# CONVERGENCE OF THE SPECTRUM OF EMPIRICAL COVARIANCE MATRICES FOR INDEPENDENT MRW PROCESSES 

Romain Allez ${ }^{1,2}$, Rémi Rhodes ${ }^{1,2}$ and Vincent Vargas $^{1,2}$


#### Abstract

We study the asymptotic of the spectral distribution for large empirical covariance matrices composed of independent lognormal Multifractal Random Walk processes. The asymptotic is taken as the observation lag shrinks to 0 . In this setting, we show that there exists a limiting spectral distribution whose Stieltjes transform is uniquely characterized by equations which we specify. We also illustrate our results by numerical simulations.


Mathematics Subject Classification. 60B20, 60G18, 60G15, 91G99.

Received December 3, 2013. Revised September 12, 2014.

## 1. Introduction

Since the seminal work of Marchenko and Pastur [12] in 1967, there has been growing interest in studying the asymptotic of large empirical covariance matrices. These studies have found applications in many fields of science: physics, telecommunications, information theory and finance, etc. The main motivation of this work stems from finance: the study of covariance matrices is a crucial tool for minimizing the risk $\mathcal{R}_{w}$ of a portfolio $w$ that invests $w_{i}$ in asset number $i$. Indeed, if we denote by $r_{i}$ the price variation of asset $i, \mathcal{R}_{w}$ can be defined as the variance of the random variable $\sum_{i} w_{i} r_{i}$ and can be computed in terms of the covariance matrix $\mathbf{R}$ of the $r_{i}$ (defined as $\left.\mathbf{R}_{i j}=\mathbb{E}\left[r_{i} r_{j}\right]\right)$ :

$$
\mathcal{R}_{w}=w^{t} \mathbf{R} w
$$

Of course, practitioners do not have access to $\mathbf{R}$; instead, they must consider a noisy empirical estimator of $\mathbf{R}$, which consists of a large empirical covariance matrix. A key tool in distinguishing noise from real correlations is the study of the eigenvalues of the empirical covariance matrix: we refer to $[6,15]$ for more extended discussions on the applications of large empirical covariance matrices in finance and in particular in portfolio theory.

We will work in a high frequency setting: we consider $N$ stock price processes $X_{i}(t)$ for $i=1, \ldots, N$ that evolve continuously with respect to time $t \in[0 ; 1]$ but we observe those prices only on a discrete finite grid $\{j / T, j=1, \ldots, T\}$ where $T$ is the number of observations. Using this discrete grid, we can compute the price

[^0]variations $r_{i}(j)$ (that we will abusively call returns) for each asset price $X_{i}$ on every time interval $[(j-1) / T$; $j / T]$ by:
$$
r_{i}(j):=X_{i}\left(\frac{j}{T}\right)-X_{i}\left(\frac{j-1}{T}\right) .
$$

Then, we define the $N \times T$ matrix $X_{N}$ such that $X_{N}(i j)=r_{i}(j)$ that enables to define the empirical covariance matrix $R_{N}$ as follows

$$
R_{N}:=X_{N} X_{N}^{t}
$$

In this work, we will be interested in the statistics of the symmetric matrix $R_{N}$ and in particular in its spectrum, or more precisely, in its limiting spectral distribution in the limit of large matrices (i.e. when $N \rightarrow \infty$ ) for different models of the i.i.d. random continuous processes $\left(X_{i}(t)\right), i \in\{1, \ldots, N\}$ (see below for precise definitions). For this purpose, the Marchenko-Pastur paper enables to deal with the case where stock prices follow independent Brownian motions. More precisely, in this case, the matrix $X_{N}$ is defined as:

$$
\begin{equation*}
X_{N}(i j)=B_{i}\left(\frac{j}{T}\right)-B_{i}\left(\frac{j-1}{T}\right) \tag{1.1}
\end{equation*}
$$

where the $B_{i}$ are i.i.d. standard Brownian motions.
If $\lambda_{1}, \ldots, \lambda_{N}$ are the eigenvalues of $R_{N}$, the empirical spectral distribution of the matrix $R_{N}$ is the probability measure defined by:

$$
\begin{equation*}
\mu_{R_{N}}=\frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_{i}} \tag{1.2}
\end{equation*}
$$

The Marchenko-Pastur (MP) result states that, in the limit of large matrices $N, T \rightarrow \infty$ with $N / T \rightarrow q \in(0,1]$, the empirical spectral distribution $\mu_{R_{N}}$ weakly converges (almost surely) to a probability measure whose density $\rho_{0}(x)$ is:

$$
\begin{equation*}
\rho_{0}(x)=\frac{1}{2 \pi q} \frac{\sqrt{\left(\gamma_{+}-x\right)\left(x-\gamma_{-}\right)}}{x} \mathbb{1}_{\left[\gamma_{-}, \gamma_{+}\right]} \mathrm{d} x \tag{1.3}
\end{equation*}
$$

where $\gamma_{ \pm}=1+q \pm 2 \sqrt{q}$.
Independently of the aforementioned work on random matrix theory, much work has been devoted to studying the statistics of financial stocks. It turns out that most financial assets (stocks, indices, etc.) possess universal features, called stylized facts. In short, one can observe empirically the following properties (the list below is obviously non exhaustive) for asset returns on financial markets:

- The returns are multifractal; in particular on short scales, they are heavy tailed but tend to have distribution closer to the Gaussian law on larger scales.
- The volatility fluctuates randomly and follows approximately a lognormal distribution.
- While the returns are rapidly decorrelated, the volatility exhibits long range correlations following a power law.

We refer to the references $[7,8]$ for a discussion on this topic. Many models have been proposed in the literature that take into account these stylized facts. Among them, there has been growing interest in the lognormal Multifractal Random Walk (MRW) model introduced in [2] (see also [1, 16]). The lognormal MRW model satisfies several of the so-called stylized facts, but a few of them remain unchecked such as asymmetry of returns and Leverage effect (see [5]). The lognormal MRW is simply defined as:

$$
\begin{equation*}
X(t)=B(M[0, t]) \tag{1.4}
\end{equation*}
$$

where $B$ is a standard Brownian motion and $M$ is an independent lognormal multifractal random measure (MRM for short) formally defined, for $t \geqslant 0$, by:

$$
\begin{equation*}
M[0 ; t]=\int_{0}^{t} \mathrm{e}^{\omega(x)-\frac{1}{2} \mathbb{E}\left[\omega(x)^{2}\right]} \mathrm{d} x \tag{1.5}
\end{equation*}
$$



Figure 1. Simulated path of a multifractal random walk with intermittency parameter $\gamma^{2}=1$ and with integral scale $\tau=1 / 4$. Note the intermittent bursts in volatility.
where $(\omega(x))_{x \in \mathbb{R}}$ is a stationary centered "Gaussian process" with a covariance function $K(x, y):=\mathbb{E}[\omega(x) \omega(y)]$ of the form

$$
K(x, y)=\gamma^{2} \ln _{+}\left(\frac{\tau}{|x-y|}\right)
$$

where $\ln _{+} x=\max (\ln x, 0)$. Note here that the variance $\mathbb{E}\left[\omega(x)^{2}\right]=K(x, x)$ is infinite so that the process $(\omega(x))$ is not well defined. We will see in Section 2.1 how to define the measure $M$ with a limiting procedure with Gaussian processes $\omega_{\varepsilon}(x)$ with covariance function $K_{\varepsilon}(x, y)$ such that $\lim _{\varepsilon \rightarrow 0} K_{\varepsilon}(x, y)=K(x, y)$. The two parameters $\gamma^{2}$ and $\tau$ are respectively called intermittency parameter and integral scale (or correlation length) of the lognormal random multifractal measure $M$.

Figure 1 represents a simulated path of a lognormal MRW $X(t)=B(M([0 ; t]))$ where $B$ is a standard Brownian motion independent of the multifractal random measure $M$ with intermittency parameter $\gamma^{2}=1$ and integral scale $\tau=1 / 4$. The reader can find a more precise reminder of the construction/definition of a more general class of Multifractal Random Measure (MRM), as well as (standard) notations used throughout the paper in Section 2.1.

We thus aim at studying the large sample covariance matrices where the underlying price processes evolve as lognormal MRW. More precisely, the matrix $X_{N}$ is defined, for $1 \leqslant i \leqslant N, 1 \leqslant j \leqslant T$, as:

$$
\begin{equation*}
X_{N}(i j)=B^{i}\left(M^{i}\left(0, \frac{j}{T}\right)\right)-B^{i}\left(M^{i}\left(0, \frac{j-1}{T}\right)\right) \tag{1.6}
\end{equation*}
$$

where the $B_{i}$ are i.i.d. Brownian motions and the $M_{i}$ are i.i.d. lognormal MRM independent of the $B_{i}$. Let us mention the work [11] which considers high frequency covariance matrices in the context of diffusion processes (see also [14] for studies of high frequency large empirical covariance matrices motivated by financial applications). The processes described by (1.6) are typically not diffusions.

In the spirit of the MP Theorem, the purpose of this work is to characterize the limit of the empirical spectral measure $\mu_{R_{N}}$ when $N, T \rightarrow \infty$ with $N / T \rightarrow q \in(0,1]$. It is interesting to understand how the long-memory volatility process affects the covariance matrix in the limit of large matrices. In particular, we will see that the intermittent volatility has the effect to spread the spectrum of the covariance matrix $R_{N}$ in a wider region
of $\mathbb{R}_{+}$. Indeed the spectral density has a compact support $\left[\gamma_{-} ; \gamma_{+}\right]$in the Marchenko-Pastur setting (in which the prices follow Brownian motions) whereas it has an infinite support with a tail that gets heavier as the intermittency parameter grows. We mention that our results can be extended to many different auto-correlated volatility processes.

The effect of the integral scale $\tau$ on the empirical covariance matrix $R_{N}$ is also very interesting in the context of price variations measured on a very short scale (high frequency). The high frequency case corresponds to large values of the parameter $\tau$ while low frequency case corresponds to small values of $\tau$. Indeed, if $X$ is a lognormal MRW with integral scale $\tau$, then the process $\widehat{X}(t)$ defined on $[0 ; 1]$ as $\widehat{X}(t)=X(t / 2)$ is a lognormal MRW with integral scale $2 \tau$. Note that this discussion on high freqency measurement is irrelevant in the MP case when asset prices follow independent Brownian motions since, in this model, the distribution of price variations is the same on any scale: It is Gaussian, only the variance will change with the scale and up to the variance parameter the limiting spectral distribution will always be the same at different scales. However, if asset prices follow lognormal MRW (or even another process with a correlated in time volatility process), the price variations measured on small scales will have a distribution with higher kurtosis (i.e. the probability mass of the tail is heavier) and therefore the spectrum of the empirical covariance matrix $R_{N}$ should be affected by decreasing the measurement scale. We therefore expect stronger right tail for the spectral distribution. The numerical analysis of our results indeed confirms this guess: The larger the integral scale is, the heavier the right tail is.

Here, we are mainly interested in the case where asset prices follow lognormal MRW but we will also present our results for two other related models where asset prices follow independent Brownian motions with a time change, which can be thought of as a volatility process with memory (i.e. the volatility process is correlated in time).

The next sections are organized as follows. In Section 2, we remind the definition of MRW and introduce the main notations of the paper. In Section 3, we state our main theorems which are characterizations of the limiting spectral measure of $R_{N}$ through its Stieltjes transform for different types of underlying processes $X$. These equations are tedious to invert analytically and it is hard to extract the properties (continuity, tails of the distribution) of the associated spectral density. In Section 4, we invert these equations numerically so as to get informations on the spectral measure of the covariance matrix $R_{N}$ as $N \rightarrow \infty$ and we check the validity and applicability of our results using numerical simulations. The proofs appear in Section 5 with some auxiliary lemmas proved in the appendix. The strategy of our proofs is classical among the random matrix literature (the so-called resolvent method) as it relies on the Schur recursion formula for the Stieltjes transform; in particular, we follow the approach of [4]. The main difficulty lies in handling the Stieltjes transforms in a multifractal setting.

## 2. BACKGRound, Notations and main Results

### 2.1. Reminder of the construction of lognormal MRM

To fix precisely the notations that we will use throughout the paper, we quickly remind the main steps of the construction of lognormal Multifractal Random Measures (MRM). Such measures attracted recently a lot of interest in various probabilistic settings and found numerous applications in natural sciences (see [13] for a general review on Gaussian multiplicative chaos). The description is necessarily concise and is restricted to the lognormal case. The reader is referred to [2] for further details on $\log$ Lévy MRM.

For $\gamma>0$, we set

$$
\varphi(p)=-i \frac{\gamma^{2}}{2} p-\frac{\gamma^{2}}{2} p^{2}
$$

We work on the half-space $S=\left\{(t, y) ; t \in \mathbb{R}, y \in \mathbb{R}_{+}^{*}\right\}$ on which we consider the positive measure

$$
\begin{equation*}
\theta(\mathrm{d} t, \mathrm{~d} y)=y^{-2} \mathrm{~d} t \mathrm{~d} y \tag{2.1}
\end{equation*}
$$

Let $\mu$ be an independently scattered Gaussian random measure on $S$ associated to the couple $(\varphi, \theta)$. It is worth recalling the properties which define an independently scattered Gaussian random measure:

- for every sequence of disjoint sets $\left(A_{n}\right)_{n}$ of $S,\left(\mu\left(A_{n}\right)\right)_{n}$ is a sequence of independent random variables and:

$$
\mu\left(\bigcup_{n} A_{n}\right)=\sum_{n} \mu\left(A_{n}\right), \quad \text { almost surely. }
$$

- for any $\theta$-measurable set $A, \mu(A)$ is a Gaussian random variable with mean $-\gamma^{2} \theta(A) / 2$ and variance $\gamma^{2} \theta(A)$. Its characteristic function is

$$
\mathbb{E}\left[\mathrm{e}^{\mathrm{i} q \mu(A)}\right]=\mathrm{e}^{\varphi(q) \theta(A)}
$$

We also introduce the Laplace exponent $\psi(p)=\varphi(-i p)=\frac{\gamma^{2}}{2}\left(p-p^{2}\right)$.
Definition 2.1 (Filtration $\left.\left(\mathcal{F}_{\varepsilon}\right)_{\varepsilon>0}\right)$.
Let $\Omega$ be the probability space on which $\mu$ is defined.
The $\sigma$-algebra $\mathcal{F}_{\varepsilon}$ is defined as the $\sigma$-algebra generated by the family of random variable $\{\mu(A \times B) ; A \in$ $\left.\mathcal{B}\left(\mathbb{R}^{2}\right), A \subset S, \operatorname{dist}\left(A, \mathbb{R}^{2} \backslash S\right) \geqslant \varepsilon\right\}$.

Given a positive parameter $\tau$, let us also define $f: \mathbb{R}_{+} \rightarrow \mathbb{R}$ by:

$$
f(r)= \begin{cases}r, & \text { if } r \leqslant \tau \\ \tau & \text { if } r \geqslant \tau\end{cases}
$$

The cone-like subset $A_{\varepsilon}(t)$ of $S$ is defined by:

$$
\begin{equation*}
A_{\varepsilon}(t)=\{(s, y) \in S ; y \geqslant \varepsilon,-f(y) / 2 \leqslant s-t \leqslant f(y) / 2\} \tag{2.2}
\end{equation*}
$$

We can then define the stationary process $\left(\omega_{\varepsilon}(t)\right)_{t \in \mathbb{R}}$ by:

$$
\begin{equation*}
\omega_{\varepsilon}(t)=\mu\left(A_{\varepsilon}(t)\right) \tag{2.3}
\end{equation*}
$$

The truncated random measures $M_{\varepsilon}$ (for which we will later check almost sure convergence to the so called lognormal multifractal random measure $M$ ) is then defined by: For all Borel sets $A$,

$$
M_{\varepsilon}(A)=\int_{A} \mathrm{e}^{\omega_{\varepsilon}(t)} \mathrm{d} t
$$

Note that $\mathbb{E}\left[M_{\varepsilon}(A)\right]=|A|^{3}$.
The covariance structure of the Gaussian process $\left(\omega_{\varepsilon}(t)\right)_{t \in \mathbb{R}}$ is

$$
\begin{aligned}
K_{\varepsilon}(s, t) & :=\mathbb{E}\left[\left(\omega_{\varepsilon}(s)+\frac{\gamma^{2}}{2}\left(\ln \left(\frac{\tau}{\varepsilon}\right)+1\right)\right)\left(\omega_{\varepsilon}(t)+\frac{\gamma^{2}}{2}\left(\ln \left(\frac{\tau}{\varepsilon}\right)+1\right)\right)\right] \\
& = \begin{cases}\gamma^{2}\left(\ln \left(\frac{\tau}{\varepsilon}\right)+1-\frac{|t-s|}{\varepsilon}\right), & \text { if }|t-s| \leqslant \varepsilon \\
\gamma^{2} \ln \left(\frac{\tau}{|t-s|}\right) & \text { if }|t-s| \geqslant \varepsilon .\end{cases}
\end{aligned}
$$

The (Radon) random measure $M$ is then defined as the almost sure limit (in the sense of weak convergence of Radon measures) by:

$$
M(A):=\lim _{\varepsilon \rightarrow 0^{+}} M_{\varepsilon}(A)
$$

[^1]for any Lebesgue measurable subset $A \subset \mathbb{R}$. The convergence is ensured by the fact that the family $\left(M_{\varepsilon}(A)\right)_{\varepsilon>0}$ is a right-continuous positive martingale with respect to the filtration $\left(\mathcal{F}_{\varepsilon}\right)_{\varepsilon>0}$.

We introduce the structure exponent $\zeta$ defined by

$$
\zeta(p):=p-\psi(p)=\left(1+\frac{\gamma^{2}}{2}\right) p-\frac{\gamma^{2}}{2} p^{2}
$$

It can be shown (see [2]) that the measure $M$ is different from 0 if and only if $\gamma^{2}<2$. If $\gamma^{2}<2$, one can check that there exists $\varepsilon>0$ such that $\zeta(1+\varepsilon)>1$. In this case, we have the following theorem.

Theorem 2.2. The measure $M$ is stationary and satisfies the exact stochastic scale invariance property: For any $\lambda \in] 0,1]$, we have the following equality in law between the two families of random variables indexed by the Borel sets $A$ included in the Euclidean ball $B(0 ; \tau)$ of radius $\tau$

$$
\begin{equation*}
(M(\lambda A))_{A \subset B(0, \tau)} \stackrel{\text { law }}{=}\left(\lambda \mathrm{e}^{\Omega_{\lambda}} M(A)\right)_{A \subset B(0, \tau)} \tag{2.4}
\end{equation*}
$$

where $\Omega_{\lambda}$ is a Gaussian random variable independent of $(M(A))_{A \subset B(0, \tau)}$, with mean $-\gamma^{2} / 2 \ln \lambda$ and variance $\gamma^{2} \ln \lambda$.

In addition, it is proven in [2] that for any $p>0$, we have $\mathbb{E}\left[M[0 ; 1]^{q}\right]<+\infty$ if and only if $\zeta(q) \geqslant 1$. For such $q$, it is easy to deduce from (2.4) that for all $\lambda \in[0 ; \tau]$, we have

$$
\begin{equation*}
\mathbb{E}\left[M[0 ; \lambda]^{q}\right]=\lambda^{\zeta(q)} \mathbb{E}\left[M[0 ; 1]^{q}\right] \tag{2.5}
\end{equation*}
$$

The function $\zeta$ is concave, the relation (2.5) exhibits the multifractal behavior of the measure $M$. The function $\zeta$ is often referred as the structure exponent in the multifractal formalism.

### 2.2. Notations

Let $N$ and $T:=T(N)$ be two integers, the aim of this paper is to compute the empirical spectral measure of the matrix $R_{N}:=X_{N}{ }^{t} X_{N}$ as $N \rightarrow \infty$, where $X_{N}$ is a $N \times T$ real matrix the entries of which are given by (1.6). Recall that the number $N$ of sampled processes is supposed to be comparable with the sample size $T:=T(N)$, and more precisely, we will suppose in the following that there exists a parameter $q \in] 0,1]$ such that:

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \frac{N}{T}=q \tag{2.6}
\end{equation*}
$$

We further set $\widetilde{R}_{N}:={ }^{t} X_{N} X_{N}$, and if $M$ is a symmetric real matrix, we will denote by $\mu_{M}$ the empirical spectral measure of $M$.

Define the $(T+N) \times(T+N)$ matrix $B_{N}$ by:

$$
B_{N}=\left(\begin{array}{cc}
0 & { }^{t} X_{N} \\
X_{N} & 0
\end{array}\right)
$$

We also define for $z \in \mathbb{C} \backslash \mathbb{R}$,

$$
A_{N}(z)=\left(z I_{T+N}-B_{N}\right)=\left(\begin{array}{cc}
z I_{T} & -{ }^{t} X_{N} \\
-X_{N} & z I_{N}
\end{array}\right)
$$

Notice that

$$
B_{N}^{2}=\left(\begin{array}{cc}
\widetilde{R}_{N} & 0 \\
0 & R_{N}
\end{array}\right)
$$

and that the eigenvalues of $\widetilde{R}_{N}$ are those of $R_{N}$ augmented with $T-N$ zero eigenvalues. We thus have:

$$
\begin{equation*}
\mu_{B_{N}^{2}}=2 \frac{N}{N+T} \mu_{R_{N}}+\frac{T-N}{N+T} \delta_{0} \tag{2.7}
\end{equation*}
$$

where $\delta_{x}$ stands for the Dirac mass at $x$. Combining this equality with the relation

$$
\begin{equation*}
\int f(x) \mu_{B_{N}^{2}}(\mathrm{~d} x)=\int f\left(x^{2}\right) \mu_{B_{N}}(\mathrm{~d} x) \tag{2.8}
\end{equation*}
$$

true for all bounded continuous functions $f$ on $\mathbb{R}$, we see that it is sufficient to study the weak convergence of the spectral measure of $B_{N}$ for the study of the convergence of the spectral measure $\mu_{R_{N}}$.

We will thus work on the (weak) convergence of the spectral measures $\mu_{B_{N}}$ and $\mathbb{E}\left[\mu_{B_{N}}\right]$ in the following. To that purpose, it is sufficient to prove the convergence of the Stieltjes transform of these two measures. Recall that, for a probability measure $\mu$ on $\mathbb{R}$, the Stieltjes transform $G_{\mu}$ of $\mu$ is defined, for all $z \in \mathbb{C} \backslash \mathbb{R}$, as:

$$
\begin{equation*}
G_{\mu}(z)=\int_{\mathbb{R}} \frac{1}{z-x} \mu(\mathrm{~d} x) \tag{2.9}
\end{equation*}
$$

and one can note that:

$$
\begin{equation*}
G_{\mu_{B_{N}}}(z)=\frac{1}{N+T} \operatorname{Trace}\left(G_{N}(z)\right) \tag{2.10}
\end{equation*}
$$

where we have set:

$$
\begin{equation*}
G_{N}(z)=\left(A_{N}(z)\right)^{-1} \tag{2.11}
\end{equation*}
$$

Hence, we have to investigate the convergence of the right-hand side of (2.10). Let us introduce the two following complex measures $L_{N}^{1, z}$ and $L_{N}^{2, z}$ such that, for all bounded and measurable function $f:[0,1] \rightarrow \mathbb{R}$ :

$$
\begin{align*}
& L_{N}^{1, z}(f)=\frac{1}{T} \sum_{k=1}^{T} f\left(\frac{k}{T}\right) G_{N}(z)_{k k}  \tag{2.12}\\
& L_{N}^{2, z}(f)=\frac{1}{N} \sum_{k=1}^{N} f\left(\frac{k}{N}\right) G_{N}(z)_{k+T, k+T} \tag{2.13}
\end{align*}
$$

Clearly, we have the relation

$$
\begin{equation*}
\frac{1}{N+T} \operatorname{Trace}\left(G_{N}(z)\right)=\frac{T}{N+T} L_{N}^{1, z}([0,1])+\frac{N}{N+T} L_{N}^{2, z}([0,1]) \tag{2.14}
\end{equation*}
$$

so that it suffices to establish the convergence of the two complex measures $L_{N}^{1, z}$ and $L_{N}^{2, z}$.

## 3. MAIN RESULTS

We will make the assumption that the intermittency parameter $\gamma^{2}$ is small enough so as to overcome in our proofs the strong correlations of the model.

Assumption 3.1. More precisely, let us suppose that:

$$
\begin{equation*}
\gamma^{2}<\frac{1}{3} \tag{3.1}
\end{equation*}
$$

Though we conjecture that our results hold as soon as the measure $M$ is non degenerated, i.e. $\gamma^{2}<2$ (see [2]), Assumption 3.1 is largely sufficient to cover most practical applications. For instance, in financial applications or in the field of turbulence, $\gamma^{2}$ is found empirically around $2 \times 10^{-2}$.

We can now state our main result about the convergence of the empirical spectral measure and mean empirical spectral measure of the matrix $R_{N}$.

## Theorem 3.2.

(i) The mean spectral measure $\mathbb{E}\left[\mu_{R_{N}}\right]$ converges weakly when $N$ goes to $\infty$ to a probability measure on $\mathbb{R}_{+}$ denoted $\rho(\mathrm{d} x)$.
(ii) The spectral measure $\mu_{R_{N}}$ converges weakly in probability to the probability measure $\rho(\mathrm{d} x)$. More precisely, for any bounded and continuous function $f, \int f(x) \mu_{R_{N}}(\mathrm{~d} x)$ converges in probability to $\int f(x) \rho(\mathrm{d} x)$ when $N$ goes to $+\infty$.
(iii) Let $N_{k}$ be an increasing sequence of integers such that $\sum_{k=1}^{\infty} N_{k}^{-1}<+\infty$, then the sequence $\mu_{R_{N_{k}}}$ converges weakly almost surely to the probability measure $\rho(\mathrm{d} x)$ when $k$ goes to $+\infty$.

We first prove Theorem 3.3 which establishes the convergence of the complex measures $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$. Theorem 3.2 is then implied by Theorem 3.3, equations (2.10) and (2.14) noting in addition that the probability measure $\rho(\mathrm{d} x)$ is the push-forward of the measure $v(\mathrm{~d} x)$ (introduced in Thm. 3.3) by the mapping $x \mapsto x^{2}$. In the sequel, we will use the notation $\rho=v \circ\left(x^{2}\right)^{-1}$.

## Theorem 3.3.

(i) The measures $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$ converge weakly towards two complex measures. More precisely, there exist a unique $\mu_{z}^{2} \in \mathbb{C}$ and a unique bounded measurable function $K_{z}(x)$ over $[0,1]$ such that, for all bounded and continuous function $f$ on $[0,1]$, we have respectively:

$$
\begin{aligned}
& \mathbb{E}\left[L_{N}^{1, z}(f)\right] \rightarrow_{N \rightarrow \infty} \int_{0}^{1} K_{z}(x) f(x) \mathrm{d} x \\
& \mathbb{E}\left[L_{N}^{2, z}(f)\right] \rightarrow_{N \rightarrow \infty} \mu_{z}^{2} \int_{0}^{1} f(x) \mathrm{d} x
\end{aligned}
$$

(ii) In addition, we have the following relation between $\mu_{z}^{2} \in \mathbb{C}$ and $K_{z}(x)$ :

$$
\begin{equation*}
\int_{0}^{1} K_{z}(x) \mathrm{d} x=q \mu_{z}^{2}+\frac{1-q}{z} \tag{3.2}
\end{equation*}
$$

(iii) Furthermore, there exists a unique probability measure $v$ on $\mathbb{R}$ whose Stieltjes transform is $\mu_{z}^{2}$, meaning that for all $z \in \mathbb{C} \backslash \mathbb{R}$,

$$
\begin{equation*}
\mu_{z}^{2}=\int_{\mathbb{R}} \frac{v(\mathrm{~d} x)}{z-x} \tag{3.3}
\end{equation*}
$$

In the following Theorem, we give a characterization of the probability measure $v$ by means of its Stieltjes transform $\mu_{z}^{2}$.

Theorem 3.4. The constant $\mu_{z}^{2}$ and the bounded function $K_{z}(x)$ are uniquely determined for all $z \in \mathbb{C} \backslash \mathbb{R}$, by the following system of equations:

$$
\begin{align*}
\mu_{z}^{2} & =\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(t) M(\mathrm{~d} t)\right)^{-1}\right]  \tag{3.4}\\
K_{z}(x) & =\left(z-q \mathbb{E}\left[\left(z-\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} K_{z}(t) M(\mathrm{~d} t)\right)^{-1}\right]\right)^{-1} \tag{3.5}
\end{align*}
$$

where ${ }^{4} M$ is the $M R M$ with structure exponent $\zeta(q)=\left(1+\gamma^{2} / 2\right) q-q^{2} \gamma^{2} / 2$.

[^2]Our approach to show the convergence of $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$ consists in proving tightness and characterizing uniquely the possible limit points. The classical Schur complement formula is our basic linear algebraic tool to study $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$ recursively on the dimension $N$, as is usual when the resolvent method is used. The original part of our proof is that we apply the Schur complement formula two times in a row to find the second equation of the system in Theorem 3.4 involving the limit point $K_{z}(x)$ of the measure $\mathbb{E}\left[L_{N}^{1, z}\right]$. We will also show that the limit points of the two complex measures $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$ satisfy a fixed point system (written in Thm. 3.4).

Let us notice that one can give a precise meaning to (3.5) for all $\gamma^{2} \in[0,2[$. Indeed, we can define for all $x \in[0,1]$ and all continuous function $f$, the following almost sure limit as a definition:

$$
\begin{equation*}
\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} f(t) M(\mathrm{~d} t)=\lim _{\eta \rightarrow 0} \int_{t \in[0,1] ;|t-x|>\eta}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} f(t) M(\mathrm{~d} t) \tag{3.6}
\end{equation*}
$$

Note that the above limit exists almost surely since, for $x$ fixed:

$$
\frac{\ln M\left[x-\varepsilon_{k}, x+\varepsilon_{k}\right]}{\ln \varepsilon_{k}} \underset{k \rightarrow \infty}{\rightarrow} 1+\frac{\gamma^{2}}{2}, \text { a.s. }
$$

where $\varepsilon_{k}=1 / 2^{k}$. One can also check with this definition that we have:

$$
\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} f(t) M(\mathrm{~d} t)=\lim _{\varepsilon \rightarrow 0} \int_{0}^{1} \mathrm{e}^{\operatorname{cov}\left(\omega_{\varepsilon}(t), \omega_{\varepsilon}(x)\right)} f(t) \mathrm{e}^{\omega_{\varepsilon}(t)} \mathrm{d} t
$$

Conjecture 3.5. With this extended definition, we conjecture that Theorem 3.4 holds in the lognormal multifractal case for all $\gamma^{2} \in[0,2[$.

## 4. Numerical Results

We want to extract information on the spectral density $v \circ\left(x^{2}\right)^{-1}$ of the covariance matrix $R_{N}$ in the limit of large matrices. This section will also give evidence that our equations are easy to use in practice for applications.

The information on the measure $v$ is entirely contained in its Stieltjes transform $\mu_{z}^{2}$ which is the unique solution of the system of equations (3.4) and (3.5). Let us admit for clarity at this point that the measure $v$ has a continuous density, at least on the set $\mathbb{R} \backslash\{0\}$. One should be able to show that this is indeed true using the two equations (3.4) and (3.5) that characterize the probability measure $v$. Under this continuity assumption for $v(x)$, we can re-find the density $v(x)$ from $\mu_{z}^{2}$ by the relation

$$
\begin{equation*}
\lim _{\varepsilon \rightarrow 0} \frac{1}{\pi} \Im\left(\mu_{x-i \varepsilon}^{2}\right)=v(x) \tag{4.1}
\end{equation*}
$$

Note that we just need to find the unique family of functions $\left(K_{z}(x)\right)_{x \in[0 ; 1]}$ for $z \in \mathbb{C} \backslash \mathbb{R}$ near the real line, that verifies the fixed point equation (3.5). Indeed, knowing $\left(K_{z}(x)\right)_{x \in[0 ; 1]}$, we can compute $\mu_{z}^{2}$ by using equation (3.4), or even simpler, the additional relation that we stated above

$$
\begin{equation*}
\int_{0}^{1} K_{z}(x) \mathrm{d} x=q \mu_{z}^{2}+\frac{1-q}{z} \tag{4.2}
\end{equation*}
$$

Let $\mathcal{C}([0 ; 1], \mathbb{C})$ be the space of bounded functions from $[0 ; 1]$ to $\mathbb{C}$. For $z \in \mathbb{C} \backslash \mathbb{R}$ fixed, the idea to find $\left(K_{z}(x)\right)_{x \in[0 ; 1]}$ is to use the fixed point method due to Picard. Let us introduce the operator $T: \mathcal{C}([0 ; 1], \mathbb{C}) \rightarrow$ $\mathcal{C}([0 ; 1], \mathbb{C})$ by setting, for $g \in \mathcal{C}([0 ; 1], \mathbb{C})$ and for all $x \in[0,1]$ :

$$
\begin{equation*}
T g(x)=\frac{1}{z-q \mathbb{E}\left[\left(z-\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} g(t) M(\mathrm{~d} t)\right)^{-1}\right]} \tag{4.3}
\end{equation*}
$$



Figure 2. Comparison between the theoretical value of the density $v \circ\left(x^{2}\right)^{-1}(x)$ and the empirical histogram computed through a sample of simulated empirical covariance matrices $R_{N}$ as defined in the introduction with $q=1$. The stock prices follow lognormal multifractal random walks with intermittency parameter $\gamma^{2}=1 / 2$.

It can easily be shown (see Sect. 5.6) that if $z \in \mathbb{C} \backslash \mathbb{R}$ is sufficiently far from the real line, then the operator $T$ is contracting and therefore admits a unique fixed point $K_{z}(\cdot)$ in $\mathcal{C}([0 ; 1], \mathbb{C})$. To find the fixed point $K_{z}$, we will iterate the operator $T$ starting from any fixed initial function $K_{z}^{(0)}$. We know that, for $z$ such that the operator $T$ is contracting, the $n$-th iteration of the function $K_{z}^{(n)}:=T\left(K_{z}^{(n-1)}\right)$ converges to the unique fixed point $K_{z}$. In fact, numerically, there is no need in applying the iteration on $T$ for $z$ such that $T$ is contracting (i.e. for $z$ far from the real line) and one can apply the Picard method directly near the real line ${ }^{5}$ and find the fixed point after a reasonable number of iterations of the operator $T$.

The multifractal lognormal random measure $M(\mathrm{~d} t)$ and multifractal random walk are simulated through the standard method by simulating first, with the use of fast Fourier transform, a Gaussian process with covariance function given for $\eta>0$ small by

$$
K_{\eta}(|t-s|)=\gamma^{2} \ln _{+}\left(\frac{\tau}{|t-s|+\eta}\right)
$$

The lognormal multifractal random measure and random walk are then constructed from this Gaussian process through the standard formulas (see e.g. $[2,16]$ ).

The results are as follows. In Figure 2, we show the comparison between the theoretical value of the density $v \circ\left(x^{2}\right)^{-1}(x)$ (computed numerically as described above) and an empirical histogram of the eigenvalues of a sample of simulated covariance matrices $R_{N}$ (defined in the introduction) for $N=1024$ and $q=1$. The agreement is excellent as expected from Theorems 3.2, 3.3 and 3.4. The figure is done for an intermittency parameter $\gamma^{2}=1 / 2$ and an integral scale $\tau=1 / 4$, suggesting that our prediction remains true for $\gamma^{2}>1 / 3$ (see conjecture 3.5 which also covers the case $\gamma^{2} \in[1,2[$ ).

[^3]

Figure 3. Log-log plot of the density $v \circ\left(x^{2}\right)^{-1}$ with $q=1, \tau=1 / 4$ for three different intermittency parameter: $\gamma^{2}=0$ (black dashed line), $\gamma^{2}=1 / 4$ (blue line) and $\gamma^{2}=1 / 2$ (red line). (Color online).

In Figure 3, we represent three curves (axis are in log-log) corresponding to the theoretical density $v \circ\left(x^{2}\right)^{-1}(x)$ for a parameter $q=1$, an integral scale $\tau=1 / 4$ and for three different values of $\gamma^{2}$. The black dashed curve corresponds to $\gamma^{2}=0$, which in fact is the Marchenko-Pastur case: asset prices are following independent Brownian motions with a trivial constant volatility process. In this case, the support is compact and the right edge of the spectrum is known to be equal to 4 . The blue curve corresponds to an intermittency parameter equal to $1 / 4$ and the red curve is for $\gamma^{2}=1 / 2$. In this way, we see precisely the distortion of the spectrum induced by the auto-correlated volatility process. The most interesting part for applications is certainly about the tails of the distribution: the higher the intermittency parameter $\gamma^{2}$ is, the heavier the tail of the distribution is.

In Figure 4, we represent four curves corresponding to the theoretical density $v \circ\left(x^{2}\right)^{-1}(x)$ but varying the integral scale $\tau$ instead of the intermittency parameter $\gamma^{2}$. We chose for this plot $q=1$ and $\gamma^{2}=1 / 4$ and represented the density $v \circ\left(x^{2}\right)^{-1}(x)$ for $\tau=0$ (corresponding to the trivial MP case) and for $\tau=1 / 4,1,2$. The result on the right tail of the distribution is the following: the higher the integral scale is, the heavier the right tail of the distribution is. As mentionned above, large integral scale corresponds to measuring price variations on small scales. On small scales, it is known that price variations will have distribution with larger kurtosis than price variations on larger scales and therefore it was expected to find heavier right tail distribution for the spectral distribution of the corresponding covariance matrix.

## 5. Proofs of the main Results

In this section, we give the proofs of Theorems $3.2,3.3$ and 3.4. The proofs are very similar when $q=1$ or when $q<1$. To simplify notations, we assume $T=N$ and hence $q=1$ in the proofs that follow.

We begin by showing tightness.


Figure 4. Log-log plot of the density $v \circ\left(x^{2}\right)^{-1}$ with $q=1, \gamma^{2}=1 / 4$ for four different integral scales $\tau: \tau=0$ (black dashed line), $\tau=1 / 4$ (red line), $\tau=1$ (blue line) and $\tau=2$ (green line). (Color online).

### 5.1. Tightness of the complex measures $\mathbb{E}\left[L_{N}^{1, z}\right], \mathbb{E}\left[\mathbf{L}_{N}^{2, z}\right]$ and limit points

Lemma 5.1. The two families of complex measures $\left(\mathbb{E}\left[L_{N}^{i, z}\right]\right)_{N \in \mathbb{N}}, i=1,2$ are tight and bounded in total variation.

Proof. Let us present the proof for $\left(\mathbb{E}\left[L_{N}^{1, z}\right]\right)_{N \in \mathbb{N}}$; the other proof is similar.
One has, for each $N$,

$$
\begin{equation*}
\left|\mathbb{E}\left[L_{N}^{1, z}\right]\right|[0,1]=\frac{1}{N} \sum_{k=1}^{N}\left|\mathbb{E}\left[G_{N}(z)_{k k}\right]\right| \leqslant \frac{1}{|\Im(z)|} \tag{5.1}
\end{equation*}
$$

and so the family of complex measures $\left(\mathbb{E}\left[L_{N}^{1, z}\right]\right)_{N \in \mathbb{N}}$ is bounded in total variation. It is obviously tight since the support of all the complex measures in the family is included in $[0,1]$, which is a compact set.

Using Prokhorov's theorem, we know that those two families of complex measures are sequentially compact in the space of complex Borel measures on $[0,1]$ equipped with the topology of weak convergence. In particular, there exists a subsequence such that, for all bounded continuous function $f$, one has, when $N$ goes to $+\infty$ along this subsequence:

$$
\begin{equation*}
\mathbb{E}\left[L_{N}^{1, z}(f)\right] \rightarrow \int_{0}^{1} f(x) \mu_{z}^{1}(\mathrm{~d} x) \tag{5.2}
\end{equation*}
$$

Lemma 5.2. The complex measure $\mu_{z}^{1}(\mathrm{~d} x)$ has Lebesgue density; more precisely, there exists a bounded measurable function $K_{z}(x)$ such that:

$$
\begin{equation*}
\mu_{z}^{1}(\mathrm{~d} x)=K_{z}(x) \mathrm{d} x \tag{5.3}
\end{equation*}
$$

Proof. One has:

$$
\begin{align*}
\left|\mathbb{E}\left[L_{N}^{1, z}(f)\right]\right| & \leqslant \frac{1}{N} \sum_{k=1}^{N}|f(k / N)| \mathbb{E}\left[G_{N}(z)_{k k}\right]  \tag{5.4}\\
& \leqslant \frac{1}{|\Im(z)|} \frac{1}{N} \sum_{k=1}^{N}|f(k / N)| \tag{5.5}
\end{align*}
$$

Letting $N \rightarrow+\infty$ along a subsequence, one obtains:

$$
\begin{equation*}
\left|\int_{0}^{1} f(x) \mu_{z}^{1}(\mathrm{~d} x)\right| \leqslant \frac{1}{|\Im(z)|} \int_{0}^{1}|f(x)| \mathrm{d} x . \tag{5.6}
\end{equation*}
$$

This proves the lemma.
Thus, there exists a subsequence such that, as $N$ tends to $+\infty$ along this subsequence:

$$
\begin{equation*}
\mathbb{E}\left[L_{N}^{1, z}(f)\right] \rightarrow \int_{0}^{1} f(x) K_{z}(x) \mathrm{d} x \tag{5.7}
\end{equation*}
$$

Lemma 5.3. There exists a subsequence and a constant $\mu_{z}^{2} \in \mathbb{C}$ such that, as $N$ goes to $+\infty$ along this subsequence:

$$
\begin{equation*}
\mathbb{E}\left[L_{N}^{2, z}(f)\right] \rightarrow \mu_{z}^{2} \int_{0}^{1} f(x) \mathrm{d} x \tag{5.8}
\end{equation*}
$$

Proof. It is easy to see that the $G_{N}(z)_{k k}, k=N+1, \ldots, 2 N$, are identically distributed. In particular, these variables have the same mean $\mu_{z}^{2}(N)$. One has, for all $N$,

$$
\begin{equation*}
\left|\mu_{z}^{2}(N)\right| \leqslant \frac{1}{|\Im(z)|} \tag{5.9}
\end{equation*}
$$

So there exists a subsequence and a complex number $\mu_{z}^{2}$ such that, as $N$ goes to $+\infty$ along this subsequence, $\mu_{z}^{2}(N) \rightarrow \mu_{z}^{2}$. One thus obtains, as $N$ goes to $+\infty$ along this subsequence:

$$
\begin{equation*}
\mathbb{E}\left[L_{N}^{2, z}(f)\right] \rightarrow \mu_{z}^{2} \int_{0}^{1} f(x) \mathrm{d} x \tag{5.10}
\end{equation*}
$$

Following the classical method as in $[3,4,10]$, we will show in the following that the limit point $\mu_{z}^{2}$ and $K_{z}(x)$ are defined uniquely and do not depend on the subsequence. We will first recall some preliminary results on resolvents.

### 5.2. Preliminary results on resolvents

We first recall the following standard and general result on resolvent matrices.
Lemma 5.4. Let $A$ be a symmetric real valued matrix of size $N$. For $z \in \mathbb{C} \backslash \mathbb{R}$, let us denote by $G(z)$ the matrix

$$
\begin{equation*}
G(z)=(z-A)^{-1} . \tag{5.11}
\end{equation*}
$$

For $z \in \mathbb{C} \backslash \mathbb{R}$ and $k \in\{1, \ldots, N\}$, we have

$$
\begin{equation*}
\Im(z) \Im\left(G(z)_{k k}\right)<0 \quad \text { and } \quad\left|G(z)_{k k}\right| \leqslant \frac{1}{|\Im(z)|} \tag{5.12}
\end{equation*}
$$

In particular, if $F \subset\{1, \ldots, N\}$ is a finite set and $\left(a_{i}\right)_{i \in F}$ a finite sequence of positive number, then:

$$
\begin{equation*}
\frac{\Im\left(z-\sum_{i \in F} a_{i} G(z)_{i i}\right)}{\Im(z)} \geqslant 1 \tag{5.13}
\end{equation*}
$$

and we also have:

$$
\begin{equation*}
\frac{1}{\left|z-\sum_{i \in F} a_{i} G(z)_{i i}\right|} \leqslant \frac{1}{|\Im(z)|} \tag{5.14}
\end{equation*}
$$

The next lemmas of this section are also standard but are applied to our particular case. For $i=1, \ldots, N$, let $X_{N}^{(i)}=\left(X_{N}(k l)\right)_{k, l \neq i}$ be the matrix obtained from $X_{N}$ by taking off the $i$-th row. Define, also for $i=1, \ldots, 2 N$ the $(2 N-1) \times(2 N-1)$ matrix $A_{N}^{(i)}(z)$ obtained from $A_{N}(z)$ by taking off the $i$-th column and row. In particular, for $i=1, \ldots, N$,

$$
A_{N}^{(N+i)}(z)=\left(\begin{array}{cc}
z I_{N} & -t X_{N}^{(i)} \\
-X_{N}^{(i)} & z I_{N-1}
\end{array}\right)
$$

For $i=1, \ldots, 2 N$, set:

$$
\begin{equation*}
G_{N}^{(i)}(z)=\left(A_{N}^{(i)}(z)\right)^{-1} \tag{5.15}
\end{equation*}
$$

Let now $\hat{X}_{N}^{(i)}$ denote the matrix $X_{N}$ with the $i$-th row set to 0 and $\hat{A}_{N}^{(i)}(z)$ denote the matrix $A_{N}(z)$ with the $i$-th column and row set to 0 excepted the diagonal term. Again we have, for $i=1, \ldots, N$ :

$$
\hat{A}_{N}^{(N+i)}(z)=\left(\begin{array}{cc}
z I_{N} & -{ }^{t} \hat{X}_{N}^{(i)} \\
-\hat{X}_{N}^{(i)} & z I_{N}
\end{array}\right)
$$

For $i=1, \ldots, 2 N$, set:

$$
\begin{equation*}
\hat{G}_{N}^{(i)}(z)=\left(\hat{A}_{N}^{(i)}(z)\right)^{-1} \tag{5.16}
\end{equation*}
$$

In the paper, we will also use the terms $A_{N}^{(k, i)}(z), G_{N}^{(k, i)}(z), \hat{A}_{N}^{(k, i)}(z), \hat{A}_{N}^{(k, i)}(z)$. The double superscript just means that you make the operations described above to the rows and columns $i$ and $k$.
Lemma 5.5. For all $k \in\{1, \ldots, N\}$ and all $t \neq N+k$, one has:

$$
\begin{equation*}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(N+k)}(z)_{t t}\right|\right] \leqslant \frac{1}{\sqrt{N}|\Im(z)|^{2}} \tag{5.17}
\end{equation*}
$$

Proof. Multiply the identity:

$$
\begin{equation*}
\hat{A}_{N}^{(N+k)}(z)-A_{N}(z)=\hat{A}_{N}^{(N+k)}(0)-A_{N}(0) \tag{5.18}
\end{equation*}
$$

to the left by $G_{N}(z)$ and to the right by $\hat{G}_{N}^{(N+k)}(z)$ to obtain

$$
\begin{equation*}
G_{N}(z)-\hat{G}_{N}^{(N+k)}(z)=G_{N}(z)\left(\hat{A}_{N}^{(N+k)}(0)-A_{N}(0)\right) \hat{G}_{N}^{(N+k)}(z) \tag{5.19}
\end{equation*}
$$

Then one has:

$$
\begin{align*}
G_{N}(z)_{t t}-\hat{G}_{N}^{(N+k)}(z)_{t t}= & \left(G_{N}(z)\left(\hat{A}_{N}^{(N+k)}(0)-A_{N}(0)\right) \hat{G}_{N}^{(N+k)}(z)\right)_{t t}  \tag{5.20}\\
= & \hat{G}_{N}^{(N+k)}(z)_{N+k, t} \sum_{i=1}^{N} G_{N}(z)_{t i} r_{k}(i)  \tag{5.21}\\
& +G_{N}(z)_{t, N+k} \sum_{j=1}^{N} r_{k}(j) \hat{G}_{N}^{(N+k)}(z)_{j t}  \tag{5.22}\\
= & G_{N}(z)_{t, N+k} \sum_{j=1}^{N} r_{k}(j) \hat{G}_{N}^{(N+k)}(z)_{j t} \tag{5.23}
\end{align*}
$$

where we have noticed that, for all $t \neq N+k, \hat{G}_{N}^{(N+k)}(z)_{N+k, t}=0$.

Therefore, we find that:

$$
\begin{equation*}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(N+k)}(z)_{t t}\right|\right] \leqslant \mathbb{E}\left[\left|G_{N}(z)_{t, N+k}\right|^{2}\right]^{1 / 2} \mathbb{E}\left[\left|\sum_{j=1}^{N} r_{k}(j) \hat{G}_{N}^{(N+k)}(z)_{j t}\right|^{2}\right]^{1 / 2} \tag{5.24}
\end{equation*}
$$

by Cauchy-Schwartz's inequality. Using then the independence of $r_{k}(j)$ and $\hat{G}_{N}^{(N+k)}(z)$ and Lemma A.1, we get:

$$
\begin{aligned}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(N+k)}(z)_{t t}\right|\right] & \leqslant \mathbb{E}\left[\left|G_{N}(z)_{t, N+k}\right|^{2}\right]^{1 / 2} \mathbb{E}\left[r_{k}(1)^{2}\right]^{1 / 2} \mathbb{E}\left[\sum_{j=1}^{N}\left|\hat{G}_{N}^{(N+k)}(z)_{j t}\right|^{2}\right]^{1 / 2} \\
& \leqslant \frac{1}{\sqrt{N}|\Im(z)|^{2}}
\end{aligned}
$$

The proof is complete.
Lemma 5.6. There exists a constant $C>0$ such that, for all $k \in\{1, \ldots, N\}$ and all $t \neq k$ :

$$
\begin{equation*}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(k)}(z)_{t t}\right|\right] \leqslant \frac{C}{|\Im(z)|^{2}} \frac{1}{N^{\frac{1-\gamma^{2}}{4}}} \tag{5.25}
\end{equation*}
$$

Proof. Again, we start from the relation:

$$
G_{N}(z)-\hat{G}_{N}^{(k)}(z)=G_{N}(z)\left(\hat{A}_{N}^{(k)}(0)-A_{N}(0)\right) \hat{G}_{N}^{(k)}(z)
$$

Thus we have

$$
\begin{align*}
G_{N}(z)_{t t}-\hat{G}_{N}^{(k)}(z)_{t t}= & \left(G_{N}(z)\left(\hat{A}_{N}^{(k)}(0)-A_{N}(0)\right) \hat{G}_{N}^{(k)}(z)\right)_{t t}  \tag{5.26}\\
= & \hat{G}_{N}^{(k)}(z)_{k, t} \sum_{i=N+1}^{N} G_{N}(z)_{t i} r_{i}(k)  \tag{5.27}\\
& +G_{N}(z)_{t, k} \sum_{j=1}^{N+1} r_{j}(k) \hat{G}_{N}^{(k)}(z)_{j t}  \tag{5.28}\\
= & G_{N}(z)_{t, k} \sum_{j=1}^{N+1} r_{j}(k) \hat{G}_{N}^{(k)}(z)_{j t} \tag{5.29}
\end{align*}
$$

where we have noticed that, for all $t \neq k, \hat{G}_{N}^{(k)}(z)_{k, t}=0$.
Therefore, we find that:

$$
\begin{equation*}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(k)}(z)_{t t}\right|\right] \leqslant \mathbb{E}\left[\left|G_{N}(z)_{t, k}\right|^{2}\right]^{1 / 2} \mathbb{E}\left[\left|\sum_{j=1}^{N} r_{j}(k) \hat{G}_{N}^{(k)}(z)_{j t}\right|^{2}\right]^{1 / 2} \tag{5.30}
\end{equation*}
$$

by Cauchy-Schwartz's inequality. We want to expand the square in the above expression. To that purpose, we first observe that, conditionally to the $M^{i}$, the variables $\left(r_{j}(k)\right)_{j}$ are independent from $\hat{G}_{N}^{(k)}(z)$ and centered. Hence we have for $j \neq j^{\prime}$,

$$
\mathbb{E}\left[r_{j}(k) r_{j^{\prime}}(k) \hat{G}_{N}^{(k)}(z)_{j t} \overline{\hat{G}_{N}^{(k)}(z)_{j^{\prime} t}}\right]=0
$$

Thus we get:

$$
\begin{aligned}
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(k)}(z)_{t t}\right|\right] & \leqslant \mathbb{E}\left[\left|G_{N}(z)_{t, k}\right|^{2}\right]^{1 / 2}\left(\sum_{j=1}^{N+1} \mathbb{E}\left[r_{j}(k)^{2}\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{2}\right]\right)^{1 / 2} \\
& \leqslant \mathbb{E}\left[\left|G_{N}(z)_{t, k}\right|^{2}\right]^{1 / 2}\left(\sum_{j=1}^{N+1} \mathbb{E}\left[r_{j}(k)^{4}\right]^{1 / 2} \mathbb{E}\left[\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{4}\right]^{1 / 2}\right)^{1 / 2} \\
& \leqslant \frac{\mathbb{E}\left[r_{1}(k)^{4}\right]^{1 / 4}}{|\Im(z)|}\left(\sum_{j=1}^{N+1} \mathbb{E}\left[\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{4}\right]^{1 / 2}\right)^{1 / 2} \\
& \leqslant \frac{\mathbb{E}\left[r_{1}(k)^{4}\right]^{1 / 4}}{|\Im(z)|}(N+1)^{1 / 4}\left(\sum_{j=1}^{N+1} \mathbb{E}\left[\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{4}\right]\right)^{1 / 4}
\end{aligned}
$$

Now we use equation (2.5) for the moments of the MRM $M$ to obtain

$$
\mathbb{E}\left[r_{j}(k)^{4}\right]=3 \mathbb{E}\left[M\left[0 ; \frac{1}{N}\right]^{2}\right]=3 \mathbb{E}\left[M[0 ; 1]^{2}\right] \frac{1}{N(2)} .
$$

Furthermore, by using Lemma A. 1 which assures that, almost surely:

$$
\begin{equation*}
\sum_{j=1}^{N+1}\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{2} \leqslant \frac{1}{|\Im(z)|^{2}} \tag{5.31}
\end{equation*}
$$

and the fact that:

$$
\begin{equation*}
\sum_{j=1}^{N+1}\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{4} \leqslant\left(\sum_{j=1}^{N+1}\left|\hat{G}_{N}^{(k)}(z)_{j t}\right|^{2}\right)^{2} \tag{5.32}
\end{equation*}
$$

we finally obtain

$$
\mathbb{E}\left[\left|G_{N}(z)_{t t}-\hat{G}_{N}^{(k)}(z)_{t t}\right|\right] \leqslant \frac{C}{|\Im(z)|^{2}}\left(\frac{1}{N}\right)^{\frac{\varsigma(2)-1}{4}}
$$

It just remains to check that $\zeta(2)=2-\gamma^{2}$.
Lemma 5.7. For each $k \in\{1, \ldots, 2 N\}$, if $t \neq k$, then

$$
\begin{equation*}
G_{N}^{(k)}(z)_{t t}=\hat{G}_{N}^{(k)}(z)_{t t} \tag{5.33}
\end{equation*}
$$

and if $t=k$, then $\hat{G}_{N}^{(k)}(z)_{k k}=z^{-1}$.
Proof. Bloc inversion for the two matrices $\left(A_{N}^{(k)}(z)\right)^{-1}$ and $\left(\hat{A}_{N}^{(k)}(z)\right)^{-1}$ gives the result.
As a consequence of Lemma 5.4, we get the following estimate on $K_{z}$.
Lemma 5.8. For all $z \in \mathbb{C}$ and Lebesgue almost every point $x \in[0,1]$, we have

$$
\begin{equation*}
\Im(z) \Im\left(K_{z}(x)\right) \leqslant 0 \tag{5.34}
\end{equation*}
$$

and

$$
\begin{equation*}
\left|\Im\left(K_{z}(x)\right)\right| \leqslant \frac{1}{\Im(z)} \tag{5.35}
\end{equation*}
$$

### 5.3. Concentration inequalities

This lemma is adapted to our case from Lemma 5.4 in [4].
Lemma 5.9. Let $f:[0,1] \rightarrow \mathbb{R}$ be a bounded measurable function. For each $i \in\{1,2\}$, we have the following concentration results:

$$
\begin{equation*}
\mathbb{E}\left[\left|L_{N}^{i, z}(f)-\mathbb{E}\left[L_{N}^{i, z}(f)\right]\right|^{2}\right] \leqslant \frac{8}{N} \frac{\|f\|_{\infty}^{2}}{|\Im z|^{4}} \tag{5.36}
\end{equation*}
$$

Proof. Define two functions $F_{N}^{1}$ and $F_{N}^{2}$ such that:

$$
\begin{align*}
& F_{N}^{1}\left(\left(X_{i j}^{(N)}\right)_{1 \leqslant j \leqslant N+1}, 1 \leqslant i \leqslant N\right)=\frac{1}{N} \sum_{k=1}^{N} f\left(\frac{k}{N}\right) G_{N}(z)_{k k}  \tag{5.37}\\
& F_{N}^{2}\left(\left(X_{i j}^{(N)}\right)_{1 \leqslant j \leqslant N+1}, 1 \leqslant i \leqslant N\right)=\frac{1}{N} \sum_{k=1}^{N+1} f\left(\frac{k}{N+1}\right) G_{N}(z)_{k+N, k+N} \tag{5.38}
\end{align*}
$$

We will prove the Lemma for $L_{N}^{1, z}$; the proof for $L_{N}^{2, z}$ is a straightforward adaptation.
Let, for $k \in\{1, \ldots, N+1\}$,

$$
\begin{equation*}
\mathcal{F}_{k}=\sigma\left(\left(X_{i j}^{(N)}\right)_{1 \leqslant j \leqslant N}, 1 \leqslant i \leqslant k\right) \tag{5.39}
\end{equation*}
$$

Denote by $P$ the law of the vector $\left(X_{1 j}^{(N)}\right)_{1 \leqslant j \leqslant N}$. With the same proof as in ([4], Lem. 5.4), we obtain

$$
\mathbb{E}\left[\left|F_{N}^{1}-\mathbb{E}\left[F_{N}^{1}\right]\right|^{2}\right] \leqslant \sum_{i=0}^{N} \sup _{\mathbb{R}^{(N+1)^{2}}}\left\|\nabla_{x_{i+1}} F_{N}\right\|^{2} \int\|x-y\|^{2} \mathrm{~d} P^{\otimes 2}(x, y)
$$

The quantity $\nabla_{x_{i+1}} F_{N}^{1}$ refers to the gradient of $F_{N}^{1}$ in the direction of the vector $x_{i+1}$.
If we consider a couple of processes $\left(\widetilde{B}^{1}, \widetilde{M}^{1}\right)$ independent from $\left(B^{1}, M^{1}\right)$ with the same law, it is easy to see that:

$$
\begin{aligned}
\int\|x-y\|^{2} \mathrm{~d} P \otimes \mathrm{~d} P(x, y) & =\sum_{j=1}^{N} \mathbb{E}\left[\left(B_{M^{1}\left(0, \frac{j}{N}\right)}^{1}-B_{M^{1}\left(0, \frac{j-1}{N}\right)}^{1}-\widetilde{B}_{\widetilde{M}^{1}\left(0, \frac{j}{N}\right)}^{1}+\widetilde{B}_{\widetilde{M}^{1}\left(0, \frac{j-1}{N}\right)}^{1}\right)^{2}\right] \\
& =2-2 \sum_{j=1}^{N} \mathbb{E}\left[\left(B_{M^{1}\left(0, \frac{j}{N}\right)}^{1}-B_{M^{1}\left(0, \frac{j-1}{N}\right)}^{1}\right)\left(\widetilde{B}_{\widetilde{M}^{1}\left(0, \frac{j}{N}\right)}^{1} \widetilde{B}_{\widetilde{M}^{1}\left(0, \frac{j-1}{N}\right)}^{1}\right)\right] \\
& =2
\end{aligned}
$$

In our case, we have, for $i \in\{1, \ldots, N+1\}, j \in\{1, \ldots, N\}$ :

$$
\begin{equation*}
\frac{\partial G_{N}(z)_{k k}}{\partial X_{i j}}=G_{N}(z)_{k, j} G_{N}(z)_{N+i, k}+G_{N}(z)_{k, N+i} G_{N}(z)_{j, k} \tag{5.40}
\end{equation*}
$$

Thus,

$$
\begin{equation*}
\nabla_{x_{i+1}} F_{N}=\frac{1}{N} \sum_{k=1}^{N} f\left(\frac{k}{N}\right) \nabla_{x_{i+1}} G_{N}(z)_{k k} \tag{5.41}
\end{equation*}
$$

It is now plain to compute:

$$
\left\|\nabla_{x_{i+1}} F_{N}\right\|^{2}=\frac{1}{N^{2}} \sum_{j=1}^{N}\left|\left(G_{N}(z) D^{1}(f) G_{N}(z)\right)_{N+i+1, j}+\left(G_{N}(z) D^{1}(f) G_{N}(z)\right)_{j, N+i+1}\right|^{2}
$$

where $D^{1}(f)$ is the $(2 N)$-dimensional diagonal matrix of entries:

$$
D^{1}(f)_{k k}=f\left(\frac{k}{N}\right) 1_{\{1 \leqslant k \leqslant N\}} .
$$

One thus has:

$$
\begin{aligned}
\left\|\nabla_{x_{i+1}} F_{N}^{1}\right\|^{2} & =\frac{4}{N^{2}} \sum_{j=1}^{N}\left|\left(G_{N}(z) D^{1}(f) G_{N}(z)\right)_{N+i+1, j}\right|^{2} \\
& \leqslant \frac{4}{N^{2}} \sum_{j=1}^{2 N}\left|\left(G_{N}(z) D^{1}(f) G_{N}(z)\right)_{N+i+1, j}\right|^{2} \\
& \leqslant \frac{4}{N^{2}} \frac{\|f\|_{\infty}^{2}}{|\Im z|^{4}} .
\end{aligned}
$$

where, in the last line, we used Lemma A. 1 and the fact that the matrix $G_{N}(z) D^{1}(f) G_{N}(z)$ has a spectral radius smaller than $\|f\|_{\infty} /|\Im z|^{2}$.

Finally,

$$
\begin{equation*}
\mathbb{E}\left[\left|F_{N}^{1}-\mathbb{E}\left[F_{N}^{1}\right]\right|^{2}\right] \leqslant \frac{8}{N} \frac{\|f\|_{\infty}^{2}}{|\Im z|^{4}} \tag{5.42}
\end{equation*}
$$

We also prove the following lemma:
Lemma 5.10. For all $\alpha>1$ such that $\zeta(2 \alpha)>1$, we have

$$
\begin{equation*}
\mathbb{E}\left[\left|\sum_{t=1}^{N} r_{k}(t)^{2}\left(\hat{G}_{N}^{(N+k)}(z)_{t t}-\mathbb{E}\left[\hat{G}_{N}^{(N+k)}(z)_{t t}\right]\right)\right|\right] \leqslant \frac{C(\ln N)^{2}}{N^{\frac{\zeta(2 \alpha)-1}{\alpha}|\Im(z)|^{4}}} \tag{5.43}
\end{equation*}
$$

for some positive constant $C$ independent from $N, z, k$.
Proof. Notice that $\left(r_{k}(t)\right)_{t}$ and $\hat{G}_{N}^{(N+k)}(z)$ are independent. Hence, by conditioning with respect to the process $\left(r_{k}(t)\right)_{t}$, we can argue along the same lines as in the previous lemma with $r_{k}(t)$ instead of $\frac{1}{N} f\left(\frac{t}{N}\right)$ and we get the formula:

$$
\mathbb{E}\left[\left|\sum_{t=1}^{N} r_{k}(t)^{2}\left(\hat{G}_{N}^{(N+k)}(z)_{t t}-\mathbb{E}\left[\hat{G}_{N}^{(N+k)}(z)_{t t}\right]\right)\right|^{2}\right] \leqslant \frac{8}{|\Im(z)|^{4}} \mathbb{E}\left[\sup _{t} r_{k}(t)^{4}\right] .
$$

We conclude with Proposition B. 1 in the appendix.
In the following, we fix $\alpha>1$ such that $\zeta(2 \alpha)>1$ (because of the expression of $\zeta$ and the inequality $\gamma^{2}<1 / 3$, it is clear that such a number exists).

### 5.4. The system verified by the limit point $\mu_{z}^{2}$ and $K_{z}(x)$ : first equation

From the Schur complement formula (see e.g. Lem. 4.2 in [4] for a reminder), one has for $k \in\{1, \ldots, N\}$ :

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{s, t=1}^{N} r_{k}(s) r_{k}(t) G_{N}^{(N+k)}(z)_{s t}\right]^{-1} \tag{5.44}
\end{equation*}
$$

Using Lemma A.2, one can write:

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} r_{k}(t)^{2} G_{N}^{(N+k)}(z)_{t t}+\varepsilon_{N, k}^{1}(z)\right]^{-1} \tag{5.45}
\end{equation*}
$$

where $\varepsilon_{N, k}^{1}(z)$ is a complex valued random variable for which there exists $C>0$ such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N$,

$$
\begin{equation*}
\mathbb{E}\left[\left|\varepsilon_{N, k}^{1}(z)\right|^{2}\right]<\frac{C}{N^{1-\gamma^{2}}} . \tag{5.46}
\end{equation*}
$$

By using Lemma 5.7, we can write:

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} r_{k}(t)^{2} \hat{G}_{N}^{(N+k)}(z)_{t t}+\varepsilon_{N, k}^{1}(z)\right]^{-1} . \tag{5.47}
\end{equation*}
$$

Lemma 5.10 applied to $\alpha>1$ such that $\zeta(2 \alpha)>1$ yields:

$$
\begin{equation*}
\mathbb{E}\left[\left|\sum_{t=1}^{N} r_{k}(t)^{2}\left(\hat{G}_{N}^{(N+k)}(z)_{t t}-\mathbb{E}\left[\hat{G}_{N}^{(N+k)}(z)_{t t}\right]\right)\right|^{2}\right] \leqslant \frac{C(\ln N)^{2}}{N^{\frac{\zeta(2 \alpha-1}{\alpha}|\Im(z)|^{4}} .} \tag{5.48}
\end{equation*}
$$

Thus, one can write:

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} r_{k}(t)^{2} \mathbb{E}\left[\hat{G}_{N}^{(N+k)}(z)_{t t}\right]+\varepsilon_{N, k}^{1}(z)+\varepsilon_{N, k}^{2}(z)\right]^{-1} \tag{5.49}
\end{equation*}
$$

where $\varepsilon_{N, k}^{2}(z)$ is a complex valued random variable such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N+1$,

$$
\begin{equation*}
\mathbb{E}\left[\left|\varepsilon_{N, k}^{2}(z)\right|^{2}\right]<\frac{C(\ln N)^{2}}{N^{\frac{\zeta(2 \alpha)-1}{\alpha}}|\Im(z)|^{4}} \tag{5.50}
\end{equation*}
$$

In addition, using Lemma 5.5, we can show:

$$
\begin{align*}
& \mathbb{E}\left[\left|\sum_{t=1}^{N} r_{k}(t)^{2}\left(\mathbb{E}\left[\hat{G}_{N}^{(N+k)}(z)_{t t}-G_{N}(z)_{t t}\right]\right)\right|\right]  \tag{5.51}\\
& \leqslant \sum_{t=1}^{N} \mathbb{E}\left[r_{k}(t)^{2}\right] \mathbb{E}\left[\left|\hat{G}_{N}^{(N+k)}(z)_{t t}-G_{N}(z)_{t t}\right|\right]  \tag{5.52}\\
& \leqslant \frac{1}{|\Im(z)|^{2} \sqrt{N}} . \tag{5.53}
\end{align*}
$$

It follows:

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} r_{k}(t)^{2} \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}^{1}(z)+\varepsilon_{N, k}^{2}(z)+\varepsilon_{N, k}^{3}(z)\right]^{-1} \tag{5.54}
\end{equation*}
$$

where $\varepsilon_{N, k}^{3}(z)$ is a complex valued random variable such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N+1$,

$$
\begin{equation*}
\mathbb{E}\left[\left|\varepsilon_{N, k}^{3}(z)\right|\right]<\frac{1}{|\Im(z)|^{2} \sqrt{N}} \tag{5.55}
\end{equation*}
$$

Let us denote by $I_{N}^{t}$ the interval $\left[\frac{t-1}{N}, \frac{t}{N}\right]$. Then we have:
Lemma 5.11. The following inequality holds:

$$
\mathbb{E}\left[\left\lvert\, \sum_{t=1}^{N}\left(r_{k}(t)^{2}-M^{k}\left(I_{N}^{t}\right)\right) \mathbb{E}\left[\left.G_{N}(z)_{t t}\right|^{2}\right] \leqslant \frac{C}{N^{1-\gamma^{2}}|\Im(z)|^{2}}\right.\right.
$$

for some positive constant $C$.
Proof. We expand the square and, because $r_{k}(t)$ and $r_{k}\left(t^{\prime}\right)$ are independent for $t \neq t^{\prime}$ conditionally to $M^{k}$, we have:

$$
\begin{aligned}
\mathbb{E}\left[\left|\sum_{t=1}^{N}\left(r_{k}(t)^{2}-M^{k}\left(I_{N}^{t}\right)\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]\right|^{2}\right] & =2 \sum_{t=1}^{N} \mathbb{E}\left[\left(M^{k}\left(I_{N}^{t}\right)\right)^{2}\right] \mathbb{E}\left[G_{N}(z)_{t t}\right]^{2} \\
& \leqslant 2 C \frac{N}{N^{\zeta(2)}|\Im(z)|^{2}} .
\end{aligned}
$$

We can thus write

$$
\begin{align*}
G_{N}(z)_{N+k, N+k}= & {\left[z-\sum_{t=1}^{N} M^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]\right.}  \tag{5.56}\\
& \left.+\varepsilon_{N, k}^{1}(z)+\varepsilon_{N, k}^{2}(z)+\varepsilon_{N, k}^{3}(z)+\varepsilon_{N, k}^{4}(z)\right]^{-1} \tag{5.57}
\end{align*}
$$

where $\varepsilon_{N, k}^{4}(z)$ is a complex valued random variable such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N+1$,

$$
\begin{equation*}
\mathbb{E}\left[\left|\varepsilon_{N, k}^{4}(z)\right|^{2}\right] \leqslant \frac{C}{N^{\zeta(2)-1}|\Im(z)|^{2}} . \tag{5.58}
\end{equation*}
$$

Set $\varepsilon_{N, k}(z)=\varepsilon_{N, k}^{1}(z)+\varepsilon_{N, k}^{2}(z)+\varepsilon_{N, k}^{3}(z)+\varepsilon_{N, k}^{4}(z)$ and rewrite:

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} M^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}(z)\right]^{-1} \tag{5.59}
\end{equation*}
$$

We now need to introduce the truncated Radon measure $M_{\varepsilon}^{k}(\mathrm{~d} x)$ with Lebesgue density $\mathrm{e}^{\omega_{\varepsilon}^{k}(x)}$ which converges almost surely as $\varepsilon$ goes to 0 , in the sense of weak convergence in the space of Radon measures, to the measure $M^{k}$ (see Sect. 2.1).

Lemma 5.12. For $\varepsilon>0$, the following uniform bound holds:

$$
\sup _{N} \mathbb{E}\left[\left|\sum_{t=1}^{N} M^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]\right|^{2}\right] \leqslant \frac{C \varepsilon^{1-\gamma^{2}}}{|\Im(z)|^{2}}
$$

Proof. We expand the square. Note that the covariance function $\rho_{\varepsilon}$ of the process $\omega_{\varepsilon}$ increases as $\varepsilon$ decreases to 0 and uniformly converges as $\varepsilon \rightarrow 0$ towards $\ln _{+} \frac{\tau}{|x|}$ over the complement of any ball centered at 0 . Thus we
have:

$$
\begin{aligned}
\sup _{N} \mathbb{E}\left[\mid \sum_{t=1}^{N}\right. & M^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[\left.G_{N}(z)_{t t}\right|^{2}\right] \\
& =\sup _{N} \sum_{t, t^{\prime}=1}^{N} \mathbb{E}\left[\left(M^{k}\left(I_{N}^{t}\right)-M_{\varepsilon}^{k}\left(I_{N}^{t}\right)\right)\left(M^{k}\left(I_{N}^{t^{\prime}}\right)-M_{\varepsilon}^{k}\left(I_{N}^{t^{\prime}}\right)\right)\right] \mathbb{E}\left[G_{N}(z)_{t t}\right] \mathbb{E}\left[G_{N}(z)_{t^{\prime} t^{\prime}}\right] \\
& =\sup _{N} \sum_{t, t^{\prime}=1}^{N} \mathbb{E}\left[\left(M^{k}\left(I_{N}^{t}\right)-M_{\varepsilon}^{k}\left(I_{N}^{t}\right)\right)\left(M^{k}\left(I_{N}^{t^{\prime}}\right)-M_{\varepsilon}^{k}\left(I_{N}^{t^{\prime}}\right)\right)\right] \mathbb{E}\left[G_{N}(z)_{t t}\right] \mathbb{E}\left[G_{N}(z)_{t^{\prime} t^{\prime}}\right] \\
& =\sup _{N} \sum_{t, t^{\prime}=1}^{N}\left(\mathbb{E}\left[M^{k}\left(I_{N}^{t}\right) M^{k}\left(I_{N}^{t^{\prime}}\right)\right]-\mathbb{E}\left[M_{\varepsilon}^{k}\left(I_{N}^{t}\right) M_{\varepsilon}^{k}\left(I_{N}^{t^{\prime}}\right)\right]\right) \mathbb{E}\left[G_{N}(z)_{t t}\right] \mathbb{E}\left[G_{N}(z)_{t^{\prime} t^{\prime}}\right] \\
& =\sup _{N} \sum_{t, t^{\prime}=1}^{N} \mathbb{E}\left[G_{N}(z)_{t t}\right] \mathbb{E}\left[G_{N}(z)_{t^{\prime} t^{\prime}}\right] \int_{I_{N}^{t}} \int_{I_{N}^{t^{\prime}}}\left(\mathrm{e}^{\left.\psi(2) \ln n_{+} \frac{\tau}{|r-u|}-\mathrm{e}^{\psi(2) \rho_{\varepsilon}(r-u)}\right) \mathrm{d} r \mathrm{~d} u}\right. \\
& \leqslant \frac{1}{|\Im(z)|^{2}} \int_{0}^{1} \int_{0}^{1}\left(\mathrm{e}^{\left.\psi(2) \ln _{+} \frac{\tau}{|r-u|}-\mathrm{e}^{\psi(2) \rho_{\varepsilon}(r-u)}\right) \mathrm{d} r \mathrm{~d} u .}\right.
\end{aligned}
$$

where, in the fourth line, we used the fact that, if $\mathcal{F}_{\varepsilon}$ is the sigma field generated by the random variables $\mu(A), A \in \mathcal{B}(\{(t, y): y \geqslant \varepsilon\})$, then $\mathbb{E}\left[M^{k}(A) \mid \mathcal{F}_{\varepsilon}\right]=M_{\varepsilon}^{k}(A)$ for all borelian set $A$. A straightforward computation leads to the relation

$$
\rho_{\varepsilon}(t)= \begin{cases}\ln \frac{\tau}{\varepsilon}+1-\frac{|t|}{\varepsilon} & \text { if }|t| \leqslant \varepsilon  \tag{5.60}\\ \ln \frac{\tau}{|t|} & \text { if } \varepsilon \leqslant|t| \leqslant \tau \\ 0 & \text { if } \tau<|t|\end{cases}
$$

By using the expression of $\rho_{\varepsilon}$, it is then plain to obtain the desired bound.
We can thus write

$$
\begin{equation*}
G_{N}(z)_{N+k, N+k}=\left[z-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}(z)+\delta(\varepsilon, N, z)\right]^{-1} \tag{5.61}
\end{equation*}
$$

where

$$
\begin{equation*}
\sup _{N} \mathbb{E}\left[|\delta(\varepsilon, N, z)|^{2}\right] \rightarrow 0 \quad \text { as } \varepsilon \rightarrow 0 \tag{5.62}
\end{equation*}
$$

and also:

$$
\begin{equation*}
\mathbb{E}\left[G_{N}(z)_{N+k, N+k}\right]=\mathbb{E}\left[\left[z-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}(z)+\delta(\varepsilon, N, z)\right]^{-1}\right] \tag{5.63}
\end{equation*}
$$

The next step is to study the convergence of the above quantity. Hence we prove (see the proof in the appendix).
Lemma 5.13. The random variable $\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]$ converges in probability as $N \rightarrow+\infty$ towards $\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x)$.

We fix $\varepsilon>0$. For that $\varepsilon$, the family of random variables $(\delta(\varepsilon, N, z))_{N}$ is bounded in $L^{2}$ so that it is tight. Even if it means extracting again a subsequence we assume that the couple $\left(\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right], \delta(\varepsilon, N, z)\right)_{N}$ converges in law towards the couple $\left(\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x), Y_{\varepsilon}\right)$. We remind the reader of (5.44) which implies that

$$
\left|\left(z-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}(z)+\delta(\varepsilon, N, z)\right)^{-1}\right| \leqslant \frac{1}{|\Im(z)|}
$$

The quantity $\left(z-\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\varepsilon_{N, k}(z)+\delta(\varepsilon, N, z)\right)^{-1}$ is therefore bounded uniformly with respect to $N, \varepsilon$ and converges in law towards

$$
\left(z-\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x)+Y_{\varepsilon}\right)^{-1}
$$

We deduce that the expectation of the former quantity converges as $\varepsilon \rightarrow 0$ towards the expectation of the latter quantity. From (5.63), we deduce that

$$
\begin{equation*}
\mu_{z}^{2}=\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x)+Y_{\varepsilon}\right)^{-1}\right] \tag{5.64}
\end{equation*}
$$

Clearly, standard arguments prove that $\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x)$ converges almost surely towards $\int_{0}^{1} K_{z}(x) M^{k}(\mathrm{~d} x)$ as $\varepsilon \rightarrow 0$ ( $K_{z}$ is deterministic (see Lem. 5.9), measurable and bounded) and, because of (5.62), $Y_{\varepsilon}$ converges almost surely towards 0 as $\varepsilon \rightarrow 0$. Again, because the quantity $\left(z-\int_{0}^{1} K_{z}(x) M_{\varepsilon}^{k}(\mathrm{~d} x)+Y_{\varepsilon}\right)^{-1}$ is bounded uniformly with respect to $\varepsilon$, we deduce that:

$$
\begin{equation*}
\mu_{z}^{2}=\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(x) M^{k}(\mathrm{~d} x)\right)^{-1}\right] \tag{5.65}
\end{equation*}
$$

### 5.5. Second equation

Now we turn our attention to the terms $G_{N}(z)_{k k}$ for $k \in\{1, \ldots, N\}$. Again, by using the Schur complement formula, we can write, for $k \in\{1, \ldots, N\}$ :

$$
\begin{align*}
G_{N}(z)_{k k} & =\left[z-\sum_{i, j=1}^{N} r_{i}(k) r_{j}(k) G_{N}^{(k)}(z)_{N+i, N+j}\right]^{-1}  \tag{5.66}\\
& =\left[z-\sum_{i=1}^{N} r_{i}(k)^{2} G_{N}^{(k)}(z)_{N+i, N+i}+\eta_{N, k}^{1}(z)\right]^{-1} \tag{5.67}
\end{align*}
$$

where, using Lemma A.3, $\eta_{N, k}^{1}(z)$ is a complex valued random variable for which there exists $c>0$ such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N, \mathbb{E}\left[\left|\eta_{N, k}^{1}(z)\right|^{2}\right]<c / N$.

With a further use of the Schur complement formula for the term $G_{N}^{(k)}(z)_{N+i, N+i}$, we obtain:

$$
\begin{equation*}
G_{N}(z)_{k k}=\left[z-\sum_{i=1}^{N} r_{i}(k)^{2}\left[z-\sum_{s, t \neq k}^{N} r_{i}(s) r_{i}(t) G_{N}^{(k, N+i)}(z)_{s t}\right]^{-1}+\eta_{N, k}^{1}(z)\right]^{-1} \tag{5.68}
\end{equation*}
$$

where $G_{N}^{(k, N+i)}(z)=A_{N}^{(k, N+i)}(z)^{-1}$. Note that $G_{N}^{(k, N+i)}(z)$ is independent of $\left(r_{i}(t)\right)_{t=1, \ldots, N}$. Using the same arguments as in the derivation of the first equation (in particular Lems. A.2, 5.7, 5.10, B.1, 5.6 and 5.5), one can show that:

$$
\begin{equation*}
G_{N}(z)_{k k}=\left[z-\sum_{i=1}^{N} \frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\delta_{N, k, i}(z)}+\eta_{N, k}^{1}(z)\right]^{-1} \tag{5.69}
\end{equation*}
$$

where $\left(\delta_{N, k, i}(z)\right)_{1 \leqslant i \leqslant N}$ are complex random variable such that

$$
\begin{equation*}
\mathbb{E}\left[\left|\delta_{N, k, i}(z)\right|\right] \leqslant \frac{C}{N^{\min \left(\frac{1-\gamma^{2}}{4}, \frac{\zeta(2 \alpha)-1}{\alpha}\right)}} \tag{5.70}
\end{equation*}
$$

for some positive constant $C$ that does not depend on $i, N$ and for $\alpha>1$ such that $\zeta(2 \alpha)>1$.

Lemma 5.14. One can write:

$$
\begin{equation*}
G_{N}(z)_{k k}=\left[z-\sum_{i=1}^{N} \frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}+\eta_{N, k}^{1}(z)+\eta_{N, k}^{2}(z)\right]^{-1} \tag{5.71}
\end{equation*}
$$

where $\eta_{N, k}^{2}(z)$ is a random variable that tends to 0 in probability as $N$ goes to $\infty$.
Proof. By using Lemma 5.4, we deduce that:

$$
\begin{align*}
& \sum_{i=1}^{N}\left|\frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]+\delta_{N, k, i}(z)}-\frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}\right| \\
& \quad \leqslant \frac{1}{|\Im(z)|^{2}} \sum_{i=1}^{N} r_{i}(k)^{2} \min \left(\left|\delta_{N, k, i}(z)\right|, 2\right) \tag{5.72}
\end{align*}
$$

We stress that the lemma is proved as soon as we can prove that the left-hand side in (5.72) converges in probability to 0 . Hence it is enough to prove that

$$
\mathbb{E}\left[\sum_{i=1}^{N} r_{i}(k)^{2} \min \left(\left|\delta_{N, k, i}(z)\right|, 2\right)\right]
$$

converges to 0 as $N$ tends to $\infty$. By noticing that:

$$
\begin{equation*}
\delta_{N, k, i}(z)=\sum_{s, t \neq k}^{N} r_{i}(s) r_{i}(t) G_{N}^{(k, N+i)}(z)_{s t}-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right] \tag{5.73}
\end{equation*}
$$

it is straightforward to see that the variables $\left(r_{i}(k)^{2} \min \left(\left|\delta_{N, k, i}(z)\right|, 2\right)\right)_{1 \leqslant i \leqslant N+1}$ are identically distributed. Thus we have

$$
\mathbb{E}\left[\sum_{i=1}^{N} r_{i}(k)^{2} \min \left(\left|\delta_{N, k, i}(z)\right|, 2\right)\right]=N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right)\right]
$$

Then for all $A>1$ and $\alpha>0$, we have

$$
\begin{aligned}
N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right)\right]= & N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right) \mathbb{1}_{\left\{N r_{1}(k)^{2} \leqslant A\right\}}\right] \\
& +N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right) \mathbb{1}_{\left\{N r_{1}(k)^{2}>A\right\}}\right] \\
\leqslant & A \mathbb{E}\left[\left|\delta_{N, k, 1}(z)\right|\right]+2 \mathbb{E}\left[N r_{1}(k)^{2} \mathbb{1}_{\left\{N r_{1}(k)^{2}>A\right\}}\right] \\
\leqslant & \frac{A C}{N^{\frac{\zeta(2)-1}{4}}}+\frac{2}{A^{\alpha}} \mathbb{E}\left[N^{1+\alpha} r_{1}(k)^{2(\alpha+1)}\right] \\
= & \frac{A C}{N^{\frac{\zeta(2)-1}{4}}}+\frac{2 N^{1+\alpha}}{A^{\alpha}} \mathbb{E}\left[M^{1}\left(0, \frac{1}{N}\right)^{\alpha+1}\right]
\end{aligned}
$$

By using the scale invariance property of the measure $M^{1}$, we have:

$$
\mathbb{E}\left[M^{1}(0,1 / N)^{\alpha+1}\right]=\frac{1}{N^{\zeta(1+\alpha)}} \mathbb{E}\left[M^{1}(0,1)^{\alpha+1}\right]
$$

in such a way that

$$
\begin{equation*}
N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right)\right] \leqslant \frac{A C}{N^{\frac{\zeta(2)-1}{4}}}+2 \mathbb{E}\left[M^{1}(0,1)^{\alpha+1}\right] \frac{N^{\psi(1+\alpha)}}{A^{\alpha}} \tag{5.74}
\end{equation*}
$$

Since $\zeta(2)>5-4 \zeta^{\prime}(1)$, we can choose $p>0$ such that

$$
\begin{equation*}
\frac{\zeta(2)-1}{4}>p>1-\zeta^{\prime}(1)=\psi^{\prime}(1) \tag{5.75}
\end{equation*}
$$

The mapping $\alpha \in] 0,+\infty\left[\mapsto p \alpha-\psi(1+\alpha)\right.$ reduces to 0 for $\alpha=0$ and, because $p>\psi^{\prime}(1)$, is strictly positive for $\alpha>0$ small enough. So we choose $\alpha<1$ such that $p \alpha-\psi(1+\alpha)>0$ and we set $A=N^{p}$. We obtain:

$$
N \mathbb{E}\left[r_{1}(k)^{2} \min \left(\left|\delta_{N, k, 1}(z)\right|, 2\right)\right] \leqslant \frac{C}{N^{\frac{\zeta(2)-1}{4}-p}}+2^{2+\alpha} \mathbb{E}\left[M^{1}(0, T)^{\alpha+1}\right] \frac{1}{N^{\alpha p-\psi(1+\alpha)}}
$$

The result follows by letting $N \rightarrow \infty$ since $\min ((\zeta(2)-1) / 4-p, \alpha p-\psi(1+\alpha))>0$.
Lemma 5.15. There exists a constant $c>0$, which does not depend on $N$, such that for each $N \in \mathbb{N}$ :

$$
\mathbb{E}\left[\left\lvert\, \sum_{i=1}^{N}\left(\frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}-\mathbb{E}\left[\frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}\right]\right)^{2}\right.\right] \leqslant \frac{c}{N^{1-\gamma^{2}}}
$$

Proof. The proof is straightforward using the fact that for $i \in\{1, \ldots, N\}$, the random variables

$$
\begin{equation*}
\frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]} \tag{5.76}
\end{equation*}
$$

are i.i.d. random variables and Lemma 5.4.
Therefore we can write

$$
\begin{equation*}
G_{N}(z)_{k k}=\left[z-\sum_{i=1}^{N} \frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}+\eta_{N, k}^{1}(z)+\eta_{N, k}^{2}(z)+\eta_{N, k}^{3}(z)\right]^{-1} \tag{5.77}
\end{equation*}
$$

with $\mathbb{E}\left[\left(\eta_{N, k}^{3}(z)\right)^{2}\right] \leqslant \frac{c}{N^{1-\gamma^{2}}}$.
Now we can take the expectation in (5.77) to obtain

$$
\begin{aligned}
\mathbb{E}\left[L_{N}^{1, z}(f)\right] & =\frac{1}{N} \sum_{k=1}^{N} f(k / N) \mathbb{E}\left[G_{N}(z)_{k k}\right] \\
& =\frac{1}{N} \sum_{k=1}^{N} f(k / N) \mathbb{E}\left[\left(z-\mathbb{E}\left[\sum_{i=1}^{N} \frac{r_{i}(k)^{2}}{z-\sum_{t=1}^{N} M^{i}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}\right]+\eta_{N, k}(z)\right)^{-1}\right] \\
& =\frac{1}{N} \sum_{k=1}^{N} f(k / N) \mathbb{E}\left[\left(z-N \mathbb{E}\left[\frac{M\left[\frac{k-1}{N} ; \frac{k}{N}\right]}{z-\sum_{t=1}^{N} M\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}\right]+\eta_{N, k}(z)\right)^{-1}\right]
\end{aligned}
$$

with $\eta_{N, k}(z)=\eta_{N, k}^{1}(z)+\eta_{N}^{2}(z)+\eta_{N, k}^{3}(z)$. Then, by introducing the truncated measure $M_{\varepsilon}$ and by using the Girsanov formula, we can approximate (uniformly in $N$ ) this last expression by:

$$
\begin{equation*}
\frac{1}{N} \sum_{k=1}^{N} f(k / N) \mathbb{E}\left[\left(z-N \mathbb{E}\left[\frac{M_{\varepsilon}\left[\frac{k-1}{N} ; \frac{k}{N}\right]}{z-\sum_{t=1}^{N} M_{\varepsilon}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]}\right]\right)^{-1}+\hat{\delta}(N, k, z, \varepsilon)\right] \tag{5.78}
\end{equation*}
$$

with $\sup _{N, k} \mathbb{E}\left[|\hat{\delta}(N, k, z, \varepsilon)|^{2}\right]$ going to 0 when $\varepsilon$ is going to 0 . Along some appropriate subsequence, this latter quantity converges as $N \rightarrow+\infty$ to:

$$
\begin{equation*}
\int_{0}^{1} f(x) \mathbb{E}\left[\left(z-\mathbb{E}\left[\frac{\mathrm{e}^{\omega_{\varepsilon}(x)}}{z-\int_{0}^{1} K_{z}(r) M_{\varepsilon}(\mathrm{d} r)}\right]\right)^{-1}+Y^{\varepsilon}\right] \mathrm{d} x \tag{5.79}
\end{equation*}
$$

where $Y^{\varepsilon}$ is such that $\mathbb{E}\left[\left(Y^{\varepsilon}\right)^{2}\right]$ converges to 0 when $\varepsilon$ is going to 0 . And, we thus obtain gathering the above arguments that:

$$
\begin{equation*}
\int_{0}^{1} f(x) K_{z}(x) \mathrm{d} x=\int_{0}^{1} f(x) \mathbb{E}\left[\left(z-\mathbb{E}\left[\frac{\mathrm{e}^{\omega_{\varepsilon}(x)}}{z-\int_{0}^{1} K_{z}(r) M_{\varepsilon}(\mathrm{d} r)}\right]\right)^{-1}+Y^{\varepsilon}\right] \mathrm{d} x \tag{5.80}
\end{equation*}
$$

It remains to pass to the limit as $\varepsilon \rightarrow 0$ in that expression. This job is carried out with the help of a Girsanov type transform in Appendix 5.7.

### 5.6. Uniqueness of the solution to the system of equations

Let $X$ be the space of bounded measurable functions $[0,1] \rightarrow \mathbb{C}$ endowed with the uniform norm defined for $f \in X$ by:

$$
\begin{equation*}
\|f\|_{\infty}=\sup _{x \in[0,1]}|f(x)| \tag{5.81}
\end{equation*}
$$

Define the operator $T: X \rightarrow X$ by setting, for $g \in X$ and for all $x \in[0,1]$ :

$$
\begin{equation*}
\operatorname{Tg}(x)=\frac{1}{z-q \mathbb{E}\left[\left(z-\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} g(t) M(\mathrm{~d} t)\right)^{-1}\right]} \tag{5.82}
\end{equation*}
$$

For $g, h \in X$ and for all $x \in[0,1]$, we have:

$$
\begin{aligned}
|T g(x)-T h(x)| & \leqslant \frac{q}{|\Im(z)|^{4}} \mathbb{E}\left[\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}}|g(t)-h(t)| M(\mathrm{~d} t)\right] \\
& \leqslant \frac{q}{|\Im(z)|^{4}} \mathbb{E}\left[\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} M(\mathrm{~d} t)\right]\|g-h\|_{\infty} \\
& \leqslant \frac{q}{|\Im(z)|^{4}} \int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} \mathrm{~d} t| | g-h \|_{\infty}
\end{aligned}
$$

Recall that $\gamma^{2}<1 / 3$, and thus it is easy to see that:

$$
\begin{equation*}
\sup _{x \in[0,1]} \int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} \mathrm{~d} t<+\infty \tag{5.83}
\end{equation*}
$$

And we can deduce that there exists a positive constant $C$ such that:

$$
\begin{equation*}
\sup _{x \in[0,1]}|T g(x)-T h(x)| \leqslant \frac{C}{|\Im(z)|^{4}}\|g-h\|_{\infty} \tag{5.84}
\end{equation*}
$$

If $z$ is such that $C /|\Im(z)|^{4}<1$, the operator $T$ is contracting and thus has a unique fixed point $g$ in the Banach space $X$. We conclude that, for each $z$ with $|\Im(z)|$ large enough, there exists a unique bounded function $K_{z}:[0,1] \rightarrow \mathbb{C}$ such that for all $x \in[0,1]$ :

$$
\begin{equation*}
K_{z}(x)=\frac{1}{z-q \mathbb{E}\left[\left(z-\int_{0}^{1}\left(\frac{\tau}{|t-x|}\right)_{+}^{\gamma^{2}} K_{z}(t) M(\mathrm{~d} t)\right)^{-1}\right]} \tag{5.85}
\end{equation*}
$$

Using the first equation, it is now plain to see that, for $z$ such that $C /|\Im(z)|^{4}<1$, the constant $\mu_{z}^{2}$ is uniquely defined by the system of equations (by the first equation, it is a function of the function $K_{z}$, which is uniquely defined for such $z$ ).

Now it remains to show that the limit point $\mu_{z}^{2}$ is uniquely defined for all $z \in \mathbb{C} \backslash \mathbb{R}$. It will be easy to see using analyticity arguments. Indeed, from the Montel theorem, every limit point $\mu_{z}^{2}$ is holomorphic on the set $\mathbb{C} \backslash \mathbb{R}$ since it is the pointwise limit of a subsequence of the sequence of holomorphic functions $L_{N}^{1, z}([0,1])$ that are uniformly bounded on each compact set of $\mathbb{C} \backslash \mathbb{R}$ (see Lem. 5.4). Thus, $\mu_{z}^{2}$ is uniquely defined for each $z \in \mathbb{C} \backslash \mathbb{R}$ by analytic extension (we have just seen that $\mu_{z}^{2}$ is uniquely defined for a set of $z$ with accumulation points).

The same argument holds for the unicity of the integral $\int_{0}^{1} K_{z}(x) \mathrm{d} x$. Indeed, every limit point $\int_{0}^{1} K_{z}(x) \mathrm{d} x$ is a holomorphic function on $\mathbb{C} \backslash \mathbb{R}$ that has some prescribed value on the set $\left\{z \in \mathbb{C} \backslash \mathbb{R}: C /|\Im(z)|^{4}<1\right\}$, which has accumulation points.

### 5.7. Proof of Theorems 3.2, 3.3 and 3.4

Let us gather the above arguments to prove the main theorems.
Proof of theorem 3.4: it is a direct consequence of Sections 5.4-5.6.
Proof of theorem 3.3 (i): The limit points $K_{z}(x) \mathrm{d} x$ and $\mu_{z}^{2} \mathrm{~d} x$ of the two complex measures $\mathbb{E}\left[L_{N}^{1, z}\right]$ and $\mathbb{E}\left[L_{N}^{2, z}\right]$ are uniquely defined because $\mu_{z}^{2}$ and $K_{z}(x)$ satisfy a fixed point system of equations (we have just seen this in Thm. 3.4).

Proof of theorem 3.3 (iii): We need to prove that $\mu_{z}^{2}$ is the Stieltjes transform of a probability measure $v$. From [9], it suffices to prove that $\mu_{z}^{2}$ is holomorphic over $\mathbb{C} \backslash \mathbb{R}$, maps $\{z \in \mathbb{C} \backslash \mathbb{R} ; \Im(z)<0\}$ to $\{z \in \mathbb{C} \backslash \mathbb{R} ; \Im(z)>0\}$ and that $\lim _{y \rightarrow \infty} i y \mu_{i y}^{2}=1(y \in \mathbb{R})$. Let us check those properties. We have already seen in Section 5.6 that $\mu_{z}^{2}$ is holomorphic. From Lemma 5.4, $\mu^{2} \operatorname{maps}\{z \in \mathbb{C} \backslash \mathbb{R} ; \Im(z)<0\}$ to $\{z \in \mathbb{C} \backslash \mathbb{R} ; \Im(z)>0\}$. Finally, from Theorem 3.4, we have

$$
z \mu_{z}^{2}=\mathbb{E}\left[\frac{1}{1-z^{-1} \int_{0}^{1} K_{z}(x) M(\mathrm{~d} x)}\right]
$$

As $\left|K_{z}(x)\right| \leqslant|\Im(z)|^{-1}$, the term $\int_{0}^{1} K_{z}(x) M(\mathrm{~d} x) / z$ converges pointwise towards 0 when $z=i y$ and $y \rightarrow \infty$. Furthermore, from Lemma 5.8, we have $\Im(z) \Im\left(K_{z}(x) \leqslant 0\right.$ in such a way that $\left|z-\int_{0}^{1} K_{z}(x) M(\mathrm{~d} x)\right|^{-1} \leqslant|\Im(z)|^{-1}$. Therefore

$$
\left|\frac{z}{z-\int_{0}^{1} K_{z}(x) M(\mathrm{~d} x)}\right| \leqslant 1
$$

when $z$ takes on the form $z=i y(y \in \mathbb{R})$. The dominated convergence theorem then implies that $\lim _{y \rightarrow \infty} i y \mu_{i y}^{2}=$ 1 and we can conclude $\mu^{2}$ is indeed the Stieltjes transform of a (unique) probability measure $v$.

Proof of theorem 3.2 (i) and 3.3 (ii): We observe that, for $z \in \mathbb{C} \backslash \mathbb{R}$ :

$$
A_{N}(z)\left(\begin{array}{cc}
z I_{T} & 0  \tag{5.86}\\
X_{N} & z I_{N}
\end{array}\right)=\left(\begin{array}{cc}
z^{2} I_{T}-{ }^{t} X_{N} X_{N}-z^{t} X_{N} \\
0 & z^{2} I_{N}
\end{array}\right)
$$

Let us rewrite the matrix $G_{N}(z)=A_{N}(z)^{-1}$ under the form:

$$
G_{N}(z)=\left(\begin{array}{cc}
G_{1}(z) & { }^{t} G_{1,2}(z)  \tag{5.87}\\
G_{1,2}(z) & G_{2}(z)
\end{array}\right)
$$

where $G_{1}(z), G_{1,2}(z), G_{2}(z)$ are respectively of size $T \times T, N \times T, N \times N$.

By taking the inverse in the relation (5.86), we obtain:

$$
\left(\begin{array}{cc}
I_{T} / z & 0  \tag{5.88}\\
-X_{N} / z^{2} & I_{N} / z
\end{array}\right)\left(\begin{array}{cc}
G_{1}(z) & { }^{t} G_{1,2}(z) \\
G_{1,2}(z) & G_{2}(z)
\end{array}\right)=\left(\begin{array}{cc}
\left(z^{2} I_{T}-{ }^{t} X_{N} X_{N}\right)^{-1} & B \\
0 & I_{N} / z^{2}
\end{array}\right)
$$

where $B=\left(z^{2} I_{T}-{ }^{t} X_{N} X_{N}\right)^{-1 t} X_{N} / z$.
It can be rewritten, using the fact that $-X_{N} G_{1}(z)+z G_{1,2}(z)=0$ and $-X_{N}{ }^{t} G_{1,2}(z)+z G_{2}(z)=I_{N}$, as:

$$
\left(\begin{array}{cc}
G_{1}(z) / z^{t} G_{1,2}(z) / z  \tag{5.89}\\
0 & I_{N} / z^{2}
\end{array}\right)=\left(\begin{array}{cc}
\left(z^{2} I_{T}-{ }^{t} X_{N} X_{N}\right)^{-1} & B \\
0 & I_{N} / z^{2}
\end{array}\right) .
$$

Therefore, taking the trace we get:

$$
\begin{equation*}
\frac{1}{T z} \sum_{k=1}^{T} G_{N}(z)_{k k}=\frac{1}{T} \operatorname{tr}\left(z^{2} I_{T}-{ }^{t} X_{N} X_{N}\right)^{-1} \tag{5.90}
\end{equation*}
$$

and, by using the fact that the eigenvalues of ${ }^{t} X_{N} X_{N}$ are those of $X_{N}{ }^{t} X_{N}$ augmented with $T-N$ zeros:

$$
\begin{equation*}
\frac{1}{T z} \sum_{k=1}^{T} G_{N}(z)_{k k}=\frac{1}{T} \operatorname{tr}\left(z^{2} I_{N}-X_{N}^{t} X_{N}\right)^{-1}+\frac{T-N}{T z^{2}} \tag{5.91}
\end{equation*}
$$

Now, taking expectation and using Theorem 3.3, we deduce:

$$
\begin{equation*}
\int_{0}^{1} K_{z}(x) \mathrm{d} x=q z \lim _{N \rightarrow \infty} \frac{1}{N} \mathbb{E}\left[\operatorname{tr}\left(z^{2} I_{N}-X_{N}^{t} X_{N}\right)^{-1}\right]+\frac{1-q}{z} \tag{5.92}
\end{equation*}
$$

Using the fact that (by (2.7)) the spectrum of $B_{N}$ contains $2 N$ eigenvalues which are the positive and negative square-roots of the spectrum of $R_{N}={ }^{t} X_{N} X_{N}$ plus $T-N$ zero eigenvalues and the fact that $1 /(z-\lambda)+1 /(z+\lambda)=$ $2 z /\left(z^{2}-\lambda^{2}\right)$, we can see that:

$$
\begin{equation*}
\frac{1}{N+T} \sum_{k=1}^{N+T} G_{N}(z)_{k k}=\frac{2 z}{N+T} \operatorname{tr}\left(z^{2} I_{N}-X_{N}{ }^{t} X_{N}\right)^{-1}+\frac{T-N}{T+N} \frac{1}{z} \tag{5.93}
\end{equation*}
$$

Using the relation 2.14 and Theorem 3.3, it is easy to see that:

$$
\begin{equation*}
\lim _{N \rightarrow+\infty} \frac{1}{N+T} \sum_{k=1}^{N+T} \mathbb{E}\left[G_{N}(z)_{k k}\right]=\frac{1}{1+q}\left(q \mu_{z}^{2}+\int_{0}^{1} K_{z}(x) \mathrm{d} x\right) \tag{5.94}
\end{equation*}
$$

Taking expectation in 5.93 and using (5.94), we get:

$$
\begin{align*}
\frac{1}{1+q}\left(q \mu_{z}^{2}+\int_{0}^{1} K_{z}(x) \mathrm{d} x\right)= & \frac{2 q z}{1+q} \lim _{N \rightarrow \infty} \frac{1}{N} \mathbb{E}\left[\operatorname{tr}\left(z^{2} I_{N}-X_{N}^{t} X_{N}\right)^{-1}\right]  \tag{5.95}\\
& +\frac{1-q}{1+q} \frac{1}{z} \tag{5.96}
\end{align*}
$$

From equations (5.92) and (5.95), we get the following relation:

$$
\begin{equation*}
\int_{0}^{1} K_{z}(x) \mathrm{d} x=q \mu_{z}^{2}+\frac{1-q}{z} \tag{5.97}
\end{equation*}
$$

and Theorem 3.3 (ii). is proved.

With (5.97), (5.94) becomes:

$$
\begin{equation*}
\lim _{N \rightarrow+\infty} \frac{1}{N+T} \sum_{k=1}^{N+T} \mathbb{E}\left[G_{N}(z)_{k k}\right]=\frac{1}{1+q}\left(2 q \mu_{z}^{2}+\frac{1-q}{z}\right) \tag{5.98}
\end{equation*}
$$

and, we note that the right hand side of (5.98) is the Stieltjes transform of the measure $2 q /(1+q) v(\mathrm{~d} x)+(1-$ $q) /(1+q) \delta_{0}(\mathrm{~d} x)$. Thus, the mean spectral measure $\mathbb{E}\left[\mu_{B_{N}}\right]$ converges weakly to the measure $2 q /(1+q) v(\mathrm{~d} x)+$ $(1-q) /(1+q) \delta_{0}(\mathrm{~d} x)$.

We have also:

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \frac{1}{N} \mathbb{E}\left[\operatorname{tr}\left(z^{2} I_{N}-X_{N}^{t} X_{N}\right)^{-1}\right]=\frac{\mu_{z}^{2}}{z} \tag{5.99}
\end{equation*}
$$

Again using the fact that, for all $x \in \mathbb{R}, 1 /\left(z^{2}-x^{2}\right)=(1 /(z-x)+1 /(z+x)) /(2 z)$ and the fact that $v(\mathrm{~d} x)$ is a symmetric measure on $\mathbb{R}\left(v(\mathrm{~d} x)\right.$ is the weak limit of $\mathbb{E}\left[\mu_{B_{N}}\right]$, which is symmetric since the spectrum of $B_{N}$ is symmetric with respect to 0 almost surely), we see that:

$$
\begin{align*}
\lim _{N \rightarrow \infty} \frac{1}{N} \mathbb{E}\left[\operatorname{tr}\left(z^{2} I_{N}-X_{N}^{t} X_{N}\right)^{-1}\right] & =\frac{1}{z} \int_{\mathbb{R}} \frac{v(\mathrm{~d} x)}{z-x}  \tag{5.100}\\
& =\int_{\mathbb{R}} \frac{v \circ\left(x^{2}\right)^{-1}(\mathrm{~d} x)}{z^{2}-x} . \tag{5.101}
\end{align*}
$$

This implies that, for each $z \in \mathbb{C} \backslash \mathbb{R}$,

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \frac{1}{N} \mathbb{E}\left[\operatorname{tr}\left(z I_{N}-X_{N}^{t} X_{N}\right)^{-1}\right]=\int_{\mathbb{R}} \frac{v \circ\left(x^{2}\right)^{-1}(\mathrm{~d} x)}{z-x} \tag{5.102}
\end{equation*}
$$

and thus, the probability measure $\mathbb{E}\left[\mu_{R_{N}}\right]$ converges weakly to the measure $v \circ\left(x^{2}\right)^{-1}(\mathrm{~d} x)$.
Proof of theorem 3.2 (ii): using relation (2.14) and Lemma 5.9, it is plain to check that $\int_{\mathbb{R}}(z-x)^{-1} \mu_{B_{N}}(\mathrm{~d} x)$ converges in probability to the Stieltjes transform of the probability measure $2 q /(1+q) v(\mathrm{~d} x)+(1-q) /(1+$ $q) \delta_{0}(\mathrm{~d} x)$. This convergence holds for finite dimensional vectors $\left.\left(\int_{\mathbb{R}}\left(z_{i}-x\right)^{-1} \mu_{B_{N}}(\mathrm{~d} x)\right), i=1, \ldots, d\right)$ as well. Using the fact that the set of functions $\left\{(z-x)^{-1}, z \in \mathbb{C} \backslash \mathbb{R}\right\}$ is dense in the set $C_{0}(\mathbb{R})$ of continuous functions on $\mathbb{R}$ going to 0 at infinity, we can show, for each $f \in C_{0}(\mathbb{R})$, that $\int f(x) \mu_{B_{N}}(\mathrm{~d} x)$ converges in probability to $\int f(x)\left(2 q /(1+q) v(\mathrm{~d} x)+(1-q) /(1+q) \delta_{0}(\mathrm{~d} x)\right)$. But, since $\mu_{B_{N}}(\mathbb{R})=2 q /(1+q) v(\mathbb{R})+(1-q) /(1+q) \delta_{0}(\mathbb{R})=1$, this vague convergence can be strengthened in a weak convergence. With the relations $\mu_{B_{N}^{2}}=2 N /(N+T) \mu_{R_{N}}+$ $(T-N) /(T+N) \delta_{0}$ and the fact that $\int f(x) \mu_{B_{N}^{2}}(\mathrm{~d} x)=\int f\left(x^{2}\right) \mu_{B_{N}}(\mathrm{~d} x)$, it is plain to conclude that $\mu_{R_{N}}$ converges weakly in probability to $v \circ\left(x^{2}\right)^{-1}(\mathrm{~d} x)$.
Proof of Theorem 3.2 (iii): again using relation (2.14) and lemma 5.9 together with Borel-Cantelli's lemma, one can show that the two spectral measures $\mu_{B_{N_{k}}}$ converges weakly almost surely to $2 q /(1+q) v(\mathrm{~d} x)+(1-q) /(1+$ $q) \delta_{0}(\mathrm{~d} x)$. It is then easy to deduce as before that $\mu_{R_{N_{k}}}$ converges weakly almost surely to $v \circ\left(x^{2}\right)^{-1}(\mathrm{~d} x)$.

## Appendix A. Auxiliary lemmas

Lemma A.1. Let $A$ be a $n \times n$ complex matrix such that the Hermitian matrix $M=A \bar{A}^{T}$ has spectral radius $\lambda_{\max }$. Then, for all $i$, we have:

$$
\begin{equation*}
\sum_{j=1}^{n}\left|A_{i j}\right|^{2} \leqslant \lambda_{\max } \tag{A.1}
\end{equation*}
$$

Proof. It is straightforward to see that all the entries of $M$ are, in modulus, smaller than $\lambda_{\max }$. On the other hand, we have:

$$
M_{i i}=\sum_{j=1}^{n}\left|A_{i j}\right|^{2}
$$

and, thus:

$$
\begin{equation*}
\sum_{j=1}^{n}\left|A_{i j}\right|^{2} \leqslant \lambda_{\max } \tag{A.2}
\end{equation*}
$$

Lemma A.2. There exists $C>0$ such that for each $N \in \mathbb{N}$ and $k \in\{1, \ldots, N\}$ :

$$
\mathbb{E}\left[\left|\sum_{s \neq t}^{N} r_{k}(s) r_{k}(t) G_{N}^{(N+k)}(z)_{s t}\right|^{2}\right] \leqslant \frac{C}{N^{1-\gamma^{2}}} .
$$

Similarly, for each $N \in \mathbb{N}$ and $k \in\{1, \ldots, N\}, i \in\{1, \ldots, N\}$, we have the following inequality concerning the conditional expectation with respect to $M^{i}$ :

$$
\mathbb{E}\left[\left|\sum_{s, t \neq k, s \neq t}^{N} r_{i}(s) r_{i}(t) G_{N}^{(k, N+i)}(z)_{s t}\right|^{2} \mid M^{i}\right] \leqslant \frac{C}{N^{1-\gamma^{2}}} .
$$

Proof. We first expand the square and use the independence of $\left(r_{k}(s)\right)_{s}$ from $G_{N}^{(N+k)}(z)$ :

$$
\mathbb{E}\left[\left|\sum_{s \neq t}^{N} r_{k}(s) r_{k}(t) G_{N}^{(N+k)}(z)_{s t}\right|^{2}\right]=2 \sum_{s \neq t}^{N} \mathbb{E}\left[r_{k}(s)^{2} r_{k}(t)^{2}\right] \mathbb{E}\left[\left|G_{N}^{(N+k)}(z)_{s t}\right|^{2}\right]
$$

Now we compute

$$
\begin{aligned}
\mathbb{E}\left[r_{k}(s)^{2} r_{k}(t)^{2}\right] & =\mathbb{E}\left[M^{k}\left(\frac{s-1}{N}, \frac{s}{N}\right) M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)\right] \\
& =\int_{\frac{s-1}{N}}^{\frac{s}{N}} \int_{\frac{t-1}{N}}^{\frac{t}{N}} \max \left(1, \frac{\tau}{|r-u|}\right)^{\psi(2)} \mathrm{d} r \mathrm{~d} u \\
& \leqslant \int_{0}^{\frac{1}{N}} \int_{\frac{1}{N}}^{\frac{2}{N}} \max \left(1, \frac{\tau}{|r-u|}\right)^{\psi(2)} \mathrm{d} r \mathrm{~d} u
\end{aligned}
$$

We consider $N$ large enough so as to make $2 / N \leqslant \tau$. The above integral is then plain to compute and we get

$$
\begin{equation*}
\mathbb{E}\left[r_{k}(s)^{2} r_{k}(t)^{2}\right] \leqslant \frac{\tau^{\psi(2)}\left(2^{2-\psi(2)}-2\right)}{(1-\psi(2))(2-\psi(2))} \frac{1}{N^{2-\psi(2)}} . \tag{A.3}
\end{equation*}
$$

Thus we have for some positive constant $C$

$$
\begin{aligned}
\mathbb{E}\left[\left|\sum_{s \neq t}^{N} r_{k}(s) r_{k}(t) G_{N}^{(N+k)}(z)_{s t}\right|^{2}\right] & \leqslant \frac{C}{N^{2-\psi(2)}} \sum_{s \neq t}^{N} \mathbb{E}\left[\left|G_{N}^{(N+k)}(z)_{s t}\right|^{2}\right] \\
& \leqslant \frac{C}{N^{1-\psi(2)}} \frac{1}{|\Im(z)|^{2}}
\end{aligned}
$$

where we have used the fact that almost surely:

$$
\frac{1}{2 N-1} \sum_{s, t \neq N+k}^{2 N}\left|G_{N}^{(N+k)}(z)_{s t}\right|^{2} \leqslant \frac{1}{|\Im(z)|^{2}}
$$

It just remains to see that $\psi(2)=\gamma^{2}$. To prove the second relation, we follow the same argument by noticing that $\left(r_{i}(t)\right)_{t}$ and $G_{N}^{(k, N+i)}(z)$ are independent conditionally to $M^{i}$.
Lemma A.3. There exists some constant $c>0$ such that for each $N \in \mathbb{N}$ and $k \in\{1, \ldots, N\}$ :

$$
\mathbb{E}\left[\left|\sum_{i \neq j}^{N} r_{i}(k) r_{j}(k) G_{N}^{(k)}(z)_{N+i, N+j}\right|^{2}\right] \leqslant \frac{c}{N}
$$

Proof. Again we expand the square and we use the fact that, conditionally to the $\left(M^{i}\right)_{i}$, the quantities $r_{i}(k), r_{j}(k), G_{N}^{(k)}(z)_{N+i, N+j}$ are independent and $r_{i}(k), r_{j}(k)$ are centered. Indeed, conditionally to the $\left(M^{i}\right)_{i}$, the variables $r_{i}(k), r_{j}(k), G_{N}^{(k)}(z)_{N+i, N+j}$ involve different increments of the Brownian motion. Thus we have

$$
\begin{aligned}
\mathbb{E}\left[\left|\sum_{i \neq j}^{N} r_{i}(k) r_{j}(k) G_{N}^{(k)}(z)_{N+i, N+j}\right|^{2}\right] & =\sum_{i \neq j}^{N} \mathbb{E}\left[r_{i}(k)^{2} r_{j}(k)^{2}\right] \mathbb{E}\left[\left|G_{N}^{(k)}(z)_{N+i, N+j}\right|^{2}\right] \\
& \leqslant \sum_{i \neq j}^{N} \mathbb{E}\left[r_{i}(k)^{2}\right] \mathbb{E}\left[r_{j}(k)^{2}\right] \mathbb{E}\left[\left|G_{N}^{(k)}(z)_{N+i, N+j}\right|^{2}\right] \\
& =N^{-2} \sum_{i \neq j}^{N} \mathbb{E}\left[\left|G_{N}^{(k)}(z)_{N+i, N+j}\right|^{2}\right] \\
& \leqslant \frac{c}{N}
\end{aligned}
$$

where we have used the fact that almost surely:

$$
\frac{1}{2 N-1} \sum_{i, j \neq k}^{2 N}\left|G_{N}^{(k)}(z)_{i, j}\right|^{2} \leqslant \frac{1}{|\Im(z)|^{2}}
$$

Proof of Lemma 5.13. We define the function $f_{N}^{k, \varepsilon}$ on the interval $[0,1]$ by

$$
f_{N}^{k, \varepsilon}(x)=N M^{k, \varepsilon}\left(I_{N}^{t}\right) \text { if } x \in I_{N}^{t}
$$

Notice the relation:

$$
\sum_{t=1}^{N} M_{\varepsilon}^{k}\left(I_{N}^{t}\right) \mathbb{E}\left[G_{N}(z)_{t t}\right]=\int_{0}^{1} f_{N}^{k, \varepsilon}(r) d \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)
$$

Then, by stationarity, we have:

$$
\begin{aligned}
& \mathbb{E}\left[\left|\int_{0}^{1} f_{N}^{k, \varepsilon}(r) d \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)-\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}(r)} \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)\right|\right] \\
& \leqslant \sum_{t=1}^{N} \mathbb{E}\left[\left|\int_{I_{N}^{t}}\left(f_{N}^{k, \varepsilon}(r)-\mathrm{e}^{\omega_{\varepsilon}^{k}(r)}\right) \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)\right|\right] \\
& \leqslant \frac{N}{|\Im(z)|} \sup _{r \in I_{N}^{1}} \mathbb{E}\left[\left|\int_{I_{N}^{1}}\left(\mathrm{e}^{\omega_{\varepsilon}^{k}(u)}-\mathrm{e}^{\omega_{\varepsilon}^{k}(r)}\right) \mathrm{d} u\right|\right] \\
& \left.\leqslant\left.\frac{N}{|\Im(z)|} \sup _{r \in I_{N}^{1}} \int_{I_{N}^{1}} \mathbb{E}\left[\mid \mathrm{e}^{\omega_{\varepsilon}^{k}(u)}-\mathrm{e}^{\omega_{\varepsilon}^{k}(r)}\right)\right|^{2}\right]^{1 / 2} \mathrm{~d} u \\
& \leqslant \frac{N}{|\Im(z)|} \sup _{r \in I_{N}^{1}} \int_{I_{N}^{1}}\left(2 \mathrm{e}^{\psi(2) \rho_{\varepsilon}(0)}-2 \mathrm{e}^{\psi(2) \rho_{\varepsilon}(r-u)}\right)^{1 / 2} \mathrm{~d} u
\end{aligned}
$$

Because of the continuity of the function $\rho_{\varepsilon}$ over $[0,1]$, we have

$$
\begin{equation*}
\mathbb{E}\left[\left|\int_{0}^{1} f_{N}^{k, \varepsilon}(r) \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)-\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}(r)} \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)\right|\right] \rightarrow 0 \quad \text { as } N \rightarrow \infty . \tag{A.4}
\end{equation*}
$$

In a quite similar way, we can prove that

$$
\begin{equation*}
\mathbb{E}\left[\left|\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}} * \phi_{p}(r) \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)-\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}(r)} \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r)\right|\right] \rightarrow 0 \quad \text { as } p \rightarrow \infty \text { uniformly w.r.t. } N \tag{A.5}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{E}\left[\left|\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}} * \phi_{p}(r) K_{z}(r) \mathrm{d} r-\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}(r)} K_{z}(r) \mathrm{d} r\right|\right] \rightarrow 0 \quad \text { as } p \rightarrow \infty \text { uniformly w.r.t. } N \tag{A.6}
\end{equation*}
$$

where $\left(\phi_{p}\right)_{p \in \mathbb{N}}$ is a regularizing sequence and $*$ stands for the convolution. Furthermore, for each fixed $p$ and because of the weak convergence of $\mathbb{E}\left[L_{N}^{1, z}\right]$ towards $K_{z}(x) \mathrm{d} x$, we have almost surely

$$
\begin{equation*}
\int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}} * \phi_{p}(r) \mathrm{d} \mathbb{E}\left[L_{N}^{1, z}\right](\mathrm{d} r) \rightarrow \int_{0}^{1} \mathrm{e}^{\omega_{\varepsilon}^{k}} * \phi_{p}(r) K_{z}(r) \mathrm{d} r \quad \text { as } N \rightarrow \infty . \tag{A.7}
\end{equation*}
$$

We prove the result by gathering (A.4)-(A.7).

## Appendix B. Sup of MRW

Here we prove
Proposition B.1. We have for all $k=1, \ldots, N+1$

$$
\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4}\right] \leqslant C \frac{(\ln N)^{2}}{N^{\frac{\zeta(2 \alpha)-1}{\alpha}}} .
$$

for some positive constant $C$.
Proof. To prove the result, we first prove
Lemma B.2. There exists a constant $C$ such that, if $\left(X_{i}\right)_{1 \leqslant i \leqslant N}$ are iid centered Gaussian random variables then:

$$
\mathbb{E}\left[\max _{1 \leqslant i \leqslant N}\left|X_{i}\right|^{4}\right] \leqslant C \max _{1 \leqslant i \leqslant N} \mathbb{E}\left[X_{i}^{2}\right]^{2}(\ln N)^{2} .
$$

Proof. By homogeneity, it suffices to assume that $\mathbb{E}\left[X_{i}^{2}\right]=1$. Then we have for all $\delta \geqslant 0$

$$
\begin{aligned}
\mathbb{E}\left[\max _{1 \leqslant i}\left|X_{i}\right|^{4}\right] & \leqslant \delta+N \int_{\delta}^{\infty} \mathbb{P}\left(\left|X_{1}\right|^{4}>t\right) \mathrm{d} t \\
& \leqslant \delta+2 N \int_{\delta}^{\infty} \mathbb{P}\left(X_{1}>t^{1 / 4}\right) \mathrm{d} t \\
& \leqslant \delta+\frac{2 N}{\sqrt{2 \pi}} \int_{\delta}^{\infty} \mathrm{e}^{-\sqrt{t}} \mathrm{~d} t \\
& \leqslant \delta+\frac{4 N}{\sqrt{2 \pi}} \int_{\sqrt{\delta}}^{\infty} \mathrm{e}^{-t} t \mathrm{~d} t \\
& \leqslant \delta+\frac{4 N}{\sqrt{2 \pi}}\left(\sqrt{\delta} \mathrm{e}^{-\sqrt{\delta}}+\mathrm{e}^{-\sqrt{\delta}}\right)
\end{aligned}
$$

and this last expression can be made smaller than $C(\ln N)^{2}$ by choosing $\delta=(\ln N)^{2}$.

We want apply the above lemma after conditioning with respect to the law of the MRM $M^{k}$ :

$$
\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4}\right]=\mathbb{E}\left[\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4} \mid M^{k}\right]\right]
$$

Notice then that, conditionally to $M^{k}\left(0, \frac{1}{N}\right)=x_{1}, \ldots, M^{k}\left(\frac{N-1}{N}, 1\right)=x_{N}$, the vector $\left(r_{k}(1), \ldots, r_{k}(N)\right.$ has the same law as the increments of $B:\left(B_{x_{1}}-B_{0}, \ldots, B_{x_{N}}-B_{x_{N-1}}\right)$. By applying Lemma B.2, we deduce that

$$
\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4} \mid M^{k}\right] \leqslant C(\ln N)^{2} \max _{t=1, \ldots, N} M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)^{2}
$$

Thus we deduce

$$
\begin{equation*}
\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4}\right] \leqslant C(\ln N)^{2} \mathbb{E}\left[\left(\max _{t=1, \ldots, N} M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)\right)^{2}\right] \tag{B.1}
\end{equation*}
$$

Finally we have for all $\delta>0$ and for $\alpha>1$ such that $\zeta(2 \alpha)>1$ :

$$
\begin{aligned}
\mathbb{E}\left[\left(\max _{t=1, \ldots, N} M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)\right)^{2}\right] & \leqslant \delta+N \int_{\delta}^{\infty} \mathbb{P}\left(M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)^{2}>x\right) \mathrm{d} x \\
& \leqslant \delta+N \int_{\delta}^{\infty} \frac{1}{x^{\alpha}} \mathbb{E}\left[M^{k}\left(\frac{t-1}{N}, \frac{t}{N}\right)^{2 \alpha}\right] \mathrm{d} x \\
& \leqslant \delta+C \delta^{1-\alpha} N^{1-\zeta(2 \alpha)}
\end{aligned}
$$

for some constant $C$ only depending on $\alpha, \tau$ and $\gamma^{2}$. Choose now $\delta=N^{\frac{1-\zeta(2 \alpha)}{\alpha}}$ so as to get

$$
\begin{equation*}
\mathbb{E}\left[\sup _{t=1, \ldots, N} r_{k}(t)^{4}\right] \leqslant(1+C) \frac{(\ln N)^{2}}{N \frac{\zeta(2 \alpha)-1}{\alpha}} \tag{B.2}
\end{equation*}
$$

## Appendix C. Girsanov transform

Lemma C.1. If the process $\omega_{\varepsilon}$ is defined as $\omega_{\varepsilon}(x)=\mu\left(A_{\varepsilon}(x)\right)$ where $\mu$ is an independently scattered random measure associated to $(\varphi, \theta)$ with $\varphi(q)=-i q \gamma^{2} / 2-q^{2} \gamma^{2} / 2$ and $\theta$ given by 2.1 , then:

$$
\lim _{\varepsilon \rightarrow 0} \mathbb{E}\left[\frac{\mathrm{e}^{\omega_{\varepsilon}(x)}}{z-\int_{0}^{1} K_{z}(r) \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r}\right]=\mathbb{E}\left[\left(z-\int_{0}^{1}\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} K_{z}(r) M(\mathrm{~d} r)\right)^{-1}\right]
$$

where $M$ is the lognormal MRM.
Proof. One can check that $\left(\omega_{\varepsilon}(x)\right)_{x \in[0 ; 1]}$ is a stationary Gaussian process with covariance given by $\gamma^{2} \rho_{\varepsilon}(x-y)$. So, using Girsanov transform, we can write:

$$
\mathbb{E}\left[\frac{\mathrm{e}^{\omega_{\varepsilon}(x)}}{z-\int_{0}^{1} K_{z}(r) \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r}\right]=\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r) \mathrm{e}^{\gamma^{2} \rho_{\varepsilon}(r-x)} \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right)^{-1}\right]
$$

We are interested in the limit when $\varepsilon$ goes to 0 of this latter term, we thus approximate it with a simpler term:

$$
\begin{align*}
& \mid \mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r) \mathrm{e}^{\gamma^{2} \rho_{\varepsilon}(r-x)} \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right)^{-1}\right] \\
& \left.-\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right)^{-1}\right] \right\rvert\, \\
& \leqslant \frac{1}{|\Im(z)|^{2}} \mathbb{E}\left[\int_{0}^{1}\left|K_{z}(r)\right| \mathrm{e}^{\omega_{\varepsilon}(r)}\left|\mathrm{e}^{\gamma^{2} \rho_{\varepsilon}(r-x)}-\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}\right| \mathrm{d} r\right] \\
& \leqslant \frac{1}{|\Im(z)|^{3}} \int_{0}^{1}\left|\mathrm{e}^{\gamma^{2} \rho_{\varepsilon}(r-x)}-\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}\right| \mathrm{d} r \tag{C.1}
\end{align*}
$$

where we have used Lemmas 5.4 and 5.8 and the normalization $\psi(1)=0$.
Because $\gamma^{2}<1$, the dominated convergence theorem implies that C. 1 converges to 0 when $\varepsilon$ goes to 0 .
We thus look at the limit when $\varepsilon$ goes to 0 of the term:

$$
\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right)^{-1}\right]
$$

The random variable

$$
\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} M(\mathrm{~d} r)
$$

is well-defined and is finite almost surely since:

$$
\mathbb{E}\left[\left|\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} M(\mathrm{~d} r)\right|\right] \leqslant \int_{0}^{1}\left|K_{z}(r)\right|\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} \mathrm{~d} r<+\infty
$$

And thus, we can compute:

$$
\begin{aligned}
& \left\lvert\, \mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right)^{-1}\right]\right. \\
& \left.-\mathbb{E}\left[\left(z-\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}} M(\mathrm{~d} r)\right)^{-1}\right] \right\rvert\, \\
& \leqslant \frac{1}{|\Im(z)|^{2}} \mathbb{E}\left[\left|\int_{0}^{1} K_{z}(r)\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}\left(\mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r-M(\mathrm{~d} r)\right)\right|\right]
\end{aligned}
$$

and, for all $n \in \mathbb{N}$, this latter term is smaller than

$$
\begin{align*}
& \mathbb{E}\left[\left|\int_{0}^{1} K_{z}(r)\left[\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}-\min \left(\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}, n\right)\right] \mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r\right|\right]  \tag{C.2}\\
& +\mathbb{E}\left[\left|\int_{0}^{1} K_{z}(r) \min \left(\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}, n\right)\left(\mathrm{e}^{\omega_{\varepsilon}(r)} \mathrm{d} r-M(\mathrm{~d} r)\right)\right|\right]  \tag{C.3}\\
& +\mathbb{E}\left[\left|\int_{0}^{1} K_{z}(r)\left[\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}-\min \left(\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}, n\right)\right] M(\mathrm{~d} r)\right|\right] . \tag{C.4}
\end{align*}
$$

The two quantities C. 2 and C. 4 are smaller than

$$
\begin{equation*}
\int_{0}^{1}\left|K_{z}(r)\right|\left[\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}-\min \left(\left(\frac{\tau}{|r-x|}\right)_{+}^{\gamma^{2}}, n\right)\right] \mathrm{d} r \tag{C.5}
\end{equation*}
$$

and thus converge to 0 , uniformly in $\varepsilon$ as $n$ goes to infinity.
For a fixed $n$, the function $\min \left((\tau /|r-x|)^{\gamma^{2}}, n\right)$ is measurable and bounded and thus it is plain to see that, for a fixed $n$, the term C. 3 goes to 0 when $\varepsilon$ goes to 0 .

The lemma follows gathering the above estimates.

## References

[1] E. Bacry, A. Kozhemyak and J.F. Muzy, Log-normal continuous cascade model of asset returns: aggregation properties and estimation. Quant. Finance 13 (2013) 795-818.
[2] E. Bacry and J.F. Muzy, Log-infinitely divisible multifractal processes. Comm. Math. Phys. 236 (2003) $449-475$.
[3] S. Belinschi, A. Dembo and A. Guionnet, Spectral measure of heavy tailed band and covariance random matrices. Comm. Math. Phys. 289 (2009) 1023-1055.
[4] G. Ben Arous and A. Guionnet, The spectrum of heavy-tailed random matrices. Comm. Math. Phys. 278 (2008) $715-751$.
[5] J.P. Bouchaud, A. Matacz and M. Potters, Leverage effect in financial markets: The retarded volatility model. Phys. Rev. Lett. 87 (2001) 228701.
[6] J.P. Bouchaud and M. Potters, Financial Applications of Random Matrix Theory: A Short Review. In Oxf. Handb. Random Matrix Theory. Oxford University Press (2011).
[7] J.P. Bouchaud and M. Potters, Theory of Financial Risk and Derivative Pricing. Cambridge University Press, Cambridge (2003).
[8] R. Cont, Empirical properties of asset returns: Stylized facts and statistical issues. Quant. Finance 1 (2001) $223-236$.
[9] J.S. Geronimo, T.P. Hill, Necessary and sufficient condition that the limit of Stieltjes transforms is a Stieltjes transform. J. Approx. Theory 121 (2003) 54-60.
[10] A. Khorunzhy, B. Khoruzhenko, L. Pastur and M. Shcherbina, The large-n limit in statistical mechanics and the spectral theory of disordered systems. Phase Transitions and Critical Phenomena. In vol. 73. Academic Press, New-York (1992).
[11] Y. Li and X. Zheng, On the estimation of integrated covariance matrices of high dimensional diffusion processes. Ann. Statist. 39 (2011) 3121-3151.
[12] V.A. Marchenko and L.A. Pastur, Distribution of eigenvalues for some sets of random matrices. Math. USSR-Sb. 1 (1967) 457-483.
[13] R. Rhodes and V. Vargas., Gaussian multiplicative chaos and applications: a review. ESAIM: PS 11 (2014) 315-392.
[14] C. Robert and M. Rosenbaum, On the limiting spectral distribution of the covariance matrices of time-lagged processes. $J$. Multivar. Anal. 101 (2010) 2434-2451.
[15] M. Potters, J.-P. Bouchaud and L. Laloux, Financial applications of random matrix theory: Old laces and new pieces. Acta Physica Polonica B $\mathbf{3 6}$ (2005) 2767.
[16] J. Duchon, R. Robert and V. Vargas, Forecasting volatility with the multifractal random walk model. Math. Finance 22 (2012) 83-108.


[^0]:    Keywords and phrases. Multifractals, Marchenko-Pastur theorem, random matrices, Gaussian multiplicative chaos.
    1 CNRS, UMR 7534, 75016 Paris, France.
    2 Université Paris-Dauphine, Ceremade, 75016 Paris, France.
    allez@ceremade.dauphine.fr; rhodes@ceremade.dauphine.fr; vargas@ceremade.dauphine.fr

[^1]:    ${ }^{3}$ In the formal definition equation (1.5) of the introduction, the "process" $(\omega(t))$ in (1.5) is centered. The process $\omega_{\varepsilon}$ is not.

[^2]:    ${ }^{4}$ The notation $(\cdot)_{+}$is a shortcut for $\max (\cdot, 1)$.

[^3]:    ${ }^{5}$ Recall that, in view of equation (4.1), we are interested in the value of the Stieltjes transform near the real line.

