf{i}t\right)|-\sum_{p\leq T}\frac{\cos(TU\log p)}{\sqrt{p}}\right| \\
+\ & \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\sum_{p\leq T}\left|\frac{\mathbf{E}\cos(TU\log p)}{\sqrt{p}}\right| \\
:=\ & I_{1,T}+I_{2,T}+I_{3,T} \qquad (39)
\end{aligned}
$$

where we used (38). We estimate the three summands. The bound for $I_{1,T}$ has been obtained in Theorem 2. From this results, we have

$$I_{1,T}\leq C\frac{1}{\log\log T}.$$

The summand $I_{2,T}$ can be majorized by using Lemma 2 with $k=1$ and $j=0$. It holds that

$$I_{2,T}\leq C\frac{1}{\sqrt{\log\log T}}.$$

From (34), for T large enough, we clearly have

$$I_{3,T}\leq C\frac{1}{T\sqrt{\log\log T}}\sum_{p\leq T}\frac{1}{\sqrt{p}\log p}\leq C\frac{1}{\sqrt{T\log T\log\log T}}$$

due to (67). These estimates (and similar estimates for the imaginary part) leads to the following result.

Theorem 3 *With ζ, U as in (1), (8) respectively, and with T large enough,*

$$d_W\left(\frac{\log\left|\zeta\left(\frac{1}{2}+iTU\right)\right|}{\sqrt{\frac{1}{2}\log\log T}}, N\right)\leq C\frac{1}{\sqrt{\log\log T}}, d_W\left(\frac{\arg\log\zeta\left(\frac{1}{2}+iTU\right)}{\sqrt{\frac{1}{2}\log\log T}}, N\right)\leq C\frac{1}{\sqrt{\log\log T}} \qquad (40)$$

where $N\sim N(0,1)$.

Remark 2 *Theorem 3 improves the error bound obtained in [13] or [17]. In these references, the right-hand bound in (40) is* $C\frac{\log\log\log T}{\sqrt{\log\log T}}$ *under the Kolmogorov metric.*

3 Multidimensional Selberg theorem and the rate of convergence

In this paragraph we give a multidimensional extension of the Selberg central limit theorem. Concretely, we consider the $d+1$ dimensional random vector

$$\mathbf{V}_T := \left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(\log\zeta\left(\frac{1}{2} + \mathbf{i}P_i\right)\right)\right)_{i=0,..,d} \tag{41}$$

and we analyze the asymptotic distribution of its real and imaginary part. As mentionned in the introduction, it has been proved in [3] that for $P_i = Ue^{(\log T)^{\lambda_i}}$, $i = 0,..,d$ with $\lambda_0 < \ldots < \lambda_d$, the vector $\mathbf{V}_T$ (41) converges in law, as $T \to \infty$, to a $d+1$-dimensional standard complex Gaussian vector. When the space between the points P_i is small then the limit of $\mathbf{V}_T$ is a Gaussian vector with correlated components. The result is due to [1].

We analyze the error bound in the multidimensional Selberg theorem in both cases: when the distance between P_i and P_j is " big" or " small". In fact, we first consider the case $P_i = T(U+i)$, $i = 0,..,d$. This is a natural multi-dimensional extension of the celebrated Selberg' s result and it seems that it has not yet proved in the literature. Here $P_i - P_j = (i-j)T$ is big enough to avoid the correlation between the components of the limit. Then we treat the case considered in [1] when the evaluation points are less distant one from each other.

The basic idea is the same: one approximates $\mathbf{V}_T$ by a random vector whose components are Dirichlet series of the form (5). Using the techniques of the Malliavin calculus, we obtain the rate of convergence of this approximation to the Gaussian limit under the Wasserstein metric. Then we deduce a bound for the Wasserstein distance between $\mathbf{V}_T$ and the Gaussian limit.

3.1 Big shifts: convergence to a standard Gaussian vector

Let us first treat the case of "big shifts", i.e. the distance between the evaluation points is big enough and the limit distribution in the multidimensional Selbergh theorem is a standard Gaussian vector.

3.1.1 Error bound for the Dirichlet series

Consider the $d+1$ dimensional random vector

$$\mathbf{X}_T = (X_T^{(0)}, ..., X_T^{(d)})$$

where for every $i = 0, .., d$,

$$X_T^{(i)} = \sum_{p \le T} \frac{1}{\sqrt{p}} \left(\cos(T \log p t_i) - \mathbf{E} \cos(T \log p t_i) \right) \tag{42}$$

where $t_i \sim \mathcal{U}[Ti, T(i+1)]$. We analyze the asymptotic limit of X_T as $T \to \infty$. Since $\mathbf{X}_T$ is close to the real part of the random vector $\mathbf{V}_T$ (41), we will then deduce the asymptotic distribution and the error bound for $\mathbf{V}_T$. In order to use the techniques of the Malliavin calculus, we will assume that

$$t_i = T(U+i), \text{ for every } i = 0, ..., d$$

where U is given by (8). Clearly $t_i \sim \mathcal{U}[Ti, T(i+1)]$.

The exists a multidimensional version of the Stein-Malliavin inequality presented in Theorem 1, see [5], [7]. This bound is given in terms of the Wasserstein distance. Namely, if $F = (F_0, ..., F_d)$ is a random vector with components in $\mathbb{D}^{1,4}$ and $N(0, \Lambda)$ denotes the $d+1$ dimensional Gaussian distribution with covariance matrix $\Lambda = (c_{i,j})_{i,j=0,..,d}$, then

$$d_W\left(F, N(0, \Lambda)\right) \le C \sum_{i=0}^{d} \mathbf{E} \left| 2c_{i,j} - \langle DF_i, D(-L)^{-1} F_j \rangle - \langle DF_j, D(-L)^{-1} F_i \rangle \right|. \tag{43}$$

Actually, the bound presented in [5] or [7] is slightly different (and uses the L^2-norm on the right-hand side of the inequality (43))but it is easy to obtain (43) by similar arguments. Recall that the Wasserstein distance between the laws of two $\mathbb{R}^d$ - valued random variables F, G is defined by

$$d_W(F, G) = \sup_{h \in \mathcal{A}} |\mathbf{E} h(F) - \mathbf{E} h(G)| \tag{44}$$

where we denote by $\mathcal{A}$ the class of all functions $h : \mathbb{R}^d \to \mathbb{R}$ such that $\|h\|_{Lip} \le 1$, where $\|h\|_{Lip} = \sup_{x, y \in \mathbb{R}^d, x \ne y} \frac{|h(x) - h(y)|}{\|x - y\|}$, with the Euclidean norm $\| \cdot \|$ in $\mathbb{R}^d$.

In order to apply (43), we need to calculate

$$\langle DX_T^{(i)}, D(-L)^{-1} X_T^{(j)} \rangle + \langle DX_T^{(j)}, D(-L)^{-1} X_T^{(i)} \rangle$$

for every $i, j = 0, .., d$. Using the series expansion of the cosinus function, we have

$$X_T^{(i)} = \sum_{p \le T} \frac{1}{\sqrt{p}} \sum_{k \ge 0} \frac{(-1)^k}{(2k)!} (T \log p)^{2k} G_{2k}^{(i)}$$

with the notation, for $i = 0, .., d$ and for $k > 0$

$$G_{2k}^{(i)} = (U+i)^{2k} - \mathbf{E}(U+i)^{2k}. \tag{45}$$

Then

$$\langle DX_T^{(i)}, D(-L)^{-1}X_T^{(j)}\rangle + \langle DX_T^{(j)}, D(-L)^{-1}X_T^{(i)}\rangle = \sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \sum_{k,l\geq 0} \frac{(-1)^{k+l}}{(2k)!(2l)!}$$
$$\times (T\log p_1)^{2k}(T\log p_2)^{2l}\left[\langle DG_{2k}^{(i)}, D(-L)^{-1}G_{2l}^{(j)}\rangle + \langle DG_{2l}^{(j)}, D(-L)^{-1}G_{2k}^{(i)}\rangle\right]. \quad (46)$$

The next step is to calculate $\langle DG_{2k}^{(i)}, D(-L)^{-1}G_{2l}^{(j)}\rangle + DG_{2l}^{(j)}, D(-L)^{-1}G_{2k}^{(i)}\rangle$. This will be done in the following lemma, based on Lemma 1.

Lemma 3 *For every $k,l>0$ and $i,j=0,..,d$ we have*

$$\begin{aligned}
&\langle DG_{2k}^{(i)}, D(-L)^{-1}G_{2l}^{(j)}\rangle + \langle DG_{2l}^{(j)}, D(-L)^{-1}G_{2k}^{(i)}\rangle \\
=\ & \frac{2k}{2l+1}(U+i)^{2k-1}U\left[(j+1)^{2l+1} - j^{2l+1} - \frac{(U+j)^{2l+1}}{U} + \frac{j^{2l+1}}{U}\right] \\
&+\frac{2l}{2k+1}(U+j)^{2l-1}U\left[(i+1)^{2k+1} - i^{2k+1} - \frac{(U+i)^{2k+1}}{U} + \frac{i^{2k+1}}{U}\right] \quad (47)
\end{aligned}$$

where $G_{2k}^{(i)}$ is given by (45).

Proof: Using the Newton formula $G_{2k}^{(i)} = \sum_{s=0}^{2k} C_{2k}^s G_s i^{2k-s}$ and

$$\begin{aligned}
&\langle DG_{2k}^{(i)}, D(-L)^{-1}G_{2l}^{(j)}\rangle + DG_{2l}^{(j)}, D(-L)^{-1}G_{2k}^{(i)}\rangle \\
=\ & \sum_{s=0}^{2k}\sum_{t=0}^{2l} C_{2k}^s C_{2l}^t i^{2k-s} j^{2l-t}\left[\langle DG_s, D(-L)^{-1}G_t\rangle + \langle DG_t, D(-L)^{-1}G_s\rangle\right]
\end{aligned}$$

with G_s from (11). By Lemma 1,

$$\langle DG_s, D(-L)^{-1}G_t\rangle + \langle DG_t, D(-L)^{-1}G_s\rangle = \frac{s}{t+1}U^s(1-U^t) + \frac{t}{s+1}U^t(1-U^s).$$

Therefore,

$$\begin{aligned}
&\langle DG_{2k}^{(i)}, D(-L)^{-1}G_{2l}^{(j)}\rangle + DG_{2l}^{(j)}, D(-L)^{-1}G_{2k}^{(i)}\rangle \\
=\ & \sum_{s=0}^{2k}\sum_{t=0}^{2l} C_{2k}^s C_{2l}^t i^{2k-s} j^{2l-t}\frac{s}{t+1}U^s(1-U^t) + \frac{t}{s+1}U^t(1-U^s) \\
=\ & \sum_{s=0}^{2k} C_{2k}^s i^{2k-s} sU^s \sum_{t=0}^{2l} C_{2l}^t j^{2l-t}\frac{1-U^t}{t+1} + \sum_{s=0}^{2k} C_{2k}^s i^{2k-s}\frac{1-U^s}{s+1}\sum_{t=0}^{2l} C_{2l}^t j^{2l-t} tU^t. \quad (48)
\end{aligned}$$

Now we calculate the sums after s and t. Notice that

$$\sum_{s=0}^{2k} C_{2k}^{s} i^{2k-s} s U^{s} = 2k(U+i)^{2k-1}U$$

and

$$\sum_{t=0}^{2l} C_{2l}^{t} j^{2l-t}\frac{1-U^{t}}{t+1} = \frac{1}{2l+1}\left[(j+1)^{2l+1} - j^{2l+1} - \frac{(U+j)^{2l+1}}{U} + \frac{j^{2l+1}}{U}\right].$$

From the above two identities and (48), we deduce the conclusion (47). ■

Remark 3 *For $i=j=0$, we retrieve the result in Lemma 1.*

We are now in position to compute the terms involving Malliavin operators that appear in the right-hand side of (43).

Lemma 4 *For $i,j=0,..,d$, let $X_T^{(i)}$ be defined by (42). Then*

$$\langle DX_T^{(i)}, D(-L)^{-1}X_T^{(j)}\rangle + \langle DX_T^{(j)}, D(-L)^{-1}X_T^{(i)}\rangle = \sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}$$

$$[\sin((U+i)T\log p_1)\sin((U+j)T\log p_2) + \sin((U+j)T\log p_1)\sin((U+i)T\log p_2)] + R_T^{(i,j)}$$

with

$$\begin{aligned} R_T^{(i,j)} &= \sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1} \\ &\quad\Big[-\frac{U}{i+1}\sin((U+j)T\log p_2)\sin((i+1)T\log p_1) - \frac{U}{j+1}\sin((U+i)T\log p_2)\sin((j+1)T\log p_1) \\ &\quad + \frac{U}{i}\sin((U+j)T\log p_2)\sin(iT\log p_1)1_{i\neq 0} + \frac{U}{j}\sin((U+i)T\log p_2)\sin(jT\log p_1)1_{j\neq 0} \\ &\quad - \sin((U+j)T\log p_2)\sin(iT\log p_1) - \sin((U+i)T\log p_2)\sin(jT\log p_1)\Big]. \end{aligned}$$

Proof: By Lemma 2 and (46),

$$
\begin{aligned}
&\langle DX_T^{(i)}, D(-L)^{-1}X_T^{(j)}\rangle + \langle DX_T^{(j)}, D(-L)^{-1}X_T^{(i)}\rangle \\
= &\sum_{p_1,p_2\le T} \frac{1}{\sqrt{p_1p_2}} \sum_{k,l\ge 0} \frac{(-1)^{k+l}}{(2k)!(2l)!}(T\log p_1)^{2k}(T\log p_2)^{2l} \\
&\left(\frac{2k}{2l+1}(U+i)^{2k-1}U\left[(j+1)^{2l+1} - j^{2l+1} - \frac{(U+j)^{2l+1}}{U} + \frac{j^{2l+1}}{U}\right]\right. \\
&\left. + \frac{2l}{2k+1}(U+j)^{2l-1}U\left[(i+1)^{2k+1} - i^{2k+1} - \frac{(U+i)^{2k+1}}{U} + \frac{i^{2k+1}}{U}\right]\right) \\
= &\sum_{p_1,p_2\le T} \frac{1}{\sqrt{p_1p_2}} \sum_{k\ge 0} \frac{(-1)^{k}}{(2k+1)!}(T\log p_1)^{2k}\left[(i+1)^{2k+1} - i^{2k+1} - \frac{(U+i)^{2k+1}}{U} + \frac{i^{2k+1}}{U}\right] \\
&\times U\sum_{l\ge 1} \frac{(-1)^{l}}{(2l-1)!}(U+j)^{2l-1}(T\log p_2)^{2l} \\
+ &\sum_{p_1,p_2\le T} \frac{1}{\sqrt{p_1p_2}} \sum_{l\ge 0} \frac{(-1)^{l}}{(2l+1)!}(T\log p_1)^{2l}\left[(j+1)^{2l+1} - i^{2l+1} - \frac{(U+j)^{2l+1}}{U} + \frac{j^{2l+1}}{U}\right] \\
&\times U\sum_{k\ge 1} \frac{(-1)^{k}}{(2k-1)!}(U+i)^{2k-1}(T\log p_2)^{2k}. \qquad (49)
\end{aligned}
$$

Next, we calculate the above sums over k and l. We have

$$\sum_{l\ge 1} \frac{(-1)^{l}}{(2l-1)!}(U+j)^{2l-1}(T\log p_2)^{2l} = -UT\log p_2 \sin((U+j)T\log p_2) \qquad (50)$$

and

$$
\begin{aligned}
&\sum_{k\ge 0} \frac{(-1)^{k}}{(2k+1)!}(T\log p_1)^{2k}\left[(i+1)^{2k+1} - i^{2k+1} - \frac{(U+i)^{2k+1}}{U} + \frac{i^{2k+1}}{U}\right] \\
= &\frac{1}{T\log p_1}\left[\frac{1}{i+1}\sin(T(i+1)\log p_1))\right. \\
&\left. -\frac{1}{i}\sin(Ti\log p_1) - \frac{1}{U}\sin(T(U+i)\log p_1) + \frac{1}{U}\sin(Ti\log p_1)\right]. \qquad (51)
\end{aligned}
$$

By plugging relations (50) and (51) into (49), we get the conclusion. ■

We measure now the Wasserstein distance between the (renormalized) sequence $\mathbf{X}_T$ and the standard $d+1$ dimensional Gaussian distribution (denoted $N(0, I_{d+1})$ in the sequel).

Proposition 3 *Let $X_T^{(i)}$ be given by (42) for $i = 0, .., d$ and let Let*

$$\mathbf{F}_T = \sqrt{\frac{1}{\frac{1}{2}\log\log T}}\mathbf{X}_T = \sqrt{\frac{1}{\frac{1}{2}\log\log T}}(X_T^{(0)}, X_T^{(1)}, ..., X_T^{(d)}).$$

Then, for large T

$$d_W(\mathbf{F}_T, N(0, I_{d+1})) \le C\frac{1}{\log\log T}. \tag{52}$$

Proof: Let $F_T^{(j)}, j = 0, .., d$ be the components of the vector $\mathbf{F}_T$. Using the Stein-Malliavin bound (43), we have

$$\begin{aligned} d_W(\mathbf{F}_T, N(0, I_{d+1})) &\le C\left[\sum_{i=0}^{d+1}\mathbf{E}\left|1 - \langle DF_T^{(i)}, D(-L)^{-1}F_T^{(i)}\rangle\right|\right. \\ &+ \left.\sum_{i=0}^{d+1}\mathbf{E}\left|\langle DF_T^{(i)}, D(-L)^{-1}F_T^{(j)}\rangle + \langle DF_T^{(j)}, D(-L)^{-1}F_T^{(i)}\rangle\right|\right]. \end{aligned}$$

The main contribution will come from the diagonal term. By Lemma 4 (with $R^{(i,j)}{}_T$ as in the statement of Lemma 4) ,

$$\begin{aligned} &\left|\langle DF_T^{(i)}, D(-L)^{-1}F_T^{(i)}\rangle - 1\right| \\ = &\frac{1}{\frac{1}{2}\log\log T}\sum_{p\le T}\frac{1}{p}\sin^2((U+i)T\log p) \\ &+\frac{1}{\frac{1}{2}\log\log T}\mathbf{E}\sum_{p_1,p_2\le T; p_1\ne p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin((U+i)T\log p_1)\sin((U+j)T\log p_2) \\ &+\frac{1}{\frac{1}{2}\log\log T}R_T^{(i,i)} - 1 \\ = &\left[\frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{2p} - 1\right] - \frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{2p}\cos(2(U+i)T\log p) \\ &+\frac{1}{\frac{1}{2}\log\log T}\sum_{p_1,p_2\le T; p_1\ne p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin((U+i)T\log p_1)\sin((U+j)T\log p_2) \\ &+\frac{1}{\frac{1}{2}\log\log T}R_T^{(i,i)} \\ := &\left[\frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{2p} - 1\right] + R_T \end{aligned}$$

where we included in the rest term R_T the summands $-\frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{2p}\cos(2(U+i)T\log p)$, $\frac{1}{\frac{1}{2}\log\log T}\sum_{p_1\ne p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\sin((U+i)T\log p_1)\sin((U+j)T\log p_2)$ and $\frac{1}{\frac{1}{2}\log\log T}R_T^{(i,i)}$.

The first term above is the one which gives the bound (52). Indeed, by (68),

$$\left| \frac{1}{\frac{1}{2}\log\log T} \sum_{p\leq T} \frac{1}{2p} - 1 \right| \leq C \frac{1}{\log\log T}.$$

On the other hand, using only the bounds (34), (35) and (36), we can easily proof that

$$\mathbf{E}|R_T| \leq C \frac{1}{\log\log T}.$$

(actually, a better estimated is possible but not necessarily for our purpose). Again by (34), (35) and (36),

$$\mathbf{E}\left| \langle DF_T^{(i)}, D(-L)^{-1}F_T^{(j)}\rangle + \langle DF_T^{(j)}, D(-L)^{-1}F_T^{(i)}\rangle \right| \leq C\frac{1}{\log\log T}$$

for every $i \neq j$. The two above estimates lead to (52). ■

3.1.2 Error bound for the multidimensional Selberg theorem

We regard now the asymptotic behavior of the vector

$$\left(\log\left|\zeta(\frac{1}{2} + \mathbf{i}P_i)\right|\right)_{i=0,..,d}$$

with the evaluation points $P_i = T(U+i), i = 0,..,d$. We show that, after normalization, it also converges to a standard Gaussian vector. Recall that U denote a standard uniform random variable defined by (8).

We have

Theorem 4 *Let*

$$\mathcal{X}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(\log|\zeta(\frac{1}{2} + \mathbf{i}UT)|, \log|\zeta(\frac{1}{2} + \mathbf{i}(U+1)T)|, \dots, \log|\zeta(\frac{1}{2} + \mathbf{i}(U+d)T)|\right) \tag{53}$$

with U from (8). Then for large T,

$$d_W(\mathcal{X}_T, N(0, I_{d+1}) \leq C \frac{1}{\sqrt{\log\log T}}.$$

Proof: As in (39), using the triangle inequality for the Wasserstein distance and the fact that for any $\mathbb{R}^d$- valued random variables F, G we have

$$d_W(F,G) \leq \mathbf{E}|F-G|_1 \leq (\mathbf{E}|F-G|_2)^{\frac{1}{2}}$$

where $|x|_1 = |x_1| + ...|x_d|$ and $|x|_2^2 = |x_1|^2 + ... + |x_d|^2$ if $x = (x_1, .., x_d)$, we can write

$$\begin{aligned} d_W(\mathcal{X}_T, N(0, I_{d+1}) &\leq d_W(\mathbf{F}_T, N(0, I_{d+1})) \\ &+ \mathbf{E}|\mathcal{X}_T - \mathbf{F}_T|_1 + \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p \leq T} \frac{1}{\sqrt{p}} \left|(\mathbf{E}(\cos(T \log p(U+i))))_{i=0,..,d}\right|_1 \end{aligned}$$

and by Proposition 3

$$d_W(\mathbf{F}_T, N(0, I_{d+1})) \leq C \frac{1}{\log\log T}$$

while the inequality (37) in Lemma 2 implies

$$\mathbf{E}|\mathcal{X}_T - \mathbf{F}_T|_1 \leq C \frac{1}{\sqrt{\log\log T}}.$$

Finally, as in the proof of Theorem 2 (using (34) and (68))

$$\frac{1}{\sqrt{\frac{1}{2}\log\log T}} \sum_{p \leq T} \frac{1}{\sqrt{p}} \left|(\mathbf{E}(\cos(T \log p(U+i))))_{i=0,..,d}\right|_1 \leq C \frac{1}{\sqrt{\log\log T}}.$$

■

Exactly the same analysis can be done for the imaginary part of $\log \zeta$ on the critical line. For $T > 0$, let us define

$$\mathbf{Y}_T = (Y_T^{(0)}, ..., Y_T^{(d)})$$

where, for every $i = 0, .., d$,

$$Y_T^{(i)} = \sum_{p \leq T} \frac{1}{\sqrt{p}} \left[\sin(T \log p(U+i)) - \mathbf{E}\sin(T \log p(U+i))\right].$$

As in the proof of Lemma 4, we can see that for every $i, j = 0, .., d$,

$$\begin{aligned} &\langle DY_T^{(i)}, D(-L)^{-1} Y_T^{(j)} \rangle + \langle DY_T^{(j)}, D(-L)^{-1} Y_T^{(i)} \rangle = \sum_{p_1, p_2 \leq T} \frac{1}{\sqrt{p_1 p_2}} \frac{\log p_2}{\log p_1} \\ &[\cos((U+i)T \log p_1) \cos((U+j)T \log p_2) + \cos((U+i)T \log p_2) \cos((U+j)T \log p_1)] \\ &+ R_T^{(i,j)} \end{aligned}$$

with

$$\begin{aligned} R_T^{(i,j)} = & \sum_{p_1, p_2 \leq T} \frac{1}{\sqrt{p_1 p_2}} \frac{\log p_2}{\log p_1} \\ & [-\cos(T(U+j) \log p_2) \cos(Ti \log p_1) - \cos(T(U+i) \log p_1) \cos(Tj \log p_1) \\ & - U \cos(T(U+j) \log p_2) \cos(T(i+1) \log p_1) - U \cos(T(U+i) \log p_1) \cos(T(j+1) \log p_1) \\ & + U \cos(T(U+j) \log p_2) \cos(Ti \log p_1) + U \cos(T(U+i) \log p_2) \cos(Tj \log p_1)]. \end{aligned}$$

Using the proof of Proposition 3, with

$$\mathbf{G}_T := \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \mathbf{Y}_T,$$

we can show that

$$d_W(\mathbf{G}_T, N(0, I_{d+1})) \leq C \frac{1}{\log\log T.}$$

We will then obtain a similar result to Theorem 4.

Theorem 5 *With U from (8) and ζ from (1), let*

$$\mathcal{Y}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(\arg\log\zeta(\frac{1}{2} + \mathbf{i}UT), \arg\log\zeta(\frac{1}{2} + \mathbf{i}(U+1)T), \ldots, \arg\log\zeta(\frac{1}{2} + \mathbf{i}(U+d)T) \right). \tag{54}$$

Then for large T,

$$d_W(\mathcal{Y}_T, N(0, I_{d+1}) \leq C \frac{1}{\sqrt{\log\log T}}.$$

3.2 Small shifts: Convergence to a Gaussian vector with correlated components

The next step is to analyze the asymptotic behavior of the vector $\mathbf{V}_T$ (1) when the evaluation points P_i are close one to each other. That is, we choose $P_i = TU + f_T^{(i)}$, $i = 0, .., d$, with U a standard uniform random variable defined by (8) and $f_T^{(i)}$ are small shifts, meaning that the difference $f_T^{(i)} - f_T^{(j)}$ is small enough for $i \neq j$. This will lead to the appearance of non trivial correlations between the components of the limit of (41) as $T \to \infty$.

We first look to the asymptotic behavior of the $d+1$-dimensional Dirichlet series that approximates $\mathbf{V}_T$. That is, we introduce

$$\mathbf{Z}_T = \left(Z_T^{(0)}, Z_T^{(1)}, \ldots, Z_T^{(d)} \right)$$

with

$$Z_T^{(i)} = \sum_{p \leq T} \cos\left((TU + f_T^{(i)}) \log p \right), \quad i = 0, .., d \tag{55}$$

We will give the rate of convergence of $\mathbf{Z}_T$ to the $d+1$-dimensional Gaussian law with suitable covariance matrix. As before, the proof will be based on the inequality (43) and it means that we need to compute $\langle DZ_T^{(i)}, D(-L)^{-1} Z_T^{(j)} \rangle$ for every $i, j = 0, .., d$. We have

Lemma 5 *For $i=1,,,,d$, let $f_T^{(i)}$ be a deterministic function and let $Z_T^{(i)}$ be given by (55). Then*

$$\begin{aligned}
&\langle DZ_T^{(i)}, D(-L)^{-1}Z_T^{(j)}\rangle + \langle DZ_T^{(j)}, D(-L)^{-1}Z_T^{(i)}\rangle\\
= &\sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\\
&\Big[\sin\Big(\log p_1(TU+f_T^{(i)})\Big)\sin\Big(\log p_2(TU+f_T^{(j)})\Big)+\sin\Big(\log p_1(TU+f_T^{(j)})\Big)\sin\Big(\log p_2(TU+f_T^{(i)})\Big)\Big]\\
&+R_T^{(i,j)}
\end{aligned}$$

with

$$\begin{aligned}
&R_T^{(i,j)} = \sum_{p_1,p_2\le T}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_2}{\log p_1}\\
&\Big[(U-1)\sin\Big((TU+f_T^{(j)})\log p_1\Big)\sin\Big(f_T^{(i)}\log p_2\Big)+(U-1)\sin\Big((TU+f_T^{(i)})\log p_1\Big)\sin\Big(f_T^{(j)}\log p_2\Big)\\
&-U\sin\Big((TU+f_T^{(j)})\log p_1\Big)\sin\Big((T+f_T^{(i)})\log p_2\Big)-U\sin\Big((TU+f_T^{(i)})\log p_1\Big)\sin\Big((T+f_T^{(j)})\log p_2\Big)\Big].
\end{aligned}$$

Proof: With arguments previously used, we can write

$$\begin{aligned}
&\langle DZ_T^{(i)}, D(-L)^{-1}Z_T^{(j)}\rangle + \langle DZ_T^{(j)}, D(-L)^{-1}Z_T^{(i)}\rangle = \sum_{p_1,p_2}\frac{1}{\sqrt{p_1p_2}}\sum_{k,l\ge 0}\frac{(-1)^{k+l}}{(2k)!(2l)!}(\log p_1)^{2k}(\log p_2)^{2l}\\
&\times\Big[DH_{2k}^{(i)}, D(-L)^{-1}H_{2l}^{(j)}\rangle + \langle DH_{2l}^{(j)}, D(-L)^{-1}H_{2k}^{(i)}\rangle\Big]
\end{aligned}$$

where we used the notation

$$H_{2k}^{(i)} = (TU+f_T^{(i)})^{2k} - \mathbf{E}(TU+f_T^{(i)})^{2k}. \tag{56}$$

The scalar product above will be computed as in the proof of Lemma 3. We get

$$\begin{aligned}
&\langle DH_{2k}^{(i)}, D(-L)^{-1}H_{2l}^{(j)}\rangle + \langle DH_{2l}^{(j)}, D(-L)^{-1}H_{2k}^{(i)}\rangle\\
= &\frac{2k}{2l+1}U(TU+f_T^{(i)})^{2k-1}\Big[(T+f_T^{(j)})^{2l+1}-(f_T^{(j)})^{2l+1}-\frac{1}{U}(T+f_T^{(j)})^{2l+1}+\frac{1}{U}(f_T^{(j)})^{2l+1}\Big]\\
+ &\frac{2l}{2k+1}U(TU+f_T^{(j)})^{2l-1}\Big[(T+f_T^{(i)})^{2k+1}-(f_T^{(i)})^{2k+1}-\frac{1}{U}(TU+f_T^{(i)})^{2k+1}+\frac{1}{U}(f_T^{(i)})^{2k+1}\Big].
\end{aligned}$$

When we compute the above sums after k and l, we obtain

$$\sum_{l\ge1}U\frac{(-1)^l}{(2l-1)!}(\log p_2)^{2l}(TU+f_T^{(j)})^{2l-1} = -U\log p_2\sin\Big(\log p_2(TU+f_T^{(j)})\Big)$$

$$\sum_{k\geq 0}\frac{(-1)^k}{(2k+1)!}(\log p_1)^{2k}\left[(T+f_T^{(i)})^{2k+1}-(f_T^{(i)})^{2k+1}-\frac{1}{U}(TU+f_T^{(i)})^{2k+1}+\frac{1}{U}(f_T^{(i)})^{2k+1}\right]$$
$$=\frac{1}{\log p_1}\left[\sin\left(\log p_1(T+f_T^{(i)})\right)-\sin\left(\log p_1(f_T^{(i)})\right)-\frac{1}{U}\sin\left(\log p_1(TU+f_T^{(i)})\right)+\frac{1}{U}\sin\left(\log p_1(f_T^{(i)})\right)\right].$$

The conclusion follows easily. ■

Before stating the main results of this section, let us recall the following technical result due to [1], which plays a key role.

Lemma 6 *Let $(\Delta_T)_T$ be bounded and positive such that $\frac{\log \Delta_T}{\log\log T}\to_T c\in[0,\infty]$. Then $\frac{1}{\log\log T}\sum_{p\leq T}\frac{\cos(\log p\log\Delta_T)}{p}\to_T c\wedge 1$ and*

$$\left|\frac{1}{\log\log T}\sum_{p\leq T}\frac{\cos(\log p\log\Delta_T)}{p}-c\wedge 1\right|\leq C\frac{1}{\log\log T}. \tag{57}$$

The bound (57) is not explicitly stated in [1], but its proof is an easy consequence of the proof of Lemma 3.4 in [1].

The main result of this paragraph states as follows.

Proposition 4 *Assume $0\leq f_T^{(0)}<f_T^{(1)}<\ldots f_T^{(d)}<C<\infty$. For every $i,j=0,..,d$ with $i\neq j$ suppose that*

$$\frac{\log|f_T^{(i)}-f_T^{(j)}|}{\log\log T}\to a_{i,j}\in[0,\infty]. \tag{58}$$

Define

$$\mathbf{A}_T:=\frac{1}{\sqrt{\frac{1}{2}\log\log T}}\mathbf{Z}_T$$

where $\mathbf{Z}_T$ is the vector with components (55). Then $\mathbf{A}_T$ converges in distribution, as $T\to\infty$, to a centered Gaussian vector with covariance matrix $\Lambda=(c_{i,j})_{i,j=0,..,d}$ with $c_{i,j}=a_{i,j}\wedge 1$. Moreover, for T large

$$d_W(\mathbf{A}_T,N(0,\Lambda))\leq C\frac{1}{\log\log T}.$$

Proof: From (43)

$$\begin{aligned}
&d_W(A_T,N(0,\Lambda))\\
\leq\;& C\sum_{i=0}^{d}\mathbf{E}\left|1-\frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{p}\sin^2\left(\log p(TU+f_T^{(i)})\right)\right|\\
&+C\sum_{i\neq j}\mathbf{E}\left|c_{i,j}-\frac{1}{\frac{1}{2}\log\log T}\sum_p\frac{1}{p}\sin\left(\log p(TU+f_T^{(i)})\right)\sin\left(\log p(TU+f_T^{(j)})\right)\right|\\
&+r_T
\end{aligned}$$

with

$$\begin{aligned} r_T &= \frac{1}{\log\log T}\sum_{i,j}\mathbf{E}|R_T^{(i,j)}| \\ &+\frac{1}{\log\log T}\sum_{i=0}^{d}\sum_{p_1\neq p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_1}{\log p_2}\left|\mathbf{E}\sin\left(\log p_1(TU+f_T^{(i)})\right)\sin\left(\log p_2(TU+f_T^{(i)})\right)\right| \\ &+\frac{1}{\log\log T}\sum_{i,j=0;i\neq j}^{d}\sum_{p_1\neq p_2}\frac{1}{\sqrt{p_1p_2}}\frac{\log p_1}{\log p_2}\left|\mathbf{E}\sin\left(\log p_1(TU+f_T^{(i)})\right)\sin\left(\log p_2(TU+f_T^{(j)})\right)\right|. \end{aligned}$$

Hence, with the trigonometric indentity $\sin^2(x)=\frac{1-\cos(2x)}{2}$

$$\begin{aligned} d_W(\mathbf{A}_T,N(0,\Lambda)) &\leq C\left|1-\frac{1}{\log\log T}\sum_{p\leq T}\frac{1}{p}\right| \\ &+C\sum_{i,j=0;i\neq j}^{d}\left|c_{i,j}-\frac{1}{\log\log T}\sum_{p\leq T}\frac{1}{p}\cos\left(\log p(f_T^{(i)}-f_T^{(j)})\right)\right| \\ &+r_{T,2}+r_T \end{aligned} \tag{59}$$

with

$$\begin{aligned} r_{T,2} &= C\frac{1}{\log\log T}\sum_{i=0}^{d}\sum_{p\leq T}\frac{1}{p}\left|\mathbf{E}\cos\left(2\log p(TU+f_T^{(i)})\right)\right| \\ &+C\frac{1}{\log\log T}\sum_{i,j=0;i\neq j}^{d}\sum_{p\leq T}\frac{1}{p}\left|\mathbf{E}\cos\left(\log p(2TU+f_T^{(i)}+f_T^{(j)})\right)\right|. \end{aligned}$$

By (68),

$$\left|1-\frac{1}{\log\log T}\sum_{p\leq T}\frac{1}{p}\right|\leq C\frac{1}{\log\log T}. \tag{60}$$

Using

$$\begin{aligned} &|\mathbf{E}\cos(C\log p(TU+a_T))|\leq C\frac{1}{T\log p}, |\mathbf{E}\sin(C\log p(TU+a_T))|\leq C\frac{1}{T\log p} \\ &|\mathbf{E}U\sin(C\log p(TU+a_T))|\leq C\frac{1}{T\log p} \end{aligned}$$

we immediately get

$$\mathbf{E}|r_T|\leq C\frac{1}{\log\log T} \text{ and } \mathbf{E}|r_{2,T}|\leq C\frac{1}{\log\log T}. \tag{61}$$

By applying Lemma 6 to $\Delta_T = f_T^{(j)} - f_T^{(i)}$ we get

$$\frac{1}{\log\log T} \sum_{i,j=0;i\neq j}^{d} \left| c_{i,j} - \sum_{p\leq T} \frac{1}{p} \cos\left(\log p(f_T^{(i)} - f_T^{(j)})\right) \right| \leq C \frac{1}{\log\log T}. \tag{62}$$

By inserting (60), (61) and (62) into (59), we obtain the desired conclusion. ■

Concerning the imaginary part of the vector $\mathbf{V}_T$ (41), we let

$$\mathbf{B}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(W_T^{(0)}, \ldots, W_T^{(d)}\right)$$

with

$$W_T^{(i)} = \sum_p \left[\sin\left(\log p(TU + f_T^{(i)})\right) - \mathbf{E}\sin\left(\log p(TU + f_T^{(i)})\right)\right]$$

Then we will get for every $i, j = 0, .., d$

$$\begin{aligned}
&\langle DW_T^{(i)}, D(-L)^{-1}W_T^{(j)}\rangle \\
= &\sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \frac{\log p_2}{\log p_1} \\
&\left[\cos\left(\log p_2(TU + f_T^{(j)})\right)\cos\left(\log p_1(TU + f_T^{(i)})\right) + \cos\left(\log p_2(TU + f_T^{(i)})\right)\cos\left(\log p_1(TU + f_T^{(j)})\right)\right] \\
&+R_T^{(i,j)}
\end{aligned}$$

where

$$\begin{aligned}
R_T^{(i,j)} = &\sum_{p_1,p_2\leq T} \frac{1}{\sqrt{p_1p_2}} \frac{\log p_2}{\log p_1} \\
&\left[-U\cos\left(\log p_2(TU + f_T^{(j)})\right)\cos\left(\log p_1(T + f_T^{(i)})\right) - U\cos\left(\log p_2(TU + f_T^{(i)})\right)\cos\left(\log p_1(T + f_T^{(j)})\right)\right. \\
&\left.+U\cos\left(\log p_2(TU + f_T^{(j)})\right)\cos\left(\log p_1 f_T^{(i)}\right) + U\cos\left(\log p_2(TU + f_T^{(i)})\right)\cos\left(\log p_1 f_T^{(j)}\right)\right].
\end{aligned}$$

This will lead to the following result:

Proposition 5 *Let the assumptions in Proposition 4 prevail and let $\mathbf{B}_T$ be as above. Then*

$$d_W\left(\mathbf{B}_T, N(0, \Lambda)\right) \leq C\frac{1}{\log\log T}.$$

Theorem 6 *Let the assumption in Proposition 4 prevail. Define*

$$\mathcal{A}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}} \left(\log\left|\zeta\left(\frac{1}{2} + \mathbf{i}(TU + f_T^{(i)})\right)\right|\right)_{i=0,..,d}$$

and

$$\mathcal{B}_T = \frac{1}{\sqrt{\frac{1}{2}\log\log T}}\left(\arg\log\zeta\left(\frac{1}{2}+\mathbf{i}(TU+f_T^{(i)})\right)\right)_{i=0,..,d}.$$

Then

$$d_W(\mathcal{A}_T; N(0,\Lambda)) \leq C\frac{1}{\sqrt{\log\log T}} \text{ and } d_W(\mathcal{B}_T; N(0,\Lambda)) \leq C\frac{1}{\sqrt{\log\log T}}.$$

Proof: The conclusion follows from Propositions 4 and 5 and from the fact that the conclusion of Lemma 2 is true if we replace $T(U+i)$ by $TU+f_T^{(i)}$ (see Section 3.1 in [1]). ■

4 Fluctuations of the zeta zeros on the critical line

An application of the multidimensional Selberg theorem is to counting zeros of the Riemann zeta functions. Denote by $N(t)$ the number of non-trivial zeros of $\zeta(s)$ on the critical line $\operatorname{Re} s = \frac{1}{2}$ with the imaginary part contained in the interval $[0,t]$. Then there is known (see e.g. [15]) that

$$N(t) = \frac{t}{2\pi}\log\frac{t}{2\pi e} + \frac{1}{\pi}\arg\log\zeta\left(\frac{1}{2}+it\right) + \mathcal{O}(\frac{1}{t}). \tag{63}$$

If $t_1 < t_2$, let

$$\Delta(t_1,t_2) = (N(t_2)-N(t_1) - \left(\frac{t_2}{2\pi}\log\frac{t}{2\pi e} - \frac{t_1}{2\pi}\log\frac{t_1}{2\pi e}\right). \tag{64}$$

The quantity $\Delta(t_1,t_2)$ is usually interpreted as the *fluctuation of the number of zeta zeros* on the critical line between the heights $\operatorname{Im} s = t_1$ and $\operatorname{Im} s = t_2$ minus its expected value.

From the results in the previous section, we can deduce the asymptotic behavior of the fluctuations of zeta zeros between random points.

Proposition 6 *Let Δ be given by (64) and U by (8). Then for every $0 < i_1 < i_2$*

$$\frac{1}{\pi\sqrt{\log\log T}}\Delta\left(UT+i_1T, UT+i_2T\right) \to^{(d)}_{T\to\infty} N(0,1)$$

and

$$d_W\left(\frac{1}{\pi\sqrt{\log\log T}}\Delta\left(UT+i_1T, UT+i_2T\right), N(0,1)\right) \leq C\frac{1}{\sqrt{\log\log T}}.$$

Proof: We proved in Theorem 6 that the random vector

$$\left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}}(\arg\log\zeta\left(\frac{1}{2}+\mathbf{i}T(U+i)\right)\right)_{i=0,\ldots,d}$$

converges in distribution to $N(0, I_{d+1})$ with speed less than $C\frac{1}{\sqrt{\log\log T}}$. This implies the conclusion. ■

In particular, by choosing $i_1 = 0$ and $i_2 = 1$, we will have

$$d_W\left(\frac{1}{\pi\sqrt{\log\log T}}\Delta(UT, UT+T), N(0,1)\right) \le C\frac{1}{\sqrt{\log\log T}}. \tag{65}$$

We also have seen that in Theorem 6 that, if $f_T^{(i)}$ are as in the statement of Proposition 4, then the random vector

$$\left(\frac{1}{\sqrt{\frac{1}{2}\log\log T}}(\arg\log\zeta\left(\frac{1}{2}+\mathbf{i}(UT+f_T^{(i)}\right)\right)_{i=0,\dots,d}$$

converges in distribution to $N(0,\Lambda)$ (the matrix Λ has been introduced in Proposition 4) at rate $C\frac{1}{\sqrt{\log\log T}}$. Consequently, we have the following result.

Proposition 7 *For $0 < i_1 < i_2$ (so $f_T^{(i_1)} < f_T^{(i_2)}$) with $f_T^{(i)}$, $i = 0,..,d$ satisfying the assumptions in Proposition 4,*

$$\frac{1}{\pi\sqrt{\log\log T}}\Delta\left(UT+f_T^{(i_1)}, UT+f_T^{(i_2)}\right) \to N(0, 1-c_{i_1,i_2}).$$

and

$$d_W\left(\frac{1}{\pi\sqrt{\log\log T}}\Delta\left(UT+f_T^{(i_1)}, UT+f_T^{(i_2)}\right), N(0,1-c_{i_1,i_2})\right) \le C\frac{1}{\sqrt{\log\log T}}.$$

If we choose $d=1$, $f_T^{(0)} = 0$ and $f_T^{(1)} = \frac{1}{(\log T)^\delta}$ with $0<\delta<1$, then

$$\frac{\log|f_T^{(1)} - f_T^{(0)}|}{-\log\log T} \to_T c_{1,0} := \delta$$

so

$$\frac{\Delta\left((UT, UT+\frac{1}{(\log T)^\delta})\right)}{\pi\sqrt{\log\log T}} \to \sqrt{1-\delta}N$$

and

$$d_W\left(\frac{\Delta\left((UT, UT+\frac{1}{(\log T)^\delta})\right)}{\pi\sqrt{\log\log T}}; \sqrt{1-\delta}N\right) \le C\frac{1}{\sqrt{\log\log T}}. \tag{66}$$

Relations (65) and (66) can be interpreted as follows. The number of zeta zeros on the critical line between the heights UT is $UT+T$ is " close" to $\sqrt{\log\log T}$. With the same approximation error, the number of zeta zeros between the horizontal lines UT and $UT + \frac{1}{(\log T)^{\delta}}$ is approximately $\sqrt{1-\delta}\sqrt{\log\log T}$ with $\delta \in (0,1)$.

A last consequence concerns the so-called *mesoscopic flutuations* of the zeta zeros.

Corollary 1 *If K_T is a deterministic sequence such that $K_T > \varepsilon > 0$ for every $T > 0$ and*

$$\frac{\log K_T}{\log\log T} \to_{T\in\infty} \delta \in [0,1)$$

then the process

$$\left(\frac{\Delta(UT + \frac{\alpha}{K_T}), \Delta(UT + \frac{\beta}{K_T})}{\frac{1}{\pi}\sqrt{\frac{1}{2}(1-\delta)\log\log T}}, 0 \le \alpha < \beta < \infty\right)$$

converges in the sense of finite dimensional distributions to the centered Gaussian process $(G(\alpha,\beta), 0 \le \alpha < \beta < \infty)$ with covariance

$$\mathbf{E}G(\alpha,\beta)G(\alpha',\beta') = 1_{((\alpha=\alpha') \text{ and } (\beta=\beta'))} + \frac{1}{2}1_{((\alpha=\alpha') \text{ and } (\beta\neq\beta'))} + \frac{1}{2}1_{((\alpha\neq\alpha') \text{ and } (\beta=\beta'))} - \frac{1}{2}1_{(\beta=\alpha')}.$$

The Wassestein distance associated to this convergence is of order less than $C\frac{1}{\sqrt{\log\log T}}$.

This is interpreted in [1] or [2] as a mesoscopic repulsion of zeros. (Recall that mesoscopic means at a scale between microscopic and macroscopic.) The result shows that the zeta zeros do not affect too much the behavior of ζ on the critical line.

5 Appendix

5.1 Elements of number theory

Let $\pi(x)$ be the prime-counting function that gives the number of primes less than or equal to x, for any real number x. The prime number theorem then states that $\pi(x)$ behaves, when x is large, as $\frac{x}{\log x}$. As a consequence of this result, certain partial sums of primes can be estimated. We list below some estimates that are needed in our work.

For every s with $\operatorname{Re} s < 1$ we have

$$\sum_{p\le x} p^{-s} \sim \frac{x^{1-s}}{(1-s)\log s} \tag{67}$$

while if $s = 1$, the sum of the reciprocals of primes diverges as

$$\sum_{p\le x} \frac{1}{p} = \log\log x + C + \mathcal{O}\left(\frac{1}{\log x}\right). \tag{68}$$

We will also use (see e.g. [16])

$$\sum_{p\leq x} \frac{\log p}{p} \sim \log x \tag{69}$$

and

$$\sum_{p\leq x} \log p \sim x. \tag{70}$$

5.2 Basics of the Malliavin calculus

We present the elements from the Malliavin calculus that we need in the paper. Consider $\mathcal{H}$ a real separable Hilbert space and $(B(\varphi), \varphi \in \mathcal{H})$ an isonormal Gaussian process on a probability space $(\Omega, \mathcal{A}, P)$, which is a centered Gaussian family of random variables such that $\mathbf{E}\left(B(\varphi)B(\psi)\right) = \langle \varphi, \psi \rangle_{\mathcal{H}}$.

We denote by D the Malliavin derivative operator that acts on smooth functions of the form $F = g(B(\varphi_1), \ldots, B(\varphi_n))$ (g is a smooth function with compact support and $\varphi_i \in \mathcal{H}, i = 1, ..., n$)

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(B(\varphi_1), \ldots, B(\varphi_n))\varphi_i.$$

It can be checked that the operator D is closable from $\mathcal{S}$ (the space of smooth functionals as above) into $L^2(\Omega; \mathcal{H})$ and it can be extended to the space $\mathbb{D}^{1,p}$ which is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p}^{p} = \mathbf{E}F^p + \mathbf{E}\|DF\|_{\mathcal{H}}^{p}.$$

By L we will denote the infinitesimal genetaror of the ornstein-Uhlenbeck semigroup and by L^{-1} its pseudo-inverse. The reader may consult the monographs [9] or [5] for the definition and the properties of this operators. What is need in this paper concerning L and L^{-1} is only the formula (15).

References

[1] P. Bourgade (2010): *Mesoscopic fluctuations of the zeta zeros.* Probability Theory and Related Fields, 148, 479-500.

[2] M. Coram and P. Diaconis (2003): *New tests of the correspondence between unitary eigenvalues and the zeros of Riemanns zeta function.* J. Phys. A: Math. Gen., 36, 2883-2906

[3] C. Hughes, A. Nikeghbali and M. Yor (2008): *An arithmetic model for the total disorder process.* Probability Theory and Related Fields, 141, 47-59.

[4] P. Kühn (2011): *On Selberg's central limit theorem.* Master thesis, ETH Zurich.

[5] I. Nourdin and G. Peccati (2012): Normal Approximations with Malliavin Calculus From Stein 's Method to Universality. Cambridge University Press.

[6] I. Nourdin and G. Peccati (2007): *Stein's method on Wiener chaos.* Probability Theory and Related Fields 145, 75-118.

[7] I. Nourdin, G. Peccati and A. Réveillac (2010): *Multivariate normal approximation using Steins method and Malliavin calculus.* Ann. Inst. H. Poincar Probab. Statist. 46 (1), 45-58.

[8] I. Nourdin and F. Viens (2009): *Density formula and concentration inequalities with Malliavin calculus.* Electronic Journal of Probability, 14, paper 78, 2287-2309.

[9] D. Nualart, Malliavin Calculus and Related Topics, second edition, Springer, 2006.

[10] A. Selberg (1943): *On the normal density of primes in small intervals, and the difference between consecutive primes.* Arch. Math. Naturvid. 47 (6).

[11] A. Selberg (1944): *On the remainder in the formula for $N(T)$, the number of zeros of $\zeta(s)$ in the strip* $0 < t < T$. Arh. Norske Vid. Akad. Oslo, No. 1.

[12] A. Selberg (1946): *Contributions to the theory of the Riemann zeta function.* Arch. Math. Naturvid. 48 (5).

[13] A. Selberg (1992): *Old and new conjectures and results about a class of Dirichlet series.* In Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori 1989), pages 367-385. Univ. Salerno, Salerno.

[14] T. Tao (2009): *Selberg's limit theorem for the Riemann zeta function on the critical line*, blog article.

[15] E.C. Titchmarsh (1986): *The theory of the Riemann zeta function, Second Edition.* Oxford University Press.

[16] K. Tsang (1984): *The distribution of the values of the Riemann zeta-function.* Thesis Princeton University.

[17] M. Wahl (2014): *On the mod-Gaussian convergence of a sum over primes.* Mathematische Zeitschrift, 276 (3), 635-654.