DAUPHINE
UNIVERSITÉ PARIS

Ecole Doctorale de Dauphine

Thèse de Doctorat
Spécialité : Mathématiques

présentée par

Romain Allez

Chaos multiplicatif Gaussien, matrices aléatoires et applications

dirigée par Massimiliano Gubinelli et Vincent Vargas

soutenue le 23 novembre 2012 devant le jury composé de

M	Jean-Philippe	Bouchaud	(Examinateur)
M	Yan V.	Fyodorov	(Rapporteur)
M	Massimiliano	Gubinelli	(Directeur de thèse)
M^{me}	Sandrine	Péché	(Examinatrice)
M	Rémi	Rhodes	(Examinateur)
M	Vincent	Vargas	(Directeur de thèse)

après avis de M. Gérard Ben Arous et de M. Yan V. Fyodorov.

Résumé

Dans ce travail, nous nous sommes intéressés d'une part à la théorie du chaos multiplicatif Gaussien introduite par Kahane en 1985 et d'autre part à la théorie des matrices aléatoires dont les pionniers sont Wigner, Wishart et Dyson. La première partie de ce manuscrit contient une brève introduction à ces deux théories ainsi que les contributions personnelles de ce manuscrit expliquées rapidement. Les parties suivantes contiennent les textes des articles publiés [5, 6, 7, 8, 9] et pré-publiés [10, 11, 12] sur ces résultats dans lesquels le lecteur pourra trouver des développements plus détaillés.

Abstract

In this thesis, we are interested on the one hand in the theory of Gaussian multiplicative chaos introduced by Kahane in 1985 and on the other hand in random matrix theory whose pioneers are Wigner, Wishart and Dyson. The first part of this manuscript constitutes a brief introduction to those two theories and also contains the personal contributions of this work rapidly explained. The following parts contain the texts of the published articles [5, 6, 7, 8, 9] and pre-prints [10, 11, 12] on those results where the reader can find more detailed developments.

Publications

1. R. Allez, J.-P. Bouchaud, S. N. Majumdar, P. Vivo, *Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marčenko-Pastur law*, to appear in J. Phys. A. (2012).

2. R. Allez, R. Rhodes, V. Vargas, *Lognormal $\star$-scale invariant random measures*, Prob. Th. Rel. Fields (online first) (2012).

3. R. Allez, J.-P. Bouchaud and A. Guionnet, *Invariant β-ensembles and the Gauss-Wigner crossover*, Phys. Rev. Lett. **109**, 094102 (2012).

4. R. Allez, J.-P. Bouchaud, *Eigenvector dynamics: general theory and some applications*, Phys. Rev. E **86**, 046202 (2012).

5. P.-A. Reigneron, R. Allez and J.-P. Bouchaud, *Principal regression analysis and the index leverage effect*, Physica A **390**, 3026 (2011).

6. R. Allez and J.-P. Bouchaud, *Individual and collective stock dynamics: intra-day seasonalities*, New J. Phys. **13**, 025010 (2011).

Articles pré-publiés / Preprints

1. R. Allez and A. Guionnet, *A diffusive matrix model for invariant β-ensembles*, submitted, (2012).

2. R. Allez, R. Rhodes and V. Vargas, *Marchenko Pastur type theorem for independent MRW processes: convergence of the empirical spectral measure*, submitted (2011).

Remerciements

Au cours de cette thèse, j'ai eu la chance d'être encadré à la fois par Jean-Philippe Bouchaud, Rémi Rhodes et Vincent Vargas. Je tiens à les remercier vivement tous les trois pour la disponibilité qu'ils m'ont accordée depuis le début de ce travail et pour toutes les choses qu'ils m'ont apprises. C'est un grand plaisir de travailler avec des personnes si dynamiques, enthousiastes et sympathiques. J'ai beaucoup aimé réfléchir sur les sujets qu'ils m'ont proposés et j'espère poursuivre mes collaborations avec eux dans les années qui viennent.

Je souhaite aussi remercier Massimiliano Gubinelli, qui a rendu ce projet possible, pour les conversations que nous avons eues au cours desquelles j'ai compris certains points importants pour mes projets de recherche.

J'ai eu également le privilège de travailler en collaboration avec Alice Guionnet, Satya N. Majumdar, Pierre-Alain Reigneron et Pierpaolo Vivo.

Je suis très reconnaissant à Alice Guionnet de m'avoir accueilli chaleureusement à l'ENS Lyon pour des courts séjours en mai et en septembre 2011, ainsi que pour toutes nos discussions très instructives en marge de conférence (à l'institut de physique des Houches, à l'IHP et à Jussieu). J'ai beaucoup apprécié cette collaboration qui m'a beaucoup apportée et qui fut un grand honneur pour moi.

Je garde aussi un excellent souvenir de notre travail avec Satya N. Majumdar et Pierpaolo Vivo, en particulier du moment clé où nous nous sommes échangés des dizaines d'emails en seulement quelques jours pour finalement converger vers une même solution par deux chemins différents. C'est un plaisir de remercier également Pierre-Alain Reigneron pour une de mes toutes premières collaborations.

It is a great honor for me that Gérard Ben Arous and Yan V. Fyodorov have accepted to write a report on my work. Yan V. Fyodorov will come to the defense and I am very grateful to him for that.

Gérard Ben Arous a guidé mes premiers pas vers la recherche au cours d'un stage de M2 en 2008 au Courant Institute à New York University pendant lequel j'ai travaillé sur un modèle de pièges pour les marches aléatoires introduit par Jean-Philippe Bouchaud. Son encadrement fut très bénéfique, en particulier ma compréhension des articles s'est considérablement améliorée au cours de ce stage. Je lui en suis très reconnaissant.

Je tiens à remercier aussi mes anciens professeurs d'Orsay et de l'ENS, en particulier Jean-François Le Gall et Wendelin Werner, pour les précieux conseils qu'ils m'ont donnés au moment où je cherchais ma voie.

Je suis ravi que Sandrine Péché fasse partie du jury de cette thèse et je la remercie d'avoir accepté ce rôle.

Je voudrais remercier Rémy Chicheportiche pour nos discussions, pour ses relectures attentives de plusieurs manuscrits et aussi pour toutes ses réponses à mes questions de physique (et de programmation en R!).

Je remercie aussi Florent Benaych Georges, Grégory Schehr et Antoine Dahlqvist qui ont porté un intêret pour mon travail et à qui j'ai pu expliquer mes recherches respectivement au cours d'écoles d'été aux Houches, au café des fontaines de la rue Soufflot et au cours de nos parties de tennis.

J'adresse mes derniers remerciements scientifiques à Laure, qui a eu la gentillesse et la patience d'écouter mes maths et qui a fait de nombreuses remarques très pertinentes sur mes recherches.

Merci aux thésards du bureau C614 ainsi qu'à tous les membres du CEREMADE et aux stagiaires/ doctorants/ chercheurs de CFM qui ont rendu la vie quotidienne sur nos lieux de travail très agréable. Un grand merci aussi aux équipes de CFM pour leur accueil si chaleureux tout au long de cette thèse.

Mes derniers remerciements vont à ma famille et à mes amis du club athlétisme et de l'ENS (et aussi les autres !). Je pense à mes parents, à ma soeur Charlotte, à mes quatre grand-parents, à ma famille parisienne et surtout, à Laure.

à Laure...

Contents

Part I

Introduction

Chapter 1

Gaussian multiplicative chaos and their scale invariance properties

Roughly observed, some random phenomena seem scale invariant. This is the case for the velocity field of turbulent flows or for the evolution in time of the logarithm of the price of a financial asset. However, a more precise empirical study of these phenomena displays in fact a generalized form of scale invariance commonly called multifractal scale invariance or intermittency or stochastic scale invariance. The exponent which governs the power law scaling of the process or field is no longer linear. More precisely, for a one dimensional stationary process we observe the following behavior for small scales h,

$$\mathbb{E}\left[|X_{t+h} - X_t|^p\right] \sim_{h\to 0_+} C_p h^{\zeta(p)}$$

where ζ is a non linear function.

The concept of nonlinear power-law scalings goes back to the Kolmogorov theory of fully developed turbulence in the sixties (see [52, 138, 143, 53, 76] and references therein), introduced to render the intermittency effects in turbulence. In 1974, Mandelbrot [111] came up with the first mathematical discrete approach of multifractality, the now celebrated multiplicative cascades, in order to build random measures describing energy dissipation and contribute explaining intermittency effects in Kolmogorov's theory of fully developed turbulence.

Despite the fact that multiplicative cascades have been widely used as reference models in many applications, they possess drawbacks related to their discrete scale invariance, mainly they involve a particular scale ratio and they do not possess stationary fluctuations (this comes from the fact that they are constructed on a dyadic tree structure).

A more refined model was then introduced by Kahane in 1985, under the name of Gaussian multiplicative chaos. Following the ideas of multiplicative cascades, Kahane's theory of Gaussian multiplicative chaos also constructs multifractal random measures by iterating products of lognormal random variables. The advantages of this construction is that the random measures possess stationary fluctuations (no particular scale ratio is involved in the construction).

This chapter is a brief introduction to this theory of multiplicative chaos. In the first subsection 1.0.1, we start by recalling the construction of Mandelbrot's multiplicative cascades and their main properties. In particular, we introduce Mandelbrot's star equation, which is a stochastic scale invariance equation, satisfied by the multiplicative cascades. In the following subsection, we introduce Kahane's theory of Gaussian multiplicative chaos. The third subsection is devoted entirely to the concept of stochastic scale invariance. We introduce two concepts: the exact stochastic scale invariance and the stochastic star-scale invariance (which is the continuous analog of the Mandelbrot's star equation). We present a brief study on the link with Gaussian multiplicative chaos and on the relation between the two concepts. In the final subsection 1.0.4, we enunciate the main result obtained in this field.

1.0.1 Multiplicative cascades and Mandelbrot's star equation

We recall the definition of multiplicative cascades, following the notations used in [34]. The fundamental properties of multiplicative cascades were first proved in [95] by Kahane and Peyrière in 1976.

Let $\mathcal{I}_n$ denotes the set of dyadic subintervals of $[0,1]$ of length 2^{-n}, i.e.

$$\mathcal{I}_n := \left\{ [k2^{-n} : (k+1)2^{-n}] : k \in \{0, 1, \dots, 2^n - 1\} \right\} .$$

Each interval in $\mathcal{I}_n$ can be divided in exactly two subintervals in $\mathcal{I}_{n+1}$. We will also set $\mathcal{I} := \bigcup_n \mathcal{I}_n$.

Let $Z = e^{X - \frac{\sigma^2}{2}}$ where X is a Gaussian variable of mean 0 and variance σ^2. Note that Z is a non negative random variable with mean 1. We choose here a lognormal distribution for Z in order to be consistent with the next subsections but the construction of multiplicative cascades can be made with any choice of non negative random variable Z with mean 1. Let also $Z_I, I \in \mathcal{I}$ be a collection of independent random variables distributed as Z.

We now define inductively a sequence of random measures $(\mu_n)_{n\in\mathbb{N}}$ on $[0;1]$. Let μ_0 denote the Lebesgue measure and set $\mu_1 := Z_{[0;1]}\mu_0$. Let μ_2 denote the measure that agrees with $Z_{[0;1/2]}\mu_1$ on $[0;1/2]$ and with $Z_{[1/2;1]}\mu_1$ on $[1/2;1]$. Inductively, define μ_{n+1} as the measure that agrees on every $I \in \mathcal{I}_n$ with $Z_I\mu_n$. Alternatively, we can also define the multiplicative cascade by setting

$$\mu_n := w_n\mu_0, \quad \text{where} \quad w_n(x) = \prod_{j=0}^{n-1} Z_{I_j(x)} ,$$

where $I_j(x)$ denotes the interval $I \in \mathcal{I}_j$ that contains x (if there is more than one, the one whose maximum is x, say).

For each Borel set $A \subset [0,1]$, note that the sequence $(\mu_n(A))_n$ is a non negative martingale. Thus, the sequence $\mu_n(A)$ converges almost surely when $n \to \infty$ to a limit that we will denote by $\mu(A)$.

Consequently the sequence of measures (μ_n) converges almost surely weakly to a measure μ on $[0;1]$. The authors of [95] also give a non-degenerescence criterion for the measure μ: the measure μ is almost surely non zero if and only if we have $\mathbb{E}[Z \log Z] < \log 2$. It is also shown that the measure μ has almost surely no atoms.

The multiplicative cascades present multifractal property in the sense that they obey the following stochastic scale invariance equation:

$$\mu(dt) \overset{law}{=} Z_{[0;1/2]}\mathbf{1}_{[0,\frac{1}{2}]}(t)\mu^0(2dt) + Z_{[1/2;1]}\mathbf{1}_{[\frac{1}{2},1]}(t)\mu^1(2dt-1) , \tag{1.1}$$

where μ^0, μ^1 are two independent copies of μ and $Z_{[0;1/2]}, Z_{[1/2;1]}$ are independent copies of Z, also independent of μ^0, μ^1. Such an equation (and its generalizations to b-adic trees for $b \geqslant 2$), the celebrated star equation introduced by Mandelbrot in [109], uniquely determines the law of the multiplicative cascade.

Let us mention here that even if we write equation (1.1) in the more general setting where $(Z_{[0;1/2]}, Z_{[1/2;1]})$ has any prescribed law with components of mean 1, Mandelbrot star's equation uniquely characterizes the law of the multiplicative cascade.

In the next subsection, we present Kahane's theory of Gaussian multiplicative chaos, which constructs, again through iterative products, random measures with properties similar to multiplicative cascades.

1.0.2 Kahane's theory of Gaussian multiplicative chaos

Kahane introduced the theory of Gaussian multiplicative chaos in 1985 in his paper [94]. This theory relies on the notion of σ-positive type kernel: a kernel $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ is of σ-positive type if there exists a sequence $K_k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$ of continuous non negative and positive definite type kernels such that:

$$K(x,y) = \sum_{k=1}^{\infty} K_k(x,y) . \tag{1.2}$$

If K is a σ-positive type kernel with decomposition (1.2), one can consider a sequence of Gaussian processes $(X_n)_{n\geqslant 1}$ defined as $X_n = \sum_{k=1}^n Y_k$ where the $(Y_k(x))_{x\in\mathbb{R}^d}, k \geqslant 1$ are independent

centered Gaussian processes on $\mathbb{R}^d$ with respective covariance functions $K_k(x,y), k \geqslant 1$. Then the covariance function of the Gaussian process $(X_n(x))_{x\in\mathbb{R}^d}$ is $\sum_{k=1}^n K_k$. It is proved in [94] that the sequence of random measures m_n defined on every Borel set $A \subset \mathbb{R}^d$ by

$$m_n(A) = \int_A e^{X_n(x) - \frac{1}{2}\mathbb{E}[X_n(x)^2]} dx$$

converges almost surely in the space of Radon measures (equipped with the topology of weak convergence) towards a (random) measure m and that the limit measure m does not depend on the sequence K_k used in the decomposition (1.2). The measure m is called Gaussian multiplicative chaos associated to the kernel K.

The main application of this theory is to give a meaning to the "limit lognormal" model introduced by Mandelbrot in [112]. We denote by $\ln_+(x)$ the quantity $\max(\ln x, 0)$. The "limit lognormal" model corresponds to the choice of a stationary kernel K given by

$$K(x,y) = \gamma^2 \ln_+ \left(\frac{R}{|x-y|} \right) + g(x,y) \tag{1.3}$$

where g is a bounded continuous function and γ^2 and R are respectively the intermittency parameter and the integral scale.

In [94], Kahane also gives a non degeneracy criterion: Gaussian multiplicative chaos m associated to a kernel K of the form (1.3) is almost surely non degenerate if and only if the intermittency parameter γ^2 is strictly less than $2d$. This phase transition shows that the logarithmic kernel is crucial in the theory of multiplicative chaos. Let us define the characteristic exponent ζ (also called structure function):

$$\zeta(p) = (d + \frac{\gamma^2}{2})p - \frac{\gamma^2 p^2}{2}\,.$$

The inequality $\gamma^2 < 2$ implies the existence of $\epsilon > 0$ such that $\zeta(1+\epsilon) > 1$ and therefore there exists a unique $p^* > 1$ such that $\zeta(p^*) = d$. Existence of positive moments for the random variable $m[0;1]$ is then characterized through p^*: the random variable $m[0;1]$ admits moments of order $q > 0$ for all $q \in]0; p^*[$. It is also shown in [94] that the Gaussian multiplicative chaos m has the following multifractal behavior when $h \to 0$,

$$\mathbb{E}[m[0;h]^p] \sim h^{\zeta(p)}$$

for all $p < p^*$. We will establish the link between Gaussian multiplicative chaos and stochastic scale invariance in the next subsections.

The reader may wonder if we can get rid of the assumption that the kernels K_k used in the decomposition are non negative (and then also of the assumption that K itself is non negative) for this construction to be valid. The answer is yes and is carried out in [129]: we can define Gaussian multiplicative chaos associated to any positive definite kernel that can be written under the form (1.3).

The theory of Gaussian multiplicative chaos has found applications in 2d quantum gravity, turbulence, finance,... We now review an application in finance.

Application in finance Let us briefly explain the application in finance for modeling the volatility of an asset. If $X(t)$ is the logarithm of the price of a financial asset, the volatility m of the asset on the interval $[0;t]$ can be defined as the quadratic variation of X:

$$m[0;t] = \lim_{n\to\infty} \sum_{k=1}^{n} (X(tk/n) - X(t(k-1)/n))^2\,.$$

The volatility can be viewed as a random measure on $\mathbb{R}_+$. The choice to model the volatility by a Gaussian multiplicative chaos m associated to the logarithmic kernel

$$K(s,t) = \gamma^2 \ln_+ \left(\frac{R}{|t-s|} \right) \tag{1.4}$$

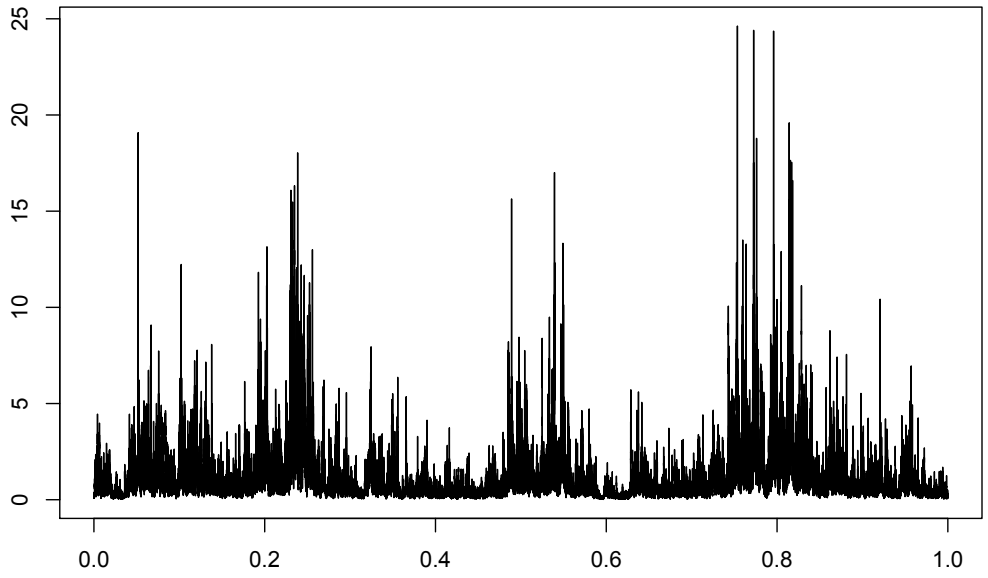


Figure 1.1: Simulated density of a Gaussian multiplicative chaos m associated to the kernel (1.3) with $g = 0$, intermittency parameter $\gamma^2 = 0.1$ and integral scale $\tau = 1/8$. The intermittency effect appears clearly.

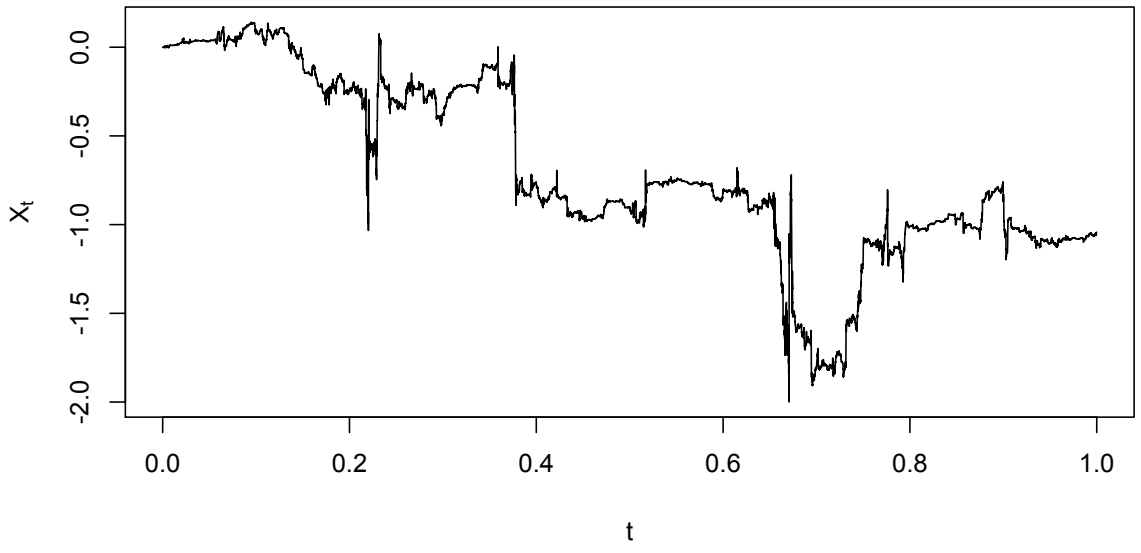


Figure 1.2: Simulated path of a multifractal random walk with intermittency parameter $\gamma^2 = 1$ and with integral scale $\tau = 1/4$. Note the intermittent bursts in volatility.

enables to reproduce important empirical properties measured on financial markets: approximate lognormality of the volatility, long range correlations (see [45, 60] for a review on empirical finance). Given the volatility m, the most natural way to construct a model for the log price X is then to set

$$X(t) = B(m[0;t]) \tag{1.5}$$

where $B(t)$ is a standard Brownian motion independent of m. Formula (1.5) defines the multifractal random walk (see [23]). A simulated path of a multifractal random walk is displayed in Fig. 8.1. We can observe the intermittent bursts of volatility which are due to the Gaussian multiplicative chaos m. We will work with this model in Chapter 8 (and also in the article [11]). □

The theory of Gaussian multiplicative chaos can be generalized by doing products of exponentials of infinitely divisible random variables. It was developed for example in [24].

We will now review stochastic scale invariance and its link with Gaussian multiplicative chaos in the next subsections. We will mainly restrict to the Gaussian case, but the following results can be extended to the case of Lévy random variables as well (see e.g. [133]).

1.0.3 Stochastic scale invariance

We will now define two different concepts of stochastic scale invariance: the exact stochastic scale invariance and the star stochastic scale invariance. We will give those two definitions for general dimension d.

Lognormal exact stochastic scale invariance

Denote by $B(0;R)$ the euclidean ball of radius R. If m is a stationary random measure on $\mathbb{R}^d$, we say that m has the exact stochastic scale invariance property if there exists $R > 0$ such that for all $\varepsilon \in]0,1]$, the following equality in law holds:

$$(m(\varepsilon A))_{A \subset B(0;R)} \overset{law}{=} e^{\Omega_\varepsilon} (m(A))_{A \subset B(0;R)} \tag{1.6}$$

where Ω_ε is a Gaussian random variable independent of m.

As already mentioned, we restrict to the case of a Gaussian random variable Ω_ε. If we do not make any particular assumption on Ω_ε, then if the measure m is not identically zero, we can easily check that the random variable Ω_ε is infinitely divisible.

Example of lognormal exact stochastic scale invariant random measures

It is proved in [129] that the logarithmic kernel K defined in (1.4) is positive definite if and only if $d \leqslant 3$.

It is then straightforward to prove that, for $d = 1, 2, 3$, the Gaussian multiplicative chaos m associated to this kernel K has the exact stochastic scale invariance property in the ball $B(0,R)$ where the random variable Ω_ε is a Gaussian random variable with mean $-(d + \gamma^2/2)\ln(1/\varepsilon)$ and variance $\gamma^2 \ln(1/\varepsilon)$.

For $d \geqslant 4$, the authors of [132] construct stationary and isotropic random measures which have the exact stochastic scale invariance property and which again fall under the scope of Kahane's theory of Gaussian multiplicative chaos introduced above.

The question of finding all the stationary random measures which satisfy the exact stochastic scale invariance property remains unsolved.

Lognormal star scale invariance

The second concept of scale invariance is stochastic star scale invariance. It is the continuous analog of the Mandelbrot star equation that we wrote in the case of 2-adic multiplicative cascades in (3.1).

A random Radon measure M is said to be lognormal star scale invariant if for all $\varepsilon < 1$, M obeys the cascading rule

$$(M(A))_{A\in\mathcal{B}(\mathbb{R}^d)} \overset{law}{=} \left(\int_A e^{\omega_\varepsilon(x)} M^\varepsilon(dx)\right)_{A\in\mathcal{B}(\mathbb{R}^d)} \tag{1.7}$$

where ω_ε is a stationary continuous Gaussian process and M^ε a random measure independent from the process ω_ε with law such that

$$(M^\varepsilon(A))_{A\in\mathcal{B}(\mathbb{R}^d)} \overset{law}{=} \varepsilon \left(M\left(\frac{A}{\varepsilon}\right)\right)_{A\in\mathcal{B}(\mathbb{R}^d)}.$$

Intuitively, this relation means that when we zoom in the measure M, we should observe the same behavior up to an independent lognormal factor. A nice feature of stochastic star scale invariance is that there is not a particular scale ratio.

Example of lognormal star scale invariant random measure

In this paragraph, we give the first known example of log-normal star scale invariant random measures, which was first described in [26] (see also [24]). In fact many other examples exist among multiplicative chaos as we will see later when explaining the results of our paper [5].

The construction of this example is very intuitive (but limited to dimension 1) : It is geometric and relies on homothetic properties of triangles in the half-plane. We also stress that this specific example of star scale invariant random measures is not restricted to the Gaussian case: The factor can be more general (log-Lévy).

Following [24], we recall the construction of this example and refer the reader to the aforementioned papers for further details. Fix $T > 0$ and let $\mathcal{S}^+$ be the state-space half plane

$$\mathcal{S}^+ = \{(t, l) : t \in \mathbb{R}, l > 0\}.$$

with which one can associate the measure

$$\mu(dt, dl) = l^{-2} dt dl.$$

Then we introduce the independently scattered Gaussian random measure P characterized for any μ-measurable set A by

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

with $\varphi(q) = -\gamma^2 q^2/2 - iq\gamma^2/2$. Under those assumptions, we can note that for any μ-measurable set A, $P(A)$ is a Gaussian variable with mean $m = -\mu(A)\gamma^2/2$ and variance $\sigma^2 = \gamma^2\mu(A)$. We can then define the Gaussian process $(\omega_l(t))_{t\in\mathbb{R}}$ for $l \geqslant 0$ by

$$\omega_l(t) = P\left(\mathcal{A}_l(t)\right)$$

where $\mathcal{A}_l(t)$ is the triangle like subset $\mathcal{A}_l(t) := \{(t', l') : l \leqslant l' \leqslant T, -l'/2 \leqslant t - t' \leqslant l'/2\}$ (see a picture of this triangle in Fig. 1.3).

Define now the random measure M_l by $M_l(dt) = e^{\omega_l(t)} dt$. Almost surely, the family of measures $(M_l(dt))_{l>0}$ weakly converges towards a random measure M. If $\gamma^2 < 2$, the measure M is not trivial.

Let us check that M is a good log-normal star scale invariant random measure. Fix $\epsilon < 1$ and define the sets $\mathcal{A}_{l,\epsilon T}(t) := \{(t', l') : l \leqslant l' \leqslant \epsilon T, -l'/2 \leqslant t - t' \leqslant l'/2\}$ and $\mathcal{A}_{\epsilon T,T}(t) := \{(t', l') : \epsilon T \leqslant l' \leqslant T, -l'/2 \leqslant t - t' \leqslant l'/2\}$ (see Fig. 1.4 for a picture of those two ensembles). Note that $\mathcal{A}_l(t) = \mathcal{A}_{l,\epsilon T}(t) \cup \mathcal{A}_{\epsilon T,T}(t)$ and that those two sets are disjoint. Thus, we can write for every μ-measurable set A

$$M_l(A) = \int_A e^{\omega_{\epsilon T,T}(t)} e^{\omega_{l,\epsilon T}(t)} dt \tag{1.8}$$

with $\omega_{\epsilon T,T}(t) = P(\mathcal{A}_{\epsilon T,T}(t))$ and $\omega_{l,\epsilon T}(t) = P(\mathcal{A}_{l,\epsilon T}(t))$.

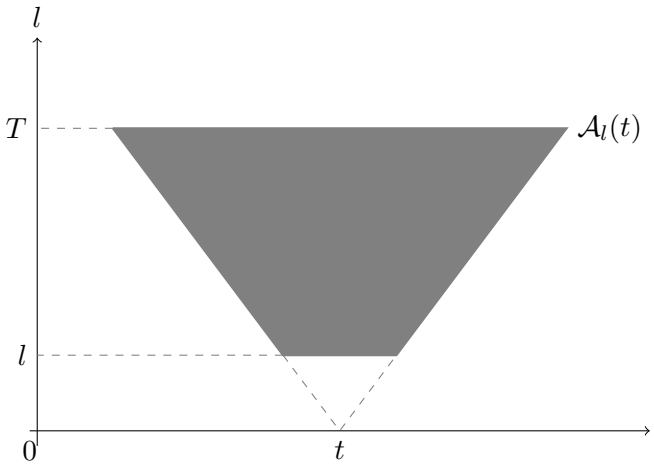


Figure 1.3: Representation of Barral-Mandelbrot's cone.

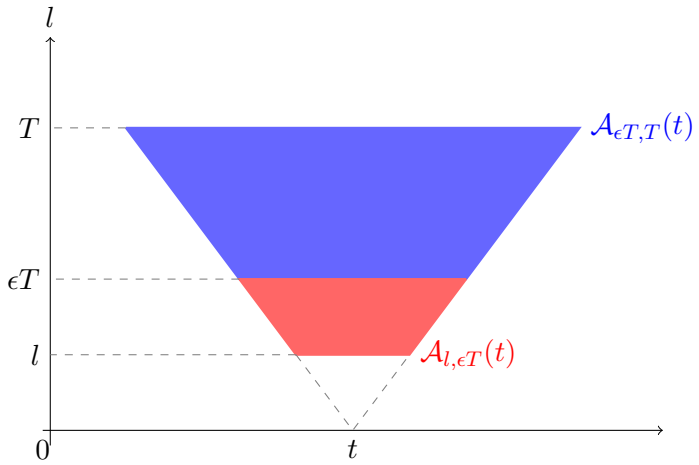


Figure 1.4: Decomposition of Barral-Mandelbrot's cone.

We then study equation (3.17) in the limit $l \to 0$; we obtain

$$M(A) = \int_A e^{\omega_{\epsilon T,T}(t)} M^\epsilon(dt) \tag{1.9}$$

where M^ϵ is the limit when $l \to 0$ of the random measure $M_l^\epsilon(dt) := e^{\omega_{l,\epsilon T}(t)} dt$. We easily verify that $M^\epsilon(\epsilon A) \overset{law}{=} \epsilon M(A)$ writing

$$M_l^\epsilon(A) = \epsilon \int_A e^{\omega_{l,\epsilon T}(\epsilon t)} dt \tag{1.10}$$

and checking that the covariance of the Gaussian process $(\omega_{l,\epsilon T}(\epsilon t))_{t\in\mathbb{R}}$ is the same as the one of $(\omega_{l,T}(t))_{t\in\mathbb{R}}$ (this comes from the fact that the red set in Figure 1.4 is homothetic to the gray set of Figure 1.3.).

In view of the result we present in the next subsection, note finally that the random measure M is a Gaussian multiplicative chaos with associated kernel

$$K(r) = \int_{|r|}^{+\infty} \frac{k(u)}{u}\, du \quad \text{with} \quad k(u) = \lambda^2(1 - \frac{|u|}{T})\mathbf{1}_{[0,T]}(|u|). \tag{1.11}$$

Before, turning to the whole description of all the random measures that are solutions of (1.7), let us make a few remarks on the link between the two notions of stochastic scale invariance we have just seen.

Note that the star scale invariance equation (1.7) is a global notion in the space $\mathbb{R}^d$ in contrast with the exact stochastic scale invariance equation (1.6) which is restricted on a compact euclidean ball (it is easily proved that the unique exact scale invariant stationary random measure on the whole space is the null measure).

The reader may wonder if the two notions are related or if there is a hierarchical relation between the two's. The answer is not yet complete but we were able to prove in [5] that if a Gaussian multiplicative chaos M has a kernel of the form $K(x,y) = \gamma^2 \ln_+(\frac{R}{|x-y|})$ (in particular we have already seen that M satisfies the exact stochastic scale invariance property), then it is *not* a lognormal star scale invariant measure.

In the next subsection, we explain the results we obtained in collaboration with R. Rhodes and V. Vargas. The reader can find a complete proof of this result in Chapter 3 (see also [5]). We characterize all the random measures that satisfy (1.7) under weak regularity assumptions for the process ω_ε and with the additional assumption that M possesses a moment of order $1+\delta$ (with $\delta > 0$).

1.0.4 Our contribution

In the following, we will say that a stationary random measure M satisfies the good lognormal star scale invariance if M is lognormal star scale invariant with some additional weak regularity assumptions on the covariance function k_ε of the process ω_ε, for all $\varepsilon < 1$.

Theorem 1.1 (R.A., R. Rhodes, V. Vargas)**.** *Let M be a good lognormal star scale invariant random measure. Assume that*

$$\mathbb{E}[M([0,1])^{1+\delta}] < +\infty$$

for some $\delta > 0$. Then M is the product of a nonnegative random variable $Y \in L^{1+\delta}$ and an independent Gaussian multiplicative chaos

$$\forall A \subset \mathcal{B}(\mathbb{R}), \quad M(A) = Y \int_A e^{X_r - \frac{1}{2}\mathbb{E}[X_r^2]}\, dr \tag{1.12}$$

with associated covariance kernel given by the improper integral

$$K(r) = \int_{|r|}^{+\infty} \frac{k(u)}{u}\, du \tag{1.13}$$

for some continuous covariance function k such that

$$k(0) \leqslant \frac{2}{1+\delta}. \tag{1.14}$$

Conversely, given some datas k and Y as above, the relation (3.11) defines a log-normal star scale invariant random measure M with finite moments of order $1+\gamma$ for every $\gamma \in [0, \delta)$.

The inequality (1.14) implies that the Gaussian multiplicative chaos associated to the kernel K defined in (1.13) has a moment of order $1+\delta$, as expected.

It seems natural to describe all the stochastic (star or exact) scale invariant random measures. Theorem 1.1 is a step in this direction as it characterizes all the (good) star scale invariant stationary random measures with a moment of order $1+\delta$. We hope this result to be useful to prove convergence to Gaussian multiplicative chaos. Indeed the star scale invariance equation (1.7) can be seen as a fixed point equation satisfied by the measure M and should be easy to verify in practice for a limiting random measure. If an unknown limiting random measure has this property, our theorem enables to identify this random measure as a Gaussian multiplicative chaos and gives a formula for its covariance structure.

Let us mention that the more general case where the lognormal factor $e^{\omega_\varepsilon(x)}$ is replaced by a log-Lévy factor has been solved later in [133].

Chapter 2

Random matrix theory and its applications

Random matrix theory (RMT) has become one of the prominent field of research, at the boundary between atomic physics, solid state physics, statistical mechanics, statistics, probability theory and number theory [3, 20, 15]. The start of the field is usually attributed to the work of Wigner, motivated by applications in nuclear physics. In 1951, Wigner suggested in [152] that the fluctuations in positions of the energy levels of heavy nuclei could be described in terms of statistical properties of eigenvalues of very large real symmetric matrices with independent identically distributed entries. This postulate has led to random matrix theory which is essentially concerned with the study of large symmetric random matrices with i.i.d. entries, called Wigner matrices but also in the study of other type of random matrices, see below. This theory aims at describing the statistical properties of large random matrices and particularly those related to the eigenvalues and eigenvectors. The first major result in this direction is due to Wigner (see [151]) in 1957: the empirical eigenvalue density of a Wigner random matrix converges almost surely in the limit of large dimension to the Wigner semicircle probability density. This probability density is compactly supported and shaped as a semi circle. Other eigenvalue statistics such as the largest eigenvalue statistics or local eigenvalues statistics have also been investigated in great details, see [15, 3, 114] for a review of RMT.

We now introduce another very popular type of random matrices of great interest for applications: the empirical covariance matrices, also called Wishart matrices. Historically, Wishart matrices were introduced before Wigner matrices by John Wishart in 1928 to study populations in biology through Principal Component Analysis (PCA). The purpose of PCA is to identify common causes (or factors) that govern the dynamics of N quantities. These quantities might be daily returns of the different stocks of the S&P 500, monthly inflation of different sectors of activity, motion of individual grains in a packed granular medium, or different biological indicators (blood pressure, cholesterol, ...) within a population. More precisely, if one is provided with a vector of (centered) correlated random variables $|x\rangle := (x_1, x_2, \cdots, x_N)$, whose covariance matrix is denoted as C and defined by $C_{ij} = \mathbb{E}[x_i x_j]$, PCA consists in writing the vector $|x\rangle$ in the orthogonal basis of the eigenvectors $|\phi_i\rangle$ of the covariance matrix C as

$$|x\rangle = \sum_{i=1}^{N} \langle x|\phi_i\rangle |\phi_i\rangle \,. \tag{2.1}$$

It is straightforward to see that the variance $\mathbb{E}[\langle x|\phi_i\rangle^2]$ of the random variable $\langle x|\phi_i\rangle$ is equal to λ_i where λ_i is the eigenvalue of the matrix C associated with the eigenvector $|\phi_i\rangle$. The decomposition (2.1) then shows that the favorite (or most likely) directions for the random vector $|x\rangle$ are the eigenvectors of C associated to the largest eigenvalues of the matrix C. Of course, practitioners do not have access to C; instead, they must consider a noisy empirical estimator E of the *true* covariance matrix C constructed from a sample of datas as

$$E_{ij} = \frac{1}{T} \sum_{t=1}^{T} x_i^t x_j^t \,. \tag{2.2}$$

where the vectors $|x^t\rangle := (x_1^t, \cdots, x_N^t)$ are independent and identically distributed as $|x\rangle$. The matrix E defines an empirical covariance matrix. Random matrix theory has thus been concerned, in the last decades, with the statistical properties of empirical covariance matrices. In particular the eigenvalue statistics (eigenvalues density, statistics of the largest eigenvalues, local eigenvalue statistics,...) for large dimensional empirical covariance matrices have been investigated in great details (see e.g. [15, 29, 113]) as in the Wigner case described above. The starting point of this active research area is the 1967 Marčenko Pastur paper [113] which gives an explicit formula for the empirical eigenvalue density in the limit of large dimension N. This result, as a new statistical tool to analyse large dimensional data sets, became very relevant in the last two decades, when the storage and handling of humongous data sets became routine in almost all fields – physics, image analysis, genomics, epidemiology, engineering, economics and finance, to quote only a few. Since their introduction, Wishart matrices have appeared in many different applications such as communication technology [134], nuclear physics [78], quantum chromodynamics [148], quantitative finance [44, 45, 124], statistical physics of directed polymers in random media [88] and non intersecting Brownian motions [135].

In this very short (and obviously far from exhaustive) introduction to the field, we introduce, in the first section, the main ensembles of random matrix theory, the so called Gaussian orthogonal ensemble (GOE) and Gaussian unitary ensemble (GUE). Those ensembles are the most classical and most studied random matrix ensembles in the literature, and were introduced by Wigner and Dyson at the starting point of the theory. We review the main statistical properties of the eigenvalues of the GOE/ GUE random matrices. In particular, we derive the joint law of the eigenvalues, establishing the link between random matrix theory and Coulomb gas repulsive interaction. For this derivation, we introduce the Dyson Brownian motion, which is a diffusive matrix process closely related to the GOE and GUE. We also introduce a generalization of those ensembles, the so called β-ensembles. At the end of the first section, we briefly explain some of our results which are related to the Gaussian ensembles. Then, in the second section, we give a fast review on the definition and main properties of empirical covariance matrices following the same line as in the first section and we sum up our contribution in the study of empirical covariance matrices at the end of the second section.

2.1 Gaussian Ensembles

We will denote by $\mathcal{H}_N^\beta$ with $\beta = 1$ (respectively $\beta = 2$) the space of symmetric real (resp. Hermitian complex) matrices of size $N \times N$. In the next subsection, we introduce the Gaussian Orthogonal and Unitary Ensembles of random matrices. The random matrices of those ensembles are random variables in the space $\mathcal{H}_N^1$ (respectively $\mathcal{H}_N^2$).

2.1.1 Classical Gaussian Ensembles

Gaussian Orthogonal Ensemble

A random matrix $\mathbf{H} \in \mathcal{H}_N^1$ is said to belong to the Gaussian Orthogonal Ensemble (GOE) if its probability law on the space $\mathcal{H}_N^1$ is given by:

$$P_N^1(\mathrm{d}H) = \frac{1}{Z_N^1} \exp\left(-\frac{N}{2}\operatorname{Tr}(H^\dagger H)\right) \mathrm{d}H \tag{2.3}$$

where $\mathrm{d}H$ denotes the Lebesgue measure on the space $\mathcal{H}_N^1$, Z_N^1 a normalization factor and where $H^\dagger$ denotes the Hermitian conjugate of H. The factor Z_N^1 can be explicitly computed (see [15, 114]).

It is clear from the definition of the law P_N^1 given by (2.3) that the random matrices in the Gaussian Orthogonal Ensemble are invariant under conjugation of orthogonal matrices, in the sense that the matrix $OHO^\dagger$ has the same law as the matrix H for any orthogonal matrix O.

If H is a real symmetric matrix, we have $\operatorname{Tr}(H^\dagger H) = \operatorname{Tr}(H^2) = \sum_{i,j=1}^N H_{ij}^2 = \sum_{i=1}^N H_{ii}^2 + 2\sum_{i<j} H_{ij}^2$. Therefore the entries of the random matrix H in the GOE are independent (up to symmetry) centered Gaussian variables with variance $1/N$ on the diagonal and variance $1/2N$ off the diagonal.

Gaussian Unitary Ensemble

The Gaussian Unitary Ensemble is similarly defined: the random matrices take values in the space $\mathcal{H}_N^\beta$ of Hermitian complex matrices of size $N \times N$ and are distributed according to the law P_N^2 defined as:

$$P_N^2(\mathrm{d}H) = \frac{1}{Z_N^2} \exp\left(-\frac{N}{2}\operatorname{Tr}(H^\dagger H)\right) \mathrm{d}H \tag{2.4}$$

where Z_N^2 is again a normalization factor (which can also be explicitly computed).

The symmetry for this ensemble is now with respect to the unitary matrices: the random matrices in the Gaussian Unitary Ensemble are invariant under conjugation of unitary matrices. The entries are independent (up to symmetry) complex centered Gaussian random variables. The diagonal entries are real centered Gaussian variables with variance $1/N$ whereas the off diagonal entries can be written as $\mathcal{N}(0, 1/2N) + \sqrt{-1}\mathcal{N}(0, 1/2N)$.

2.1.2 Dyson Brownian motion

Definition

In this subsection, the parameter β is equal to 1 or 2. In the following, the process $(H^\beta(t))_{t \geqslant 0}$ will denote a Dyson Brownian motion, i.e. a process with values in the set of $N \times N$ symmetric real ($\beta = 1$) or Hermitian complex ($\beta = 2$) matrices with entries $\left(H_{ij}^\beta(t), t \geqslant 0, 1 \leqslant i \leqslant j \leqslant d\right)$ constructed via independent real valued Brownian motions $\left(B_{ij}(t), \widetilde{B}_{ij}(t), 1 \leqslant i \leqslant j \leqslant d\right)$ by

$$H_{ij}^\beta(t) = \begin{cases} \frac{1}{\sqrt{2N}}(B_{ij}(t) + \sqrt{-1}(\beta - 1)\widetilde{B}_{ij}(t)) & \text{if } i < j\,, \\ \frac{1}{\sqrt{N}} B_{ii}(t) & \text{otherwise}\,. \end{cases} \tag{2.5}$$

The process $H^\beta(t)$ was first introduced by Dyson in [69].

We now want to define another diffusive matrix process that would converge in law in the limit of large time to the law of the Gaussian Ensembles defined previously. The idea is simply to define this matrix process $X^\beta(t)$ as solution of the following *Ornstein-Uhlenbeck* type equation

$$\mathrm{d}X^\beta(t) = -\frac{1}{2}X^\beta(t)\mathrm{d}t + \mathrm{d}H^\beta(t) \tag{2.6}$$

where $H^\beta(t)$ is a Dyson Brownian motion of size $N \times N$.

It is well known that the stationary law of a Ornstein-Uhlenbeck process is the Gaussian law and therefore, the real symmetric (resp. complex hermitian) matrix process $X^\beta(t)$ converges in law when $t \to \infty$ to the law of a GOE (resp. GUE) random matrix.

The study of the eigenvalues diffusion process of $X^\beta(t)$ will enable us to derive, in the next section, the joint law of the eigenvalues of the random matrices in the GOE and GUE.

Eigenvalues and eigenvectors diffusion processes

For each $t \geqslant 0$, the matrix $X^\beta(t)$ is a real symmetric matrix if $\beta = 1$ (resp. hermitian complex if $\beta = 2$) and therefore is diagonalizable in an orthonormal basis. The eigenvalues of $X^\beta(t)$ will be denoted[1] in increasing order as $\lambda_1(t) \leqslant \ldots \leqslant \lambda_N(t)$ and the associated orthonormal eigenvectors as $\psi_1(t), \ldots, \psi_N(t)$. Sometimes we will also use the following decomposition for the matrix $X^\beta(t)$:

$$X^\beta(t) = O^\beta(t)\Delta^\beta(t)O^\beta(t)^\dagger$$

where $\Delta^\beta(t)$ is the diagonal matrix $\mathrm{Diag}(\lambda_1(t), \ldots, \lambda_N(t))$ and where $O^\beta(t)$ is the orthogonal matrix (resp. unitary if $\beta = 2$) whose columns are (in respective order) given by $\psi_1(t), \ldots, \psi_N(t)$.

To find the stochastic differential system of equations verified by the eigenvalues and eigenvectors processes, a direct method is provided by using perturbation theory. Let us briefly recall the main ideas of this theory before coming back to our eigenvalues and eigenvectors processes.

[1] To simplify notations, we omit the subscript β for the eigenvalues and eigenvectors of $X^\beta(t)$.

Perturbation Theory. We are given a symmetric (resp. hermitian) matrix H_0 that is perturbed by the adding of a small symmetric (resp. Hermitian) matrix εP. Perturbation theory enables to find approximations of the eigenvalues and eigenvectors of the matrix H_1 defined as

$$H_1 = H_0 + \varepsilon P \tag{2.7}$$

in the limit $\varepsilon \to 0$. To second order in ε for the eigenvalues it gives

$$\lambda_i^1 = \lambda_i^0 + \varepsilon P_{ii} + \varepsilon^2 \sum_{j \neq i} \frac{|P_{ij}|^2}{\lambda_i^0 - \lambda_j^0} + o(\varepsilon^2) \tag{2.8}$$

where[2] $P_{ij} := \langle \phi_j^0 | P | \phi_i^0 \rangle$. For the eigenvectors, perturbation theory to second order writes as

$$\psi_i^1 = \left(1 - \frac{\varepsilon^2}{2} \sum_{j \neq i} \left(\frac{|P_{ij}|}{\lambda_i^0 - \lambda_j^0} \right)^2 \right) \psi_i^0 + \varepsilon \sum_{j \neq i} \frac{P_{ij}}{\lambda_i^0 - \lambda_j^0} \, \psi_j^0 + O(\varepsilon^2) \,, \tag{2.9}$$

where $O(\varepsilon^2)$ contains the second order transverse term (due to the non zero overlap of ϕ_i^1 with the non perturbed eigenvectors ψ_j^0 for $j \neq i$) which will turn out to be negligible in the context of Itô's stochastic calculus (see below). □

Physical derivation of the eigenvalues process. Coming back to our purpose, we can re-interpret the stochastic differential equation (2.6) verified by $X^\beta(t)$ as a perturbation equation by writting $X^\beta(t + \mathrm{d}t) = X^\beta(t) - \frac{1}{2} X^\beta(t) \mathrm{d}t + \mathrm{d}H^\beta(t)$. It is now straightforward to check that $\langle \psi_i(t) | X^\beta(t) | \psi_i(t) \rangle = \lambda_i(t)$ and that $\langle \psi_i(t) | X^\beta(t) | \psi_j(t) \rangle = 0$ for $i \neq j$ due to the orthogonality of the family $\{\psi_k(t)\}$. On the other hand, conditionally on $(\psi_k(t))_k$, the random variables $\langle \psi_i(t) | \mathrm{d}H^\beta(t) | \psi_j(t) \rangle, i < j$ are centered Gaussian (real or complex whether $\beta = 1$ or 2) random variables (as they are linear combination of independent Gaussian variables) which can be written as $\mathcal{N}(0, \frac{1}{2N}(1 + \delta_{i=j}) \, \mathrm{d}t)$ if $\beta = 1$ (respectively $\mathcal{N}(0, \mathrm{d}t/N)$ for $i = j$ and $\mathcal{N}(0, 1/2N) + \sqrt{-1}\mathcal{N}(0, 1/2N)$ for $i \neq j$ if $\beta = 2$). In addition we can check that they are independent (their covariance is zero again because of orthogonality). The second order terms in those perturbative equations are of order of the square of those random variables, i.e. of order $\mathrm{d}t$. Therefore, it is standard argument in Itô's calculus that the fluctuations of those second order terms are negligible: they can be replaced by their mean and we only keep the fluctuations of the first order terms.

Gathering the above arguments, we conclude that the eigenvalues verify the following Stochastic Differential System (SDS)

$$\mathrm{d}\lambda_i = -\frac{1}{2} \lambda_i \mathrm{d}t + \frac{1}{\sqrt{N}} \mathrm{d}b_i + \frac{\beta}{2N} \sum_{i \neq j} \frac{\mathrm{d}t}{\lambda_i - \lambda_j} \tag{2.10}$$

where the b_i are independent standard Brownian motions and with $\beta = 1$ or 2 depending on the symmetry class (symmetric or hermitian). □

For the eigenvectors, the evolution is described as follows. Let $w_{ij}^\beta(t), i < j$ be a family of real or complex (whether $\beta = 1$ or 2) Brownian motions (i.e. $w_{ij}^\beta(t) = \frac{1}{\sqrt{2}}(B_{ij}^1(t) + \sqrt{-1}\,(\beta - 1) B_{ij}^2(t))$ where the B_{ij}^1, B_{ij}^2 are standard Brownian motions on $\mathbb{R}$), independent of the family of Brownian motions $\{b_i\}$ involved in (2.10). For $i < j$, set in addition $w_{ji}^\beta(t) := \bar{w}_{ij}^\beta(t)$. The system of Stochastic Differential Equation (SDE) verified by the eigenvectors can then be written as

$$\mathrm{d}\psi_i = -\frac{\beta}{4N} \sum_{j \neq i} \frac{\mathrm{d}t}{(\lambda_i - \lambda_j)^2} \, \psi_i + \frac{1}{\sqrt{N}} \sum_{j \neq i} \frac{\mathrm{d}w_{ij}^\beta}{\lambda_i - \lambda_j} \, \psi_j \,. \tag{2.11}$$

To write this evolution in terms of the matrix $O^\beta(t)$, define the skew Hermitian matrix (i.e. such that $R^\beta = -(R^\beta)^*$) by setting for $i \neq j$,

$$dR_{ij}^\beta(t) = \frac{1}{\sqrt{N}} \frac{dw_{ij}^\beta(t)}{\lambda_i(t) - \lambda_j(t)}, \quad R_{ij}^\beta(0) = 0 \,.$$

[2] We use the classical bracket notations for the (hermitian) scalar product.

Then, with $\lambda_i(t)$ being the solution of the SDS (2.10), the matrix $O^\beta(t)$ evolves as

$$dO^\beta(t) = O^\beta(t)dR^\beta(t) - \frac{\beta}{4}O^\beta(t)d\langle (R^\beta)^\dagger, R^\beta\rangle_t\,. \tag{2.12}$$

To show more rigorously that Equations (2.10) and (2.11) (or equivalently (2.10) and (2.12)) hold, the proof usually goes *backward*, through the following steps (see [15, Lemmas 4.3.3 and 4.3.4]):

- One first has to show that the SDS (2.10) is well defined for all time $t \geqslant 0$. One can indeed show that the λ_i solution of (2.10) almost surely never collide, in the sense that the first collision time $T_1 := \inf\{t \geqslant 0 : \exists i \neq j, \lambda_i(t) = \lambda_j(t)\}$ is almost surely infinite.

- Then one can show that the unique solution of the SDE (2.12) (this SDE indeed has a unique solution as it is linear in O^β and R^β is a well defined martingale) is a matrix process with values in the space of orthogonal (respectively unitary) matrices.

- The last step uses standard Itô's calculus to show that the matrix process $Y^\beta(t)$ defined as $Y^\beta(t) := O^\beta(t)\Delta^\beta(t)O^\beta(t)^\dagger$, with O^β the solution of sde (2.12) and with $\Delta^\beta(t)$ the diagonal matrix $\mathrm{Diag}(\lambda_1(t), \ldots, \lambda_N(t))$ where the λ_i are solution of the system (2.10), indeed verifies the stochastic differential equation (2.6) (with Y instead of X).

2.1.3 Joint law of the eigenvalues

As already mentioned, the matrix $X^\beta(t)$ converges in law when $t \to \infty$ to the law P_N^β of the random matrices in the GOE if $\beta = 1$ and in the GUE if $\beta = 2$. Therefore a simple way to determine the joint probability of the eigenvalues of the random matrices in the GOE (resp. GUE) is to find the limiting law of the eigenvalues process of $X^\beta(t)$ for $\beta = 1$ (resp. $\beta = 2$). The classical way to do this is to compute a stationary probability of the SDS (2.10) through the Fokker-Planck equation (or through the infinitesimal generator associated to (2.10)) and then to show that this stationary probability is the unique such measure.

The Fokker Planck equation gives a partial differential equation satisfied by the probability density transition function of a diffusion given its stochastic differential equation. For the process $(\lambda_1(t), \ldots, \lambda_N(t))$ that verifies (2.10), the transition function $P(\lambda_1, \ldots, \lambda_N; t)$ verifies:

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{N}\frac{\partial}{\partial \lambda_i}\left[\left(-\frac{\lambda_i}{2} + \frac{\beta}{2N}\sum_{j\neq i}\frac{1}{\lambda_i - \lambda_j}\right)P\right] + \frac{1}{2N}\sum_{i=1}^{N}\frac{\partial^2 P}{\partial \lambda_i^2}\,. \tag{2.13}$$

The stationary solutions of (6.81) are found by setting the time derivative to 0. It is easy using elementary algebra to show that a stationary solution to equation (6.81) is given by:

$$P_\beta(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_\beta}\prod_{1\leqslant i<j\leqslant N}|\lambda_i - \lambda_j|^\beta \exp(-\frac{N}{2}\sum_{i=1}^{N}\lambda_i^2)\prod_{i=1}^{N}\mathrm{d}\lambda_i\,. \tag{2.14}$$

Furthermore it is the unique stationary probability measure for the process $(\lambda_1(t), \ldots, \lambda_N(t))$ since if there was another invariant distribution Q_β, we could reconstruct a Hermitian Ornstein-Uhlenbeck process $\widetilde{X}^\beta(t)$ and a matrix $\widetilde{X}_0^\beta$ whose eigenvalues would follow Q_β so that $\widetilde{X}^\beta(0) := \widetilde{X}_0^\beta$ and

$$\mathrm{d}\widetilde{X}^\beta(t) = -\frac{1}{2}\widetilde{X}^\beta(t)\mathrm{d}t + \mathrm{d}H^\beta(t)\,.$$

But this gives a contradiction since as time goes to infinity, the law of $\widetilde{X}^\beta(t)$ is a Gaussian law, independently of the law Q_β.

Other derivations of formula (6.5) exist and can be found for example in [15].

2.1.4 Matrix model for β-ensembles

The probability measure P_β introduced in (6.5) is a Gibbs measure of a one-dimensional repulsive Coulomb gas confined in an harmonic well with inverse temperature β and is in fact a well defined probability measure for all $\beta > 0$. In the previous section, we have introduced two random matrix ensembles whose eigenvalues are distributed according to the law $P_{\beta=1}$ for the GOE symmetric real matrices and $P_{\beta=2}$ for the GUE hermitian complex matrices. Another random matrix ensemble, which was not introduced above and called the Gaussian Symplectic Ensemble (GSE), is invariant under conjugation by a symplectic matrix and has eigenvalues distributed according to the law $P_{\beta=4}$. Therefore, depending on the symmetry of the random matrix, only three values are allowed $\beta = 1, 2$ and 4 up to now. This is known as Dyson's "threefold way". The existence of matrix ensembles that would lead to general values of $\beta > 0$, is a very natural question, and the quest for such ensembles probably goes back to Dyson himself.

Ten years ago, Dumitriu and Edelman [68] have proposed the following explicit construction of tri-diagonal random matrices, with eigenvalues distributed according to P_β for general $\beta > 0$,

$$H^\beta = \frac{1}{\sqrt{2N}} \begin{pmatrix} g_1 & \chi_{(N-1)\beta} & & & \\ \chi_{(N-1)\beta} & g_2 & \chi_{(N-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & g_{N-1} & \chi_\beta \\ & & & \chi_\beta & g_N \end{pmatrix} \tag{2.15}$$

where the g_k are independent Gaussian random variables with variance 2 and where the $\chi_{k\beta}$ are independent χ distributed random variables with parameters $k\beta$ and scale parameter 2.

Another construction is proposed in [3, page 426-427] (see also [74]) and uses a bordering procedure to construct iteratively on the dimension a sequence of matrices with eigenvalues distributed according to P_β for general $\beta > 0$. The advantage of this construction is that it gives not just the eigenvalue probability density of one matrix but also the joint eigenvalue probability density of all matrices of the sequence.

2.1.5 Eigenvalues density

Let H be an $N \times N$ symmetric (or hermitian) random matrix and denote by λ_i the (real) eigenvalues of H, with $\lambda_1 \leqslant \cdots \leqslant \lambda_N$. The empirical distribution of the eigenvalues of H is the following (random) probability measure on $\mathbb{R}$

$$\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i) \, .$$

The measure μ_N is sometimes called spectral measure or density of states and is one of the central object of study in random matrix theory. In the following subsection, we briefly recall the main classical results on the empirical distribution of the eigenvalues in the limit of large random Wigner matrices, which are symmetric or hermitian random matrices with independent entries (up to symmetry).

Wigner matrices

A Wigner symmetric real (respectively hermitian complex) random matrix H can be defined from a family of independent and identically distributed (i.i.d.) real (respectively complex) centered random variables $\{h_{ij}\}_{1 \leqslant i \leqslant j < \infty}$ living on a common probability space $(\Omega, \mathbb{P})$ by setting

$$H_N(ij) = \begin{cases} \frac{1}{\sqrt{2N}} h_{ij}, & \text{if} \quad i < j, \\ \frac{1}{\sqrt{N}} h_{ii}, & \text{otherwise} \, . \end{cases} \tag{2.16}$$

If the entries have a finite second moment $\sigma^2 = \mathbb{E}[h_{ij}^2]$, then Wigner's theorem (see [151]) asserts that the empirical spectral measure of the matrix $H_N(ij)$ converges weakly almost surely to the semicircle distribution

$$\rho(\mathrm{d}x) = \frac{1}{\pi\sigma^2} \sqrt{2\sigma^2 - x^2} \, \mathrm{d}x \, . \tag{2.17}$$

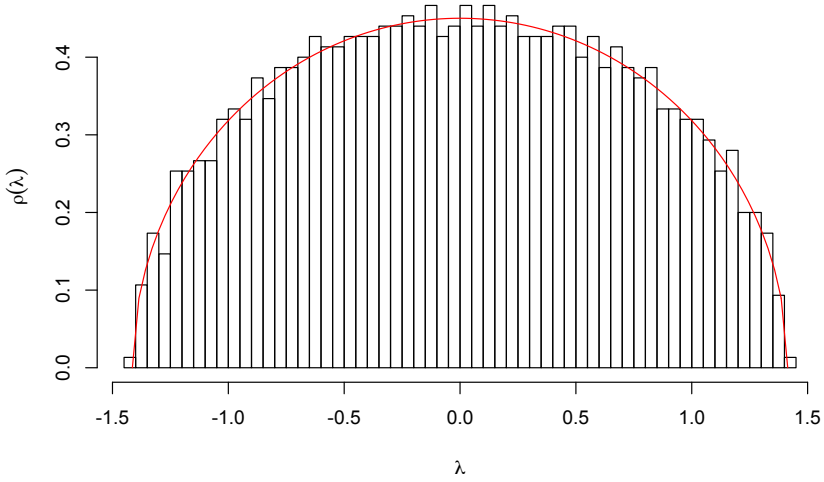


Figure 2.1: The histogram represents the eigenvalue density of a simulated GOE random matrix of size $N = 1500$. The (red) curve is the Wigner semicircle density.

An illustration of this convergence is displayed in Fig. 2.1.

The case where the entries have infinite second moment was also treated in [43, 33]. The entries are assumed to be in the domain of attraction of an α- stable law, for $\alpha \in (0; 2)$, i.e. they verify the following asymptotic for the tail

$$\mathbb{P}[h_{ij} \geqslant u] = \frac{L(u)}{u^\alpha},$$

where L is a slowly varying function. Under this assumption, the authors of [33] show that the empirical spectral measure μ_N of the matrix H_N (where this time the entries of the matrix are renormalized by a sequence a_N, instead of $\sqrt{N}$ in (2.16), of normalizing constants which roughly grows as $N^{1/\alpha}$) converges weakly (in probability, say) to a measure μ_α. Moreover it is also shown in [33] that the measure μ_α is symmetric, has unbounded support and a smooth density $\rho_\alpha(x)$ outside a small subset. Even the asymptotic of the tails of $\mu_\alpha(\mathrm{d}x) = \rho_\alpha(x)\ \mathrm{d}x$ is known: there exists a constant $L_\alpha > 0$ such that

$$\rho_\alpha(x) \sim \frac{L_\alpha}{x^{\alpha+1}}.$$

Note that the GOE and GUE random matrices introduced above are particular cases of Wigner matrices with Gaussian entries. In the next subsection, we will show a possible way to recover the asymptotic of the empirical spectral distribution for those ensembles and more generally in the case of general β-ensembles.

Proof of the Wigner semicircle law for β-ensembles

We first need to introduce the Stieltjes transform of a probability measure. If μ is a probability measure on $\mathbb{R}$, its Stieltjes transform is defined for all $z \in \mathbb{C} \setminus \mathbb{R}$ as

$$G(z) = \int \frac{\mu(\mathrm{d}x)}{x - z}.$$

The measure μ is characterized by its Stieltjes transform and its values on bounded continuous test functions can be recovered from $G(z)$ by a limiting procedure $z \to x \in \mathbb{R}$. In particular, if the measure μ has a continuous density $\rho(x)$ with respect to Lebesgue measure, we have the classical inversion formula

$$\lim_{\epsilon \to 0} \Im(G(x - i\epsilon)) = \pi\rho(x). \tag{2.18}$$

We now show that the empirical spectral distribution converges weakly to the Wigner semicircle density for general β-ensembles. Let $(\lambda_1(t), \dots, \lambda_N(t))$ be the diffusion process that verify the stochastic differential system (2.10) (with $\beta > 0$) and denote by $G_N(z,t)$ the associated Stieltjes transform of the empirical distribution of the $\lambda_i(t)$. The following relation holds

$$G_N(z,t) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i(t) - z} \,. \tag{2.19}$$

The idea of the proof is to study the evolution of $G_N(z,t)$ with respect to time t and in particular its convergence to equilibrium when $t \to \infty$. We already know that the diffusion process $(\lambda_1(t), \dots, \lambda_N(t))$ converges in law in the limit of large time t to P_β. Therefore, $G_N(z, t = \infty)$ corresponds to the Stieltjes transform of the empirical spectral distribution of a random matrix H^β in the β-ensembles.

Applying Itô's formula to the functional $\frac{1}{N}\sum_{i=1}^{N} f(\lambda_i(t))$ where f is the smooth function $f(\lambda) = 1/(\lambda - z)$ and using the SDS (2.10) verified by the λ_i, we obtain the following Langevin equation for $G_N(z,t)$

$$2\,\mathrm{d}G_N = \frac{\beta}{2}\frac{\partial G_N^2}{\partial z}\,\mathrm{d}t + \frac{\partial z G_N}{\partial z}\,\mathrm{d}t + \frac{1}{2N}(2-\beta)\,\frac{\partial^2 G_N}{\partial z^2}\,\mathrm{d}t + \mathrm{d}M_t^N \tag{2.20}$$

where $\mathrm{d}M_t^N = -\frac{2}{N}\sum_{i=1}^{N} \frac{1}{(\lambda_i - z)^2}\,\frac{\mathrm{d}b_i}{\sqrt{N}}$ with quadratic variation

$$d\langle M^N \rangle_t = \frac{4}{N^3} \sum_{i=1}^{N} \frac{1}{(\lambda_i - z)^4}\,\mathrm{d}t \,.$$

In the large N limit, we easily see that the two last diffusion and martingale terms are negligible compared to the other terms and therefore we can rewrite the Eq. (2.20) in the limit $N \to \infty$ as a deterministic Burgers evolution equation

$$2\frac{\partial G}{\partial t} = \frac{\beta}{2}\frac{\partial G^2}{\partial z} + \frac{\partial z G}{\partial z} \,. \tag{2.21}$$

To leading order, the stationary equation associated to (2.21) (where the time derivative is set to 0 and which is satisfied by $G_N(z, t = \infty)$) can be integrated with respect to z as:

$$\frac{\beta}{2} G^2 + zG + 1 = 0 \tag{2.22}$$

where the integration constant comes from the boundary condition $G_N \sim -1/z$ when $|z| \to \infty$. It is then easy to solve this equation (2.22) to find the equilibrium Stieltjes transform

$$G(z, \infty) = \frac{1}{\beta}\left[\sqrt{z^2 - 2\beta} - z\right] \,.$$

It can now be checked with the inversion formula (2.18) that the associated empirical spectral distribution is indeed the Wigner semicircle density

$$\rho(\lambda) = \frac{1}{\pi\beta}\sqrt{2\beta - \lambda^2}, \quad -\sqrt{2\beta} \leqslant \lambda \leqslant \sqrt{2\beta} \,. \tag{2.23}$$

2.1.6 Our contribution

Let us briefly explain our results related to Gaussian ensembles. For further details on those results, we refer to the following chapters.

In the first paragraph, we define a new diffusive matrix model converging towards the β-Dyson Brownian motion for all $\beta \in [0, 2]$ that provides an explicit construction of β-ensembles of random matrices that is invariant under the orthogonal/unitary group. We show in the second paragraph that, for small values of β, our process allows one to interpolate smoothly between the Gaussian distribution and the Wigner semicircle. The interpolating limit distributions form a one parameter family and can be explicitly computed. In the third paragraph, we use the

previous construction to compute the corrections for the empirical eigenvalue distribution of β-ensembles to the Wigner semicircle density for large but finite dimension. The last paragraph concerns the eigenvectors of the GOE random matrices. We propose a general framework to study the stability of the subspace spanned by P consecutive eigenvectors of a generic symmetric matrix $\mathbf{H}_0$, when a small perturbation is added. This problem is relevant in various contexts, including quantum dissipation (H_0 is then the Hamiltonian) and financial risk control (in which case $\mathbf{H}_0$ is the assets return covariance matrix). We state our results later in chapter 7 in the case where H_0 is a GOE random matrix or when H_0 is a covariance matrix.

A diffusive matrix model for invariant β-ensembles [joint work with Alice Guionnet, see also chapter 5 or [10]].

The goal of this work is to provide a natural interpretation of β-ensembles in terms of random matrices for $\beta \in [0, 2]$. Dumitriu and Edelman [68] already proposed the tridiagonal matrix introduced above with eigenvalues distributed according to P_β. However, this tridiagonal matrix lacks the invariant property of the classical ensembles (GOE is invariant under conjugation of an orthogonal matrix whereas GUE is invariant under the conjugation of a unitary matrix). The construction introduced in [10] has this property and moreover is constructive as it is based on a dynamical scheme. It was proposed by JP Bouchaud, and [10] provides rigorous proofs of the results stated in [6]. The idea is to interpolate between the Dyson Brownian motion and the standard Brownian motion by throwing a coin at every infinitesimal time step to decide whether our matrix will evolve according to a Dyson Brownian motion (with probability p) or will keep the same eigenvectors but with eigenvalues diffusing according to independent Brownian motions. When the size of the infinitesimal time steps goes to zero, we prove that the dynamics of the eigenvalues of this matrix valued process converges towards the β-Dyson Brownian motion as defined in (2.10) with $\beta = p$. The same construction with a hermitian Brownian motion leads to the same limit with $\beta = 2p$.

More precisely, our model is defined as follows: we divide time into small intervals of length $1/n$ and for each interval $[k/n; (k+1)/n]$, we choose independently Bernoulli random variables $\epsilon_k^n, k \in \mathbb{N}$ such that $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$. Then, setting $\epsilon_t^n = \epsilon_{[nt]}^n$, our diffusive matrix process simply evolves as:

$$\mathrm{d}M_n(t) = -\frac{1}{2}M_n(t)\mathrm{d}t + \epsilon_t^n \mathrm{d}H(t) + (1 - \epsilon_t^n)\mathrm{d}Y(t) \tag{2.24}$$

where $\mathrm{d}H(t)$ is a Dyson real Brownian increment as defined in (5.2) and where $\mathrm{d}Y(t)$ is a symmetric matrix that is co-diagonalizable with $M_n(t)$ (i.e. the two matrices have the same eigenvectors) but with a spectrum given by N independent Brownian increments of variance $\mathrm{d}t/N$. It is clear that the eigenvalues of the matrix $M_n(t)$ will cross at some points but only in intervals $[k/n; (k+1)/n]$ for which $\epsilon_k^n = 0$ (in the other intervals where they follow Dyson Brownian motion with parameter $\beta = 1$, it is well known that the repulsion is too strong and that collisions are avoided). In such a case, the eigenvalues are re-numbered at time $t = (k+1)/n$ in increasing order.

Now, using again standard perturbation theory, it is easy to derive the evolution of the eigenvalues of $M_n(t)$ denoted as $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t)$:

$$\mathrm{d}\lambda_i^n = -\frac{1}{2}\lambda_i^n \mathrm{d}t + \frac{\epsilon_t^n}{2N}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i^n - \lambda_j^n} + \frac{1}{\sqrt{N}}\mathrm{d}b_i \tag{2.25}$$

where the b_i are independent Brownian motions also independent of the $\epsilon_k^n, k \in \mathbb{N}$.

Recall that Cépa and Lépingle showed in [54] the uniqueness and existence of the strong solution to the stochastic differential system

$$d\lambda_i(t) = -\frac{1}{2}\lambda_i(t)dt + \frac{p}{2N}\sum_{j\neq i}\frac{1}{\lambda_i(t) - \lambda_j(t)}dt + \frac{1}{\sqrt{N}}\,db_i \tag{2.26}$$

starting from $\lambda(0) = (\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_d)$ and such that for all $t \geqslant 0$

$$\lambda_1(t) \leq \lambda_2(t) \leq \cdots \leq \lambda_d(t) \quad a.s. \tag{2.27}$$

For the scaling limit of the *ordered* eigenvalues, we prove in Chapter 5 (and also in [10]) that

Theorem 2.1 (R. A., A. Guionnet). *Let M_0^β be a symmetric (resp. Hermitian) matrix if $\beta = 1$ (resp. $\beta = 2$) with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_d$ and $(M_n^\beta(t))_{t\geq 0}$ be the matrix process defined in Definition 5.1. Let $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_d^n(t)$ be the ordered eigenvalues of the matrix $M_n^\beta(t)$. Let also $(\lambda_1(t), \ldots, \lambda_d(t))_{t \geqslant 0}$ be the unique strong solution of (5.5) with initial conditions in $t = 0$ given by $(\lambda_1, \lambda_2, \ldots, \lambda_d)$.*

Then, for any $T < \infty$, the process $(\lambda_1^n(t), \ldots, \lambda_d^n(t))_{t\in[0,T]}$ converges in law as n goes to infinity towards the process $(\lambda_1(t), \ldots, \lambda_d(t))_{t\in[0,T]}$ in the space of continuous functions $\mathcal{C}([0,T], \mathbb{R}^d)$ embedded with the uniform topology.

One of the difficulty of the proof comes from the fact that when $p < 1$, there is a positive probability for eigenvalues verifying (2.10) to collide in finite time (the ordering constraint is therefore useful at those points to restart). The idea is then to show that collisions are in a sense sufficiently rare for the SDS (2.10) (with $\beta = p < 1$) to make sense (see [10, 54] or chapter 5 for further details).

The corresponding scaling limit of the matrices $M_n(t)$, denoted as $M(t)$, is furthermore *invariant* under the orthogonal (or unitary) group. This is intuitively clear, since both alternatives in the evolution of $M_n(t)$ (adding a free slice or adding a commuting slice) respect this invariance, and lead to a Haar probability measure for the eigenvectors (i.e. uniform over the orthogonal/unitary group). We have also proved in [10] that a collision leads to a complete randomization of the eigenvectors within the two-dimensional subspace corresponding to the colliding eigenvalues.

To check numerically Theorem 2.1, a possible way is to check that indeed the matrix process $M_n(t)$ share the known properties of the spectrum of β ensembles, for large time t (and large n for $M_n(t)$ to be near its scaling limit). Indeed we have seen that the eigenvalue density of β-ensembles converges for large matrices to the Wigner semicircle density. Moreover the behavior of the nearest neighbor spacing distribution (NNSD) $P(s)$ is expected to behave as s^β near 0. We simulated numerically the matrix $M_n(t)$ with $N = 200$ for a very small step $1/n$ and until a large value of t so as to reach the stationary distribution for the eigenvalues. Then we started recording the spectrum and the nearest neighbor spacings (NNS) every 100 steps so as to sample the ensemble. We verified that the spectral density of $M_n(t = \infty)$ is indeed in very good agreement with the Wigner semi-circle distribution for $\beta = 1/2$ (see Fig. 4.1). Our sample histogram for the NNS distribution is displayed in Fig.4.1. We also added the corresponding Wigner surmise (which is expected to provide a good approximate description of the NNSD).

Invariant β-ensembles and the Gauss-Wigner crossover [joint work with Jean-Philippe Bouchaud and Alice Guionnet, see also chapter 4 or [6]].

The construction (6.13) of the previous subsection leads to the β-ensembles (with in fact $\beta = p$) introduced above as soon as the parameter $p > 0$ and we have seen that in this case the empirical spectral distribution is given by the Wigner semicircle law with edges at $\pm\sqrt{2p}$.

On the other hand, if $p = 0$, the eigenvalues process of the diffusive matrix verifying (6.13) is given by N independent Ornstein-Uhlenbeck processes. The empirical eigenvalues distribution is therefore in this case given by the Gaussian distribution.

This raises the question of the existence of an interpolation between these two regimes. A continuous cross-over indeed takes place for $\beta = p = 2c/N$ with c strictly positive and independent of N. The SDE for the limiting eigenvalues process $(\lambda_i(t))$ is again given by (2.10) with the additional ordering constraint $\lambda_1(t) \leqslant \ldots \leqslant \lambda_N(t)$ (necessary to restart after collisions occurred) and the stationary joint probability density function (pdf) is still given by (6.5) but with now the vanishing repulsion coefficient $\beta = 2c/N$.

We can proceed along the same steps as in subsection 2.1.5. We restart from Equation (2.20). In the present scaling, we expect the spectrum to have a width of order $\sqrt{p} \propto 1/\sqrt{N}$ and therefore we can easily check that in this scaling, the martingale term of (2.20) is negligible compared to the other terms which are now all of the same order (the second derivative term is no longer negligible!). The stationary differential equation derived from equation (2.20) can be integrated with respect to z and after a further rescaling to make the support of the eigenvalues

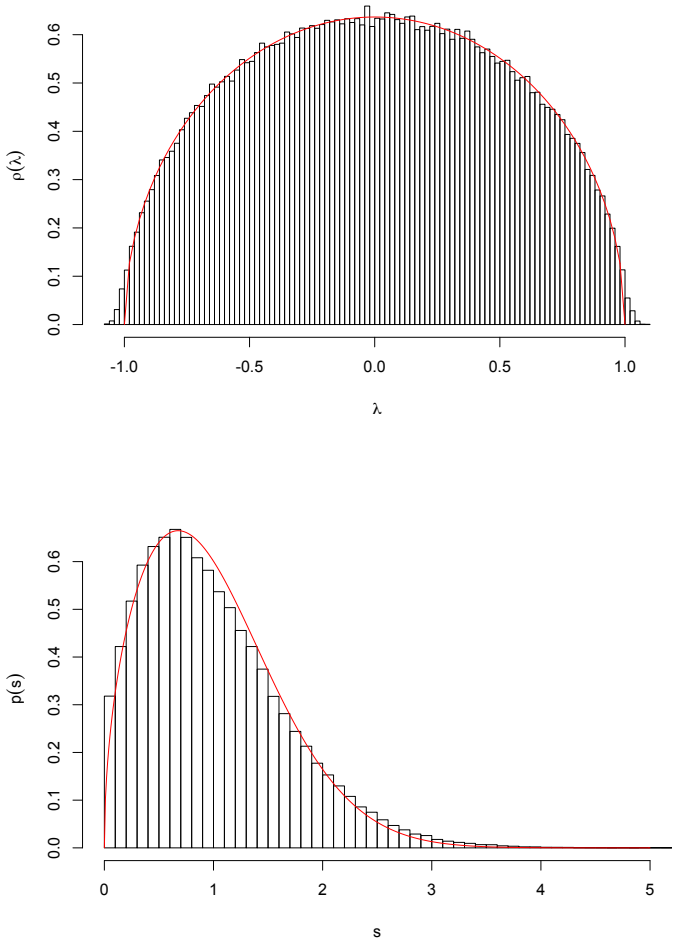


Figure 2.2: Up: Empirical eigenvalue distribution for the matrix $M_n(t = \infty)$ with the semicircle density for $N = 200, \beta = 1/2$. Down: Empirical NNSD $P(s)$ for the matrix $M_n(t = \infty)$ for $\beta = p = 1/2$ with the Wigner surmise (red curve) corresponding to $\beta = \frac{1}{2}$, which behaves as s^β when $s \to 0$.

density of order 1, we obtain:

$$c\,G^2 + zG + \frac{\mathrm{d}G}{\mathrm{d}z} = -1, \tag{2.28}$$

where the integration constant comes from the boundary condition $G \sim -1/z$ for $z \to \infty$. Equation (4.12) can be also recovered directly from the saddle point equation route starting from the joint pdf P_β with $\beta = 2c/N$ (this method is also presented in [6]).

Equation (4.12) can be explicitly solved and leads to the following eigenvalues density

$$\rho_c(\lambda) = \frac{1}{\sqrt{2\pi}\Gamma(1+c)} \frac{1}{|D_{-c}(i\lambda)|^2}; \tag{2.29}$$
$$D_{-c}(z) = \frac{e^{-z^2/4}}{\Gamma(c)} \int_0^\infty \mathrm{d}x e^{-zx-\frac{x^2}{2}} x^{c-1}.$$

The probability density ρ_c is the asymptotic eigenvalues density of a matrix H^β defined as in (2.15) with $\beta = 2c/N$ or equivalently the asymptotic empirical density of a random vector $(x_1, \ldots, x_N)$ with distribution P_β (with again $\beta = 2c/N$). See the progressive deformation of the Gaussian towards Wigner's semi-circle in Fig. 2.3.

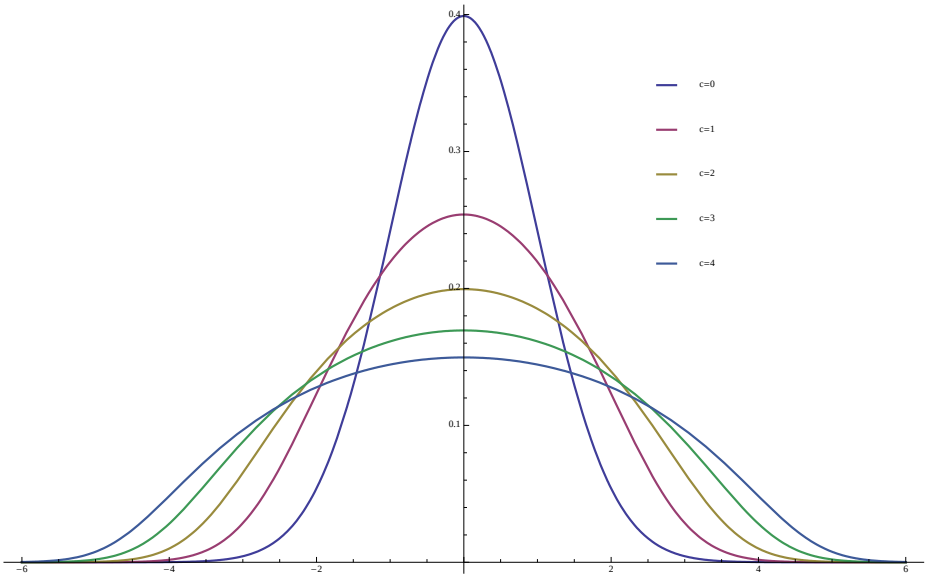


Figure 2.3: Density $\rho_c(u)$ for $c = 0, 1, 2, 3, 4$, showing the progressive deformation of the Gaussian towards Wigner's semi-circle.

This family of distributions is indeed a crossover (or an interpolation) between the Wigner semicircle and the Gaussian distributions, as we have on the one hand, for $c = 0$,

$$\rho_0(\lambda) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{\lambda^2}{2})$$

and on the other hand, for $c \to \infty$,

$$\rho_c(\lambda) \sim \frac{1}{2\pi c} \sqrt{4c - \lambda^2}\,.$$

Wigner correction for large but finite dimension [joint work with Jean-Philippe Bouchaud, Satya N. Majumdar and Pierpaolo Vivo, see also chapter 6 or [12]].

We now explain how to derive a $1/N$ expansion for the asymptotic mean empirical eigenvalues distribution at order $1/N^2$ for matrices in the GOE and in the GUE but only at order $1/N$ for general β-ensembles. More precisely, by denoting $\rho_N(\lambda)$ the density (with respect to Lebesgue measure) of the probability density $\rho_N^\beta(\lambda) := \mathbb{E}[\frac{1}{N}\sum_{i=1}^N \delta(\lambda - \lambda_i)]$ where the λ_i are the eigenvalues of a GOE or GUE random matrix, we want to determine explicitly the $1/N$ expansion of ρ_N written in the form

$$\rho_N(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2} + \frac{1}{N}\rho_1(\lambda) + \frac{1}{N^2}\rho_2(\lambda) + O(\frac{1}{N^3})\,. \tag{2.30}$$

We describe in the following how to find explicit formulas for the correction functions (not necessarily positive) ρ_1 and ρ_2.

Let us return to (2.20) for $\beta = 1$ or 2. We consider the stationary differential equation associated to (2.20) (i.e. when the derivative with respect to time is set to 0) as we have done in the previous subsection. By rescaling[3] the eigenvalues as $\lambda \leftarrow \lambda\sqrt{2}/\sqrt{\beta}$ and by taking expectation, it is easy to see that equation (2.20) can be rewritten as

$$\frac{\mathrm{d}\mathbb{E}[G_N]^2}{\mathrm{d}z} + \frac{\mathrm{d}z\mathbb{E}[G_N]}{\mathrm{d}z} + \frac{1}{N}\frac{2-\beta}{\beta}\frac{\mathrm{d}^2\mathbb{E}[G_N]}{\mathrm{d}z^2} + \frac{\mathrm{d}}{\mathrm{d}z}\left(\mathbb{E}[G_N^2] - \mathbb{E}[G_N]^2\right) = 0\,. \tag{2.31}$$

It turns out that the limit when $N \to \infty$ of the last term has been explicitly computed in [105]. It is shown in [105] that

$$\lim_{N\to\infty} N^2\left(\mathbb{E}[G_N^2] - \mathbb{E}[G_N]^2\right) = \frac{1}{2\beta\pi^2}\int_{-2}^{2}\int_{-2}^{2}\frac{1}{(\lambda-z)^2(\mu-z)^2}\frac{4-\lambda\mu}{\sqrt{4-\lambda^2}\sqrt{4-\mu^2}}\mathrm{d}\lambda\mathrm{d}\mu \tag{2.32}$$

for $\beta \in \{1, 2\}$. By denoting V_β the right hand side of (2.32), we can integrate (2.31) with respect to z, neglecting terms of order N^{-3}, as

$$\mathbb{E}[G_N]^2 + z\mathbb{E}[G_N] + \frac{1}{N}\frac{2-\beta}{\beta}\frac{\mathrm{d}\mathbb{E}[G_N]}{\mathrm{d}z} + \frac{1}{N^2}V_\beta = -1 \tag{2.33}$$

where the integration constant is again chosen according to the boundary condition $G \sim -1/z$ when $|z| \to \infty$. The interesting fact here is that the integral V_β can be computed analytically as a function of z. We find

$$V_\beta = \frac{2}{\beta}\frac{1}{(z^2-4)^2}\,.$$

Then, using perturbation theory in (2.33), we can compute explicitly the coefficients ρ_1 and ρ_2 in expansion (2.30).

The result reads:

$$\rho_N^\beta(\mathrm{d}\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}\,\mathrm{d}\lambda \tag{2.34}$$

$$+\frac{1}{N}\left(\frac{1}{\beta}-\frac{1}{2}\right)\left(\frac{1}{2}\left(\delta(\lambda-2)+\delta(\lambda+2)\right) - \frac{1}{\pi}\frac{1}{\sqrt{4-\lambda^2}}\right)\mathrm{d}\lambda \tag{2.35}$$

$$+\frac{1}{\pi}\left(\frac{1}{\beta}-\frac{1}{2}\right)^2\frac{1}{\sqrt{4-\lambda^2}}\left[\frac{3}{2}\frac{1}{\lambda^2-4} + \frac{5}{4}\left(\frac{1}{(\lambda+2)^2}+\frac{1}{(\lambda-2)^2}\right)\right]\frac{1}{N^2} \tag{2.36}$$

$$-\frac{1}{\beta\pi}\frac{\mathrm{d}\lambda}{(4-\lambda^2)^{5/2}}\frac{1}{N^2} + O(\frac{1}{N^3})\,. \tag{2.37}$$

A numerical evidence for this formula is shown in 2.4. The issue of extracting subleading corrections to the asymptotic semicircular density was addressed in quite a few papers [75, 89, 63, 96]. Whereas the results in [89] coincide with those derived here, other works [75, 63, 96] found oscillatory corrections of order $1/N$. Those oscillatory terms contribute to higher order after integration. In [75], the authors discuss the origin of the Dirac mass at the edges.

Stability of eigenspaces [joint work with Jean-Philippe Bouchaud, see also chapter 7 or [7]].

We are now interested in the eigenvectors stability of GOE random matrices when a small GOE perturbation matrix εP is added. Our aim is to understand the overlap between the eigenvectors of a non-perturbed GOE matrix H_0 with the eigenvectors of the perturbed matrix H_1 obtained from H_0 by adding a small perturbation εP as in equation (2.7) where P is a GOE matrix and ε is a small parameter.

It is quite clear, in view of the perturbation equation for the eigenvectors (7.2), that it will be difficult to follow the evolution of one single eigenvector when the perturbation is added if the perturbation entries are too large compared to the eigenvalue spacing of the GOE matrix

[3] This rescaling is chosen so that the eigenvalue density, in the limit of large matrices, is the Wigner semicircle density with the classical support $[-2; 2]$ instead of $[-\sqrt{2\beta}; \sqrt{2\beta}]$.

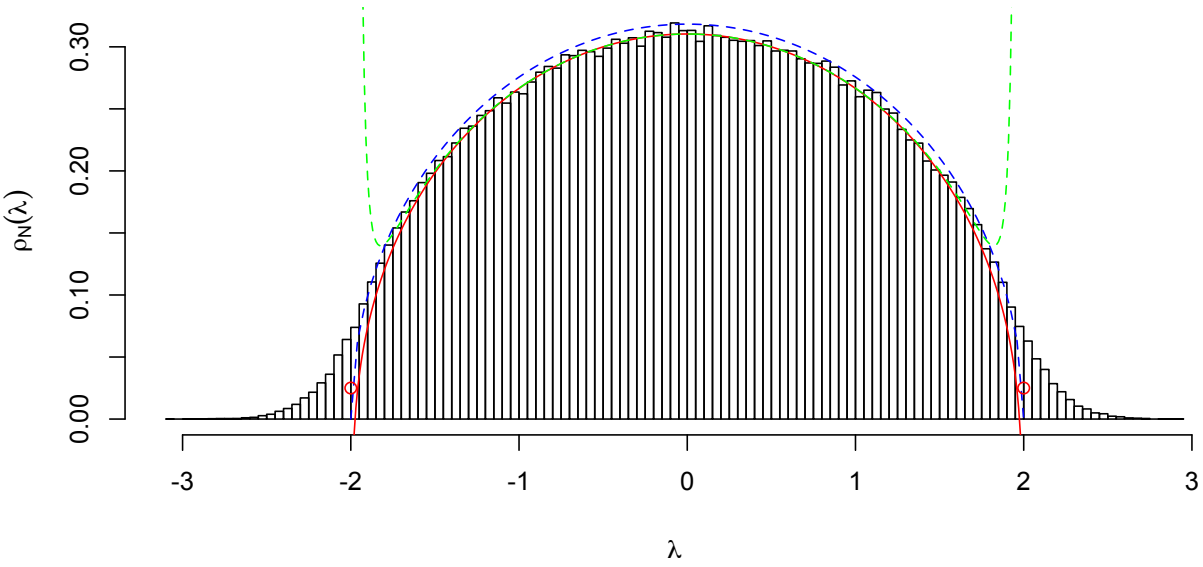


Figure 2.4: Numerical simulations of the state density of a GOE matrix of dimension $N = 10$. Histogram: sample density. Red curve: *Finite size correction* (2.34) until order $1/N$, which coincides almost perfectly with the sample density *inside* the bulk. The Dirac is pictured with a single point. Green dashed curve: *Finite size correction* (2.34) until order $1/N^2$. Blue curve: $N \to \infty$ Wigner semi-circle density.

H_0 (this level spacing is typically of order $1/(N\rho(\lambda))$ where ρ is the density of eigenvalues for a GOE matrix renormalized by $1/\sqrt{N}$). Indeed the small denominators, due to eigenvalues at very near distance, will lead to divergence in the individual overlaps $\langle \psi_i^1 | \psi_j^0 \rangle$ between the perturbed and non perturbed eigenvectors.

The idea to avoid this problem is to study the overlap between a whole subspace of non-perturbed eigenvectors with a whole subspace of perturbed eigenvectors. More precisely, we study the (not necessarily square) overlap matrix with entries

$$G_{ij} = \langle \psi_i^1 | \psi_j^0 \rangle$$

obtained by taking the scalar product between all the non-perturbed eigenvectors ψ_j^0 whose eigenvalues lie in the interval[4] $[a; b] \subset [-2; 2]$ with all the perturbed eigenvectors ψ_i^1 whose associated eigenvalues lie in an interval $[a - \delta; b + \delta]$ (where $\delta > 0$).

We are then able to extract precise informations on the overlap matrix G. In particular, we characterize its spectrum in the limit of large matrices and we extract the main information on the shape of this spectrum. We also define the overlap distance $D(V_0, V_1)$ between the two subspaces V_0 and V_1, which are respectively the subspaces generated by the ψ_j^0 and by the ψ_i^1, as follows

$$D(V_0, V_1) = -\frac{1}{P} \sum_{i=1}^{P} \log(s_i)$$

where $s_1, \ldots, s_P$ are the singular values of G (with P being the smallest dimension of the rectangular matrix G)[5]. We also compute this distance $D(V_0, V_1)$ in the limit of large dimension. This establishes some stability properties of the subspace generated by the eigenvectors of a GOE matrix and finds applications in quantum dissipation and in the study of the physical phenomenon of singular x-ray absorption in metals.

[4] In this paper, the renormalization of the matrices are such that the support of the spectrum when $N = \infty$ is $[-2; 2]$.

[5] In the limit of large matrices the former definition of G through the interval $[a; b]$ implies that $P \sim N \int_a^b \rho$ where ρ is the eigenvalue density.

2.2 Empirical Covariance matrices

In this section, we first review the classical ensembles of random covariance matrices and we recall as in the previous section their main spectral properties. Then, in the final subsection, we present our contribution in this direction. Part of those results are theoretical results but we have also worked on some applications in quantitative finance.

2.2.1 Real and complex Gaussian Wishart Ensembles

Let $\mathbf{X}$ be a real (respectively complex) Gaussian random matrix of size $M \times N$, i.e. a random matrix chosen in the space of $M \times N$ real (resp. complex) matrices according to the law:

$$P(\mathrm{d}\mathbf{X}) \propto \exp\left(-\frac{1}{2}\mathrm{Tr}(\mathbf{X}^\dagger \mathbf{X})\right) \mathrm{d}\mathbf{X}\,, \tag{2.38}$$

where $\mathbf{X}^\dagger$ is the Hermitian conjugate of $\mathbf{X}$. In the following, we will denote the real (resp. complex) Wishart ensemble by $\mathcal{W}^\beta$ with $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case).

The real (resp. complex) Wishart Ensemble is the ensemble of $(N \times N)$ square matrices of the product form $\mathbf{W} := \mathbf{X}^\dagger \mathbf{X}$ where $\mathbf{X}$ is a real (resp. complex) Gaussian random matrix of size $N \times M$.

The spectral properties of the Wishart matrices have been studied extensively and it is known [87] that for $M \geqslant N$, all N positive eigenvalues of $\mathbf{W}$ are distributed via the joint pdf

$$P_\beta(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z} e^{-\frac{1}{2}\sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{\beta}{2}(M-N+1)-1} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \tag{2.39}$$

where Z is a constant normalization factor and where $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case). Note that the distribution P_β defined in (2.39) is in fact defined for every $\beta > 0$.

2.2.2 Continuous processes for real and complex Wishart ensembles

We wish to define here a diffusive matrix process depending on a fictitious time $t \geqslant 0$ that will converge to the Wishart Ensembles in the limit of large time. The idea is simply to set

$$\mathbf{W}_t := \mathbf{X}_t^\dagger \mathbf{X}_t \tag{2.40}$$

where $\mathbf{X}_t$ is a real (resp. complex) random matrix process (of size $M \times N$) following the Ornstein-Uhlenbeck law,

$$\mathrm{d}\mathbf{X}_t = -\frac{1}{2}\mathbf{X}_t \mathrm{d}t + \mathrm{d}\mathbf{B}_t$$

where $\mathbf{B}_t$ is a real Brownian (resp. complex) random matrix, i.e. a matrix whose entries are given by independent standard Brownian motions.

It is well known that the stationary law of a Ornstein-Uhlenbeck process is the Gaussian law and therefore, the real (resp. complex) matrix process $\mathbf{X}_t$ converges in law when $t \to \infty$ to the law of a Gaussian real (resp. complex) random matrix. Hence, we deduce that the real (resp. complex) matrix process $\mathbf{W}_t$ defines a diffusive matrix process that converges in law to $\mathbf{W} \in \mathcal{W}^\beta$.

It is also easy to check that the positive definite matrix process $\mathbf{W}_t$ verifies the following stochastic differential equation [49]:

$$\mathrm{d}\mathbf{W}_t = -\mathbf{W}_t \mathrm{d}t + \sqrt{\mathbf{W}_t}\mathrm{d}\mathbf{B}_t + \mathrm{d}\mathbf{B}_t^\dagger \sqrt{\mathbf{W}_t} + M\beta\,\mathbf{I}\,\mathrm{d}t \tag{2.41}$$

where $\mathbf{B}_t$ is a real (resp. complex) Brownian random matrix and with $\beta = 1$ (resp. $\beta = 2$).

The evolution of the eigenvalue process $\lambda_1(t) \leqslant \lambda_2(t) \ldots \leqslant \lambda_N(t)$ is also easy to derive [50] using perturbation theory to second order as in the second section of this chapter

$$\mathrm{d}\lambda_i = -\lambda_i \mathrm{d}t + 2\sqrt{\lambda_i}\,\mathrm{d}b_i + \beta\left(M + \sum_{k\neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k}\right)\mathrm{d}t \tag{2.42}$$

where the b_i are independent standard Brownian motions. The stationary distribution of the process $(\lambda_1, \dots, \lambda_N)(t)$ is necessarily the joint pdf $P_\beta(\lambda_1, \dots, \lambda_N)$ defined in (2.39) (this is true for any $\beta > 0$ and can also be recovered using the Fokker-Planck equation for the multivariate diffusion (2.42)).

2.2.3 Eigenvalues density

As in the second section, we are interested in the limiting eigenvalues density for large empirical covariance matrices, which we define here as matrices of the product form $\mathbf{E} := \mathbf{X}^\dagger \mathbf{X}/M$ where $\mathbf{X}$ is an $M \times N$ matrix with i.i.d. entries.

Real and complex Gaussian Wishart random matrices are a particular case of what we call an empirical covariance matrix.

Under the assumption that the entries, denoted as x_{ij}, of the random matrix $\mathbf{X}$ have finite second moment (i.e. such that $\mathbb{E}[x_{ij}^2] = \sigma^2 < \infty$), the Marčenko Pastur theorem states that the eigenvalues empirical distribution of the random covariance matrix $\mathbf{E}$ converges weakly almost surely, in the limit $N, M \to \infty$ with $N/M \to q$ (where q is a fixed parameter), to a deterministic probability measure whose density with respect to Lebesgue measure is

$$\rho(\lambda) = \frac{1}{2\pi q} \frac{\sqrt{(\gamma_+ - \lambda)(\lambda - \gamma_-)}}{\lambda}, \quad \gamma_- < \lambda < \gamma_+ \tag{2.43}$$

where $\gamma_\pm$ are the edges of the compactly supported spectrum given by

$$\gamma_- = (1 - \sqrt{q})^2, \quad \gamma_+ = (1 + \sqrt{q})^2 .$$

This result can be derived in the case of real and complex Gaussian Wishart random matrices by using the diffusion process defined in (2.42) along the same line as in the second section (see subsection 2.1.5). It is also done in [12].

We mention in passing that the case of heavy tail entries for the matrix $\mathbf{X}$ is treated in [32], but as in the Wigner case, the limiting spectral density is not the Marčenko Pastur density but is fully characterized in [32]. As in [33], the authors of [32] also derive the main properties of the limiting probability measure (continuous density with respect to Lebesgue measure, tails behavior).

2.2.4 Our contribution

This subsection is devoted to our main contribution in the study of Wishart matrices. In the first paragraph, we construct a diffusive matrix model for the β-Wishart (or Laguerre) ensemble for general $\beta \in [0, 2]$, which preserves invariance under the orthogonal/unitary group. Scaling the Dyson index β with the largest size M of the data matrix as $\beta = 2c/M$ (with c a fixed positive constant), we obtain a family of spectral densities interpolating continuously between the Marčenko-Pastur and the Gamma laws as c is varied. We obtain as a byproduct the correction to the Marčenko-Pastur density until order $1/M$ for all β and until order $1/M^2$ for the particular cases $\beta = 1, 2$. In the second paragraph, we are interested in the empirical covariance matrix composed from the increments of independent multifractal random walks and in particular in the eigenvalue empirical density of this matrix. This study is motivated by applications in risk control and portfolio optimization in finance. In the third and last paragraph, we are interested in empirical covariance matrices composed from datas which have a non trivial "true" covariance matrix with one eigenvalue much larger than the other ones. In particular, we study the evolution of the top eigenvalue and associated eigenvector of the empirical matrix when it is measured through a sliding widow along a time series of datas.

Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marčenko-Pastur law [joint work with Jean-Philippe Bouchaud, Satya N. Majumdar and Pierpaolo Vivo, see also chapter 6 or [12]].

This paragraph is related to the work [6] but concerns the Gaussian Wishart model instead of the Gaussian Orthogonal ensemble. We define a diffusive matrix model for invariant β-Wishart

Ensembles for all β and we find the interpolation family of spectral measures when taking a vanishing repulsion coefficient $\beta = 2c/M$. As a by product, we can also find the $1/N$ and $1/N^2$ correction terms to the Marčenko Pastur density (for the convergence of the empirical spectral distribution).

We first need to introduce a family of real diffusion processes. Let $\delta > 0$ be a fixed parameter. The CIR process (named after its creators John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross) is the diffusion process $x(t)$ defined by $x(0) := x_0 > 0$ and for $t \geqslant 0$ by

$$\mathrm{d}x(t) = -x(t)\ \mathrm{d}t + 2\sqrt{x(t)}\ \mathrm{d}b_t + \delta\ \mathrm{d}t\,. \tag{2.44}$$

Using the assumption $\delta > 0$, it is easy to see that the process $x(t)$ will remain non negative for all times $t \geqslant 0$. It is also easy to verify that the stationary pdf of the Langevin equation (6.11) is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ defined as

$$p_\delta(x) = \frac{1}{2^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})} x^{\frac{\delta}{2}-1} e^{-\frac{x}{2}}\,. \tag{2.45}$$

In analogy with squared Bessel processes, the parameter δ will be called the *dimension* of the process $x(t)$.

Following [6, 10], our goal is to construct a diffusive matrix process whose eigenvalues process is asymptotically distributed according to P_β for general $\beta \in [0; 2]$.

The idea is to slice the time interval into small intervals of length $1/n$ and for each interval $[k/n; (k+1)/n]$, to choose independently Bernoulli random variables $\epsilon_k^n, k \in \mathbb{N}$ such that $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$. Then, setting $\epsilon_t^n = \epsilon_{[nt]}^n$, our diffusive matrix process evolves as:

$$\mathrm{d}\mathbf{W}_t^n = -\mathbf{W}_t^n\,\mathrm{d}t + \mathrm{d}\mathbf{\Delta}_t^n \tag{2.46}$$

where the increment matrix $\mathrm{d}\mathbf{\Delta}_t^n$ now depends on the value of the additional random process ϵ_t^n:

- if $\epsilon_t^n = 1$, then
$$\mathrm{d}\mathbf{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\,\mathrm{d}\mathbf{B}_t + \mathrm{d}\mathbf{B}_t^\dagger\,\sqrt{\mathbf{W}_t^n} + M\ \mathbf{I}\,\mathrm{d}t.$$
where $\mathrm{d}\mathbf{B}_t$ is an $N \times N$ real[6] Brownian increment matrix whose entries have variance $\mathrm{d}t$.
- if $\epsilon_t^n = 0$, then
$$\mathrm{d}\mathbf{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\,\mathrm{d}\mathbf{Y}_t + \mathrm{d}\mathbf{Y}_t^\dagger\,\sqrt{\mathbf{W}_t^n} + \delta\ \mathbf{I}\,\mathrm{d}t.$$
with $\delta > 0$ and where $\mathrm{d}\mathbf{Y}_t$ is a symmetric matrix that is co-diagonalizable with $\mathbf{W}_t^n$ (i.e. the two matrix have the same eigenvectors) but with a spectrum given by N independent real Brownian increments of variance $\mathrm{d}t$.

It is clear that the eigenvalues of the matrix $\mathbf{W}_t^n$ will cross at some points but only in intervals $[k/n; (k+1)/n]$ for which $\epsilon_k^n = 0$ (in the other intervals where they follow the SDE (2.42) with parameter $\beta = 1$, it is well known that the repulsion is too strong and thus collisions are avoided). In this case, the eigenvalues are re-numbered at time $t = (k+1)/n$ in increasing order. With this procedure, when ordered $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t)$, we can again check as in [6, 10], using perturbation theory, that the eigenvalues will remain always non-negative and will verify the Stochastic Differential System (SDS):

$$\mathrm{d}\lambda_i^n = -\lambda_i^n \mathrm{d}t + 2\sqrt{\lambda_i^n}\,\mathrm{d}b_i + \left(\epsilon_t^n M + (1-\epsilon_t^n)\delta + \epsilon_t^n \sum_{k\neq i} \frac{\lambda_i^n + \lambda_k^n}{\lambda_i^n - \lambda_k^n}\right)\mathrm{d}t \tag{2.47}$$

where the b_i are independent standard Brownian motions, which are also independent of the process ϵ_t^n.

Note that when $\epsilon_t^n = 0$, the particles λ_i^n are evolving as independent CIR processes of dimension $\delta > 0$ as defined above. Therefore, the particles can cross in those time intervals,

[6] Obviously, this construction can be done with complex Brownian motions, corresponding to the parameter $\beta = 2$ instead of $\beta = 1$ in the present case.

breaking the increasing order so that they will be re-ordered at time $([nt]+1)/n$ but they *will* remain non-negative as the dimension δ is strictly positive. Therefore the SDS (6.14) remains well defined at all times $t \geqslant 0$.

One can follow the proof of [10] to prove that the scaling limit (i.e. when $n \to \infty$) of the process $(\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t))$ verifies the following SDS

$$\mathrm{d}\lambda_i = -\lambda_i \mathrm{d}t + 2\sqrt{\lambda_i}\,\mathrm{d}b_i + \left(pM + (1-p)\delta + p\sum_{k\neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) \mathrm{d}t\,. \tag{2.48}$$

One can deduce from the above equation (6.15) the Fokker-Planck equation for the joint density $P(\{\lambda_i\}, t)$, for which the stationary joint pdf is readily found to be

$$P^*(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z} e^{-\frac{1}{2}\sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{p}{2}(M-N+1-\delta)-(1-\frac{\delta}{2})} \prod_{i<j} |\lambda_i - \lambda_j|^p\,. \tag{2.49}$$

The probability P_β introduced in (2.39) is recovered here by taking the values $p = \beta$ and $\delta = 0$. If $p = \beta > 0$, the eigenvalue probability density in the large N, M limit is the Marčenko-Pastur law. Note that with the above normalizations, the spectrum is spread over a region of $\mathbb{R}_+$ of width of order pM. On the other hand, if $p = 0$, the large N, M-limit of the spectral density is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ (recall that it is the stationary pdf of the CIR process of dimension δ):

$$\rho_0(\mathrm{d}\lambda) = \frac{1}{2^{\delta/2}\Gamma(\frac{\delta}{2})} \lambda^{\frac{\delta}{2}-1} e^{-\frac{\lambda}{2}}\,\mathrm{d}\lambda\,. \tag{2.50}$$

It is quite natural to ask whether a crossover regime may be found, interpolating between the Marčenko-Pastur density ($p > 0$, independent of the dimension M) and the Gamma distribution ($p = 0$). A good candidate for triggering such a transition is clearly a parameter p vanishing with M as $p = 2c/M$ where c is a positive fixed constant.

The derivation of the crossover density can be made with two different (and perhaps surprisingly equivalent) methods: the Itô method using Stochastic calculus as was first done in [6] and the Saddle point route (which was also presented in [6]). We explain the link between those two methods in [12].

The interpolating family of probability densities that we finally find is the three parameters c (such that $p = 2c/M$), $q = N/M$ and $\delta \geqslant 0$ as

$$\rho_{c,q,\delta}(\lambda) = \frac{1}{2\Gamma(\mu+\zeta+\frac{1}{2})\Gamma(\zeta-\mu+\frac{3}{2})} \frac{1}{|W_{-\zeta,\mu}(-\frac{\lambda}{2})|^2} \tag{2.51}$$

where $W_{-\zeta,\mu}$ is a Whittaker function and with the following values for the parameters

$$\zeta = cq - \frac{\alpha}{4}; \quad \mu = \frac{1}{4}|\alpha - 2| \quad \text{with} \quad \alpha = (2-\delta) - 2c(1-q); \quad .$$

For $c = 0$, the probability density $\rho_{0,q,\delta}$ is indeed given by the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$. For $c \to +\infty$, one can also check that the distribution $\rho_{c=+\infty,q,\delta}$ indeed corresponds to the Marčenko Pastur distribution with parameter q (the parameter δ is irrelevant in this regime). We also checked expression (2.51) numerically, with very good agreement, see Fig. 2.5.

We mention also that following a method similar to the one explained in paragraph **Wigner correction for large but finite dimension**, we can derive the correction term to the Marčenko Pastur distribution for the empirical eigenvalue distribution.

Marčenko Pastur theorem for MRW processes [joint work with Rémi Rhodes and Vincent Vargas, see also chapter 8 or [11]].

This work is inspired from applications. It stems from finance and focuses on the study of covariance matrices which is a crucial tool for minimizing the risk $\mathcal{R}_w$ of a portfolio w that

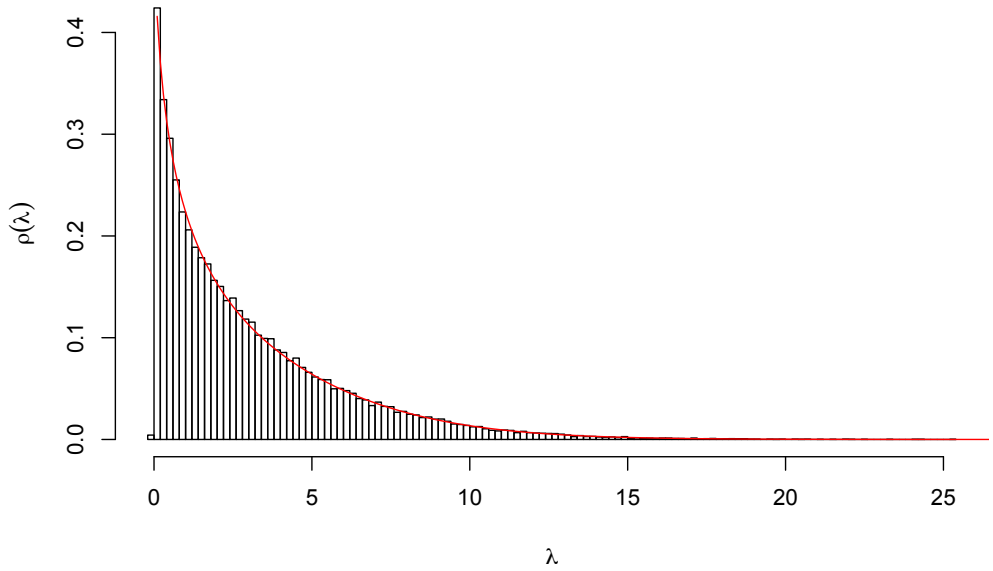


Figure 2.5: The histogram represents the empirical eigenvalue distribution of the matrix $\mathbf{W}_{t=\infty}^{n}$ for the following values of the parameters $c = 1, p = \beta = 2c/M, M = 100, N = 50, \delta = 1, q = 1/2$. The (red) curve is our theoretical prediction for the limiting eigenvalue density given by (2.51).

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 $\mathbf{R}_{ij} = \mathbb{E}[r_i r_j]$):

$$\mathcal{R}_w = w^t \mathbf{R} w \,.$$

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 [44, 124] for more extended discussions on the applications of large empirical covariance matrices in finance and in particular in portfolio theory.

We consider here N stock price processes $X_i(t)$ for $i = 1, \ldots, N$ that evolve continuously with respect to time $t \in [0; 1]$ and 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 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 $X_N(ij) = r_i(j)$ that enables to define the empirical covariance matrix R_N as follows

$$R_N := X_N X_N^t \,.$$

The Marčenko Pastur theorem enables to compute the limiting eigenvalue density of the empirical covariance matrix R_N if the stock price processes $X_i(t)$ are evolving as independent standard Brownian motions. Indeed, in this case, the random variables $r_i(j)$ are i.i.d. Gaussian variables and the matrix R_N is precisely an empirical covariance matrix as defined in subsection 2.2.3 (more precisely, R_N is a real Gaussian Wishart matrix).

In [11], we have been interested in the case where the stock price processes $X_i(t)$ are independent lognormal multifractal random walk (LMRW), as defined in the first section in equation (1.5), which are typically not diffusions and which present many very interesting properties for

finance as they respect several of the universal features, called stylized facts, observed for the price of assets on financial markets (see [60] for a review on stylized facts). We have been able to prove convergence of the eigenvalue density and to characterize the limiting distribution (which is a deformation of the Marčenko Pastur distribution due to the (long) memory volatility process of the LMRW process) through its Stieltjes transform. We have verified numerically our result: the agreement between the simulated eigenvalue density and the numerical value of the theoretical density obtained by inverting our equations on the Stieltjes transform is excellent. We also give some numerical properties of the limiting eigenvalue density in view of applications.

Empirical measurement of a covariance matrix with one isolated top eigenvalue [joint work with Jean-Philippe Bouchaud, see also chapter 7 or [7]].

We are concerned here with the empirical measurement of a covariance matrix denoted as C which has a very large dimension N and a spectrum of the form $(\lambda_1, 1, \ldots, 1)$ where $\lambda_1 \gg 1$ and we are given a sample of i.i.d. centered Gaussian vectors $(r_1(t), \ldots, r_N(t)), t \in \mathbb{N}$ whose covariance matrix is C. There are two ways of measuring the empirical covariance matrix E. The first one is to pick a large integer T (typically such that $N/T = q \in [0:1]$ where q is a fixed parameter) and to compute E as before through the classical formula,

$$E_{ij} = \frac{1}{T} \sum_{i=1}^{T} r_i(t) r_j(t) \tag{2.52}$$

for each entries i, j of the matrix E.

The other way of measuring the empirical covariance matrix is through the exponential moving average estimator. Letting $\epsilon > 0$ (this parameter will play the role of T), the matrix E is measured through

$$E_{ij} = \epsilon \sum_{t=0}^{\infty} (1-\epsilon)^t r_i(t) r_j(t)\,. \tag{2.53}$$

The two different formulas (2.52) and (2.53) lead to two different matrices with different properties (although related). We can also use a sliding window to have sequences of matrices $E(t)$ which evolve with time as, if we consider for example the second model defined in (2.53),

$$E_{ij}(t) = (1-\epsilon) E_{ij}(t-1) + \epsilon r_i(t) r_j(t)\,. \tag{2.54}$$

We always consider the evolution equation (2.54) in its stationary regime.

In [7], we are interested in the dynamic evolution of the top eigenvalue and eigenvector of the matrix $E(t)$ which evolves with time through (2.54). In particular, we give the Langevin equations followed by the top eigenvalue and the angle θ_t between the top eigenvector of $E(t)$ with the top eigenvector of the matrix C.The dynamics of the angle θ_t defines an interesting new class of random processes. We also compare our results with numerical simulations with very good agreement.

We mention in passing that the study of the top eigenvalue of the empirical covariance matrix E defined in (2.52) can be found in [29] in a more general setting where an arbitrary finite number of spikes (isolated eigenvalues different from 1) are authorized. The authors prove convergence of the top eigenvalue of the matrix E and they characterize the fluctuations for large but finite N. In particular, they exhibit a phase transition depending on the value of the top eigenvalue of the covariance matrix C. If this top eigenvalue of C is above their explicit threshold, the fluctuation of the empirical top eigenvalue is given by a (generalized) Gaussian distribution with the classical scaling in $\sqrt{N}$ although it is, as expected given by a Tracy Widom law if the top eigenvalue of C lies below the threshold with the usual scaling $N^{2/3}$.

In [7], let us mention that we are able to recover the value of the limit of the top eigenvalue of E when $N, T \to \infty$ with $N/T \to q$ but not the fluctuations.

Applications to empirical finance [Work in collaboration with Jean-Philippe Bouchaud, see Chapters 7, 9 and 10]. We have also been interested in applications of random matrix theory to empirical and statistical finance. We would not go too much into the details in this paragraph

and we refer the reader to the corresponding chapters. A part of Chapter 7 is devoted to the study of the evolution in time of the sectors (or top eigenvectors of the correlation matrix) in financial markets. This study uses the results we establish on eigenvectors stability in the same chapter. In Chapter 9, we revisit the Leverage effect in financial market. The Leverage effect is among the best known stylized facts of financial markets: negative price returns induce increased future volatilities. Using random matrix theory, we study this effect and its mechanisms in great details. Random matrix theory enables us in particular to compute the noise measurement and to check the significancy of our empirical results. In Chapter 10, we study the intra-day dynamics of stocks returns and we draw the picture of the co-movement of stocks. Again, this study relies partly on random matrix theory.

Part II

Gaussian multiplicative chaos and their scale invariance properties

Chapter 3

Lognormal $\star$-scale invariant random measures

Résumé

Cet article est publié dans le journal *Probability theory and Related Fields* et est écrit en collaboration avec Rémi Rhodes et Vincent Vargas. Nous considérons l'analogue continu de la célébre équation étoile de Mandelbrot (avec des poids lognormals). Mandelbrot a introduit cette équation afin de caractériser la loi des cascades multiplicatives. Nous montrons l'existence et l'unicité des mesures qui satisfont cette équation continue; ces mesures aléatoires appartiennent à l'ensemble des chaos multiplicatifs Gaussiens introduit par J.-P. Kahane en 1985 (voir à des extensions de cet ensemble). Nous obtenons aussi une caractérisation explicite de la structure de covariance de ces mesures. Nous prouvons de plus que certaines propriétés qualitatives telles que l'indépendance à longue portée ou l'isotropie peuvent être déduites de cette équation.

Abstract

In this article, we consider the continuous analog of the celebrated Mandelbrot star equation with lognormal weights. Mandelbrot introduced this equation to characterize the law of multiplicative cascades. We show existence and uniqueness of measures satisfying the aforementioned continuous equation; these measures fall under the scope of the Gaussian multiplicative chaos theory developed by J.P. Kahane in 1985 (or possibly extensions of this theory). As a by product, we also obtain an explicit characterization of the covariance structure of these measures. We also prove that qualitative properties such as long-range independence or isotropy can be read off the equation.

3.1 Introduction

Fractality and the related concept of scale invariance is nowadays well introduced in many fields of applications ranging from physics, finance, information or social sciences. The scale-invariance property of a stochastic process is usually quantified by the scaling exponents $\xi(q)$ associated with the power-law behavior of the order q moments of the fluctuations at different scales. More precisely, if X_t is a 1-d process with stationary increments, we can consider the q-th moments of its fluctuations at scale l:

$$\mathbb{E}\big[|X_{t+l} - X_t|^q\big].$$

The scaling exponents $\xi(q)$ are defined through the following power-law scaling:

$$\mathbb{E}\big[|X_{t+l} - X_t|^q\big] = C_q l^{\xi(q)} \quad \forall l < T.$$

When $\xi(q) = qH$ is linear, the process is said to be monofractal. Famous examples of such processes are (fractional) Brownian motion, α-stable Lévy processes or Hermitte processes.

When ξ is nonlinear, the process is said to be multifractal. The concept of nonlinear power-law scalings goes back to the Kolmogorov theory of fully developed turbulence in the sixties (see [52, 138, 143, 53, 76] and references therein), introduced to render the intermittency effects in turbulence. Mandelbrot [110] came up with the first mathematical discrete approach of multifractality, the now celebrated multiplicative cascades. Roughly speaking, a (dyadic) multiplicative cascade is a positive random measure M on the unit interval $[0,1]$ that obeys the following decomposition rule:

$$M(dt) \overset{law}{=} Z^0 \mathbf{1}_{[0,\frac{1}{2}]}(t) M^0(2dt) + Z^1 \mathbf{1}_{[\frac{1}{2},1]}(t) M^1(2dt-1), \tag{3.1}$$

where M^0, M^1 are two independent copies of M and (Z^0, Z^1) is a random vector with prescribed law and positive components of mean 1 independent from M^0, M^1. Such an equation (and its generalizations to b-adic trees for $b \geqslant 2$), the celebrated star equation introduced by Mandelbrot in [109], uniquely determines the law of the multiplicative cascade. Despite the fact that multiplicative cascades have been widely used as reference models in many applications, they possess many drawbacks related to their discrete scale invariance, mainly they involve a particular scale ratio and they do not possess stationary fluctuations (this comes from the fact that they are constructed on a dyadic tree structure).

Much effort has been made to develop a continuous parameter theory of suitable stationary multifractal random measures ever since, stemming from the theory of multiplicative chaos introduced by Kahane [94, 26, 138, 24, 129, 132]. The construction of such measures is now well understood and they are largely used in mathematical modeling since they obey a so-called stochastic scale invariance property, namely the property of being equal in law at different scales up to an independent stochastic factor. This is some kind of continuous parameter generalization of (3.1). Stochastic scale invariance property is observed in many experimental and theoretical problems, like turbulence (see [76, 52] and many others), quantum gravity (see [97, 66, 131]), mathematical finance, etc... and this is the main motivation for introducing multifractal random measures. However, as far as we know, the following question has never been solved: are these measures the only existing stochastic scale invariant object? This is fundamental since a positive answer gives a full justification to their intensive use. In this paper, we characterize stochastic scale invariant measures when the stochastic factor is assumed to be log-normal. We prove that the class of such objects is made up of Gaussian multiplicative chaos with a specific structure of the covariation kernel, which turns out to be larger than described in the literature.

3.2 Background

Let us first remind the reader of the main definitions we will use throughout the paper. We denote by $\mathcal{B}(E)$ the Borelian sigma field on a topological space E. A random measure M is a random variable taking values into the set of positive Radon measures defined on $\mathcal{B}(\mathbb{R}^d)$ such that $\mathbb{E}[M(K)] < +\infty$ for every compact set K. A random measure M is said to be stationary if for all $y \in \mathbb{R}^d$ the random measures $M(\cdot)$ and $M(y+\cdot)$ have the same law.

3.2.1 Gaussian multiplicative chaos

We remind the reader of the notion of Gaussian multiplicative chaos as introduced by Kahane [94]. Consider a sequence $(X^n)_n$ of independent centered stationary Gaussian processes with associated nonnegative covariance kernel $k_n(r) = \mathbb{E}[X^n_r X^n_0] \geqslant 0$. For each $N \geqslant 1$, we can define a Radon measure M^N on the Borelian subsets of $\mathbb{R}^d$ by

$$M^N(A) = \int_A e^{\sum_{n=0}^N X^n_r - \frac{1}{2}\mathbb{E}[(X^n_r)^2]}\, dr.$$

For each Borelian set A, the sequence $(M^N(A))_N$ is a positive martingale. Thus it converges almost surely towards a random variable denoted by $M(A)$. One can deduce that the sequence of measures $(M^N)_N$ weakly converges towards a Radon measure M, commonly denoted by

$$M(A) = \int_A e^{X_r - \frac{1}{2}\mathbb{E}[X_r^2]}\, dr \tag{3.2}$$

and called Gaussian multiplicative chaos associated to the kernel

$$K(r) = \sum_{n=0}^{+\infty} k_n(r). \tag{3.3}$$

Roughly speaking, (3.2) can be understood as a measure admitting as density the exponential of a Gaussian process X with covariance kernel K. Of course, this is purely formal because X can only be understood as a (random Gaussian) distribution in the sense of Schwartz because of the possible singularities of the kernel K.

Of special interest is the situation when the function K can be rewritten as (for some $\lambda^2 > 0$)

$$K(r) = \lambda^2 \ln_+ \frac{T}{|r|} + g(r) \tag{3.4}$$

for some bounded function g (and $\ln_+(x) = \max(0, \ln(x))$). In that case, Kahane proved that the martingale $(M^N(A))_N$, for some Borelian set A with non-null finite Lebesgue measure, is uniformly integrable if and only if $\lambda^2 < 2d$. This condition is necessary and sufficient in order for the limiting measure M to be non identically null. For kernels of the form (3.4) which can not be written as a sum of nonnegative terms as (3.3), we refer to the extended Gaussian multiplicative theory developed in [129]. We remind that Gaussian multiplicative chaos with kernel given by (3.4) has found applications in many fields in science:

- In dimension 1, the measure M is called the lognormal Multifractal Random Measure (MRM). It is used to model the volatility of a financial asset (see [25], [65]).
- In dimension 2, the measure M is a probabilistic formulation of the quantum gravity measure (more precisely, the quantum gravity measure is defined as the exponential of the Gaussian Free Field and therefore is defined in a bounded domain). We refer to references [34], [66], [131] for probabilistic papers on this topic.
- In dimension 3, the measure M is called the Kolmogorov-Obhukov model (see textbook [76]): it is a model of energy dissipation in the statistical theory of fully developed turbulence.

3.3 Main results

3.3.1 Definitions

In this paper we are interested in stationary random measures satisfying the following scale invariance property:

Definition 3.1. Log-normal $\star$-scale invariance. *A random measure M is said to be lognormal $\star$-scale invariant if for all $\epsilon < 1$, M obeys the cascading rule*

$$\big(M(A)\big)_{A\in\mathcal{B}(\mathbb{R}^d)} \overset{law}{=} \big(\int_A e^{\omega_\epsilon(r)} M^\epsilon(dr)\big)_{A\in\mathcal{B}(\mathbb{R}^d)} \tag{3.5}$$

where ω_ϵ is a stationary Gaussian process with continuous sample paths and M^ϵ is a random measure independent from ω_ϵ satisfying the relation

$$\big(M^\epsilon(\epsilon A)\big)_{A\in\mathcal{B}(\mathbb{R}^d)} \overset{law}{=} \epsilon^d \big(M(A)\big)_{A\in\mathcal{B}(\mathbb{R}^d)}. \tag{3.6}$$

□

Intuitively, this relation means that when you zoom in the measure M, you should observe the same behavior up to an independent log-normal factor. This relation is the continuous parameter analog of the celebrated Mandelbrot star equation.

Remark. In order for a measure M satisfying (3.5) with a moment of order 1 to be non trivial, it is obvious to check that the Gaussian process ω_ϵ must be normalized so that $\mathbb{E}[e^{\omega_\epsilon}(r)] = 1$.

Definition 3.2. *We will say that a stationary random measure M satisfies the good lognormal $\star$-scale invariance if M is lognormal $\star$-scale invariant and for each $\epsilon < 1$, the covariance kernel k_ϵ of the process ω_ϵ involved in (3.5) is continuous and satisfies:*

$$|k_\epsilon(r)| \to 0 \quad as \quad |r| \to +\infty, \tag{3.7}$$

$$\forall r, r' \in \mathbb{R}^d \setminus \{0\}, \qquad |k_\epsilon(r) - k_\epsilon(r')| \leqslant C_\epsilon \theta\big(\min(|r|, |r'|)\big)|r - r'| \tag{3.8}$$

for some positive constant C_ϵ and some decreasing function $\theta :]0, +\infty[\to \mathbb{R}_+$ such that

$$\int_1^{+\infty} \theta(u) \ln(u)\, du < +\infty. \tag{3.9}$$

□

Though we would like to solve (3.5) in great generality, we must make a few technical assumptions to avoid pathological situations (a pathological example is given at the very end of Section 3.4). This is basically the purpose of the above definition 3.2. Let us make a few comments on its content.

Equation (3.8) mainly expresses that the kernel k_ϵ is Lipschitzian with a local Lipschitz constant that decays at most like θ when approaching infinity. By combining (3.7) and (3.8), it is plain to see that

$$\forall r \neq 0, \qquad |k_\epsilon(r)| \leqslant C_\epsilon \int_{|r|}^{+\infty} \theta(u)\, du. \tag{3.10}$$

This is a very weak decorrelation property for the process ω_ϵ, which describes how fast the covariance function decays at infinity. In our proofs, it will be the key tool to investigate the mixing properties of the measure M.

3.3.2 Results

In what follows, we are mainly interested in the one-dimensional case $d = 1$. We have the following description of the solutions to (3.5), which is the main result of the paper:

Theorem 3.3. *Let M be a good lognormal $\star$-scale invariant random measure. Assume that*

$$\mathbb{E}[M([0,1])^{1+\delta}] < +\infty$$

for some $\delta > 0$. Then M is the product of a nonnegative random variable $Y \in L^{1+\delta}$ and an independent Gaussian multiplicative chaos

$$\forall A \subset \mathcal{B}(\mathbb{R}), \quad M(A) = Y \int_A e^{X_r - \frac{1}{2}\mathbb{E}[X_r^2]}\, dr \tag{3.11}$$

with associated covariance kernel given by the improper integral

$$K(r) = \int_{|r|}^{+\infty} \frac{k(u)}{u}\, du \tag{3.12}$$

for some continuous covariance function k such that $k(0) \leqslant \frac{2}{1+\delta}$.

Conversely, given some datas k and Y as above, the relation (3.11) defines a log-normal $\star$-scale invariant random measure M with finite moments of order $1 + \gamma$ for every $\gamma \in [0, \delta)$.

Let us also state the following result giving a sufficient (and not far from being necessary) condition in terms of k for the measure M as constructed in Theorem 3.3 to be good:

Proposition 3.4. *Let M be a log-normal $\star$-scale invariant random measure as constructed in Theorem 3.3. If*

$$\int_1^{+\infty} \ln r \sup_{|u| \geqslant r} \frac{|k(u)|}{u}\, dr < +\infty \tag{3.13}$$

then M is a good lognormal $\star$-scale invariant random measure.

Let us comment on Theorem 3.3. First we point out that Y is deterministic as soon as the random measure M is ergodic. Second, good lognormal $\star$-scale invariant measures exhibit a multifractal behaviour. More precisely, if we consider a measure M as in Theorem 3.3, we define its structure exponent

$$\forall q > 0, \quad \xi(q) = (1 + \frac{k(0)}{2})q - \frac{k(0)}{2}q^2.$$

Then we have the following asymptotic power-law spectrum, for $q < 1 + \delta$:

$$\mathbb{E}\big[M([0,t])^q\big] \simeq C_q t^{\xi(q)} \quad \text{as } t \to 0,$$

for some positive constant C_q.

We also stress that the intermittency parameter $k(0)$ is explicit when one knows K because of the relation

$$K(r) \sim k(0) \ln\left(\frac{1}{r}\right), \quad \text{when } r \to 0. \tag{3.14}$$

The covariance function K can also be recovered from the two sets marginals of the measure M thanks to formula (3.48).

Finally, Theorem 3.3 has the following consequence about the regularity of good lognormal $\star$-scale invariant measures:

Corollary 3.5. *Almost surely, a good log-normal $\star$-scale invariant random measure M does not possess any atom on $\mathbb{R}$, that is:*

$$\text{almost surely}, \forall x \in \mathbb{R}, \quad M(\{x\}) = 0.$$

Now we investigate long-range independence for good lognormal $\star$-scale invariant random measures. So we introduce the related notion of cut-off:

Definition 3.6. *We will say that a stationary random measure M admits a cut-off $d > 0$ if, for $t < s$, the σ-algebras $\mathcal{H}^t_{-\infty} = \sigma\{M(A); A \in \mathbb{B}(\mathbb{R}), A \subset (-\infty, t]\}$ and $\mathcal{H}^{+\infty}_s = \sigma\{M(A); A \in \mathbb{B}(\mathbb{R}), A \subset [s, +\infty)\}$ are independent, conditionally to the asymptotic σ-algebra of M, as soon as $s - t > d$.* □

Of course, if the measure M is ergodic then the asymptotic σ-algebra of M is trivial and we can remove the sentence "conditionally to the asymptotic σ-algebra of M" from the definition. For instance the measure constructed in subsection 3.3.4 admits a cut-off T and is ergodic. It results from the proof of Theorem 3.3 that the cut-off property can be read off the cascading rule (3.5):

Proposition 3.7. *Let M be a good lognormal $\star$-scale invariant random measure with finite $1+\delta$ moment. Then M admits a cutoff if and only if, for some $\epsilon < 1$ (or equivalently for all $\epsilon < 1$), the covariance kernel k_ϵ of the process ω_ϵ in (3.5) reduces to 0 outside a compact set.*

Finally, we mention that another notion of stochastic scale invariance has been studied in the literature before: it is called the exact stochastic scale invariance (see [24, 52, 132]). Let us recall the main result: if the Gaussian multiplicative chaos M admits a covariance kernel K such that $K(x) = \lambda^2 \ln \frac{T}{|x|} + C$ for some constant C and for all x in a ball $B(0, R)$ then M satisfies the "exact stochastic scale invariance":

$$\forall \alpha \in (0,1), \quad (M(\alpha A))_{A \subset B(0,R)} \overset{law}{=} \alpha e^{Y_\alpha - \frac{1}{2}\mathbb{E}[Y_\alpha^2]} (M(A))_{A \subset B(0,R)}$$

where Y_α is a centered Gaussian random variable with variance $\lambda^2 \ln \frac{1}{\alpha}$.

The reader may wonder if we can construct random measures that are both exactly stochastically scale invariant and good lognormal $\star$-scale invariant. Let us show that

Proposition 3.8. *Let M be a Gaussian multiplicative chaos whose covariance kernel K is such that, for $|r| \leqslant R$, $K(r) = \lambda^2 \ln \frac{T}{|r|} + C$ for some constant C (in particular, M satisfies the "exact stochastic scale invariance"), then M is not a good lognormal $\star$-scale invariant random measure.*

3.3.3 Multidimensional results

We stress that our results remain true in higher dimensions without changes in the proofs. For the sake of completeness, we state the main result.

Theorem 3.9. *Let M be a good lognormal $\star$-scale invariant random measure such that for each $\epsilon < 1$, the covariance kernel k_ϵ of the process ω_ϵ is continuous and differentiable on $\mathbb{R}^d \setminus \{0\}$. Assume that*

$$\mathbb{E}[M([0,1]^d)^{1+\delta}] < +\infty$$

for some $\delta > 0$. Then M is the product of a nonnegative random variable $Y \in L^{1+\delta}$ and an independent Gaussian multiplicative chaos:

$$\forall A \subset \mathcal{B}(\mathbb{R}^d), \quad M(A) = Y \int_A e^{X_r - \frac{1}{2}\mathbb{E}[X_r^2]}\, dr \tag{3.15}$$

with associated covariance kernel given by the improper integral

$$\forall x \in \mathbb{R}^d \setminus \{0\}, \quad K(x) = \int_1^{+\infty} \frac{k(xu)}{u}\, du \tag{3.16}$$

for some continuous covariance function k such that $k(0) \leqslant \frac{2d}{1+\delta}$.

Conversely, given some datas k and Y as above, the relation (3.11) defines a lognormal $\star$-scale invariant random measure M with finite moments of order $1+\gamma$ for every $\gamma \in [0, \delta)$.

It turns out that Proposition 3.4 remains true in dimension $d \geqslant 1$. When the dimension is greater than 1, it may be interesting to focus on the isotropy properties. In the same spirit as Proposition 3.7, for a good lognormal $\star$-scale invariant measure M with a finite moment of order $1+\delta$, the following assertions are equivalent:

1. M is isotropic,
2. its covariance kernel K (or equivalently k in (3.16)) is isotropic,
3. the covariance kernel k_ϵ is isotropic for some $\epsilon < 1$,
4. the covariance kernels k_ϵ are isotropic for all $\epsilon < 1$.

3.3.4 Classical example

As far as we know, there exists only one example of good log-normal $\star$-scale invariant random measures in the literature, which was first described in [26] (see also [24]). Its construction is very intuitive: it is geometric and relies on homothetic properties of triangles in the half-plane. We also stress that this specific example of $\star$-scale invariant random measures is not restricted to the Gaussian case: the factor can be more general (log-Lévy).

Following [24], we recall the construction of this example and refer the reader to the aforementioned papers for further details. Fix $T > 0$ and let $\mathcal{S}^+$ be the state-space half plane

$$\mathcal{S}^+ = \{(t,l) : t \in \mathbb{R}, l > 0\}.$$

with which one can associate the measure

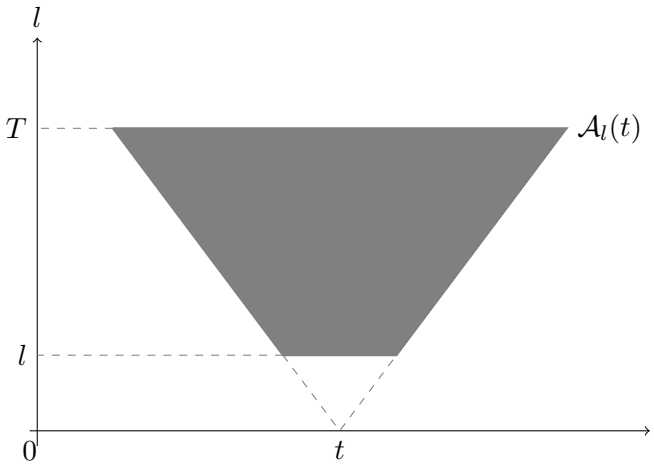
$$\mu(dt, dl) = l^{-2} dt dl.$$

Then we introduce the independently scattered Gaussian random measure P defined for any μ-measurable set A by

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

with $\varphi(q) = -\lambda^2 q^2/2 - iq\lambda^2/2$. Under those assumptions, we can note that for any μ-measurable set A, $P(A)$ is a Gaussian variable with mean $m = -\mu(A)\lambda^2/2$ and variance $\sigma^2 = \lambda^2\mu(A)$. We can then define the Gaussian process $(\omega_l(t))_{t\in\mathbb{R}}$ for $l \geqslant 0$ by

$$\omega_l(t) = P\left(\mathcal{A}_l(t)\right)$$



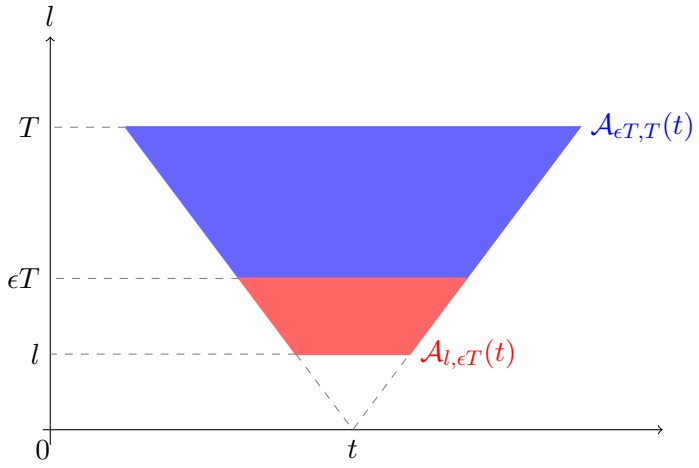
where $\mathcal{A}_l(t)$ is the triangle like subset $\mathcal{A}_l(t) := \{(t', l') : l \leqslant l' \leqslant T, -l'/2 \leqslant t - t' \leqslant l'/2\}$.

Define now the random measure M_l by $M_l(dt) = e^{\omega_l(t)}dt$. Almost surely, the family of measures $(M_l(dt))_{l>0}$ weakly converges towards a random measure M. If $\lambda^2 < 2$, this measure is not trivial.

Let us check that M is a good log-normal $\star$-scale invariant random measure. Fix $\epsilon < 1$ and define the sets $\mathcal{A}_{l,\epsilon T}(t) := \{(t', l') : l \leqslant l' \leqslant \epsilon T, -l'/2 \leqslant t - t' \leqslant l'/2\}$ and $\mathcal{A}_{\epsilon T,T}(t) := \{(t', l') : \epsilon T \leqslant l' \leqslant T, -l'/2 \leqslant t - t' \leqslant l'/2\}$. Note that $\mathcal{A}_l(t) = \mathcal{A}_{l,\epsilon T}(t) \cup \mathcal{A}_{\epsilon T,T}(t)$ and that those two sets are disjoint. Thus, we can write for every μ-measurable set A

$$M_l(A) = \int_A e^{\omega_{\epsilon T,T}(t)} e^{\omega_{l,\epsilon T}(t)} dt \tag{3.17}$$

with $\omega_{\epsilon T,T}(t) = P(\mathcal{A}_{\epsilon T,T}(t))$ and $\omega_{l,\epsilon T}(t) = P(\mathcal{A}_{l,\epsilon T}(t))$.



We then study equation (3.17) in the limit $l \to 0$; we obtain

$$M(A) = \int_A e^{\omega_{\epsilon T,T}(t)} M^\epsilon(dt) \tag{3.18}$$

where M^ϵ is the limit when $l \to 0$ of the random measure $M_l^\epsilon(dt) := e^{\omega_{l,\epsilon T}(t)}dt$. We easily verify that $M^\epsilon(\epsilon A) \overset{law}{=} \epsilon M(A)$ writing

$$M_l^\epsilon(A) = \epsilon \int_A e^{\omega_{l,\epsilon T}(\epsilon t)} dt \tag{3.19}$$

and checking that the covariance of the Gaussian process $(\omega_{l,\epsilon T}(\epsilon t))_{t\in\mathbb{R}}$ is the same as the one of $(\omega_{l,T}(t))_{t\in\mathbb{R}}$.

The covariance kernel of the stationary Gaussian process $\omega_{\epsilon T,T}(t)$ is given by

$$k_\epsilon(r) = \begin{cases} 0 & \text{if} \quad |r| \geqslant T \\ \lambda^2(\ln \frac{T}{|r|} + \frac{|r|}{T} - 1) & \text{if} \quad \epsilon T \leqslant |r| \leqslant T \\ \lambda^2(\ln \frac{1}{\epsilon} + \frac{|r|}{T} - \frac{|r|}{\epsilon T}) & \text{if} \quad |r| \leqslant \epsilon T\,. \end{cases} \tag{3.20}$$

Since k_ϵ reduces to 0 outside a compact set, it is straightforward to check (3.7) and (3.8). We further stress that this measure admits a cut-off in the sense of Definition 3.6.

Remark. In view of Theorem 3.3, note that the random measure M is a Gaussian multiplicative chaos with associated kernel

$$K(r) = \int_{|r|}^{+\infty} \frac{k(u)}{u}\, du \quad \text{with} \quad k(u) = \lambda^2(1 - \frac{|u|}{T})\mathbf{1}_{[0,T]}(|u|). \tag{3.21}$$

and that we have

$$k_\epsilon(r) = \int_{|r|}^{\frac{|r|}{\epsilon}} \frac{k(u)}{u}\, du.$$

3.4 Construction of log-normal $\star$-scale invariant random measures

This section is devoted to the existence part of Theorem 3.3: we give an explicit construction of lognormal $\star$-scale invariant random measures.

We are given a positive random variable $Y \in L^{1+\delta}$ (for some $\delta > 0$) and a continuous covariation kernel k such that $k(0) \leqslant \frac{2}{1+\delta}$. Let F be the (symmetric) spectral measure associated to k, that is

$$k(t) = \int_{\mathbb{R}} e^{i\lambda t} F(d\lambda),$$

and we assume that the improper integral

$$K(r) = \int_r^{+\infty} \frac{k(u)}{u}\, du$$

converges for $r > 0$.

Let μ, ν be two i.i.d. independently scattered Gaussian random measures (independent of Y) distributed on the half plane $\mathbb{R} \times \mathbb{R}_+^*$ such that:

$$\forall A \in \mathcal{B}(\mathbb{R} \times \mathbb{R}_+^*), \quad \mathbb{E}[e^{q\mu(A)}] = e^{\frac{1}{2}q^2\theta(A)}$$

where

$$\theta(A) = \int_{\lambda\in\mathbb{R}} \int_{y\in\mathbb{R}_+^*} \mathbf{1}_A(\lambda, y) \frac{1}{y} dy F(d\lambda).$$

Let $\epsilon < 1$, we define the centered Gaussian process

$$\forall t \in \mathbb{R}, \quad X_\epsilon(t) = \int_{\lambda\in\mathbb{R}} \int_{y\in[1,\frac{1}{\epsilon}[} \cos(\lambda t y)\mu(d\lambda, dy) + \int_{\lambda\in\mathbb{R}} \int_{y\in[1,\frac{1}{\epsilon}[} \sin(\lambda t y)\nu(d\lambda, dy).$$

It is plain to compute its covariation kernel, call it k_ϵ, by using the symmetry of the spectral measure $F(d\lambda)$:

$$\begin{aligned}
k_\epsilon(t-s) &= \mathbb{E}[X_\epsilon(s)X_\epsilon(t)] \\
&= \int_{\lambda\in\mathbb{R}} \int_{y\in[1,\frac{1}{\epsilon}[} \cos(\lambda t y)\cos(\lambda t s)\frac{1}{y} dy F(d\lambda) + \int_{\lambda\in\mathbb{R}} \int_{y\in[1,\frac{1}{\epsilon}[} \sin(\lambda t y)\sin(\lambda s y)\frac{1}{y} dy F(d\lambda) \\
&= \int_{\lambda\in\mathbb{R}} \int_{y\in[1,\frac{1}{\epsilon}[} \cos(\lambda(t-s)y)\frac{1}{y} dy F(d\lambda) \\
&= \int_{y\in[1,\frac{1}{\epsilon}[} \int_{\lambda\in\mathbb{R}} e^{i\lambda(t-s)y} F(d\lambda)\frac{1}{y} dy \\
&= \int_{y\in[1,\frac{1}{\epsilon}[} \frac{k(|t-s|y)}{y} dy \\
&= \int_{|t-s|}^{\frac{1}{\epsilon}|t-s|} \frac{k(y)}{y} dy.
\end{aligned}$$

For all $A \in \mathcal{B}(\mathbb{R})$, the process

$$M_{1/l}(A) = Y \int_A \exp\left(X_{1/l}(r) - \frac{1}{2}\mathbb{E}[X^2_{1/l}(r)]\right) dr$$

is obviously a positive martingale and thus converges as $l \to \infty$ towards a random variable $M(A)$. The stationary random measure $(M(A))_{A\in\mathcal{B}(\mathbb{R})}$ is a Gaussian multiplicative chaos in the sense of [129] with associated kernel K.

Note that for $l > 1/\epsilon$, we have $\forall t \in \mathbb{R}$:

$$\begin{aligned} X_{1/l}(t) =& X_\epsilon(t) + \int_{\lambda\in\mathbb{R}} \int_{y\in[\frac{1}{\epsilon},l[} \cos(\lambda t y)\mu(d\lambda, dy) + \int_{\lambda\in\mathbb{R}} \int_{y\in[\frac{1}{\epsilon},l[} \sin(\lambda t y)\nu(d\lambda, dy) \\ &\stackrel{def}{=} X_\epsilon(t) + \bar{X}_{\epsilon,1/l}(t), \end{aligned} \tag{3.22}$$

where $\bar{X}_{\epsilon,1/l}$ is a centered stationary Gaussian process independent from X_ϵ with covariance kernel given by:

$$\bar{k}_{\epsilon,1/l}(t-s) = \mathbb{E}[\bar{X}_{\epsilon,1/l}(s)\bar{X}_{\epsilon,1/l}(t)] = \int_{\frac{1}{\epsilon}|t-s|}^{l|t-s|} \frac{k(y)}{y} dy.$$

As above, we can define the random measure M^ϵ as the limit as $l \to +\infty$ of the random measures

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad M^\epsilon_{1/l}(A) = Y \int_A \exp\left(\bar{X}_{\epsilon,1/l}(r) - \frac{1}{2}\mathbb{E}[\bar{X}^2_{\epsilon,1/l}(r)]\right) dr.$$

The stationary random measure $(M^\epsilon(A))_{A\in\mathcal{B}(\mathbb{R})}$ is a Gaussian multiplicative chaos in the sense of [129] with associated covariance $K(\cdot\frac{1}{\epsilon})$. We deduce that $\frac{1}{\epsilon}M^\epsilon(\epsilon\,\cdot)$ is a Gaussian multiplicative chaos in the sense of [129] with associated covariance $K(\cdot)$. The measure $\frac{1}{\epsilon}M^\epsilon(\epsilon\,\cdot)$ thus has the same law as M. From (3.22), we obviously have:

$$M(A) = \int_A \exp\left(X_\epsilon(r) - \frac{1}{2}\mathbb{E}[X^2_\epsilon(r)]\right) M^\epsilon(dr)$$

in such a way that (3.5) holds. Finally we point out that M admits a moment of order $1+\gamma$ for all $0 \leqslant \gamma < \delta$ (see [94]).

Remark. By focusing on the above construction, we see that the covariance kernel k can be intuitively interpreted as some kind of infinitesimal stochastic generator. We may look X_ϵ as a sum

$$X_\epsilon(r) = \sum_{1\leqslant y\leqslant\frac{1}{\epsilon}} a_y Z^y_r$$

where $(Z^y)_y$ are independent centered Gaussian processes with kernel $k(y\cdot)$ and $(a_y)_y$ are independent random Gaussian variables with variance $\frac{dy}{y}$. So, when ϵ decreases infinitesimally, we "add" an independent Gaussian process with kernel $k(\frac{1}{\epsilon}\cdot)$ times an independent Gaussian factor of variance $\frac{-d\epsilon}{\epsilon}$.

Proof of Proposition 3.4 We show that the measure M is good under assumption (3.13). Because k is continuous, the kernel $k_\epsilon(r) = \int_{|r|}^{|r|/\epsilon} \frac{k(u)}{u}\, du$ is of class C^1 on $\mathbb{R}^*$. Thus, we have:

$$|k_\epsilon(r) - k_\epsilon(r')| \leqslant \sup_{u \geqslant \min(|r|,|r'|)} |k'_\epsilon(u)|.$$

Because we have

$$k'_\epsilon(u) = \frac{1}{u}(k(u/\epsilon) - k(u)),$$

it is plain to see that a reasonable choice for θ is $\theta(x) = \sup_{u\geqslant|x|} |\frac{k(u)}{u}|$ and $C_\epsilon = 2/\epsilon$.

$$\int_1^{+\infty} \ln r \sup_{|u|\geqslant r} \frac{|k(u)|}{u}\, dr < +\infty \Rightarrow \int_1^{+\infty} \ln r\, \theta(r)\, dr < +\infty,$$

so that the measure is good. □

3.4.1 Practical examples

In this subsection, we give practical examples of log-normal $\star$-scale invariant random measures. Using Theorem 3.3, **good** log-normal $\star$-scale invariant random measures are Gaussian multiplicative chaos whose covariance structure is given by

$$K(s) = \int_{|s|}^{+\infty} \frac{k(u)}{u} du \tag{3.23}$$

where k is a continuous covariance function satisfying $k(0) < 2$ and some weak decay assumptions (ensuring (3.13) for instance). Therefore, to define explicit examples, we just need to exhibit suitable kernels k. The decay assumptions can be read off the spectral measure of k. For instance, if k is the Fourier transform of some positive even integrable function f, which possesses an integrable derivative, it is a simple application of the Riemann theorem to prove that (3.13) is satisfied. Actually, for (3.13) to be satisfied, the assumptions on the regularity of the spectral measure can be much weakened. For instance, we can consider a kernel k that is the Fourier transform of some positive even integrable function f with integrable α-fractional derivative for $0 < \alpha < 1$:

$$\partial_\alpha f = \int_{\mathbb{R}^*} \frac{f(x+z) - f(x)}{|z|^{1+\alpha}} dz \in L^1(\mathbb{R}).$$

In that case, the Riemann theorem implies $|u|^\alpha k(u) \to 0$ as $|u| \to \infty$ and it is then plain to se that (3.13) is satisfied.

Below are listed a few examples of such kernels:

- the function $k(s) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{|s|^2}{2\sigma^2}}$ (where $\sigma > 0$) is continuous and positive-definite since its Fourier transform $\hat{k}(r) = e^{-\sigma^2 r^2/2}$ is positive.
- the covariance function of the stationary Orstein-Uhlenbeck process which takes on the form $k(s) = \frac{\sigma^2}{2\theta} e^{-\theta|s|}$ where $\theta > 0, \sigma > 0$.
- we can consider k as the Fourier transform of the function ($\lambda > 0$)

$$f(x) = \int_0^{+\infty} e^{-\lambda t} \mathbb{E}[g(x + X_t)]\, dt \tag{3.24}$$

  where $g \in L^1(\mathbb{R})$ is any positive integrable function and X is a pure jump Lévy process with Lévy symbol

$$\eta(u) = \int_{\mathbb{R}^*} (e^{iuz} - 1) \frac{1}{|z|^{1+\alpha}} dz$$

  for some $0 < \alpha < 1$. It is well know that the Lebesgue measure is invariant for the semi-group generated by X so that $k(0) = \|f\|_1 = \|g\|_1/\lambda$: this gives a condition on the norm $\|g\|_1$ for having $k(0) < 2$. Furthermore, f admits an integrable α-fractional derivative so that (3.13) is satisfied. Actually, it turns out that all the functions in $L^1(\mathbb{R})$ with an integrable α-fractional derivative admit a representation as (3.24). The reader may consult [16] for further details.

We stress that, as soon as they are not trivial (i.e. $k(0) < 2$), the Gaussian multiplicative chaos of the first two above examples do not have cut off in the sense of Definition 3.6. Obviously, many other examples exist.

Let us mention another example of log-normal $\star$-scale invariant random measures which does not present the goodness property of Definition 3.2. From Theorem 3.3, the Gaussian multiplicative chaos associated to the covariance function

$$K(s) = \int_{|s|}^{+\infty} \frac{\cos(u)}{u} du. \tag{3.25}$$

is log-normal $\star$-scale invariant in the sense of Definition 3.1. The function $k(r) = \cos(r)$ is indeed positive definite since its spectral measure is the positive measure $(\delta_1(dx) + \delta_{-1}(dx))/2$. The kernel $k_\epsilon(r) = \int_{|r|}^{|r|/\epsilon} \frac{\cos(u)}{u} du$ does not satisfy (3.8) so that the associated measure M is not good. Note that this Gaussian multiplicative chaos falls under the scope of [129] since the function K does not have a constant positive sign.

3.5 Characterization of star scale invariance

This section is devoted to the proof of the first statement of Theorem 3.3. For the sake of readability, some proofs of auxiliary results are gathered in the appendix.

Let M be a good log-normal scale invariant random measure defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We introduce as usually the spaces L^p on $(\Omega, \mathcal{F}, \mathbb{P})$ for $1 \leqslant p \leqslant \infty$. Recall that the measure M satisfies, for all $\epsilon \in (0,1)$

$$\big(M(A)\big)_{A\in\mathcal{B}(\mathbb{R})} \overset{law}{=} \big(\int_A e^{\omega_\epsilon(r)} M^\epsilon(dr)\big)_{A\in\mathcal{B}(\mathbb{R})} \tag{3.26}$$

where ω_ϵ is a Gaussian process independent from M^ϵ, with $M^\epsilon(dr) = \epsilon M(\frac{dr}{\epsilon})$ in law. k_ϵ denotes the covariation kernel of the process ω_ϵ. Furthermore, we assume that the measure M is non trivial ($M \neq 0$) with a moment of order $1+\delta$ so that the process ω_ϵ is necessarily normalized, that is $\mathbb{E}[e^{\omega_\epsilon}] = 1$.

Now we introduce some definitions and tools that will be used throughout this section. For each $\epsilon \in (0,1)$, define

$$\forall r \neq 0, \quad K^\epsilon(r) = \sum_{n=0}^{+\infty} k_\epsilon\big(\frac{r}{\epsilon^n}\big). \tag{3.27}$$

The uniform convergence of the series on the sets $\{r \in \mathbb{R}; |r| > \rho\}$ for any $\rho > 0$ is ensured by (3.10) since for $|r| > \rho$:

$$\begin{aligned}
\sum_{n=0}^{+\infty} |k_\epsilon\big(\frac{r}{\epsilon^n}\big)| &\leqslant C_\epsilon \sum_{n=0}^{+\infty} \int_{\frac{|r|}{\epsilon^n}}^{+\infty} \theta(u)\, du \leqslant C_\epsilon \sum_{n=0}^{+\infty} \int_{\frac{\rho}{\epsilon^n}}^{+\infty} \theta(u)\, du \\
&\leqslant C_\epsilon \int_0^{+\infty} \int_{\rho\epsilon^{-y+1}}^{+\infty} \theta(u)\, du\, dy \\
&= C_\epsilon \int_{\rho\epsilon}^{+\infty} \theta(u) \int_0^{\frac{\ln\frac{u}{\rho}}{-\ln\epsilon}+1} dy\, du \\
&= \frac{C_\epsilon}{-\ln\epsilon} \int_{\epsilon\rho}^{+\infty} \theta(u) \ln\frac{u}{\epsilon\rho}\, du
\end{aligned} \tag{3.28}$$

and this last integral is assumed to be converging (3.9). Furthermore, (3.8) also ensures that K^ϵ is Lipschitzian over each set $\{z \in \mathbb{R}; |z| > \rho\}$ for any $\rho > 0$ because:

$$\begin{aligned}
|K^\epsilon(r) - K^\epsilon(r')| &\leqslant \sum_{n=0}^{+\infty} |k_\epsilon\big(\frac{r}{\epsilon^n}\big) - k_\epsilon\big(\frac{r'}{\epsilon^n}\big)| \\
&\leqslant C_\epsilon \sum_{n=0}^{+\infty} \theta\big(\frac{\min(|r|,|r'|)}{\epsilon^n}\big) \big|\frac{r-r'}{\epsilon^n}\big| \\
&\leqslant C_\epsilon \int_0^{+\infty} \theta\big(\frac{\rho}{\epsilon^{y-1}}\big) \big|\frac{r-r'}{\epsilon^y}\big|\, dy \\
&\leqslant \frac{C_\epsilon}{-\rho\epsilon\ln\epsilon} |r-r'| \int_{\rho\epsilon}^{+\infty} \theta(u)\, du.
\end{aligned}$$

We let $(X^n)_n$ denote a sequence of independent centered stationary Gaussian processes with respective covariance kernels

$$\mathbb{E}[X^n_r X^n_s] = k_\epsilon\big(\frac{r-s}{\epsilon^n}\big) \overset{\text{def}}{=} \overline{k}_n(r-s).$$

Clearly X^n depends on ϵ but this parameter is omitted from the notations for the sake of readability. We assume that the whole sequence $(X^n)_n$ and the measure M are constructed on the same probability space and are mutually independent. We further define the measure M^N for $N \geqslant 0$ by

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad M^N(A) = \epsilon^{N+1} M\big(\frac{1}{\epsilon^{N+1}} A\big).$$

Note that $\mathbb{E}[M^N(A)] = |A|$ where $|A|$ stands for the Lebesgue measure of the set A.

By iterating the scale invariance relation (3.5), it is plain to see that, for each $N \geqslant 0$, the measure $\widetilde{M}^N$ defined by

$$\widetilde{M}^N(A) = \int_A \exp\Big(\sum_{n=0}^{N} X_r^n - \frac{1}{2}\mathbb{E}[(X_r^n)^2]\Big)\, M^N(dr) \tag{3.29}$$

has the same law as the measure M.

3.5.1 Ergodic properties

First we investigate the immediate properties of M resulting from the definitions.

Lemma 3.10. *Let M be a stationary random measure on $\mathbb{R}$ admitting a moment of order $1+\delta$. There is a nonnegative integrable random variable $Y \in L^{1+\delta}$ such that, for every bounded interval $I \subset \mathbb{R}$,*

$$\lim_{T\to\infty} \frac{1}{T} M(TI) = Y|I| \quad \textit{almost surely and in } L^{1+\delta},$$

where $|\cdot|$ stands for the Lebesgue measure on $\mathbb{R}$. As a consequence, almost surely the random measure

$$A \in \mathcal{B}(\mathbb{R}) \mapsto \frac{1}{T} M(TA)$$

weakly converges towards $Y|\cdot|$ and $\mathbb{E}_Y[M(A)] = Y|A|$ ($\mathbb{E}_Y[\cdot]$ denotes the conditional expectation with respect to Y).

Proof. If M is a stationary random measure, the Birkhoff ergodic theorem implies the following convergence, for $n \in \mathbb{N}, n \to \infty$,

$$\frac{1}{n} M([0,n]) = \frac{1}{n}\sum_{i=1}^{n} M([i-1,i]) \to Y \quad \text{almost surely and in } L^{1+\delta} \tag{3.30}$$

where $Y \in L^{1+\delta}$ is a nonnegative random variable. Using monotonicity of the mapping $t \mapsto M([0,t])$, one can show that $\frac{1}{T}M([0,T]) \to Y$ almost surely and in $L^{1+\delta}$. For $a > 0, b > a$, it is clear that $\frac{1}{T}M(T[0,a]) \to aY$ and that $\frac{1}{T}M(T[a,b]) \to (b-a)Y$ almost surely and in $L^{1+\delta}$. So, for every bounded interval $I \subset \mathbb{R}_+$, the following convergence holds $\frac{1}{T}M(TI) \to |I|Y$ almost surely and in $L^{1+\delta}$. Along the same lines, one can show the same convergence for every bounded interval $I \subset \mathbb{R}_-$ involving some nonnegative random variable $Y' \in L^{1+\delta}$. Stationarity implies that $\frac{1}{T}M(T[-1,1])$ has the same law as $\frac{1}{T}M(T[0,2])$. By letting T go to ∞, we find that $Y+Y'$ has the same law as $2Y$. Stationarity also implies that Y' has the same law as Y. Let $0 < \alpha < 1$. We prove

$$\mathbb{E}[Y^\alpha] = \mathbb{E}\left[\left(\frac{Y+Y'}{2}\right)^\alpha\right] \geqslant \frac{1}{2}\left(\mathbb{E}[Y^\alpha] + \mathbb{E}[Y'^\alpha]\right) = \mathbb{E}[Y^\alpha] \tag{3.31}$$

by using the Jensen inequality for the concave function $x \mapsto x^\alpha$. So the above inequality turns out to be an equality and thus $Y = Y'$ almost surely. We have shown that $\frac{1}{T}M(TI) \to |I|Y$ almost surely and in $L^{1+\delta}$ when $T \to \infty$ for every bounded interval $I \subset \mathbb{R}$.

Finally, by the portemanteau theorem, the convergence of the measure $A \in \mathcal{B}(\mathbb{R}) \mapsto \frac{1}{T}M(TA)$ on the intervals towards $Y|\cdot|$ is enough to ensure the weak convergence. □

3.5.2 Mixing properties

This section is devoted to study of the mixing properties of the measure M, which can be read off the structure of the kernel K^ϵ.

We first draw attention to the following relation, which will be used throughout the paper:

$$\mathbb{E}_Y\big[F\big(M(A_1),\dots,M(A_n)\big)\big] = \mathbb{E}_Y\big[F\big(\widetilde{M}^N(A_1),\dots,\widetilde{M}^N(A_n)\big)\big] \quad a.s.$$

for every positive measurable function $F : \mathbb{R}^n \to \mathbb{R}$. The proof is deferred to appendix 3.6 (see Lemma 3.23).

Lemma 3.11. *Let A, B be two disjoint sets such that* $\mathrm{dist}(A,B) > 0$. *Then the random variable $M(A)M(B)$ is integrable under $\mathbb{E}_Y[.]$ and*

$$\mathbb{E}_Y[M(A)M(B)] = Y^2 \int_{A\times B} e^{K^\epsilon(r-u)} dr\, du.$$

Proof. We fix $R > 0$ and denote by $\mathcal{G}$ the σ-field generated by M. Because the function $x \in \mathbb{R}_+ \mapsto \min(R,x)$ is concave, we have

$$\begin{aligned}\mathbb{E}_Y\big[\min\big(R, M(A)M(B)\big)\big] &= \mathbb{E}_Y\big[\min\big(R, \widetilde{M}^N(A)\widetilde{M}^N(B)\big)\big]\\ &= \mathbb{E}_Y\big[\mathbb{E}\big[\min\big(R, \widetilde{M}^N(A)\widetilde{M}^N(B)\big)|\mathcal{G}\big]\big]\\ &\leqslant \mathbb{E}_Y\Big[\min\Big(R, \mathbb{E}\big[\widetilde{M}^N(A)\widetilde{M}^N(B)|\mathcal{G}\big]\Big)\Big].\end{aligned}$$

Since $\widetilde{M}^N$ is given by (3.29), it is straightforward to compute:

$$\mathbb{E}\big[\widetilde{M}^N(A)\widetilde{M}^N(B)|\mathcal{G}\big] = \int_{A\times B} e^{\sum_{n=0}^N \bar{k}_n(r-u)} M^N(dr)M^N(du). \tag{3.32}$$

Because of the uniform convergence of the series $\big(\sum_{n=0}^N \bar{k}_n(r-u)\big)_N$ on the set $\{(r,u)\in\mathbb{R}^2; |r-u| \geqslant d\}$ towards K^ϵ and the weak convergence of the measure M^N towards $Y|\cdot|$ (cf. Lemma 3.10), the random variable

$$\int_{A\times B} e^{\sum_{n=0}^N \bar{k}_n(r-u)} M^N(dr)M^N(du)$$

almost surely converges towards

$$Y^2 \int_{A\times B} e^{K^\epsilon(r-u)} dr\, du.$$

The dominated convergence theorem then yields:

$$\mathbb{E}_Y\big[\min\big(R, M(A)M(B)\big)\big] \leqslant \mathbb{E}_Y\Big[\min\Big(R, Y^2\int_{A\times B} e^{K^\epsilon(r-u)} dr\, du\Big)\Big].$$

By letting $R\to\infty$, the monotone convergence theorem yields

$$\mathbb{E}_Y\big[M(A)M(B)\big] \leqslant Y^2 \int_{A\times B} e^{K^\epsilon(r-u)} dr\, du.$$

On the other hand, we also have

$$\mathbb{E}_Y\big[M(A)M(B)\big] = \mathbb{E}_Y\big[\widetilde{M}^N(A)\widetilde{M}^N(B)\big] = \mathbb{E}_Y\big[\mathbb{E}\big[\widetilde{M}^N(A)\widetilde{M}^N(B)|\mathcal{G}\big]\big]. \tag{3.33}$$

By gathering (3.32) and (3.33) and by using the Fatou's lemma, we deduce

$$\mathbb{E}_Y\big[M(A)M(B)\big] \geqslant Y^2 \int_{A\times B} e^{K^\epsilon(r-u)} dr\, du.$$

This completes the proof. □

Lemma 3.12. *We have*

$$\sup_{|r|\geqslant d} |K^\epsilon(r)| \to 0 \quad \text{as } d\to\infty.$$

Proof. By using (3.28), we have for $|r| \geqslant d$:

$$|K^\epsilon(r)| \leqslant \frac{C_\epsilon}{-\ln\epsilon}\int_{\epsilon d}^{+\infty} \theta(u)\ln\frac{u}{\epsilon d}\, du$$

Now, if $\epsilon d \geqslant 1$, we have:

$$\sup_{|r|\geqslant d} |K^\epsilon(r)| \leqslant \frac{C_\epsilon}{-\ln\epsilon}\int_{\epsilon d}^{+\infty} \theta(u)\ln u\, du$$

Hence the result follows from the convergence of the last integral. □

Proposition 3.13. *The measure M possesses the following mixing property: given two disjoint sets A, B such that $\mathrm{dist}(A,B) = d > 0$ we have:*

$$\left|\mathbb{E}_Y\left[M(A)M(B)\right] - Y^2|A||B|\right| \leqslant Y^2\xi(d)|A||B| \tag{3.34}$$

for some function $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{d\to\infty}\xi(d) = 0$.

As a consequence, for any Lebesgue integrable function ϕ on $\mathbb{R}^2$ and $d > 0$, we have:

$$\left|\mathbb{E}_Y\Big[\int_{|u-r|>d}\phi(u,r)M(dr)M(du)\Big] - Y^2\int_{|u-r|>d}\phi(u,r)\,du\,dr\right| \leqslant Y^2\xi(d)\int_{|u-r|>d}|\phi(u,r)|\,du\,dr. \tag{3.35}$$

Proof. From Lemma 3.11, we have

$$\begin{aligned}\left|\mathbb{E}_Y\left[M(A)M(B)\right] - Y^2|A||B|\right| &= Y^2\int_{A\times B}(e^{K^\epsilon(r-u)} - 1)dr\,du\\ &\leqslant Y^2\varepsilon(d)|A||B|\end{aligned}$$

where we have set $\xi(d) = \sup_{|r|\geqslant d}|e^{K^\epsilon(r)} - 1|$. From Lemma 3.12, we have $\lim_{d\to\infty}\xi(d) = 0$. It is then plain to derive (3.35). □

As a direct consequence, we obtain:

Corollary 3.14. *For any Lebesgue integrable function ϕ on $\mathbb{R}^2$ and $d > 0$, we have for all $N \in \mathbb{N}\setminus\{0\}$:*

$$\begin{aligned}\Big|\int_{|u-r|>d}\phi(u,r)\mathbb{E}_Y\left[M^N(dr)M^N(du)\right] - Y^2\int_{|u-r|>d}|\phi(u,r)|\,du\,dr\Big|&\\ \leqslant Y^2\xi\big(\frac{d}{\epsilon^N}\big)\int_{|u-r|>d}|\phi(u,r)|\,du\,dr.&\end{aligned}$$

3.5.3 Characterization of the measure M

Having in mind that the measure M^N weakly converges towards $Y|\cdot|$ as N goes to infinity, it is very tantalizing to think that the solution of our problem reduces to taking the limit in (3.29) as $N \to \infty$. However, multiplicative chaos badly behaves with respect to weak convergence of measures. So we want to get rid of the measure M^N and have the Lebesgue measure instead in order to deal with a multiplicative chaos in the sense of Kahane. This is the main difficulty of the proof. For that purpose, it is appropriate to take the conditional expectation of $\widetilde{M}^N$ with respect to the σ-algebra $\mathcal{F}_N = \sigma(X^0,\dots,X^N,Y)$. Therefore, for any Borelian subset A of $\mathbb{R}$, we define

$$G_N(A) = \mathbb{E}[\widetilde{M}^N(A)|\mathcal{F}_N]$$

and we claim

Lemma 3.15. *The following relation holds for each $N \geqslant 0$:*

$$G_N(A) = Y\int_A \exp\Big(\sum_{n=0}^N X_r^n - \frac{1}{2}\mathbb{E}[(X_r^n)^2]\Big)\,dr. \tag{3.36}$$

Furthermore, for each bounded Borelian set A, the sequence $(G_N(A))_N$ is a positive martingale bounded in $L^{1+\delta}$.

Proof. If A has infinite Lebesgue measure, both sides of (3.36) are infinite. So we focus on the case when A has finite Lebesgue measure. First observe that for each $s < t$ and $A \in \mathcal{F}_N$, we have from Lemma 3.10

$$\mathbb{E}[\int_{\mathbb{R}}\mathbf{1}_{[s,t]}(r)\mathbf{1}_A M^N(dr)|\mathcal{F}_N] = \mathbf{1}_A\mathbb{E}_Y[M^N([s,t])] = \mathbf{1}_A Y(t-s).$$

By using density arguments and Fatou's lemma, we establish that, for each positive $\mathcal{F}_N \otimes \mathcal{B}(\mathbb{R})$-measurable function $\varphi \in L^1(\Omega \times \mathbb{R}; \mathbb{P} \otimes dt)$, we have

$$\mathbb{E}\Big[\int_{\mathbb{R}} \varphi(\omega, r) M^N(dr) \big| \mathcal{F}_N\Big] = \int_{\mathbb{R}} \varphi(\omega, r) Y \, dr.$$

So (3.36) is proved.

Finally, for each bounded set A we have $\mathbb{E}[M(A)^{1+\delta}] < +\infty$ for some $\delta > 0$. The Jensen inequality then yields

$$\mathbb{E}[(G_N(A))^{1+\delta}] = \mathbb{E}[(\mathbb{E}[\widetilde{M}^N(A)|\mathcal{F}_N])^{1+\delta}] \leqslant \mathbb{E}[(\widetilde{M}^N(A))^{1+\delta}] = \mathbb{E}[M(A)^{1+\delta}] < +\infty.$$

The martingale $(G_N(A))_N$ is thus bounded in $L^{1+\delta}$. □

Being bounded in $L^{1+\delta}$, the martingale converges almost surely and in $L^{1+\delta}$ towards a random variable $Q(A)$, which can be formally thought of as

$$Q(A) = Y \int_A \exp\left(X_r - \frac{1}{2}\mathbb{E}[X_r^2]\right) dr$$

where $(X_r)_{r\in\mathbb{R}}$ is a "Gaussian process" with covariance kernel $K^\epsilon(r)$, that is a Gaussian multiplicative chaos. The remaining part of our argument can be roughly summed up as follows. First, we obtain estimates on the kernel K^ϵ derived from the fact that the Gaussian multiplicative chaos Q admits a moment of order $1+\delta$. Second, we use these estimates to prove that Q has the same law as M. Finally, since Q has the same law as M, which does not depend on ϵ, the kernel K^ϵ should not depend on ϵ either. This is a strong constraint on K^ϵ, from which we derive the specific structure of K^ϵ given by (3.12).

So we claim

Proposition 3.16. *For each $0 < \gamma < \delta$, we can find $\rho > 0$ such that:*

$$\sup_n n^{1+\rho}\mathbb{E}[M([0, \frac{1}{n}])^{1+\gamma}] < +\infty. \tag{3.37}$$

Proof. The proof relies on the following bound (see the proof below):

Lemma 3.17. *The existence of a moment of order $1+\delta$ for the measure M implies the following bound:*

$$k_\epsilon(0) \leqslant \frac{2}{1+\delta} \ln \frac{1}{\epsilon}.$$

Since we have for all $r \in \mathbb{R}$: $k_\epsilon(r) \leqslant k_\epsilon(0)$, the covariance kernel of the process ω_ϵ is dominated by that of the constant process $\omega_\epsilon(0)$. Hence, by using (3.5) and Lemma 3.24, it is plain to see that, for each $\gamma > 0$:

$$\begin{aligned}
\mathbb{E}[M([0, \tfrac{1}{n}])^{1+\gamma}] &= \mathbb{E}\left[\Big(\int_0^{1/n} e^{\omega_{1/n}(r)} M^{1/n}(dr)\Big)^{1+\gamma}\right] \\
&\leqslant \mathbb{E}\left[\Big(\int_0^{1/n} e^{\omega_{1/n}(0)} M^{1/n}(dr)\Big)^{1+\gamma}\right] \\
&\leqslant \mathbb{E}\left[e^{(1+\gamma)\omega_{1/n}(0)}\right] \mathbb{E}\left[\Big(M^{1/n}([0, \tfrac{1}{n}])\Big)^{1+\gamma}\right] \\
&= e^{\frac{(1+\gamma)^2}{2} k_{1/n}(0) - \frac{1+\gamma}{2} k_{1/n}(0)} \mathbb{E}\Big[\Big(M([0,1])\Big)^{1+\gamma}\Big] \frac{1}{n^{1+\gamma}}.
\end{aligned}$$

Since $k_{1/n}(0) \leqslant \frac{2}{1+\delta} \ln n$, we deduce

$$\begin{aligned}
\mathbb{E}[M([0, \tfrac{1}{n}])^{1+\gamma}] &\leqslant e^{\left(\frac{\gamma^2+\gamma}{1+\delta} - \gamma - 1\right) \ln n} \mathbb{E}\Big[\Big(M([0,1])\Big)^{1+\gamma}\Big] \\
&= \frac{1}{n^{1+\rho}} \mathbb{E}\Big[\Big(M([0,1])\Big)^{1+\gamma}\Big]
\end{aligned}$$

where we have set

$$\rho \stackrel{def}{=} -\frac{\gamma^2+\gamma}{1+\delta}+\gamma.$$

Clearly, we have $\rho > 0$ provided that $0 < \gamma < \delta$. The proof of Proposition 3.16 is complete. □

Proof of Lemma 3.17. Let $n \in \mathbb{N}$.

$$\mathbb{E}\left[M[0;t]^{1+\delta}\right] = \mathbb{E}\left[\left(M[0;\frac{t}{n}]+M[\frac{t}{n};\frac{2t}{n}]+\cdots+M[\frac{(n-1)t}{n};t]\right)^{1+\delta}\right] \tag{3.38}$$

$$\geqslant \mathbb{E}\left[\left(M[0;\frac{t}{n}]\right)^{1+\delta}+\left(M[\frac{t}{n};\frac{2t}{n}]\right)^{1+\delta}+\cdots+\left(M[\frac{(n-1)t}{n};t]\right)^{1+\delta}\right] \tag{3.39}$$

$$= n\mathbb{E}\left[\left(M[0;\frac{t}{n}]\right)^{1+\delta}\right] \tag{3.40}$$

We used the stationarity of the measure M in the second line. Now write, for $h > 0$:

$$g(h) = \sup_{r \leqslant h} \mid k_{1/n}(0) - k_{1/n}(r) \mid \tag{3.41}$$

We have, for every $r \in (0, t/n]$ and n large enough:

$$\mid k^{1/n}(0) - g(t/n) \mid \leqslant k^{1/n}(r).$$

So, using classical Gaussian inequality (see Lemma 3.24):

$$\begin{aligned}
\mathbb{E}\left[M[0;\frac{t}{n}]^{1+\delta}\right] &= \mathbb{E}\left[\left(\int_0^{t/n} e^{\omega_{1/n}(r)} M^{1/n}(dr)\right)^{1+\delta}\right] \\
&\geqslant \mathbb{E}\left[\left(\int_0^{t/n} e^{\sqrt{|k_{1/n}(0)-g(t/n)|}Z_n - \frac{1}{2}|k_{1/n}(0)-g(t/n)|} M^{1/n}(dr)\right)^{1+\delta}\right] \\
&= \mathbb{E}\left[\left(e^{\sqrt{|k_{1/n}(0)-g(t/n)|}Z_n - \frac{1}{2}|k_{1/n}(0)-g(t/n)|}\right)^{1+\delta}\right] \mathbb{E}\left[\left(M^{1/n}[0;\frac{t}{n}]\right)^{1+\delta}\right] \\
&= e^{-\frac{1+\delta}{2}|k_{1/n}(0)-g(t/n)|} e^{\frac{(1+\delta)^2}{2}|k_{1/n}(0)-g(t/n)|} \frac{1}{n^{1+\delta}} \mathbb{E}\left[(M[0;t])^{1+\delta}\right]
\end{aligned} \tag{3.42}$$

We used Lemma 13 in the second line. Using equations (3.40) and (3.42), one gets

$$e^{-\frac{1+\delta}{2}|k_{1/n}(0)-g(t/n)|} e^{\frac{(1+\delta)^2}{2}|k_{1/n}(0)-g(t/n)|} \frac{1}{n^{\delta}} \leqslant 1 \tag{3.43}$$

As h goes to 0, $g(h)$ goes to 0 (the function $k_{1/n}$ is continuous). Letting t goes to 0 in (3.43), one gets

$$k_{1/n}(0) \leqslant \frac{2}{1+\delta} \ln n.$$

and the lemma is proved. □

We are now in position to tackle the main step of the proof:

Proposition 3.18. *The random measures $(Q(A))_{A\in\mathcal{B}(\mathbb{R})}$ and $(M(A))_{A\in\mathcal{B}(\mathbb{R})}$ have the same law.*

Proof. Let F be some function defined on $\mathbb{R}_+$ such that:

- F is convex,
- $F(x) \leqslant Cx^{1+\gamma}$ for some constants $C > 0$ and $0 < \gamma < \delta$,
- $F \circ \sqrt{\ }$ is concave, nondecreasing and sub-additive.

Let f be a lower semi-continuous positive function on $\mathbb{R}$ with compact support. We have by Jensen's inequality:

$$\begin{aligned}\mathbb{E}\Big[F\Big(\int_{\mathbb{R}} f(x)\,M(dx)\Big)\Big] &= \mathbb{E}\Big[F\Big(\int_{\mathbb{R}} f(x)\,\widetilde{M}^N(dx)\Big)\Big]\\ &= \mathbb{E}\Big[\mathbb{E}\Big[F\Big(\int_{\mathbb{R}} f(x)\,\widetilde{M}^N(dx)\Big)|\mathcal{F}_N\Big]\Big]\\ &\geqslant \mathbb{E}\Big[F\Big(\int_{\mathbb{R}} f(x)\,G_N(dx)\Big)\Big].\end{aligned}$$

We let N go to $+\infty$. By using the weak convergence of $G_N(dr)$ towards $Q(dr)$, we obtain:

$$\mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\,M(dr)\big)\big] \geqslant \mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\,Q(dr)\big)\big]. \tag{3.44}$$

Now we want to establish the converse inequality. We set $\widetilde{F} = F\circ\sqrt{\ }$. For any $\tau>0$, we have by using the sub-additivity of $\widetilde{F}$:

$$\begin{aligned}\mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\,M(dr)\big)\big] =& \mathbb{E}\Big[\widetilde{F}\Big(\big(\int_{\mathbb{R}} f(r)\,\widetilde{M}^N(dr)\big)^2\Big)\Big]\\ =& \mathbb{E}\Big[\widetilde{F}\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} f(r)f(u)\,\widetilde{M}^N(dr)\widetilde{M}^N(du)\Big)\Big]\\ \leqslant& \mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|\leqslant\tau} f(r)f(u)\,\widetilde{M}^N(dr)\widetilde{M}^N(du)\Big)\Big]\\ &+\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau} f(r)f(u)\,\widetilde{M}^N(dr)\widetilde{M}^N(du)\Big)\Big].\end{aligned}$$

Then, by conditioning with respect to $\mathcal{F}_N$ and by using the Jensen inequality in the second term of the latter inequality, we deduce:

$$\begin{aligned}&\mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\,M(dr)\big)\big] \qquad (3.45)\\ &\leqslant \mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|\leqslant\tau} f(r)f(u)\,\widetilde{M}^N(dr)\widetilde{M}^N(du)\Big)\Big]\\ &\quad+\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau} f(r)f(u)\exp\big(\sum_{k=0}^{N} X_r^n + X_u^n - k_n(0)\big)\,\mathbb{E}_Y[M^N(dr)M^N(du)]\Big)\Big]\\ &\overset{def}{=} C(1,\tau,N)+C(2,\tau,N). \qquad (3.46)\end{aligned}$$

We claim:

Lemma 3.19. *For each fixed $\tau>0$, $C(2,\tau,N)$ converges as $N\to\infty$ towards*

$$\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau} f(r)f(u)\,Q(dr)Q(du)\Big)\Big].$$

Furthermore, this latter quantity converges, as $\tau\to 0$, towards

$$\mathbb{E}\Big[F\Big(\int f(r)\,Q(dr)\Big)\Big].$$

Finally, the quantity $C(1,\tau,N)$ converges to 0 as $\tau\to 0$ uniformly with respect to $N\in\mathbb{N}^*$.

Let us admit for a while the above lemma to finish the proof of Proposition 3.18. By gathering (3.46) and Lemma 3.19, we deduce

$$\begin{aligned}\mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\,M(dr)\big)\big] &\leqslant \liminf_{\tau\to 0}\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau} f(r)f(u)\,Q(dr)Q(du)\Big)\Big]\\ &= \mathbb{E}\Big[F\Big(\int f(r)\,Q(dr)\Big)\Big].\end{aligned}$$

Hence we have proved

$$\mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\, M(dr)\big)\big] = \mathbb{E}\big[F\big(\int_{\mathbb{R}} f(r)\, Q(dr)\big)\big]. \tag{3.47}$$

The basic choice for F is the function $x \mapsto x^{1+\gamma}$ with $0 < \gamma < \delta$. Thus we have proved that the mappings

$$\mathbb{E}\big[\exp\big(z \ln \int_{\mathbb{R}} f(r)\, M(dr)\big)\big] \quad \text{and} \quad \mathbb{E}\big[\exp\big(z \ln \int_{\mathbb{R}} f(r)\, Q(dr)\big)\big]$$

coincide for $z \in]1, 1+\delta[$. By analyticity arguments, we deduce that $\int_{\mathbb{R}} f(x)\, M(dx)$ and $\int_{\mathbb{R}} f(x)\, Q(dx)$ have the same law. This is enough to prove that the random measures M and Q have the same law. Indeed, if we consider two families $(\lambda_i)_{1 \leqslant i \leqslant n}$ of positive real numbers and $(A_i)_{1 \leqslant i \leqslant n}$ of bounded open subsets of $\mathbb{R}$, we define the lower semi-continuous function

$$f(x) = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{A_i}(x)$$

and we obtain

$$\sum_{i=1}^{n} \lambda_i M(A_i) \overset{law}{=} \sum_{i=1}^{n} \lambda_i Q(A_i).$$

It turns out that the law of a random vector $(Y_1, \ldots, Y_n)$ made up of positive random variables is characterized by the combinations

$$\sum_{i=1}^{n} \lambda_i Y_i$$

where $(\lambda_i)_{1 \leqslant i \leqslant n}$ is a family of positive real numbers. The proof of Proposition 3.18 is complete. □

Proof of Lemma 3.19. Let us first investigate the quantity $C(1, \tau, N)$. Assume the function f has its support included in the ball $B(0, R)$ for some $R > 0$. We can cover the set

$$\{(x, y) \in \mathbb{R}^2; |x - y| \leqslant \tau \text{ and } \max(|x|, |y|) \leqslant R\}$$

by the squares

$$A_j^n = [t_j^n, t_{j+2}^n] \times [t_j^n, t_{j+2}^n] \quad \text{where } t_j^n = -R + 2\tau j, \quad \text{for } j = 0, \ldots, E(\frac{R}{\tau}).$$

We set $S = \sup_{\mathbb{R}} f$. Because $\widetilde{F}$ is sub-additive and increasing, we have:

$$\begin{aligned}
C(1, \tau, N) &\leqslant \mathbb{E}\Big[\widetilde{F}\Big(\sum_{0 \leqslant j \leqslant E(\frac{R}{\tau})} \int_{A_j^n} f(r) f(u)\, \widetilde{M}^N(dr) \widetilde{M}^N(du)\Big)\Big] \\
&\leqslant \sum_{0 \leqslant j \leqslant E(\frac{R}{\tau})} \mathbb{E}\Big[\widetilde{F}\Big(\int_{A_j^n} f(r) f(u)\, \widetilde{M}^N(dr) \widetilde{M}^N(du)\Big)\Big] \\
&\leqslant \sum_{0 \leqslant j \leqslant E(\frac{R}{\tau})} \mathbb{E}\Big[\widetilde{F}\Big(S^2 \int_{A_j^n} \widetilde{M}^N(dr) \widetilde{M}^N(du)\Big)\Big] \\
&= \sum_{0 \leqslant j \leqslant E(\frac{R}{\tau})} \mathbb{E}\Big[\widetilde{F}\Big(S^2 (\widetilde{M}^N([t_j^n, t_{j+2}^n]))^2\Big)\Big] \\
&= \sum_{0 \leqslant j \leqslant E(\frac{R}{\tau})} \mathbb{E}\Big[F\Big(SM([t_j^n, t_{j+2}^n])\Big)\Big].
\end{aligned}$$

By stationarity, we deduce

$$\begin{aligned}
C(1, \tau, N) &\leqslant \frac{2R}{\tau} \mathbb{E}\Big[F\Big(SM([0, 2\tau])\Big)\Big] \\
&\leqslant \frac{2R}{\tau} S^{1+\gamma} \mathbb{E}\big[M([0, 2\tau])^{1+\gamma}\big].
\end{aligned}$$

It results from Proposition 3.16 that the last quantity converges towards 0 as τ goes to 0 uniformly with respect to N.

Now we investigate the quantity $C(2,\tau,N)$. Since $\widetilde{F}$ is sub-additive and increasing, we have $|\widetilde{F}(a)-\widetilde{F}(b)|\leqslant\widetilde{F}(|b-a|)$ for all positive real numbers a,b. This together with Corollary 3.14 yields

$$\begin{aligned}
&\Big|C(2,\tau,N)-\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\exp\big(\sum_{k=0}^{N}X_r^n+X_u^n-k_n(0)\big)\,Y^2dr\,du\Big)\Big]\Big|\\
&\qquad\leqslant\mathbb{E}\Big[\widetilde{F}\Big(Y^2\xi\big(\tfrac{\tau}{\epsilon^N}\big)\int_{|r-u|>\tau}f(r)f(u)\exp\big(\sum_{k=0}^{N}X_r^n+X_u^n-k_n(0)\big)\,dr\,du\Big)\Big]\\
&\qquad\leqslant\mathbb{E}\Big[\widetilde{F}\Big(\xi\big(\tfrac{\tau}{\epsilon^N}\big)S^2G_N([-R,R])^2\Big)\Big]\\
&\qquad\leqslant\mathbb{E}\Big[F\Big(S\xi\big(\tfrac{\tau}{\epsilon^N}\big)^{1/2}G_N([-R,R])\Big)\Big]\\
&\qquad\leqslant\xi\big(\tfrac{\tau}{\epsilon^N}\big)^{\frac{1+\gamma}{2}}S^{1+\gamma}\mathbb{E}\Big[G_N([-R,R])^{1+\gamma}\Big].
\end{aligned}$$

Obviously, the last quantity converges to 0 as N goes to ∞. Furthermore, the quantity

$$\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\exp\big(\sum_{k=0}^{N}X_r^n+X_u^n-k_n(0)\big)\,Y^2dr\,du\Big)$$

almost surely converges towards

$$\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\,Q(dr)\,Q(du)\Big)$$

and is uniformly integrable because $F(x)\leqslant Cx^{1+\gamma}$ and Q is a multiplicative chaos admitting a moment of order $1+\delta$ with $\delta>\gamma$. The Lebesgue convergence theorem then yields:

$$\begin{aligned}
&\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\exp\big(\sum_{k=0}^{N}X_r^n+X_u^n-k_n(0)\big)\,Y^2dr\,du\Big)\Big]\\
&\qquad\to\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\,Q(dr)\,Q(du)\Big)\Big]\quad\text{as } N\to\infty.
\end{aligned}$$

Gathering the above relations yields

$$C(2,\tau,N)\to\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\,Q(dr)\,Q(du)\Big)\Big]\quad\text{as } N\to\infty.$$

Similar arguments as those used above allow to establish that

$$\begin{aligned}
\liminf_{\tau\to0}\mathbb{E}\Big[\widetilde{F}\Big(\int_{|r-u|>\tau}f(r)f(u)\,Q(dr)Q(du)\Big)\Big]&=\mathbb{E}\Big[\widetilde{F}\Big(\int_{\mathbb{R}^2}f(r)f(u)\,Q(dr)Q(du)\Big)\Big]\\
&=\mathbb{E}\Big[F\Big(\int_{\mathbb{R}}f(r)\,Q(dr)\Big)\Big].
\end{aligned}$$

Indeed, by proceeding as for $C(1,\tau,N)$, we can prove that the "diagonal contribution" goes to 0 as $\tau\to0$. Details are left to the reader. The proof of the Lemma is complete. □

The final step of our argument is now to prove that the kernel K^ϵ defined by (3.27) does not depend on ϵ. Expressing the kernel K^ϵ as a function of the marginals of the measure M is enough for that purpose. So we remind the reader of Lemma 3.11, which states

$$\mathbb{E}_Y[M(A)M(B)]=Y^2\int_{A\times B}e^{K^\epsilon(r-u)}\,drdu.$$

We deduce that, for any $s\neq0$ and on the set $\{Y>0\}$,

$$K^\epsilon(s)=\lim_{h\to0}\ln\Big(\frac{1}{h^2}\mathbb{E}_Y[M([0,h])M([s,s+h])]\Big)-2\ln Y.\tag{3.48}$$

As a straightforward consequence, the kernel K^ϵ defined by (3.27) does not depend on ϵ since the left-hand side in (3.48) does not either. So we can define the quantity

$$\forall r \neq 0, \quad K(r) = K^\epsilon(r)$$

for some $\epsilon \in (0,1)$ and this relation is also valid for any $\epsilon \in (0,1)$. It is also plain to see that for each $\epsilon \in (0,1)$ we have:

$$\forall r \neq 0, \quad K(r) = k_\epsilon(r) + K(\frac{r}{\epsilon}) \tag{3.49}$$

since K^ϵ satisfies such a relation. Such a specific functional equation implies a precise structure for the function K:

Proposition 3.20. *For $r > 0$, we have*

$$K(r) = \int_r^{+\infty} \frac{k(u)}{u} du \tag{3.50}$$

where $k(u)$ is a positive-definite continuous function $\mathbb{R}_+ \to \mathbb{R}$.

Proof. Because K is Lipschitzian on the compact subsets of $\mathbb{R} \setminus \{0\}$, there exists a locally bounded measurable function f on $(0; +\infty)$ such that for all $r, s > 0$,

$$K(s) - K(r) = \int_r^s f(t)dt.$$

Define, for $r \in \mathbb{R}$,

$$\phi(r) = K(e^r)$$

It is straightforward to derive from (3.49) that, for all $r \in \mathbb{R}, \alpha \geqslant 0$,

$$\phi(r+\alpha) - \phi(r) = -k_{e^{-\alpha}}(e^r) \tag{3.51}$$

Note that $k_1(e^r) = 0$. From equation (3.51), one obtains :

$$\frac{1}{\alpha} \int_r^{r+\alpha} e^u f(e^u) du = -\frac{k_{e^{-\alpha}}(e^r)}{\alpha} \tag{3.52}$$

For almost every r, the left-hand side of equation (3.52) tends to $e^r f(e^r)$ when α goes to 0. Thus, the right-hand side of (3.52) converges also for almost every r to $e^r f(e^r)$ when α goes to 0.

We define the function g by the following limit for almost every r:

$$g(r) = \lim_{\alpha \to 0} -\frac{1}{\alpha} \int_r^{r+\alpha} e^u f(e^u) du = \lim_{\alpha \to 0} \frac{k_{e^{-\alpha}}(e^r)}{\alpha} \tag{3.53}$$

As defined, the function g is measurable with respect to the Borelian σ-field of $\mathbb{R}$. For almost every $x \in (0, +\infty)$, define

$$h(x) = g(\ln(x)),$$

and $h(0)$ by $h(0) = \frac{k_{e^{-\alpha}}(0)}{\alpha}$ for some $\alpha > 0$. Note that the definition of $h(0)$ does not depend on α because:

Lemma 3.21. *We have the following asymptotic behaviour of K around 0:*

$$K(r) \simeq \frac{k_\epsilon(0)}{\ln \epsilon} \ln r \qquad \text{as } r \to 0.$$

Thus h is well defined at 0 and we can now prove that it is positive definite:

Lemma 3.22. *The function $h(|.|)$ is positive definite (as a tempered distribution in the sense of Schwartz, see [80] or [139]). One can also find a symmetric positive measure μ on $\mathbb{R}$ (with $\mu(\mathbb{R}) < \infty$) such that for almost every $x \in \mathbb{R}$:*

$$h(|x|) = \int_{\mathbb{R}} e^{ix\xi} \mu(d\xi)$$

Proof. For almost every $x \in \mathbb{R}$, $h(|x|) = \lim_{\alpha \to 0} \frac{k_{e^{-\alpha}}(|x|)}{\alpha}$ and $\frac{k_{e^{-\alpha}}(|x|)}{\alpha} \leqslant h(0)$ uniformly in α. Thus, if φ is a smooth function with compact support, we get using the dominated convergence theorem:

$$\int_{\mathbb{R}} \int_{\mathbb{R}} h(|y-x|)\varphi(x)\overline{\varphi}(y)dxdx = \lim_{\alpha \to 0} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{k_{e^{-\alpha}}(|y-x|)}{\alpha} \varphi(x)\overline{\varphi}(y)dxdx \geqslant 0.$$

We conclude that $h(|.|)$ is positive definite. By the Bochner-Schwartz theorem, the Fourier transform of $h(|.|)$ is a symmetric positive measure $\mu(d\xi)$ such that there exists $p \geqslant 0$ with:

$$\int_{\mathbb{R}} \frac{\mu(d\xi)}{(1+|\xi|)^p} < \infty.$$

In order to conclude, it is sufficient to prove that $\mu(\mathbb{R}) < \infty$. We note $\theta(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$ and $\theta^\epsilon = \frac{1}{\epsilon}\theta(./\epsilon)$ for $\epsilon > 0$. By the inverse Fourier theorem, we get:

$$(\theta^\epsilon * h)(0) = \int_{\mathbb{R}} e^{-\epsilon^2 \xi^2/2} \mu(d\xi).$$

Thus the right hand side of the above equality is bounded by $h(0)$ and we conclude by letting ϵ go to 0. □

Integrating with respect to the Lebesgue measure the relation $g(t) = -e^t f(e^t)$ which is true for almost every $t \in \mathbb{R}$, one gets

$$K(s) - K(r) = -\int_r^s \frac{h(u)}{u} du.$$

Because $K(s) \to 0$ as $s \to +\infty$, the function $u \mapsto \frac{h(u)}{u}$ is integrable at the vicinity of $+\infty$ in the generalized sense. We deduce:

$$K(r) = \int_r^{+\infty} \frac{h(u)}{u} du.$$

By the previous lemma, there exists a finite symmetric positive measure μ on $\mathbb{R}$ such that, for almost every $x \in \mathbb{R}$,

$$h(x) = \int_{\mathbb{R}} e^{ix\xi} \mu(d\xi)$$

For simplicity, define for all $x \in \mathbb{R}$, $k(x) = \int_{\mathbb{R}} e^{ix\xi} \mu(d\xi)$. The function k is continuous on $\mathbb{R}$. We get finally,

$$K(r) = \int_r^{+\infty} \frac{k(u)}{u} du. \tag{3.54}$$

The proof of Proposition 3.20 is complete. □

Proof of Proposition 3.7. This is just a direct consequence of Theorem 3.3 and equation (3.48). □

3.6 Proofs of some auxiliary lemmas

Lemma 3.23. *Let $F : \mathbb{R}^n \mapsto \mathbb{R}$ be a measurable function. Then, for all bounded Borelian sets $A_1, \ldots, A_n \subset \mathbb{R}$, the following relation holds almost surely:*

$$\mathbb{E}_Y \left[F(M(A), \ldots, M(A_n)) \right] = \mathbb{E}_Y \left[F(\widetilde{M}^N(A), \cdots, \widetilde{M}^N(A_n)) \right]$$

Proof. By using the Jensen inequality, we have

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\frac{1}{T}\widetilde{M}^N[0;T]-\frac{1}{T}M^N[0;T]\Big|\Big]\\
&=\mathbb{E}\Big[\Big(\Big|\frac{1}{T}\widetilde{M}^N[0;T]-\frac{1}{T}M^N[0;T]\Big|^2\Big)^{1/2}\Big]\\
&\leqslant \mathbb{E}\Big[\Big(\mathbb{E}\Big[\Big|\frac{1}{T}\widetilde{M}^N[0;T]-\frac{1}{T}M^N[0;T]\Big|^2|M\Big]\Big)^{1/2}\Big]\\
&=\mathbb{E}\Big[\Big(\frac{1}{T^2}\int_0^T\int_0^T\mathbb{E}\Big[\big(e^{\sum_{n=0}^N X_r^n-\frac{1}{2}\mathbb{E}[(X_r^n)^2]}-1\big)\big(e^{\sum_{n=0}^N X_u^n-\frac{1}{2}\mathbb{E}[(X_u^n)^2]}-1\big)\Big]M^N(dr)M^N(du)\Big)^{1/2}\Big]\\
&=\mathbb{E}\Big[\Big(\frac{1}{T^2}\int_0^T\int_0^T\big(e^{\sum_{n=0}^N \bar{k}_n(r-u)}-1\big)M^N(dr)M^N(du)\Big)^{1/2}\Big]
\end{aligned}$$

The integrand in the above expectation converges almost surely towards 0 because, for each $0\leqslant n\leqslant N$, $\bar{k}_n$ is bounded and converges to 0 in the vicinity of ∞. Furthermore, it is uniformly integrable because

$$\sup_T\mathbb{E}\Big[\Big(\frac{1}{T}M^N([0;T])\Big)^{1+\delta}\Big]<+\infty.$$

We deduce that

$$\mathbb{E}\Big[\Big|\frac{1}{T}\widetilde{M}^N[0;T]-\frac{1}{T}M^N[0;T]\Big|\Big]\to 0\quad\text{as } T\to+\infty.$$

As a consequence, $\frac{1}{T}\widetilde{M}^N[0;T]$ converges almost surely along a subsequence towards Y.

One has, for any function h bounded and continuous,

$$\mathbb{E}\left[F(M(A_1),\dots,M(A_n))h\left(\frac{1}{T}M[0;T]\right)\right]=\mathbb{E}\left[F(\widetilde{M}(A_1),\dots,\widetilde{M}(A_n))h\left(\frac{1}{T}\widetilde{M}^N[0;T]\right)\right]$$

Sending T to $+\infty$ along the subsequence, we get by the bounded convergence theorem

$$\mathbb{E}\left[F(M(A_1),\dots,M(A_n))h(Y)\right]=\mathbb{E}\left[F(\widetilde{M}(A_1),\dots,\widetilde{M}(A_n))h(Y)\right]$$

and the lemma is proved. □

Lemma 3.24. *Let $F:\mathbb{R}_+\to\mathbb{R}$ be some convex function such that*

$$\forall x\in\mathbb{R}_+,\quad |F(x)|\leqslant M(1+|x|^\beta),$$

for some positive constants M,β, and σ be a Radon measure on the Borelian subsets of $\mathbb{R}$. Given $a<b$, let $(X_r)_{a\leqslant r\leqslant b},(Y_r)_{a\leqslant r\leqslant b}$ be two continuous centered Gaussian processes with continuous covariance kernels k_X and k_Y such that

$$\forall u,v\in[a,b],\quad k_X(u,v)\leqslant k_Y(u,v).$$

Then

$$\mathbb{E}\Big[F\Big(\int_a^b e^{X_r-\frac{1}{2}\mathbb{E}[X_r^2]}\sigma(dr)\Big)\Big]\leqslant\mathbb{E}\Big[F\Big(\int_a^b e^{Y_r-\frac{1}{2}\mathbb{E}[Y_r^2]}\sigma(dr)\Big)\Big].$$

Proof. For each $N\in\mathbb{N}$, we define the smooth subdivision $t_p^N=a+p\frac{b-a}{N}$, $p=0,\dots,N$, of the interval $[a,b]$. We also introduce the random variables

$$S_N^X=\sum_{p=0}^{N-1}e^{X_{t_p^N}-\frac{1}{2}\mathbb{E}[X_{t_p^N}^2]}\sigma([t_p^N,t_{p+1}^N))\quad\text{and}\quad S_N^Y=\sum_{p=0}^{N-1}e^{Y_{t_p^N}-\frac{1}{2}\mathbb{E}[Y_{t_p^N}^2]}\sigma([t_p^N,t_{p+1}^N)).$$

By classical Gaussian inequalities (see [129, corollary 6.2] for instance), we have

$$\forall N\geqslant 1,\quad\mathbb{E}\Big[F\Big(S_N^X\Big)\Big]\leqslant\mathbb{E}\Big[F\Big(S_N^Y\Big)\Big].$$

So it just remains to pass to the limit as $N \to \infty$ by using the dominated convergence theorem. By continuity of the processes X, Y the random variables S_N^X, S_N^Y converge almost surely respectively towards $\int_a^b e^{X_r - \frac{1}{2}\mathbb{E}[X_r^2]}\,\sigma(dr)$, $\int_a^b e^{Y_r - \frac{1}{2}\mathbb{E}[Y_r^2]}\,\sigma(dr)$. Clearly, we have:

$$|F(S_N^X)| \leqslant M\big(1 + |S_N^X|^\beta\big),$$

so that we just have to prove that $|S_N^X|^\beta$ is uniformly integrable (the same argument holds for $|S_N^Y|^\beta$). It is enough to establish that for each $d \in \mathbb{N}$,

$$\sup_N \mathbb{E}\big[(S_N^X)^d\big] < +\infty.$$

We have

$$\begin{aligned}
\mathbb{E}\big[(S_N^X)^d\big] =& \mathbb{E}\Big[\Big(\sum_{p=0}^{N-1} e^{X_{t_p^N} - \frac{1}{2}\mathbb{E}[X^2_{t_p^N}]}\sigma([t_p^N, t_{p+1}^N))\Big)^d\Big] \\
=& \sum_{p_1,\dots,p_d=0}^{N-1} \mathbb{E}\Big[e^{X_{t_{p_1}^N} + \cdots + X_{t_{p_d}^N}}\Big] e^{-\frac{1}{2}(\mathbb{E}[X^2_{t_{p_1}^N}] + \cdots + \mathbb{E}[X^2_{t_{p_d}^N}])}\sigma([t_{p_1}^N, t_{p_1+1}^N)) \times \cdots \times \sigma([t_{p_d}^N, t_{p_d+1}^N)) \\
=& \sum_{p_1,\dots,p_d=0}^{N-1} e^{\frac{1}{2}\sum_{i,j=1}^d k_X(t_{p_i}^N, t_{p_j}^N)} e^{-\frac{1}{2}(\mathbb{E}[X^2_{t_{p_1}^N}] + \cdots + \mathbb{E}[X^2_{t_{p_d}^N}])}\sigma([t_{p_1}^N, t_{p_1+1}^N)) \times \cdots \times \sigma([t_{p_d}^N, t_{p_d+1}^N)) \\
\to& \int_a^b \cdots \int_a^b e^{\frac{1}{2}\sum_{i\neq j}^d k_X(u_i,u_j)}\sigma(du_1)\cdots\sigma(du_d)
\end{aligned}$$

as $N \to \infty$. This completes the proof. □

Proof of Lemma 3.21. We choose any $\epsilon < 1$ and consider $|r| \leqslant 1$. Since k_ϵ is continuous at 0, we can find, for $\alpha > 0$, some $\eta > 0$ such that $k_\epsilon(0) - \alpha \leqslant k_\epsilon(u) \leqslant k_\epsilon(0)$ for $|u| \leqslant \eta$. Then we decompose K as

$$\begin{aligned}
K^\epsilon(r) &= \sum_{n=0}^{+\infty} k_\epsilon\big(\frac{r}{\epsilon^n}\big) \\
&= \sum_{n=0}^{\frac{\ln\frac{r}{\eta}}{\ln\epsilon}-1} k_\epsilon\big(\frac{r}{\epsilon^n}\big) + \sum_{n=\frac{\ln\frac{r}{\eta}}{\ln\epsilon}}^{+\infty} k_\epsilon\big(\frac{r}{\epsilon^n}\big) \\
&\overset{def}{=} \sum_{n=0}^{\frac{\ln\frac{r}{\eta}}{\ln\epsilon}-1} k_\epsilon\big(\frac{r}{\epsilon^n}\big) + g_\epsilon(r)
\end{aligned}$$

Let us prove that g_ϵ is bounded over a neighborhood of 0. By using (3.10) and following the computations of (3.28), we have for $p \in \mathbb{N}$:

$$\sum_{n=p}^{+\infty} |k_\epsilon\big(\frac{r}{\epsilon^n}\big)| \leqslant \frac{2C_\epsilon}{-\ln\epsilon}\int_{\frac{r}{\epsilon^{p-1}}}^{+\infty} \theta(u)\ln u\,du.$$

We deduce by taking $p = \frac{\ln\frac{r}{\eta}}{\ln\epsilon}$:

$$|g_\epsilon(r)| \leqslant \frac{2C_\epsilon}{-\ln\epsilon}\int_{\eta\epsilon}^{+\infty} \theta(u)\ln u\,du.$$

Hence g_ϵ is bounded. By noticing that $\frac{r}{\epsilon^n} \leqslant \eta \Leftrightarrow n \leqslant \frac{\ln\frac{r}{\eta}}{\ln\epsilon}$, we deduce

$$\frac{\ln\frac{r}{\eta}}{\ln\epsilon}(k_\epsilon(0) - \alpha) + g_\epsilon(r) \leqslant K^\epsilon(r) \leqslant \frac{\ln\frac{r}{\eta}}{\ln\epsilon}k_\epsilon(0) + g_\epsilon(r).$$

By taking the lim sup and lim inf in the above inequality, we have proved that for each $\alpha > 0$:

$$\frac{k_\epsilon(0) - \alpha}{\ln \frac{1}{\epsilon}} \leqslant \liminf_{r\to 0} \frac{K^\epsilon(r)}{\ln \frac{1}{r}} \leqslant \limsup_{r\to 0} \frac{K^\epsilon(r)}{\ln \frac{1}{r}} \leqslant \frac{k_\epsilon(0)}{\ln \frac{1}{\epsilon}},$$

which completes the proof. □

Proof of Corollary 3.5. By stationarity, it is enough to prove that, almost surely, the measure M does not possess any atom on the segment $[0, 1]$. From [61, Corollary 9.3 VI], it is enough to check that for each $\alpha > 0$:

$$\sum_{k=1}^{n} \mathbb{P}\Big(M[\frac{k-1}{n}; \frac{k}{n}] > \alpha\Big) = n\mathbb{P}\Big(M[\frac{0}{n}; \frac{1}{n}] > \alpha\Big) \to 0 \quad \text{as } n \to \infty.$$

This is a direct consequence of the Markov inequality and Lemma 3.16:

$$n\mathbb{P}\Big(M[\frac{0}{n}; \frac{1}{n}] > \alpha\Big) \leqslant \frac{n}{\alpha^{1+\gamma}} \mathbb{E}[M([0, \frac{1}{n}])^{1+\gamma}] \to 0 \text{ as } n \to \infty. \quad \square$$

Proof of Proposition 3.8. Otherwise, if M is a good lognormal $\star$-scale invariant random measure, then using Theorem 3.3, we know that there exists k a continuous covariance function such that, for all $|r| \leqslant R$:

$$K(r) = \int_{|r|}^{\infty} \frac{k(u)}{u} du = \lambda^2 \ln \frac{T}{|r|} + C. \tag{3.55}$$

By differentiating this equality with respect to r, we obtain $k(r) = \lambda^2$ for all $|r| \leqslant R$. Then, let $(X_t)_{t\in\mathbb{R}}$ be a centered stationary Gaussian process with covariance kernel k. For all $s, t \in \mathbb{R}$ such that $|t - s| < R$, we have $cov(X_t, X_s) = k(|t - s|) = k(0) = var[X_t]$ which implies (by Cauchy-Schwarz inequality) that $X_t = X_s$ almost surely. The process X being stationary, this shows that it is a constant process. Hence $k(r) = \lambda^2$ for all $r \in \mathbb{R}$. Because of equation (3.55), this is a contradiction since it would imply $K(r) = +\infty$ for all r. □

Part III

Contribution in Random matrix theory

Chapter 4

Invariant β-ensembles and the Gauss-Wigner crossover

Résumé

Cet article est publié dans le journal *Physical Review Letters* et est écrit en collaboration avec Jean-Philippe Bouchaud et Alice Guionnet. Nous définissons un nouveau modèle de diffusion matricielle qui converge vers le mouvement Brownien de Dyson avec un paramètre $\beta \in [0, 2]$. En temps long, le modèle limite donne une construction explicite de matrices aléatoires appartenant aux ensembles β qui sont invariantes par conjugaison par des matrices orthogonales ou unitaires. Pour des petites valeurs de β, la densité limite en grande dimension des valeurs propres de ces matrices définit une interpolation continue entre la distribution Gaussienne et le demi-cercle de Wigner. Cette famille de distribution est une famille à un paramètre et les distributions admettent des densités explicitement calculables. Un prolongement de ce calcul nous permet de trouver les corrections de tailles finies (quand la dimension est grande mais pas infinie) dans la convergence vers la loi du demi cercle de Wigner.

Abstract

We define a new diffusive matrix model converging towards the β -Dyson Brownian motion for all $\beta \in [0, 2]$ that provides an explicit construction of β-ensembles of random matrices that is invariant under the orthogonal/unitary group. For small values of β, our process allows one to interpolate smoothly between the Gaussian distribution and the Wigner semi-circle. The interpolating limit distributions form a one parameter family that can be explicitly computed. This also allows us to compute the finite-size corrections to the semi-circle.

Since Wigner's initial intuition that the statistical properties of the eigenvalues of random matrices should provide a good description of the excited states of complex nuclei, Random Matrix Theory has become one of the prominent field of research, at the boundary between atomic physics, solid state physics, statistical mechanics, statistics, probability theory and number theory [3, 20, 15]. It is well known that the joint distribution of the eigenvalues of a large Gaussian random matrix can be expressed as the Boltzmann-Gibbs equilibrium weight of a one-dimensional repulsive Coulomb gas confined in an harmonic well. However, the effective "inverse temperature" β of the system cannot take arbitrary values but is *quantized* (in units of the repulsive Coulomb potential). Depending on the symmetry of the random matrix, only three values are allowed $\beta = 1$ for symmetric real matrices, $\beta = 2$ for Hermitian matrices and $\beta = 4$ for the symplectic ensemble. This is known as Dyson's "threefold way". The existence of matrix ensembles that would lead to other, possibly continuous, values of β, is a very natural question, and the quest for such ensembles probably goes back to Dyson himself. Ten years ago, Dumitriu and Edelman [68] have proposed an explicit construction of tri-diagonal matrices with non-identically distributed elements whose joint law of the eigenvalues is the one of

β-ensembles for general β. Another construction is proposed in [3, p. 426-427] (see also [74]) and uses a bordering procedure to construct recursively a sequence of matrices with eigenvalues distributed as β-ensembles. This construction gives not just the eigenvalue probability density of one matrix of the sequence but also the joint eigenvalue probability density of all matrices. This has lead to a renewed interest for those ensembles, that have connections with many problems, both in physics and in mathematics, see e.g. [72, 15]. The aim of the paper is to provide another construction of β-ensembles that is, at least to our eyes, natural and transparent, and respects by construction the orthogonal/unitary symmetry [149]. Another motivation for our work comes from the recent development of free probability theory. "Freeness" for random matrices is the natural extension of independence for classical random variables. Very intuitively, two real symmetric matrices $\mathbf{A}, \mathbf{B}$ are mutually free in the large N limit if the eigenbasis of $\mathbf{B}$ can be thought of as a random rotation of the eigenbasis of $\mathbf{A}$ (see e.g. [145] for an accessible introduction to freeness and for more rigorous statements). "Free convolution" then allows one to compute the eigenvalue distribution of the sum $\mathbf{A} + \mathbf{B}$ from the eigenvalue distribution of $\mathbf{A}$ and $\mathbf{B}$, much in the same way as convolution allows one to compute the distribution of the sum of two independent random variables. In this context, the Wigner semi-circle distribution appears as the limiting distribution for the sum of a large number of free random matrices, exactly as the Gaussian is the limiting distribution for the sum of a large number of *iid* (independent and identically distributed) random variables. A natural question, from this perspective, is whether one can build a natural framework that interpolates between these two limits.

Let us first recall Dyson's Brownian motion construction of the GOE [69] (for the sake of simplicity, we will only consider here extensions of the $\beta = 1$ ensemble, but similar considerations hold for $\beta = 2$ Hermitian matrices see [10] for full details). It is defined as the real $N \times N$ symmetric matrix process $\mathbf{M}(t)$ solution of the stochastic differential equation (SDE):

$$\mathrm{d}\mathbf{M}(t) = -\frac{1}{2}\mathbf{M}(t)\mathrm{d}t + \mathrm{d}\mathbf{H}(t) \tag{4.1}$$

where $\mathrm{d}\mathbf{H}(t)$ is a symmetric Brownian increment (i.e. a symmetric matrix whose entries above the diagonal are independent Brownian increments with variance $\langle \mathrm{d}\mathbf{H}_{ij}^2(t)\rangle = \frac{1}{2}(1+\delta_{ij})\mathrm{d}t$). Standard second order perturbation theory allows one to write the evolution equation for the eigenvalues λ_i of the matrix $\mathbf{M}(t)$:

$$\mathrm{d}\lambda_i = -\frac{1}{2}\lambda_i \mathrm{d}t + \frac{1}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_j} + \mathrm{d}b_i, \tag{4.2}$$

where $b_i(t)$ are independent standard Brownian motions. This defines Dyson's Coulomb gas model, i.e. "charged" particles on a line, with positions λ_i, interacting via a logarithmic potential, subject to some thermal noise and confined by a harmonic potential. One can deduce from the above equation the Fokker-Planck equation for the joint density $P(\{\lambda_i\}, t)$, for which the stationary joint probability density function (pdf) is readily found to be:

$$P^*(\{\lambda_i\}) = Z\prod_{i<j}|\lambda_i - \lambda_j|^\beta \exp\left[-\frac{1}{2}\sum_i \lambda_i^2\right], \tag{4.3}$$

with $\beta \equiv 1$ and where Z is a normalization factor. The above expression is the well known joint distribution of the eigenvalues of an $N \times N$ random GOE matrix. The Wigner distribution can be recovered either by a careful analysis of the mean marginal univariate distribution $\rho(\lambda) = \int \ldots \int \mathrm{d}\lambda_2 \ldots \mathrm{d}\lambda_N P^*(\lambda = \lambda_1, \lambda_2, \ldots, \lambda_N)$ in the large N limit [114], or by using the above SDE (4.2) to derive a dynamical equation for the Stieltjes transform $G(z,t)$ of $\rho(\lambda, t)$:

$$G(z,t) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i(t) - z}, \quad z \in \mathbb{C}. \tag{4.4}$$

With this scaling, the spectrum is spread out in a region of width of order $\sqrt{N}$ and therefore $z \sim \sqrt{N}$ and $G \sim 1/\sqrt{N}$. Applying Itô's formula to $G(z,t)$ and using (4.2), we obtain the following Burgers equation for G [130]:

$$2\frac{\partial\langle G\rangle}{\partial t} = \frac{\alpha N}{2}\frac{\partial\langle G\rangle^2}{\partial z} + \frac{\partial z\langle G\rangle}{\partial z} + (2-\alpha)\frac{1}{2}\frac{\partial^2\langle G\rangle}{\partial z^2} \tag{4.5}$$

where α is introduced for later convenience, with $\alpha = 1$ for now. Note that we have neglected in Eq. (4.5) a term of order $N^{-5/2}$. Indeed in agreement with [20]: $\langle G^2\rangle - \langle G\rangle^2 \sim N^{-3}$. The neglected term is thus $1/N$ smaller than the diffusion term in Eq. (4.5).

For large N, the last (diffusion) term of Eq. (4.5) is of order $1/N$ smaller than the other ones. To leading order, the stationary solution (where the time derivative is set to 0) can be integrated with respect to z:

$$\frac{1}{2}\alpha N G_\infty^2(z) + zG_\infty(z) = -1\,, \tag{4.6}$$

where the integration constant comes from the boundary condition $G(z) \sim -1/z$ when $z \to \infty$. It is then easy to solve this equation to find the Stieltjes transform that indeed corresponds to the Wigner semi-circle density:

$$\begin{aligned} G_\infty(z) &= \frac{1}{\alpha N}\left[\sqrt{z^2 - 2\alpha N} - z\right] \\ \longrightarrow \rho(\lambda) &= \frac{1}{\pi\alpha N}\sqrt{2\alpha N - \lambda^2}\mathbb{1}_{\{|\lambda| \leqslant \sqrt{2\alpha N}\}}. \end{aligned} \tag{4.7}$$

Now let us turn to the central idea of the present paper. In Dyson's construction, the extra Gaussian slice $\mathrm{d}\mathbf{M}(t)$ that is added to $\mathbf{H}(t)$ is chosen to be independent of $\mathbf{M}(t)$ itself. The eigenbasis of $\mathrm{d}\mathbf{H}(t)$ is a random rotation, taken uniformly over the orthogonal group. As mentioned above, this corresponds to free addition of matrices, and Eq. (4.5) can indeed be derived (for $N = \infty$) using free convolution [145]. If instead we choose to add a random matrix $\mathrm{d}\mathbf{Y}(t)$ that is *always diagonal in the same basis* as that of $\mathbf{M}(t)$, the process becomes trivial. The diagonal elements of $\mathbf{M}(t)$ are all sums of *iid* random variables, and the eigenvalue distribution converges towards the Gaussian. The construction we propose is to alternate randomly the addition of a "free" slice and of a "commuting" slice. More precisely, our model is defined as follows: we divide time into small intervals of length $1/n$ and for each interval $[k/n; (k+1)/n]$, we choose independently Bernoulli random variables $\epsilon_k^n, k \in \mathbb{N}$ such that $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$. Then, setting $\epsilon_t^n = \epsilon_{[nt]}^n$, our diffusive matrix process simply evolves as:

$$\mathrm{d}\mathbf{M}_n(t) = -\frac{1}{2}\mathbf{M}_n(t)\mathrm{d}t + \epsilon_t^n \mathrm{d}\mathbf{H}(t) + (1 - \epsilon_t^n)\mathrm{d}\mathbf{Y}(t) \tag{4.8}$$

where $\mathrm{d}\mathbf{H}(t)$ is a symmetric Brownian increment as above and where $\mathrm{d}\mathbf{Y}(t)$ is a symmetric matrix that is co-diagonalizable with $\mathbf{M}_n(t)$ (i.e. the two matrix have the same eigenvectors) but with a spectrum given by N independent Brownian increments of variance $\mathrm{d}t$. It is clear that the eigenvalues of the matrix $\mathbf{M}_n(t)$ will cross at some points but only in intervals $[k/n; (k+1)/n]$ for which $\epsilon_k^n = 0$ (in the other intervals where they follow Dyson Brownian motion with parameter $\beta = 1$, it is well known that the repulsion is too strong and that collisions are avoided). In such a case, the eigenvalues are re-numbered at time $t = (k+1)/n$ in increasing order.

Now, using again standard perturbation theory, it is easy to derive the evolution of the eigenvalues of $\mathbf{M}_n(t)$ denoted as $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t)$:

$$\mathrm{d}\lambda_i^n = -\frac{1}{2}\lambda_i^n \mathrm{d}t + \frac{\epsilon_t^n}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i^n - \lambda_j^n} + \mathrm{d}b_i \tag{4.9}$$

where the b_i are independent Brownian motions also independent of the $\epsilon_k^n, k \in \mathbb{N}$.

A mathematically rigorous derivation provided in [10] allows one to show that the scaling limits $\lambda_i(t)$, when $n \to \infty$, of the eigenvalues $\lambda_i^n(t)$ obey the following modified Dyson SDE:

$$\mathrm{d}\lambda_i = -\frac{1}{2}\lambda_i \mathrm{d}t + \frac{p}{2}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_j} + \mathrm{d}b_i, \tag{4.10}$$

with the additional ordering constraint $\lambda_1(t) \leqslant \ldots \leqslant \lambda_N(t)$ for all t. One of the difficulty of the proof comes from the fact that when $p < 1$, there is a positive probability for eigenvalues to collide in finite time (the ordering constraint is therefore useful at those points to re-start). The idea is then to show that collisions are in a sense sufficiently rare for the above SDE to

make sense (see [10, 54] for further details). Using the SDE (4.10), one can derive as above the stationary distribution for the joint distribution of eigenvalues, which is still given by Eq. (4.3) but with now $\beta = \alpha = p \leqslant 1$. A very similar construction can be achieved in the GUE case, leading to $\beta = 2p$. As announced, our dynamical procedure, that alternates standard and free addition of random matrices, can lead to any β-ensemble with $\beta \leqslant 2$. The corresponding matrices $\mathbf{M}(t)$ are furthermore *invariant* under the orthogonal (or unitary) group. This is intuitively clear, since both alternatives (adding a free slice or adding a commuting slice) respect this invariance, and lead to a Haar probability measure for the eigenvectors (i.e. uniform over the orthogonal/unitary group). We have also proved that a collision leads to a complete randomization of the eigenvectors within the two-dimensional subspace corresponding to the colliding eigenvalues, see again [10].

It is well known that the eigenvalue density corresponding to the measure P^* given by (4.3) is the Wigner semi-circle for any $\beta > 0$. In fact, using (4.5) with now $\alpha = \beta = p$, one immediately finds that the eigenvalue density is a semi-circle with edges at $\pm\sqrt{2\beta N}$. We simulated numerically the matrix $\mathbf{M}_n(t)$ with $N = 200$ for a very small step $1/n$ and until a large value of t so as to reach the stationary distribution for the eigenvalues. Then we started recording the spectrum and the nearest neighbor spacings (NNS) every 100 steps so as to sample the ensemble. We verified that the spectral density of $\mathbf{M}_n(t = \infty)$ is indeed in very good agreement with the Wigner semi-circle distribution for $\beta = 1/2$. Our sample histogram for the NNS distribution is displayed in Fig.4.1. We also added the corresponding Wigner surmise (which is expected to provide a good approximate description of the NNSD).

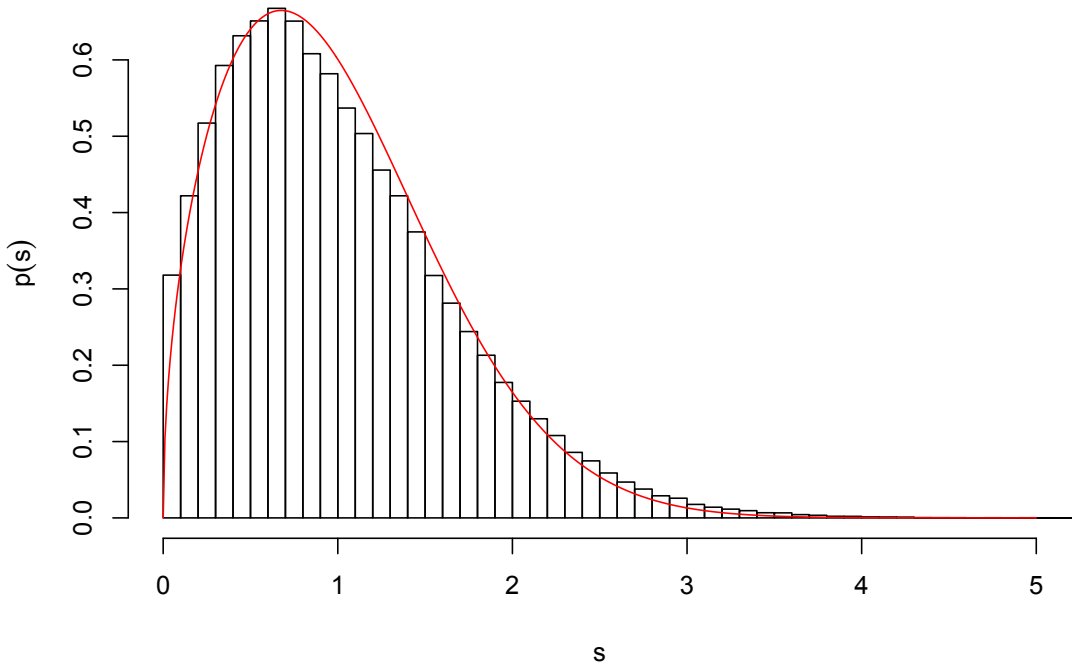


Figure 4.1: Empirical NNSD $P(s)$ for the matrix $\mathbf{M}_n(t = \infty)$ for $\beta = p = 1/2$ with the Wigner surmise (red curve) corresponding to $\beta = \frac{1}{2}$, which behaves as s^β when $s \to 0$.

From the point of view of a cross-over between the standard Gaussian central limit theorem for random variables and the Wigner central limit theorem for random matrices, we see that as soon as the probability p for a non-commuting slice is positive, the asymptotic density is the Wigner semi-circle, with a width of order $\sqrt{pN}$. A continuous cross-over therefore takes place for $p = 2c/N$ with c strictly positive and independent of N. When $c = 0$, $\rho(\lambda)$ is a Gaussian of rms 1, which indeed corresponds to the solution of Eq. (4.5) for $\alpha = 0$. The SDE for the system

$(\lambda_i(t))$ becomes

$$\mathrm{d}\lambda_i = -\frac{1}{2}\lambda_i \mathrm{d}t + \frac{c}{N}\sum_{j\neq i}\frac{\mathrm{d}t}{\lambda_i - \lambda_j} + \mathrm{d}b_i, \tag{4.11}$$

with the additional ordering constraint $\lambda_1(t) \leqslant \ldots \leqslant \lambda_N(t)$ and the stationary joint pdf is still given by (4.3) but with now a vanishing repulsion coefficient $\beta = 2c/N$. In order to elicit the cross-over, we study Eq. (4.5) with $\alpha = 2c/N$. The stationary differential equation corresponding to (4.5) (note this time that all terms are of the same order and the second derivative term is not negligible) can be integrated with respect to z again as:

$$cG^2 + zG + \frac{\mathrm{d}G}{\mathrm{d}z} = -1, \tag{4.12}$$

where the integration constant comes from the boundary condition $G \sim -1/z$ for $z \to \infty$. Note that (4.12) can be recovered directly from the saddle point equation route: under the measure P^* with $\beta = 2c/N$, the energy of a configuration of the λ_i's can be expressed in term of the continuous state density ρ, neglecting terms $\ll 1$, as:

$$\mathcal{E}[\rho] = \frac{1}{2}\int \lambda^2 \rho(\lambda)\mathrm{d}\lambda - c\int\int \ln(|\lambda - \lambda'|)\rho(\lambda)\rho(\lambda')\mathrm{d}\lambda\mathrm{d}\lambda'\,.$$

The probability density P^* therefore rewrites in term of ρ as:

$$P^*[\rho] = Z\exp\left(-N\left[\mathcal{E}[\rho] + \int \rho\ln(\rho)\right]\right)\delta(\int\rho - 1)\,,$$

where the entropy term, which is negligible when $\beta = p$ is of order 1, is now of the same order as the energy term (see [62] for a detailed discussion on the origin of the entropy term). We now need to minimize the quantity $\mathcal{E}[\rho] + \int \rho\ln(\rho)$ with respect to ρ. It is easy to see that the unique minimizer ρ_c satisfies:

$$\int \frac{\lambda\rho_c(\lambda)}{\lambda - z}\mathrm{d}\lambda - 2c\int\int \frac{\rho_c(\lambda)\rho_c(\lambda')}{(\lambda - z)(\lambda - \lambda')}\mathrm{d}\lambda\mathrm{d}\lambda' + \int \frac{\rho_c'(\lambda)}{\lambda - z}\mathrm{d}\lambda + \nu = 0$$

where ν is an integration constant. It is now straightforward to derive (4.12) from this last equation by identifying each term and choosing the constant ν so as to have the correct boundary condition for the Stieltjes transform of a probability measure. As expected physically, the diffusion term in (4.12) corresponds exactly to the entropy contribution to the saddle-point.

Eq. (4.12) was studied in detail by Askey & Wimp [17] and Kerov [92] (see also [31]). Set $G(z) := u'(z)/cu(z)$ to obtain a second order equation on u:

$$u''(z) + zu'(z) + cu(z) = 0\,. \tag{4.13}$$

It follows from the asymptotic behavior of $G(z)$ that, for $|z| \to \infty$,

$$u(z) \sim \frac{A_1}{z^c}\,. \tag{4.14}$$

Eq. (4.13) can in turn be transformed with the change of function $u(z) := e^{-z^2/4}y(z)$ into a Schrodinger equation on $y(z)$:

$$y''(z) + [c - \frac{1}{2} - \frac{1}{4}z^2]y(z) = 0\,. \tag{4.15}$$

The solutions of (4.15) are known (see [83]) to write as $y(z) = A_2 D_{c-1}(z) + A_3 D_{-c}(iz)$ where D_{c-1}, D_{-c} are parabolic cylinder functions and where A_2 and A_3 are two constants. The general solution for u therefore is $u(z) = e^{-z^2/4}(A_2 D_{c-1}(z) + A_3 D_{-c}(iz))$ and the correct asymptotic behavior of u is fulfilled for $A_2 = 0$. Now, one can recover the spectral density $\rho_c(\lambda)$ associated

to G by the classical inversion formula and various elegant tricks [106]. The final result for $\rho(\lambda)$ reads, for all $c > 0$:

$$\rho_c(\lambda) = \frac{1}{\sqrt{2\pi}\Gamma(1+c)} \frac{1}{|D_{-c}(i\lambda)|^2}; \tag{4.16}$$
$$D_{-c}(z) = \frac{e^{-z^2/4}}{\Gamma(c)} \int_0^\infty \mathrm{d}x e^{-zx-\frac{x^2}{2}} x^{c-1}.$$

Expression (4.16) was again checked with numerical simulations with very good agreement. The integral representation for $D_{-c}(z)$ does not hold for $c = 0$, but the function $D_{-c}(iu)$ is still well defined for all $c \in (-1; 0]$ (see [17]). It is easy to check that $\rho_0(u) = e^{-u^2/2}/\sqrt{2\pi}$ when $c = 0$, as expected. When $c \to \infty$, the Wigner semi-circle law is recovered

$$\rho_c(u) \approx \frac{1}{2\pi c}\sqrt{4c - u^2}\,. \tag{4.17}$$

Standard results [83] on D_{-c} enable to find the tails of ρ_c:

$$\rho_c(u) \sim u^{2c} e^{-u^2/2} \quad (|u| \to \infty). \tag{4.18}$$

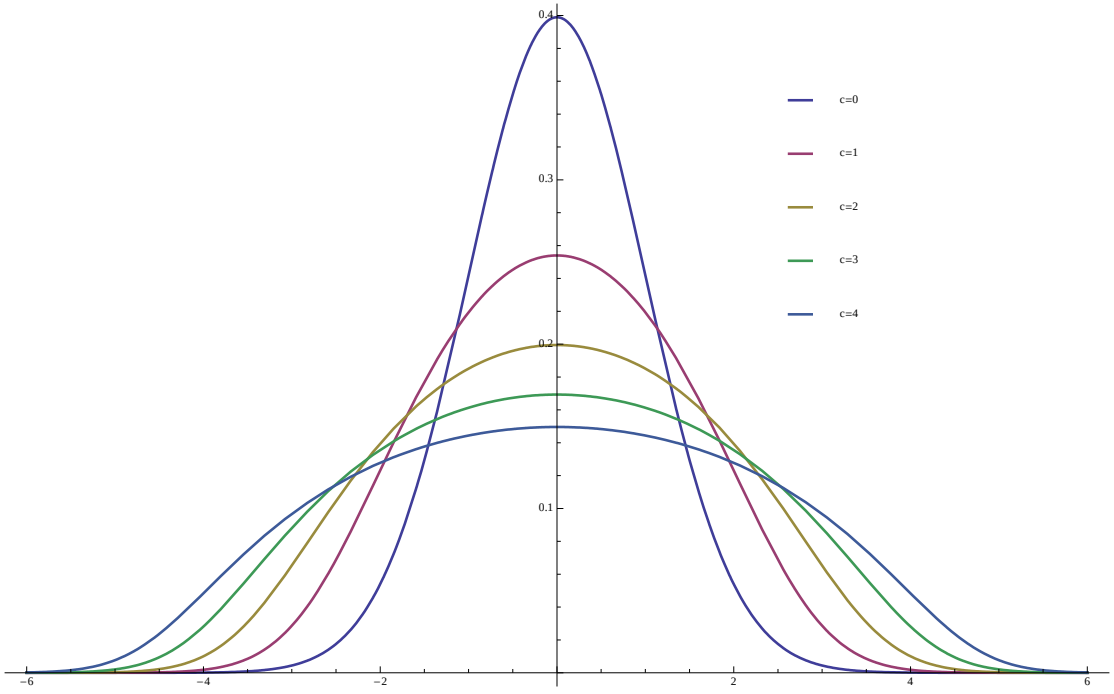


Figure 4.2: Density $\rho_c(u)$ for $c = 0, 1, 2, 3, 4$, showing the progressive deformation of the Gaussian towards Wigner's semi-circle.

Let us return to (4.5) for $\beta = \alpha \in (0; 2)$. Interestingly, our method allows us to compute the correction to the Wigner semicircle inside the support of the spectral density for large but finite N due to the last diffusion term, which is usually neglected. Indeed one can solve as above the stationary equation of (4.5) keeping every term. This leads to the following *corrected spectral density*, valid for large but finite N:

$$\rho(\lambda) = \frac{\sqrt{\alpha}}{\sqrt{2\pi}\Gamma(1+c)} \frac{1}{|D_{-c}(i\sqrt{\alpha}\lambda)|^2}\,, \tag{4.19}$$

where $\alpha = 2/(2-\beta)$ and $c = \beta N/(2-\beta)$. Note that this correction is valid only inside the spectrum and *does not* describe the edge scaling behavior nor the Tracy-Widom tails.

The above discussion can also be formally extended to $-1 \leqslant c < 0$, corresponding to a weakly attracting Coulomb gas (also mentioned in [149]; see also [77] for an application). We conjecture that the stationary density for large system is again given by the above Askey-Wimp-Kerov

distributions ρ_c but for the parameter range $c \in (-1; 0]$. For $c = -1$, the stationary density ρ_{-1} is a Dirac mass at 0. Beyond this level, the attraction is too strong and the gas completely collapses on itself.

As a conclusion, we have provided here the first explicit construction of invariant β-ensembles of random matrices, for arbitrary $\beta \leqslant 2$. The stationary distribution for the eigenvectors is the Haar probability measure on the orthogonal group if $0 < \beta \leqslant 1$, respectively unitary group if $1 < \beta \leqslant 2$. We have found a natural scaling limit that allows one to interpolate smoothly between the Gaussian distribution, relevant for sums of independent random variables, and the Wigner semi-circle distribution, relevant for sums of free random matrices. The interpolating limit distributions form a one parameter family that can be explicitly computed. The statistics of the largest eigenvalue is also very interesting (and now well known for $\beta > 0$, see [67, 41, 42, 73]): one should be able to interpolate smoothly, as a function of c, between the well-known Gumbel distribution of extreme value statistics and the Tracy-Widom(β) distributions. Whether this can be mapped into a generalized KPZ/Directed polymer problem remains to be seen.

Acknowledgements

We are grateful to S. N. Majumdar for many important suggestions. We thank C. Garban, R. Rhodes, V. Vargas, F. Benaych-Georges, L. Dumaz, R. Chicheportiche, G. Schehr and P. P. Vivo for useful comments and discussions. A. G. and J.-P. B. thank the organizers of the NYU-Abu-Dhabi conference in January 2011, where this work was initiated (on way to buying dates). A. G. and R. A. acknowledge financial support of ANR GranMa ANR-08-BLAN-0311-01.

Chapter 5

A diffusive matrix model for invariant β-ensembles

Résumé

Cet article est soumis dans le journal *Electronic Journal of probability* et est écrit en collaboration avec Alice Guionnet. Nous définissons un nouveau modèle de diffusion matricielle qui converge vers le mouvement Brownien de Dyson avec un paramètre β quelconque appartenant à l'intervalle $[0, 2]$. En temps long, le modèle limite donne une construction explicite de matrices aléatoires appartenant aux ensembles β qui sont invariantes par conjugaison par des matrices orthogonales ou unitaires. Nous décrivons aussi la dynamique des vecteurs propres du processus matriciel limite; nous montrons que lorsque $\beta < 1$ et que deux valeurs propres collisionnent, les vecteurs propres associés à ces deux valeurs propres fluctuent très fortement et finissent par être distribués uniformément sur le sous-espace supplémentaire orthogonal du sous espace engendré par les vecteurs propres associés aux autres valeurs propres.

Abstract

We define a new diffusive matrix model converging towards the β-Dyson Brownian motion for all $\beta \in [0, 2]$ that provides an explicit construction of β-ensembles of random matrices that is invariant under the orthogonal/unitary group. We also describe the eigenvector dynamics of the limiting matrix process; we show that when $\beta < 1$ and that two eigenvalues collide, the eigenvectors of these two colliding eigenvalues fluctuate very fast and take the uniform measure on the orthocomplement of the eigenvectors of the remaining eigenvalues.

5.1 Introduction

It is well known that the law of the eigenvalues of the classical Gaussian matrix ensembles are given by a Gibbs measure of a Coulomb gas interaction with inverse temperature $\beta = 1$ (resp. 2, resp. 4) in the symmetric (resp. Hermitian, resp. symplectic) cases;

$$dP_\beta(\lambda) = \frac{1}{Z^\beta} \prod_{i<j} |\lambda_i - \lambda_j|^\beta e^{-\frac{1}{2}\sum \lambda_i^2} \prod d\lambda_i \,.$$

Such measures are associated with symmetric Langevin dynamics, the so-called Dyson Brownian motion, which describe the random motion of the eigenvalues of a symmetric (resp. Hermitian, resp. symplectic) Brownian motion. They are given by the stochastic differential system

$$d\lambda_i(t) = \sqrt{2} db_i(t) - \lambda_i(t)dt + \beta \sum_{j \neq i} \frac{1}{\lambda_i(t) - \lambda_j(t)} dt \tag{5.1}$$

with iid Brownian motions (b_i). These laws and dynamics have been intensively studied, and both local and global behaviours of these eigenvalues have been analyzed precisely, starting from the reference book of Mehta [114].

More recently, the generalization of these distributions and dynamics to all $\beta \geq 0$, the so-called β-ensembles, was considered. As for $\beta = 1, 2, 4$, the Langevin dynamics converge to their unique invariant Gibbs measure P_β as times goes to infinity. Indeed, the stochastic differential system under study is a set of Brownian motions in interaction according to a strictly convex potential. Thus, one can then show by a standard coupling argument that two solutions driven by the same Brownian motion but with different initial data will soon be very close to each others. This entails the uniqueness of the invariant measure as well as the convergence to this Gibbs measure. It turns out that the case $\beta \in [0, 1)$ and the case $\beta \in [1, \infty)$ are quite different, as in the first case the eigenvalues process can cross whereas in the second the repulsion is strong enough so that the eigenvalues do not collide with probability one in finite time. However, the diffusion was shown to be well defined, even for $\beta < 1$, by Cépa and Lépingle [54], at list once reordered.

The goal of this article is to provide a natural interpretation of β-ensembles in terms of random matrices for $\beta \in [0, 2]$. Dumitriu and Edelman [68] already proposed a tridiagonal matrix with eigenvalues distributed according to the β-ensembles. However, this tridiagonal matrix lacks the invariant property of the classical ensembles. Our construction has this property and moreover is constructive as it is based on a dynamical scheme. It was proposed by JP Bouchaud, and this article provides rigorous proofs of the results stated in [**?**]. The idea is to interpolate between the Dyson Brownian motion and the standard Brownian motion by throwing a coin at every infinitesimal time step to decide whether our matrix will evolve according to a Hermitian Brownian motion (with probability p) or will keep the same eigenvectors but has eigenvalues diffusing according to a Brownian motion. When the size of the infinitesimal time steps goes to zero, we will prove that the dynamics of the eigenvalues of this matrix valued process converges towards the β-Dyson Brownian motion with $\beta = 2p$. The same construction with a symmetric Brownian motion leads to the same limit with $\beta = p$. This result is more precisely stated in Theorem 5.2. We shall not consider the extension to the symplectic Brownian motion in this paper, but it is clear that the same result holds with $\beta = 4p$. Our construction can be extended to other matrix models such as Wishart matrices, Circular and Ginibre Gaussian Ensembles and will lead to similar results.

We thus deduce from our construction that β-ensembles can be interpreted as an interpolation between free convolution (obtained by adding a Hermitian Brownian motion) and standard convolution (arising when the eigenvalues evolve following standard Brownian motions). It is natural to wonder whether a notion of β-convolution could be more generally defined.

Moreover we shall study the eigenvectors of our matrix-valued process. In the case where $\beta \geq 1$, their dynamics is well known and is similar to the dynamics of the eigenvectors of the Hermitian or Symplectic Brownian motions, see e.g. [15]. When $\beta < 1$ the question is to determine what happens at a collision. It turns out that when we approach a collision, the eigenvectors of the non-colliding eigenvalues converge to some orthogonal family B of $d - 2$ vectors whereas the eigenvectors of the colliding eigenvalues oscillate very fast and take the uniform distribution on the ortho-complement of B, see Proposition 5.6.

5.2 Statement of the results

Let $\mathcal{H}_d^\beta$ be the space of $d \times d$ symmetric (respectively Hermitian) matrices if $\beta = 1$ (resp. $\beta = 2$) and $\mathcal{O}_d^\beta$ be the space of $d \times d$ orthogonal (respectively unitary) matrices if $\beta = 1$ (resp. $\beta = 2$).

We consider the matrix-valued process defined as follows. Let γ be a positive real number and $M_0^\beta \in \mathcal{H}_d^\beta$ with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_d$. For each $n \in \mathbb{N}$, we let $(\epsilon_k^n)_{k \in \mathbb{N}}$ be a sequence of i.i.d $\{0, 1\}$-valued Bernoulli variables with mean p in the sense that

$$\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]\,.$$

Furthermore, for $t \geqslant 0$, we set $\epsilon_t^n := \epsilon_{[nt]}^n$.

In the following, the process $(H^\beta(t))_{t \geqslant 0}$ will denote a symmetric Brownian motion, i.e. a process with values in the set of $d \times d$ symmetric matrices (respectively Hermitian if $\beta = 2$) with entries $H_{ij}^\beta(t), t \geqslant 0, i \leqslant j$ constructed via independent real valued Brownian motions

$(B_{ij}, \widetilde{B}_{ij}, 1 \leqslant i \leqslant j \leqslant d)$ by

$$H_{ij}^{\beta}(t) = \begin{cases} B_{ij}(t) + i(\beta-1)\widetilde{B}_{ij}(t) & \text{if } i < j \\ \sqrt{2}\, B_{ii}(t) & \text{otherwise} \end{cases} \tag{5.2}$$

Definition 5.1. *For each $n \in \mathbb{N}$, we define a diffusive matrix process $(M_n^{\beta}(t))_{t \geqslant 0}$ such that $M_n^{\beta}(0) := M_0^{\beta}$ and for $t \geqslant 0$*

$$dM_n^{\beta}(t) = -\gamma M_n^{\beta}(t)dt + \epsilon_t^n dH_t^{\beta} + (1-\epsilon_t^n)dY_t \tag{5.3}$$

where $(H_t^{\beta})_{t \geqslant 0}$ is a $d \times d$ symmetric (resp. Hermitian) as defined in (5.2) whereas

$$dY_t = \sqrt{2}\sum_{i=1}^{d} \chi_i^n\left(\frac{[nt]}{n}\right) dB_t^i$$

with i.i.d Brownian motions $(B_t^i)_{t \geqslant 0}$ and where $\chi_i^n([nt]/n)$ is the spectral projector associated to the i-th eigenvalue $\lambda_i([nt]/n)$ of the matrix $M_n^{\beta}([nt]/n)$ if the eigenvalues are numbered as $\lambda_1([nt]/n) < \lambda_2([nt]/n) < \cdots < \lambda_d([nt]/n)$ (we shall see that the above is possible as the eigenvalues are almost surely distinct at the given times $\{k/n, k \in \mathbb{N}\}$).

As for all t, the matrix $M_n^{\beta}(t)$ is in the space $\mathcal{H}_d^{\beta}$, we know that it can be decomposed as

$$M_n^{\beta}(t) = O_n^{\beta}(t)\Delta_n^{\beta}(t)O_n^{\beta}(t)^*$$

where $\Delta_n^{\beta}(t)$ is the diagonal matrix whose diagonal is the vector of the ordered eigenvalues of $M_n^{\beta}(t)$ and where $O_n^{\beta}(t)$ is in the space $\mathcal{O}_d^{\beta}$ for all $t \in \mathbb{R}_+$. We also introduce a matrix $O^{\beta}(0)$ to be the initial orthogonal matrix (resp. unitary if $\beta = 2$) such that $M_0^{\beta}(t) = O^{\beta}(0)\Delta_0 O^{\beta}(0)^*$ where $\Delta_0 := \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$.

The evolution of the eigenvalues of $M_n^{\beta}(t)$ during the time interval $[k/n; (k+1)/n]$ is given by independent Brownian motions if $\epsilon_k^n = 0$ and by Dyson Brownian motions if $\epsilon_k^n = 1$.

The eigenvectors of $M_n^{\beta}(t)$ do not evolve on intervals $[k/n; (k+1)/n]$ such that $\epsilon_k^n = 0$ and evolve with the classical diffusion of the eigenvectors of Dyson Brownian motions if $\epsilon_k^n = 1$ (see [15] for a review on Dyson Brownian motion).

Our main theorems describe the asymptotic properties of the ordered eigenvalues of the matrix $M_n^{\beta}(t)$ denoted in the following as

$$(\lambda_1^n(t) \leqslant \lambda_2^n(t) \leqslant \cdots \leqslant \lambda_d^n(t)) \tag{5.4}$$

and also those of the matrix $O_n^{\beta}(t)$ defined above, as n goes to infinity.

Let $(b_t^i)_{t \geqslant 0}, i \in \{1, \ldots, d\}$ be a family of independent Brownian motions on $\mathbb{R}$. Recall that Cépa and Lépingle showed in [54] the uniqueness and existence of the strong solution to the stochastic differential system

$$d\lambda_i(t) = -\gamma\lambda_i(t)dt + \sqrt{2}db_t^i + \beta p \sum_{j \neq i} \frac{1}{\lambda_i(t) - \lambda_j(t)}dt \tag{5.5}$$

starting from $\lambda(0) = (\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_d)$ and such that for all $t \geqslant 0$

$$\lambda_1(t) \leq \lambda_2(t) \leq \cdots \leq \lambda_d(t) \quad a.s. \tag{5.6}$$

For the scaling limit of the *ordered* eigenvalues, we shall prove that

Theorem 5.2. *Let M_0^{β} be a symmetric (resp. Hermitian) matrix if $\beta = 1$ (resp. $\beta = 2$) with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_d$ and $(M_n^{\beta}(t))_{t \geq 0}$ be the matrix process defined in Definition 5.1. Let $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_d^n(t)$ be the ordered eigenvalues of the matrix $M_n^{\beta}(t)$. Let also $(\lambda_1(t), \ldots, \lambda_d(t))_{t \geqslant 0}$ be the unique strong solution of (5.5) with initial conditions in $t = 0$ given by $(\lambda_1, \lambda_2, \ldots, \lambda_d)$.*

Then, for any $T < \infty$, the process $(\lambda_1^n(t), \ldots, \lambda_d^n(t))_{t \in [0,T]}$ converges in law as n goes to infinity towards the process $(\lambda_1(t), \ldots, \lambda_d(t))_{t \in [0,T]}$ in the space of continuous functions $\mathcal{C}([0,T], \mathbb{R}^d)$ embedded with the uniform topology.

In the case where $\beta p \geqslant 1$, the eigenvalues almost never collide and we will see (see section 5.6.1) in this case that it is easy to construct a coupling of λ and λ^n so that λ^n almost surely converges towards λ.

We shall also describe the scaling limit of the matrix $O_n^\beta(t)$ (the columns of $O_n^\beta(t)$ are the eigenvectors of $M_n^\beta(t)$) when n tends to infinity, at least until the first collision time for the eigenvalues, i.e. until the time T_1 defined as $T_1 := \inf\{t \geqslant 0 : \exists i \in \{2, \dots, d\}, \lambda_i(t) = \lambda_{i-1}(t)\}$.

Let $w_{ij}^\beta(t), 1 \leqslant i < j \leqslant d$ be a family of real or complex (whether $\beta = 1$ or 2) standard Brownian motions (i.e. $w_{ij}^\beta(t) = B_{ij}^1(t) + \sqrt{-1}\,(\beta - 1)B_{ij}^2(t)$ where the B_{ij}^1, B_{ij}^2 are standard Brownian motions on $\mathbb{R}$), independent of the family of Brownian motions $(b_t^i)_{t \geqslant 0}, i \in \{1, \dots, d\}$. For $i < j$, set in addition $w_{ji}^\beta(t) := \bar{w}_{ij}^\beta(t)$ and define the skew Hermitian matrix (i.e. such that $R^\beta = -(R^\beta)^*$) by setting for $i \neq j$,

$$dR_{ij}^\beta(t) = \frac{dw_{ij}^\beta(t)}{\lambda_i(t) - \lambda_j(t)}, \quad R_{ij}^\beta(0) = 0\,.$$

Then, with $\lambda_i(t), 0 \leqslant t \leqslant T_1, i \in \{1, \dots, d\}$ being the solution of (5.5) until its first collision time, there exists a unique strong solution $(O^\beta(t))_{0 \leqslant t \leqslant T_1}$ to the stochastic differential equation

$$dO^\beta(t) = \sqrt{p} O^\beta(t) dR^\beta(t) - \frac{p}{2} O^\beta(t) d\langle (R^\beta)^*, R^\beta \rangle_t \tag{5.7}$$

This solution exists and is unique since it is a linear equation in O^β and R^β is a well defined martingale at least until time T_1. It can be shown as in [15, Lemma 4.3.4] that $O^\beta(t)$ is indeed an orthogonal (resp. unitary if $\beta = 2$) matrix for all $t \in [0; T_1]$.

We mention at this point that the matrix $O_n^\beta(t)$ is not uniquely defined, *even* when we impose the diagonal matrix to have a non-decreasing diagonal $\lambda_1^n(t) \leqslant \dots \leqslant \lambda_n(t)$. Indeed, the matrix $O_n^\beta(t)$ can be replaced, for example, by $-O_n^\beta(t)$ (other possible matrices exist). The following proposition overcomes this difficulty.

Define $T_n(1)$ to be the first collision time of the process $(\lambda_1^n(t), \dots, \lambda_d^n(t))$.

Proposition 5.3. *There exists a continuous process $(O_n^\beta(t))_{0 \leqslant t \leqslant T_1}$ in $\mathcal{O}_d^\beta$ with a uniquely defined law and such that for each $t \in [0; T_n(1)]$, we have*

$$O_n^\beta(t) \Delta_n^\beta(t) O_n^\beta(t)^* \overset{law}{=} M_n^\beta(t)\,,$$

where $\Delta_n^\beta(t)$ is the diagonal matrix of the ordered (as in (5.4)) eigenvalues of $M_n^\beta(t)$.

Proposition 5.3 is proved in Section 5.7. We are now ready to state our main result for the convergence in law of the matrix $O_n^\beta(t)$.

Theorem 5.4. *Let η and T be positive real numbers. Then, conditionally on the sigma-algebra generated by $(\lambda_1^n(s), \dots, \lambda_d^n(s))$, $0 \leqslant s \leqslant T_1 \wedge T$, the matrix process $(O_n^\beta(t))_{0 \leqslant t \leqslant (T_1 - \eta) \wedge T}$ introduced in Proposition 5.3 converges in law in the space of continuous functions $\mathcal{C}([0; T], \mathcal{O}_d^\beta)$ towards the unique solution of the stochastic differential equation (5.7).*

Theorem 5.4 gives a convergence result as n goes to infinity for the eigenvectors of the matrix process $(M_n^\beta(t))$ but only until the first collision time T_1. If $p\beta \geqslant 1$, the result is complete as one can show (see [15] and section 5.6.1) that the process $(\lambda_1(t), \dots, \lambda_d(t))$ is a non colliding process (i.e. almost surely $T_1 = \infty$). However, if $p\beta < 1$, it would be interesting to have a convergence on all compact sets $[0; T]$ even after collisions occurred. Our next results describe the behavior of the columns of the matrix $O^\beta(t)$ denoted as $(\phi_1(t), \dots, \phi_d(t))$ when $t \to T_1$ with $t < T_1$.

We first need to describe the behavior of the eigenvalues $(\lambda_1(t), \dots, \lambda_d(t))$ in the left vicinity of T_1.

Proposition 5.5. *If $p\beta < 1$ then almost surely $T_1 < \infty$ and there exists a unique index $i^* \in \{2, \dots, d\}$ such that $\lambda_{i^*}(T_1) = \lambda_{i^*-1}(T_1)$. While we have, for all $t \geqslant 0$ and almost surely,*

$$\int_0^t \frac{ds}{(\lambda_{i^*} - \lambda_{i^*-1})(s)} < +\infty\,,$$

the following divergence occurs almost surely

$$\int_0^{T_1} \frac{ds}{(\lambda_{i^*} - \lambda_{i^*-1})^2(s)} = +\infty\,. \tag{5.8}$$

The first part of Proposition 5.5 is proved in subsections 5.3.1 and 5.3.2, the last statement is proved in 5.7. Hence equality (5.8) implies the existence of diverging integrals in the SDE (5.7). Because of this singularity, we will show

Proposition 5.6. *Conditionally on $(\lambda_1(t), \dots, \lambda_d(t)), 0 \leqslant t \leqslant T_1$, we have:*

1. *For all $j \neq i^*, i^* - 1$, the eigenvector $\phi_j(t)$ for the eigenvalue $\lambda_j(t)$ converges almost surely to a vector denoted $\widetilde{\phi}_j$ as t grows to T_1. The family $\{\widetilde{\phi}_j, j \neq i^*, i^* - 1\}$ is an orthonormal family of $\mathbb{R}^d$ (respectively $\mathbb{C}^d$) if $\beta = 1$ (resp. $\beta = 2$). We denote by V the corresponding generated subspace and by W its two dimensional orthogonal complementary in $\mathbb{R}^d$ (resp. $\mathbb{C}^d$).*

2. *The family $\{\phi_{i^*}(t), \phi_{i^*-1}(t)\}$ converges weakly to the uniform law on the orthonormal basis of W as t grows to T_1.*

The paper is organized as follows. In Section 5.3, we review and establish some new properties for the limiting eigenvalues process $(\lambda_1(t), \dots, \lambda_d(t))$ defined in 5.5 that will be useful later in our proof of Theorems 5.2 and 5.4. We also introduce, in subsection 5.3.4, a process with fewer collisions that approximates the limiting eigenvalue process. In fact this gives a new construction of the limiting eigenvalues process already constructed in [54], perhaps simpler and more intuitive using only standard Itô's calculus. We give some useful estimates on the processes of eigenvalues and matrix entries of M_n^β in Section 5.4. In Section 5.5, we prove the almost sure convergence of the process $(\lambda_1^n, \dots, \lambda_d^n)$ to the limiting eigenvalues process $(\lambda_1, \dots, \lambda_d)$ until the first hitting time of two particles with a coupling argument. In Section 5.6, we finish the proof of Theorem 5.2 by approximating in the same way the process $(\lambda_1^n, \dots, \lambda_d^n)$ with the same idea of separating the particles which collide by a distance $\delta > 0$. At this point, it suffices to apply that the result of Section 5.5 to show that the two approximating processes are close in the large n limit. In Section 5.7, we prove Theorem 5.4, the last statement of Proposition 5.5 and Propositions 5.3 and 5.6.

5.3 Properties of the limiting eigenvalues process

In this section we shall study the unique strong solution of (5.5) introduced by Cépa and Lépingle in [54]. We first derive some boundedness and smoothness properties. In view of proving the convergence of λ^n towards this process, and in particular to deal with possible collisions, we construct it for $p\beta < 1$ as the limit of a process which is defined similarly except when two particles hit, when we separate them by a (small) positive distance, see Definition 5.12.

5.3.1 Regularity properties of the limiting process

Lemma 5.7. *Let $\lambda = (\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_d)$. Then there exists a unique strong solution of (5.5). Moreover, it satisfies*

- *For all $T < \infty$, there exists $\alpha, M_0 > 0$ finite so that for $M \geqslant M_0$*

$$\mathbb{P}\left[\max_{1\leq i\leq d} \sup_{0 \leqslant t \leqslant T} |\lambda_i(t)| \geqslant M\right] \leqslant e^{-\alpha(M-M_0)^2}\,. \tag{5.9}$$

- *For all $T < \infty$, all $i, j \in \{1, \dots, d\}$, $i \neq j$,*

$$\mathbb{E}\left[\int_0^T \frac{ds}{|\lambda_i(s) - \lambda_j(s)|}\right] < \infty\,.$$

Furthermore, there exists $\alpha, M_0 > 0$ finite so that for $M \geqslant M_0$ and $i \neq j$, we have

$$\mathbb{P}\left[\int_0^T \frac{ds}{|\lambda_i(s) - \lambda_j(s)|} \geqslant M\right] \leqslant e^{-\alpha(M-M_0)^2}\,.$$

Proof. The existence and unicity of the strong solution is [54, Proposition 3.2].

For the first point, we choose a twice continuously differentiable symmetric function ϕ, increasing on $\mathbb{R}^+$, which approximates smoothly $|x|$ in the neighborhood of the origin so that $\phi(0) = 0$, $x\phi'(x) \geq 0$, $|\phi'(x)| \leq c$ and $|\phi''(x)| \leq c$, whereas $|\phi(x)| \geq |x| \times |x| \wedge 1$ (take e.g $\phi(x) = x^2(1+x^2)^{-1/2}$) to obtain by Itô's Lemma

$$d(\phi(\lambda_i(t))) = -\gamma\lambda_i(t)\phi'(\lambda_i(t))dt + \sqrt{2}\phi'(\lambda_i(t))db_t^i + p\beta\sum_{j\neq i}\phi'(\lambda_i(t))\frac{dt}{\lambda_i(t)-\lambda_j(t)} + \phi''(\lambda_i(t))dt.$$

For all t, we have $\lambda_i(t)\phi'(\lambda_i(t)) \geqslant 0$, and also

$$\sum_{i=1}^{d}\sum_{j\neq i}\frac{\phi'(\lambda_i(t))}{\lambda_i(t)-\lambda_j(t)} = \frac{1}{2}\sum_{i=1}^{d}\sum_{j\neq i}\frac{\phi'(\lambda_i(t))-\phi'(\lambda_j(t))}{\lambda_i(t)-\lambda_j(t)} \leqslant \frac{d(d-1)}{2}\,||\,\phi''\,||_\infty\,.$$

We deduce from the above arguments that there exists $C > 0$ such that

$$\sum_{i=1}^{d}\phi(\lambda_i(t)) \leqslant \sqrt{2}\sum_{i=1}^{d}\int_0^t \phi'(\lambda_i(s))db_s^i + Ct + \sum_{i=1}^{d}\phi(\lambda_i)\,.$$

By usual martingales inequality, as ϕ' is uniformly bounded we know that, see e.g. [15, Corollary H.13],

$$P\left[\sup_{0\leq t\leq T}|\sum_{i=1}^{d}\int_0^t \phi'(\lambda_i(t))db_i(t)\,| \geq M\right] \leqslant \exp(-\frac{M^2}{2cT})$$

and therefore using the fact that $|\phi(x)| \geq |x| \times |x| \wedge 1$, we deduce the first point with $M_0 = |\sum_{i=1}^{d}\phi(\lambda_i)| + CT$ and $\alpha = 1/2CT$.

For the second point, we first remark as in the proof of [54, Lemma 3.5] that for all $i < d$

$$\begin{aligned} p\beta\int_0^T \frac{dt}{|\,\lambda_d(t)-\lambda_i(t)\,|} &\leqslant p\beta\sum_{j<d}\int_0^T \frac{dt}{|\,\lambda_d(t)-\lambda_j(t)\,|} \\ &= p\beta\sum_{j<d}\int_0^T \frac{dt}{\lambda_d(t)-\lambda_j(t)} \\ &= \lambda_d(T)-\lambda_d(0)-\sqrt{2}b_T^d + \gamma\int_0^T\lambda_d(t)dt\,. \end{aligned}$$

so that the first point gives the claim fo $j = d$. We then continue recursively. □

5.3.2 Estimates on collisions

To obtain regularity estimates on the process λ, we need to control the probability that more than two particles are close together. We shall prove, building on an idea from Cépa and Lépingle [55], that

Lemma 5.8. *For $r \geq 3$ and $I \subset \{1,\dots,d\}$ with $|I| = r$, set*

$$S_t^I = \sum_{i,j\in I}(\lambda_i(t)-\lambda_j(t))^2\,.$$

We let, for $\varepsilon > 0$,

$$\tau_\varepsilon^r := \inf\{t \geq 0 : \min_{|I|=r} S_t^I \leq \varepsilon\}$$

Then, for any $T > 0$ and $\eta > 0$, for any $r \geq 3$ there exists $\varepsilon_r > 0$ which only depends on $\{S_0^I, |I| \geq 3\}$ so that

$$\mathbb{P}\left(\tau_{\varepsilon_r}^r \leq T\right) \leq \eta\,.$$

Proof. The proof is done by induction over r and we start with the case $r = d$, $I = \{1, \dots, d\}$. Then, S verifies the following SDE (see e.g. [55, Theorem 1]):

$$dS_t = -2\gamma S_t dt + 4\sqrt{d}\sqrt{S_t}d\beta_t + adt$$

where β_t is a a standard brownian motion and $a = 2d(d-1)(2+p\beta d)$. The square root of $\rho_t := \sqrt{S_t}$ verifies the SDE

$$d\rho_t = -\gamma\rho_t dt + 2\sqrt{d}\, d\beta_t + (\frac{a}{2} - 2d)\frac{dt}{\rho_t}.$$

In particular, one can check that, if $\alpha = 2 - \frac{a}{4d} = 2 - (d-1)(1 + p\beta d/2)$

$$d\rho_t^\alpha = -\alpha\gamma\rho_t^\alpha dt + 2\sqrt{d}\,\alpha\rho_t^{\alpha-1}d\beta_t.$$

Thus, as $\alpha < 0$ for $d \geq 3$, for any $\varepsilon > 0$, $\rho^{\alpha-1}_{t\wedge\tau_\varepsilon^d}$ is bounded so that $\int_0^{\cdot} \rho^{\alpha-1}_{s\wedge\tau_\varepsilon^d}d\beta_s$ is a martingale and therefore

$$\mathbb{E}[\rho^\alpha_{T\wedge\tau_\varepsilon^d}] \leq \rho_0^\alpha - \alpha\gamma\int_0^T \mathbb{E}[\rho^\alpha_{t\wedge\tau_\varepsilon^d}]dt$$

By Gronwall's lemma, since $\sup_t \mathbb{E}[\rho^\alpha_{t\wedge\tau_\varepsilon^d}]$ is finite, we deduce that

$$\mathbb{E}[\rho^\alpha_{T\wedge\tau_\varepsilon^d}] \leq \rho_0^\alpha(1 - \frac{1}{\alpha\gamma})e^{-\alpha\gamma T} + \frac{\rho_0^\alpha}{\alpha\gamma}.$$

As a consequence, since $\alpha < 0$, we have

$$\varepsilon^{\alpha/2}\mathbb{P}(\tau_\varepsilon^d \leq T) \leqslant \mathbb{E}[S^{\alpha/2}_{T\wedge\tau_\varepsilon^d}] = \mathbb{E}[\rho^\alpha_{T\wedge\tau_\varepsilon^d}] \leqslant \rho_0^\alpha(1 - \frac{1}{\alpha\gamma})e^{-\alpha\gamma T} + \frac{\rho_0^\alpha}{\alpha\gamma}\,.$$

We can take ε small enough to obtain the claim for $r = d$.

We next assume that we have proved the claim for $u \geqslant r+1$ and choose ε_{r+1} so that the probability that the hitting time is smaller than T is smaller than $\eta/2$. We can choose I to be connected without loss of generality as the λ^i are ordered. We let $R = \min\{\tau_\varepsilon^I, \tau^{r+1}_{\varepsilon_{r+1}}\}$ when τ_ε^I is the first time where S^I reaches ε. Again following [55], we have

$$\begin{aligned}
\log S^I_{T\wedge R} \;=\; & \log S_0^I - 2\gamma T + 4\sqrt{2}\sum_{k,j\in I}\int_0^{T\wedge R}\frac{\lambda_j(t)-\lambda_k(t)}{S_t^I}db_t^j \\
& + 2\beta p\sum_{j,k\in I}\sum_{l\notin I}\int_0^{T\wedge R}\frac{\lambda_j(t)-\lambda_k(t)}{S_t^I}[\frac{1}{\lambda_j(t)-\lambda_l(t)} - \frac{1}{\lambda_k(t)-\lambda_l(t)}]dt \\
& + 4r[(r-1)(\frac{p\beta}{2}r+1) - 2]\int_0^{T\wedge R}\frac{dt}{S_t^I} \qquad (5.10)
\end{aligned}$$

Note that $M_t = 4\sqrt{2}\sum_{k,j\in I}\int_0^{t\wedge R}\frac{\lambda_j(s)-\lambda_k(s)}{S_s^I}db_s^j$ is a martingale with bracket $A_t = 16r\int_0^{t\wedge R}\frac{ds}{S_s^I}$. For $r \geq 3$, $4r[(r-1)(rp\beta/2+1) - 2] \geq 2p\beta > 0$ and therefore we deduce

$$\begin{aligned}
\mathbb{E}[\log S^I_{T\wedge R}] \geqslant\; & \log S_0^I - 2\gamma T + 2\beta p\mathbb{E}\left[\int_0^{T\wedge R}\frac{dt}{S_t^I}\right] \\
& + \mathbb{E}\left[2\beta p\sum_{j,k\in I}\sum_{l\notin I}\int_0^{T\wedge R}\frac{\lambda_j(t)-\lambda_k(t)}{S_t^I}[\frac{1}{\lambda_j(t)-\lambda_l(t)} - \frac{1}{\lambda_k(t)-\lambda_l(t)}]dt\right]
\end{aligned}$$

For $j, k \in I$, we cut the last integral over times

$$\Omega_{j,k} = \{t \leq T\wedge R : \sum_{l\notin I}\frac{1}{\lambda_j(t)-\lambda_l(t)}\frac{1}{\lambda_k(t)-\lambda_l(t)} \leq \frac{1}{S_t^I}\}$$

so that

$$-\sum_{j,k\in I}\int_{\Omega_{j,k}}\frac{(\lambda_j(t)-\lambda_k(t))^2}{S_t^I}\sum_{l\notin I}\Big[\frac{1}{(\lambda_j(t)-\lambda_l(t))(\lambda_k(t)-\lambda_l(t))}\Big]dt\geq-\int_0^{T\wedge R}\frac{dt}{S_t^I}$$

This term will therefore be compensated by the third term in (5.10). For the remaining term, if $l\notin I$ is such that $\min_{i\in I}|\lambda_l-\lambda_i|\leq\min_{i\in I}|\lambda_k-\lambda_i|$ for all $k\notin I$ then if $t\in\Omega_{j,k}^c$ and $i^*\in I$ is so that $\min_{i\in I}|\lambda_l-\lambda_i|=|\lambda_l-\lambda_{i^*}|$, we get

$$\frac{d-r}{(\lambda_l(t)-\lambda_{i^*}(t))^2}\geq\frac{1}{S_t^I}$$

and therefore on $\tau_{\varepsilon_{r+1}}^{r+1}\geq t$,

$$\varepsilon_{r+1}\leq S_t^I+\sum_{j\in I}(\lambda_j(t)-\lambda_l(t))^2\leq S_t^I+2r(\lambda_{i*}(t)-\lambda_l(t))^2+2S_t^I\leq(3+2r(d-r))S_t^I\,.$$

As a consequence, we have the bound for all $j,k\in I$, all $t\in\Omega_{j,k}^c$, $t\leq R$,

$$\frac{\lambda_j(t)-\lambda_k(t)}{S_t^I}\geqslant-1/\sqrt{S_t^I}\geqslant-\sqrt{3+2r(d-r)}/\sqrt{\varepsilon_{r+1}}$$

which entails the existence of a finite constant c so that

$$\sum_{j,k\in I}\sum_{l\notin I}\int_{\Omega_{j,k}^c}\frac{\lambda_j(t)-\lambda_k(t)}{S_t^I}\Big[\frac{1}{\lambda_j(t)-\lambda_l(t)}-\frac{1}{\lambda_k(t)-\lambda_l(t)}\Big]dt$$
$$\geqslant-\frac{c}{\sqrt{\varepsilon_{r+1}}}\sum_{i\in I}\sum_{l\notin I}\int_0^T\frac{dt}{|\,\lambda_i(t)-\lambda_l(t)\,|}\,.$$

Using Lemma 5.7 we hence conclude that there exists a universal finite constant c' depending only on T so that

$$\mathbb{E}[\log S_{T\wedge R}^I]\geqslant\log S_0^I-2\gamma T-\frac{c'}{\sqrt{\varepsilon_{r+1}}}\,.\tag{5.11}$$

On the other hand, we have

$$\mathbb{E}[\log S_{T\wedge R}^I]\leqslant\mathbb{P}(\tau_\varepsilon^I\leq T)\log(\varepsilon)+\mathbb{E}[\sup_{0\leq t\leq T}\log S_t^I]$$

where the last term is bounded above by (5.9). We deduce that

$$\mathbb{P}(\tau_\varepsilon^I\leq T)\leq\frac{|\log S_0^I|}{|\log(\varepsilon)|}+\frac{c''}{\sqrt{\varepsilon_{r+1}}|\log(\varepsilon)|}+\frac{c}{|\log(\varepsilon)\,|}+\frac{2\gamma T}{|\log(\varepsilon)\,|}\,.$$

We finally choose ε small enough so that the right hand side is smaller than $\eta/2$ to conclude. □

We next show that not only collisions of three particles are rare but also two collisions of different particles rarely happen around the same time.

Lemma 5.9. *For all i,j such that $i+1<j$, set*

$$\tau_{\varepsilon'}^{ij}=\inf\{t\geqslant0:(\lambda_i(t)-\lambda_{i-1}(t))^2+(\lambda_j(t)-\lambda_{j-1}(t))^2\leqslant\varepsilon'\}.$$

Then, for any $T>0$ and $\eta>0$, there exists ε' such that

$$\mathbb{P}\left[\tau_{\varepsilon'}^{ij}\leqslant T\right]\leqslant\eta.$$

Proof. Using Itô's formula, it is easy to see that

$$\begin{aligned}d\left((\lambda_i-\lambda_{i-1})^2+(\lambda_j-\lambda_{j-1})^2\right)&=8(1+p\beta)dt\\&-2\gamma\left[(\lambda_i-\lambda_{i-1})^2+(\lambda_j-\lambda_{j-1})^2\right]dt\\&+2\sqrt{2}\left[(\lambda_i-\lambda_{i-1})(db_t^i-db_t^{i-1})+(\lambda_j-\lambda_{j-1})(db_t^j-db_t^{j-1})\right]\\&-2p\beta\left[\sum_{k\neq i-1,i}\frac{(\lambda_i-\lambda_{i-1})^2}{(\lambda_i-\lambda_k)(\lambda_{i-1}-\lambda_k)}+\sum_{k\neq j-1,j}\frac{(\lambda_j-\lambda_{j-1})^2}{(\lambda_j-\lambda_k)(\lambda_{j-1}-\lambda_k)}\right]dt\,.\end{aligned}$$

Set $X_t := (\lambda_i(t) - \lambda_{i-1}(t))^2 + (\lambda_j(t) - \lambda_{j-1}(t))^2$ and note that the quadratic variation of

$$\int_0^t \frac{(\lambda_i - \lambda_{i-1})(db_s^i - db_s^{i-1}) + (\lambda_j - \lambda_{j-1})(db_s^j - db_s^{j-1})}{\sqrt{X_s}}$$

is $2t$. Thus there exists a standard Brownian motion B so that

$$\begin{aligned} dX_t &= 8(1+p\beta)dt - 2\gamma X_t dt + 4\sqrt{X_t}dB_t \\ &\quad - 2p\beta \left[\sum_{k\neq i-1,i} \frac{(\lambda_i - \lambda_{i-1})^2}{(\lambda_i - \lambda_k)(\lambda_{i-1} - \lambda_k)} + \sum_{k\neq j-1,j} \frac{(\lambda_j - \lambda_{j-1})^2}{(\lambda_j - \lambda_k)(\lambda_{j-1} - \lambda_k)} \right] dt\,. \end{aligned}$$

Note that, by the previous Lemma 5.8, we can choose ε such that

$$\mathbb{P}[\tau_\varepsilon^3 < T] \leqslant \frac{\eta}{2}. \tag{5.12}$$

Moreover, for all $t \leqslant \tau_\varepsilon^3$ such that $X_t \leqslant \varepsilon/4$, we have for all $k \neq i-1, i$,

$$(\lambda_i - \lambda_k)(\lambda_{i-1} - \lambda_k)(t) \geqslant \frac{\varepsilon}{8}.$$

The same property holds for j. To finish the proof, we will use the fact that the sum in the last term is bounded for all $t \leqslant \tau_\varepsilon^3$ such that $X_t \leqslant \varepsilon/4$. We thus need to introduce the process Y_t defined by $Y_t = \min(X_t, \frac{\varepsilon}{4})$. Let us set $f(x) := \min(x, \varepsilon/4)^{-p\beta}$. Note that f is a convex function $\mathbb{R}_+ \to \mathbb{R}_+$ and that the left-hand derivative of f is given by

$$f'_-(x) = -p\beta x^{-p\beta-1} 1_{\{x \leqslant \frac{\varepsilon}{4}\}}.$$

Its second derivative in the sense of distributions is the positive measure

$$f''(dx) = p\beta \left(\frac{\varepsilon}{4}\right)^{-p\beta-1} \delta_{\frac{\varepsilon}{4}} + \frac{p\beta(p\beta+1)}{x^{p\beta+2}} 1_{\{x \leqslant \frac{\varepsilon}{4}\}}\, dx\,.$$

Thus, by Itô-Tanaka formula, see e.g. [91, Theorem 6.22], we have

$$\begin{aligned} Y_t^{-p\beta} &= Y_0^{-p\beta} - p\beta \int_0^t X_s^{-p\beta-1} 1_{\{X_s \leqslant \frac{\varepsilon}{4}\}} dX_s \\ &\quad + \frac{1}{2}\left(p\beta \left(\frac{\varepsilon}{4}\right)^{-p\beta-1} L_t^{\frac{\varepsilon}{4}}(X) + \int_0^{\frac{\varepsilon}{4}} \frac{p\beta(p\beta+1)}{x^{p\beta+2}} L_t^x(X) dx \right), \end{aligned}$$

where $L_t^x(X)$ is the local time of X in x. By definition we have

$$\int_0^{\frac{\varepsilon}{4}} \frac{p\beta(p\beta+1)}{x^{p\beta+2}} L_t^x(X) dx = \int_0^t \frac{p\beta(p\beta+1)}{X_s^{p\beta+2}} 1_{\{X_s \leqslant \frac{\varepsilon}{4}\}} d\langle X, X\rangle_s,$$

and thus, we obtain

$$\begin{aligned} Y_t^{-p\beta} &= Y_0^{-p\beta} + \int_0^t 1_{\{X_s \leqslant \frac{\varepsilon}{4}\}} \left(p\beta\gamma Y_s^{-p\beta} dt + 4 Y_s^{-p\beta-\frac{1}{2}} dB_s \right) \\ &+ 2p^2\beta^2 \int_0^t Y_s^{-p\beta-1} \left[\sum_{k\neq i-1,i} \frac{((\lambda_i - \lambda_{i-1})(s))^2}{((\lambda_i - \lambda_k)(s))((\lambda_{i-1} - \lambda_k)(s))} \right. \\ &\left. + \sum_{k\neq j-1,j} \frac{((\lambda_j - \lambda_{j-1})(s))^2}{((\lambda_j - \lambda_k)(s))((\lambda_{j-1} - \lambda_k)(s))} \right] 1_{X_s \leq \varepsilon/4} ds + \frac{1}{2} p\beta \left(\frac{\varepsilon}{4}\right)^{-p\beta-1} L_t^{\frac{\varepsilon}{4}}(X)\,. \end{aligned} \tag{5.13}$$

The definition of local time implies that, almost surely, $L_t^x(X) \leqslant t$. We thus deduce from (5.13) that

$$\mathbb{E}\left[Y^{-p\beta}_{T\wedge\tau^{ij}_{\varepsilon'}\wedge\tau^3_\varepsilon} \right] \leqslant Y_0^{-p\beta} + \frac{1}{2} p\beta \left(\frac{\varepsilon}{4}\right)^{-p\beta-1} T + C \int_0^T \mathbb{E}\left[Y^{-p\beta}_{t\wedge\tau^{ij}_{\varepsilon'}} \right] dt\,.$$

with $C = (p\beta\gamma + 4p^2\beta^2(d-1)\frac{8}{\varepsilon})$. Gronwall's Lemma implies that

$$\mathbb{E}\left[Y^{-p\beta}_{T\wedge\tau^{ij}_{\varepsilon'}\wedge\tau^3_\varepsilon}\right] \leqslant \left(Y_0^{-p\beta} + \frac{1}{2}p\beta\left(\frac{\varepsilon}{4}\right)^{-p\beta-1} T\right)\exp(CT). \tag{5.14}$$

If $\varepsilon' < \varepsilon/4$, equation (5.14) implies that

$$(\varepsilon')^{-p\beta}\mathbb{P}\left[\tau^{ij}_{\varepsilon'} \leqslant T\wedge\tau^3_\varepsilon\right] \leqslant Y_0^{-p\beta}\exp(CT), \tag{5.15}$$

Taking ε' small enough gives the result with (5.12). □

As a direct consequence, we deduce the uniqueness of the i^* of Proposition 5.5.

Lemma 5.10. *With the same notations as in the previous Lemma 5.9, we have almost surely*

$$\inf_{(k,\ell):k+1<\ell} \tau_0^{k\ell} = +\infty.$$

In particular, this gives the unicity of the i^* in Proposition 5.5.

Proof. It is enough to write that for all $\varepsilon > 0$

$$\mathbb{P}\left(\inf_{k+1<\ell}\tau_0^{k\ell} \leq T\right) \leq d^2\{\max_{k+1<\ell}\mathbb{P}\left(\tau_0^{k\ell} \leq T\wedge\tau^3_\varepsilon\right) + \mathbb{P}\left(\tau^3_\varepsilon \leq T\right)\}$$

and deduce from Lemmas 5.9 and 5.8 that the right hand side is as small as wished when ε goes to zero.

□

5.3.3 Smoothness properties of the limiting process

Lemma 5.11. *We have the following smoothness properties:*

- *For all $T < \infty$ and $\varepsilon > 0$, there exists C, c', c finite positive constants so that for all δ, η positive real numbers so that $\eta \leq c'(\varepsilon^2 \wedge \delta\varepsilon)$ we have*

$$\mathbb{P}\left[\max_{1\leq i\leq d}\sup_{\substack{s\leq t\leq(s+\eta)\wedge\tau^3_\varepsilon\\ 0\leqslant t\leqslant T}} |\lambda_i(s) - \lambda_i(t)| \geq \delta\right] \leq \frac{C}{\eta}\left(e^{-c\delta^4/2\eta} + e^{-c\varepsilon^4/\eta}\right). \tag{5.16}$$

- *For all $T < \infty$ and $\varepsilon > 0$, there exists C, c', c finite positive constants so that for all δ, η positive real numbers so that $\eta \leq c'(\varepsilon^2 \wedge \delta\varepsilon)$ we have*

$$\mathbb{P}\left[\max_{i\neq j}\sup_{\substack{s\leq t\leq(s+\eta)\wedge\tau^3_\varepsilon\\ 0\leqslant t\leqslant T}}\int_s^t \frac{du}{|\lambda_i(u) - \lambda_j(u)|} \geq \delta\right] \leq \frac{C}{\eta}\left(e^{-c\delta^4/2\eta} + e^{-c\varepsilon^4/\eta}\right). \tag{5.17}$$

Proof. Let us first fix $s \in [0,T]$ and set $I = \{i \in \{2,\dots,d\} : |\lambda_i(s) - \lambda_{i-1}(s)| \leqslant \varepsilon/3\}$ and note that on the event $\{s \leqslant \tau^3_\varepsilon\}$, the connected subsets of I contain at most one element. Let $T_\varepsilon = \inf\{t \geq s : \inf_{i\notin I}|\lambda_i(t) - \lambda_{i-1}(t)| \leqslant \varepsilon/4\}$. The continuity of the λ_i implies that T_ε is almost surely strictly positive.

If $i \notin I \cup \{I-1\}$, then we have, for $t \in [s; (s+\eta)\wedge\tau^3_\varepsilon \wedge T_\varepsilon]$

$$\begin{aligned}|\lambda_i(t) - \lambda_i(s)| &\leqslant \gamma\int_s^t |\lambda_i(u)|du + \sqrt{2}|b^i_t - b^i_s| + p\beta\int_s^t\sum_{j\neq i}\frac{du}{|\lambda_i(u) - \lambda_j(u)|}\\ &\leqslant \gamma\int_s^t |\lambda_i(u)|du + \sqrt{2}|b^i_t - b^i_s| + 4p\beta(d-1)\frac{t-s}{\varepsilon}.\end{aligned}$$

Using (5.9) and [15, Corollary H.13], it is easy to deduce that there exists a constant $c > 0$ such that for $\eta < \varepsilon\delta/(8p\beta(d-1))$

$$\mathbb{P}\left[\max_{i\notin I\cup\{I-1\}}\sup_{t\in[s;(s+\eta)\wedge\tau^3_\varepsilon\wedge T_\varepsilon]} |\lambda_i(t) - \lambda_i(s)| \geqslant \delta\right] \leqslant cde^{-\frac{\delta^2}{2\eta}}. \tag{5.18}$$

Now, if $i \in I$, with the same argument as for (5.18) (the drift term in the SDE satisfied by $\lambda_i + \lambda_{i-1}$ is also bounded), we can show that there exists a constant $c > 0$ such that

$$\mathbb{P}\left[\sup_{t\in[s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]} |(\lambda_i+\lambda_{i-1})(t) - (\lambda_i+\lambda_{i-1})(s)| \geqslant \delta\right] \leqslant ce^{-c\frac{\delta^2}{2\eta}}. \tag{5.19}$$

On the other hand, the process $x_i(t) := (\lambda_i - \lambda_{i-1})(t)$ verifies

$$\begin{aligned} dx_i^2(t) &= 4(1+p\beta)dt - \gamma x_i^2(t)dt + 2x_i(t)(db_t^i - db_t^{i-1}) \\ &\quad - 2p\beta \sum_{k\neq j-1,j} \frac{(\lambda_i(t)-\lambda_{i-1}(t))^2}{(\lambda_i(t)-\lambda_k(t))(\lambda_{i-1}(t)-\lambda_k(t))}dt\,. \end{aligned}$$

The denominator in the last term of the above r.h.s is bounded below on the interval $t \in [s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]$ by $2p\beta(d-2)\frac{1}{\varepsilon}$. Thus, using again (5.9) and [15, Corollary H.13], we can show that for $\delta > c\eta/\varepsilon$,

$$\mathbb{P}\left[\sup_{t\in[s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]} |x_i(t)-x_i(s)| \geqslant \sqrt{\delta}\right] \leq \mathbb{P}\left[\sup_{t\in[s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]} |x_i^2(t)-x_i^2(s)| \geqslant \delta\right] \leqslant ce^{-c\frac{\delta^2}{2\eta}} \tag{5.20}$$

where the first inequality is due to the fact that x_i is non-negative. Using (5.19) and (5.20) gives for $\eta < \delta\varepsilon/c$

$$\mathbb{P}\left[\max_{i\in I\cup\{I-1\}} \sup_{t\in[s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]} |\lambda_i(t)-\lambda_i(s)| \geqslant \delta\right] \leqslant 2cde^{-c\frac{\delta^4}{2\eta}}.$$

Thus, with (5.18), we deduce that for $\eta < \delta\varepsilon/c$

$$\mathbb{P}\left[\max_{i} \sup_{t\in[s;(s+\eta)\wedge\tau_\varepsilon^3\wedge T_\varepsilon]} |\lambda_i(t)-\lambda_i(s)| \geqslant \delta\right] \leqslant 2cde^{-c\frac{\delta^4}{2\eta}}.$$

In particular, there exists $c' > 0$ so that if $\varepsilon^2 > c\eta$,

$$\mathbb{P}\left[T_\varepsilon < (s+\eta)\wedge\tau_\varepsilon^3\right] \leq \mathbb{P}\left[\max_{i} \sup_{s\leq t\leq(s+\eta)\wedge T_\varepsilon\wedge\tau_\varepsilon^3} |\lambda_i(t)-\lambda_i(s)| \geqslant 5\varepsilon/12\right] \leq \frac{4cdT}{\eta}e^{-c'\varepsilon^4/2\eta},$$

which is as small as wished provided η is chosen small enough. This allows to remove the stopping time and get for any fixed $s < T$, and $\delta > c\eta/\varepsilon$

$$\mathbb{P}\left[\max_{i} \sup_{s\leq t\leq(s+\eta)\wedge\tau_\varepsilon^3} |\lambda_i(t)-\lambda_i(s)| \geqslant \delta\right] \leq 2cde^{-c\delta^4/2\eta} + 2dce^{-c'\varepsilon^4/2\eta}.$$

The uniform estimate on s is obtained as usual by taking s in a grid with mesh $\eta/2$ up to divise δ by two and to multiply the probability by $2T/\eta$. Thus we find constant c, c', and C so that if $\eta \leq c(\varepsilon^2 \wedge \delta\varepsilon)$ we have

$$\mathbb{P}\left[\max_{i} \sup_{\substack{s\leq t\leq(s+\eta)\wedge\tau_\varepsilon^3\\ 0\leq s,t\leq T}} |\lambda_i(t)-\lambda_i(s)| \geqslant \delta\right] \leq \frac{CT}{\eta}\left(e^{-c\delta^4/2\eta} + e^{-c'\varepsilon^4/\eta}\right).$$

The second control is a direct consequence of the first as we can first consider the cas $j = d$ to deduce that for $i < d$

$$\left|\int_s^t \frac{du}{\lambda_d(u)-\lambda_i(u)}\right| \leq |\lambda_d(t)-\lambda_d(s)| + \sqrt{2}|b_d(t)-b_d(s)|$$

where the right hand side is continuous. We then consider recursively the other indices. □

5.3.4 Approximation by less colliding processes

When $p\beta \geqslant 1$, it is well known [15, Lemma 4.3.3] that the process λ has almost surely no collision. In this case, the singularity of the drift which defines the SDE is not really important as it is almost always avoided. In the case $p\beta < 1$, we know that collisions occur and in fact can occur as much as for a Bessel process with small parameter. The singularity of the drift becomes important, in particular when we will show the convergence in law of the process of the eigenvalues λ^n towards λ. To this end, we show that λ can be approximated by a process which does not spend too much time in collisions.

For $\delta > 0$, we define a new process $(\lambda_i^\delta(t))_{t \geqslant 0}$ as follows.

Definition 5.12. *Let $T_1 := \inf\{t \geqslant 0 : \exists i \neq j, \lambda_i(t) = \lambda_j(t)\}$ and for all $t < T_1$, set $\lambda_i^\delta(t) := \lambda_i(t)$. For $t > T_1$, we define the process recursively by setting for all $\ell \geqslant 2$, $\lambda_i^\delta(T_\ell^\delta) := \lambda_i^\delta(T_\ell^\delta -) + i\delta$ and for $t > T_\ell^\delta$, the process $\lambda_i^\delta(t)$ is defined up to time $T_{\ell+1}^\delta := \inf\{t > T_\ell^\delta : \exists i \neq j, \lambda_i^\delta(t) = \lambda_j^\delta(t)\}$ as the unique strong solution of the system*

$$d\lambda_i^\delta(t) = -\gamma\lambda_i^\delta(t)dt + \sqrt{2}db_t^i + p\beta \sum_{j\neq i} \frac{dt}{\lambda_i^\delta(t) - \lambda_j^\delta(t)}\,. \tag{5.21}$$

The main result of this section is that

Theorem 5.13. *Construct the process λ with the same Brownian motion b. Then, for any time $T > 0$, any $\xi \in (0, p\beta/4)$*

$$\lim_{\delta\downarrow 0} \mathbb{P}\left(\sup_{0\leq t\leq T} \max_{1\leq i\leq d} |\lambda_i(t) - \lambda_i^\delta(t)| \leq \delta^\xi \right) = 1\,.$$

The theorem is a direct consequence of the following lemma and proposition.

Lemma 5.14. *Let $\delta > 0$. Construct the process λ with the same Brownian motion b than λ^δ. There exists a constant $c > 0$ such that, almost surely, for all $\ell \in \mathbb{N}$*

$$\max_{1\leqslant i\leqslant d} \sup_{0\leqslant t\leqslant T_\ell^\delta} |\lambda_i^\delta(t) - \lambda_i(t)| \leqslant c\delta\ell\,.$$

To finish the proof it is enough to show that T_ℓ^δ goes to infinity for $\ell \ll 1/\delta$. This is the content of the next proposition.

Proposition 5.15. *Let $T < \infty$, $0 < \xi < p\beta/4$ and $L = [1/\delta^{1-\xi}]$. Then the probability $\mathbb{P}\left[T_L^\delta \leqslant T\right]$ vanishes when δ goes to zero.*

Proof of Lemma 5.14. We proceed by induction over ℓ to show that, for each ℓ,

$$\sup_{0\leqslant t\leqslant T_\ell^\delta} \left(\sum_{i=1}^d (\lambda_i^\delta - \lambda_i)^2(t) \right)^{1/2} \leqslant c\delta\ell$$

with $c = (\sum_{i=1}^d i^2 = d(d+1)(2d+1)/6)^{\frac{1}{2}}$.

• We treat the case $\ell = 1$. By definition of the processes, $\lambda^\delta = \lambda$ on $[0, T_1^\delta)$. At time $t = T_1^\delta$, the separation procedure implies that

$$\sum_{i=1}^d (\lambda_i^\delta - \lambda_i)^2(T_1^\delta) = \sum_{i=1}^d ((\lambda_i^\delta - \lambda_i)(T_1^\delta -) + i\delta)^2 = c^2\delta^2\,.$$

The property is true for $\ell = 1$.

• Suppose it is true for ℓ. For $t \in [T_\ell^\delta, T_{\ell+1}^\delta)$, since λ^δ and λ are driven by the same Brownian motion, we get

$$\begin{aligned} d\sum_{i=1}^d (\lambda_i^\delta(t) - \lambda_i(t))^2 &= -2\gamma \sum_{i=1}^d (\lambda_i^\delta(t) - \lambda_i(t))^2 dt \\ &\quad + 2p\beta \sum_{i=1}^d \sum_{j\neq i} (\lambda_i^\delta(t) - \lambda_i(t)) \left(\frac{1}{\lambda_i^\delta(t) - \lambda_j^\delta(t)} - \frac{1}{\lambda_i(t) - \lambda_j(t)} \right) dt\,. \end{aligned}$$

Observe that

$$\begin{aligned}
&\sum_{i=1}^{d}\sum_{j\neq i}(\lambda_i^\delta(t)-\lambda_i(t))\left(\frac{1}{\lambda_i^\delta(t)-\lambda_j^\delta(t)}-\frac{1}{\lambda_i(t)-\lambda_j(t)}\right) \qquad (5.22)\\
&=\frac{1}{2}\sum_{i=1}^{d}\sum_{j\neq i}(\lambda_i^\delta(t)-\lambda_j^\delta(t)-(\lambda_i(t)-\lambda_j(t)))\left(\frac{1}{\lambda_i^\delta(t)-\lambda_j^\delta(t)}-\frac{1}{\lambda_i(t)-\lambda_j(t)}\right)\\
&=\frac{1}{2}\sum_{i=1}^{d}\sum_{j\neq i}\Big(\lambda_i^\delta(t)-\lambda_j^\delta(t)-(\lambda_i(t)-\lambda_j(t))\Big)^2\frac{1}{(\lambda_i^\delta(t)-\lambda_j^\delta(t))(\lambda_i(t)-\lambda_j(t))}\\
&\leqslant 0
\end{aligned}$$

as the $(\lambda_i)_{1\leq i\leq d}$ and the $(\lambda_i^\delta)_{1\leq i\leq d}$ are ordered. Thus,

$$\sup_{t\in[T_\ell^\delta,T_{\ell+1}^\delta)}\sum_{i=1}^{d}(\lambda_i^\delta(t)-\lambda_i(t))^2\leqslant\sum_{i=1}^{d}(\lambda_i^\delta(T_\ell^\delta)-\lambda_i(T_\ell^\delta))^2. \qquad (5.23)$$

In addition, because of the separation procedure at time $T_{\ell+1}^\delta$, we have

$$\begin{aligned}
\left(\sum_{i=1}^{d}(\lambda_i^\delta-\lambda_i)^2(T_{\ell+1}^\delta)\right)^{1/2}&=\left(\sum_{i=1}^{d}\Big((\lambda_i^\delta-\lambda_i)(T_{\ell+1}^\delta-)+i\delta\Big)^2\right)^{1/2}\\
\leqslant\left(\sum_{i=1}^{d}(\lambda_i^\delta-\lambda_i)^2(T_{\ell+1}^\delta-)\right)^{1/2}&+\delta c\leqslant\delta(\ell+1)c\,,
\end{aligned}$$

where we used the induction hypothesis in the last line. The proof is thus complete. □

Proof of Proposition 5.15. In the case $p\beta\geq 1$, it is well known [15, p. 252] that T_1 is almost surely infinite and therefore the proposition is trivial. We hence restrict ourselves to $p\beta\leq 1$. Let $\eta>0$. Let us define the stopping times

$$\begin{aligned}
\tau_\varepsilon^{3,\delta}&:=\inf\{t\geqslant 0:\min_{|I|=3}S_t^{I,\delta}\leqslant\varepsilon\}\,,\\
\tau_\varepsilon^{2,\delta}&:=\inf\{t\geqslant 0:\min_{1\leqslant i,j\leqslant d}((\lambda_i^\delta-\lambda_{i-1}^\delta)^2+(\lambda_j^\delta-\lambda_{j-1}^\delta)^2)(t)\leqslant\varepsilon\},
\end{aligned}$$

where $S_t^{I,\delta}:=\sum_{i,j\in I}(\lambda_i^\delta-\lambda_j^\delta)^2(t)$. Set also $\tau_\varepsilon^\delta:=\tau_\varepsilon^{2,\delta}\wedge\tau_\varepsilon^{3,\delta}$. We know from Lemmas 5.8 and 5.9 that we can choose ε small enough so that

$$\mathbb{P}\left[\tau_{2\varepsilon}^3\wedge\tau_{2\varepsilon}^2\leqslant T\right]\leqslant\eta.$$

The number ε being fixed, it is then straightforward to see from Lemma 5.14 that there exists δ_0 small enough so that for all $\delta\leqslant\delta_0$, we have

$$\mathbb{P}\left[\tau_\varepsilon^\delta\leqslant T\right]\leqslant\eta.$$

Now, we have

$$\mathbb{P}\left[T_L^\delta\leqslant T\right]\leqslant\eta+\mathbb{P}\left[\delta^\xi\sum_{\ell=1}^{L}1_{\{T_{\ell+1}^\delta-T_\ell^\delta\geqslant\delta^\xi\}}\leqslant T;\tau_\varepsilon^\delta\geqslant T_L^\delta\right].$$

We need to show that the second term goes to 0 when $\delta\to 0$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the filtration of the driving Brownian motion. We will prove in Lemma 5.18, there exists a constant $c>0$ such that, on the event $\{\tau_\varepsilon^\delta\geqslant T_L^\delta\}$, almost surely

$$\sum_{\ell=1}^{L}\mathbb{P}\left[T_{\ell+1}^\delta-T_\ell^\delta\geqslant\delta^\xi\mid\mathcal{F}_{T_\ell^\delta}\right]\geqslant c\,\delta^{-p\beta+\xi}\,.$$

In the following, we suppose that δ is small enough so that $c\,\delta^{-p\beta+\xi} \geqslant \delta^{-p\beta+2\xi}$ and $\delta^{-\xi}\,T - \delta^{-p\beta+\xi} \leqslant -\delta^{-p\beta+2\xi}$. For such δ, we have

$$\begin{aligned}
&\mathbb{P}\left[\sum_{\ell=1}^{L} 1_{\{T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi\}} \leqslant \delta^{-\xi}\,T; \tau^\delta_\varepsilon \geqslant T^\delta_L\right] \\
&\leqslant \mathbb{P}\left[\sum_{\ell=1}^{L} 1_{\{T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi\}} - \mathbb{P}\left[T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right] \leqslant -\delta^{-p\beta+2\xi}; \tau^\delta_\varepsilon \geqslant T^\delta_L\right] \\
&\leqslant \mathbb{P}\left[\left|\sum_{\ell=1}^{L} 1_{\{T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi\}} - \mathbb{P}\left[T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right]\right| \geqslant \delta^{-p\beta+2\xi}; \tau^\delta_\varepsilon \geqslant T^\delta_L\right] \\
&\leqslant \delta^{2p\beta-4\xi} \sum_{\ell=1}^{L} \mathbb{P}\left[T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi; \tau^\delta_\varepsilon \geqslant T^\delta_L\right]
\end{aligned}$$

where we used the Tchebychev inequality in the last line. Using Lemma 5.16, we get that there exists a constant $C>0$ such that

$$\mathbb{P}\left[\sum_{\ell=1}^{L} 1_{\{T^\delta_{\ell+1}-T^\delta_\ell \geqslant \delta^\xi\}} \leqslant \delta^{-\xi}\,T; \tau^\delta_\varepsilon \geqslant T^\delta_L\right] \leqslant C\,\delta^{2p\beta-4\xi}\,L\,\delta^{(1-p\beta)(1-2^{-1}\xi)} \leqslant C\,\delta^{p\beta-4\xi}$$

which goes to 0 when δ goes to 0. The proposition is proved. □

Lemma 5.16. *Let $\xi \in (0;2)$. Then there exists a constant $C>0$ such that, almost surely, on ; $\tau^\delta_\varepsilon \geqslant T^\delta_L$*

$$\mathbb{P}\left[\delta^\xi \leqslant T^\delta_{\ell+1}-T^\delta_\ell \mid \mathcal{F}_{T^\delta_\ell}\right] \leqslant C\delta^{(1-p\beta)(1-2^{-1}\xi)}\,. \tag{5.24}$$

Proof. We know that there are no multiple collisions nor simultaneous collisions (because of Lemmas 5.8 and 5.9) and therefore we can denote by i the unique element such that $\lambda^\delta_i(T^\delta_\ell-) = \lambda^\delta_{i-1}(T^\delta_\ell-)$ and $(\lambda^\delta_i-\lambda^\delta_{i-1})(T^\delta_\ell)=\delta$. We have by Itô's formula

$$\begin{aligned}
d(\lambda^\delta_i-\lambda^\delta_{i-1})(t) &= -\gamma(\lambda^\delta_i-\lambda^\delta_{i-1})(t)dt + \sqrt{2}(db^i_t - db^{i-1}_t) \\
&+ 2p\beta\frac{dt}{(\lambda^\delta_i-\lambda^\delta_{i-1})(t)} - \beta p\sum_{k\neq i,i-1}\frac{(\lambda^\delta_i-\lambda^\delta_{i-1})(t)}{(\lambda^\delta_i-\lambda^\delta_k)(t)(\lambda^\delta_{i-1}-\lambda^\delta_k)(t)}dt\,.
\end{aligned} \tag{5.25}$$

Let us define the Bessel like process $(X_t)_{t\geqslant 0}$ by $X_0=\delta$ and for $t\geqslant 0$,

$$dX_t = \sqrt{2}(db^i_{T^\delta_\ell+t} - db^{i-1}_{T^\delta_\ell+t}) + 2p\beta\frac{dt}{X_t}. \tag{5.26}$$

Using the comparison theorem for SDE [91, Proposition 2.18] (note that the drifts are smooth before $T^\delta_{\ell+1}-T^\delta_\ell$), we know that for all $t\in[0,T^\delta_{\ell+1}-T^\delta_\ell)$, we have almost surely

$$(\lambda^\delta_i-\lambda^\delta_{i-1})(T^\delta_\ell+t) \leqslant X_t. \tag{5.27}$$

Let us define $T^\delta_X := \inf\{t\geqslant 0 : X_t=0\}$. It is clear that almost surely $T^\delta_{\ell+1}-T^\delta_\ell \leqslant T^\delta_X$. We thus have on $\tau^\delta_\varepsilon \geqslant T^\delta_L$

$$\mathbb{P}\left[\delta^\xi \leqslant T^\delta_{\ell+1}-T^\delta_\ell \mid \mathcal{F}_{T^\delta_\ell}\right] \leqslant \mathbb{P}\left[T^\delta_X \geqslant \delta^\xi\right].$$

We finally conclude using a classical result for Bessel process, see e.g. [**?**, (13)]; the density with respect to the Lebesgue measure on $\mathbb{R}_+$ of the law of the random variable T^δ_X is

$$p_\delta(t) = \frac{1}{\Gamma(\frac{1-p\beta}{2})}\frac{1}{t}\left(\frac{\delta^2}{2t}\right)^{\frac{1-p\beta}{2}} e^{-\frac{\delta^2}{2t}}\,.$$

Hence we deduce that for $\xi\leq 2$ there exists a constant $c>0$ such that

$$\mathbb{P}\left[T^\delta_X \geqslant \delta^\xi\right] \leqslant c\,\delta^{(1-p\beta)(1-2^{-1}\xi)}.$$

□

For time $t \in [0;T]$, we define the random set

$$I_t := \{i \in \{2,\dots,d\} : |\lambda_i^\delta - \lambda_{i-1}^\delta|(t) \leqslant \sqrt{\varepsilon}/3\}. \tag{5.28}$$

Note that, on the event $\Omega := \{\tau_\varepsilon^\delta \geqslant T\}$, for each $t \leqslant T$, the set I_t contains at most one element. For each $\ell \in \{1,\dots,L\}$, and $i \in \{1,\dots,d\}$, we define the stopping times

$$t_\ell^\delta(\sqrt{\varepsilon}/3) := \inf\{t \geqslant T_\ell^\delta : \min_j |\lambda_j^\delta - \lambda_{j-1}^\delta|(t) \geqslant \sqrt{\varepsilon}/3\}\,,$$
$$\bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6) := \inf\{t \geqslant T_\ell^\delta : \min_{j\neq i} |\lambda_j^\delta - \lambda_{j-1}^\delta|(t) \leqslant \sqrt{\varepsilon}/6\}\,.$$

If i denotes the unique index such that $\lambda_i^\delta(T_\ell^\delta-) = \lambda_{i-1}(T_\ell^\delta-)$, note that if $T_\ell^\delta \leqslant \tau_\varepsilon^\delta$ then $\min_{j\neq i} |\lambda_j^\delta - \lambda_{j-1}^\delta|(T_\ell^\delta) \geqslant \sqrt{\varepsilon}/3$.

Lemma 5.17. *If $T_\ell^\delta \leqslant \tau_\varepsilon^\delta$ and if i denotes the (unique) index such that $\lambda_i^\delta(T_\ell^\delta-) = \lambda_{i-1}^\delta(T_\ell^\delta-)$, then there exists a constant $c>0$ and $\delta_0>0$ such that for all $\delta \leqslant \delta_0$, we have*

$$c\delta^{1-p\beta} \leqslant \mathbb{P}\left[t_\ell^\delta(\sqrt{\varepsilon}/3) \wedge \bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6) \leqslant T_{\ell+1}^\delta | \mathcal{F}_{T_\ell^\delta}\right]. \tag{5.29}$$

Proof. Note that i is the unique element of the set $I_{T_\ell^\delta}$ defined by (5.28) for which $|\lambda_i^\delta - \lambda_{i-1}^\delta|(T_\ell^\delta) = \delta$. For $\alpha = 1-p\beta$ and $t \in [T_\ell^\delta; T_{\ell+1}^\delta)$, we have by Itô's formula

$$\begin{aligned} d(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t) &= -\gamma\alpha(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t)dt \\ &+ \alpha(\lambda_i^\delta - \lambda_{i-1}^\delta)^{\alpha-1}(t)\sqrt{2}(db_t^i - db_t^{i-1}) - \beta p \sum_{k\neq i,i-1} \frac{(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t)}{(\lambda_i^\delta - \lambda_k^\delta)(t)(\lambda_{i-1}^\delta - \lambda_k^\delta)(t)}dt\,. \end{aligned} \tag{5.30}$$

For $t \in [T_\ell^\delta, \tau_\varepsilon^\delta]$, we deduce that

$$d(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t) \geq \alpha(\lambda_i^\delta - \lambda_{i-1}^\delta)^{\alpha-1}(t)\sqrt{2}(db_t^i - db_t^{i-1}) - c'(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t)dt$$

where $c' = \alpha\gamma + \beta p(d-2)36/\varepsilon$. Let $T_{\ell+1}^{\delta,\kappa}$ be the first time after T_ℓ^δ so that $\lambda_i^\delta - \lambda_i^{\delta-1}$ reaches $\kappa < \delta$. Then, as $\int_0^{\cdot\wedge T_{\ell+1}^{\delta,\kappa}} (\lambda_i^\delta - \lambda_{i-1}^\delta)^{\alpha-1}(t)\sqrt{2}(db_t^i - db_t^{i-1})$ is a martingale, we find that

$$\mathbb{E}\left[(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t_\ell^\delta(\sqrt{\varepsilon}/3) \wedge \bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6) \wedge T_{\ell+1}^{\delta,\kappa}) \mid \mathcal{F}_{T_\ell^{\delta,\kappa}}\right] \geqslant \delta^\alpha \exp\left(-c'\,T\right). \tag{5.31}$$

Before time $\bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6)$, $(\lambda_j^\delta - \lambda_{j-1}^\delta)(t)$ can not cancel if $j \neq i$. Therefore we can choose κ small enough so that the last inequality implies

$$\mathbb{E}\left[(\lambda_i^\delta - \lambda_{i-1}^\delta)^\alpha(t_\ell^\delta(\sqrt{\varepsilon}/3) \wedge \bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6))\,1_{\{t_\ell^\delta(\sqrt{\varepsilon}/3)\wedge\bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6) \leqslant T_{\ell+1}^\delta\}} \mid \mathcal{F}_{T_\ell^\delta}\right] \geqslant \frac{1}{2}\delta^\alpha \exp\left(-c'\,T\right).$$

which can be rewriten using the fact that $|\lambda_i^\delta - \lambda_{i-1}^\delta|(t_\ell^\delta(\sqrt{\varepsilon}/3) \wedge \bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6)) \leqslant \sqrt{\varepsilon}/3$, as follows

$$\mathbb{P}\left[t_\ell^\delta(\sqrt{\varepsilon}/3) \wedge \bar{t}_\ell^\delta(i,\sqrt{\varepsilon}/6) \leqslant T_{\ell+1}^\delta \mid \mathcal{F}_{T_\ell^\delta}\right] \geqslant \delta^\alpha \left(\frac{3}{\sqrt{\varepsilon}}\right)^\alpha \exp(-c'\,T)\,.$$

The lemma follows with $c = (\frac{3}{\sqrt{\varepsilon}})^\alpha \exp(-c'\,T)$. □

Lemma 5.18. *Let $\xi, T > 0$. There exists a constant $c>0$ and $\delta_0>0$ so that if $\delta \leq \delta_0$, on $T_\ell^\delta \leqslant \tau_\varepsilon^\delta \wedge T$,*

$$\mathbb{P}\left[\delta^\xi \leqslant T_{\ell+1}^\delta - T_\ell^\delta \mid \mathcal{F}_{T_\ell^\delta}\right] \geqslant c\delta^{1-p\beta}\,. \tag{5.32}$$

Proof. We assume in the sequel that $\delta \leq 1$. The proof is based on Lemma 5.17. It implies

$$\begin{aligned}&\mathbb{P}\left[\delta^\xi \leqslant T^\delta_{\ell+1} - T^\delta_\ell \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\quad\geqslant \mathbb{P}\left[t^\delta_\ell(\sqrt{\varepsilon}/3) \wedge \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) \leqslant T^\delta_{\ell+1}; \delta^\xi \leqslant T^\delta_{\ell+1} - T^\delta_\ell \leq 1 \mid \mathcal{F}_{T^\delta_\ell}\right].\end{aligned}$$

By Lemma 5.17, we deduce that

$$\begin{aligned}&\mathbb{P}\left[\delta^\xi \leqslant T^\delta_{\ell+1} - T^\delta_\ell \leq 1 \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\quad\geqslant c\delta^{1-p\beta} - \mathbb{P}\left[t^\delta_\ell(\sqrt{\varepsilon}/3) \wedge \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) \leqslant T^\delta_{\ell+1} \leqslant T+1; \delta^\xi \geqslant T^\delta_{\ell+1} - T^\delta_\ell \mid \mathcal{F}_{T^\delta_\ell}\right].\end{aligned}$$

But

$$\begin{aligned}&\mathbb{P}\left[t^\delta_\ell(\sqrt{\varepsilon}/3) \wedge \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) \leqslant T^\delta_{\ell+1} \leq T+1; T^\delta_{\ell+1} - T^\delta_\ell \leqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\quad\leqslant \mathbb{P}\left[t^\delta_\ell(\sqrt{\varepsilon}/3) \leqslant \ \wedge T+1; T^\delta_{\ell+1} - t^\delta_\ell(\sqrt{\varepsilon}/3) \leqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\quad+ \mathbb{P}\left[\bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) \leqslant t^\delta_\ell(\sqrt{\varepsilon}/3); \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) - T^\delta_\ell \leqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right].\end{aligned}$$

Let us handle the first term of the previous right hand side

$$\begin{aligned}&\mathbb{P}\left[t^\delta_\ell(\sqrt{\varepsilon}/3) \leqslant T^\delta_{\ell+1} \wedge (T+1); T^\delta_{\ell+1} - t^\delta_\ell(\sqrt{\varepsilon}/3) \leqslant \delta^\xi \mid \mathcal{F}_{t^\delta_\ell(\sqrt{\varepsilon}/3)}\right]\\ &\leqslant \mathbb{P}\left[\max_j \sup_{t^\delta_\ell(\sqrt{\varepsilon}/3) \leqslant s \leqslant (t^\delta_\ell(\sqrt{\varepsilon}/3)+\delta^\xi)\wedge t^\delta_\ell(\sqrt{\varepsilon}/12)\wedge(T+1)} |\lambda^\delta_j(s) - \lambda^\delta_j(t^\delta_\ell(\sqrt{\varepsilon}/3))| \geqslant \frac{\sqrt{\varepsilon}}{24} \mid \mathcal{F}_{t^\delta_\ell(\sqrt{\varepsilon}/3)}\right]\\ &\leqslant C\exp(-\frac{c\varepsilon^2}{\delta^\xi})\end{aligned}$$

where we used Lemma 5.11 for the last line (actually the proof since we used the estimate for a fixed s). For the second term, the idea is similar

$$\begin{aligned}&\mathbb{P}\left[\bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) \leqslant t^\delta_\ell(\sqrt{\varepsilon}/3); \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6) - T^\delta_\ell \leqslant \delta^\xi \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\leqslant \mathbb{P}\left[\max_{j\neq i} \sup_{T^\delta_\ell \leqslant s \leqslant (T^\delta_\ell+\delta^\xi)\wedge \bar{t}^\delta_\ell(i,\sqrt{\varepsilon}/6)\wedge(T+1)} |\lambda^\delta_j(s) - \lambda^\delta_j(T^\delta_\ell)| \geqslant \frac{\sqrt{\varepsilon}}{12} \mid \mathcal{F}_{T^\delta_\ell}\right]\\ &\leqslant C\exp(-\frac{c\varepsilon^2}{\delta^\xi}),\end{aligned}$$

by Lemma 5.11. As for all $\xi > 0$, $\exp(-\frac{c}{\delta^{\xi/4}}) \ll \delta^{1-p\beta}$ for small enough δ, the proof is complete. □

5.4 Properties of the eigenvalues of M^β_n

In this section, we will study the regularity and boundedness properties of the eigenvalues of M^β_n.

Definition 5.19. *Let M^β_0 be a symmetric (resp. Hermitian) matrix if $\beta = 1$ (resp. $\beta = 2$) with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_d$ and $(M^\beta_n(t))_{t\geqslant 0}$ be the matrix process defined in Definition 5.1. For all $t \geqslant 0$, the ordered eigenvalues of the matrix $M^\beta_n(t)$ will be denoted by $\lambda^n_1(t) \leqslant \lambda^n_2(t) \leqslant \ldots \leqslant \lambda^n_d(t)$.*

The following proposition characterizes the evolution of the process $\lambda^n(t)$ until its first collision time.

Proposition 5.20. *Let $(\lambda^n_1(t), \ldots, \lambda^n_d(t))$ be the process defined in Definition 5.19 and set $T_n(1) := \inf\{t \geqslant 0 : \exists i \neq j, \lambda^n_i(t) = \lambda^n_j(t)\}$. Then, almost surely, the process $(\lambda^n_1(t), \ldots, \lambda^n_d(t))$ verifies for every $k \in \mathbb{N}$, the following strict inequality*

$$\lambda^n_1(k/n) < \lambda^n_2(k/n) < \cdots < \lambda^n_d(k/n). \tag{5.33}$$

In addition, there exist a sequence of Bernoulli random variables $(\epsilon_k^n)_{k\in\mathbb{N}}$ with mean p and a sequence of independent (standard) Brownian motions $(b_t^i)_{t\geqslant 0}$, $i\in\{1,\dots,d\}$ also independent of the Bernoulli random variables $(\epsilon_k^n)_{k\in\mathbb{N}}$ such that, the process $(\lambda_1^n(t),\dots,\lambda_d^n(t))_{t\geqslant 0}$ is the re-ordering of the process $(\mu_1^n(t),\dots,\mu_d^n(t))_{t\geqslant 0}$ defined for $t\geqslant 0$ by

$$d\mu_i^n(t) = -\gamma\mu_i^n(t)\,dt + \sqrt{2}db_t^i + \beta\sum_{j\neq i}\frac{\epsilon_t^n}{\mu_i^n(t)-\mu_j^n(t)}\,dt\,. \tag{5.34}$$

with initial conditions in $t=0$ given by $(\mu_1^n(0),\dots,\mu_d^n(0)) = (\lambda_1,\dots,\lambda_d)$. In particular, up to time $T_n(1)$, the process λ^n verifies

$$d\lambda_i^n(t) = -\gamma\lambda_i^n(t)\,dt + \sqrt{2}db_t^i + \beta\sum_{j\neq i}\frac{\epsilon_t^n}{\lambda_i^n(t)-\lambda_j^n(t)}\,dt\,.$$

Remark here that we use the property that $\epsilon_t^n = (\epsilon_t^n)^2$.

Proof. Let us show first that for each $k\in\mathbb{N}$ such that $k/n < T_n(1)$, we have almost surely the strict inequality (5.33). We will proceed by induction over k. Note that under our assumptions, it is true for $k=0$. Suppose it is true at rank k and let us show it is then true at rank $k+1$. From Definition 5.1, if the eigenvalues of $M_n^\beta(k/n)$ are denoted as $\lambda_1^n(k/n) < \cdots < \lambda_d^n(k/n)$, then, depending on the value of the Bernoulli random variable ϵ_k^n, the dynamic for $t\in[k/n;(k+1)/n]$ is

- if $\epsilon_k^n = 1$, the process $(\lambda_1^n(t),\dots,\lambda_d^n(t))$ follows the Dyson Brownian motion with initial conditions $(\lambda_1^n(k/n),\dots,\lambda_d^n(k/n))$ (see [15, Theorem 4.3.2]); More precisely, we have for $t\in[k/n;(k+1)/n)$

$$d\lambda_i^n(t) = -\gamma\lambda_i^n(t)\,dt + \sqrt{2}dW_t^i + \beta\sum_{j\neq i}\frac{dt}{\lambda_i^n(t)-\lambda_j^n(t)}\,.$$

  where the $(W_t^i)_{t\geqslant 0}, i\in\{1,\dots,d\}$ are independent Brownian motions. In particular, this process is non-colliding in the sense that the $\lambda_i^n(t)$ will almost surely remain strictly ordered for all $t\in[k/n;(k+1)/n)$ (see [15, Theorem 4.3.2]). Thus, we will almost surely have $\lambda_1^n((k+1)/n) < \cdots < \lambda_d^n((k+1)/n)$.

- on the other hand, if $\epsilon_k^n = 0$, we need to define a new process $(\mu_1^n(t),\dots,\mu_d^n(t))$ of independent Ornstein-Uhlenbeck processes with initial conditions $(\lambda_1^n(k/n),\dots,\lambda_d^n(k/n))$; More precisely, the evolution for $t\in[k/n;(k+1)/n]$ is given by

$$d\mu_i^n(t) = -\gamma\mu_i^n(t)dt + \sqrt{2}dB_t^i \tag{5.35}$$

  where the Brownian motions B^i are the ones of Definition 5.1. Note that, before time $T_n(1)$, the two processes λ^n and μ^n coincide. In this case, the $\mu_i^n(t)$ can cross and the ordering can be broken in the interval $[k/n;(k+1)/n]$. However, if crossing for the process μ^n happen before time $t=(k+1)/n$ still we know that $e^{\gamma(k+1)/n}\mu_i^n((k+1)/n)$ are almost surely distinct. The re-ordering of the μ_i^n thus always gives $\lambda_1^n((k+1)/n) < \cdots < \lambda_d^n((k+1)/n)\,a.s.$

The induction is complete and proves equality (5.33) for all $k\in\mathbb{N}$. We deduce from the above arguments that for k such that $k/n < T_n(1)$, the evolution of $\lambda^n(t)$ for $t\in[k/n;(k+1)/n\wedge T_n(1))$ is

$$d\lambda_i^n(t) = -\gamma\lambda_i^n(t)\,dt + \sqrt{2}(\epsilon_t^n dW_t^i + (1-\epsilon_t^n)dB_t^i) + \beta\sum_{j\neq i}\frac{\epsilon_t^n}{\lambda_i^n(t)-\lambda_j^n(t)}\,dt\,.$$

with initial conditions in $t=k/n$ given by $(\lambda_1^n(k/n),\dots,\lambda_d^n(k/n))$. Let us define the process b^i for $t\geqslant 0$ by $b_t^i := \int_0^t(\epsilon_s^n dW_s^i + (1-\epsilon_s^n)dB_s^i)$. Using the fact that the Brownian motions $(W_t^i)_{t\geqslant 0}, i\in\{1,\dots,d\}$ are mutually independent and independent of the Brownian motions $(B_t^i)_{t\geqslant 0}, i\in\{1,\dots,d\}$ (also mutually independent), it is straightforward to check that the

processes $(b_t^i)_{t\geqslant 0}, i\in\{1,\dots,d\}$ are mutually independent Brownian motions. It is also easy to see that, for all $s,t\in[k/n;(k+1)/n]$, the random variables $\epsilon_k^n(W_t^i-W_s^i)+(1-\epsilon_k^n)(B_t^i-B_s^i)$ and ϵ_k^n are independent. Therefore, we deduce that the brownian motions $(b_t^i)_{t\geqslant 0}, i\in\{1,\dots,d\}$ are independent of the sequence $(\epsilon_k^n)_{k\in\mathbb{N}}$. □

The following regularity properties will be useful later on.

Lemma 5.21. *Let $T<\infty$. Then there exist constants $C, A_0, c, c', \alpha>0$ which depend only on T, d such that for all $n\in\mathbb{N}$, all $A\geqslant A_0$ and all $\varepsilon>0$*

$$\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{0\leqslant t\leqslant T}|M_n^\beta(t)_{ij}|>A\right]\leqslant C\exp(-\alpha A^2)\,, \tag{5.36}$$

$$\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{\substack{0\leqslant s,t\leqslant T,\\ |t-s|\leqslant\delta}}|M_n^\beta(t)_{ij}-M_n^\beta(s)_{ij}|>\varepsilon\right]\leqslant\frac{c}{\delta}\exp(-\frac{\varepsilon^2}{c'\delta})\,. \tag{5.37}$$

Proof. Using Itô's formula, we can check that

$$e^{\gamma t}M_n^\beta(t)-e^{\gamma s}M_n^\beta(s)=\int_s^t e^{\gamma s}\left(\epsilon_s^n dH_s^\beta+(1-\epsilon_s^n)\sqrt{2}\sum_{i=1}^d\chi_i^n(\frac{[ns]}{n})dB_s^i\right).$$

Let us set $\Delta_n(s,t):=e^{\gamma t}M_n^\beta(t)-e^{\gamma s}M_n^\beta(s)$. The entries of $\Delta_n(s,.)$ are martingales with respect to the filtration of the Brownian motions conditionally to the Bernoulli random variables $(\epsilon_k^n)_{k\in\mathbb{N}}$ (this is due to the independence between the Brownian motions $(B_t^i)_{t\geqslant 0}, (H_t^\beta(ij))_{t\geqslant 0}, 1\leqslant i,j\leqslant d$ and the sequence of Bernoulli random variables $(\epsilon_k^n)_{k\in\mathbb{N}}$. Using the fact that $|\chi_i^n([ns]/n)_{ij}|\leqslant 1$ for all i,j, we can check that there exists a constant $C(d,T)$ which does not depend on n such that for all $n\in\mathbb{N}$

$$|\langle\Delta_n(s,\cdot)_{ij},\Delta_n(s,\cdot)_{kl}\rangle_t|\leqslant C(T,d)|t-s|\,.$$

Let $A>0$, using [15, corollary H.13], we have

$$\begin{aligned}
&\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{0\leqslant t\leqslant T}|(e^{\gamma t}M_n^\beta(t))_{ij}|>A\right]\\
&\leqslant d^2\max_{1\leqslant i,j\leqslant d}\mathbb{P}\left[\sup_{0\leqslant t\leqslant T}|(e^{\gamma t}M_n^\beta(t)-M_0^\beta)_{ij}|>A-\max_{i,j}|M_0^\beta(i,j)|\right]\\
&=d^2\max_{1\leqslant i,j\leqslant d}\mathbb{P}\left[\sup_{0\leqslant t\leqslant T}|\Delta_n(0,t)_{ij}|>A-\max_{i,j}|M_0^\beta(i,j)|\right]\\
&\leqslant d^2\exp\left(-\frac{(A-\max_{i,j}|M_0^\beta(i,j)|)^2}{C(d,T)T}\right).
\end{aligned}\tag{5.38}$$

Similarly, for any given $s\in[0,T]$, for $\varepsilon>0$, using [15, Corollary H.13], we have, for each entry ij and for every $\delta>0$:

$$\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{t\in[s-\delta,s+\delta]}|(e^{\gamma t}M_n^\beta(t)-e^{\gamma s}M_n^\beta(s))_{ij}|>\varepsilon\right]\leqslant d^2\exp\left(-\frac{\varepsilon^2}{2C\delta}\right).$$

and therefore there exists a positive constant c' so that

$$\begin{aligned}
&\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{\substack{0\leqslant s,t\leqslant T,\\ |t-s|\leqslant\delta}}|(e^{\gamma t}M_n^\beta(t)-e^{\gamma s}M_n^\beta(s))_{ij}|>\varepsilon\right]\\
&\leqslant\sum_{i=1}^{[2T/\delta]+1}\mathbb{P}\left[\max_{1\leqslant i,j\leqslant d}\sup_{|t-\frac{i\delta}{2}|\leqslant\delta/2}|(e^{\gamma t}M_n^\beta(t)-e^{\gamma i\delta/2}M_n^\beta(i\delta/2))_{ij}|>\varepsilon/2\right]\\
&\leqslant d^2\frac{2T}{\delta}\exp\left(-\frac{\varepsilon^2}{c'\delta}\right).
\end{aligned}$$

□

Lemma 5.22. *Let $T < \infty$. Then there exist constants $C', A_0, c', c'', \alpha, \epsilon_0 > 0$ which depend only on T, d such that for all $n \in \mathbb{N}$, all $A \geqslant A_0$ and all $\varepsilon > 0$*

$$\mathbb{P}\left[\max_{1 \leqslant i \leqslant d} \sup_{0 \leqslant t \leqslant T} |\lambda_i^n(t)| > A\right] \leqslant C' \exp(-\alpha A^2)\,, \tag{5.39}$$

$$\mathbb{P}\left[\max_{1 \leqslant i \leqslant d} \sup_{\substack{0 \leqslant s,t \leqslant T,\\ |t-s| \leqslant \delta}} |\lambda_i^n(t) - \lambda_i^n(s)| > \varepsilon\right] \leqslant \frac{c''}{\delta} \exp(-\frac{\varepsilon^2}{c'\delta})\,. \tag{5.40}$$

Proof. This lemma is a consequence of Lemma 5.21 and the inequalities

$$\begin{aligned} \max_{1 \leq k \leq d} |\lambda_k^n(t) - \lambda_k^n(s)| &\leqslant \left(\sum_{i=1}^{d} |\lambda_i^n(t) - \lambda_i^n(s)|^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{i,j=1}^{d} |M_n^\beta(t)_{ij} - M_n^\beta(s)_{ij}|^2\right)^{1/2} \\ &\leqslant d \max_{1 \leq i,j \leq d} |M_n^\beta(t)_{ij} - M_n^\beta(s)_{ij}| \end{aligned} \tag{5.41}$$

where, for the second inequality, we used [15, lemma 2.1.19] and the fact that the λ_i^n are ordered. □

5.5 Convergence till the first hitting time

Proposition 5.23. *Take $\lambda(0) = (\lambda_1 < \lambda_2 < \cdots < \lambda_d)$. Construct μ^n, strong solution of (5.34), with the same Brownian motion than λ, strong solution of (5.5), both starting from $\lambda(0)$. λ^n equals μ^n till $T_n(1)$. For all $T > 0$, we have the following almost sure convergence*

$$\lim_{n \to \infty} \max_{1 \leqslant i \leqslant d} \sup_{t \leqslant T \wedge T_n(1) \wedge \tau_\varepsilon^3} |\lambda_i^n(t) - \lambda_i(t)| = 0\,.$$

As a consequence, if we let $T_1 = \inf\{t > 0, \exists i \neq j,\ \lambda_i(t) = \lambda_j(t)\}$, we have almost surely

$$T_1 \leqslant \liminf T_n(1)\,.$$

We point out that this convergence does not happen on a trivial interval since we have

Remark. For any $\eta > 0$, there exists $\tau(\eta) > 0$ so that

$$\lim_{n \to \infty} \mathbb{P}\left[T_n(1) \geqslant \tau(\eta)\right] \geqslant 1 - \eta\,.$$

Proof of Remark 5.5. By the same arguments developed in (5.41), we find that

$$\begin{aligned} \mathbb{P}\left[\sup_{t \leqslant T} \max_{1 \leqslant i \leqslant d} |\lambda_i^n(t)e^{\gamma t} - \lambda_i(0)| \geqslant \epsilon\right] &\leqslant \mathbb{P}\left[\sup_{t \leqslant T} |\mathrm{tr}((M^n(t)e^{\gamma t} - M_0)^2)| \geqslant \epsilon^2\right] \\ &\leqslant d^2 \exp(-\frac{\epsilon^2}{2C(d,T)T})\,. \end{aligned}$$

But since also the λ_i^n are uniformly bounded with high probability, we can choose for any $\eta > 0$ the parameter T small enough so that

$$\mathbb{P}\left[\max_{1 \leqslant i \leqslant d} \sup_{t \leqslant T} |\lambda_i^n(t) - \lambda_i(0)| \geq \min_{1 \leqslant i \leqslant d} |\lambda_i - \lambda_{i+1}|/3\right] \leqslant \eta$$

This implies that $P(T_n(1) \leq T) \leq \eta$. □

Proof of Proposition 5.23 Using Itô's formula, we can compute

$$\sum_{i=1}^{d}(\lambda_i^n(t)-\lambda_i(t))^2 = -2\gamma\int_0^t\sum_{i=1}^{d}(\lambda_i^n(s)-\lambda_i(s))^2\,ds \tag{5.42}$$
$$+2\beta\int_0^t\epsilon_s^n\sum_{i=1}^{d}\sum_{j\neq i}(\lambda_i^n(s)-\lambda_i(s))\left(\frac{1}{\lambda_i^n(s)-\lambda_j^n(s)}-\frac{1}{\lambda_i(s)-\lambda_j(s)}\right)ds$$
$$+2\beta\int_0^t(\epsilon_s^n-p)\sum_{i=1}^{d}\sum_{j\neq i}\frac{\lambda_i^n(s)-\lambda_i(s)}{\lambda_i(s)-\lambda_j(s)}ds\,.$$

By the same argument as in (5.22) the second term in the right hand side is non positive. Thus using equations 5.42, we find for $t\leqslant T_n(1)$

$$\sum_{i=1}^{d}(\lambda_i^n(t)-\lambda_i(t))^2 \leqslant 2\beta\int_0^t(\epsilon_s^n-p)\sum_{i=1}^{d}\sum_{j\neq i}\frac{\lambda_i^n(s)-\lambda_i(s)}{\lambda_i(s)-\lambda_j(s)}ds := R_n(t)\,.$$

We next prove that

$$\lim_{n\to\infty}\sup_{0\leqslant t\leqslant T\wedge\tau_\varepsilon^3}R_n(t)=0\quad a.s. \tag{5.43}$$

Write $R_n(t)$ as $R_n(t)=P_n(t)+Q_n(t)$ where

$$P_n(t) := \int_0^t(\epsilon_s^n-p)\sum_{i=1}^{d}\sum_{j\neq i}\frac{\lambda_i^n([ns]/n)-\lambda_i(s)}{\lambda_i(s)-\lambda_j(s)}ds\,,$$
$$Q_n(t) := \int_0^t(\epsilon_s^n-p)\sum_{i=1}^{d}\sum_{j\neq i}\frac{\lambda_i^n(s)-\lambda_i^n([ns]/n)}{\lambda_i(s)-\lambda_j(s)}ds\,.$$

We first handle the convergence of $Q_n(t)$. Set $\Omega_1=\{\sup_{\substack{|s-t|\leq 1/n\\ t\leq T}}\max_{1\leq i\leq d}|\lambda_i^n(t)-\lambda_i^n(s)|\leqslant n^{-1/2+\epsilon}\}$. On the event Ω_1, we have

$$|Q_n(t)|\leq n^{-1/2+\epsilon}\sum_{i=1}^{d}\sum_{j\neq i}\int_0^t\frac{ds}{|\,\lambda_i(s)-\lambda_j(s)\,|}\,.$$

Following (5.41), we know that

$$P(\Omega_1^c)\leq ce^{-cn^{2\epsilon}}\,.$$

We thus deduce from Lemma 5.7 that

$$\mathbb{P}\left[\sup_{t\leqslant T}|Q_n(t)|>\delta\right]\leqslant\mathbb{P}\left[\sum_{i=1}^{d}\sum_{j\neq i}\int_0^T\frac{ds}{|\,\lambda_i(s)-\lambda_j(s)\,|}>\delta n^{1/2-\epsilon}\right]+\mathbb{P}\left[\Omega_1^c\right]$$
$$\leqslant c\,e^{-c\,\delta^2\,n^{1-2\epsilon}}+c\,e^{-c\,n^{2\epsilon}}\,.$$

Hence, Borel Cantelli's Lemma insures the almost sure convergence of Q_n to zero. We now turn to the convergence of $P_n(t)$. Let $\eta>0$ small and write

$$P_n(t)=-\frac{d(d-1)}{2}\int_0^t(\epsilon_s^n-p)ds+\widetilde{P}_n(t)$$

with

$$\widetilde{P}_n(t)=\int_0^t(\epsilon_s^n-p)\sum_{i=1}^{d}\sum_{j<i}\frac{\lambda_i^n([ns]/n)-\lambda_j^n([ns]/n)}{\lambda_i(s)-\lambda_j(s)}ds\,.$$

The process $\int_0^t(\epsilon_s^n-p)ds$ is a martingale and by Azuma-Hoeffding inequality, for any $\delta>0$

$$\mathbb{P}\left(\max_{t\leq T}|\int_0^t(\epsilon_s^n-p)ds|\geq\delta\right)\leq 2\exp(-\frac{\delta^2 n}{2})\,.$$

We now use the independence between the brownian motions $(b_t^i)_{0\leqslant t\leqslant T}, i=1,\dots,d$ and the Bernoulli random variables $\epsilon_k^n, k=1,\dots,[nT]$. Conditionally on the $(b_t^i)_{0\leqslant t\leqslant T}, i=1,\dots,d$, the processes $\lambda_i(t), i=1,\dots,d$ are deterministic and the process $\widetilde{P}_n$ is a martingale with respect to the filtration of the ϵ_k^n. We let

$$A_k^n = \sum_{i=1}^{d}\sum_{j<i}\int_{k/n}^{k+1/n}\frac{\lambda_i^n([ns]/n)-\lambda_j^n([ns]/n)}{\lambda_i(s)-\lambda_j(s)}ds.$$

By Lemma 5.11 and Lemma 5.22, the set

$$\Omega=\{\sup_{k\leq nT\wedge\tau_\epsilon^3}|A_k^n|\leq n^{-1/8}\}$$

has probability larger than $1-e^{-cn^{1/16}}$. Moreover, by martingale property it is easy to see that for all $\lambda\geq 0$,

$$\mathbb{E}[1_\Omega e^{\lambda\widetilde{P}_n(k/n)-\frac{1}{2}\lambda^2\sum_{\ell=0}^{k-1}(A_{k/n}^n)^2}]\leq 1\,.$$

Taking $\lambda=n^{1/16}$, since on Ω, $-n^{1/16}|A_k^n|+n^{1/8}|A_k^n|^2/2\leq 0$, Tchebychev's inequality yields

$$\mathbb{P}\left(\{|\widetilde{P}_n(k/n\wedge\tau_\varepsilon^3)|\geq n^{-1/16}(\sum_{\ell=0}^{[Tn]}|A_k^n|+t)\}\cap\Omega\right)\leq e^{-t}$$

As by Lemma 5.7, $\sum_{\ell=0}^{[Tn]}|A_k^n|$ is bounded by $n^{1/32}$ with probability greater than $1-e^{-n^{1/16}}$ we conclude that

$$\mathbb{P}\left(|\widetilde{P}_n(k/n\wedge\tau_\varepsilon^3)|\geq n^{-1/32}\right)\leq Ce^{-n^{1/32}}\,.$$

The uniform estimate is obtained easily by controlling the increments of $\widetilde{P}_n$ in between the times $k/n, k\leq[nT]$ by $\sup_{k\leq[nT]}|A_k^n|$ which we have already bounded.

□

5.6 Proof of Theorem 5.2

5.6.1 Non colliding case $p\beta\geqslant 1$

It is straightforward to deduce Theorem 5.2 when $p\beta\geqslant 1$. Indeed if $\beta p\geqslant 1$ we know that there are no collisions for the limiting process and more precisely, see e.g [15, p. 252],

$$\mathbb{P}(\tau_\varepsilon^2\leqslant T)\leqslant c(\lambda_0)T/|\log\varepsilon|$$

with some finite constant $c(\lambda_0)$ which only depends on the spacings of the eigenvalues at the initial time. This implies in particular that

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\mathbb{P}(T_\varepsilon^n\leqslant T)=0$$

from which we easily deduce Theorem 5.2 from Proposition 5.23.

5.6.2 Colliding case $p\beta<1$

We now define the process $(\lambda_i^{n,\delta}(t))_{t\geqslant 0}$ which will depend on the sequence $(T_\ell^\delta)_{\ell\in\mathbb{N}}$ defined in Definition 5.12. To unify notations, set $T_1^\delta:=T_1$ and $T_n^\delta(1):=T_n(1)$.

Definition 5.24. *For $t<T_1^\delta$, set $\lambda_i^{n,\delta}(t):=\lambda_i^n(t)$. For time $t>T_1^\delta$, we define the process recursively by setting for each $\ell\geqslant 1$, $\lambda_i^{n,\delta}(T_\ell^\delta)=\lambda_i^{n,\delta}(T_\ell^\delta-)+i\delta$ for all $i\in\{1,\dots,d\}$ and for $t>T_\ell^\delta$, the process $\lambda_i^{n,\delta}$ is defined up to time $T_{\ell+1}^\delta$ by ordering the process $(\mu_1^{n,\delta}(t),\dots,\mu_d^{n,\delta}(t))_{T_\ell^\delta\leqslant t\leqslant T_{\ell+1}^\delta}$ which is defined for $t\geqslant T_\ell^\delta$ as*

$$d\mu_i^{n,\delta}(t)=-\gamma\mu_i^{n,\delta}(t)\,dt+\sqrt{2}db_t^i+\beta\sum_{j\neq i}\frac{\epsilon_t^n}{\mu_i^{n,\delta}(t)-\mu_j^{n,\delta}(t)}\,dt\,. \tag{5.44}$$

with initial conditions in $t=T_\ell^\delta$ given by $(\lambda_1^{n,\delta}(T_\ell^\delta),\dots,\lambda_d^{n,\delta}(T_\ell^\delta))$.

Lemma 5.25. *Let $T < \infty$ and $\delta > 0$. We have the following convergence in probability, for all $\ell \in \mathbb{N}$,*

$$\lim_{n\to\infty} \max_{1 \leqslant i \leqslant d} \sup_{0 \leqslant t \leqslant T_\ell^\delta \wedge T} |\lambda_i^\delta(t) - \lambda_i^{n,\delta}(t)| = 0\,.$$

In particular, for every ℓ, if T_n^δ is the first collision time for $\lambda^{n,\delta}$ after $T_{\ell-1}^\delta$,

$$T_\ell^\delta \wedge T \leqslant \liminf T_n^\delta(\ell) \wedge T \quad a.s.$$

Proof Again, we prove this Lemma by induction over ℓ.

• We begin with the case $\ell = 1$. Proposition 5.23 yields that the random variable $\max_{1 \leqslant i \leqslant d} \sup_{0 \leqslant t \leqslant T_n(1)\wedge T} \lambda_i^n(t)| = 0$ converges to 0 in probability as by Lemma 5.8, $P(\tau_\varepsilon^3 \geq T)$ goes to one as ε vanishes. Since we have the almost sure inequality $T_1^\delta \leqslant \liminf T_n^\delta(1)$, the continuity of the $\lambda_i, 1 \leqslant i \leqslant d$, the regularity property of the λ_i^n given by Lemma 5.22, Lemma 5.11 and Proposition 5.23, we can check that since before T_1^δ $\lambda_i^\delta = \lambda_i$ and $\lambda_i^{n,\delta} = \lambda_i^n$, if $T_n^\delta(1) < T_1^\delta \wedge T$,

$$\max_{1 \leqslant i \leqslant d} \sup_{T_n^\delta(1) \leqslant t < T_1^\delta \wedge T} |\lambda_i^\delta(t) - \lambda_i^{n,\delta}(t)| \tag{5.45}$$

$$\leqslant \max_{1 \leqslant i \leqslant d} \sup_{T_n^\delta(1) \leqslant t < T_1^\delta \wedge T} \{|\lambda_i^n(t) - \lambda_i^n(T_n^\delta(1))| + |\lambda_i(t) - \lambda_i(T_n^\delta(1))|\} \tag{5.46}$$

$$+ |\lambda_i^n(T_n^\delta(1)) - \lambda_i(T_n^\delta(1))|$$

goes to zero in probability, when n goes to infinity.

• Suppose the property is true for ℓ and let us show that it is then true for $\ell+1$. By the same argument as in the proof of Proposition 5.23, we can show that, for all $t \in [T_\ell^\delta; T_n^\delta(\ell+1) \wedge T_{\ell+1}^\delta]$, we have

$$\sum_{i=1}^d \left(\lambda_i^{n,\delta} - \lambda_i^\delta\right)^2 (t) \leqslant \sum_{i=1}^d \left(\lambda_i^{n,\delta} - \lambda_i^\delta\right)^2 (T_\ell^\delta) \tag{5.47}$$

$$+ 2\beta \int_{T_\ell^\delta}^t (\epsilon_s^n - p) \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i^{n,\delta}(s) - \lambda_i^\delta(s)}{\lambda_i^\delta(s) - \lambda_j^\delta(s)} ds.$$

The same proof as in Proposition 5.23 shows that, if $\tau_\epsilon^{3,\ell}$ is the stopping time τ_ℓ^3 for the process $\lambda^\delta(t), t \geq T_\ell^\delta$,

$$\lim_{n\to\infty} \sup_{t\in[T_\ell^\delta; T_n^\delta(\ell+1)\wedge T_{\ell+1}^\delta \wedge \tau_\epsilon^{3,\ell}]} \int_{T_\ell^\delta}^t (\epsilon_s^n - p) \sum_{i=1}^d \sum_{j\neq i} \frac{\lambda_i^{n,\delta}(s) - \lambda_i^\delta(s)}{\lambda_i^\delta(s) - \lambda_j^\delta(s)} ds = 0 \quad a.s. \tag{5.48}$$

Thus, because of (5.47), the following convergence in holds

$$\lim_{n\to\infty} \max_i \sup_{t\in[T_\ell^\delta; T_n^\delta(\ell+1)\wedge T_{\ell+1}^\delta \wedge \tau_\epsilon^3]} |\lambda_i^{n,\delta}(t) - \lambda_i^\delta(t)| = 0 \quad a.s\,. \tag{5.49}$$

Because of (5.49), we have $T_{\ell+1}^\delta \wedge \tau_\epsilon^3 \leqslant \liminf_{n\to\infty} T_n^\delta(\ell+1) \wedge \tau_\epsilon^3$. Since the probability that τ_ϵ^3 is larger than T goes to one as ϵ vanishes, we can show as in (5.45) (note that Lemma 5.22, Lemma 5.11 and Proposition 5.23 extend to $\{\lambda_t^{n,\delta}, \lambda_t^\delta, t \geq T_\ell^\delta\}$) that in probability,

$$\lim_{n\to\infty} \max_{1 \leqslant i \leqslant d} \sup_{T_n^\delta(\ell+1) \leqslant t \leqslant T_{\ell+1}^\delta} |\lambda_i^\delta(t) - \lambda_i^{n,\delta}(t)| = 0.$$

The property at rank $\ell + 1$ is established. The Lemma is proved. □

Lemma 5.26. *There exists a constant $c > 0$ such that for all $L \in \mathbb{N}$, we have the following almost sure estimate*

$$\max_{1 \leqslant j \leqslant d} \sup_{0 \leqslant t \leqslant T_L^\delta} |\lambda_j^{n,\delta}(t) - \lambda_j^n(t)| \leqslant \delta\, L\, \sqrt{c}\,.$$

Proof. Note that the estimate is striaghtforward on $[0, T_1^\delta]$. We then proceed by induction on the time intervals $[T_\ell^\delta, T_{\ell+1}^\delta]$ as in the proof of Lemma 5.14 until the first collision time

$$t_1 := \inf\{t \geqslant T_\ell^\delta \ : \exists i,\ \lambda_i^n(t) = \lambda_{i-1}^n(t) \text{ or } \lambda_i^{n,\delta}(t) = \lambda_{i-1}^{n,\delta}(t)\}\,.$$

We next claim that, at a given time, almost surely the eigenvalues λ^n are different. Indeed, this is clear if the eigenvalues follows Brownian motion and even more when they follow Dyson Brownian motion. Moreover the probability that more than two eigenvalues collide at some time vanishes. Indeed, this can only happen if the eigenvalues follow the Brownian motion. But the probability that 3 Brownian motions collide vanishes and hence the result.

Hence, there are almost surely at most two eigenvalues which can collide. Hence, let $i(t_1)$ be the unique integer in $\{1, \dots, d\}$ such that $\lambda_i^n(t_1) = \lambda_{i-1}^n(t_1)$ (respectively $\lambda_i^{n,\delta}(t_1) = \lambda_j^{n,\delta}(t_1)$) and let $\tau_1 = ([nt_1]+1)/n$. Notice that, for $t \in [[nt_1]/n; ([nt_1]+1)/n)$, we necessarily have $\epsilon_t^n = 0$. Let $\mu_i^{n,\delta}$ and μ_i^n for $i \in \{1, \dots, d\}$ be the processes such that for $t \in [t_1; \tau_1]$

$$\begin{aligned} d\mu_i^{n,\delta}(t) &= -\gamma\mu_i^{n,\delta}(t)dt + \sqrt{2}db_t^i \\ d\mu_i^n(t) &= -\gamma\mu_i^n(t)dt + \sqrt{2}db_t^i \end{aligned}$$

with initial conditions at $t = t_1$ respectively given by $\mu^{n,\delta}(t_1) = \lambda^{n,\delta}(t_1)$ and $\mu^n(t_1) = \lambda^n(t_1)$. We know that the $\lambda_i^{n,\delta}$, respectively the λ_i^n, are just a re-ordering of the processes $\mu_i^{n,\delta}$ and μ_i^n

By definition, for $t \in [t_1; \tau_1]$, we find that :

$$(\mu_j^{n,\delta} - \mu_j^n)(t) = e^{-\gamma(t-t_1)}(\mu_j^{n,\delta} - \mu_j^n)(t_1)\,.$$

As a consequence, we deduce that

$$\sum_{j=1}^d (\mu_j^{n,\delta} - \mu_j^n)^2(t) \leqslant \sum_{j=1}^d (\lambda_j^{n,\delta} - \lambda_j^n)^2(t_1)\,.$$

Moreover, as the λ's are ordered but the set of the values of the λ's and the μ's are the same, using for instance [15, lemma 2.1.19], we have that

$$\sum_{j=1}^d (\lambda_j^{n,\delta} - \lambda_j^n)^2(t) \leqslant \sum_{j=1}^d (\mu_j^{n,\delta} - \mu_j^n)^2(t)\,.$$

Gathering the above inequalities, we have shown that

$$\sup_{t\in[0,\tau_1]} \sum_{j=1}^d (\lambda_j^{n,\delta} - \lambda_j^n)^2(t) \leqslant \sum_{j=1}^d (\lambda_j^{n,\delta} - \lambda_j^n)^2(T_\ell^\delta)\,.$$

We can continue inductively until we reach the time $T_{\ell+1}^\delta$ to finish the proof.

5.7 Asymptotic properties of the eigenvectors

Recall that $w_{ij}^\beta, i < j$ are real (respectively complex) standard Brownian motions if $\beta = 1$ (resp. $\beta = 2$) with quadratic variation βt and that we also set for $i < j$, $w_{ji}^\beta := \bar{w}_{ij}^\beta$. In addition we also defined the skew Hermitian matrix $R^\beta = -(R^\beta)^*$ by setting for $i < j$,

$$dR_{ij}^\beta(t) = \frac{dw_{ij}^\beta(t)}{\lambda_i^n(t) - \lambda_j^n(t)}, \quad R_{ij}^\beta(0) = 0\,.$$

Proof of Proposition 5.3

It is classical to check that the unique strong solution of the stochastic differential equation

$$dO_n^\beta(t) = \epsilon_t^n O_n^\beta(t)dR^\beta(t) - \frac{\epsilon_t^n}{2} O_n^\beta(t)d\langle (R^\beta)^*, R^\beta\rangle_t\,, \tag{5.50}$$

with initial condition $O_n^\beta(0) := O^\beta(0)$ (defined at the end of Section 5.1), is in the space $\mathcal{O}_d^\beta$ for all time t (see e.g. [15, Lemma 4.3.4]) and is such that, with $\Delta_n^\beta(t)$ being the diagonal matrix of the ordered (as in (5.4)) eigenvalues of $M_n^\beta(t)$, we have

$$O_n^\beta(t)\Delta_n^\beta(t)O_n^\beta(t)^* \overset{law}{=} M_n^\beta(t)\,.$$

The law of the continuous process O_n^β is uniquely determined as the unique strong solution of (5.50). □

One can thus define the eigenvectors of $M_n^\beta(t)$, denoted as $\phi_i^n(t)$, so that they satisfy the stochastic differential system

$$d\phi_i^n(t) = \epsilon_t^n \sum_{j\neq i} \frac{dw_{ij}^\beta(t)}{\lambda_i^n(t)-\lambda_j^n(t)}\phi_j^n(t) - \frac{\epsilon_t^n}{2}\sum_{j\neq i}\frac{\beta}{(\lambda_i^n(t)-\lambda_j^n(t))^2}dt\phi_i^n(t) \tag{5.51}$$

where $w_{ij}^\beta, i<j$ is a family of i.i.d. Brownian motions (on $\mathbb{R}$ if $\beta=1$, $\mathbb{C}$ if $\beta=2$), independent of the eigenvalues $\lambda_i^n, 1\leqslant i\leqslant d$.

Proof of Theorem 5.4

This proof is classical and uses the theory of stability for stochastic differential equations.

For $\eta>0$ fixed, we deduce from Proposition 5.23 and Lemma 5.8 that the process $(\lambda_1^n(t),\dots,\lambda_d^n(t))$ converges almost surely in the space of continuous functions $\mathcal{C}([0;(T_1-\eta)\wedge T],\mathbb{R}^d)$ (respectively $\mathbb{C}^d$) if $\beta=1$ (resp. $\beta=2$) endowed with the uniform norm towards $(\lambda_1(t),\dots,\lambda_d(t))_{0\leqslant t\leqslant (T_1-\eta)\wedge T}$ where the λ_i's are the unique strong solutions of (5.5) (with the same Brownian motions b^i) and where T_1 is the first collision time of the $\lambda_i, 1\leqslant i\leqslant d$. In the sequel we will work conditionally to the (λ_i^n,λ_i)'s satisfying the above convergence.

Define for $i\neq j$ the processes $w_{ij}^{\beta,n}$ by setting

$$w_{ij}^{\beta,n}(t) = \int_0^t \epsilon_s^n dw_{ij}^\beta(s)\,. \tag{5.52}$$

Note that the quadratic variation of this continuous martingale converges almost surely towards βpt so that by Rebolledo's theorem $(w_{ij}^{\beta,n}, i<j)$ converges towards $(\sqrt{p}w_{ij}^\beta, i<j)$.

Moreover, if T_1^ϵ is the first time at which two eigenvalues are at distance less than ϵ, the drift coefficients being bounded, we see, with a proof similar to the proof of Proposition 5.23, that for $i\neq j$

$$\int_0^{t\wedge T_1^\epsilon}\frac{\epsilon_s^n}{(\lambda_i^n-\lambda_j^n)^2(s)}ds$$

converges towards $p\int_0^{t\wedge T_1^\epsilon}(\lambda_i(s)-\lambda_j(s))^{-2}ds$ uniformly almost surely. Since T_1^ϵ converges towards T_1 as ϵ goes to zero, the convergence holds till $(T_1-\eta)\wedge T$ for any $\eta>0$.

Gathering the above arguments, the result follows from [86, Theorem 6.9, p. 578]. □

We now turn to the analysis of the behavior of the columns $\phi_i(t)$ of the matrix $O^\beta(t)$ when $t\to T_1$ with $t<T_1$. Those vectors $\phi_i(t)$ form an orthonormal basis of $\mathbb{R}^d$ (respectively $\mathbb{C}^d$) if $\beta=1$ (resp. $\beta=2$) and it is easy to check that they verify the following stochastic differential system

$$d\phi_i(t) = \sum_{j\neq i}\frac{\sqrt{p}}{\lambda_i(t)-\lambda_j(t)}dw_{ij}^\beta(t)\phi_j(t) - \frac{p\beta}{2}\sum_{j\neq i}\frac{dt}{(\lambda_i(t)-\lambda_j(t))^2}\phi_i(t)\,. \tag{5.53}$$

In the following of this section, we will denote by i^* the unique (because of Lemma 5.10) index such that $\lambda_{i^*}(T_1)=\lambda_{i^*-1}(T_1)$.

The main issue we meet at this point in the presence of collisions (that *will* occur if $p\beta<1$; see [54]) lies in the divergence of the integral 5.8 that we now prove.

We now describe the behavior of the $d-2$ vectors $\phi_j(t), j\neq i^*, i^*-1$ just before the first collision time T_1.

Proof of the first statement of Proposition 5.6

We will denote by $\phi_{j\ell}(t)$ the ℓ-th entry of the d-dimensional vector $\phi_j(t)$. For $0\leqslant t<T_1$, we have

$$d\phi_j(t) = \sum_{k\neq j}\frac{\sqrt{p}}{\lambda_j(t)-\lambda_k(t)}dw_{jk}^\beta(t)\phi_k(t) - \frac{p}{2}\sum_{k\neq j}\frac{\beta}{(\lambda_j-\lambda_k)^2}\phi_j(t)dt\,. \tag{5.54}$$

We recall from section 5.3.2 that there are no multiple collisions nor two collisions at the same time for the system $(\lambda_1(t), \lambda_2(t), \dots, \lambda_d(t))_{0 \leqslant t \leqslant T_1}$ verifying (5.5), and therefore we may assume without loss of generality that for $j \neq i^*, i^*-1$, every diffusions and drift terms of (5.54) remains almost surely bounded for $t \in [0; T_1]$. To prove the lemma, we just need to prove that almost surely

$$\lim_{\substack{s \to T_1; \\ s < T_1}} \sup_{s \leqslant t < T_1} \|\phi_j(t) - \phi_j(s)\|_2 = 0\,.$$

The drift terms appearing in (5.54) are obvious to deal with since $1/(\lambda_j - \lambda_k)(t)$ is bounded in the vicinity of T_1 and that $|\phi_{j\ell}(t)| \leqslant 1$ for all $t < T_1$. For the diffusion terms, we have for every $\ell \in \{1, \dots, d\}$ and for every $s \in [0; T_1]$ the following estimate

$$\mathbb{P}\left[\sup_{s \leqslant t < T_1} \Big| \int_s^t \sum_{k \neq j} \frac{\sqrt{p}}{\lambda_j(u) - \lambda_k(u)} dw^\beta_{jk}(u) \phi_{k\ell}(u) \Big| > \eta \right] \leqslant \exp\Big(-\frac{\eta^2}{2\beta p(d-1) M(T_1 - s)}\Big)\,,$$

where $M = \sup_{t \in [0;T_1]} \max_{k \neq j} \frac{1}{(\lambda_j - \lambda_k)^2(t)}$. Using the Borel-Cantelli Lemma, we deduce the result. □

For $\delta > 0$, we want to define a process $(\widetilde{\phi}_1(t), \widetilde{\phi}_2(t), \dots, \widetilde{\phi}_d(t))_{T_1 - \delta \leqslant t < T_1}$ that will be a good approximation of the process $(\phi_1(t), \phi_2(t), \dots, \phi_d(t))_{T_1 - \delta \leqslant t < T_1}$ on the time interval $[T_1 - \delta; T_1]$. Hence for $j \neq i^*, i^* - 1$, we set $\widetilde{\phi}_j(t) = \widetilde{\phi}_j$ (the vectors do not depend of time). It remains to define the evolution for $(\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t))$ that will depend of time t.

Let V be the $(d-2)$-dimensional subspace spanned by the orthonormal family $\{\widetilde{\phi}_j; j \neq i^*, i^* - 1\}$ and W its orthogonal complement in $\mathbb{R}^d$. Let us define the "diffusive orthonormal basis" in the space W that will describe the evolution of the two vectors $(\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t))$ on the interval $[T_1 - \delta; T_1]$ (up to the initial conditions at time $t = T_1 - \delta$ we will explicit later).

Lemma 5.27. *Let $\delta > 0$ and (u, v) an orthonormal basis of the two-dimensional subspace W. We consider the following stochastic differential system*

$$\begin{aligned} d\widetilde{\phi}_{i^*}(t) &= \frac{\sqrt{p}}{(\lambda_{i^*} - \lambda_{i^*-1})(t)} dw^\beta_{i^*-1,i^*}(t)\ \widetilde{\phi}_{i^*-1}(t) - \frac{p\beta}{2} \frac{dt}{(\lambda_{i^*} - \lambda_{i^*-1})^2(t)} \widetilde{\phi}_{i^*}(t)\,, \\ d\widetilde{\phi}_{i^*-1}(t) &= -\frac{\sqrt{p}}{(\lambda_{i^*} - \lambda_{i^*-1})(t)} d\bar{w}^\beta_{i^*-1,i^*}(t)\ \widetilde{\phi}_{i^*}(t) - \frac{p\beta}{2} \frac{dt}{(\lambda_{i^*} - \lambda_{i^*-1})^2(t)} \widetilde{\phi}_{i^*-1}(t) \end{aligned} \tag{5.55}$$

with initial conditions $(\widetilde{\phi}_{i^*-1}(T_1 - \delta), \widetilde{\phi}_{i^*}(T_1 - \delta)) = (u, v)$.

This stochastic differential system has a unique strong solution defined on the interval $[T_1 - \delta; T_1)$ such that for each $t \in [T_1 - \delta; T_1)$, $\{\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t)\}$ is an orthonormal basis of W.

Proof. For all $\epsilon > 0$, the function $t \to 1/(\lambda_{i^*} - \lambda_{i^*-1})(t)$ is bounded on the interval $[T_1 - \delta; T_1^\epsilon]$ and therefore there is a unique strong solution to the stochastic differential system (5.55) till the time T_1^ϵ where $|\lambda_{i^*} - \lambda_{i^*-1}| < \epsilon$ as it is driven by bounded linear drifts. As T_1^ϵ grows to T_1 the proof is complete.

To show that for all $t \in [T_1 - \delta; T_1)$ the family $\{\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t)\}$ is an orthonormal basis of W, we proceed along the same line as in the proof of [15, Lemma 4.3.4]. □

In the following lemma, we show that we can choose a constant $\delta > 0$ small enough and an initial condition $(u, v) \in W$ such that the processes $(\widetilde{\phi}_1(t), \dots, \widetilde{\phi}_1(t))_{t \in [T_1 - \delta; T_1)}$ defined by Lemma 5.27 is indeed a good approximation of the process $(\phi_1(t), \dots, \phi_d(t))_{t \in [T_1 - \delta; T_1)}$. The advantage of the process $(\widetilde{\phi}_1(t), \dots, \widetilde{\phi}_1(t))_{t \in [T_1 - \delta; T_1)}$ is that it is simpler to study in the vicinity of T_1 (see Lemma 5.29 below).

Lemma 5.28. *Let $\eta > 0$ and $\kappa > 0$. Then there exists an orthonormal basis (u, v) of W and $\delta > 0$ small enough such that if we denote by $(\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t))_{t \in [T_1 - \delta; T_1)}$ the unique strong solution of the stochastic differential system (5.55) with initial conditions given in $t_0 = T_1 - \delta$ by $(\widetilde{\phi}_{i^*-1}(t_0), \widetilde{\phi}_{i^*}(t_0)) = (u, v)$, we have*

$$\mathbb{P}\left(\sup_{t \in [t_0; T_1)} ||\phi_{i^*}(t) - \widetilde{\phi}_{i^*}(t)||_2^2 + ||\phi_{i^*-1}(t) - \widetilde{\phi}_{i^*-1}(t)||_2^2 \geqslant \eta\right) \leq \kappa\,.$$

Proof. Using Itô's formula, we find[1] for all $t \in [t_0; T_1)$,

$$||\phi_{i^*}(t) - \widetilde{\phi}_{i^*}(t)||_2^2 + ||\phi_{i^*-1}(t) - \widetilde{\phi}_{i^*-1}(t)||_2^2 = ||\phi_{i^*}(t_0) - u||_2^2 + ||\phi_{i^*-1}(t_0) - v||_2^2 - 2\int_{t_0}^t \sum_{i\in\{i^*,i^*-1\}} \sum_{j\neq i^*,i^*-1} \frac{\sqrt{p}}{(\lambda_i - \lambda_j)(s)} dw_{ij}^\beta(s) \langle \widetilde{\phi}_i(s), \phi_j(s)\rangle \,. \tag{5.56}$$

As for $i \in \{i^*, i^*-1\}$ and $j \notin \{i^*, i^*-1\}$ the terms $1/(\lambda_i - \lambda_j)^2(t)$ have almost surely a finite integral with respect to Lebesgue measure on the interval $[t_0; T_1)$ (in fact those terms are almost surely bounded as the corresponding particles remain at finite distance), the quadratic variation of the last term is of order δ and therefore is smaller than $\eta/2$ with probability greater that $1-\kappa$ for δ small enough.

It remains to check that we can choose (u, v) an orthonormal basis of W and $\delta > 0$ such that

$$||\phi_{i^*}(T_1 - \delta) - u||_2^2 + ||\phi_{i^*-1}(T_1 - \delta) - v||_2^2 \leqslant \eta/2 \,. \tag{5.57}$$

This is a straightforward: Indeed we can approximate the $\phi_j(T_1 - \delta)$ for $j \notin \{i^*, i^* - 1\}$ by the $\widetilde{\phi}_j$ because of the first point of Proposition 5.6, thus we can choose two vectors $\{u, v\}$ in the two dimensional space W so that (5.57) holds. This completes the proof. □

We now turn to the study of the couple $(\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t))$ for $t \in [T_1 - \delta; T_1)$ and in particular when $t \to T_1, t < T_1$. A crucial point is equation 5.8 which we now prove.

Itô's Formula gives for $t < T_1$

$$\ln(\lambda_i - \lambda_{i-1})(t) = (-\gamma + 2p\beta)t + \int_0^t \sqrt{2}\frac{db_s^{i^*} - db_s^{i^*-1}}{(\lambda_{i^*} - \lambda_{i^*-1})(s)} - p\beta \int_0^t \sum_{j\neq i^*,i^*-1} \frac{ds}{(\lambda_{i^*} - \lambda_j)(\lambda_{i^*-1} - \lambda_j)(s)} - \int_0^t \frac{2\,ds}{(\lambda_{i^*} - \lambda_{i^*-1})^2(s)} \,.$$

If we suppose that $\int_0^{T_1} dt/(\lambda_{i^*} - \lambda_{i^*-1})^2(t) < +\infty$ and since $T_1 < \tau_\epsilon^3$ for some $\epsilon > 0$ small enough, we obtain a contradiction letting $t \to T_1$: under this assumption, the right hand side tends to $-\infty$ whereas the left hand side is almost surely bounded in this limit. □

The next Lemma 5.29 shows that the orthonormal basis $(\widetilde{\phi}_{i^*-1}(t), \widetilde{\phi}_{i^*}(t))$ of the subspace W is in fact uniformly distributed in the set of all orthonormal basis of W in the limit $t \to T_1, t < T_1$.

As W is two dimensional, up to a change basis, we can suppose that the two vectors $\widetilde{\phi}_{i^*-1}(t)$ and $\widetilde{\phi}_{i^*}(t)$ are two dimensional (we just study the evolution of their coordinates in an orthonormal basis of W). Let us define the two by two matrix $\widetilde{\phi}(t)$ whose first line is the vector $\widetilde{\phi}_{i^*}(t)$ and second line is the vector $\widetilde{\phi}_{i^*-1}(t)$:

$$\widetilde{\phi}(t) := \begin{pmatrix} \widetilde{\phi}_{i^*}(t) \\ \widetilde{\phi}_{i^*-1}(t) \end{pmatrix} .$$

Lemma 5.29. *The matrix $\widetilde{\phi}(t)$ converges in law when $t \to T_1, t < T_1$ to the Haar probability measure on the orthogonal group (respectively unitary group if $\beta = 2$.)*

Proof. To simplify notations, we do the proof in the case $\beta = 1$.

Set $t_0 := T_1 - \delta$ and define for $t \in [0; \delta)$ the function

$$\varphi(t) := \int_{t_0}^{t_0+t} \frac{ds}{(\lambda_{i^*} - \lambda_{i^*-1})^2(s)}$$

and denote by φ^{-1} its functional inverse. We now proceed to a change of time by setting for $t \in [0; \delta)$

$$\widetilde{\psi}_{i^*}(t) = \widetilde{\phi}_{i^*}(\varphi^{-1}(t)), \quad \widetilde{\psi}_{i^*-1}(t) = \widetilde{\phi}_{i^*-1}(\varphi^{-1}(t)) \,.$$

[1] Note that all the diverging terms in T_1 cancel in this expression.

As $\varphi^{-1}(t) \to +\infty$ when $t \to \delta, t < \delta$ (because of (5.8)), the two by two matrix $\widetilde{\psi}(t)$ whose first line is $\widetilde{\psi}_{i^*}(t)$ and second line is $\widetilde{\psi}_{i^*}(t)$:

$$\widetilde{\psi}(t) := \begin{pmatrix} \widetilde{\psi}_{i^*}(t) \\ \widetilde{\psi}_{i^*-1}(t) \end{pmatrix}$$

is now defined for all $t \in \mathbb{R}_+$ and verifies the following stochastic differential equation

$$d\widetilde{\psi}(t) = \sqrt{p}\, A\, \widetilde{\psi}(t)\, dB_t - \frac{p\beta}{2}\, \widetilde{\psi}(t)\, dt\,. \tag{5.58}$$

where B is a standard Brownian motion on $\mathbb{R}$ and where A is the two by two matrix defined by

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,.$$

Note in particular that $A^2 = -I$.

It is clear that there is pathwise uniqueness in the stochastic differential equation (5.58) (it is linear in $\widetilde{\psi}$). Therefore to solve entirely this equation, we just need to exhibit one solution. Using Itô's Formula, one can check that the solution is

$$\begin{aligned} \widetilde{\psi}(t) &= \exp\left(\sqrt{p}\, A\, B_t\right) \widetilde{\psi}(0) \\ &= \begin{pmatrix} \cos(\sqrt{p}B_t) & \sin(\sqrt{p}B_t) \\ -\sin(\sqrt{p}B_t) & \cos(\sqrt{p}B_t) \end{pmatrix} \widetilde{\psi}(0)\,. \end{aligned}$$

Note that for all $t \in \mathbb{R}_+$, the matrix $\widetilde{\psi}(t)$ is indeed in the space of orthogonal matrices.

But $(\cos(\sqrt{p}B_t), \sin(\sqrt{p}B_t))$ converges in law as time goes to infinity towards the law of $(\theta, \varepsilon\sqrt{1-\theta^2})$ with θ uniformly distributed on $[-1,1]$ and $\varepsilon = \pm 1$ with probability $1/2$, from which the result follows. □

Lemmas 5.28 and 5.29 give the second statement of Proposition 5.6.

Acknowledgments

We thank J.-P. Bouchaud for proposing this subject, and working with us to understand it, which led to [6]; the present article gives complete proofs of some of the results stated in [6]. We are very grateful to C. Garban who proposed us to attack the analysis of the case $p\beta < 1$ by introducing the auxiliary process λ^δ of section 5.6.2 and for many enlightening discussions. We thank L. Dumaz, R. Rhodes, V. Vargas, G . Schehr, R. Chicheportiche and F. Benaych-Georges for useful comments and discussions.

Chapter 6

Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marčenko-Pastur law

Résumé

Cet article est à paraître dans *Journal of Physics A: Mathematical and Theoretical* et est écrit en collaboration avec Jean-Philippe Bouchaud, Satya N. Majumdar et Pierpaolo Vivo. Nous construisons un modèle de matrice diffusante dont l'équilibre stationnaire correspond à des ensembles de matrices aléatoires invariantes par conjugaison orthogonale ou unitaire avec des valeurs propres distribuées suivant les ensembles β-Wishart (ou Laguerre) avec β quelconque appartenant à l'intervalle $[0, 2]$. Dans le régime où le paramètre β dépend de la plus grande dimension M de la matrice des données par la relation $\beta = 2c/M$ (où c est une constante positive), les densités limites des valeurs propres forment une famille de distributions indéxée par le paramètre c qui interpole continument entre la loi de Marčenko Pastur (correspondant au cas où $\beta > 0$ ne dépend pas de M) et la loi Gamma (qui correspond au cas $\beta = 0$). Un prolongement de ce calcul nous permet de trouver les corrections de tailles finies (quand la dimension est grande mais pas infinie) dans le théoreème de Marčenko Pastur.

Abstract

We construct a diffusive matrix model for the β-Wishart (or Laguerre) ensemble for general $\beta \in [0, 2]$, which preserves invariance under the orthogonal/unitary group. Scaling the Dyson index β with the largest size M of the data matrix as $\beta = 2c/M$ (with c a fixed positive constant), we obtain a family of spectral densities interpolating continuously between the Marčenko-Pastur and the Gamma laws as c is varied. Analyzing the full resolvent equation, we obtain as a byproduct the correction to the Marčenko-Pastur density until order $1/M$ for all β and until order $1/M^2$ for the particular cases $\beta = 1, 2$.

6.1 Introduction

The theory of matrices with random entries, originally devised as a tool to understand and predict the spectra of heavy nuclei for which a detailed account of the interactions between particles is too complicated, has seen a spectacular resurgence of interest in recent years, with a number of unexpected and often surprising applications (see [3, 15, 20, 72, 145] for a recent overview). While Wigner and Dyson are usually regarded as the pioneers in the field, John Wishart had already introduced random matrices in 1928 in his studies of multivariate populations [155]. The

Wigner-Dyson (Gaussian) and Wishart ensembles (together with a few others) lie at the core of the *classical* world of invariant matrices, characterized by the following main features:

1. The joint probability distribution (jpd) of matrix entries, collectively denoted by $P[\mathbf{X}]$, remains unaltered if one performs a similarity transformation $\mathbf{X} \to \mathbf{U}\mathbf{X}\mathbf{U}^{-1}$, with $\mathbf{U}$ being orthogonal (real symmetric $\mathbf{X}$), unitary (complex hermitian $\mathbf{X}$) or symplectic (quaternion self-dual $\mathbf{X}$) matrix. As a consequence, the eigenvectors of such matrices are Haar (uniform) distributed in their respective groups.

2. The joint distribution of the N real eigenvalues $P(\lambda_1, \ldots, \lambda_N)$ can be generically written in the Gibbs-Boltzmann form,

$$P(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} \exp\left(-\mathcal{H}(\lambda_1, \ldots, \lambda_N)\right) \tag{6.1}$$

with the Hamiltonian $\mathcal{H}(\lambda_1, \ldots, \lambda_N)$ given by:

$$\mathcal{H}(\lambda_1, \ldots, \lambda_N) = \sum_{i=1}^{N} V(\lambda_i) - \beta \sum_{j<k} \ln|\lambda_j - \lambda_k| \tag{6.2}$$

where $V(x)$ is a confining potential that depends on the precise form of the joint distribution of matrix entries $P[\mathbf{X}]$. For example, if the entries of $\mathbf{X}$ are independent, the only allowed potential is quadratic $V(x) = \beta x^2/2$, which correspond to the Gaussian ensembles. If correlations among the entries are allowed, then different potentials (all corresponding to rotationally invariant weights) are possible. For example, in the Wishart case, $V(x) = \infty$ for $x < 0$ (so that all the eigenvalues are non-negative) and $V(x) = x/(2\sigma^2) - \alpha \log x$ for $x \geq 0$.

The normalization constant Z_N is called the partition function and is simply given by the multiple integral

$$Z_N = \int \cdots \int \prod_i d\lambda_i \exp\left(-\mathcal{H}(\lambda_1, \ldots, \lambda_N)\right). \tag{6.3}$$

From (6.1), one easily deduces that the system of N eigenvalues of a classically invariant ensemble behaves as a thermodynamic system of charged particles arranged on the real line, in equilibrium at inverse temperature β under competing interactions (the confining potential $V(x)$ and the logarithmic all-to-all repulsion term) in (6.2). In contrast with the usual canonical ensemble in statistical mechanics, however, the so-called Dyson index β is quantized and can only assume the values $\beta = 1, 2, 4$ for real symmetric, complex hermitian and quaternion self-dual matrices respectively.

Lifting the quantization of β (Dyson's *threefold way*) has been a major theoretical challenge in view of possible applications e.g. to the quantum Hall effect. Dumitriu and Edelman [68] were eventually able to construct ensembles of tridiagonal matrices with independent entries whose eigenvalues are distributed as (6.1) with general $\beta > 0$. Their ensemble is however *not* invariant under similarity transformations, and the eigenvectors are not Haar distributed in the appropriate symmetry group. After an earlier attempt in the case of 2×2 matrices [149], the explicit construction of an ensemble of $N \times N$ matrices displaying at once rotational invariance and a continuous β was put forward in [6, 10] for the Gaussian ensemble. It was further shown in [6] that only by letting the Dyson index β of that ensemble scale with the matrix size N in an appropriate way (namely $\beta = c/N$) one obtains a continuous family of deformed spectral densities parametrized by c, interpolating between Wigner's semicircle (typical for $\beta \sim \mathcal{O}(1)$ invariant ensembles) and a Gaussian law (properly describing the non-interacting limit $\beta \to 0$). This result can be established in two alternative ways:

1. Starting from the stationary joint distribution of eigenvalues (eq. (6.1)), setting $\beta = c/N$, and then finding the average density of eigenvalues $\rho(\lambda) = (1/N)\langle\sum_i \delta(\lambda - \lambda_i)\rangle$. In the limit of large N, this average density can be obtained by a saddle point analysis of the partition function Eq. (6.3) in a standard way. Usually, when $\beta \sim \mathcal{O}(1)$, only the energy term $\sim \mathcal{O}(N^2)$ dominates and the entropy term $\sim \mathcal{O}(N)$ is subleading. However, when $\beta \sim c/N$, *both* the energy and the entropy terms are of the same order ($\sim \mathcal{O}(N)$), which leads to a nontrivial modification of the density. The complicated nonlinear singular integro-differential equation for the saddle density, reduces very nicely to a Riccati equation for the Stieltjes transform of the density, which can then be subsequently solved exactly. Finally one obtains the density by taking imaginary part of the Stieltjes transform (see section 6.3.3 for a detailed discussion).

2. Starting from the dynamical equation of motion of the eigenvalues, one first derives directly the equation of motion of the Stieltjes transform of the density via Itô's calculus, finds the stationary solution and then obtains the average density by taking the imaginary part of the Stieltjes transform (see detailed discussion in section 6.3.4).

   In section 6.3.3 and 6.3.4, in the context of our model, we show that both methods lead to the same solution.

In this paper, we handle the Wishart case, i.e. the case where the confining potential introduced above is given by $V(x) = x/(2\sigma^2) - \alpha \log x$ for $x \geq 0$. We first explicitly construct a random matrix model which is invariant under similarity transformations and whose jpd of eigenvalues is exactly given by the β-Wishart ensemble of random matrices with a continuous $\beta > 0$. Then, letting the Dyson index β of the ensemble scale inversely with the size of the matrix, we analytically derive the density of states for this crossover model, written in terms of the Whittaker hypergeometric function (see eq. (6.59)), and we show that it continuously interpolates between the Marčenko-Pastur law[1] (corresponding to constant values of $\beta > 0$) and a certain type of Gamma distribution (see subsection 6.3.3). Finally, for any constant Dyson index $\beta > 0$ (and not necessarily quantized to $\beta = 1, 2, 4$), we analyze the full Stieltjes transform equation and we can compute the $1/N$ correction to the Marčenko-Pastur asymptotic density for the β-Wishart ensemble. Furthermore, using results obtained by Pastur and Lytova in [105] on the noise in the Marčenko-Pastur law, we are also able to derive the $1/N^2$ correction term in the particular cases $\beta = 1$ and 2.

The rest of the paper is outlined as follows. In section 6.2 we introduce the main features of the classical Wishart ensemble along with the evolution law for the eigenvalue process. In section 6.3, we construct a 3-parameters matrix model (and the respective evolution law for the eigenvalues) that at large times interpolates between the Wishart ensemble and so-called CIR processes whose stationary pdf is a certain Gamma distribution. The corresponding parametrical density of states is computed exactly in the two ways described above (from the saddle point route on the partition function in section 6.3.3 and from Itô's calculus in section 6.3.4) and constitutes a continuous deformation of the Marčenko-Pastur distribution (see below). In section 6.4 we compute the systematic $1/N$ (for all $\beta > 0$) and $1/N^2$ (for $\beta = 1, 2$) corrections to the Marčenko-Pastur law for the smoothed density in the bulk for the (scaled) β-Wishart ensemble. For the special cases $\beta = 1$, 2 and 4, the $\mathcal{O}(1/N)$ correction term was computed in [75, 73]. Our result generalizes this to arbitrary β for the $\mathcal{O}(1/N)$ term and in addition, we obtain the $\mathcal{O}(1/N^2)$ correction for $\beta = 1$ and $\beta = 2$. We conclude with a summary and discussion in section 6.5.

[1] Some other types of deformations of the Marčenko-Pastur distribution for Wishart-like matrix models were reported in the literature (see e.g. [4, 51])

6.2 Wishart ensembles

6.2.1 Real and complex Wishart ensembles

Let $\mathbf{X}$ be a real (respectively complex) Gaussian random matrix of size $M \times N$, i.e. a random matrix chosen in the space of $M \times N$ real (resp. complex) matrices according to the law:

$$P(\mathbf{X})d\mathbf{X} \propto \exp\left(-\frac{1}{2\sigma^2}\mathrm{Tr}(\mathbf{X}^\dagger\mathbf{X})\right) d\mathbf{X}\,, \tag{6.4}$$

where $\mathbf{X}^\dagger$ is the Hermitian conjugate of $\mathbf{X}$. In the following, we will denote the real (resp. complex) Wishart ensemble by $\mathcal{W}^\beta$ with $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case).

The real (resp. complex) Wishart Ensemble is the ensemble of $(N \times N)$ square matrices of the product form $\mathbf{W} := \mathbf{X}^\dagger\mathbf{X}$ where $\mathbf{X}$ is a real (resp. complex) Gaussian random matrix of size $N \times M$. They have appeared in many different applications such as communication technology [134], nuclear physics [78], quantum chromodynamics [148], statistical physics of directed polymers in random media [88] and non intersecting Brownian motions [135, 117, 125], as well as Principal Component Analysis of large datasets [108, 107].

The spectral properties of the Wishart matrices have been studied extensively and it is known [87] that for $M \geqslant N$, all N positive eigenvalues of $\mathbf{W}$ are distributed via the joint probability density function (pdf)

$$P_\beta(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{\beta}{2}(M-N+1)-1} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \tag{6.5}$$

where Z is a constant normalization factor and where $\beta = 1$ in the real case (resp. $\beta = 2$ in the complex case). Note that the joint distribution P_β defined in (6.5) is in fact well behaved for every $\beta > 0$.

Another classical result of Random Matrix Theory concerns the asymptotic density of states (or spectral measure) for the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ of a real Wishart matrix $\mathbf{W} \in \mathcal{W}^1$ in the limit of large matrices, i.e. when $N, M \to \infty$ with $N/M = q \in (0; 1]$ where q is a fixed parameter. Let us recall that the density of states of the matrix $\mathbf{W}$ is simply the probability measure ρ_N^β defined as

$$\rho_N^\beta = \frac{1}{N}\sum_{i=1}^N \delta(\lambda - \lambda_i) \tag{6.6}$$

where β is introduced for later convenience ($\beta \equiv 1$ in the present case) and where $(\lambda_1, \ldots, \lambda_N)$ are the eigenvalues of $\mathbf{W}$. Setting $\lambda = M\hat{\lambda}$, the Marčenko-Pastur Theorem (see [113]) states that, in the limit $N, M \to \infty$ with $N/M = q \in (0; 1]$, the spectral measure of a Wishart matrix $\mathbf{W} \in \mathcal{W}^1$ converges to a continuous probability density (with compact support) given by

$$\rho_\beta(\hat{\lambda}) = \frac{1}{2\pi\sigma^2\beta q}\frac{\sqrt{(b-\hat{\lambda})(\hat{\lambda}-a)}}{\hat{\lambda}}, \quad a < \hat{\lambda} < b \tag{6.7}$$

where the edges a, b of the spectrum are given by

$$a = \sigma^2\beta\,(1-\sqrt{q})^2, \quad b = \sigma^2\beta\,(1+\sqrt{q})^2$$

with again $\beta = 1$.

For general $\beta > 0$, the probability measure ρ_N^β is defined again as in (6.6) where this time the vector $(\lambda_1, \ldots, \lambda_N)$ is distributed according to the law P_β in (6.5). The Marčenko-Pastur theorem remains in fact valid for all $\beta > 0$ in the sense that the probability law ρ_N^β converges when $N, M \to \infty$ with $N/M = q \in (0; 1]$ to the continuous probability density ρ_β in (6.7) for every $\beta > 0$.

The probability measure ρ_N^β will sometimes be referred to as the spectral density as it corresponds to the spectral density of random matrices $\mathbf{W} \in \mathcal{W}^\beta$ at least when $\beta = 1$ or 2.

6.2.2 Continuous processes for real and complex Wishart ensembles

We wish to define here a diffusive matrix process depending on a fictitious time $t \geqslant 0$ that will converge to the Wishart Ensembles in the limit of large time. The idea is simply to set

$$\mathbf{W}_t := \mathbf{X}_t^\dagger \mathbf{X}_t \tag{6.8}$$

where $\mathbf{X}_t$ is a real (resp. complex) random matrix diffusion process (of size $M \times N$), starting at time $t = 0$ from an initial arbitrary real (resp. complex) fixed matrix $\mathbf{X}_0$ and then evolving for $t \geqslant 0$ as

$$d\mathbf{X}_t = -\frac{1}{2}\mathbf{X}_t dt + \sigma\, d\mathbf{B}_t$$

where $\mathbf{B}_t$ is a real Brownian (resp. complex) random matrix, i.e. a matrix whose entries are given by independent standard Brownian motions. By a standard Brownian motion, one means a centered (zero-mean) Gaussian process with covariance function $\langle B_t B_{t'} \rangle = \min(t, t')$. Note that the entries of the matrix process $\mathbf{X}_t$ are independent one dimensional Ornstein Uhlenbeck processes.

It is well known that the stationary law of a Ornstein-Uhlenbeck process is the Gaussian law and therefore, the real (resp. complex) matrix process $\mathbf{X}_t$ converges in law when $t \to \infty$ to the law of a Gaussian real (resp. complex) random matrix. Hence, we deduce that the real (resp. complex) matrix process $\mathbf{W}_t$ defines a diffusive matrix process that converges in law to $\mathbf{W} \in \mathcal{W}^\beta$ with $\beta = 1$ or $\beta = 2$.

It is also easy to check that the positive definite matrix process $\mathbf{W}_t$ has initial condition $\mathbf{W}_0 = \mathbf{X}_0^\dagger \mathbf{X}_0$ in $t = 0$ and satisfies, for $t \geqslant 0$, the following stochastic differential equation studied by Bru [49]:

$$d\mathbf{W}_t = -\mathbf{W}_t dt + \sigma \sqrt{\mathbf{W}_t}\, d\mathbf{B}_t + \sigma\, d\mathbf{B}_t^\dagger \sqrt{\mathbf{W}_t} + M\sigma^2 \beta\, \mathbf{I}\, dt \tag{6.9}$$

where $\mathbf{B}_t$ is a real (resp. complex) Brownian random matrix and with $\beta = 1$ in the real (resp. $\beta = 2$ for complex) case.

The evolution of the eigenvalue process $\lambda_1(t) \leqslant \lambda_2(t) \ldots \leqslant \lambda_N(t)$ is also easy to derive [50] using perturbation theory to second order

$$d\lambda_i = -\lambda_i dt + 2\sigma\sqrt{\lambda_i}\, db_i + \sigma^2 \beta \left(M + \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) dt \tag{6.10}$$

where b_i's are independent standard Brownian motions and with $\beta = 1$ in the real (resp. $\beta = 2$ for complex) case. The stationary distribution of the process $(\lambda_1, \ldots, \lambda_N)(t)$ is necessarily the jpdf $P_\beta(\lambda_1, \ldots, \lambda_N)$ defined in (6.5) (this is true for any $\beta > 0$ and can also be recovered using the Fokker-Planck equation for the multivariate diffusion (6.10)).

6.3 Crossover between Wishart and CIR processes

Following [6, 10], we aim at defining a diffusive matrix process $\mathbf{W}_t$ which converge in the limit of large time to a general β-Wishart matrix, i.e. a matrix whose eigenvalues are distributed according to P_β in (6.5) for general $\beta > 0$ and with Haar distributed eigenvectors. In this paper, we will restrict ourselves to the description of the eigenvalues process but the interested reader can find a study of the eigenvectors for a related model in [10]. To simplify notations, we will take in this section $\sigma = 1$.

6.3.1 Preliminary definition: CIR diffusion process

We first need to introduce a family of real diffusion processes. Let $\delta > 0$ be a fixed parameter. The CIR process (named after its creators Cox, Ingersoll, and Ross [**?**] and widely used in finance to model short term interest rate) is a diffusion process $x(t)$ defined by $x(0) := x_0 > 0$ and for $t \geqslant 0$ by

$$dx(t) = -x(t)\ dt + 2\sqrt{x(t)}\ db_t + \delta\ dt\,. \tag{6.11}$$

Using the assumption $\delta > 0$, it is easy to see that the process $x(t)$ will remain non negative for all times $t \geqslant 0$. It is also easy to verify that the stationary pdf of the Langevin equation (6.11) is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ defined as

$$p_\delta(x) = \frac{1}{2^{\frac{\delta}{2}}\Gamma(\frac{\delta}{2})} x^{\frac{\delta}{2}-1} e^{-\frac{x}{2}} . \tag{6.12}$$

This stationary pdf p_δ is in fact the unique stationary pdf for the process $x(t)$ and hence does not depend on the initial condition x_0. In analogy with squared Bessel processes, the parameter δ will be called the *dimension* of the process $x(t)$.

6.3.2 Diffusive matrix process for general β-Wishart matrices

Following [6, 10], our goal in this paragraph is to construct a diffusive matrix process whose eigenvalues process is asymptotically distributed according to P_β for general $\beta \in [0, 1]$. This construction can be extended respectively for general $\beta \in [0, 2]$ (resp. $\beta \in [0, 4]$) by using complex (resp. symplectic) Brownian motions instead of real Brownian motions in the following.

The idea is to slice the time interval into small chops of length $1/n$ and for each interval $[k/n; (k+1)/n]$, to choose independently Bernoulli random variables $\epsilon_k^n, k \in \mathbb{N}$ such that $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$. Then, setting $\epsilon_t^n = \epsilon_{[nt]}^n$, our diffusive matrix process evolves as:

$$d\mathbf{W}_t^n = -\mathbf{W}_t^n \, dt + d\mathbf{\Delta}_t^n \tag{6.13}$$

where the increment matrix $d\mathbf{\Delta}_t^n$ now depends on the value of the additional random process ϵ_t^n:

- if $\epsilon_t^n = 1$, then
$$d\mathbf{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\, d\mathbf{B}_t + d\mathbf{B}_t^\dagger \sqrt{\mathbf{W}_t^n} + M \, \mathbf{I}\, dt.$$
where $d\mathbf{B}_t$ is an $N \times N$ real [2] Brownian increment matrix whose entries have variance dt.
- if $\epsilon_t^n = 0$, then
$$d\mathbf{\Delta}_t^n = \sqrt{\mathbf{W}_t^n}\, d\mathbf{Y}_t + d\mathbf{Y}_t^\dagger \sqrt{\mathbf{W}_t^n} + \delta \, \mathbf{I}\, dt.$$
with $\delta > 0$ and where $d\mathbf{Y}_t$ is a symmetric matrix that is co-diagonalizable with $\mathbf{W}_t^n$ (i.e. the two matrix have the same eigenvectors) but with a spectrum given by N independent real Brownian increments of variance dt.

An algorithmic description of how to build (approximatively on a discrete grid) the matrix process $\mathbf{W}_t^n$ can be found in Appendix 6.5.

It is clear that the eigenvalues of the matrix $\mathbf{W}_t^n$ will cross at some points but only in intervals $[k/n; (k+1)/n]$ for which $\epsilon_k^n = 0$ (in the other intervals where they follow the SDE (6.10) with parameter $\beta = 1$, it is well known that the repulsion is too strong and thus collisions are avoided). In this case, the eigenvalues are re-numbered at time $t = (k+1)/n$ in increasing order. With this procedure, when ordered $\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t)$, we can again check as in [6, 10], using perturbation theory, that the eigenvalues will remain always non-negative and will verify the Stochastic Differential System (SDS):

$$d\lambda_i^n = -\lambda_i^n dt + 2\sqrt{\lambda_i^n}\, db_i + \left(\epsilon_t^n M + (1 - \epsilon_t^n)\delta + \epsilon_t^n \sum_{k \neq i} \frac{\lambda_i^n + \lambda_k^n}{\lambda_i^n - \lambda_k^n} \right) dt \tag{6.14}$$

where the b_i are independent standard Brownian motions, which are also independent of the process ϵ_t^n.

Note that when $\epsilon_t^n = 0$, the particles λ_i^n are evolving as independent CIR processes of dimension $\delta > 0$ as defined in paragraph 6.3.1. Therefore, the particles can cross in those time intervals, breaking the increasing order so that they will be re-ordered at time $([nt] + 1)/n$ but they *will* remain non-negative as the dimension δ is strictly positive. Therefore the SDS (6.14) remains well defined at all times $t \geqslant 0$.

[2] Here one can use use complex Brownian motions instead to extend the interval of β to $[0, 2]$.

One can follow the proof of [10] to prove that the scaling limit (i.e. the limiting process when $n \to \infty$) of the process $(\lambda_1^n(t) \leqslant \ldots \leqslant \lambda_N^n(t))$ satisfies the following SDS

$$d\lambda_i = -\lambda_i dt + 2\sqrt{\lambda_i}\, db_i + \left(pM + (1-p)\delta + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) dt\,. \tag{6.15}$$

One can deduce from the above equation (6.15) the Fokker-Planck equation for the joint density $P(\{\lambda_i\}, t)$, for which the stationary jpdf is readily found to be [see the derivation in appendix 6.5]

$$P^*(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z} e^{-\frac{1}{2}\sum_{i=1}^N \lambda_i} \prod_{i=1}^N \lambda_i^{\frac{p}{2}(M-N+1-\delta)-(1-\frac{\delta}{2})} \prod_{i<j} |\lambda_i - \lambda_j|^p\,. \tag{6.16}$$

The probability P_β introduced in (6.5) is recovered here by taking the values $p = \beta$ and $\delta = 0$. The corresponding large N, M-limit spectral probability density is therefore given by the Marčenko-Pastur law in the case where $p = \beta > 0$ *independent of* M. Note that with the above normalizations, the spectrum is spread over a region of $\mathbb{R}_+$ of width of order $pM = \beta M$. On the other hand, if $p = 0$, the large N, M-limit of the spectral density is the Gamma distribution with shape and scale parameters $k = \delta/2$ and $\theta = 2$ (recall that it is the stationary pdf of the CIR process of dimension δ):

$$\rho_0(\lambda) d\lambda = \frac{1}{2^{\delta/2}\Gamma(\frac{\delta}{2})} \lambda^{\frac{\delta}{2}-1} e^{-\frac{\lambda}{2}}\, d\lambda\,. \tag{6.17}$$

It is quite natural to ask whether a crossover regime may be found, interpolating between the Marčenko-Pastur density (p independent of M) and the Gamma distribution ($p = 0$). A good candidate for triggering such a transition is clearly a parameter p vanishing with M as $p = 2c/M$ where c is a positive fixed constant. We discuss this case in the following subsection.

In the next two subsections we compute the crossover density interpolating between the Marčenko-Pastur law and the Gamma distribution with shape parameter $\delta/2$. This family of probability densities is indexed by the three parameters c (such that $p = 2c/M$), $q = N/M$ and $\delta \geqslant 0$. More precisely, we compute the limiting density of the probability measure $\rho_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$, when $N, M \to \infty$ with $N/M = q \in (0; 1]$ and where $(\lambda_1, \ldots, \lambda_N)$ is distributed according to the law P^* defined in (6.16) with $p = \beta = 2c/M$. As mentioned in the introduction, the crossover density can be computed via two alternative methods: (1) by a saddle point method as shown in subsection 6.3.3 and (2) by analyzing directly the stochastic differential systems introduced in 6.3.2 above following the analogous route for the Gaussian case in Ref. [6]. This is done in section 6.3.4. We will see that both methods yield identical result.

6.3.3 Crossover for the spectral density via the saddle point method

Our starting point in the joint probability law of eigenvalues in Eq. (6.16), where the normalization constant (partition function) Z is given by the N-fold integral

$$\begin{aligned} Z &= \int_{[0,\infty]^N} \prod_i d\lambda_i e^{-\frac{1}{2}\sum_i \lambda_i} \prod_{i<j} |\lambda_i - \lambda_j|^p \prod_i \lambda_i^{\frac{p}{2}(M-N+1-\delta)-(1-\delta/2)} \\ &= \int_{[0,\infty]^N} \prod_i d\lambda_i e^{-E[\{\lambda_i\}]} \end{aligned} \tag{6.18}$$

where the energy function $E[\{\lambda_i\}]$ is given by

$$E[\{\lambda_i\}] = \frac{1}{2}\sum_i \lambda_i - \left(\frac{p}{2}(M-N+1-\delta) - (1-\delta/2)\right) \sum_i \ln \lambda_i - \frac{p}{2}\sum_{i \neq j} \ln|\lambda_i - \lambda_j|. \tag{6.19}$$

Written in this form, Eq. (6.18) is the Gibbs-Boltzmann canonical weight of a system of charged particles on the positive half-line in equilibrium at inverse temperature $\beta = 1$ under the effect

of competing interactions. We wish to compute the average density of states $\langle \frac{1}{N}\sum_{i=1}^{N}\delta(\lambda-\lambda_i)\rangle$ where $\langle O\rangle$ denote the expectation value of O with respect to the probability distribution P^* in (6.16). There are many ways to compute this average density, but the one rather convenient for large N is the saddle point route. This was originally done by Dyson [69] for the Gaussian random matrices and a physically more transparent derivation can be found in Ref. [62].

The main idea behind the saddle point method is as follows. In the large N limit, the most dominant contribution to the partition function emerges indeed from a set of configurations of λ_i's that correspond to a particular density $\rho^*(\lambda)$. Naturally then, the average computed over the ensemble of λ_i's, in this large N limit, will also be given by the saddle point density $\langle \frac{1}{N}\sum_{i=1}^{N}\delta(\lambda-\lambda_i)\rangle \approx \rho^*(\lambda)$. It thus suffices to analyze just the partition function Z in the large N limit and find, in particular, the saddle point density $\rho^*(\lambda)$ that maximizes the partition function Z for large N.

To analyze Z in the large N limit, one first defines a 'local' smooth density function

$$\rho(\lambda) = \frac{1}{N}\sum_{i=1}^{N}\delta\left(\lambda-\lambda_i\right) \tag{6.20}$$

which is normalized to unity. The main idea then is to split the multiple integration in Eq. (6.18) in two parts: First fix the local density $\rho(\lambda)$ and sum over all microstates (i.e., configurations of λ_i's consistent with the local density defined in (6.20)) and then, sum (functionally) over all possible local density functions. Roughly speaking, the first step corresponds to a partial tracing over microstates by fixing the local density. Notationally, on can express this by

$$Z = \int \mathcal{D}[\rho]\int_{[0,\infty]^N}\prod_i d\lambda_i e^{-E[\{\lambda_i\}]}\, I\left[\rho(\lambda),\{\lambda_i\}\right] \tag{6.21}$$

where $\mathcal{D}[\rho]$ denotes a functional integration over the function space and $I\left[\rho(\lambda),\{\lambda_i\}\right]$ is an indicator function that is 1 if the microstate $\{\lambda_i\}$ is compatible with a given $\rho(\lambda)$, normalized to unity, as defined in (6.20) and otherwise $I=0$. The energy function $E[\{\lambda_i\}]$ associated with a microstate can then be expressed in terms of the local density $\rho(\lambda)$ using the identity $\sum_i f(\lambda_i) = N\int d\lambda f(\lambda)\rho(\lambda)$ and one gets [62]

$$\begin{aligned} E[\rho(\lambda)] = \frac{N}{2}\int d\lambda\lambda\rho(\lambda) - \left[\frac{p}{2}\left((\frac{1}{q}-1)N+1-\delta\right)-\left(1-\frac{\delta}{2}\right)\right]N\int d\lambda\rho(\lambda)\ln\lambda \\ -\frac{p}{2}N^2\int\int d\lambda d\lambda'\rho(\lambda)\rho(\lambda')\ln|\lambda-\lambda'| + \frac{p}{2}N\int d\lambda\rho(\lambda)\ln\frac{1}{\rho(\lambda)} + C_1\left(\int d\lambda\rho(\lambda)-1\right) \end{aligned} \tag{6.22}$$

where the last term includes a Lagrange multiplier C_1 that enforces the normalization of the local density to 1. The next-to-last term accounts for the self-energy term ($\lambda\to\lambda'$) that needs to be subtracted. Note indeed that in the original discrete sum $\sum_{i\neq j}\ln|\lambda_i-\lambda_j|$, the eigenvalues do not coincide. This means that the integral over λ and λ' should exclude the region where $|\lambda-\lambda'|$ is less than the typical spacing between eigenvalues, which is proportional to $1/N\rho(\lambda)$ [69]. The contribution of this thin sliver is the next-to-last term, up to an additional contribution that can be absorbed into C_1.

Once this is done, equation (6.21) simplifies further and one gets

$$Z \approx \int \mathcal{D}[\rho]\, e^{-E[\rho(\lambda)]}\, J[\rho(\lambda)] \tag{6.23}$$

where $J[\rho(\lambda)] = \int_{[0,\infty]^N}\prod_i d\lambda_i I\left[\rho(\lambda),\{\lambda_i\}\right]$ is an entropic factor that just counts how many microstates are compatible with a given local density function $\rho(\lambda)$. This can be estimated very simply by the following combinatorial argument involving arrangement of N particles in K boxes. Let us first divide our one dimensional line into K small boxes of equal width. We have N particles that need to be distributed into the K boxes with occupation numbers $\{n_1,n_2,\ldots,n_K\}$. The number of ways this can be done is simply

$$\frac{N!}{n_1!n_2!\cdots n_K!}\ . \tag{6.24}$$

Setting $\rho_i = n_i/N$ (the local density in box i) and using Stirling's approximation $N! \sim N^{N+1/2}\, e^{-N}$ (using the fact that $\sum_i n_i = N$), we have:

$$\frac{N!}{n_1! n_2! \cdots n_K!} \sim e^{-\sum_i n_i \ln n_i} \tag{6.25}$$

which in the continuum limit, becomes $\sim e^{-N \int d\lambda \rho(\lambda) \ln \rho(\lambda)}$. Thus, the entropic factor can also be expressed as a simple functional of the local density $\rho(\lambda)$. Inserting this expression in the functional integral over the density, yields:

$$Z = \int \mathcal{D}[\rho] e^{-E[\rho(\lambda)]} e^{-N \int d\lambda \rho(\lambda) \ln \rho(\lambda)} = \int \mathcal{D}[\rho] e^{-N F[\rho(\lambda)]} \tag{6.26}$$

where the free energy $F[\rho(\lambda)]$ is given by:

$$\begin{aligned} F[\rho(\lambda)] = \frac{1}{2} \int d\lambda \lambda \rho(\lambda) - \left[\frac{p}{2} \left((\frac{1}{q} - 1) N + 1 - \delta \right) - \left(1 - \frac{\delta}{2} \right) \right] \int d\lambda \rho(\lambda) \ln \lambda \\ - \frac{p}{2} N \int \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| + \left(1 - \frac{p}{2} \right) \int d\lambda \rho(\lambda) \ln \rho(\lambda) + C_1 \left(\int d\lambda \rho(\lambda) - 1 \right) \end{aligned} \tag{6.27}$$

Note that for $p \sim \mathcal{O}(1/N)$ the entropy term becomes of the same order of the energy term, while in the usual case $p \sim \mathcal{O}(1)$ the entropy contribution is subdominant in the large N limit and is therefore disregarded.

Setting now $p = 2c/M = 2cq/N$, we get:

$$\begin{aligned} F[\rho(\lambda)] = \frac{1}{2} \int d\lambda \lambda \rho(\lambda) - \left[cq \left(\frac{1}{q} - 1 \right) - \left(1 - \frac{\delta}{2} \right) \right] \int d\lambda \rho(\lambda) \ln \lambda \\ - cq \int \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| + \left(1 - \frac{cq}{N} \right) \int d\lambda \rho(\lambda) \ln \rho(\lambda) + C_1 \left(\int d\lambda \rho(\lambda) - 1 \right) \end{aligned} \tag{6.28}$$

We set $a = cq(1/q - 1) - (1 - \delta/2)$ and take $N \to \infty$ (so that the term cq/N drops out). Now, the saddle point density $\rho^*(\lambda)$ is obtained by minimizing the free energy $F[\rho(\lambda)$, i.e., by taking the functional derivative $\frac{\delta F}{\delta \rho} = 0$ and $\rho(\lambda) == \rho^*(\lambda)$. This gives the saddle point equation

$$\frac{\lambda}{2} - a \ln \lambda - 2cq \int d\lambda' \rho^*(\lambda') \ln |\lambda - \lambda'| + \ln \rho^* + C_2 = 0 \tag{6.29}$$

where $C_2 = C_1 + 1$ is just a constant. For notational simplicity, in the rest of the subsection we will denote the saddle point density $\rho^*(\lambda)$ simply by $\rho(\lambda)$.

Differentiating Eq. (6.29) with respect to λ, we get

$$\frac{1}{2} - \frac{a}{\lambda} - 2cq \ \mathrm{Pr} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda' + \frac{\rho'(\lambda)}{\rho(\lambda)} = 0 \tag{6.30}$$

where Pr stands for Principal Value.

Next, we define the Stieltjes transform:

$$H(z) = \int \frac{\rho(\lambda)}{\lambda - z} d\lambda \tag{6.31}$$

for z complex and outside the support of ρ. By definition, for large $|z|$, $H(z) \to -1/z$. Multiplying eq. (6.30) by $\rho(\lambda)/(\lambda - z)$ and integrating over λ, we have:

$$\frac{1}{2} \int \frac{\rho(\lambda)}{\lambda - z} d\lambda - a \int \frac{\rho(\lambda) d\lambda}{\lambda(\lambda - z)} - 2cq \ \mathrm{Pr} \int \frac{\rho(\lambda) d\lambda}{\lambda - z} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda' + \int \frac{\rho'(\lambda) d\lambda}{\lambda - z} = 0 \tag{6.32}$$

and we analyze each of the four contributions separately.

1. $T_1 = \frac{1}{2}\int \frac{\rho(\lambda)}{\lambda - z} d\lambda = \frac{1}{2} H(z)$

2. $T_2 = -a \int \frac{\rho(\lambda) d\lambda}{\lambda(\lambda - z)}$. We rewrite this as:

$$T_2 = -a \int \rho(\lambda) d\lambda \left[\frac{1}{\lambda - z} - \frac{1}{\lambda} \right] \frac{1}{z} \tag{6.33}$$

   implying:

$$T_2 = -\frac{a}{z} H(z) + \frac{b_1}{z} \tag{6.34}$$

   where $b_1 = a \int d\lambda \frac{\rho(\lambda)}{\lambda}$.

3. $T_3 = -2cq \ \mathrm{Pr} \int \frac{\rho(\lambda) d\lambda}{\lambda - z} \int \frac{\rho(\lambda')}{\lambda - \lambda'} d\lambda'$ which we rewrite as:

$$T_3 = 2cq \left\{ \mathrm{Pr} \int d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \left[\frac{1}{\lambda - z} - \frac{1}{\lambda - \lambda'} \right] \frac{1}{\lambda' - z} \right\} \tag{6.35}$$

$$= 2cq H^2(z) - 2cq \ \mathrm{Pr} \int \frac{d\lambda d\lambda' \rho(\lambda) \rho(\lambda')}{(\lambda - \lambda')(\lambda' - z)} \tag{6.36}$$

   By renaming $\lambda \to \lambda'$ and $\lambda' \to \lambda$, we get:

$$T_3 = 2cq H^2(z) + 2cq \mathrm{Pr} \int \frac{d\lambda d\lambda' \rho(\lambda) \rho(\lambda')}{(\lambda - \lambda')(\lambda - z)} \tag{6.37}$$

$$= 2cq H^2(z) - T_3 \tag{6.38}$$

   Solving for T_3 we get:

$$T_3 = cq H^2(z) \tag{6.39}$$

4. $T_4 = \int \frac{\rho'(\lambda) d\lambda}{\lambda - z}$, which we integrate by parts, obtaining:

$$T_4 = \frac{1}{\lambda - z} \rho(\lambda) \Big|_0^\infty + \int \frac{\rho(\lambda)}{(\lambda - z)^2} d\lambda$$

$$= \frac{c_1}{z} + H'(z) \tag{6.40}$$

In the derivation above, we assumed b_1 and c_1 to be finite. This is not completely obvious, because $\rho(\lambda)$ at an edge point may diverge. However, by imposing that for large z, $H(z) \to -1/z$, it is immediate to derive that $b_1 + c_1 = 1/2$. Thus, one may regularize the density near the edge points so that b_1 and c_1 exist individually, but eventually their sum is universally $1/2$ and hence is independent of the specific regularization near the edge.

Adding up the four contributions, we get the equation:

$$\frac{dH}{dz} - \frac{a}{z} H + \frac{b_1 + c_1}{z} + cq \, H^2 + \frac{1}{2} H = 0$$

Thus we find the following differential equation for the Stieltjes transform H:

$$\boxed{\frac{dH}{dz} + \gamma \, H^2 + \frac{1}{2} \left(1 + \frac{\alpha}{z} \right) H + \frac{1}{2z} = 0} \tag{6.41}$$

where we have set

$$\alpha = (2 - \delta) - 2c(1 - q), \quad \gamma = cq \, .$$

In the next subsection, we will derive the same equation via Itô's stochastic calculus route.

The density $\rho(\lambda)$ (normalized to unity) can then be read off from

$$\rho(\lambda) = \frac{1}{\pi} \mathrm{Im}[H(z \to \lambda)] \tag{6.42}$$

where $z \to \lambda$ occurs inside the cut on the real axis.

To solve the Riccati equation (6.41), we make a standard substitution

$$H(z) = \frac{1}{\gamma}\frac{u'(z)}{u(z)} = \frac{1}{\gamma}\,\partial_z \ln u(z)\,. \tag{6.43}$$

This gives a second order differential equation for $u(z)$

$$u''(z) + \frac{1}{2}\left[1 + \frac{\alpha}{z}\right] u'(z) + \frac{\gamma}{2z} u(z) = 0\,. \tag{6.44}$$

It follows from Eq. (6.43) and the asymptotic behavior of $H(z)$ that

$$u(z) \xrightarrow[|z|\to\infty]{} \frac{A_1}{z^\gamma} \tag{6.45}$$

where A_1 is a constant.

To reduce Eq. (6.44) to a Schrödinger like differential equation, we make the substitution

$$u(z) = e^{-z/4}\, z^{\alpha/4}\, \psi(z)\,, \tag{6.46}$$

and we find the following equation for ψ

$$\psi''(z) + \left[-\frac{1}{16} + \frac{1}{z}\frac{4\gamma - \alpha}{8} + \frac{\alpha}{4}\left(1 - \frac{\alpha}{4}\right)\frac{1}{z^2}\right]\psi(z) = 0\,.$$

Making further a rescaling $\psi(z) = y(z/2)$, it reduces to the standard form of the Whittaker differential equation [83]

$$y''(z) + \left[-\frac{1}{4} + \frac{\lambda}{z} + \frac{\frac{1}{4} - \mu^2}{z^2}\right] y(z) = 0\,, \tag{6.47}$$

where

$$\zeta = \gamma - \frac{\alpha}{4}, \quad \mu = \frac{1}{2}|1 - \frac{\alpha}{2}|\,. \tag{6.48}$$

Note that the solution of this differential equation does not depend on the sign of μ, hence we take the absolute value. The differential equation (6.47) has two linearly independent solutions $W_{\zeta,\mu}(z)$ and $W_{-\zeta,\mu}(-z)$. The Whittaker function $W_{\zeta,\mu}(z)$ has the following asymptotic behavior [83]

$$W_{\zeta,\mu}(z) \xrightarrow[|z|\to\infty]{} z^\zeta\, e^{-z/2}\,. \tag{6.49}$$

Thus the general solution of $u(z)$, using Eq. (6.46), reads

$$u(z) = e^{-z/4}\, z^{\alpha/4}\, \left[\mathcal{C}_1\, W_{\zeta,\mu}(z/2) + \mathcal{C}_2\, W_{-\zeta,\mu}(-z/2)\right] \tag{6.50}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are arbitrary constants. Using the asymptotic behavior in Eq. (6.49) it is easy to check that only the second solution has the right asymptotic behavior in Eq. (6.45). Thus, finally, we have our solution

$$u(z) = \mathcal{C}_2\, e^{-z/4}\, z^{\alpha/4}\, W_{-\zeta,\mu}(-z/2) \tag{6.51}$$

where ζ and μ are given in Eq. (6.48).

By plugging this solution (6.51) into Eq. (6.43) and using Eq. (6.42), we find the following expression

$$\rho(\lambda) = \frac{\mathcal{C}_2}{2\pi\gamma}\frac{(\mathrm{Im}(W'_{-\zeta,\mu})\mathrm{Re}(W_{-\zeta,\mu}) - \mathrm{Im}(W_{-\zeta,\mu})\mathrm{Re}(W'_{-\zeta,\mu}))(-\lambda/2)}{|W_{-\zeta,\mu}(-\lambda/2)|^2}\,. \tag{6.52}$$

where Re and Im denote respectively the real and imaginary parts. Using the linear differential equation verified by the Whittaker functions (6.47), it is easy to see that the derivative with

respect to λ of the Wronskian type function $(\mathrm{Im}(W'_{-\zeta,\mu})\mathrm{Re}(W_{-\zeta,\mu}) - \mathrm{Im}(W_{-\zeta,\mu})\mathrm{Re}(W'_{-\zeta,\mu}))$ is equal to 0. Hence the Wronskian appearing in the numerator in (6.52) is simply a constant.

Collecting all the constants together, we get:

$$\rho(\lambda) = \frac{A}{|W_{-\zeta,\mu}(-\lambda/2)|^2}\,. \tag{6.53}$$

The overall normalization constant A has to be fixed from $\int_0^\infty \rho(\lambda)\,d\lambda = 1$. Thus we get, after rescaling $\lambda/2 \to \lambda$,

$$\frac{1}{A} = 2\int_0^\infty \frac{d\lambda}{|W_{-\zeta,\mu}(-\lambda)|^2}\,. \tag{6.54}$$

This integral in Eq. (6.54) can be done in closed form. First, we first use the well known identity [1]

$$W_{\zeta,\mu}(z) = z^{\mu+1/2}\,e^{-z/2}\,U(\mu-\zeta+1/2, 1+2\mu; z) \tag{6.55}$$

where $U(a,b;z)$ is the Tricomi hypergeometric function (or Kummer function) that behaves for large z as $U(z) \sim z^{-a}$. Using this in Eq. (6.54) gives

$$\frac{1}{A} = 2\int_0^\infty d\lambda\,\lambda^{-2\mu-1}\,e^{-\lambda}\,|U(\mu+\zeta+1/2, 1+2\mu; -\lambda)|^{-2}\,. \tag{6.56}$$

It turns out that there exists an interesting integral representation in a paper by Ismail and Kelker [85]

$$\int_0^\infty \frac{dt\; e^{-t}\,t^{-b}}{z+t}\,|U(a,b;-t)|^{-2} = \Gamma(a)\Gamma(a-b+2)\,\frac{1}{z}\frac{U(a,b-1;z)}{U(a,b;z)}; \quad \text{for } a>0,\ 1<b<a+1 \tag{6.57}$$

Note that in Ref. [85] they use the notation $\psi(a,b,z)$ instead of $U(a,b;z)$, but it is the same function. Our μ and ζ satisfy the condition of validity of this identity: $a>0$ and $1<b<a+1$. Taking $z\to\infty$ limit on both sides and using $U(z)\sim z^{-a}$, we arrive at the following exact expression of the normalization constant

$$\frac{1}{A} = 2\,\Gamma(\mu+\zeta+1/2)\Gamma(\zeta-\mu+3/2)\,. \tag{6.58}$$

This leads to the following final expression for the spectral density, which is the central result of our work[3]:

$$\boxed{\rho_c(\lambda) = \frac{1}{2\Gamma(\mu+\zeta+\frac{1}{2})\Gamma(\zeta-\mu+\frac{3}{2})}\,\frac{1}{|W_{-\zeta,\mu}(-\frac{\lambda}{2})|^2}} \tag{6.59}$$

with the following values for the parameters

$$\alpha = (2-\delta) - 2c(1-q); \quad \zeta = cq - \frac{\alpha}{4}; \quad \text{and} \quad \mu = \frac{1}{4}|\alpha-2|\,.$$

The above expression is the analogue, in the present context, of the Askey-Wimp-Kerov one-parameter family of models found in [6], that smoothly interpolates between the Gaussian distribution and Wigner's semi-circle.

Let us now consider the limiting case $c\to 0$ first. In this case, we have $\alpha = (2-\delta)$, $\zeta = \delta/4 - 1/2$ and $\mu = \delta/4$. Thus, $W_{-\zeta,\mu}(-\lambda/2) = W_{1/2-\delta/4,\delta/4}(-\lambda/2)$. It turns out that for these special values of the indices, the Whittaker function simply reduces to $W_{1/2-\delta/4,\delta/4}(-\lambda/2) \propto \lambda^{-\delta/4+1/2}\,e^{\lambda/4}$ up to a proportionality constant [83]. Substituting this in (6.59), we then recover the CIR density in Eq. (6.12). The limit $c\to\infty$ is more tricky as one needs to rescale $\lambda \to \hat{\lambda}\,c$ and take the large c limit carefully. This can be done and one recovers the Marčenko-Pastur law. This can be rewritten (without rescaling λ) as

$$\rho_c(\lambda) \sim_{c\to\infty} \frac{1}{4\pi cq}\,\frac{\sqrt{(\gamma_+-\lambda)(\lambda-\gamma_-)}}{\lambda}\;\mathbf{1}_{\{\gamma_-<\lambda<\gamma_+\}}\,, \tag{6.60}$$

[3] We add the subscript c in the notation of the density ρ to recall the dependence in c; the density ρ depends also on the two parameters δ and q but we omit to subscript those.

where $\gamma_\pm = 2c\,(1 \pm \sqrt{q})^2$.

Using standard results on Whittaker functions (see e.g. [83, 1]), it is easy to compute the asymptotic behavior of $\rho_c(\lambda)$ for $\lambda \to 0$ and $\lambda \to +\infty$. Up to multiplicative constants, we have

$$\rho_c(\lambda) \sim_{\lambda \to +\infty} \lambda^{2\zeta}\ e^{-\frac{\lambda}{2}}\,,$$

and

$$\rho_c(\lambda) \sim_{\lambda \to 0_+} \lambda^{2\mu-1}\,.$$

We plotted in Fig. 6.1 the density ρ_c for $c = 0, 1, 2, 3, 4, 5, 10$ and $q = 1/2, \delta = 1$, showing the progressive deformation of the Gamma distribution with shape parameter $\delta/2$ towards the Marčenko-Pastur distribution (6.60) with parameter $q = 1/2$. The critical value of c at which the divergence at $\lambda \to 0_+$ changes to convergence is $c^* = (2-\delta)/(2(1-q)) = 1$. As expected, in Fig. 6.1, the curve with second highest value at the origin corresponds to $c = 1$ and converges when $\lambda \to 0_+$ to $1/2$. The curve with highest value at the origin is the Gamma distribution with shape parameter $\delta/2 = 1/2$ and diverges at 0_+ to $+\infty$. The other curves corresponding to $c = 2, 3, 4, 5, 10$ converge to 0 when $\lambda \to 0_+$.

We have also verified our analytical result for the crossover density in (6.59) numerically for the sample value $c = 1$ and found very good agreement (see Fig. 6.2).

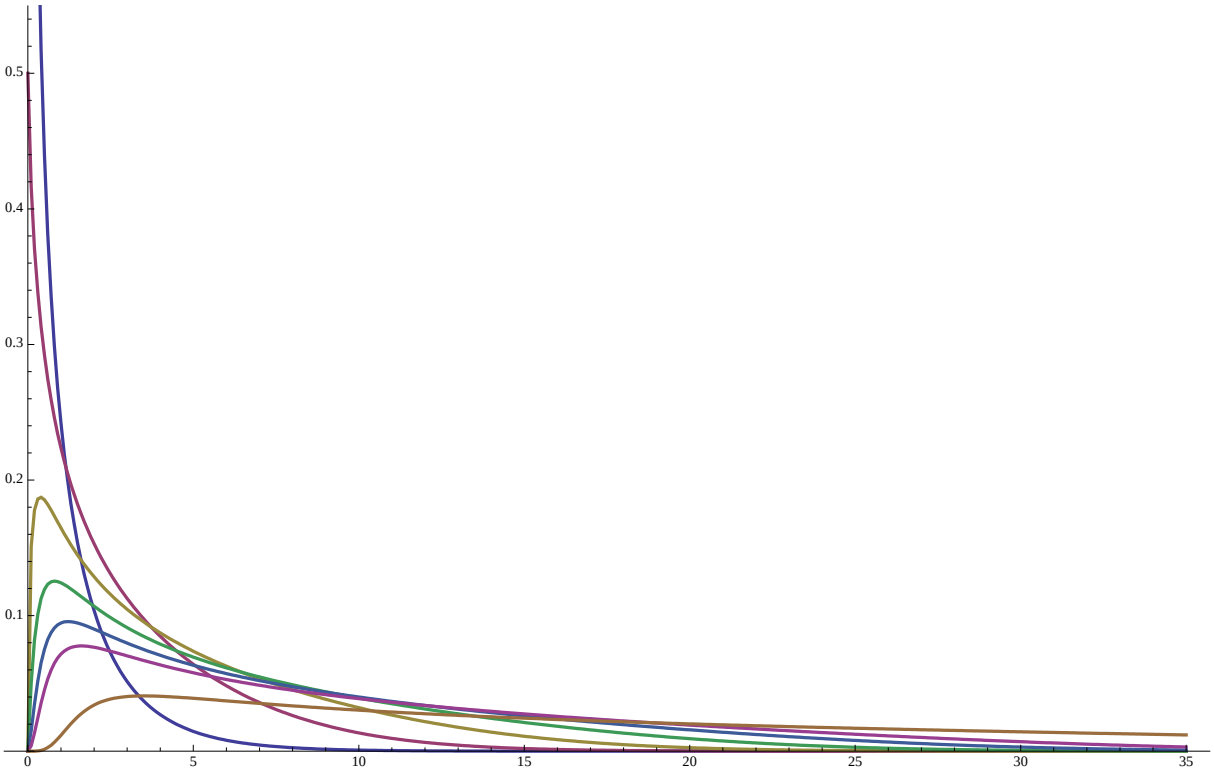


Figure 6.1: Density $\rho_c(\lambda)$ for $c = 0, 1, 2, 3, 4, 5, 10$ of Eq. (6.59) showing the progressive deformation of the Gamma distribution (6.12) with parameter $\delta = 1$ towards the Marčenko-Pastur distribution with parameter $q = 1/2$. The value $\rho_c(0)$ at the origin decreases when c increases.

6.3.4 Crossover for the spectral density via Itô's stochastic calculus

In this subsection, we want to re-obtain the result Eq. (6.41) of the previous subsection via Itô's calculus. We therefore consider the process $(\lambda_1(t), \cdots, \lambda_N(t))$ which verifies the stochastic differential system (6.15) with the scaling relation $p = 2c/M$. The idea is to work out the evolution equation of the probability measure

$$\rho_N^t(dx) := \frac{1}{N}\sum_{i=1}^{N} \delta(x - \lambda_i(t)) \tag{6.61}$$

in the large N limit. We expect the equilibrium of this evolution equation to be the solution of (6.41).

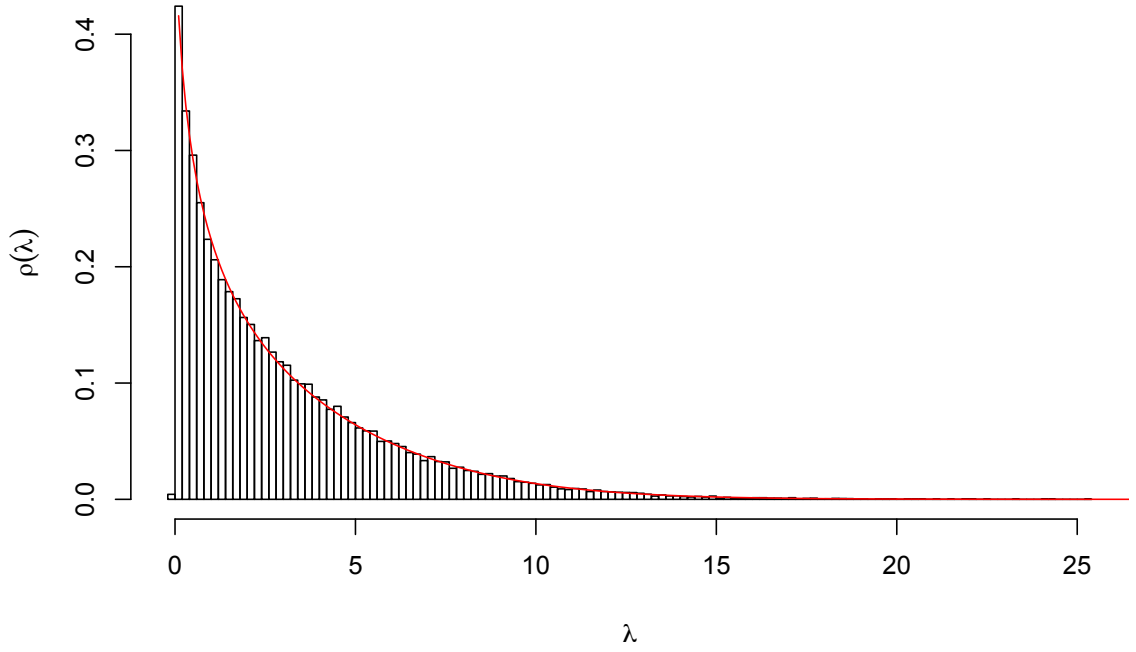


Figure 6.2: Numerical simulations of the state density of the random matrix $\mathbf{W}^n_{t=\infty}$ defined in Eq. (6.13) for $c = 1, p = \beta = 2c/M, M = 100, N = 50, \delta = 1, q = 1/2$.

In the following, f is a smooth function. Using Itô's formula for $\int f(x)\rho^t_N(dx)$, Eq. (6.15) and the scaling relation $p = 2c/M$, we obtain (see [130] for similar calculations)

$$\begin{aligned} d\int f(x)\rho^t_N(dx) = &\int \left(-x + 2c + \left(1 - \frac{2c}{M}\right)\delta\right) f'(x)\rho^s_N(dx)dt \\ &+ 2\left(1 - \frac{c}{M}\right)\int x f''(x)\rho^s_N(dx)dt \\ &+ cq\int_0^t \int\int \frac{f'(x) - f'(y)}{x - y}(x + y)\rho^s_N(dx)\rho^s_N(dy)dt + dM^N_t\,, \end{aligned} \tag{6.62}$$

where $dM^N_t = \frac{2}{N}\sum_{i=1}^N \sqrt{\lambda_i}\, f'(\lambda_i)db_i$ is a noise term of variance $\frac{4}{N}\int \lambda f'(\lambda)^2\rho^t_N(d\lambda)dt$. When $N, M \to \infty$ with $N/M = q$, this noise term is of order $1/\sqrt{N}$.

In the large N, M limit, the *stationary* probability measure ρ solution of Eq. (6.62) therefore satisfies to leading order (keeping only the terms of order 1)

$$\begin{aligned} &\int (-x + 2c + \delta)\, f'(x)\rho(dx) + 2\int x f''(x)\rho(dx) \\ &+cq\int\int \frac{f'(x) - f'(y)}{x - y}(x + y)\rho(dx)\rho(dy) = 0 \end{aligned} \tag{6.63}$$

Applying Eq. (6.63) to the particular function $f(x) = \frac{1}{x-z}$ for $z \in \mathbb{C}\setminus\mathbb{R}$ and denoting $H(z)$ the Stieltjes transform of the probability measure ρ, we obtain the following differential equation for H

$$\begin{aligned} \left[H(z) + zH'(z)\right] - (2c + \delta)\, H'(z) + 2\left[2H'(z) + zH''(z)\right] \\ +2cq\left[H'(z) + H^2(z) + 2zH(z)H'(z)\right] = 0\,. \end{aligned} \tag{6.64}$$

Eq. (6.64) can be rearranged as

$$\begin{aligned} 2cqH(z)[H(z) + 2zH'(z)] + \frac{1}{2}[H(z) + 2zH'(z)] + \frac{1}{2}H(z) + [3H'(z) + 2zH''(z)] \\ + [(1 - \delta) - 2c(1 - q)]\, H'(z) = 0\,. \end{aligned} \tag{6.65}$$

Eq. (6.65) can be integrated easily by doing the change of function $G(z) = zH(z^2)$. Indeed, we just need to write (6.65) for z^2 instead of z and then multiply the corresponding equation by z to obtain the following equation

$$2cqG(z)G'(z) + \frac{1}{2}[zG'(z) + G(z)] + \frac{1}{2}G''(z) + \left[\frac{1-\delta}{2} - c(1-q)\right] 2zH'(z^2) = 0\,,$$

which can be integrated as

$$2cqG^2(z) + \left[z + \frac{(1-\delta) - 2c(1-q)}{z}\right] G(z) + G'(z) = -1 \tag{6.66}$$

where the integration constant is chosen so that $zG(z) \sim -1$ when $|z| \to \infty$. Note that the asymptotic behavior for H is therefore also $zH \sim -1$ when $|z| \to \infty$ as is expected for the Stieltjes transform of a probability measure. Rewriting now Equation (6.66) in term of the function H, we obtain exactly Eq. (6.41).

6.4 Correction to the Marčenko-Pastur law for large but finite dimension

In this section, we come back to the case of generalized Wishart matrices for which particles are distributed according to the jpdf P_β with general parameter $\beta > 0$ (*not* scaling with M). We want to compute the first correction terms to the Marčenko-Pastur density ρ_N^β (defined in Eq. (6.6)) for large but finite N, M with $N/M = q \in (0; 1]$.

We are therefore interested in computing functionals of the form $\langle \int f(\lambda)\rho_N^\beta(d\lambda)\rangle$ where f is a test function and where $(\lambda_1, \lambda_2, \dots, \lambda_N)$ is distributed according to the jpdf P_β defined in Eq. (6.5) for $\beta > 0$. The idea is to use the stochastic process $\lambda_1(t) \leqslant \dots \leqslant \lambda_N(t)$ following the SDE (6.10) that converges in law when $t \to \infty$ to the vector $(\lambda_1 \leqslant \dots \leqslant \lambda_N)$ distributed according to P_β.

Note that in order to have a well behaved limiting spectral distribution with edges that do not depend on M or on β, we will choose in this section $\sigma = 1/\sqrt{M\beta}$ (or alternatively rescale all eigenvalues by $1/M\beta$).

Using again Itô's formula for $\int f(\lambda)\rho_N^t(d\lambda)$ (where ρ_N^t is still defined by Eq. (6.61)) and Eq. (6.10), we obtain

$$\begin{aligned} d\int f(x)\rho_N^t(dx) = &\int (-x+1)\, f'(x)\rho_N^t(dx)dt + \frac{1}{M\beta}(2-\beta)\int x f''(x)\rho_N^s(dx)dt \\ &+ \frac{q}{2}\int\int \frac{f'(x)-f'(y)}{x-y}(x+y)\rho_N^t(dx)\rho_N^t(dy)dt + dM_t^N \end{aligned} \tag{6.67}$$

where $dM_t^N = \frac{1}{N}\sum_{i=1}^N 2\sqrt{\frac{\lambda_i}{M\beta}}\, f'(\lambda_i)db_i$ is a noise term of variance $\frac{4}{q\beta M^2}\int \lambda f'(\lambda)^2\rho_N^t(d\lambda)dt$.

Let us introduce the Stieltjes transform $H_t(z)$ of the probability measure ρ_N^t defined as

$$H_t(z) = \int \frac{\rho_N^t(dx)}{x-z}\,.$$

We now apply (6.67) to the particular function $f(x) = 1/(x-z)$ and we take the expectation with respect to the b_i; Eq. (6.67) rewrites as

$$\begin{aligned} \frac{\partial\langle H_t\rangle}{\partial t} = &\left[\langle H_t\rangle + z\frac{\partial\langle H_t\rangle}{\partial z}\right] - \frac{\partial\langle H_t\rangle}{\partial z} + \frac{1}{M\beta}(2-\beta)\left[2\frac{\partial\langle H_t\rangle}{\partial z} + z\frac{\partial^2\langle H_t\rangle}{\partial z^2}\right] \\ &+ q\left[\frac{\partial\langle H_t\rangle}{\partial z} + \langle H_t\rangle^2 + 2z\langle H_t\rangle\frac{\partial\langle H_t\rangle}{\partial z}\right] + q\left[\langle H_t^2\rangle - \langle H_t\rangle^2\right] + zq\frac{\partial}{\partial z}\left[\langle H_t^2\rangle - \langle H_t\rangle^2\right] \end{aligned} \tag{6.68}$$

where $\langle\cdot\rangle$ denotes the expectation with respect to the Brownian motions b_i. The two last terms come from the replacement of $\langle H_t^2\rangle$ by $\langle H_t\rangle^2$ in the third term of the right hand side of (6.68). By setting

$$F_t(z) = \langle H_t(z)^2\rangle - \langle H_t(z)\rangle^2\,, \tag{6.69}$$

Eq. (6.68) can be rewritten as

$$\begin{aligned}\frac{\partial\langle H_t\rangle}{\partial t} &= \left[\langle H_t\rangle + z\frac{\partial\langle H_t\rangle}{\partial z}\right] - \frac{\partial\langle H_t\rangle}{\partial z} + \frac{1}{M\beta}(2-\beta)\left[2\frac{\partial\langle H_t\rangle}{\partial z} + z\frac{\partial^2\langle H_t\rangle}{\partial z^2}\right] \\ &+ q\left[\frac{\partial\langle H_t\rangle}{\partial z} + \langle H_t\rangle^2 + 2z\langle H_t\rangle\frac{\partial\langle H_t\rangle}{\partial z}\right] + q\left[F_t + z\frac{\partial F_t}{\partial z}\right].\end{aligned} \tag{6.70}$$

To simplify notations, we will now omit the $\langle\cdot\rangle$ and write H instead of $\langle H\rangle$. The stationary solution of (6.70) writes simply as

$$\begin{aligned}&\left[H + z\frac{dH}{dz}\right] - \frac{dH}{dz} + \frac{1}{M\beta}(2-\beta)\left[2\frac{dH}{dz} + z\frac{d^2H}{dz^2}\right] \\ &+q\left[\frac{dH}{dz} + H^2 + 2zH\frac{dH}{dz}\right] + q\left[F + z\frac{dF}{dz}\right] = 0\,.\end{aligned} \tag{6.71}$$

which can be rewritten as

$$\begin{aligned}&qH(z)[H(z) + 2z\frac{dH}{dz}] + \frac{1}{2}[H(z) + 2z\frac{dH}{dz}] + \frac{1}{2}H(z) + \frac{1}{2M\beta}(2-\beta)[3\frac{dH}{dz} + 2z\frac{d^2H}{dz^2}] \\ &+\left[\frac{1}{2M\beta}(2-\beta) + (-1+q)\right]\frac{dH}{dz} + q\left[F + z\frac{dF}{dz}\right] = 0\,.\end{aligned} \tag{6.72}$$

Eq. (6.72) can be integrated easily by doing the change of function $G(z) = zH(z^2)$. Indeed, we just need to write (6.72) for z^2 instead of z and then multiply the corresponding equation by z to obtain the following equation

$$\begin{aligned}&qG\frac{dG}{dz} + \frac{1}{2}\left[z\frac{dG}{dz} + G\right] + \frac{1}{4M\beta}(2-\beta)\frac{d^2G}{dz^2} \\ &+\frac{1}{2}\left[\frac{1}{2M\beta}(2-\beta) + (-1+q)\right]2z\frac{dH}{dz}(z^2) + qz\left[F(z^2) + z^2\frac{dF}{dz}(z^2)\right] = 0\,.\end{aligned} \tag{6.73}$$

Equation (6.73) can be straightforwardly integrated with respect to z as

$$qG^2 + zG + \frac{1}{2M\beta}(2-\beta)\frac{dG}{dz} + \left[\frac{1}{2M\beta}(2-\beta) + \beta(-1+q)\right]\frac{G(z)}{z} + qz^2F(z^2) = -1 \tag{6.74}$$

where the integration constant is chosen so that $zG(z) \sim -1$ when $|z| \to \infty$. Note that the asymptotic behavior for H is therefore also $zH \sim -1$ when $|z| \to \infty$ as is expected for the Stieltjes transform of a probability measure. Rewriting now Equation (6.74) in term of the function H, we obtain

$$qH^2 + H\left[1 + \frac{1}{z}\left(q - 1 + \frac{1}{M}\frac{2-\beta}{\beta}\right)\right] + \frac{1}{M}\frac{2-\beta}{\beta}\frac{dH}{dz} + \frac{1}{z} + q\,F(z) = 0\,. \tag{6.75}$$

Now, using the result about the noise in the Marčenko-Pastur law obtained in [105] by Lytova and Pastur, we know that in the limit of large M, we have for $\beta = 1$ or 2,

$$\begin{aligned}&F(z) \sim_{M\to\infty} \\ &\frac{1}{M^2}\frac{1}{q^2}\frac{1}{2\beta\pi^2}\int_{\gamma_-}^{\gamma_+}\int_{\gamma_-}^{\gamma_+}\frac{d\lambda d\mu}{(\lambda-z)^2(\mu-z)^2}\frac{4q-(\lambda-(1+q))(\mu-(1+q))}{\sqrt{4q-(\lambda-(1+q))^2}\sqrt{4q-(\mu-(1+q))^2}}\end{aligned} \tag{6.76}$$

with $\gamma_\pm = (1 \pm \sqrt{q})^2$.

The idea to obtain the correction to the Marčenko-Pastur law is to use perturbation theory in Eq. (6.75). More precisely, we want to compute explicitly the coefficients ρ_0, ρ_1 and ρ_2 such that the eigenvalue density of a β-Wishart matrix writes under the form, in the limit of large N, M with $N/M = q$,

$$\rho(\lambda) = \rho_0(\lambda) + \frac{1}{M}\rho_1(\lambda) + \frac{1}{M^2}\rho_2(\lambda) + o\left(\frac{1}{M^2}\right). \tag{6.77}$$

Note that this asymptotic expansion (6.77) is obtained by perturbation theory and therefore is valid only for the values of λ such that the correction terms $\rho_1(\lambda)/M$ and $\rho_2(\lambda)/M^2$ are negligible compared to the leading term $\rho_0(\lambda)$ in the limit of large M, i.e. for the values of λ such that $\rho_0(\lambda) \neq 0$. The expansion (6.77) is not valid outside the Marčenko-Pastur sea, i.e., it breaks down near the edges (see below). In addition, here we are talking about smoothed density, hence it contains no oscillatory term in the finite N bulk corrections [75].

To this purpose, we first write $H(z)$ under the form

$$H(z) = H_0(z) + \frac{1}{M} H_1(z) + \frac{1}{M^2} H_2(z) + o\left(\frac{1}{M^2}\right) \tag{6.78}$$

and we plug Eq. (6.78) into Eq. (6.75). By solving the equation to leading order, we find the following expression for $H_0(z)$

$$\begin{aligned} H_0(z) &= \frac{1}{2q} \frac{-(z+q-1) + \sqrt{(z-\gamma_-)(z-\gamma_+)}}{z} \\ &= \frac{1}{2q} \frac{-(z+q-1) + \sqrt{(z-(1+q))^2 - 4q}}{z} \,. \end{aligned} \tag{6.79}$$

We deduce from this the famous Marčenko-Pastur result: the eigenvalue density converges in the limit of large N, M with $N/M = q$ to the Marčenko-Pastur density as expected given by

$$\rho_0(\lambda) = \frac{1}{2\pi q} \frac{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}}{\lambda} \,.$$

For all value of $\beta > 0$, we can now compute the $1/M$ correction to the Marčenko-Pastur density by plugging Eq. (6.78) into Eq. (6.75) and solve to order $1/M$. This gives the following expression for $H_1(z)$

$$H_1(z) = -\left(\frac{1}{\beta} - \frac{1}{2}\right) \frac{1}{q} \left[\frac{1}{2} \left(\frac{1}{z-\gamma_+} + \frac{1}{z-\gamma_-} \right) - \frac{1}{\sqrt{(z-\gamma_+)(z-\gamma_-)}} \right]$$

and the corresponding $1/N$ correction to the density is then given (for all β) by

$$\rho_1(\lambda) = \left(\frac{1}{\beta} - \frac{1}{2}\right) \frac{1}{q} \left[\frac{1}{2} \left(\delta(\lambda-\gamma_+) + \delta(\lambda-\gamma_-) \right) - \frac{1}{\pi} \frac{d\lambda}{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}} \right] . \tag{6.80}$$

Comparing $H_1(z)/M$ with $H_0(z)$, we see that the correction term ceases to be negligible when $|\lambda - \gamma_\pm| \sim M^{-2/3}$, as expected: this is indeed the standard edge scaling that defines the Tracy-Widom region. Note that this $\mathcal{O}(1/N)$ correction term was derived earlier [[75] and references therein] for $\beta = 1$, 2 and 4, but our result is valid for general β.

For the particular value $\beta = 1$ or 2, we can use the result of Lytova and Pastur stated above in Eq. (6.76) to compute the $1/M^2$ correction with the same method by solving the equation until order $1/M^2$.

Let us first compute an explicit expression for $F(z)$ from the integral representation in Eq.

(6.76)

$$
\begin{aligned}
q^2M^2F(z) &= \frac{1}{2\beta\pi^2}\left[4q\left(\int_{\gamma_-}^{\gamma_+}\frac{1}{(\lambda-z)^2}\frac{d\lambda}{\sqrt{4q-(\lambda-(1+q))^2}}\right)^2\right. \\
&\left.-\left(\int_{\gamma_-}^{\gamma_+}\frac{d\lambda}{(\lambda-z)^2}\frac{\lambda-(1+q)}{\sqrt{4q-(\lambda-(1+q))^2}}\right)^2\right]+o(1) \\
&= \frac{1}{2\beta\pi^2}\left[4q\left(-\frac{\pi}{(z-\gamma_-)(z-\gamma_+)}\frac{2z-\gamma_--\gamma_+}{\sqrt{(z-\gamma_-)(z-\gamma_+)}}\right)^2\right. \\
&\left.-\frac{\pi^2}{(z-\gamma_-)(z-\gamma_+)}\left(1-\frac{1}{2}\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)(z-\gamma_+)}\right)^2\right]+o(1) \\
&= \frac{1}{2\beta}\frac{1}{(z-\gamma_-)(z-\gamma_+)}\left[4q\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)^2(z-\gamma_+)^2}-\left(1-\frac{1}{2}\frac{(2z-\gamma_--\gamma_+)^2}{(z-\gamma_-)(z-\gamma_+)}\right)^2\right]+o(1) \\
&= \frac{1}{2\beta}\frac{1}{(z-\gamma_-)(\gamma_+-z)}+o(1)\,.
\end{aligned}
$$

Then we can turn to compute $H_2(z)$ and deduce from this computation the expression for $\rho_2(\lambda)$

$$
\begin{aligned}
\rho_2(\lambda) = -2\left(\frac{1}{2}-\frac{1}{\beta}\right)^2\frac{1}{q}\frac{1}{\sqrt{(\lambda-\gamma_-)(\gamma_+-\lambda)}}\left[\frac{1}{2}\left(\frac{1}{\lambda-\gamma_-}+\frac{1}{\lambda-\gamma_+}\right)\right. \\
\left.-\frac{\lambda}{2}\left(\frac{1}{(\lambda-\gamma_-)^2}+\frac{1}{(\lambda-\gamma_+)^2}\right)\right]+\frac{1}{2q\beta}\frac{\lambda}{(\lambda-\gamma_-)^{3/2}(\gamma_+-\lambda)^{3/2}}\,.
\end{aligned}
$$

Again, the comparison of this correction term with the dominant term indicates that our perturbation expansion breaks down when $|\lambda-\gamma_\pm|\sim M^{-2/3}$.

6.5 Conclusions

In summary, we proposed a random matrix model (invariant under similarity transformations) whose joint density of eigenvalues is given by the classical β-Wishart ensemble where the quantization of the Dyson index β is lifted. The procedure is constructive and is described in section 6.3.2. The resulting ensemble is by construction invariant under similarity transformations with Haar distributed eigenvectors The diffusive evolution equation for the eigenvalues involves the Dyson index of the ensemble as a free parameter. Letting it scale with the size M of the matrix, the spectral density of the ensemble becomes a one-parameter continuous family interpolating between the familiar Marčenko-Pastur distribution and a certain type of Gamma distribution. On the other hand, keeping the Dyson index unscaled but not quantized, we showed that a careful analysis of the full Stieltjes transform equation lead naturally to $1/N$ and $1/N^2$ corrections (and possibly systematically to any order) to the average spectral density (Marčenko-Pastur) for all $\beta\neq 2$. This then extends the previous work [75, 73] on the $\mathcal{O}(1/N)$ correction term for $\beta=1$, 2 and 4. To order $\mathcal{O}(1/N^2)$, our result is valid for $\beta=1$ and $\beta=2$. It would be interesting to see (or conjecture) if this formula to $\mathcal{O}(1/N^2)$ term is valid for general β.

In this work, we have computed the crossover density as a function of the interpolating parameter c. In the limit $c\to\infty$, it reduces to the standard Marčenko-Pastur density, whereas the opposite limit $c\to 0$ corresponds to the Gamma laws associated with the CIR process. It would be interesting to extend our analysis to the distribution of the largest eigenvalue. As in the case of bulk density, we would expect a c-dependent distribution for the largest eigenvalue, properly centered and scaled, interpolating between the Tracy-Widom distribution ($c\to\infty$ limit) and Gumbel distribution (as $c\to 0$).

Derivation of (6.16)

The Fokker-Planck equation for the transition probability density $P(\lambda_1, \cdots, \lambda_N; t)$ of the process $(\lambda_1(t), \cdots, \lambda_N(t))$ which satisfies the stochastic differential system (6.15) reads

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial \lambda_i} \left[P \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) \right] + 2 \sum_{i=1}^{N} \frac{\partial^2}{\partial \lambda_i^2} \left[\lambda_i P \right] . \tag{6.81}$$

The stationary solution is the solution which does not depend on time t, satisfying

$$-\sum_{i=1}^{N} \frac{\partial}{\partial \lambda_i} \left[P \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) \right] + 2 \sum_{i=1}^{N} \frac{\partial^2}{\partial \lambda_i^2} \left[\lambda_i P \right] = 0 . \tag{6.82}$$

It is easy to check using elementary algebra that the jpdf P^* defined in (6.16) verifies Eq. (6.81) as in fact we can verify that for all i,

$$2 \frac{\partial}{\partial \lambda_i} \left[\lambda_i P^* \right] = P^* \left(-\lambda_i + pM + (1-\delta)p + p \sum_{k \neq i} \frac{\lambda_i + \lambda_k}{\lambda_i - \lambda_k} \right) . \tag{6.83}$$

Algorithmic description of how to build the process $\mathbf{W}_t^n$ in practice

Let us describe shortly an algorithmic description of how to build the process $\mathbf{W}_t^n$ in practice, on a discrete grid. First note that this algorithmic description needs a discrete grid and that it does not reproduce exactly the process $\mathbf{W}_t^n$ but only a discretized approximation of it. Choose a large value of n and an initial symmetric matrix $\mathbf{W}_0$. The construction is iterative. Suppose that the process is constructed until time k/n and let us explain how to compute the matrix $\mathbf{W}_{(k+1)/n}^n$ at the next discrete time of the grid, $(k+1)/n$.

1. Step 1. We first need to compute the matrix $\sqrt{\mathbf{W}_{k/n}^n}$. It suffices to compute the orthogonal matrix $\mathbf{O}_{k/n}^n$ such that
$$\mathbf{W}_{k/n}^n = \mathbf{O}_{k/n}^n \mathbf{\Sigma}_{k/n}^n {\mathbf{O}_{k/n}^n}^\dagger$$
where $\mathbf{\Sigma}_{k/n}^n$ is the diagonal matrix composed of the eigenvalues of $\mathbf{W}_{k/n}^n$ (in increasing order). The eigenvalues of the matrix $\mathbf{W}_{k/n}^n$ should be non negative as the eigenvalues process of $\mathbf{W}_t^n$ are almost surely non negative at all time t. However, due to the discretization scheme necessary for algorithmic procedure, the non-negativity can fail. To avoid this problem, we define $\sqrt{\mathbf{W}_{k/n}^n}$ as
$$\sqrt{\mathbf{W}_{k/n}^n} = \mathbf{O}_{k/n}^n \sqrt{\mathbf{\Sigma}_{k/n}^n} {\mathbf{O}_{k/n}^n}^\dagger \tag{6.84}$$
where $\sqrt{\mathbf{\Sigma}_{k/n}^n}$ is the diagonal matrix composed of the square roots of the absolute values of the eigenvalues of $\mathbf{W}_{k/n}^n$ (again in increasing order).

2. Step 2. We sample the Bernoulli random variable ϵ_k^n with $\mathbb{P}[\epsilon_k^n = 1] = p = 1 - \mathbb{P}[\epsilon_k^n = 0]$.

3. Step 3. It depends on the value of ϵ_k^n:

   - if $\epsilon_k^n = 1$, we sample a $N \times N$ matrix $\mathbf{G}_n$ filled with independent Gaussian variables with mean 0 and variance $1/n$ and then we compute the matrix $\mathbf{W}_{(k+1)/n}^n$ by the formula
$$\mathbf{W}_{(k+1)/n}^n = \left(1 - \frac{1}{n}\right) \mathbf{W}_{k/n}^n + \sqrt{\mathbf{W}_{k/n}^n}\, \mathbf{G}_n + \mathbf{G}_n^\dagger \sqrt{\mathbf{W}_{k/n}^n} + \frac{1}{n} M\, \mathbf{I} .$$

- if $\epsilon_k^n = 0$, we sample N independent Gaussian variables $(z_1, \cdots, z_N)$ with mean 0 and variance $1/n$. We then compute the matrix $\mathbf{Y}_n$, which is co diagonalizable with the matrix $\mathbf{W}_{k/n}^n$, defined as the product

$$\mathbf{Y}_n := \mathbf{O}_{k/n}^n \mathrm{Diag}\left(z_1, z_2, \ldots, z_N\right) {\mathbf{O}_{k/n}^n}^{\dagger}. \tag{6.85}$$

  Finally we obtain the matrix $\mathbf{W}_{(k+1)/n}^n$ by

$$\mathbf{W}_{(k+1)/n}^n = \left(1 - \frac{1}{n}\right) \mathbf{W}_{k/n}^n + \sqrt{\mathbf{W}_{k/n}^n}\, \mathbf{Y}_n + \mathbf{Y}_n^{\dagger} \sqrt{\mathbf{W}_{k/n}^n} + \frac{1}{n}\delta\, \mathbf{I}.$$

Acknowledgments

We would like to thank A. Comtet and A. Guionnet for useful discussions. SNM acknowledges support by ANR grant 2011-BS04-013-01 WALKMAT.

Chapter 7

Eigenvector dynamics: general theory and some applications

Résumé

Cet article est publié dans *Physical Review E* et est écrit en collaboration avec Jean-Philippe Bouchaud. Nous proposons un cadre de travail général pour étudier la stabilité du sous espace engendré par P vecteurs propres associés à des valeurs propres consécutives d'une matrice symmétrique $\mathbf{H}_0$ sous l'effet d'une petite perturbation additive. Ce problème intervient dans divers contextes, notamment la dissipation quantique (dans ce cas, $\mathbf{H}_0$ est l'Hamiltonien du système) et le contrôle du risque en finance (dans ce cas la matrice $\mathbf{H}_0$ est la matrice de covariance des rendements des actions). L'idée est de calculer les valeurs singulières de la matrice des chevauchements ỳartir desquelles on peut calculer une distance mesurant le chevauchement de deux sous-espaces vectoriels. Nous nous concentrons sur le cas particulier où la matrice $\mathbf{H}_0$ est une matrice aléatoire de l'ensemble orthogonal Gaussien. Dans ce cas, le spectre des valeurs singulières admet une forme limite explicite dans la limite des grandes matrices. Nous nous intéressons aussi au cas où $\mathbf{H}_0$ est une matrice de covariance et nous illustrons l'utilité pratique de nos résultats par des applications en finance. Le cas particulier où cette matrice de covariance a une valeur propre beaucoup plus grande que toutes les autres est traité très précisément. En particulier, la dynamique de l'angle entre les plus grands vecteurs propres de la vraie matrice de covariance et de la matrice de covariance empirique fait apparaître une nouvelle classe de processus stochastiques.

Abstract

We propose a general framework to study the stability of the subspace spanned by P consecutive eigenvectors of a generic symmetric matrix $\mathbf{H}_0$, when a small perturbation is added. This problem is relevant in various contexts, including quantum dissipation ($\mathbf{H}_0$ is then the Hamiltonian) and financial risk control (in which case $\mathbf{H}_0$ is the assets return covariance matrix). We argue that the problem can be formulated in terms of the singular values of an overlap matrix, which allows one to define an overlap distance. We specialize our results for the case of a Gaussian Orthogonal $\mathbf{H}_0$, for which the full spectrum of singular values can be explicitly computed. We also consider the case when $\mathbf{H}_0$ is a covariance matrix and illustrate the usefulness of our results using financial data. The special case where the top eigenvalue is much larger than all the other ones can be investigated in full detail. In particular, the dynamics of the angle made by the top eigenvector and its true direction defines an interesting new class of random processes.

7.1 Introduction

Random Matrix Theory (RMT) is extraordinarily powerful at describing the eigenvalues statistics of large random, or pseudo-random, matrices [145, 3, 15, 114]. Eigenvalue densities, two-

point correlation functions, level spacing distributions, etc. can be characterized with exquisite details. The "dynamics" of these eigenvalues, i.e. the way these eigenvalues evolve when the initial matrix $\mathbf{H}_0$ is perturbed by some small matrix $\varepsilon\mathbf{P}$, is also well understood [140]. The knowledge of the corresponding eigenvectors is comparatively much poorer (but see [153]). One reason is that many RMT results concern rotationally invariant matrix ensembles, such that by definition the statistics of eigenvectors is featureless. Still, as we will show below, some interesting results can be derived for the dynamics of these eigenvectors. Let us give two examples for which this question is highly relevant.

One problem where the evolution of eigenvectors is important is Quantum Dissipation [154] (see also the related recent strand of the literature on Quantum "Fidelity" [71]). As the parameters of the Hamiltonian $\mathbf{H}_t = \mathbf{H}_0 + \varepsilon\mathbf{P}_t$ of a system evolve with time t, the average energy changes as well. One term corresponds to the average (reversible) change of the Hamiltonian which leads to a shift of the energy levels (the eigenvalues). But if the external perturbation is not infinitely slow, some transitions between energy levels will take place, leading to a dissipative (irreversible) term in the evolution equation of the average energy of the system. The adiabaticity condition which ensures that no transition takes place amounts to comparing the speed of change of the perturbation $\varepsilon\mathbf{P}_t$ with a quantity proportional to the typical spacing between energy levels. For systems involving a very large number N of degrees of freedom, the average level spacing of the $N \times N$ Hamiltonian $\mathbf{H}$ goes to zero as N^{-1}. For $N \to \infty$, any finite speed of change therefore corresponds to the "fast" limit, where a large number of transitions between states is expected. In fact, if the quantum system is in state $|\phi_i^0\rangle$ at time $t = 0$, which corresponds to the ith eigenvector of $\mathbf{H}_0$, the probability to jump to the jth eigenvector of $\mathbf{H}_1$, $|\phi_j^1\rangle$, at time $t = 1$ is given by $|\langle\phi_j^1|\phi_i^0\rangle|^2$, where we use the bra-ket notation for vectors and scalar products. The way energy is absorbed by the system will therefore be determined by the perturbation-induced distortion of the eigenvectors. More precisely, if $|\phi_i^0\rangle$ is different from $|\phi_i^1\rangle$, some transitions must take place in the non-adiabatic limit, which involve all the states j that have a significant overlap with the initial state.

Another very relevant situation is Quantitative Finance, where the covariance matrix $\mathbf{C}$ between the returns of N assets (for example stocks) plays a major role in risk control and portfolio construction [44]. More precisely, the risk of a portfolio which invests w_α in asset α is given by $\mathcal{R}^2 = \sum_{\alpha\beta} w_\alpha \mathbf{C}_{\alpha\beta} w_\beta$. Constructing low risk portfolios requires the knowledge of the n largest eigenvalues of $\mathbf{C}$ (n is often chosen empirically, keeping only the statistically meaningful eigenvalues which lie outside the Marchenko-Pastur sea, see [45] for details), $\lambda_1 \geqslant \ldots \geqslant \lambda_n$ and their corresponding eigenvectors $|\phi_1\rangle, \ldots, |\phi_n\rangle$. The top eigenvalues and eigenvectors represent the most risky directions in a financial context. A portfolio such that the vector of weights $|w\rangle$ has zero overlap with the first n eigenvectors of $\mathbf{C}$ has a risk which is bounded from above by λ_{n+1}. The problem with this idea is that it relies on the assumption that the covariance matrix $\mathbf{C}$ is perfectly known and *constant in time*. The observation of a sufficiently long time series of past returns would thus allow one, in such a stable world, to determine $\mathbf{C}$ and to immunize the portfolio against risky investment modes.

Unfortunately, this idea is thwarted by two (inter-related) predicaments: a) time series are always of finite length, and lead to substantial "noise" in empirical estimates of $\mathbf{C}$ [44] and b) the world is clearly not stationary and there is no guarantee that the covariance matrix corresponding to the pre-crisis period 2000-2007 is the same as the one corresponding to the period 2008-2011. For one thing, some companies disappear and others are created in the course of time. But even restricted to companies which exist throughout the whole period, it is by no means granted that the correlation between stock returns do not evolve in time. This is why it is common practice in the financial industry to restrict the period used to determine the covariance matrix to windows of a few years into the recent past. This leads to the measurement noise problem alluded above. Now, if the "future" large eigenvectors do not coincide with the past ones, a supposedly low risk portfolio will in fact be exposed to large risks directions in the future. Denoting as $|\phi_i^0\rangle$ the past eigenvectors and $|\phi_j^1\rangle$ the future ones, the total risk of the portfolio $|w\rangle = |\phi_i^0\rangle$ can be defined as $\sum_{j=1}^N \lambda_j^1 \langle\phi_j^1|\phi_i^0\rangle^2$. Therefore, as for the quantum dissipation problem, the statistics of the overlaps $\langle\phi_j^1|\phi_i^0\rangle$ is a crucial piece of information.

In practice, one computes the empirical covariance matrix $\mathbf{E}$ using past stock returns, which

is defined as:

$$\mathbf{E}_{ij} = \frac{1}{T}\sum_{t=1}^{T} r_i^t r_j^t \,,$$

where T is the length of the period on which the measurement is done and r_i^t is the return of stock number i on day t. If the true covariance matrix $\mathbf{C}$ exists and is stable in this period, the empirical matrix $\mathbf{E}$ can be seen as a perturbation of $\mathbf{C}$, since one can write $\mathbf{E} = \mathbf{C} + \mathcal{E}$ where $\mathcal{E}$ is a matrix whose elements are of order $1/\sqrt{T}$ (by the central limit theorem) to be considered small as T is usually quite large. In this sense, the problem falls in the more general context introduced above.

The paper is organized as follows. In the next section 7.2, we introduce the main statistical tools and objects studied in different contexts in the following sections and we also briefly recall standard perturbation theory. In the next two sections, we turn to two explicit illustrations, first in the context of matrices in the Gaussian Orthogonal ensemble (GOE), and then in the context of covariance matrices. More precisely, in section 7.3, we study the stability of the eigenvectors for a matrix $\mathbf{H}_0$ in the GOE by computing the "overlap distance" between the perturbed space and the non-perturbed space in the limit of large matrices $\mathbf{H}_0$ when the perturbation matrix $\mathbf{P}$ is also in the GOE. Furthermore, we are able to compute the full spectrum of the overlap matrix in this limit, which gives a precise idea of the perturbation induced distortion for the eigenvectors. In section 7.4, we go through the same steps in the context of covariance matrices. We study the link between the *population* eigenvectors (the eigenvectors of the *true* covariance matrix) and the *sample* eigenvectors (the eigenvectors of the empirical covariance matrix). Then, in section 7.5, we analyse more precisely the case of a population covariance matrix with an isolated top eigenvalue much larger than the other ones. We measure the empirical covariance matrix with an exponential moving average estimator and characterize the temporal evolution of the angle made by the top eigenvector and its true direction which defines an interesting new class of random processes. Finally, in section 7.6, we apply our ideas to the analysis of financial market correlations. Our purpose here is to study whether correlations between stock returns evolve or not. In particular, is there a constant in time correlation matrix (population correlation matrix)? Do the economical sectors (eigenvectors of the correlation matrix) evolve or not ? We find that there is indeed a genuine evolution of the correlation matrix of stocks returns for different markets in the U.S, in Europe and in Japan, a result which confirms recent studies (see e.g. [27, 8, 116]). We also give a partial description of this temporal evolution.

7.2 Perturbation theory and Statistical tools

In this section, we first recall the perturbation theory for the eigenvalues and eigenvectors (see Eq. (7.3) and (7.2)) when the perturbed matrix $\mathbf{H}_1$ writes as in Eq. (7.1). Then we define the two main objects of the paper, the overlap matrix $\mathbf{G}$ and the overlap distance D (see Eq. (7.4) and (7.5)) useful for the comparison of the two perturbed and non perturbed eigen subspaces, that will be studied in different contexts in sections 7.3, 7.4 and also in 7.6 for financial applications. At the end of this section, we compute asymptotic expressions for the two objects $\mathbf{G}$ and D using the perturbation theory's equation for the eigenvectors, when the perturbation is of the form (7.1) (see Eq. (7.7), (7.8) for the matrix $\mathbf{G}^\dagger\mathbf{G}$ and (7.9) for the distance D). Those computations will be very useful later in sections 7.3 and 7.4. The last definition of the matrix $\mathbf{\Sigma}$ defined in Eq. (7.10) and the expression (7.11) for its entries in the perturbative regime will be convenient and used later.

In the whole paper, we will mainly be interested in the eigenvectors of a matrix $\mathbf{H}_1$ which can be written as

$$\mathbf{H}_1 = \mathbf{H}_0 + \varepsilon\mathbf{P} \tag{7.1}$$

where $\mathbf{H}_0$ and $\mathbf{P}$ are two $N \times N$ symmetric matrices and ε a small (positive) parameter. The matrix $\mathbf{H}_0$ is the *true signal* which is perturbed by the adding of the small term $\varepsilon\mathbf{P}$. The matrix $\mathbf{H}_1$ will be referred as the *perturbed* matrix. The eigenvalues of the matrix $\mathbf{H}_i, i = 0, 1$ will be denoted as $\lambda_1^i \geqslant \lambda_2^i \geqslant \ldots \geqslant \lambda_N^i$ and the corresponding eigenvectors $|\phi_1^i\rangle, \ldots, |\phi_N^i\rangle$.

Our aim is to describe the relation between the perturbed eigenvectors $|\phi_i^1\rangle$ and the non-perturbed eigenvectors $|\phi_i^0\rangle$ when the parameter ε tends to 0.

When trying to follow the evolution of a given eigenvector $|\phi_i\rangle$ when the small perturbation $\varepsilon\mathbf{P}$ is added, one immediately faces a problem if the neighbouring eigenvalues of λ_i^0 are too close to λ_i^0. For example, if the distance between the eigenvalues λ_i^0 and λ_{i+1}^0 is very small, the eigenvectors $|\phi_i\rangle$ and $|\phi_{i+1}\rangle$ will strongly hybridize (this phenomenon was observed for example in [147, Fig. 1]). The eigenvector $|\phi_i^0\rangle$ will in fact hybridize with all the perturbed eigenvectors $|\phi_j^1\rangle$, with stronger overlaps for those associated to eigenvalues $\lambda_j^0, j \neq i$ which are close to λ_i^0. This idea can be made precise by using standard perturbation theory to second order in ε: the perturbed eigenvectors can be expressed in terms of the initial eigenvectors, for small ε, as:

$$|\phi_i^1\rangle = \left(1 - \frac{\varepsilon^2}{2}\sum_{j\neq i}\left(\frac{P_{ij}}{\lambda_i^0-\lambda_j^0}\right)^2\right)|\phi_i^0\rangle + \varepsilon\sum_{j\neq i}\frac{P_{ij}}{\lambda_i^0-\lambda_j^0}|\phi_j^0\rangle \\ + \varepsilon^2\sum_{j\neq i}\frac{1}{\lambda_i^0-\lambda_j^0}\left(\sum_{\ell\neq i}\frac{P_{j\ell}P_{\ell i}}{\lambda_i^0-\lambda_\ell^0} - \frac{P_{ii}P_{ij}}{\lambda_i^0-\lambda_j^0}\right)|\phi_j^0\rangle \tag{7.2}$$

where $P_{ij} \equiv \langle\phi_j^0|\mathbf{P}|\phi_i^0\rangle$. The denominators $\lambda_i^0 - \lambda_j^0$ remind us that the eigenvector $|\phi_i^0\rangle$ can have very large overlaps with the eigenvectors associated to the closest eigenvalues to λ_i^0. This fact makes difficult to follow the evolution of one single eigenvector in the case of small spacings between the eigenvalues (this will happen when the dimension of the matrix is large, see below). We mention in passing that perturbation theory to second order in ε for the eigenvalues gives

$$\lambda_i^1 = \lambda_i^0 + \varepsilon P_{ii} + \varepsilon^2\sum_{j\neq i}\frac{P_{ij}^2}{\lambda_i^0-\lambda_j^0}\,. \tag{7.3}$$

The reader can find a study of perturbed eigenvalues statistics in [150, 79].

It is important to note at this point that equations (7.2) and (7.3) are *a priori* only valid in the perturbative regime, i.e. when the entries of the perturbation matrix $\varepsilon\mathbf{P}$ are small compared to the level spacing of the non-perturbed matrix $\mathbf{H}_0$. This condition ensures that the asymptotic correction terms appearing in (7.2) and (7.3) are small compared to the leading term of order 1 corresponding to the non-perturbed system.

The idea is then to study the stability of a whole subspace V_0 spanned by $2p+1$ several consecutive eigenvalues: $\{|\phi_{k-p}^0\rangle, \ldots |\phi_k^0\rangle, \ldots, |\phi_{k+p}^0\rangle\}$. Motivated by the above examples, we ask the following question: how should one choose $q \geqslant p$ such that the subspace V_1 of dimension $2q+1$ spanned by the set $\{|\phi_{k-q}^1\rangle, \ldots |\phi_k^1\rangle, \ldots, |\phi_{k+q}^1\rangle\}$ has a significant overlap with the initial subspace? In order to answer this question, we consider the $(2q+1)\times(2p+1)$ rectangular matrix of overlaps $\mathbf{G}$ with entries:

$$G_{ij} := \langle\phi_i^1|\phi_j^0\rangle\,. \tag{7.4}$$

The $(2p+1)$ non zero singular values $1 \geqslant s_1 \geqslant s_2 \geqslant \ldots \geqslant s_{2p+1} \geqslant 0$ of $\mathbf{G}$ give full information about the overlap between the two spaces. For example, the largest singular value s_1 indicates that there is a certain linear combination of the $(2q+1)$ perturbed eigenvectors that has a scalar product s_1 with a certain linear combination of the $(2p+1)$ unperturbed eigenvectors. If $s_{2p+1} = 1$, then the initial subspace is entirely spanned by the perturbed subspace. If on the contrary $s_1 \ll 1$, it means that the initial and perturbed eigenspace are nearly orthogonal to one another since even the largest possible overlap between any linear combination of the original and perturbed eigenvectors is very small. A good measure of what can be called an overlap distance $D(V_0, V_1)$ between the two spaces V_0 and V_1 is provided by the average of the logarithm of the singular values:

$$D(V_0, V_1) := -\frac{\sum_i \ln s_i}{2p+1}, \tag{7.5}$$

but alternative measures, such as $1 - \sum_i s_i/(2p+1)$, can be considered as well. Since the singular values s are obtained as the square-root of the eigenvalues of the matrix $\mathbf{G}^\dagger\mathbf{G}$, one has $D \equiv -\ln\det\mathbf{G}^\dagger\mathbf{G}/2P$, where we henceforth introduce for convenience the notations $P = 2p+1$, $Q = 2q+1$. The overlap distance D was originally studied for $P = Q$ in [14], see e.g. [84, 147],

where a fundamental effect observed in many body systems, called the *Anderson Orthogonality catastrophe* (AOC) is introduced. Anderson [14] addressed the ground state of a finite system consisting of P noninteracting electrons. Upon the introduction of a finite rank perturbation matrix $\varepsilon\mathbf{P}$, this ground state gets modified. It is then shown that the overlap between the original and the modified P-electron ground state, which is in fact exactly given by our overlap distance $D(V_0, V_1)$ between the two subspaces V_0 and V_1 (with $P = Q$), is proportional to a negative power of P, and vanishes in the thermodynamic $P \to +\infty$ limit, hence the catastrophe. We will see that our idea of introducing a rectangular $Q \times P$ overlap matrix $\mathbf{G}$ enables to avoid this orthogonality catastrophe. Our objects introduced here will also allow us to revisit the AOC in the case of square matrices $\mathbf{G}$ showing that it occurs for the random matrix model studied in section 7.3 (AOC for this RM model is also studied in [147]). In [84, 147], the AOC is also investigated through random matrix models as in our paper. The main difference with [84] is that we consider here full rank perturbation instead of a localized perturbation of rank 1, for which one can do explicit computations (and so treat the non-perturbative regime).

As an interesting benchmark, consider the case when two subspaces W_0 and W_1 respectively of dimensions P and Q are constructed using randomly chosen orthonormal vectors in a space of dimension N. In this case, one expects accidental overlaps, such that the s_i are in fact non zero, and therefore $D(W_0, W_1)$ is finite. This distance can be calculated exactly using Random Matrix Theory tools in the limit $N, P, Q \to \infty$, with $\alpha = P/N$ and $\beta = Q/N$ held fixed. The result is [46]:

$$D_{RMT}(W_0, W_1) = -\int_0^1 \mathrm{d}s \ln(s) \frac{\sqrt{(s^2-\gamma_-)_+(\gamma_+-s^2)_+}}{\beta\pi s(1-s^2)}$$

where $\gamma_\pm = \alpha+\beta-2\alpha\beta \pm 2\sqrt{\alpha\beta(1-\alpha)(1-\beta)}$. In other words, in that limit, the full density of singular values is known; all singular values are within the interval $[\sqrt{\gamma_-}, \sqrt{\gamma_+}]$. This provides a benchmark to test whether the two eigenspaces are accidentally close ($D \approx D_{RMT}$), or if they are genuinely similar ($D \ll D_{RMT}$).

Endowed with the above formalism, we can now proceed to compute $D(V_0, V_1)$ in the case where the perturbation is small. Indeed equation (7.2) allows us to obtain the overlap matrix $\mathbf{G}$. Keeping only the relevant terms to order ε^2, we find:[1]

$$G_{ij} = \begin{cases} 1 - \frac{\varepsilon^2}{2} \sum_{\ell \neq i} \left(\frac{P_{i\ell}}{\lambda_i^0 - \lambda_\ell^0} \right)^2 & \text{if } i = j, \\ \frac{\varepsilon P_{ij}}{\lambda_i^0 - \lambda_j^0} + \frac{\varepsilon^2}{\lambda_i^0 - \lambda_j^0} \left(\sum_{\ell \neq i} \frac{P_{j\ell} P_{\ell i}}{\lambda_i^0 - \lambda_\ell^0} - \frac{P_{ii} P_{ij}}{\lambda_i^0 - \lambda_j^0} \right) & \text{if } i \neq j. \end{cases} \tag{7.6}$$

Using (7.6), we can also compute the matrix $\mathbf{G}^\dagger\mathbf{G}$ to second order in ε, we obtain for $i \neq j$:

$$(G^\dagger G)_{ij} = -\varepsilon^2 \sum_{\ell \notin \{k-q;\ldots;k+q\}} \frac{P_{\ell i} P_{\ell j}}{(\lambda_i^0 - \lambda_\ell^0)(\lambda_j^0 - \lambda_\ell^0)}, \tag{7.7}$$

and, for $i = j$:

$$(G^\dagger G)_{ii} = 1 - \varepsilon^2 \sum_{j \notin \{k-q;\ldots;k+q\}} \left(\frac{P_{ij}}{\lambda_i^0 - \lambda_j^0} \right)^2. \tag{7.8}$$

It is then easy to derive the central result of our study: to second order in ε, the distance $D(V_0, V_1)$ between the initial and perturbed eigenspaces is:

$$D(V_0, V_1) = \frac{\varepsilon^2}{2P} \sum_{i=k-p}^{k+p} \sum_{j \notin \{k-q,\ldots,k+q\}} \left(\frac{P_{ij}}{\lambda_j^0 - \lambda_i^0} \right)^2. \tag{7.9}$$

The matrices $\mathbf{G}$ and $\mathbf{G}^\dagger\mathbf{G}$ are both close to the identity matrix as they should. Let us define the matrix $\mathbf{\Sigma}$ by

$$\mathbf{\Sigma} = \frac{1}{\varepsilon^2} \left(\mathbf{I} - \mathbf{G}^\dagger\mathbf{G} \right) \tag{7.10}$$

[1] see [153] for similar calculations.

whose elements write, using the previous perturbation equations for $\mathbf{G}^\dagger\mathbf{G}$ (Eq. (7.7) and (7.8)), as

$$\Sigma_{ij} = \sum_{\ell \notin \{k-q;\dots;k+q\}} \frac{P_{\ell i} P_{\ell j}}{(\lambda_i^0 - \lambda_\ell^0)(\lambda_j^0 - \lambda_\ell^0)} \,. \tag{7.11}$$

One can note that the matrix $\mathbf{\Sigma}$ is positive definite and that its matrix elements are of order 1 when ε goes to 0.

7.3 Eigenvector stability in the GOE ensemble

We will now define a random matrix model for which we will apply the results of the previous section. Let $\mathbf{H}_0$ be a random matrix of the Gaussian Orthogonal Ensemble (GOE), i.e. a matrix of size $N \times N$ with gaussian entries randomly chosen with the probability measure on the space of real symmetric matrices

$$P(\mathrm{d}H_0) = \exp(-\frac{N}{2\sigma^2}\mathrm{tr}(H_0^2))\,\mathrm{d}H_0\,.$$

This definition implies that the matrix $\mathbf{H_0}$ is symmetric with independent Gaussian entries above the diagonal with variance σ^2/N on the diagonal and $\sigma^2/2N$ off diagonal.

The perturbation matrix is similarly defined as a random matrix of the GOE, independent of $\mathbf{H}_0$ with the same variance profile for the entries.

We then define the *perturbed* matrix $\mathbf{H}_1$ as before:

$$\mathbf{H}_1 = \mathbf{H}_0 + \varepsilon\mathbf{P}. \tag{7.12}$$

It is very well known that the density of $\mathbf{H}_0$-eigenvalues $\rho_N(\lambda) := 1/N \sum_{i=1}^N \delta_{\lambda_i}$ tends in the large N limit to the Wigner semi-circle law

$$\rho(\mathrm{d}\lambda) \equiv \frac{1}{2\pi}\sqrt{4\sigma^2 - \lambda^2}\mathrm{d}\lambda\,. \tag{7.13}$$

For simplicity, we take $\sigma^2 = 2$ in the following.

Remark. Here the choice of a GOE random matrix for $\mathbf{H_0}$ is made to get an explicit expression for the density of states in the limit of large matrices. But our theory developed in the following would apply for sequences of matrices $(\mathbf{H_0}(N))_N$ such that the density of states converges to a general (not necessarily the semi-circle density) continuous density $\rho(\lambda)\mathrm{d}\lambda$. Moreover, the sequence $(\mathbf{H_0}(N))_N$ can be supposed deterministic or random. The matrix $\mathbf{H_0}$ can be seen as the *true signal* to which a small noisy perturbation $\varepsilon\mathbf{P}$ is added.

In this whole current section, $\overline{\cdots}$ denotes an averaging over the random matrix $\mathbf{P}$ [2].

In the following subsection 7.3.1, we study the overlap distance $D(V_0, V_1)$ between two eigenspaces V_0 and V_1 (see below for definition) of the matrices $\mathbf{H}_0$ and $\mathbf{H}_1$ and we consider in particular its limit when the dimension N of the (GOE) random matrices $\mathbf{H}_0$, $\mathbf{H}_1$ and $\mathbf{P}$ tends to infinity. In subsection 7.3.2, we study the spectrum of the matrix $\mathbf{G}^\dagger\mathbf{G}$ introduced above in this context, and we characterize the limiting eigenvalue empirical distribution of $\mathbf{G}^\dagger\mathbf{G}$ in the limit of large N. The characterization appears in equation (7.20) which is an equation (with a unique solution) satisfied by the Stieltjes transform (or resolvent) of the limiting probability measure. Then, in the following subsection 7.3.3, we analyse equation (7.20) and its solution in great details so as to extract informations on the density of this distribution (it has a compact support, values of its edges,...).

[2] There is no need in averaging over the random matrix $\mathbf{H}_0$ for the following results to be valid.

7.3.1 Distance between subspaces of perturbed and non-perturbed eigenvectors

We consider the subspace V_0 of initial eigenvectors corresponding to all the eigenvalues λ contained in a certain finite interval $[a, b]$ included in the Wigner sea $[-2, 2]$. We want to compute the mean overlap distance $\overline{D}(V_0, V_1)$ between V_0 and the subspace V_1 spanned by the perturbed eigenvectors of $\mathbf{H}_1$, corresponding to all eigenvalues contained in $[a - \delta, b + \delta]$, where δ is a positive parameter.

Using formula (7.9), which is valid if the entries of the perturbation matrix $\varepsilon\mathbf{P}$ (of order $\varepsilon N^{-1/2}$) are much smaller than the mean level spacing of the matrix $\mathbf{H}_0$, of order $(N\rho(\lambda))^{-1}$, we can write for $\varepsilon \ll N^{-1/2}$:

$$\overline{D}(V_0, V_1) = \frac{\varepsilon^2}{2P} \sum_{\lambda_i^0 \in [a;b]} \sum_{\lambda_j^0 \notin [a-\delta;b+\delta]} \frac{1}{(\lambda_j^0 - \lambda_i^0)^2}. \tag{7.14}$$

It is easily seen that Eq. (7.14) becomes, in the large N limit:

$$\overline{D}(V_0, V_1) = \frac{\varepsilon^2}{2\int_a^b \rho(\lambda)\mathrm{d}\lambda} \int_a^b \mathrm{d}\lambda \int_{[-2;2]\setminus[a-\delta;b+\delta]} \mathrm{d}\lambda' \, \frac{\rho(\lambda)\rho(\lambda')}{(\lambda - \lambda')^2}, \tag{7.15}$$

where ρ is the Wigner semicircle density (7.13).

Formula (7.15) is a priori only rigorously valid in the perturbative regime where $\varepsilon \ll N^{-1/2}$. We argue that in fact it remains valid in a wider regime where $\varepsilon \ll 1$. Indeed although perturbation theory for the eigenvectors fails for $\mathbf{H}_0$ eigenvalues which are at distance of order of the mean level spacing of $\mathbf{H}_0$, it remains valid in the limit $\varepsilon \ll 1$ for eigenvalues at distance large compared to the order of the perturbation entries $\varepsilon N^{-1/2}$ and in particular for two eigenvalues lying respectively in the two well separated intervals $[a; b]$ and $[a-\delta; b+\delta]$ for which this distance is larger than δ (which indeed is $\gg \varepsilon N^{-1/2}$). We see that every terms appearing in (7.2) corresponding to overlap between eigenvectors associated to eigenvalues which are at distance smaller than δ disappear in formulas (7.9) (and also in (7.7), (7.8)). Therefore, we expect (7.15), as well as our results below, to remain valid in the regime $N^{-1/2} \ll \varepsilon \ll 1$, provided the computed distance $\overline{D}(V_0, V_1)$ itself remains much smaller[3] than unity. We checked formula (7.15) using numerical simulations, with very good agreement for different values of $a, b, \delta, N, \varepsilon$. In those numerical tests we chose the parameters N, ε, δ so as to approach the regime $N^{-1/2} \ll \varepsilon \ll 1$ (for example, $N = 4000, \varepsilon = 0.1, \delta = 0.5$).

We will now write $\overline{D}(a, b; \delta, \varepsilon)$ instead of $\overline{D}(V_0, V_1)$.

It is interesting to study the above expression in the double limit $\delta \to 0$ and $\Delta = b - a \to 0$. One finds:

$$\frac{1}{\varepsilon^2}\overline{D}(a, a + \Delta; \delta, \varepsilon) \approx \begin{cases} \frac{\rho(a)\ln(\Delta/\delta)}{\Delta} & \text{if } \delta \ll \Delta \ll 1, \\ \frac{\rho(a)}{\delta} & \text{if } \Delta \ll \delta \ll 1. \end{cases} \tag{7.16}$$

In the second case where $\Delta \ll \delta \ll 1$, this last expression shows that when the width Δ of interval $[a, b]$ tends to zero, the corresponding eigenvectors are scattered in a region of width δ much larger than Δ itself as soon as $\varepsilon \gg \sqrt{\delta}$. In the first case, it shows that for fixed Δ, the distance $\overline{D}$ diverges logarithmically when $\delta \to 0$. This is a consequence of the small spacings between the (non-perturbed) eigenvalues close to the boundaries of the two intervals $[a; b]$ and $[a - \delta; b + \delta]$. When $\delta > 0$, these spacings remains larger than the fixed distance $\delta > 0$ and D remains finite.

When $\delta = 0$, we can do a more precise analysis of the right hand side of (7.14). One can show, for large N, that the following result holds at least in the regime $\varepsilon \ll N^{-1/2}$:

$$\frac{1}{\varepsilon^2}\overline{D}(a, b; \delta = 0, \varepsilon) \approx \ln N \, \frac{\rho(a)^2 + \rho(b)^2}{2\int_a^b \rho(\lambda)d\lambda} + A(a, b) \tag{7.17}$$

where $A(a, b)$ is a constant independent of N which can be explicitly computed, and involves the well known two-point function $g(r)$ which describes the level-level correlations in the GOE

[3] For this condition to be valid, δ has to be fixed independent of ε, or at least such that $\varepsilon^2 |\ln(\delta)| \ll 1$.

(see Appendix A for the details of this computation). The $\ln N$ term can be guessed from the logarithmic behavior of D when $\delta \to 0$, since one indeed expects the divergence to be cut off when δ becomes of the order of the level spacing, i.e. $\delta \sim (N\rho)^{-1}$. Eq. (7.17) is precisely the Anderson orthogonality catastrophe as first introduced in [14] in the case of finite rank perturbation matrices. We recover here exactly the result of [147] (see their Eq. (31)) by taking $a = -2, b = 0$ in our Eq. (7.17). [4]

As a side remark, we note that Eq. (7.16) predicts that when $\delta \gg \Delta$, a fraction ε^2/δ of the original eigenspace gets shoved away at distances larger than δ (in eigenvalue space). In the context of the non adiabatic evolution of a quantum system [30], this implies that the energy of the system makes jump with a power-law distribution of sizes that decays as δ^{-2}, since by the above argument the cumulative distribution decays as δ^{-1}. This means that under an extreme non-adiabatic process, the energy is not diffusive but rather performs a "Cauchy flight" (i.e. a Lévy flight with a tail exponent equal to 2), see [154].

7.3.2 Full distribution of the singular values of the overlap matrix

To order ε^2, the distance D computed in the previous subsection is proportional to the mean position of the singular values. One can actually be much more precise and compute, for $N \to \infty$, the *full distribution of all singular values*, giving an indication of their scatter around the mean position $\langle s \rangle$. The computation of the density of states (DOS) can be straightforwardly performed using free random matrices techniques.

We have already seen in Eq. (7.11) that the entries of the matrix $\mathbf{\Sigma}$, defined in (7.10), write in the perturbative regime $\varepsilon\sqrt{N} \ll 1$ as: [5]

$$\Sigma_{ij} = \sum_{\ell \notin \{k-q;\dots;k+q\}} \frac{P_{\ell i} P_{\ell j}}{(\lambda_i - \lambda_\ell)(\lambda_j - \lambda_\ell)} \tag{7.18}$$

Denote, for each $\ell \notin \{k-q;\dots;k+q\}$ by $\mathbf{v}_\ell$ the random Gaussian vectors of $\mathbb{R}^P$

$$\mathbf{v}_\ell = \left(\frac{P_{\ell,k-p}}{\lambda_{k-p} - \lambda_\ell}, \frac{P_{\ell,k-p+1}}{\lambda_{k-p+1} - \lambda_\ell}, \dots, \frac{P_{\ell,k+p}}{\lambda_{k+p} - \lambda_\ell} \right)^\dagger .$$

It is easily seen that in fact (changing to the equivalent notation for the summation on ℓ in term of a, b)

$$\Sigma = \sum_{\ell : \lambda_\ell \notin [a-\delta; b+\delta]} \mathbf{v}_\ell \mathbf{v}_\ell^\dagger$$

This matrix $\mathbf{v}_\ell \mathbf{v}_\ell^\dagger$ is clearly the matrix of a projector on $\mathbf{v}_\ell$ and has only one non-zero eigenvalue which is equal to

$$\sigma(\lambda_\ell) = ||\mathbf{v}_\ell||_2^2 = \sum_{j \in \{k-p;\dots;k+p\}} \left(\frac{P_{\ell j}}{\lambda_j - \lambda_\ell} \right)^2$$

which can be approximated in the limit of large matrices ($P \to \infty$) by

$$\sigma(\lambda_\ell) \to \int_a^b d\lambda \frac{\rho(\lambda)}{(\lambda - \lambda_\ell)^2}.$$

The resolvent $Z^\ell(z) \equiv \frac{1}{P}\mathrm{tr}((z - \mathbf{v}_\ell \mathbf{v}_\ell^\dagger)^{-1})$ of the matrix $\mathbf{v}_\ell \mathbf{v}_\ell^\dagger$ is equal to: [6]

$$Z^\ell(z) = \frac{1}{P} \left(\frac{1}{z - \sigma(\lambda_\ell)} + \frac{P-1}{z} \right) .$$

[4] The authors of [147] expect deviations in (7.17) when the parameter $x := \varepsilon\sqrt{N}$ (which has to be $\ll 1$ for (7.17) to be fully valid) is increased. However their numerical results (presented in Fig. 2 of [147]) show that the discrepancies are only noticeable for x close to 1. In addition, the authors of [147] explain that the failure of (7.17) for not small enough x is due to the first-order perturbation theory estimate that breaks down when used for levels in the vicinity of the edges a, b of the initial interval. This problem was avoided previously by the use of rectangular matrices with $Q > P$ and the introduction of the δ margin at the edges a and b.

[5] We skip the subscript 0 on the eigenvalues λ_is.

[6] Resolvents are usually denoted by the letter G, but we do not want to confuse the reader with the overlap matrix G of which we compute the singular value spectrum.

The Blue function, which by definition is the functional inverse of the resolvent $B^\ell(Z^\ell(z)) = z$, can be computed to first order in $1/P$:

$$B^\ell(z) = \frac{1}{z} + \frac{1}{P}\frac{\sigma(\lambda_\ell)}{1-\sigma(\lambda_\ell)z}$$

Finally, the Red function, defined as $R^\ell(z) \equiv B^\ell(z) - \frac{1}{z}$, is given by:

$$R^\ell(z) = \frac{1}{P}\frac{\sigma(\lambda_\ell)}{1-\sigma(\lambda_\ell)z}$$

The trick, coming from the theory of free matrices, is to use the additive property of the Red function (also called R-transform) for the asymptotically free matrices $\mathbf{v}_\ell\mathbf{v}_\ell^\dagger$. Essentially, the R-transform of the matrix $\mathbf{\Sigma}$ can be computed as the sum of the R-transforms of the matrices $\mathbf{v}_\ell\mathbf{v}_\ell^\dagger$:

$$R(z) = \sum_{\ell\notin\{k-q;\dots;k+q\}} R^\ell(z) = \frac{1}{P}\sum_{\ell\notin\{k-q;\dots;k+q\}} \frac{\sigma(\lambda_\ell)}{1-\sigma(\lambda_\ell)z}$$

Finally, the Blue function of $\mathbf{\Sigma}$ is:

$$B(z) = \frac{1}{z} + \frac{1}{P}\sum_{\ell\notin\{k-q;\dots;k+q\}} \frac{\sigma(\lambda_\ell)}{1-\sigma(\lambda_\ell)z}$$

which can be approximated in the limit of large P as:

$$B(z) = \frac{1}{z} + \frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]} d\lambda \frac{\rho(\lambda)\sigma(\lambda)}{1-\sigma(\lambda)z}. \tag{7.19}$$

where we note here and below $N_a^b := \int_a^b \rho(\lambda)d\lambda$. Rewriting equation (7.19) in terms of the resolvent gives our central result:

$$\boxed{z = \frac{1}{Z(z)} + \frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]} d\lambda \frac{\rho(\lambda)\sigma(\lambda)}{1-\sigma(\lambda)Z(z)}\,.} \tag{7.20}$$

Equation (7.20) characterizes the density of states of the matrix $\mathbf{\Sigma}$ in the limit of large dimension. We ran numerical simulations to test the validity of Eq. (7.20) in the regime $1/\sqrt{N} \ll \varepsilon \ll 1$, see Fig. 7.1. The agreement is excellent. It would be interesting to run this numerical test for very large values of N (here we took $N = 4000$) so as to fully reach the regime $1/\sqrt{N} \ll \varepsilon \ll 1$. However, this becomes numerically demanding, and we leave this study for future work.

We now want to extract the qualitative informations about the distribution of all singular values of the matrix $\mathbf{G}$ from this equation. In particular, we will show in the next subsection that the density of singular values has a compact support for which we characterize the left and right edges. We also study the shape of this distribution in the two asymptotic regimes $\Delta \ll \delta$ and $\delta \ll \Delta \ll 1$.

7.3.3 Qualitative properties of the spectrum of $\mathbf{\Sigma}$

Right and Left edges

The relation between the resolvent Z and the density of states $r(s)$ of the matrix $\mathbf{\Sigma}$ is $\lim_{\omega\to 0}\Im Z(s-i\omega) = \pi r(s), s\in\mathbb{R}$. Note that one should not confuse the density of states $\rho(\lambda)$ of the original matrix $\mathbf{H}_0$ with the density of eigenvalues $r(s)$ of $\mathbf{\Sigma}$.

From equation (7.20), we can derive a system of equations for the real ($g(s)$) and imaginary ($r(s)$) parts of $Z(s)$, for $\omega\to 0$:

$$s = \frac{g(s)}{g(s)^2+\pi^2 r(s)^2} + \frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)(1-\sigma(x)g(s))}{(1-\sigma(x)g(s))^2+\sigma(x)^2\pi^2 r(s)^2}\,, \tag{7.21}$$

$$0 = r(s)\left(\frac{-1}{g(s)^2+\pi^2 r(s)^2} + \frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)^2}{(1-\sigma(x)g(s))^2+\sigma(x)^2\pi^2 r(s)^2}\right). \tag{7.22}$$

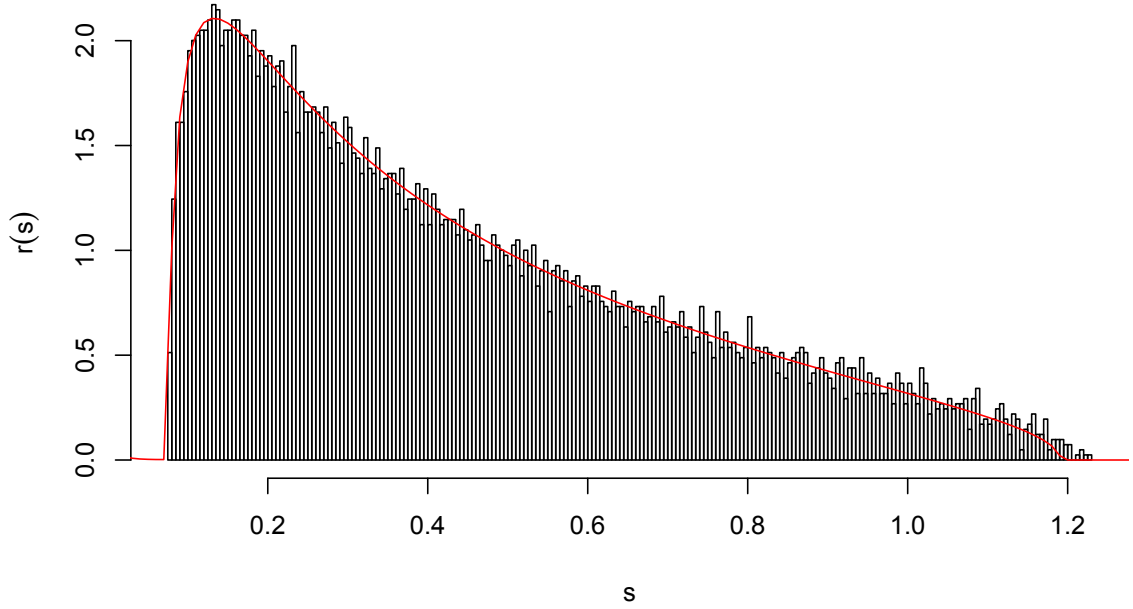


Figure 7.1: The histogram represents a numerical simulation of the density of states of the matrix $\mathbf{\Sigma}$ (computed with 15 independent samples). The red curve is the theoretical corresponding density for $r(s)$ obtained by solving numerically (7.20). For this figure, we chose $a = 0, b = 0.5, \delta = 0.5$. We chose the parameters N and ε so as to approach the "less perturbative" regime where $1/\sqrt{N} \ll \varepsilon \ll 1$ for this figure as $N = 4000$ and $\varepsilon = 0.1$.

The second equation (7.22) always admits the solution $r = 0$. Plugging $r = 0$ into the first equation gives:

$$s = \frac{1}{g(s)} + \frac{1}{N_a^b} \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)}{1-\sigma(x)g(s)} \tag{7.23}$$

Equation (7.23) implies the asymptotic relation $g(s) \sim_{s\to\infty} 1/s$ and therefore large positive values of s correspond to small values of $g(s)$. Set

$$m_0 \equiv \max_{x\in[-2;2]\setminus[a-\delta;b+\delta]} \sigma(x), \tag{7.24}$$

the Right Hand Side (RHS) of the above equation is well defined provided $g(s) \in (0; 1/m_0)$. However, when $g(s) \to 0^+$ or when $g(s) \to (1/m_0)^-$, the RHS tends to $+\infty$. Thus, on the interval $g(s) \in (0; 1/m_0)$, the RHS must reach a minimum which corresponds to the right edge of the density of states. The point $\bar{g} \in (0; 1/m_0)$ for which this minimum is reached verifies:

$$-\frac{1}{\bar{g}^2} + \frac{1}{N_a^b} \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)^2}{(1-\sigma(x)\bar{g})^2} = 0\,, \tag{7.25}$$

and we can compute the right edge of the spectrum s_{max} from:

$$s_{max} = \frac{1}{\bar{g}} + \frac{1}{N_a^b} \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)}{1-\sigma(x)\bar{g}}. \tag{7.26}$$

We can now turn to the left edge of the spectrum. Equation (7.23) implies also the asymptotic relation $g(s) \to -\infty$ when $s \to 0$ and therefore small positive values of s correspond to large negative values of $g(s)$. The RHS of equation (7.23) is well defined for negative values of $g(s)$; it goes to 0^- for very large and negative values of $g(s)$, and goes to $-\infty$ for $g(s) = 0^-$, so it has a positive maximum somewhere in between. The value of this maximum corresponds to the left

edge of the density of states and can be computed numerically like for the right edge. The point $\widetilde{g} \in (-\infty; 0)$ for which this maximum is reached verifies the same equation as $\bar{g}$ above, and the left edge s_{min} is now given by:

$$s_{min} = \frac{1}{\widetilde{g}} + \frac{1}{N_a^b} \int\limits_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)\sigma(x)}{1-\sigma(x)\widetilde{g}}. \tag{7.27}$$

Small fluctuations regime $\Delta \ll \delta$

We first consider the case where $\Delta \equiv b - a \ll \delta$, corresponding to $P \ll Q$, in particular the dimension of the perturbed subspace is much larger than the dimension of the unperturbed space and so the perturbed space almost surely spans the unperturbed subspace. We therefore expect small fluctuations in this regime. Equation (7.20) can be solved explicitly in this case. It is in fact possible to perform an asymptotic expansion in $\sigma(x)$, which is very small compared to 1 for all $x \in [-2; 2] \setminus [a - \delta; b + \delta]$ and then to solve equation (7.20).

More precisely, in this regime, we have for all $x \in [-2; 2] \setminus [a - \delta; b + \delta]$:

$$\sigma(x) \approx \frac{\rho(a)}{(x-a)^2}\Delta\,.$$

We plug this approximation in equation (7.20) to obtain

$$\begin{aligned} z &= \frac{1}{Z(z)} + \frac{\Delta \times \rho(a)}{N_a^b} \int\limits_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)}{(x-a)^2 - \Delta \times \rho(a) Z(z)} \\ &\approx \frac{1}{Z(z)} + \int\limits_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)}{(x-a)^2} + \Delta \times \rho(a) Z(z) \int\limits_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)}{(x-a)^4}\,. \end{aligned}$$

Now setting $A \equiv \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)}{(x-a)^2}$ and $B \equiv \Delta \times \rho(a) \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx \frac{\rho(x)}{(x-a)^4}$, we see that $Z(z)$ is solution of the polynomial equation of degree two:

$$BZ(z)^2 + (A - z)Z(z) + 1 = 0. \tag{7.28}$$

For $z = s \in \mathbb{R}$, this equation has solutions with non-zero imaginary part only if $s \in [A - 2\sqrt{B}; A + 2\sqrt{B}]$, which are given by

$$Z(z) = \frac{-A + s \pm i\sqrt{4B - (A - s^2)}}{2B}.$$

Using the relation $\lim_{\omega\to 0} \Im Z(s - i\omega) = \pi r(s)$ for $s \in \mathbb{R}$, we find that $r(s)$ in this regime is given by the semi-circle law

$$r(s) = \frac{1}{2B\pi}\sqrt{4B - (A - s)^2}, \quad A - 2\sqrt{B} < s < A + 2\sqrt{B}. \tag{7.29}$$

This result is consistent with (7.15) since, in this regime, $D(a, b; \delta) = \varepsilon^2 A$.

Note that in the particular regime $\Delta \ll \delta \ll 1$, the quantity B is proportional to Δ/δ^3 and is therefore much smaller than $A^2 \propto 1/\delta^2$, meaning that $r(s)$ becomes concentrated around $s = A$, with fluctuations of order $\sqrt{\Delta/\delta^3}$. This result is also consistent with the direct calculation of the root-mean squared fluctuations of s, as obtained in Appendix B, see equation (7.61).

Strong fluctuations regime $\delta \ll \Delta \ll 1$

To simplify notations, we will suppose in the following that a and b are such that $\rho(a) \geqslant \rho(b)$. Let us first consider the right edge s_{max} as given by (7.26). We need to find an asymptotic expansion in this regime of the $\bar{g} \in (0; 1/m_0)$ which verifies (7.25). So we start by defining $\alpha := \bar{g}m_0 \in (0; 1)$ and investigate its behavior when $\delta \ll \Delta \ll 1$. Since $m_0 \sim_{\delta\to 0} \rho(a)/\delta$, equation (7.25) now rewrites as

$$\frac{\alpha^2\delta^2}{\rho(a)^2 N_a^b} \int\limits_{[-2;2]\setminus[a-\delta;b+\delta]} dx\, \frac{\rho(x)\sigma(x)^2}{(1 - \alpha\frac{\sigma(x)}{m_0})^2} \sim 1\,. \tag{7.30}$$

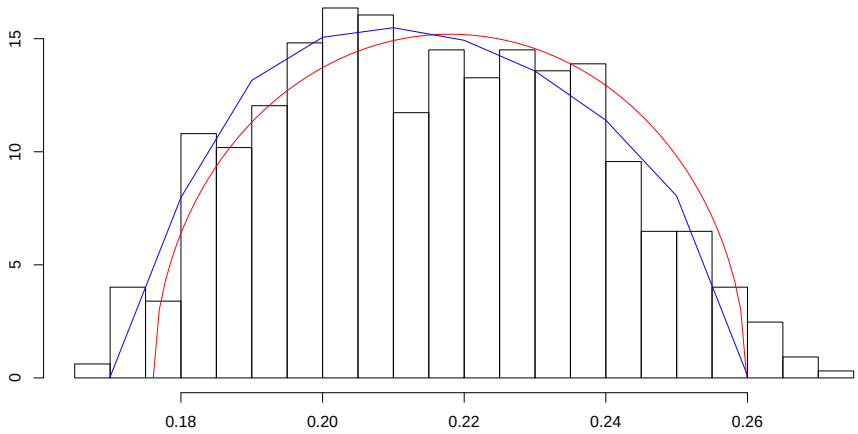


Figure 7.2: The histogram represents a numerical simulation of the density of states of the matrix $\mathbf{\Sigma}$ (computed with 100 independent samples). The red curve is the theoretical corresponding density for $r(s)$ in small fluctuations regime given by (7.29). The blue curve represents also the theoretical density $r(s)$ but computed numerically by solving directly the system (7.21) and (7.22). For this figure, we chose $a = 0, b = 0.01, \Delta = 0.01, \delta = 1$.

In the limit $\Delta \ll 1$, it is easy to see that the function σ can be written for $x < a$ as

$$\sigma(x) = \frac{\rho(a)}{a-x} f\left(\frac{a-x}{\Delta}\right), \tag{7.31}$$

where the function f verifies $f(u) \sim_{u\to 0} 1$ and $f(u) \sim_{u\to\infty} 1/u$. Using (7.31), we can write

$$\begin{aligned}\int_{-2}^{a-\delta} dx\, \frac{\rho(x)\sigma(x)^2}{(1-\alpha\sigma(x)\frac{\delta}{\rho(a)})^2} &= \rho(a)^2 \int_{-2}^{a-\delta} dx\, \frac{\rho(x) f^2(\frac{a-x}{\Delta})}{(a-x-\alpha f(\frac{a-x}{\Delta})\delta)^2} \\ &= \frac{\rho(a)^2}{\Delta} \int_{\frac{\delta}{\Delta}}^{\frac{a+2}{\Delta}} du\, \frac{\rho(a-u\Delta) f^2(u)}{(u-\alpha\frac{\delta}{\Delta} f(u))^2},\end{aligned}$$

where we did the change of variables $u = (a-x)/\Delta$ for the last line. In the limit $\delta \ll \Delta \ll 1$, this last integral is dominated by the region where u is small and $f(u) \sim 1$. We thus have

$$\begin{aligned}\int_{-2}^{a-\delta} dx\, \frac{\rho(x)\sigma(x)^2}{(1-\alpha\sigma(x)\frac{\delta}{\rho(a)})^2} &\sim \frac{\rho(a)^3}{\Delta} \int_0^{+\infty} \frac{du}{(u-\alpha\frac{\delta}{\Delta})^2} \\ &\sim \frac{\rho(a)^3}{\delta} \frac{1}{1-\alpha}.\end{aligned}$$

Then, using (7.30) and with the same argument now for $x > b$, we get

$$\alpha = 1 - \frac{2\delta}{\Delta}.$$

The corresponding $\bar{g}$ is $\bar{g} = \delta/\rho(a)(1-2\delta/\Delta)$ and plugging this value of $\bar{g}$ in (7.26) gives

$$s_{max} \sim \frac{\rho(a)}{\delta} + \frac{1}{N_a^b} \int_{[-2;2]\setminus[a-\delta;b+\delta]} dx\, \frac{\rho(x)\sigma(x)}{1-\sigma(x)\bar{g}}.$$

But, it is plain to check that the second term is of order at most $\ln(\delta/\Delta)/\Delta \ll 1/\delta$ in the limit $\delta \ll \Delta \ll 1$.

The determination of s_{min} proceeds similarly, and the calculations are detailed in Appendix C. The final result is that

$$s_{min} = \frac{c\rho(a)}{\Delta},$$

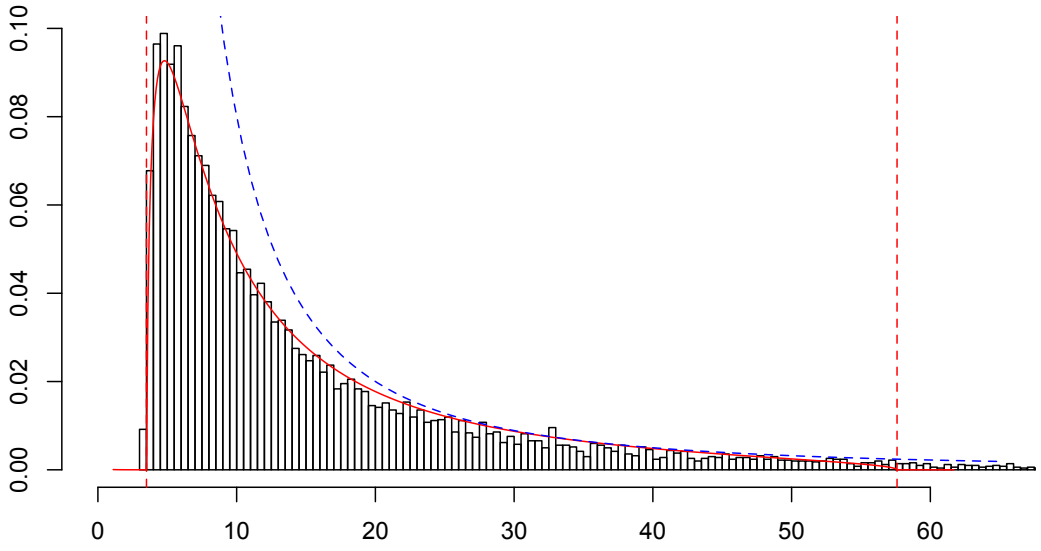


Figure 7.3: The histogram represents a numerical simulation of the density of states of the matrix $\mathbf{\Sigma}$ (computed with 20 independent samples). The red curve is the theoretical corresponding density for $r(s)$, it is computed numerically by solving the system (7.21) and (7.22). The red dotted vertical lines show the left and right edges of the density $r(s)$. The blue dotted curve is the graph of the function $8/x^2$. For this figure, we chose $a = 0, b = 0.1, \Delta = 0.1, \delta = 0.01$.

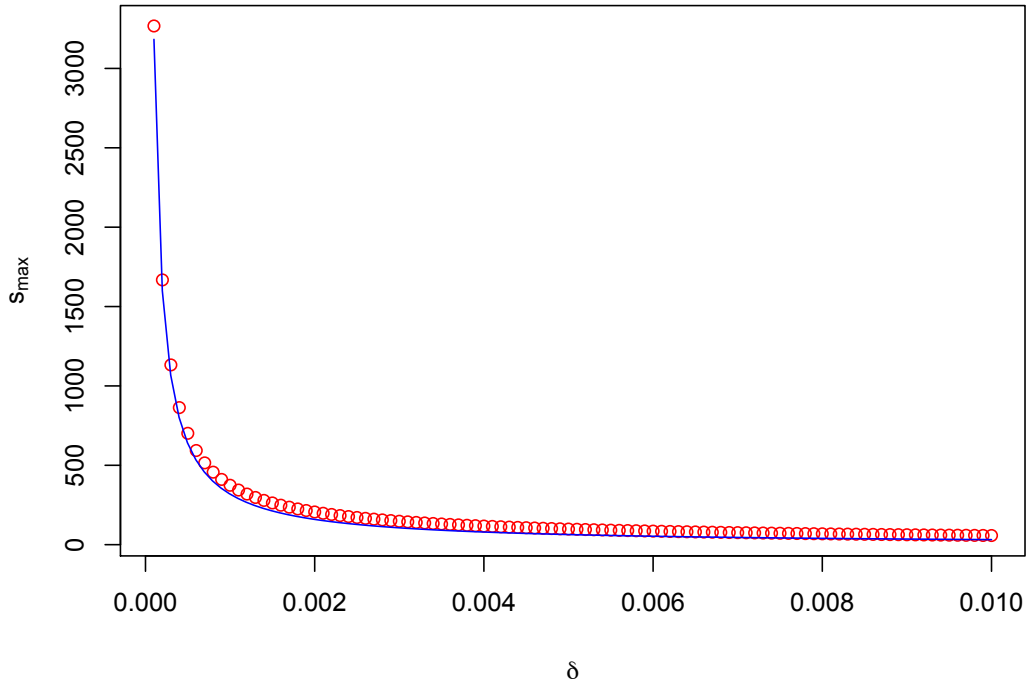


Figure 7.4: The points represent the function $s_{max}(\delta)$ as a function of δ. They are computed numerically through equations (7.25) and (7.26). The blue dotted line is the function $\delta \to \rho(a)/\delta$. In this figure, we chose $a = 0, b = 0.1, \Delta = 0.1$.

where $c > 0$ is a number of order unity which can be determined if needed.

To summarize, in this regime, the minimum and maximum eigenvalues s_{min} and s_{max} of the random matrix Σ are asymptotically given by:

$$s_{min} \sim \frac{c\rho(a)}{\Delta} \ll s_{max} \sim \frac{\rho(a)}{\delta}.$$

We verified the result for $s_{max}(\delta)$ with numerical simulations (see Fig. 7.4).

Together with the exact result on the average value of $r(s)$ in this regime (given by $D(a, a+\Delta; \delta)/\varepsilon^2$) and its variance computed in Appendix B, we conjecture that the asymptotic behaviour of $r(s)$ in the region $s_{min} \ll s \ll s_{max}$ is given by:

$$r(s) \propto \frac{s_{min}}{s^2}. \tag{7.32}$$

Since the integral of $sr(s)$ is logarithmically divergent (but cut-off at s_{min} and s_{max}), it is easy to see that this form reproduces exactly the logarithmic behavior of $D(a, a+\Delta; \delta)$ in this regime, see equation (7.16). On the other hand, the integral of $s^2 r(s)$ is dominated by its upper bound, leading to a variance of the spectrum given by $s_{min} \times s_{max}$, in agreement with the exact result obtained in Appendix B, see equation 7.60. Therefore, in this regime, the situation is particularly interesting: while most eigenvalues are close to s_{min}, there is a slow power-law tail in $r(s)$ which makes the average of s logarithmically divergent when $\delta \to 0$. This is why we call this a strong fluctuation regime: the 'overlap' distance D between the initial and the target spaces is large because a relatively small number of directions are completely lost.

7.4 Eigenvector stability for covariance matrices

In the next subsection, we do essentially the same study as in the previous section 7.3 for another random matrix model. There is a small difference with the previous study as the spectrum of the matrices introduced below has some isolated eigenvalues as well as a continuous part. The study of the eigenvectors associated to eigenvalues in the continuous part (that we call the "sea") is very similar to the previous study. The isolated eigenvectors have to be treated separately (see the paragraph *Isolated eigenvectors*).

In the next subsection 7.4.2, we compute the overlap distance D between two eigenspaces generated by the top isolated eigenvectors of the empirical covariance matrix and the true covariance matrix. The formulas that we obtain are (7.38) and (7.39) and will be used later in the section 7.6 on application to financial data.

We end this subsection 7.4.2 by doing a different analysis of the stability of eigenspaces through the spectral projectors. The formulas that we will use later for applications are about the spectrum of the mean spectral projectors (7.42) and (7.43).

7.4.1 Eigenvectors of Spiked matrices

In this subsection, we will assume that $(\mathbf{C}_N)$ is a sequence of positive definite matrices. We will denote by $\lambda_1^N, \ldots, \lambda_N^N$ the eigenvalues of $(\mathbf{C}_N)$ in decreasing order and we will suppose that

- there exists a fixed number $k < N$, $q \in (0; 1)$ and $(\lambda_1 > \cdots > \lambda_k > (1+\sqrt{q})^2)$ such that

$$(\lambda_1^N, \ldots, \lambda_k^N) \to_{N\to\infty} (\lambda_1, \ldots, \lambda_k).$$

- the empirical measure $\mu_N \equiv \frac{1}{N}\sum_{i=k+1}^{N} \delta_{\lambda_i}$ converges in the limit of large N, T with $N/T = q$ to the Marchenko-Pastur distribution whose density with respect to Lebesgue measure is given by

$$\rho(x) \equiv \frac{1}{2\pi q x}\sqrt{(\gamma_+ - x)(x - \gamma_-)}, \quad a < x < b,$$

where $\gamma_- \equiv (1-\sqrt{q})^2$ and $\gamma_+ \equiv (1+\sqrt{q})^2$.

For each N, $\mathbf{C}_N$ is the true covariance matrix (also called "population covariance matrix"). This particular choice for the shape of the matrices $\mathbf{C}_N$ is rather natural in view of applications. For example, in financial market, the correlation (or covariance) matrix has k isolated eigenvalues well separated from the other eigenvalues which form the noisy part of the spectrum (*Marchenko-Pastur sea* or the *bulk*).

We now consider the associated empirical covariance matrix $\mathbf{E}_N$ defined as:

$$\mathbf{E}_{N,ij} \equiv \frac{1}{T}\sum_{t=1}^{T} r_i^t r_j^t$$

where the $(r_1^t, \ldots, r_N^t), 1 \leqslant t \leqslant T$ are i.i.d. Gaussian vectors of covariance $\mathbf{C}_N$.

The question we ask in this subsection is: how close are the eigenvectors of $\mathbf{E}_N$ to those of the matrix $\mathbf{C}_N$? In the following two paragraphs, we treat the two cases of the eigenvectors associated to eigenvalues in the Marchenko-Pastur sea and of those associated to the isolated eigenvalues $\lambda_1, \ldots, \lambda_k$.

This question falls under the scope of section 7.2 since the matrix $\mathbf{E}_N$ can be written as a perturbation of the matrix $\mathbf{C}_N$. Indeed we have:

$$\mathbf{E}_N = \mathbf{C}_N + \mathcal{E}_N, \quad \text{with} \quad \mathcal{E}_{N,ij} = \frac{1}{T}\sum_{t=1}^{T} r_i^t r_j^t - C_{N,ij}. \tag{7.33}$$

and the matrix elements of $\mathcal{E}$ are (because of the Central Limit Theorem) of order $1/\sqrt{T}$ which is much smaller than 1 as T is large. However, this problem is of different nature than the one treated in section 7.3 because of the non-trivial dependance structure for the matrix elements of the perturbation matrix $\mathcal{E}$. It is given by

$$\overline{\mathcal{E}_{N,ij}\mathcal{E}_{N,k\ell}} = (C_{N,ik}C_{N,j\ell} + C_{N,i\ell}C_{N,jk})/T. \tag{7.34}$$

In the whole current section, $\overline{\cdots}$ denotes an averaging over the r_i^t.

Eigenvectors in the Marchenko-Pastur sea

The results of subsections 7.3.1, 7.3.2 and 7.3.3 can be extended to this context. We consider the subspace of eigenvectors of $\mathbf{C}_N$ corresponding to all the eigenvalues λ contained in a certain finite interval $[a, b]$ included in the Marchenko-Pastur sea $[\gamma_-, \gamma_+]$. We want to compute the distance D between this subspace and the subspace spanned by the perturbed eigenvectors of $\mathbf{E}_N$ corresponding to all eigenvalues of $\mathbf{E}_N$ contained in $[a-\delta, b+\delta]$, where δ is a positive parameter. Using formula (7.9) as before, we find that in the limit of large N, T (with $N/T = q$), as soon as $\delta > 0$, the mean overlap distance $\overline{D}$ is given (using (7.34) for the averaging) by:

$$\overline{D}(a, b; \delta) \sim \frac{1}{2TN_a^b}\int_a^b d\lambda \int_{[-2;2]\setminus[a-\delta;b+\delta]} d\lambda' \, \frac{\lambda\lambda'\rho(\lambda)\rho(\lambda')}{(\lambda-\lambda')^2}, \tag{7.35}$$

where $\rho(\lambda)$ is now the Marchenko-Pastur distribution of parameter $q = N/T$. Obviously, when T is infinite, $D = 0$ since $\mathbf{E}_N = \mathbf{C}_N$.

For the singular value density of states $r(s)$, the resolvent of the matrix $\mathbf{\Sigma}$ defined as $\mathbf{\Sigma} \equiv T(\mathbf{I} - \mathbf{G}\mathbf{G}^\dagger)$ now verifies:

$$z = \frac{1}{Z(z)} + \frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]} d\lambda \frac{\rho(\lambda)\nu(\lambda)}{1-\nu(\lambda)Z(z)} \tag{7.36}$$

where ν is defined as $\nu(\lambda) \equiv \lambda\int_a^b dx \frac{\rho(x)}{(x-\lambda)^2}$. As before, it is easy to show that the density of states of $\mathbf{\Sigma}$ is compactly supported and to find numerical evaluations of the left and right edges. One can also study the limit shape of the density of states in the two regimes $\Delta \ll \delta$ and $\delta \ll \Delta \ll 1$, with results very similar to the GOE ones above.

The matrix $\mathbf{G}\mathbf{G}^\dagger$ in this case gives a precise information on the relationship between the eigenvectors of the *population covariance matrix* (or true covariance matrix) $\mathbf{C_N}$ and the eigenvectors of the sample covariance matrix $\mathbf{E_N}$. Previous works along these lines can be found in [99, 20].

Isolated eigenvectors

In this paragraph, we now consider the case of eigenvectors associated to isolated eigenvalues $\lambda_1, \ldots, \lambda_k$. We denote by $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_k\rangle$ the corresponding eigenvectors of $\mathbf{C}_N$ and by $|\phi'_1\rangle, |\phi'_2\rangle, \ldots, |\phi'_k\rangle$ [7] the corresponding eigenvectors of $\mathbf{E}_N$.

To understand precisely how the $|\phi'_i\rangle$ decompose in the basis of the $|\phi_j\rangle$ in the limit of large N, we want to compute the limit of the average local density of states for each state $|\phi'_i\rangle$ $(1 \leqslant i \leqslant k)$, that is the probability measure

$$\nu_N^{(i)}(\lambda) \equiv \frac{1}{N} \sum_{j=1}^{N} \overline{\langle \phi'_i | \phi_j \rangle^2}\, \delta(\lambda - \lambda_j)$$

where $\overline{\cdots}$ denotes an average over $\mathbf{E}_N$. This expresses the way $|\phi'_i\rangle$ is scattered over the unperturbed eigenvectors.

Perturbation theory again allows to compute the quantities $\langle \phi'_i | \phi_j \rangle^2$ for $i \neq j$:

$$\overline{\langle \phi'_i | \phi_j \rangle^2} = \frac{\overline{\langle \phi_i | \mathcal{E} | \phi_j \rangle^2}}{(\lambda_i^N - \lambda_j^N)^2} = \frac{1}{T} \frac{\lambda_i^N \lambda_j^N}{(\lambda_i^N - \lambda_j^N)^2} \sim_{N\to\infty} \frac{1}{T} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2},$$

and for $i = j$,

$$\overline{\langle \phi'_i | \phi_i \rangle^2} = 1 - \sum_{k \neq i} \frac{\overline{\langle \phi_i | \mathcal{E} | \phi_k \rangle^2}}{(\lambda_i^N - \lambda_k^N)^2} \sim_{N\to\infty} 1 - \frac{1}{T} \sum_{k \neq i} \frac{\lambda_i \lambda_k}{(\lambda_i - \lambda_k)^2}.$$

Note that the random variables $\langle \phi_i | \mathcal{E} | \phi_j \rangle, i \neq j$ are uncorrelated. Thus, the local density of states $\nu_N^{(i)}$ has k atoms and (for large N, T with $N/T = q$) admits a continuous density in the Marchenko-Pastur sea. The atom are localized on the $\lambda_j, j = 1, \ldots, k$ and have weights $\frac{1}{T} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2}$ for $j \neq i$. The continuous density in the Marchenko-Pastur sea $[\gamma_-, \gamma_+]$ is given by:

$$\frac{1}{T} \frac{\lambda_i \lambda}{(\lambda_i - \lambda)^2} \rho(\lambda) d\lambda, \quad \gamma_- < \lambda < \gamma_+. \tag{7.37}$$

This asymptotic for the probability measure $\nu_N^{(i)}$ has been verified with numerical simulations.

7.4.2 Stability of eigenspaces

We now want to characterize the stability of the subspace spanned by the eigenvectors associated to the (largest) isolated eigenvalues. The theory we develop here provides a precise estimate of the amount of eigenspace instability induced by measurement noise. This sets a benchmark that will allow us to detect any extra dynamics of the eigenvectors of the correlation matrix of stock returns in financial markets not explained by measurement noise and therefore attributable to a genuine evolution of the market (see section 7.6).

As shown by Eq. (7.33) above, the sample covariance matrix $\mathbf{E}$[8] is a perturbed version of $\mathbf{C}$. Using again the framework of section 7.2, one can calculate the distance (or overlap) between the top P eigenvectors of the true correlation matrix $\mathbf{C}$ and the top Q[9] eigenvectors of the empirical correlation matrix $\mathbf{E}$.

Provided T is large enough for the above perturbation theory to be valid, and upon averaging over the measurement noise, one gets the following expression for the overlap distance D:

$$\overline{D}(P, Q) = \frac{1}{2TP} \sum_{i=1}^{P} \sum_{j=Q+1}^{N} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2}, \tag{7.38}$$

where the λ_is are the eigenvalues of $\mathbf{C}$, in decreasing order.

[7] The vectors $|\phi_i\rangle$ and $|\phi'_i\rangle$ depend on N but to simplify notations, we drop the subscript.

[8] As N does not have to be necessarily large in this subsection and in the next section, we drop the subscript N for the matrices $\mathbf{C}$ and $\mathbf{E}$.

[9] We take $Q \geqslant P$ as before in section 7.2 .

Note that one can extend the previous result (7.38) to the case where the vectors $(r_1^t, \dots, r_N^t), t \geqslant 0$ are distributed according to a multivariate Student distribution with ν-degrees of freedom and covariance matrix $\mathbf{C}$. In this case[10], Eq. (7.38) becomes

$$\overline{D}(P,Q) = \left(\frac{\nu-2}{\nu-4}\right) \frac{1}{2TP} \sum_{i=1}^{P} \sum_{j=Q+1}^{N} \frac{\lambda_i \lambda_j}{(\lambda_i - \lambda_j)^2} \,. \tag{7.39}$$

Note that the Gaussian case corresponds to $\nu \to \infty$. For $\nu \to 4^+$, on the other hand, fluctuations become divergent.

In practice for applications (see section 7.6), one does not know the true correlation matrix $\mathbf{C}$ and thus it is in fact not possible to compute empirically the overlap distance between the eigenvectors of $\mathbf{C}$ and the eigenvectors of the empirical correlation matrix $\mathbf{E}$. However, if one is given a time series of empirical correlation matrix $(\mathbf{E}^t)_{t \geqslant 0}$ defined for all t as

$$E_{ij}^t = \frac{1}{T} \sum_{u=1}^{T} r_i^{t+u} r_j^{t+u} \,, \tag{7.40}$$

where, $(r_1^v, \dots, r_N^v), v \geqslant 0$ are independent Gaussian vectors of covariance matrix $\mathbf{C}$, one can similarly define the distance between the eigenspaces of two independent sample covariance matrices $\mathbf{E}^s$ and $\mathbf{E}^t$ (determined on two non overlapping time periods, i.e. such that $|t-s| > T$). In this case, the above formula Eq. (7.38) is simply multiplied by a factor 2.

For the comparison between the eigenvalues of $\mathbf{E}^s$ and $\mathbf{E}^t$, one can show using perturbation theory (see equation (7.3) and also equation (7.34) for the averaging) that the measurement noise is, for T large enough, given by:

$$\overline{(\lambda_i^s - \lambda_i^t)^2}_{|t-s|>T} \approx \frac{4\lambda_i^2}{T}. \tag{7.41}$$

where the λ_i are the eigenvalues of the matrix $\mathbf{C}$ measured empirically using the whole period of time and where $\overline{\cdots}_{|t-s|>T}$ denotes an empirical average over all s, t such that $|t-s| > T$. As before, if the vectors $(r_1^v, \dots, r_N^v), v \geqslant 0$ are distributed according to a multivariate Student distribution with ν-degrees of freedom and covariance matrix $\mathbf{C}$, one finds an extra multiplicative term $(\nu-2)/(\nu-4)$ in (7.41).

Another characterization of the stability of eigenspaces was proposed by Zumbach [157]. The idea here is to study the stability of the *spectral projectors* associated to the top k eigenvalues. The spectral projector of rank k associated to the top k eigenvalues is defined as follows:

$$\chi_k = \sum_{i=1}^{k} |\phi_i\rangle\langle\phi_i| \,,$$

where the $|\phi_i\rangle, i \in \{1, \dots, k\}$ are the eigenvectors of $\mathbf{C}$. As before the *true* spectral projector χ_k is measured through an empirical covariance matrix $\mathbf{E}$ and the resulting spectral projector χ_k' will be affected by measurement noise. The aim is again to compute properties of this spectral projector χ_k, so as to be able to separate the measurement noise effect from a true temporal evolution of the matrix $\mathbf{C}$.

Using perturbation theory in Eq. (7.33), we have:

$$\begin{aligned}
\chi_k' = \sum_{i=1}^{k} |\phi_i'\rangle\langle\phi_i'| &= \sum_{i=1}^{k} \left(1 - \sum_{j\neq i} \frac{\langle\phi_i|\mathcal{E}|\phi_j\rangle^2}{(\lambda_i-\lambda_j)^2}\right) |\phi_i\rangle\langle\phi_i| \\
&+ \sum_{i=1}^{k} \sum_{j\neq i} \frac{\langle\phi_i|\mathcal{E}|\phi_j\rangle}{\lambda_i - \lambda_j} \left(|\phi_i\rangle\langle\phi_j| + |\phi_j\rangle\langle\phi_i|\right) \\
&+ \sum_{i=1}^{k} \sum_{j\neq i} \alpha_{i,j} (|\phi_i\rangle\langle\phi_j| + |\phi_j\rangle\langle\phi_j|) + \sum_{i=1}^{k} \sum_{j\neq i} \sum_{\ell\neq i} \frac{\langle\phi_i|\mathcal{E}|\phi_j\rangle\langle\phi_i|\mathcal{E}|\phi_\ell\rangle}{(\lambda_i-\lambda_j)(\lambda_i-\lambda_\ell)} |\phi_j\rangle\langle\phi_\ell|
\end{aligned}$$

[10] see e.g. Eq. (9.28) p. 154 of [45] that replaces Eq. (7.34) above.

where

$$\alpha_{i,j} = \frac{1}{\lambda_i - \lambda_j}\left(\sum_{\ell\neq i}\frac{\langle\phi_j|\mathcal{E}|\phi_\ell\rangle\langle\phi_\ell|\mathcal{E}|\phi_i\rangle}{\lambda_i-\lambda_\ell} - \frac{\langle\phi_i|\mathcal{E}|\phi_i\rangle\langle\phi_i|\mathcal{E}|\phi_j\rangle}{\lambda_i-\lambda_j}\right).$$

Using again equation (7.34),

$$\overline{\langle\phi_j|\mathcal{E}|\phi_i\rangle\langle\phi_\ell|\mathcal{E}|\phi_i\rangle} = \begin{cases} 0 & \text{if } \ell\neq j,\\ \lambda_j\lambda_i/T & \text{if } j=\ell, j\neq i,\\ 2\lambda_i^2/T & \text{otherwise,}\end{cases}$$

we get:

$$\overline{\chi'_k} = \sum_{i=1}^{k}\left(1-\frac{1}{T}\sum_{j\neq i}\frac{\lambda_i\lambda_j}{(\lambda_i-\lambda_j)^2}\right)|\phi_i\rangle\langle\phi_i| + \frac{1}{T}\sum_{i=1}^{k}\sum_{j\neq i}\frac{\lambda_i\lambda_j}{(\lambda_i-\lambda_j)^2}|\phi_j\rangle\langle\phi_j|\,.$$

We see that the vectors $\phi_i, i\in\{1,\dots,N\}$ are also eigenvectors of $\overline{\chi'_k}$, but with shifted eigenvalues. More precisely, we have, for $i\leqslant k$

$$\overline{\chi'_k}|\phi_i\rangle = \left(1-\frac{1}{T}\sum_{j=k+1}^{N}\frac{\lambda_i\lambda_j}{(\lambda_i-\lambda_j)^2}\right)|\phi_i\rangle\,, \tag{7.42}$$

and, for $i>k$,

$$\overline{\chi'_k}|\phi_i\rangle = \frac{1}{T}\sum_{j=1}^{k}\frac{\lambda_i\lambda_j}{(\lambda_i-\lambda_j)^2}|\phi_i\rangle\,. \tag{7.43}$$

Therefore, in the absence of measurement noise (i.e. for $T\to\infty$), $\overline{\chi'_k}$ has k eigenvalues exactly equal to unity, and $N-k$ eigenvalues equal to zero, as expected since in this case $\overline{\chi'_k}=\chi_k$. All the above results will be compared with empirical data (for the case of financial markets) in section 7.6.1 below.

7.5 The case of an isolated top eigenvalue

In the first subsection 7.5.1, we consider the case where the covariance matrix $\mathbf{C}$ has only one isolated eigenvalue much larger than all the other ones. By measuring the covariance matrix through an exponential moving average estimator along a time series of multivariate gaussian vectors with covariance matrix $\mathbf{C}$, we find the time evolution of the angle θ_t between the top eigenvector of the empirical covariance matrix and the top eigenvector of the covariance matrix $\mathbf{C}$. The result is a Langevin equation for $x_t\equiv 1-\cos\theta_t$ given in (7.48), for which we can compute the stationary distribution and even more information on the transition probability density. We also find the Langevin equation for the top eigenvalue of the empirical covariance matrix (see Eq. (7.46)). This enables to compute variograms of different related quantities in subsection 7.5.2, which will be useful later on for applications. In the final subsection (7.5.3), we analyse the transverse fluctuations of the top eigenvector of the empirical covariance matrix through a random matrix.

7.5.1 A Langevin equation for the top eigenvalue and eigenvector

A more detailed characterization of the dynamics of the top eigenvalue and eigenvector can be given in the case where this top eigenvalue is well separated from all the others, as is well known to be the case for financial covariance matrices. The financial interpretation of this large eigenvalue is the so-called 'market mode': in a first approximation, all stocks move together, up or down. In this subsection, we assume that the *true* covariance matrix $\mathbf{C}$ has one large eigenvalue λ_1 of order N well separated from the other ones, which are *all* equal to λ_2. We suppose that $\lambda_1\gg\lambda_2$.

Let $\left(r_i^t\right)_{1 \leqslant i \leqslant N}, 1 \leqslant t \leqslant T$ be i.i.d. Gaussian vectors of covariance $\mathbf{C}$. Both for technical convenience and to follow market practice, we suppose that the covariance matrix is now measured through an exponential moving average of the r_i^t. This means that the matrix $\mathbf{E}$ evolves in time as:

$$\mathbf{E}_{ij,t} = (1-\varepsilon)\mathbf{E}_{ij,t-1} + \varepsilon r_i^t r_j^t. \tag{7.44}$$

We address the following question: what is the dynamics of the top eigenvalue $\lambda_1(t)$ and of the top eigenvector ϕ_1^t of the empirical covariance matrix $\mathbf{E}_t$? Of course, the largest eigenvalue and eigenvector of the empirical covariance matrix will be, as discussed at length above, affected by measurement noise. Can one make predictions about the fluctuations of both the largest eigenvalue and the corresponding eigenvector induced by measurement noise? We shall see that such a decomposition is indeed possible in the limit where $\lambda_1 \gg \lambda_2$. The calculations in this section and in Appendix D follow closely those made in [124] which were slightly incorrect (see below).

We keep the same notations as in the previous section for the eigenvalues of $\mathbf{C}$. The eigenvalues and eigenvectors of $\mathbf{E}_t$ will be respectively denoted as $\lambda_1^t, \ldots, \lambda_N^t$ and $\phi_1^t, \ldots, \phi_N^t$.

Standard perturbation theory, valid for $\varepsilon \ll 1$, gives:

$$\lambda_1^t = (1-\varepsilon)\lambda_1^{t-1} + \varepsilon\langle\phi_1^{t-1}|C|\phi_1^{t-1}\rangle + \varepsilon\langle\phi_1^{t-1}|\eta_t|\phi_1^{t-1}\rangle,$$

with $\eta_{ij} = r_i r_j - \mathbf{C}_{ij}$. Because the returns are Gaussian, we have:

$$\overline{\eta_{ij}\eta_{k\ell}} = C_{ik}C_{j\ell} + C_{i\ell}C_{jk}.$$

In the limit where λ_1 becomes much larger than all other eigenvalues, the above equation simplifies to:

$$\lambda_1^t \approx (1-\varepsilon)\lambda_1^{t-1} + \varepsilon\cos^2(\theta_{t-1})\lambda_1\left[1+\xi_t\right], \tag{7.45}$$

where $\cos(\theta_t) \equiv \langle\phi_1^t|\phi_1\rangle$ and ξ_t is a random noise term of mean zero and variance equal to 2. In the limit of large matrices and $\varepsilon \to 0$, the above difference equation can be written as a Langevin (or stochastic differential) equation, in the Itô sense:

$$\mathrm{d}\lambda_1^t = \varepsilon\left[(\lambda_1 - \cos^2(\theta_t)\lambda_1^t)\mathrm{d}t + \sqrt{2}\lambda_1\cos^2(\theta_t)\,\mathrm{d}B_t\right]. \tag{7.46}$$

where B_t is a standard Brownian motion. We have neglected in the above equation a deterministic term equal to $\varepsilon\sin^2(\theta_t)\lambda_2$, which will turn out to be a factor λ_2/λ_1 smaller than the terms retained in Eq. (7.46). As we shall show below, the angle θ_t turns out to be small, so that one can replace $\cos(\theta_t)$ by unity in the above equation, which becomes a simple Ornstein-Uhlenbeck process. We therefore find for the variogram of λ_1:

$$\left\langle\left(\lambda_1^s - \lambda_1^t\right)^2\right\rangle \approx 2\varepsilon\lambda_1^2\left(1-\exp(-\varepsilon|t-s|)\right), \tag{7.47}$$

a result that we mentioned in the above section 7.4.2.

A similar SDE can be written for the projection of the instantaneous eigenvector $|\phi_1^t\rangle$ on the true eigenvector $|\phi_1\rangle$. This can again be done using perturbation theory, as is detailed in Appendix D. The quantity $\cos(\theta_t)$ is found to be close to 1 when ε is small, so we set $x_t \equiv 1 - \cos(\theta_t)$.

Keeping only the leading term in the three small parameters $\varepsilon, \lambda_2/\lambda_1$ and x_t, we finally find the following Langevin equation for x_t (in the Itô sense):

$$\mathrm{d}x_t = 2\varepsilon\left(\mu - x_t\right)\mathrm{d}t + \varepsilon\sqrt{2x_t(4x_t + \frac{\lambda_2}{\lambda_1})}\,\mathrm{d}B_t \tag{7.48}$$

with, for $N \to \infty$, $\varepsilon \to 0$,

$$\mu := \frac{q}{4}\frac{\lambda_2}{\lambda_1}, \quad \text{with} \quad q \equiv \varepsilon N.$$

Equation (7.48) defines a very interesting class of random processes, that we call "Pöschl-Teller" processes, on which we say more in Appendix E.

In the continuous time limit, we have therefore established two coupled Langevin equations (SDEs) for the top eigenvalue λ_1^t and x_t. To leading order and for $N \to \infty$, $\varepsilon \to 0$, the stationary solution for the "angle" x_t can be computed to be:

$$P(x) \propto \left(\frac{4x}{4x + \frac{\lambda_2}{\lambda_1}}\right)^{\frac{N}{2}} \left(\frac{1}{4x + \frac{\lambda_2}{\lambda_1}}\right)^{\frac{1}{2\varepsilon}},$$

which corrects the result obtained in [124], and is plotted in Fig. 7.5. From the above Langevin equation, it is immediate to see that the average value of x is given by $\overline{x} = \mu$. It is nicer to rewrite the stationary distribution in terms of $\hat{x} = x/\mu$. The interesting regime is when q remains of order unity when $N \to \infty$ and $\varepsilon \to 0$, in which case:

$$P(\hat{x}) \approx Z e^{-\frac{N f(\hat{x})}{2}}, \qquad f(\hat{x}) = \frac{\ln(1+q\hat{x})}{q} + \ln\left(1 + \frac{1}{q\hat{x}}\right),$$

where Z is a normalisation. It is easy to see that $f(\hat{x})$ has a minimum for $\hat{x} = 1$, or $x = \mu$ (corresponding to the most probable value), and that $f''(1) = 1/(1+q)$. This shows that the fluctuations of $\hat{x}$ around $\hat{x} = 1$ are of order $\sqrt{(1+q)/N}$ and thus very small in the large N limit.

Note finally that according to Eq. (7.46), the largest eigenvalue is on average *shifted upwards* compared to the true value λ_1, by a factor $\approx (1 + 2\mu) = (1 + \frac{q}{2}\frac{\lambda_2}{\lambda_1})$. This is the analogue of a similar well-known result for flat-window averages of empirical covariance matrices – see [19, 29].

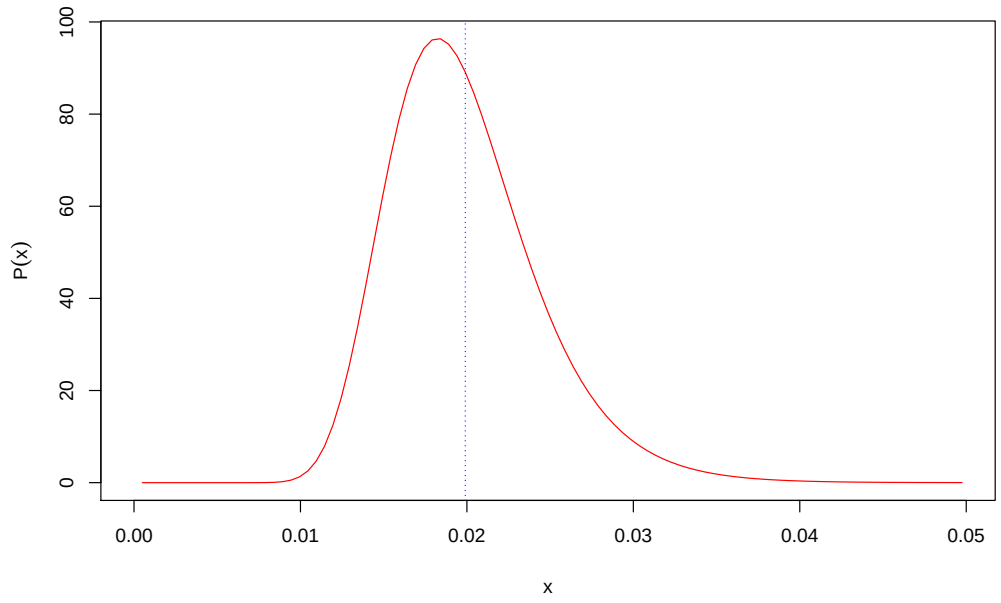


Figure 7.5: A picture of the stationary probability density $P(x)$ of the process x_t verifying (7.48). The parameters are: $\varepsilon = 1/50$, $N = 200$ (corresponding to $q = 4$) and $\lambda_2/\lambda_1 = 0.02$. The vertical blue dotted line shows the position of $\mu \approx 0.02$ for this choice of parameters.

7.5.2 Variograms

From the Langevin equation one can easily compute the second moment $\overline{x_t^2}$ with as initial condition $x_0 = 0$. Indeed, using Itô's formula and taking expectations, we get:

$$\overline{x_t^2} = x_0^2 + 4\varepsilon\left(\mu + \varepsilon\frac{\lambda_2}{2\lambda_1}\right)\int_0^t \overline{x_s}\mathrm{d}s - 4\varepsilon(1-2\varepsilon)\int_0^t \overline{x_s^2}\,\mathrm{d}s\,.$$

Computing $\overline{x_t}$ with the same technique, we can solve this ordinary differential equation to obtain that

$$\overline{x_t^2} = x_0^2 e^{-4\varepsilon(1-2\varepsilon)t} + \frac{\mu(\mu + \varepsilon\frac{\lambda_2}{2\lambda_1})}{1-2\varepsilon}\left(1 - e^{-4\varepsilon(1-2\varepsilon)t}\right) + \frac{(2\mu + \varepsilon\frac{\lambda_2}{\lambda_1})(x_0 - \mu)}{1-4\varepsilon}\left(e^{-2\varepsilon t} - e^{-4\varepsilon(1-2\varepsilon)t}\right).$$

In order to characterize the dynamics of the angle fluctuations, we want to compute the variogram of x_t, defined as $\upsilon(\tau) := \overline{(x_{t+\tau} - x_t)^2}$ for $\tau \geqslant 0$, and in the limit $t \to \infty$. Using the previous computations, we obtain, in the scaling limit:

$$\upsilon(\tau) \approx \frac{q^2(1+q)}{4N}\left(\frac{\lambda_2}{\lambda_1}\right)^2 \left(1 - e^{-2\varepsilon\tau}\right).$$

We show in Fig. 7.6 a numerical simulation of the dynamics of the top eigenvector of a fixed matrix $\mathbf{C}$ such that $\lambda_2/\lambda_1 = 0.033$. The resulting variogram compares very well with the above prediction.

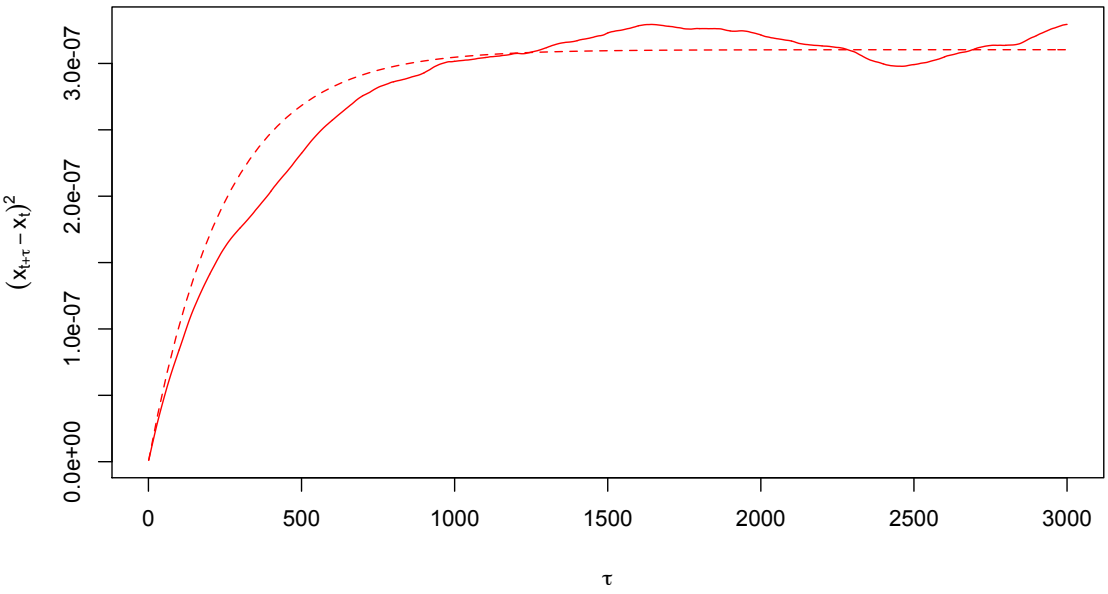


Figure 7.6: The plain line represents the function $\upsilon(\tau)$ as a function of τ for $\varepsilon = 0.002, N = 200$ ($q = 0.4$) and $\lambda_2/\lambda_1 = 0.033 = 30$. The dotted line is a numerical simulation of the semivariogram of x_t in the benchmark case where there is a constant in time correlation matrix $\mathbf{C}$.

However, the above calculation is not particularly useful for financial applications, since the "true" top eigenvector $|\phi_1\rangle$, needed to define the angle θ_t, is in general not known. A more appropriate quantity to describe the dynamical fluctuations of $|\phi_1\rangle$ is, as suggested in [124], the function $\tau \to \overline{\langle \phi_1^t | \phi_1^{t+\tau} \rangle}$, which we now study analytically. Let us write $|\phi_1^t\rangle$ as

$$|\phi_1^t\rangle = \cos(\theta_t)|\phi_1\rangle + |\varphi_\perp^t\rangle, \tag{7.49}$$

where φ_2^t is a vector in the eigenspace corresponding to the small eigenvalues λ_2. Therefore:

$$\langle \phi_1^t | \phi_1^{t+\tau} \rangle = \cos(\theta_t)\cos(\theta_{t+\tau}) + \langle \varphi_\perp^t | \varphi_\perp^{t+\tau} \rangle.$$

Now, it is easy to have an explicit expression for φ_2^t by considering the empirical covariance (or correlation) matrix $\mathbf{E}^t$ as a perturbation of the true covariance matrix $\mathbf{C}$, as we did above. Standard perturbation theory then gives

$$|\phi_1^t\rangle = \left(1 - \frac{1}{2}\sum_{i\neq 1}\frac{\langle\phi_1|\mathcal{E}^t|\phi_i\rangle^2}{(\lambda_1-\lambda_2)^2}\right)|\phi_1\rangle + \sum_{i\neq 1}\frac{\langle\phi_1|\mathcal{E}^t|\phi_i\rangle}{\lambda_1-\lambda_2}|\phi_i\rangle$$

where

$$\mathcal{E}_{ij}^t = \varepsilon \sum_{s=0}^{+\infty} (1-\varepsilon)^s \left(r_i^{t-s} r_j^{t-s} - C_{ij} \right) .$$

It is clear that the last term of the above expression is exactly $|\varphi_2^t\rangle$, which enables us to obtain:

$$\langle \varphi_\perp^t | \varphi_\perp^{t+\tau} \rangle = \frac{1}{(\lambda_1 - \lambda_2)^2} \sum_{i \neq 1} \langle \phi_1 | \mathcal{E}^t | \phi_i \rangle \langle \phi_1 | \mathcal{E}^{t+\tau} | \phi_i \rangle .$$

But, by noting that:

$$\mathcal{E}_{ij}^{t+\tau} = (1-\varepsilon)^\tau \mathcal{E}_{ij}^t + \varepsilon \sum_{s=0}^{\tau-1} (1-\varepsilon)^s \left(r_i^{t+\tau-s} r_j^{t+\tau-s} - C_{ij} \right)$$

and with the fact that $\overline{\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle^2} = \varepsilon \lambda_1 \lambda_2 / 2$, we get that:

$$\overline{\langle \varphi_\perp^t | \varphi_\perp^{t+\tau} \rangle} \approx 2\,\mu\, e^{-\varepsilon\tau} ,$$

and hence, our final result, to lowest order in μ:

$$\overline{\langle \phi_1^t | \phi_1^{t+\tau} \rangle} = \overline{(1-x_t)(1-x_{t+\tau})} + \overline{\langle \varphi_\perp^t | \varphi_\perp^{t+\tau} \rangle} \tag{7.50}$$

$$\approx 1 - 2\mu \left(1 - e^{-\varepsilon\tau} \right) . \tag{7.51}$$

which is similar to the result obtained in [124], except that the coefficient μ was a factor N too small in that paper. This result will be compared with empirical data in section 7.6.2.

7.5.3 Transverse fluctuations of the top eigenvector

In order to go further and describe the evolution of the top eigenvector of $\mathbf{E}$ (the so-called "market mode" in the context of financial markets), we need to study the statistics of the transverse component $|\varphi_\perp^t\rangle$. In order to make sense of the pattern created by these transverse fluctuations, we propose to introduce the correlation matrix of the components of $|\varphi_\perp^t\rangle$ in the eigen-basis of the true correlation matrix. We therefore define the following $N-1 \times N-1$ matrix:

$$\mathbf{F}_{ij} = \frac{1}{T} \sum_{t=1}^{T} \langle \varphi_\perp^t | \phi_i \rangle \langle \varphi_\perp^t | \phi_j \rangle \, (i,j \geqslant 2)$$

The eigenvalues and eigenvectors of this new correlation matrix (not to be confused with the empirical correlation matrix $\mathbf{E}$ needed to define $|\varphi_2^t\rangle$!) will entirely characterize the transverse fluctuations of the "market mode".

In the benchmark case where there is a true correlation matrix $\mathbf{C}$ stable in time, one can check that:

$$\mathbf{F}_{ij} = \frac{1}{T} \sum_{t=1}^{T} \frac{\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle}{\lambda_1 - \lambda_i} \frac{\langle \phi_1 | \mathcal{E}^t | \phi_j \rangle}{\lambda_1 - \lambda_j} .$$

What is the eigenvalue spectrum of $\mathbf{F}$ for this benchmark case? In our case where for all $i \neq 1$, $\lambda_i = \lambda_2$, the density of states of this type of random matrix has been studied before in the literature (see [121]). Indeed the random variables $\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle$ are uncorrelated for $i \neq j$, their mean is 0 and their variance is given by:

$$\overline{\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle^2} = \frac{\varepsilon \lambda_1 \lambda_2}{2} . \tag{7.52}$$

However, the random variables $\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle$ *are correlated in time* and thus the density of states in the limit of large matrices will not be given by the usual Marchenko-Pastur law. Rather, $\langle \phi_1 | \mathcal{E}^t | \phi_i \rangle$ follows an auto-regressive linear process, for which the authors of [121], give a precise way to compute the density of states in the limit of large matrices by mean of its Stieltjes transform. This probability density depends as expected on the parameter N/T but also on the parameter ε of the auto-regression. In the case where $\lambda_{i>1} = \lambda_2$, one furthermore expects that

the eigenvectors of $\mathbf{F}$ are isotropically distributed in the $N-1$ dimensional subspace spanned by $|\phi_2\rangle, \ldots |\phi_N\rangle$. This means that the transverse fluctuations $|\varphi_\perp\rangle$ of the top eigenvector have no particular structure.

In the more general context where the λ_i for $i \neq 1$ are not all equal to λ_2, the eigenvalue spectrum of $\mathbf{F}$ must be characterized numerically, see below.

7.6 Empirical results

For the following analysis, we have used the daily returns of several pools of stocks belonging to 4 major indices: SP500, Nikkei, DAX & CAC 40. The number of stocks are respectively $N = 500, 204, 30, 39$ and the period of interest is $2000 - 2010$ (11 years of data, corresponding to ≈ 2750 days). The main issue, as alluded to above, is that the empirical determination of correlation[11] matrices requires some measurement time T. If this time is too short, the empirical correlation matrix will appear to evolve with time, but this may just be due to the measurement noise which one would like to distinguish from a genuine evolution of the underlying structure of correlation. If the measurement time is too long, on the other hand, one may miss important correlation shifts and get exposed to unwanted sources of risk.

7.6.1 Stability of eigenspaces

We first determined the empirical variograms $\langle(\lambda_i^s - \lambda_i^t)^2\rangle_{|t-s|=\tau}$ for $i = 1, 2$, the result (for $i = 1$) is shown in Figure 7.7 and is found to be much larger than the above theoretical prediction, i.e. $4\lambda_i^2/T$, shown as a horizontal plain line. The fact that the empirical (red) curve starts from 0 for $\tau = 0$ and increases to reach the stationary noise level at time $\tau = T$ is simply due to the overlapping between the sliding periods. For those figures, we computed the time series of correlation matrices using a sliding window of size $T = N$ (recall N is the number of stocks). Thus, for small markets like DAX and CAC40, this value is quite small (respectively 30 and 40) and we find that the first eigenvalue of the correlation matrix does not evolve too much during the following (non overlapping) period $\tau \in [T; 250]$ days. After this time period, the evolution appears and from this point, the difference between the two non overlapping periods increases significantly with the time lag. For larger markets such as SPX and Nikkei, the value of T is quite large as N is respectively equal to 500 and 200. So it is not very surprising the temporal evolution shows up immediately. This clearly shows that there is a genuine evolution of the eigenvalues of $\mathbf{C}$ with time. For the top eigenvalue, this is a well known effect (see [124] and section 7.5.2, Fig. 7.6 below): both the volatility of individual stocks and the average correlation between stocks are indeed time dependent, and tend to increase in crisis periods [8, 27]. We see that the same is true for smaller eigenvalues too, reflecting the instability of intra-sector correlations (data not shown).

But what about the eigenvectors? One could be in a "mixed" situation where the eigenvectors of the true underlying covariance $\mathbf{C}$ keep a fixed direction through time[12] while its eigenvalues are moving around. But if the eigenvalues of the matrix $\mathbf{C}$ (which was always supposed not to depend of time in the previous sections) themselves are evolving with time, the formulas derived in the theoretical section above need to be upgraded. Let us assume that the true covariance matrix $\mathbf{C}_t$ has time dependent eigenvalues $\lambda_1^t, \ldots, \lambda_N^t$ but with constant eigenvectors which will be denoted $|\phi_1\rangle, \ldots, |\phi_N\rangle$ as above. For times $s < t$ with $|t-s| \geqslant T$, we define the overlap matrix $\mathbf{G}^{s,t}$ as: $G_{ij}^{s,t} = \langle\phi_i^s|\phi_j^t\rangle$. Under the assumption that the eigenvalues are varying

[11] Our results in the previous sections hold for empirical covariance matrices. Hence we centered and normalized our empirical time series of returns so as to use them.

[12] Here we mean that the non-perturbed (or population) eigenvectors do not evolve with time; obviously we do not talk about the sample eigenvectors of the empirical covariance matrix $\mathbf{E}$ which will be affected by measurement noise, evolving around the population eigenvectors.

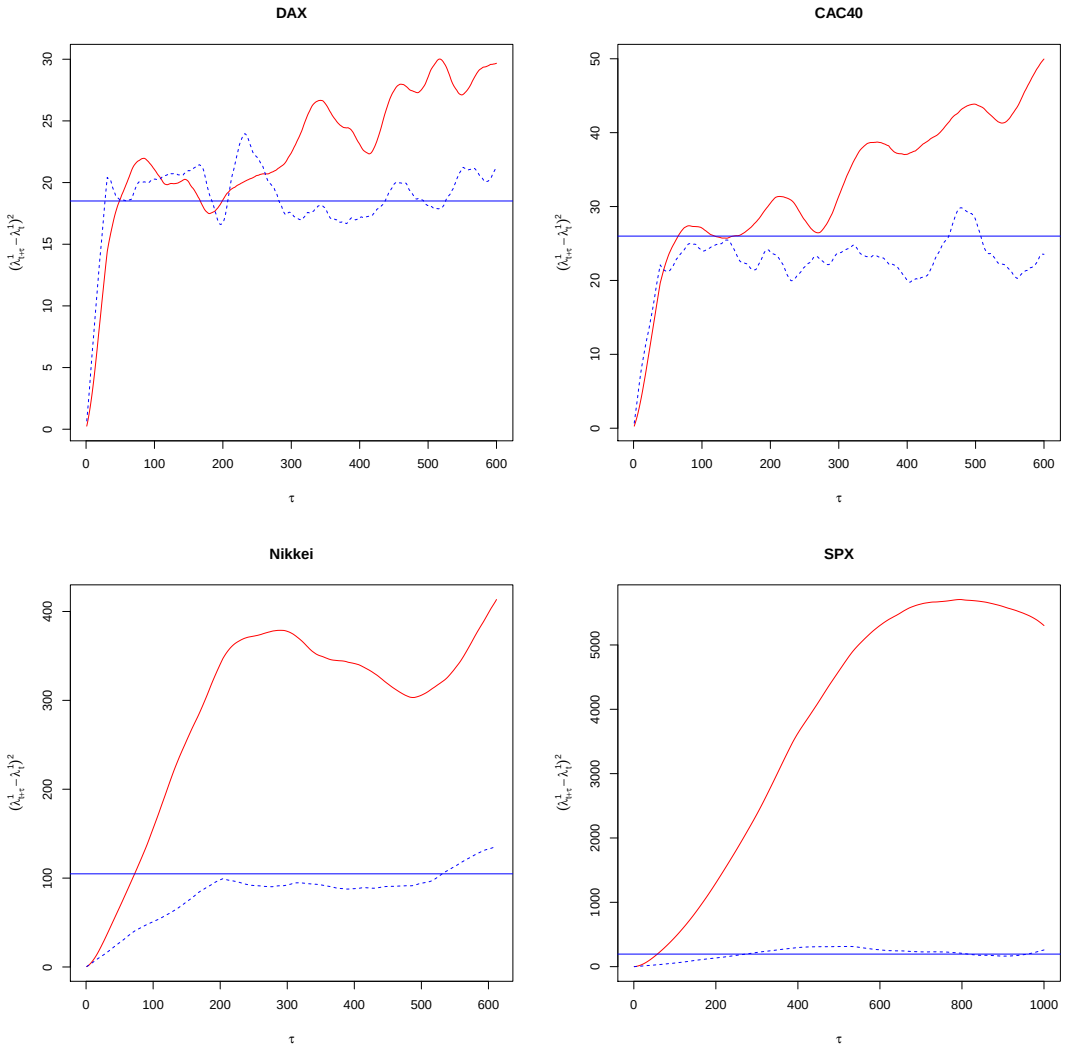


Figure 7.7: Plot of $\langle(\lambda_1^s - \lambda_1^t)^2\rangle_{|t-s|=\tau}$ as a function of τ for the four different indexes of our sample. The empirical correlation matrices are computed on a sliding window of size $T = N$. The red line corresponds to the empirical datas from our pools of stocks, the plain blue line is the theoretical prediction $4\lambda_1^2/T$ (valid in the limit of large T) and the dotted blue line represents a numerical simulation of the benchmark case. Very similar curves hold for the second and third eigenvalues as well.

sufficiently slowly with time, one now finds that:

$$D(P,Q;s,t) = -\frac{1}{2P}\left\langle \ln|\det(\mathbf{G}^{s,t\dagger}\mathbf{G}^{s,t})|\right\rangle \approx \frac{1}{2TP}\sum_{i=1}^{P}\sum_{j=Q+1}^{N}\left(\frac{\lambda_i^s\lambda_j^s}{(\lambda_i^s-\lambda_j^s)^2}+\frac{\lambda_i^t\lambda_j^t}{(\lambda_i^t-\lambda_j^t)^2}\right). \tag{7.53}$$

Up to corrections of order $T^{-3/2}$, one can replace in the above formulas the $\lambda^{s,t}$ by their empirical estimates. We finally compute the theoretical distance $D_{th}(P,Q,\tau)$ as an average over all s,t such that $|t-s|=\tau$ of the above quantity.

We now compare our null hypothesis formula, Eq. (7.53) with (a) an empirical determination of $D_{emp}(P,Q,\tau)$ using financial data and (b) a numerical determination of $D_{num}(P,Q,\tau)$ using synthetic time series of returns which abide to the hypothesis of a covariance matrix $\mathbf{C}_t$ with *fixed* eigenvectors, but time dependent eigenvalues. To achieve this, we choose an arbitrary (but fixed) set of orthonormal vectors $|\psi_1\rangle,\ldots,|\psi_N\rangle$ and define $\mathbf{C}_t$ as $\mathbf{C}_t=\sum_{i=1}^N\lambda_i^t|\psi_i\rangle\langle\psi_i|$, where the λ^t are the empirical eigenvalues obtained on the financial return time series. We then use $\mathbf{C}_t$ to generate synthetic Gaussian multivariate returns $\{r_i(u)\}$. We show the corresponding results in Fig. 7.8, with the choice $P=5, Q=10$, as a function of τ and for $T=N$ days. As above, the study concerns the same 4 different pools of stocks corresponding to 4 major indices: SP500, Nikkei, DAX, CAC 40. We conclude that (i) the theoretical formula Eq. (7.53) is indeed in very good agreement with the numerical results obtained with synthetic data: $D_{num}\approx D_{th}$; whereas (ii) the financial data clearly departs from the null hypothesis of constant eigenvectors, since $D_{emp}>D_{th}$. The same conclusion holds for different values of P,Q.

We have also computed the value $D_{emp}(\tau=T)$ for different values of T for every pool of stocks, the result is shown in Fig. 7.9. We compare the empirical function $T\to D_{emp}(T)$ with the theoretical value $D_{th}(T)$ in the benchmark case where the stock returns are distributed as Gaussian vectors of constant covariance matrix $\mathbf{C}$. At first sight, the noise contribution appears to be too small to explain the value of $D_{emp}(T)$ at small Ts, at least for the pool of the CAC40 and DAX indices. Nevertheless, if we now compare the value of $D_{emp}(\tau=T)$ for small value of T with the value of $D_{th}(\tau=T)$ in the benchmark case where the stock returns are distributed with a multivariate Student distribution with ν-degrees of freedom and with a constant covariance matrix $\mathbf{C}$, we see that we can make the two curves coincide for small values of T. Therefore, the initial decline as T increases indeed follows from a reduction of the measurement noise. However, when T becomes very large, the "true" evolution of the eigenvectors starts being visible, and leads to an increase of D_{emp}. This plot suggests that the optimal time scale to measure the empirical eigenspaces is around $T^*=600$ days for the stocks from the Nikkei index, $T^*=400$ days for the ones from CAC40, $T^*=450$ days for the ones from DAX and $T^*=700$ days for the ones from the SP500 index.

The above results are fully confirmed, and made more precise, by the spectral projector analysis proposed by Zumbach. In Fig. 7.10 we plot, as in [157], the eigenvalues of the average spectral projector $\overline{\chi'_k}$ as a function of its theoretical rank k, for several values of k. We show in plain lines the eigenvalues of the empirically determined $\overline{\chi'_k}$ for the Nikkei idex, where the averaging is made over (overlapping) periods of length $T=600$ days, and in dotted lines the corresponding theoretical predictions Eqs. (7.42) and (7.43) for the benchmark case where the eigenspaces are fixed in time, but are blurred by measurement noise. Here again we find clear signals of a true evolution of the eigenspaces. The results for other stock indices are very similar.

7.6.2 The dynamics of the top eigenvector

As explained above, one expects in general the top eigenvector to wobble around its "true" direction $|\phi_1\rangle$. The fluctuations around $|\phi_1\rangle$ have two possible origins: one is measurement noise, the other is the presence of a systematic rotation of the top eigenvector due to some financial mechanism.

As a further check that measurement noise is not enough to explain the observed dynamics of $|\phi_1^t\rangle$, we have studied numerically the average overlap of the top eigenvector measured a time τ apart: $\overline{\langle\phi_1^t|\phi_1^{t+\tau}\rangle}$. This is an interesting quantity because it does not require the knowledge

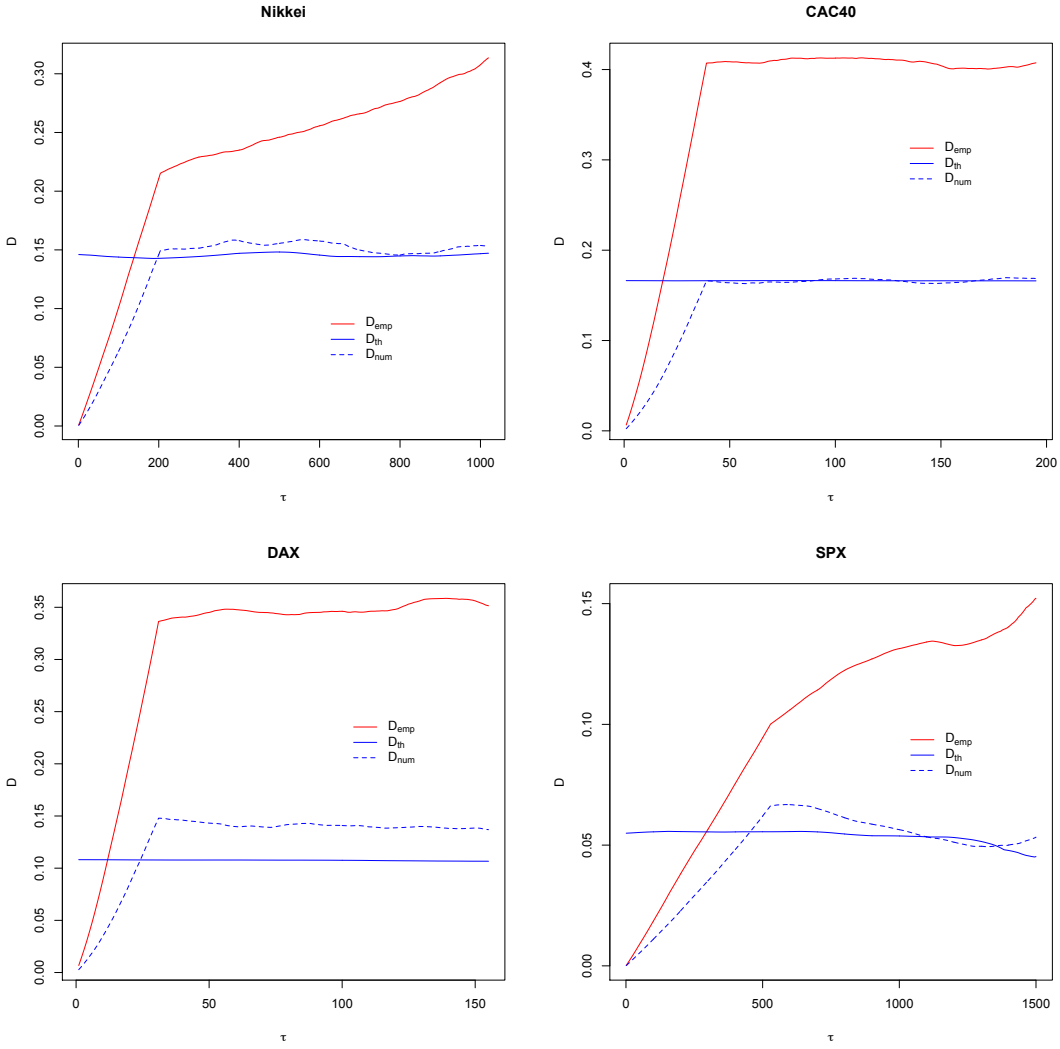


Figure 7.8: Plot of D_{th}, D_{num}, D_{emp} for $T = N$, $P = 5, Q = 10$ for the four indices considered here. The blue lines are theoretical benchmark results for fixed eigenvector directions (plain line: analytical result, dotted line: numerical simulations, while the red line is the empirical result). These plots clearly show that the subspace spanned by the 5 top eigenvectors evolve with time.

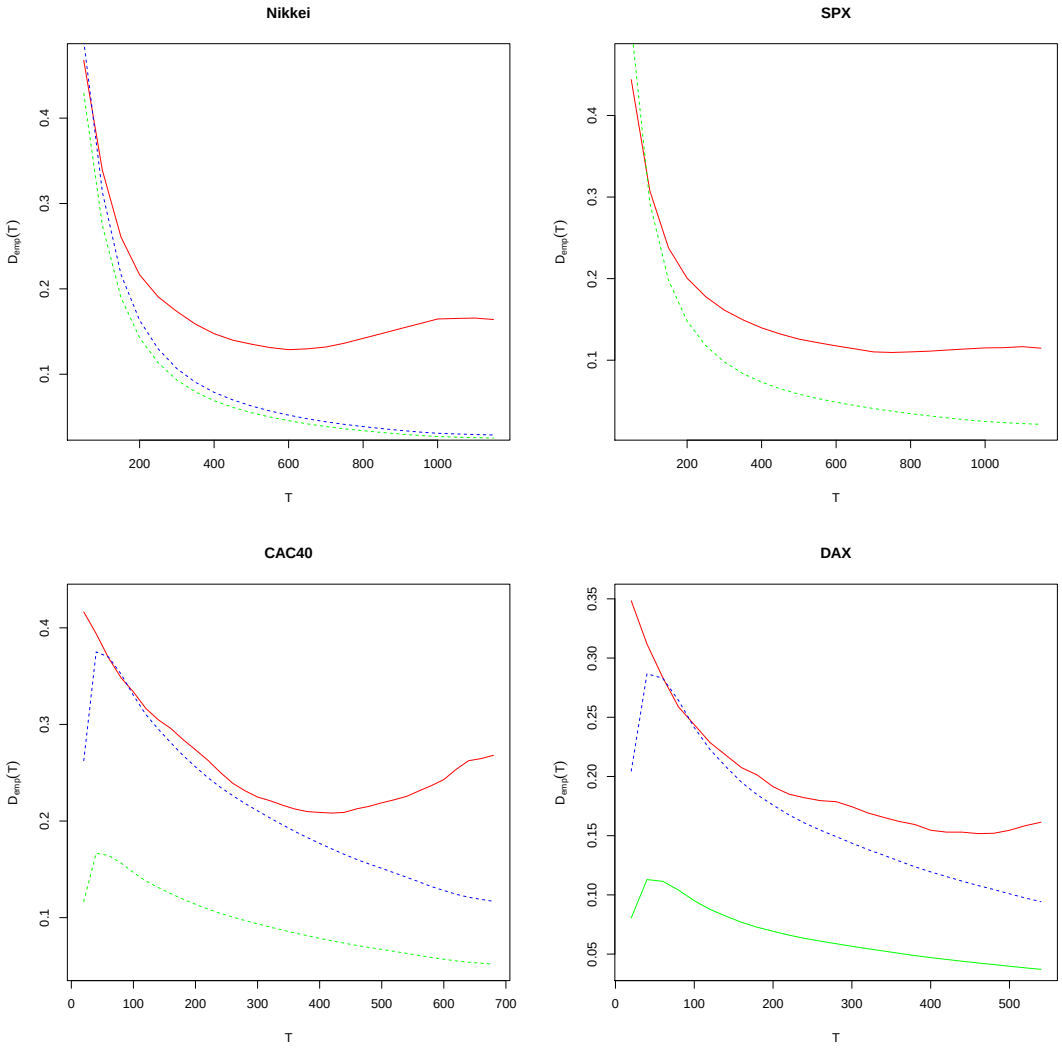


Figure 7.9: Plot of $D_{emp}(\tau = T)$ (red line) and $D_{th}(\tau = T)$ as a function of T, $P = 5, Q = 10$ for the four indices considered here. The dotted green line represents $D_{th}(\tau = T)$ in the benchmark case where the returns are Gaussian with constant covariance matrix $\mathbf{C}$ and the dotted blue line represents $D_{th}(\tau = T)$ in the benchmark case where the returns are distributed with a multivariate Student distribution with ν-degrees of freedom and with a constant covariance matrix $\mathbf{C}$. The constant ν is chosen equal to 5.5 for the CAC40 and DAX indexes and to 18 for the Nikkei index. The initial decline as T increases follows from reducing the measurement noise. However, when T becomes very large, the "true" evolution of the eigenvectors is being felt, and leads to an increase of D_{emp}. This plot suggests that the optimal time scale to measure the empirical eigenspaces is around two years ($T^* = 500$ days).

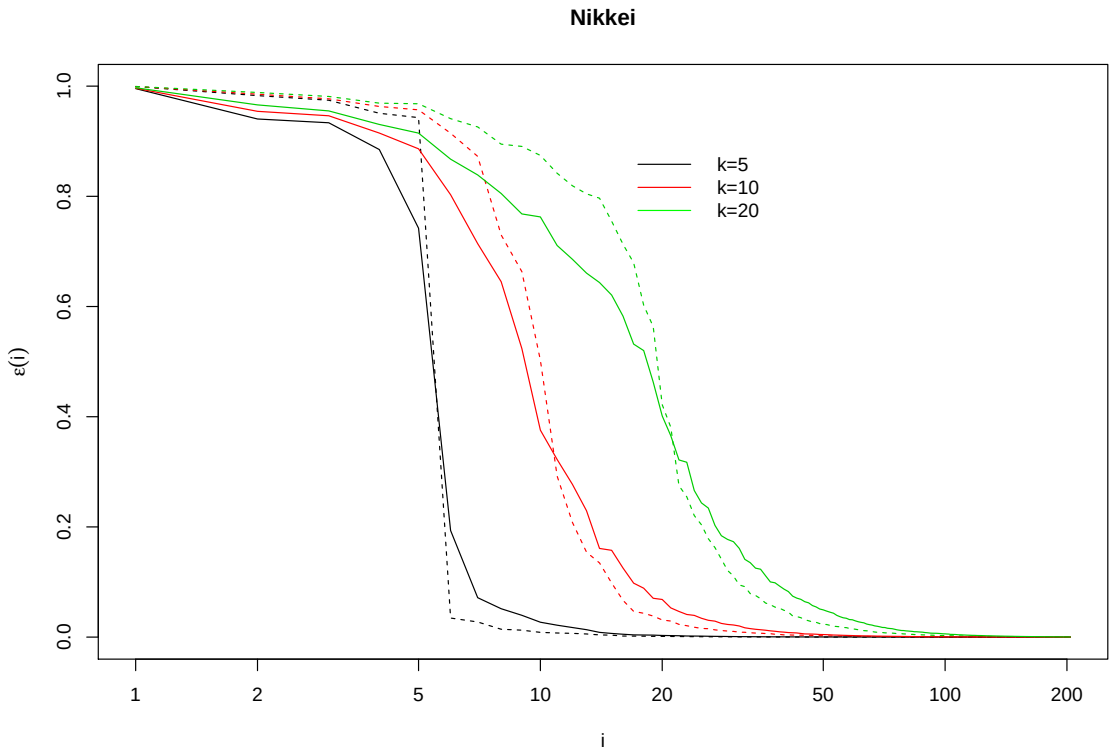


Figure 7.10: The dotted lines represent the eigenvalues ϵ_i of the spectral projector of rank k (for $k = 5, 10, 20$ see the legend) as a function of $\log(i), i = 1, \ldots, M$ in the benchmark case where the true correlation matrix $\mathbf{C}$ is not evolving but dressed by measurement noise. The plain lines represent the same function for the empirical data from the Nikkei index (204 stocks between $2000 - 2010$). Here $T = 600$. In the ideal case (constant correlation matrix, $T \to \infty$), these functions should be step functions: $\epsilon_{i \leqslant k} = 1$ and $\epsilon_{i>k} = 0$.

of the true direction $|\phi_1\rangle$. As shown above, this quantity should be approximately given by $1 - 2\mu\,(1 - e^{-\varepsilon\tau})$ if measurement noise is the only source of fluctuations. We show in Fig. 7.11 a comparison between this prediction and empirical data on the market mode of the Nikkei index. Here again, we find that the decorrelation of the top eigenvector is much stronger than the benchmark. The deviation from unity is, for $\tau = 350$, more than three times larger than the benchmark case, with no signs of saturation.

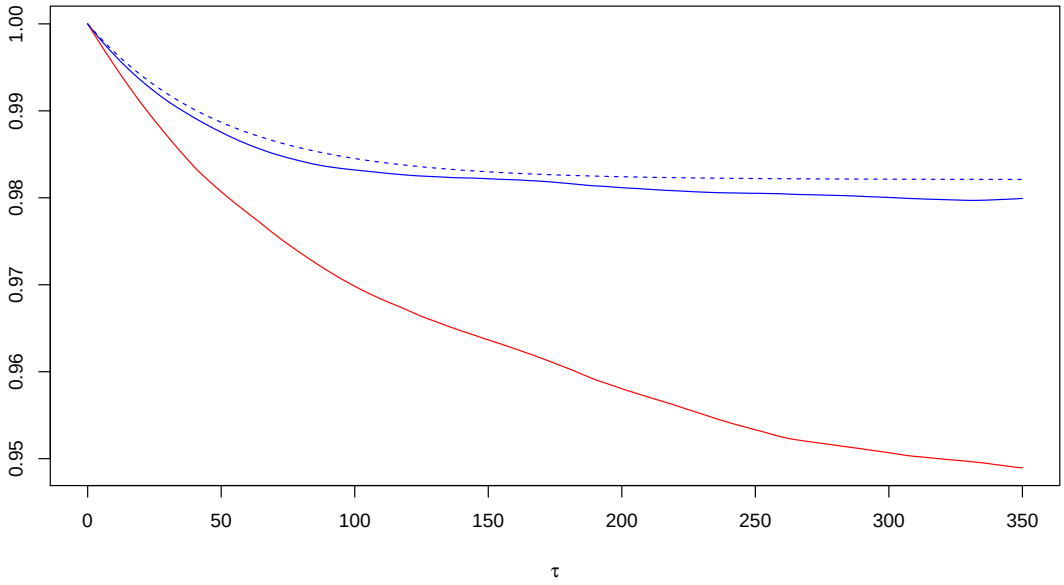


Figure 7.11: The plain line represents the empirical function $\langle \phi_1^t | \phi_1^{t+\tau} \rangle$ as a function of τ. The period on which the average is performed has 2336 days starting 01/01/2000. There are $N = 204$ stocks from the Nikkei index. The exponential moving average is made with a parameter $\epsilon = 1/50$. The true empirical correlation matrix $\mathbf{C}$ is chosen to be the empirical correlation matrix computed using the data on the whole period. For this $\mathbf{C}$, we have $\lambda_1 \approx 73$ and $\lambda_2 \approx 0.7$. The beginning of the period is used to initialize the exponential moving average. The plain blue is a numerical simulation in the benchmark case. The dotted line represents the function $\tau \to 1 - 2\mu(1 - \exp(-\varepsilon\tau))$ which corresponds to the benchmark case when there is a constant in time correlation matrix.

So there is a genuine motion of the top eigenvector in time. This was already pointed out in [8], where we established empirically that the top eigenvector rotates towards the uniform vector $|e\rangle = (1, 1, \ldots, 1)/\sqrt{N}$ when the market goes down, and away from $|e\rangle$ when the market goes up. In order to be more comprehensive and understand in details the dominant transverse fluctuations of the top eigenvector, we have studied the correlation matrix $\mathbf{F}$ defined in subsection 7.5.3 above. We first determined the eigenvalue spectrum of $\mathbf{F}$ numerically, both for the benchmark case (with only measurement noise) and for real empirical data, see Fig. 7.12. From this figure, we conclude that, for the Nikkei index during the period $2000 - 2010$, there are 3 (maybe 4) eigenvalues of the empirical matrix $\mathbf{F}$ that reside outside the spectrum of the corresponding benchmark matrix. This suggests that these 3 or 4 modes are real and correspond to true fluctuations of the market mode, which contribute to the discrepancy displayed in Fig. 7.11 above. We are now in a position to identify the corresponding eigenvectors, i.e. the directions in which the market mode most likely to tilt.

It is natural to think that these directions should themselves correspond to large eigenvectors of the correlation matrix $\mathbf{C}$. Therefore we look for the decomposition of the top three eigenvectors of $\mathbf{F}$ (that we call $|\omega_1\rangle, |\omega_2\rangle, |\omega_3\rangle$) in terms of $|\phi_i\rangle, i \in \{2, 3, 4, 5\}$. A singular value analysis of the 3×4 overlap matrix shows that one can indeed explain $\approx 85\%$ of these three eigenvectors

in this way, with:

$$\begin{aligned} |\omega_1\rangle &\approx -0.34\,|\phi_2\rangle + 0.29\,|\phi_3\rangle + 0.30\,|\phi_4\rangle + 0.84\,|\phi_5\rangle \\ |\omega_2\rangle &\approx 0.53\,|\phi_2\rangle + 0.45\,|\phi_3\rangle + 0.47\,|\phi_4\rangle - 0.54\,|\phi_5\rangle \\ |\omega_3\rangle &\approx 0.77\,|\phi_2\rangle + 0.40\,|\phi_3\rangle + 0.48\,|\phi_4\rangle\,. \end{aligned} \tag{7.54}$$

This means that all the four top eigenvectors of $\mathbf{C}$ contribute to the "tilt motion" of the market mode. To check that this result is significant, we ran numerical simulations for this singular value decomposition in the benchmark case with a constant correlation matrix $\mathbf{C}$ chosen as before to be the empirical correlation matrix computed using the whole period of time (here the decade $2000-2010$). The 3×4 singular values analysis now give an explanatory power of $\approx 70\%$, which is clearly less than the 85% obtained above. Still, a large part of this explanatory power seems to trivially come from the non random structure of $\mathbf{C}$ itself.

In order to revisit the result found in [8], we need to understand the link between the uniform vector $|e\rangle$ and the eigenvectors $|\phi_2\rangle,\ldots,|\phi_5\rangle$ of the correlation matrix $\mathbf{C}$. Thus, we look at the orthogonal projection $|e_\perp\rangle := (|e\rangle - \langle e|\phi_1\rangle|\phi_1\rangle)/\mathcal{N}$ ($\mathcal{N}$ is chosen such that $\langle e_\perp|e_\perp\rangle^2 = 1$) of the uniform vector $|e\rangle$ in the space generated by the $|\phi_i\rangle, i \geqslant 2$. The overlap $\langle e_\perp|\phi_i\rangle$ for all $i > 2$ are shown in fig. 7.13 for the Nikkei index during the period $2000-2010$. We see that $|e_\perp\rangle$ has indeed very strong overlap with $|\phi_2\rangle, |\phi_3\rangle, |\phi_4\rangle, |\phi_5\rangle$, and hence, from the above results, also with $|\omega_1\rangle, |\omega_2\rangle, |\omega_3\rangle$. Therefore, the fact that the main fluctuation modes of $|\phi_1\rangle$ are along these three ω directions is compatible with the tilt motion towards $|e\rangle$. However, other modes, not mentionned in [8], are detected by the present analysis.

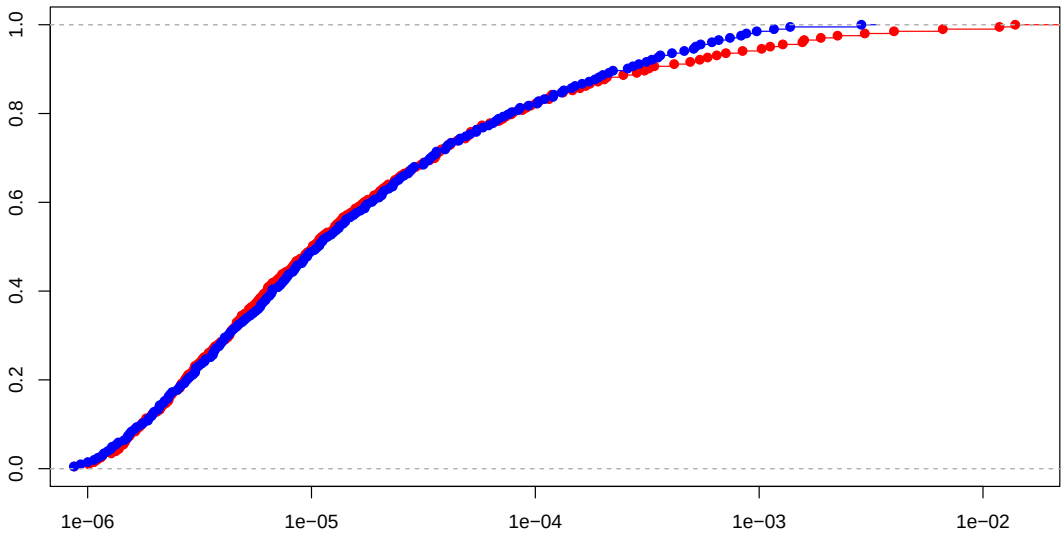


Figure 7.12: The red curve represents the cumulative distribution of the density of states of the matrix $\mathbf{F}$ for the Nikkei index with $N = 204$ stocks, in the period $2000-2010$, with $\varepsilon = 1/50$. The blue curve is a numerical simulation for the benchmark case with the true correlation matrix $\mathbf{C}$ chosen to be the empirical correlation matrix using the whole period. For this period and pool of stocks, we have $\lambda_1 \approx 73$ and $\lambda_2 \approx 0.7$.

7.7 Conclusion & Open problems

Let us try to summarize what we have achieved in this paper. We have developed general tools to describe the dynamics of eigenvectors under the influence of small random perturbations and to study the stability of the subspace spanned by P consecutive eigenvectors of a generic symmetric matrix. This problem is relevant in various contexts, including quantum dissipation

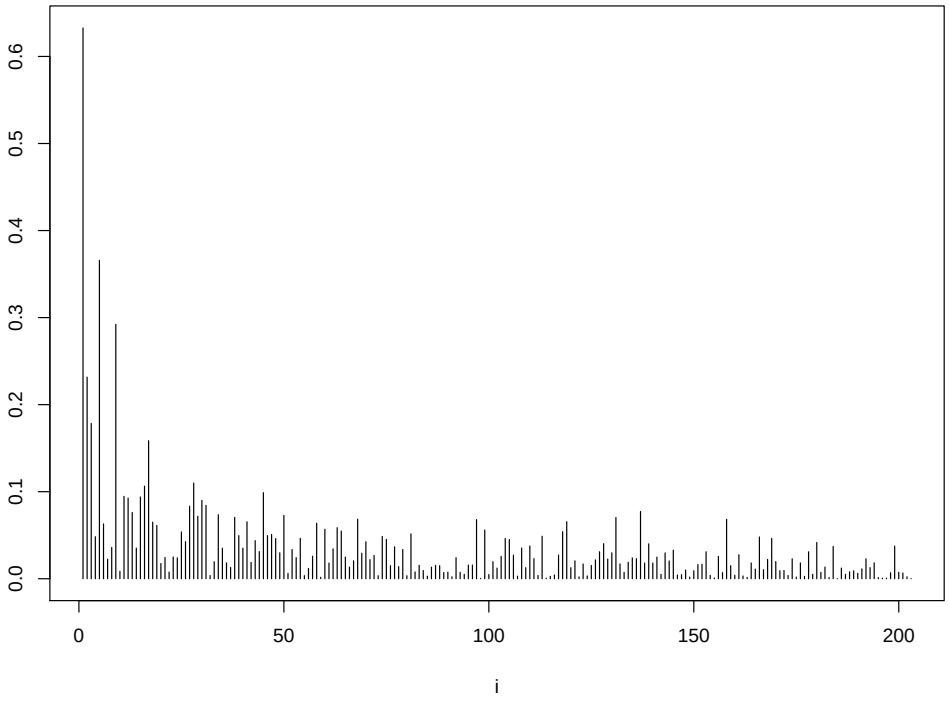


Figure 7.13: Plot of the overlap $|\langle e_\perp|\phi_i\rangle|$ as a function of i for $i \geqslant 2$. This graph shows that the main contribution to $|e_\perp\rangle$ comes from the top eigenvectors of the correlation matrix $\mathbf{C}$.

and financial risk control, but hopefully the ideas and methods introduced here can be used in a much broader context.

We argue that the problem can be formulated in terms of the singular values of the overlap matrix between the initial eigenspace and the target eigenspace, which allows one to define an overlap distance, which is small if most of the initial information is conserved. We first specialize our results for the case of a Gaussian Orthogonal Ensemble, for which the full spectrum of singular values can be explicitly computed in the limit of large matrices under the regime where the entries of the perturbation are very small compared to the mean level spacing of the non-perturbed matrix. We argue that our setting with rectangular $Q \times P$ overlap matrices $\mathbf{G}$ allows to extend our results to perturbations with entries larger than the mean level spacing. We provide some numerical evidences that it is indeed true. We find two regimes, depending on the dimension of the target space Q compared to that of the initial space P. If $Q \gg P$, all singular values are close to one another, and their distribution is given by Wigner's semi-circle. If on the other hand $(Q-P)/P \ll 1$, the singular values s are distributed according to a very broad law which decays as s^{-2}. These results are actually universal, and apply for other matrix ensemble as well – for example the case of empirical covariance matrices – provided one is interested in eigenspaces deep in the bulk.

We have also studied the case of isolated eigenvalues, which are usually very important for applications, for example in finance. In most cases, empirical correlation matrices are noisy measurements of the true covariance matrix and this can lead to an apparent evolution of the top eigenspace, whereas in reality the underlying process is stationary. We have derived exact expressions both for the overlap distance and for the average spectral projectors (introduced by Zumbach [157]) which can be directly compared to empirical results. The special case where the top eigenvalue is much larger than all the other ones can be investigated in full detail. In particular, the dynamics of the angle made by the top eigenvector and its true direction defines an interesting new class of random processes, for which we have provided explicit analytical results.

When compared to empirical correlation matrices of several major stock markets, our results allow us to unambiguously conclude that there is a genuine evolution in time of the true underly-

ing correlation matrix: measurement noise in itself is unable to explain the observed variability (in time) of the top eigenspaces. We have found that the overlap distance is minimized when the measurement time is on the order of two to three years. Both for shorter and longer averaging times, measurement noise and the genuine evolution of the market leads to an instability of the correlation matrix, and to exposure to unwanted sources of risk.

The case of the top eigenvector of the correlation matrix, usually called the market mode, is particularly interesting. We have suggested a characterization of the evolution of its direction through a new correlation matrix, which measures the amplitude of its fluctuations transverse to its average direction. We found that the dominant modes are in the space spanned by the largest eigenvectors of the correlation matrix itself.

Now the genuine evolution of the correlation structure of stock returns is well characterized, one should aim at devising quantitative models for this evolution. As usual, there are two ways to do this. One is to postulate an econometric model and try to calibrate it on data. In this line of thought, extensions of the GARCH framework have been proposed: multivariate GARCH, BEKK model, etc.[28], but they often lack intuition (to say the least) and are very hard to calibrate (the a priori number of parameters is of order N^4!).

The second approach is to think about mechanisms which can lead to changes of the correlation structure. For example, market drops may lead to panic sell-offs, which increase the top eigenvalue of the correlation matrix and tilt the top eigenvector towards uniformity, as reported in [27, 8]. The impact of rebalancing or deleveraging complex portfolios can also lead to substantial changes in the correlation matrix – see the insightful work of Cont and Wagalath [60] in this direction. We hope that the tools provided in this paper will help building financially motivated, more efficient models of dynamical correlations and, correspondingly, second generation risk models where impact and feedback effects are accounted for [47].

7.8 Proof of the formula for $\delta = 0$

We need to introduce the two level density of states

$$\rho_2^N(\lambda, \lambda') = \frac{1}{N^2} \sum_{i,j=1}^{N} \delta(\lambda - \lambda_i, \lambda' - \lambda_j),$$

and to note from equation (7.9) that

$$D(V_0; V_1) = \frac{N\varepsilon^2}{2P} \int_a^b \int_{[-2;2]\setminus[a;b]} \frac{\rho_2^N(\lambda, \lambda')}{(\lambda - \lambda')^2}. \tag{7.55}$$

From [114], we know the asymptotic behavior of the two level density of states in the limit of large matrices; more precisely, there exists a function g such that, in the limit of large N,

$$\rho_2^N(\lambda, \lambda') = g(N\rho(\lambda)|\lambda - \lambda'|)\rho(\lambda)\rho(\lambda')d\lambda d\lambda' \tag{7.56}$$

which is defined as $g(r) = 1 - \left(\frac{1}{2} - \int_0^r s(t)dt\right) s'(r) + s(r)^2$ with $s(r) = \frac{\sin(\pi r)}{\pi r}$. One can check that:

- $g(r) \leqslant 1$ for all r,
- in the neighborhood of 0, $g(r) \sim \frac{\pi^2}{2} r$,
- $g(r)$ tends to 1 when r goes to ∞,
- $g'(r) = O(1/r^2)$ in the neighborhood of ∞.

We can write:

$$D(a, b; \delta = 0) = \frac{N\varepsilon^2}{2P} \int_a^b \int_{[-2;2]\setminus[a;b]} \frac{g(N\rho(x)|x - y|)}{(x - y)^2} \rho(x)\rho(y)dxdy.$$

We want to do an asymptotic expansion of the right hand side when $N \to \infty$.

First, note that N/P tends to $1/\int_a^b \rho$.

For the integral, we begin by doing an integration by part, we get for $x \in [-2; a]$:

$$\int_a^b \frac{g(N\rho(x)|x-y|)}{(x-y)^2}\,dy = \frac{\rho(a)g(N\rho(x)(a-x))}{a-x} - \frac{\rho(b)g(N\rho(x)(b-x))}{b-x} \tag{7.57}$$

$$+ \int_a^b \frac{dy}{y-x}\left[\rho'(y)g\left(N\rho(x)(y-x)\right) + N\rho(x)\rho(y)g'\left(N\rho(x)(y-x)\right)\right]. \tag{7.58}$$

We need to integrate equation (7.57) between -2 and a and between b and 2 and to compute the asymptotic of every integrals of the right hand side. We will decompose each integral into two terms so as to take advantages of the asymptotic property of g around 0 and ∞.

Set $\eta = N^{-1+\alpha}$ with $\alpha > 0$. First we consider the integral:

$$\int_{a-\eta}^{a} \rho(x)\frac{g(N\rho(x)(a-x))}{a-x}dx = \int_0^{N^\alpha} \frac{dx}{x}\rho(a-\frac{x}{N})g(\rho(a-\frac{x}{N})x)$$
$$\sim \rho(a)\int_0^{N^\alpha}\frac{dx}{x}g(\rho(a)x) = \rho(a)\left[\ln\left(\rho(a)N^\alpha\right) - \int_0^\infty \ln(x)g'(x)dx\right].$$

Using the fact that $g(r)$ tends to 1 when r goes to ∞, we easily get that, in the limit $N \to \infty$:

$$\int_{-2}^{a-\eta} \rho(x)\frac{g(N\rho(x)(a-x))}{a-x}dx$$
$$\sim \int_{-2}^{a-\delta}\frac{\rho(x)}{a-x}dx = -\rho(a)\ln(N^{-1+\alpha}) + \int_{-2}^{a}\rho'(x)\ln(a-x)dx\,.$$

Moreover, we easily find that, when N tends to ∞:

$$\int_{-2}^{a-\eta} \rho(x)\frac{g(N\rho(x)(b-x))}{b-x}dx \to \int_{-2}^{a}\frac{\rho(x)}{b-x}dx\,,$$

and

$$\int_{a-\eta}^{a} \rho(x)\frac{g(N\rho(x)(b-x))}{b-x}dx \leqslant \int_{a-\eta}^{a}\frac{\rho(x)}{b-x}dx\,,$$

which goes to 0 as η goes to 0.

The next term is easy to control using the fact that $g(r)$ goes to 1 when r goes to ∞; as $N \to \infty$:

$$\int_{-2}^{a-\eta} dx\rho(x)\int_a^b \frac{\rho'(y)dy}{y-x}g\left(N\rho(x)(y-x)\right) \to \int_{-2}^{a} dx\rho(x)\int_a^b\frac{\rho'(y)dy}{y-x}\,.$$

Using the fact that $g'(r)$ is of order $1/r^2$ for large r, it is easy to check that

$$N\int_{-2}^{a-\eta} dx\rho(x)\int_a^b \frac{dy}{y-x}\rho(y)g'\left(N\rho(x)(y-x)\right)$$

is of order $N^{-\alpha}$.

The remaining term

$$\int_{a-\eta}^{a} dx\rho(x)\int_a^b \frac{dy}{y-x}\left[\rho'(y)g\left(N\rho(x)(y-x)\right) + N\rho(x)\rho(y)g'\left(N\rho(x)(y-x)\right)\right]$$

is of order $N^{-1+\alpha}$.

One has to go through the same steps to compute the asymptotic of the integrals between b and 2.

Finally, we get:

$$D(a,b;\delta=0) \approx \ln N\ \varepsilon^2\ \frac{\rho(a)^2+\rho(b)^2}{2\int_a^b\rho(\lambda)d\lambda} + A(a,b) \tag{7.59}$$

where

$$
\begin{aligned}
A(a,b) = \frac{\varepsilon^2}{2\int_a^b \rho}\Bigg[& \left(\rho(a)^2+\rho(b)^2\right)\left(1-\int_0^\infty \ln(x)g'(x)dx\right) \\
& + \rho(a)^2\ln(\rho(a)) + \rho(b)^2\ln(\rho(b)) \\
& + \rho(a)\int_{-2}^{a}\rho'(x)\ln(a-x)dx - \rho(b)\int_b^2 \rho'(x)\ln(x-b)dx \\
& - \rho(b)\int_{-2}^{a}\frac{\rho(x)dx}{b-x} - \rho(a)\int_b^2\frac{\rho(x)dx}{x-a} \\
& + \int_{-2}^{a}dx\rho(x)\int_a^b\frac{\rho'(y)dy}{y-x} - \int_b^2 dx\rho(x)\int_a^b\frac{\rho'(y)dy}{x-y}\Bigg].
\end{aligned}
$$

7.9 Derivation of the standard deviation $\sigma(r)$

We have

$$
\begin{aligned}
\sigma^2(r) &= \langle r(s)^2\rangle - \langle r(s)\rangle^2 \\
&\approx \frac{1}{P}\langle \mathrm{tr}\left(\Sigma^2\right)\rangle - \langle\left(\frac{1}{P}\mathrm{tr}(\Sigma)\right)^2\rangle .
\end{aligned}
$$

But the two quantities are computable easily in the limit of large matrices using the convergence of the density of states for $\mathbf{H}_0$; We obtain

$$
\begin{aligned}
\frac{1}{P}\langle \mathrm{tr}\left(\Sigma^2\right)\rangle \approx & \frac{\varepsilon^4}{N_a^b}\int_{[a;b]}d\lambda\int_{[-2;2]\setminus[a-\delta;b+\delta]}d\lambda'\int_{[-2;2]\setminus[a-\delta;b+\delta]}d\lambda''\frac{\rho(\lambda)\rho(\lambda')\rho(\lambda'')}{(\lambda-\lambda')^2(\lambda-\lambda'')^2} \\
& + \frac{\varepsilon^4}{N_a^b}\int_{[a;b]}d\lambda\int_{[a;b]}d\lambda'\int_{[-2;2]\setminus[a-\delta;b+\delta]}d\lambda''\frac{\rho(\lambda)\rho(\lambda')\rho(\lambda'')}{(\lambda-\lambda'')^2(\lambda'-\lambda'')^2}
\end{aligned}
$$

and

$$
\langle\left(\frac{1}{P}\mathrm{tr}(\Sigma)\right)^2\rangle \approx \frac{\varepsilon^4}{N_a^{b\,2}}\int_{[a;b]}d\lambda\int_{[a;b]}d\lambda'\int_{[-2;2]\setminus[a-\delta;b+\delta]}d\lambda''\int_{[-2;2]\setminus[a-\delta;b+\delta]}d\lambda'''\frac{\rho(\lambda)\rho(\lambda')\rho(\lambda'')\rho(\lambda''')}{(\lambda-\lambda'')^2(\lambda'-\lambda''')^2}.
$$

Those two expressions give in the regime $\delta \ll \Delta \ll 1$

$$
\sigma(r) \approx \frac{\rho(a)}{\sqrt{\delta\Delta}}, \tag{7.60}
$$

and in the regime $\Delta \ll \delta \ll 1$,

$$
\sigma(r) \approx \rho(a)\sqrt{\frac{2\Delta}{3\delta^3}}. \tag{7.61}
$$

7.10 s_{min} in the strong fluctuation limit

It is given by (7.27) and we have to compute the $\widetilde{g} \in (-\infty;0)$ which verifies (7.25). For simplicity, we set $\hat{g} = -\widetilde{g}$ and we aim to compute $\hat{g} \geqslant 0$ such that

$$
\frac{1}{N_a^b}\int_{[-2;2]\setminus[a-\delta;b+\delta]}dx\frac{\rho(x)\sigma(x)^2}{(1+\sigma(x)\hat{g})^2} = \frac{1}{\hat{g}^2},
$$

As $\hat{g}$ is non-negative, the integral on the left hand side converges when δ goes to 0 and hence $\hat{g}$ verifies in fact

$$
\frac{1}{N_a^b}\int_{[-2;2]\setminus[a;b]}dx\frac{\rho(x)\sigma(x)^2}{(1+\sigma(x)\hat{g})^2} = \frac{1}{\hat{g}^2}, \tag{7.62}
$$

We now need to estimate the integral in the limit $\Delta \ll 1$. As before, we can write using (7.31)

$$\int_{-2}^{a} dx\, \frac{\rho(x)\sigma(x)^2}{(1+\sigma(x)\hat{g})^2} = \frac{\rho(a)^2}{\Delta} \int_0^{\frac{a+2}{\Delta}} du\, \frac{\rho(a-u\Delta) f^2(u)}{(u+\frac{\rho(a)\hat{g}}{\Delta} f(u))}\,.$$

In the limit $\Delta \ll 1$, this integral is dominated by the region where u is small and $f(u) \sim 1$ and hence we have the following estimate

$$\begin{aligned}\int_{-2}^{a} dx\, \frac{\rho(x)\sigma(x)^2}{(1+\sigma(x)\hat{g})^2} &\sim \frac{\rho(a)^3}{\Delta} \int_0^{+\infty} \frac{du}{(u+\frac{\rho(a)\hat{g}}{\Delta})^2} \\ &\sim \frac{\rho(a)^2}{\hat{g}}\,.\end{aligned}$$

Then we deduce from (7.62) and with the same argument for the integral between b and 2 that

$$\hat{g} = \frac{\Delta}{2\rho(a)}\,.$$

Now we have to plug this $\hat{g}$ into equation (7.27) to obtain

$$s_{min} = -\frac{2\rho(a)}{\Delta} + \frac{1}{N_a^b} \int_{[-2;2]\setminus[a;b]} dx\, \frac{\rho(x)\sigma(x)}{1+\sigma(x)\frac{\Delta}{2\rho(a)}}\,.$$

To evaluate the integral, we need to cut it into two parts. The first part is handled by

$$\begin{aligned}\int_{-2}^{a} \frac{\rho(x)\sigma(x)}{1+\sigma(x)\frac{\Delta}{2\rho(a)}} &= \rho(a) \int_0^{\frac{a+2}{\Delta}} du\, \frac{\rho(a-u\Delta) f(u)}{u+f(u)/2} \\ &\to \rho(a)^2 \int_0^{+\infty} du\, \frac{f(u)}{u+f(u)/2}\,.\end{aligned}$$

Finally, we can deduce that

$$s_{min} = \frac{2\rho(a)}{\Delta} \left(\int_0^{+\infty} du\, \frac{f(u)}{u+f(u)/2} - 1 \right)\,.$$

7.11 SDE for x_t

Using perturbation theory, one gets, in braket notation:

$$\begin{aligned}|\phi_1^t\rangle &= \left(1 - \frac{\varepsilon^2}{2} \sum_{i\neq 1} \frac{\langle\phi_1^{t-1}|r^t r^{t*}|\phi_i^{t-1}\rangle^2}{(\lambda_1^{t-1}-\lambda_i^{t-1})^2}\right)|\phi_1^{t-1}\rangle + \varepsilon \sum_{i\neq 1} \frac{\langle\phi_1^{t-1}|r^t r^{t*}|\phi_i^{t-1}\rangle}{\lambda_1^{t-1}-\lambda_i^{t-1}}|\phi_i^{t-1}\rangle \\ &\approx \left(1 - \frac{\varepsilon^2}{2(\lambda_1^{t-1})^2} \sum_{i\neq 1} \langle\phi_1^{t-1}|r^t r^{t*}|\phi_i^{t-1}\rangle^2\right)|\phi_1^{t-1}\rangle + \frac{\varepsilon}{\lambda_1^{t-1}} \sum_{i\neq 1} \langle\phi_1^{t-1}|r^t r^{t*}|\phi_i^{t-1}\rangle|\phi_i^{t-1}\rangle \\ &= \left(1 - \frac{\varepsilon^2}{2(\lambda_1^{t-1})^2} \left(\langle\phi_1^{t-1}|(r^t r^{t*})^2|\phi_1^{t-1}\rangle - \langle\phi_1^{t-1}|r^t r^{t*}|\phi_1^{t-1}\rangle^2\right)\right)|\phi_1^{t-1}\rangle \\ &+ \frac{\varepsilon}{\lambda_1^{t-1}} \left(r^t r^{t*}|\phi_1^{t-1}\rangle - \langle\phi_1^{t-1}|r^t r^{t*}|\phi_1^{t-1}\rangle|\phi_1^{t-1}\rangle\right)\,.\end{aligned}$$

Since $\cos(\theta_t) = \langle\phi_1^t|\phi_1\rangle$, we can write

$$\phi_1^t = \cos(\theta_t)|\phi_1\rangle + \sin(\theta_t)|\varphi_\perp^t\rangle$$

where $|\varphi_\perp\rangle$ is a vector lying in the subspace spanned by the vectors $|\phi_2\rangle, \ldots, |\phi_N\rangle$. We want to describe the dynamic of $\cos(\theta_t)$; we deduce from the previous equation that

$$\cos(\theta_t) = \left(1 - \frac{\varepsilon^2}{2(\lambda_1^{t-1})^2} \left(\langle\phi_1^{t-1}|(r^t r^{t*})^2|\phi_1^{t-1}\rangle - \langle\phi_1^{t-1}|r^t r^{t*}|\phi_1^{t-1}\rangle^2\right)\right)\cos(\theta_{t-1}) \tag{7.63}$$

$$+ \frac{\varepsilon}{\lambda_1^{t-1}} \left(\langle\phi_1|r^t r^{t*}|\phi_1^{t-1}\rangle - \langle\phi_1^{t-1}|r^t r^{t*}|\phi_1^{t-1}\rangle \cos(\theta_{t-1})\right)\,. \tag{7.64}$$

Since we have:

$$\begin{aligned}
\langle\phi_1|\mathbf{C}|\phi_1^{t-1}\rangle &= \lambda_1\cos(\theta_{t-1})\,,\\
\langle\phi_1^{t-1}|\mathbf{C}|\phi_1^{t-1}\rangle &= \lambda_1\cos^2(\theta_{t-1}) + \lambda_2\sin^2(\theta_{t-1})\,,\\
\overline{\langle\phi_1|\eta_t|\phi_1^{t-1}\rangle^2} &= 2\cos^2(\theta_{t-1})\lambda_1^2 + \sin^2(\theta_{t-1})\lambda_1\lambda_2\,,\\
\overline{\langle\phi_1^{t-1}|\eta_t|\phi_1^{t-1}\rangle^2} &= 2\left(\lambda_1\cos^2(\theta_{t-1}) + \lambda_2\sin^2(\theta_{t-1})\right)^2\,,\\
\overline{\langle\phi_1^{t-1}|r^t r^{t*}|\phi_1^{t-1}\rangle^2} &= \lambda_1^2\cos^4(\theta_{t-1}) + 2\cos^2(\theta_{t-1})\lambda_1^2 + \sin^2(\theta_{t-1})\lambda_1\lambda_2\,,\\
\overline{\langle\phi_1^{t-1}|(r^t r^{t*})^2|\phi_1^{t-1}\rangle} &= \cos^2(\theta_{t-1})(3\lambda_1^2 + (N-1)\lambda_1\lambda_2) + \sin^2(\theta_{t-1})((N+1)\lambda_2^2 + \lambda_1\lambda_2)\,,
\end{aligned}$$

equation (7.63) can be rewritten, in the asymptotic regime where $\varepsilon \ll 1, N \gg 1$ with $q = \varepsilon N$ fixed and $\lambda_2 \ll \lambda_1$, keeping up to terms of order 2 for "drift" terms and of order 1 for noise terms in ε and λ_2/λ_1[13]:

$$\begin{aligned}
d(\cos(\theta_t)) = &-\frac{\varepsilon^2}{2}\frac{1}{\lambda_1^2}\left[(\lambda_1^2 + N\lambda_1\lambda_2)\cos^2(\theta_t) - \lambda_1^2\cos^4(\theta_t)\right]\cos(\theta_t)dt\\
&+ \varepsilon\cos(\theta_t)\sin^2(\theta_t)dt + \sigma_t dB_t
\end{aligned}$$

where

$$\overline{\sigma_t^2} = \frac{\varepsilon^2}{\lambda_1^2}\left[2\lambda_1^2\cos^2(\theta_t)\sin^2(\theta_t) + \lambda_1\lambda_2\cos^2(2\theta_t)\right]\sin^2(\theta_t). \tag{7.65}$$

When $\theta_t \ll 1$, this leads to Eq. 7.48 given in the main text for $x_t = 1 - \cos(\theta_t)$.

7.12 Transition probability of x_t

In this appendix, we show that the function $P(x,t)$ giving the probability that the "particle" x_t verifying (7.48) is in x at time t can be computed explicitly. More generally, we will show that one can compute explicitly this transition density $P(x,t)$ for a process x_t with initial condition in $t = 0$ given by $x_0 \geqslant 0$ verifying the Langevin equation

$$\mathrm{d}x_t = \theta(\mu - x_t)\mathrm{d}t + \sigma\sqrt{x_t(x_t+b)}\mathrm{d}B_t \tag{7.66}$$

where θ, μ, σ and b are positive constants and B_t a standard Brownian motion. One can proceed to the change of variables

$$y_t = \cosh^{-1}\left(\frac{2}{b}x_t + 1\right) \quad \Leftrightarrow \quad x_t = \frac{b}{2}\left(\cosh(y_t) - 1\right)\,,$$

and find that the process y_t verifies

$$\mathrm{d}y_t = \left(\theta(1 + \frac{2\mu}{b})\frac{1}{\sinh(y_t)} - (\theta + \frac{\sigma^2}{2})\frac{\cosh(y_t)}{\sinh(y_t)}\right)\mathrm{d}t + \sigma\,\mathrm{d}B_t\,. \tag{7.67}$$

We will denote by $F(y)$ the drift coefficient of the previous stochastic differential equation (7.67) and denote by U its potential, which verifies $U' = -F$. The transition density $\bar{P}(y,t)$ verifies the Fokker-Planck equation

$$\frac{\partial\bar{P}}{\partial t} = -\frac{\partial(F\bar{P})}{\partial y} + \frac{\sigma^2}{2}\frac{\partial^2\bar{P}}{\partial y^2}\,.$$

By setting $\bar{P}(y,t) := e^{-U(y)/\sigma^2}\psi(y,t)$, this equation becomes a Schrodinger equation:

$$\frac{\partial\psi}{\partial t} = \frac{\sigma^2}{2}\frac{\partial^2\psi}{\partial y^2} - V(y)\psi\,,$$

[13] Note that $\sin^2(\theta_t) \approx 2\mu$ is of order λ_2/λ_1 and that $1 - \cos(\theta_t) \approx \mu$ is also of order λ_2/λ_1.

with the so-called Pöschl-Teller potential $V(y)$:

$$\begin{aligned} V(y) &= \frac{1}{2}\left(\frac{F^2(y)}{\sigma^2} + F'(y)\right) \\ &= \frac{1}{2}\left(\frac{\alpha}{\sinh^2(y)} - \frac{\beta\,\cosh(y)}{\sinh^2(y)} + \gamma\right) \end{aligned}$$

with:

$$\begin{aligned} \alpha &= \left(\theta + \frac{\sigma^2}{2}\right)\left(\frac{3}{2} + \frac{\theta}{\sigma^2}\right) + \frac{\theta^2}{\sigma^2}\left(1 + \frac{2\mu}{b}\right)^2 , \\ \beta &= 2\theta\left(1 + \frac{2\mu}{b}\right)\left(1 + \frac{\theta}{\sigma^2}\right) , \\ \gamma &= \frac{1}{\sigma^2}\left(\theta + \frac{\sigma^2}{2}\right)^2 . \end{aligned}$$

Since the evolution of $\psi(y,t)$ is governed by a self adjoint operator

$$\mathcal{H} := \frac{\sigma^2}{2}\frac{\partial^2 \cdot}{\partial y^2} - V(y)\cdot$$

we can use its eigenfunctions to construct an orthonormal basis (ψ_n) with corresponding eigenvalues $(-\lambda_n)$. The general solution $\psi(y,t)$ can thus be expanded in the following form

$$\psi(y,t) = \sum_n c_n \psi_n(y) e^{-\lambda_n t} .$$

The general solution for $\bar{P}$ is thus given by

$$\bar{P}(y,t) = e^{-U(y)/\sigma^2} \sum_n c_n \psi_n(y) e^{-\lambda_n t} .$$

The initial conditions for y_t determines the sequence (c_n). In particular, if at time $t = 0$, the probability $\bar{P}(y,0) = \delta(y - y_0)$ with $y_0 := \arg\cosh(\frac{2}{b}x_0 + 1)$, then it is straightforward to see that

$$c_n = e^{U(y_0)/\sigma^2}\psi_n(y_0) .$$

The spectrum of $\mathcal{H}$ consists of a discrete and a continuous branch. The discrete energy levels (eigenvalues) are computed in, e.g. [64] and are given for all $n \in \mathbb{N}, n \leqslant g/2$ with $g = 1 + 2\theta\sigma^2$, by

$$\lambda_n = \frac{\sigma^2}{2} n\,(g - n) . \tag{7.68}$$

The corresponding eigenvectors are also computed in [64] and are expressed in terms of Jacobi polynomials. To the best of our knowledge, the continuous branch of the spectrum has not been fully characterized in the literature. We should also mention that in the limit $b \to 0$ the corresponding process has been studied in details (see [136] and the appendix of [115]). The problem can now be mapped into the Morse potential, which has exactly the same discrete spectrum as above (as expected since b does not appear), with eigenfunctions that can be expressed in terms of Laguerre polynomials. However, we have not been able to directly match the eigenfunctions in the two cases, and understand the $b \to 0$ limit in details. The limit $b \to \infty$ with $\sigma^2 b$ fixed, on the other hand, boils down to the standard Bessel process with mean-reversion.

Chapter 8

Marčenko Pastur theorem for independent MRW

Résumé

Cet article est soumis au journal *Annals of Applied Probability* et est écrit en collaboration avec Rémi Rhodes et Vincent Vargas. Nous prouvons la convergence de la densité empirique des valeurs propres d'une matrice de covariance empirique construite à partir des accroissements de marches aléatoires multifractales indépendantes. Nous caractérisons la mesure de probabilité limite par sa transformée de Stieltjes, qui vérifie une équation explicite admettant une unique solution. Nous illustrons nos résultats numériquement.

Abstract

We study the asymptotic of the spectral distribution for large empirical covariance matrices composed of independent 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.

8.1 Introduction

Since the seminal work of Marčenko and Pastur [113] 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 $\mathbf{R}_{ij} = \mathbb{E}[r_i r_j]$):

$$\mathcal{R}_w = w^t \mathbf{R} w \,.$$

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 [44], [124] 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 $i = 1, \dots, N$ that evolve continuously with respect to time $t \in [0; 1]$ but we observe those prices *only on a discrete finite grid* $\{j/T, j = 1, \dots, T\}$ where T is the number of observations. Using

this discrete grid, we can compute the price 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:

$$r_i(j) := X_i(\frac{j}{T}) - X_i(\frac{j-1}{T}).$$

Then, we define the $N \times T$ matrix X_N such that $X_N(ij) = r_i(j)$ 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 $N \to \infty$) for different models of the *i.i.d.* random continuous processes $(X_i(t)), i \in \{1, \dots, N\}$ (see below for precise definitions). For this purpose, the Marčenko-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:

$$X_N(ij) = B_i\left(\frac{j}{T}\right) - B_i\left(\frac{j-1}{T}\right) \tag{8.1}$$

where the B_i are *i.i.d.* standard Brownian motions.

If $\lambda_1, \dots, \lambda_N$ are the eigenvalues of R_N, the empirical spectral distribution of the matrix R_N is the probability measure defined by:

$$\mu_{R_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}. \tag{8.2}$$

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

$$\rho(x) = \frac{1}{2\pi q}\frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{x}\mathbb{1}_{[\gamma-,\gamma+]}dx \tag{8.3}$$

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 [45, 60] 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 [**?**] (see also [21, **?**]). 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 [48]). The lognormal MRW is simply defined as:

$$X(t) = B\left(M[0,t]\right) \tag{8.4}$$

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:

$$M[0;t] = \int_0^t e^{\omega(x) - \frac{1}{2}\mathbb{E}[\omega(x)^2]} \mathrm{d}x \,,$$

where $(\omega(x))_{x\in\mathbb{R}}$ is a "gaussian field" whose covariance kernel K is

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

where $\ln_+ x = \max(\ln x, 0)$. The two parameters γ^2 and τ are respectively called intermittency parameter and integral scale (or correlation length) of the lognormal random multifractal measure M.

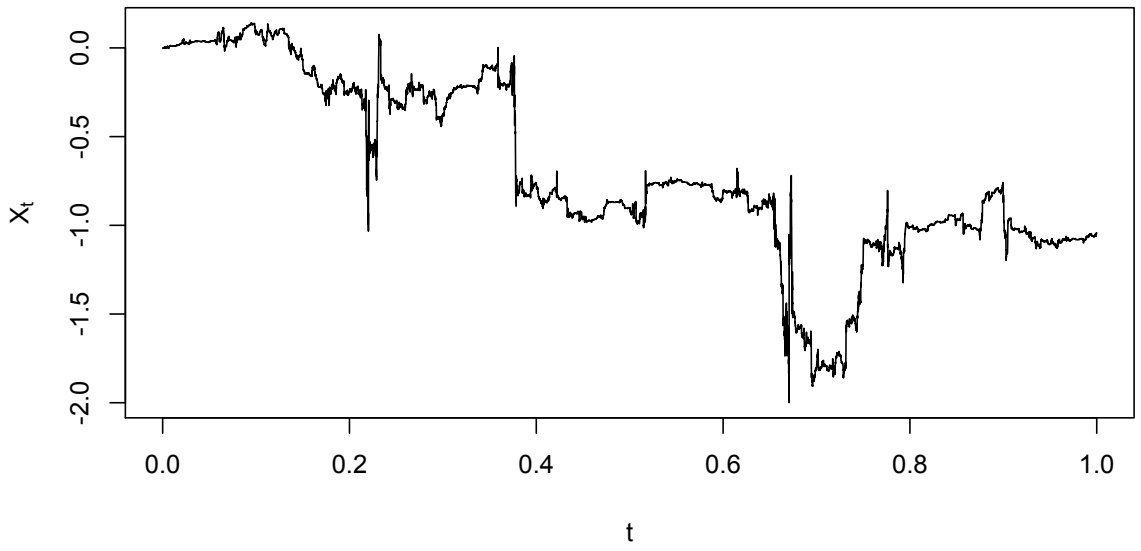


Figure 8.1: Simulated path of a multifractal random walk with intermittency parameter $\gamma^2 = 1$ and with integral scale $\tau = 1/4$. Note the intermittent bursts in volatility.

Fig. 8.1 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 8.2.1.

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:

$$X_N(ij) = B^i(M^i(0, \frac{j}{T})) - B^i(M^i(0, \frac{j-1}{T})) \tag{8.5}$$

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 [100] which considers high frequency covariance matrices in the context of diffusion processes (see also [128] for studies of high frequency large empirical covariance matrices motivated by financial applications). The processes described by (8.5) 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 μ_{R_N} when $N, T \to \infty$ with $N/T \to q \in (0, 1]$. 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 $[\gamma_-;\gamma_+]$ in the Marčenko-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 τ 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 τ while low frequency case corresponds to small values of τ. Indeed, if X is a lognormal MRW with integral scale τ, then the process $\widehat{X}(t)$ defined on $[0;1]$ as $\widehat{X}(t) = X(t/2)$ is a lognormal MRW with integral scale 2τ. 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 is the right tail.

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 $N \to \infty$ 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 [33]. The main difficulty lies in handling the Stieltjes transforms in a multifractal setting.

8.2 Background, notations and main results

8.2.1 Reminder of the construction of MRM

To fix precisely the notations that we will use throughout the paper, we quickly remind the main steps of the construction of Multifractal Random Measures (MRM). The description is necessarily concise and the reader is referred to [24] for further details. In particular, we use the same notations as in [24] to facilitate the reading. We consider the characteristic function of an infinitely divisible random variable Z, which can be written as $\mathbb{E}[e^{ipZ}] = e^{\varphi(p)}$ where (Lévy-Khintchine's formula):

$$\varphi(p) = imp - \frac{1}{2}\gamma^2 p^2 + \int_{\mathbb{R}^*} (e^{ipx} - 1)\,\nu(dx) \tag{8.6}$$

and $\nu(dx)$ is a so-called Lévy measure (ie satisfying $\int_{\mathbb{R}^*} \min(1, x^2)\,\nu(dx) < +\infty$) together with the following additional assumption:

$$\int_{[-1,1]} |x|\,\nu(dx) < +\infty, \tag{8.7}$$

so that its characteristic function perfectly makes sense as written in (8.6). We also introduce the Laplace exponent ψ of Z by $\psi(p) = \varphi(-ip)$ for each p such that both terms of the equality make sense, and we assume that the following renormalization condition holds: $\psi(1) = 0$.

We further consider the half-space $S = \{(t, y); t \in \mathbb{R}, y \in \mathbb{R}_+^*\}$, with which we associate the measure (on the Borel σ-algebra $\mathcal{B}(S)$):

$$\theta(dt, dy) = y^{-2} dt\, dy. \tag{8.8}$$

Then we consider an independently scattered infinitely divisible random measure μ associated to (φ, θ) and distributed on S.

Then we define a process ω_ϵ for $\epsilon > 0$ by the following. Given a positive parameter τ, let us define the function $f : \mathbb{R}_+ \to \mathbb{R}$ by:

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

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

$$A_\epsilon(t) = \{(s, y) \in S; y \geqslant \epsilon, -f(y)/2 \leqslant s - t \leqslant f(y)/2\}. \tag{8.9}$$

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

$$\omega_\epsilon(t) = \mu\left(A_\epsilon(t)\right). \tag{8.10}$$

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

$$M(A) = \lim_{\epsilon\to 0^+} M_\epsilon(A) = \lim_{\epsilon\to 0^+} \int_A e^{\omega_\epsilon(r)}\, dr$$

for any Lebesgue measurable subset $A \subset \mathbb{R}$. The convergence is ensured by the fact that the family $(M_\epsilon(A))_{\epsilon>0}$ is a right-continuous positive martingale. The structure exponent of M is defined by:

$$\forall p \geqslant 0, \quad \zeta(p) = p - \psi(p)$$

for all p such that the right-hand side makes sense. The measure M is different from 0 if and only if there exists $\epsilon > 0$ such that $\zeta(1 + \epsilon) > 1$, (or equivalently $\psi'(1) < 1$). In that case, we have:

Theorem 8.1. *The measure M is stationary and satisfies the* **exact stochastic scale invariance property***: for any $\lambda \in]0, 1]$,*

$$(M(\lambda A))_{A\subset B(0,\tau)} \overset{\text{law}}{=} (\lambda e^{\Omega_\lambda} M(A))_{A\subset B(0,\tau)},$$

where Ω_λ is an infinitely divisible random variable, independent of $(M(A))_{A\subset B(0,T)}$, the law of which is characterized by:

$$\mathbb{E}[e^{ip\Omega_\lambda}] = \lambda^{-\varphi(p)}.$$

8.2.2 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 $R_N := X_N {}^tX_N$ as $N \to \infty$, where X_N is a $N \times T$ real matrix the entries of which are given by (8.5). 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:

$$\lim_{N\to\infty} \frac{N}{T} = q. \tag{8.11}$$

We further set $\widetilde{R}_N := {}^tX_N X_N$, and if M is a symmetric real matrix, we will denote by μ_M the empirical spectral measure of M.

Define the $(T+N)\times(T+N)$ matrix B_N by:

$$B_N = \begin{pmatrix} 0 & {}^tX_N \\ X_N & 0 \end{pmatrix}.$$

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

$$A_N(z) = (zI_{T+N} - B_N) = \begin{pmatrix} zI_T & -{}^tX_N \\ -X_N & zI_N \end{pmatrix}.$$

Notice that

$$B_N^2 = \begin{pmatrix} \widetilde{R}_N & 0 \\ 0 & R_N \end{pmatrix}$$

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

$$\mu_{B_N^2} = 2\frac{N}{N+T}\mu_{R_N} + \frac{T-N}{N+T}\delta_0, \tag{8.12}$$

where δ_x stands for the Dirac mass at x. Combining this equality with the relation

$$\int f(x)\mu_{B_N^2}(dx) = \int f(x^2)\mu_{B_N}(dx) \tag{8.13}$$

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 μ_{R_N}.

We will thus work on the (weak) convergence of the spectral measures μ_{B_N} and $\mathbb{E}[\mu_{B_N}]$ 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 μ on $\mathbb{R}$, the Stieltjes transform G_μ of μ is defined, for all $z\in\mathbb{C}\setminus\mathbb{R}$, as:

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z-x}\mu(dx). \tag{8.14}$$

and one can note that:

$$G_{\mu_{B_N}}(z) = \frac{1}{N+T}\mathrm{Trace}(G_N(z)), \tag{8.15}$$

where we have set:

$$G_N(z) = (A_N(z))^{-1}. \tag{8.16}$$

Hence, we have to investigate the convergence of the right-hand side of (8.15). 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 $f:[0,1]\to\mathbb{R}$:

$$L_N^{1,z}(f) = \frac{1}{T}\sum_{k=1}^{T} f\left(\frac{k}{T}\right) G_N(z)_{kk}$$
$$L_N^{2,z}(f) = \frac{1}{N}\sum_{k=1}^{N} f\left(\frac{k}{N}\right) G_N(z)_{k+T,k+T}$$

Clearly, we have the relation

$$\frac{1}{N+T}\mathrm{Trace}(G_N(z)) = \frac{T}{N+T}L_N^{1,z}([0,1]) + \frac{N}{N+T}L_N^{2,z}([0,1]) \tag{8.17}$$

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

8.3 Main results

8.3.1 Lognormal multifractal random walk

We first present our results when the process $X(t)$ is a lognormal multifractal random walk, i.e. $X(t) = B(M[0;t])$ where M is the MRM whose characteristic and structure exponent (see section 8.2.1) are respectively given by:

$$\varphi(q) = -i\frac{\gamma^2}{2}q - \frac{\gamma^2}{2}q^2,$$
$$\zeta(q) = (1+\frac{\gamma^2}{2})q - \frac{\gamma^2}{2}q^2.$$

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

Assumption 8.2. *More precisely, let us suppose that:*

$$\gamma^2 < \frac{1}{3}. \tag{8.18}$$

Though we conjecture that our results hold as soon as the measure M is non degenerated, i.e. $\gamma^2 < 2$ (see [**?**]), Assumption 8.2 is largely sufficient to cover most practical applications. For instance, in financial applications or in the field of turbulence, γ^2 is found empirically around 2.10^{-2}.

We can now state our main result about the convergence of the empirical spectral measures and mean empirical spectral measures of the matrices B_N and R_N:

Theorem 8.3. *i) There exists a probability measure υ on $\mathbb{R}$ such that the two mean spectral measures $\mathbb{E}[\mu_{B_N}]$ and $\mathbb{E}[\mu_{R_N}]$ converge weakly respectively towards the two probability measures $\frac{2q}{1+q}\upsilon + \frac{1-q}{1+q}\delta_0$ and $\upsilon \circ (x^2)^{-1}$ as N goes to ∞, where $\upsilon \circ (x^2)^{-1}$ is the push-forward of the measure υ by the mapping $x \mapsto x^2$.*

ii) The two spectral measures μ_{B_N} and μ_{R_N} converge weakly in probability respectively to the two probability measures $\frac{2q}{1+q}\upsilon + \frac{1-q}{1+q}\delta_0$ and $\upsilon \circ (x^2)^{-1}$ as N goes to ∞. More precisely, for any bounded and continuous function f, $\int f(x)\mu_{R_N}(dx)$ converges in probability to $\int f(x)\upsilon \circ (x^2)^{-1}(dx)$.

iii) Let N_k be an increasing sequence of integers such that $\sum_{k=1}^{\infty} N_k^{-1} < +\infty$, then the two sequences $\mu_{B_{N_k}}$ and $\mu_{R_{N_k}}$ converge weakly almost surely to the two probability measures $\frac{2q}{1+q}\upsilon + \frac{1-q}{1+q}\delta_0$ and $\upsilon \circ (x^2)^{-1}$ as k goes to ∞.

Theorem 8.3 is implied by (8.15), (8.17) and by Theorem 8.4:

Theorem 8.4. *i) The measures $\mathbb{E}[L_N^{1,z}]$ and $\mathbb{E}[L_N^{2,z}]$ 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:*

$$\mathbb{E}\left[L_N^{1,z}(f)\right] \to_{N\to\infty} \int_0^1 K_z(x)f(x)\,dx,$$
$$\mathbb{E}\left[L_N^{2,z}(f)\right] \to_{N\to\infty} \mu_z^2 \int_0^1 f(x)\,dx.$$

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

$$\int_0^1 K_z(x)\,dx = q\mu_z^2 + \frac{1-q}{z} \tag{8.19}$$

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

$$\mu_z^2 = \int_{\mathbb{R}} \frac{\upsilon(dx)}{z-x}. \tag{8.20}$$

It is important to state a characterization of the probability measure υ: it is done by means of its Stieltjes transform μ_z^2:

Theorem 8.5. *The constant μ_z^2 and the bounded function $K_z(x)$ are uniquely determined for all $z \in \mathbb{C} \setminus \mathbb{R}$, by the following system of equations:*

$$\mu_z^2 = \mathbb{E}\left[\left(z - \int_0^1 K_z(t)M(dt)\right)^{-1}\right], \tag{8.21}$$

$$K_z(x) = \left(z - q\mathbb{E}\left[\left(z - \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} K_z(t)M(dt)\right)^{-1}\right]\right)^{-1} \tag{8.22}$$

where[1] *M is the MRM with structure exponent $\zeta(q) = (1+\gamma^2/2)q - q^2\gamma^2/2$.*

Let us notice that one can give a precise meaning to (8.22) 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:

$$\int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} f(t)M(dt) = \lim_{\eta\to 0} \int_{t\in[0,1];|t-x|>\eta} \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} f(t)M(dt) \tag{8.23}$$

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

$$\frac{\ln M[x-\epsilon_k, x+\epsilon_k]}{\ln \epsilon_k} \underset{k\to\infty}{\to} 1 + \frac{\gamma^2}{2}, \ a.s.$$

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

$$\int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} f(t)M(dt) = \lim_{\epsilon\to 0} \int_0^1 e^{cov(\omega_\epsilon(t),\omega_\epsilon(x))} f(t)e^{\omega_\epsilon(t)}dt$$

Conjecture 8.6. *With this extended definition, we conjecture that theorem 8.5 holds in the lognormal multifractal case for all $\gamma^2 \in [0,2[$ and thus that the limiting equations can be obtained by the ones of theorem 8.10 (see below) with $2W = \omega_\epsilon$ as $\epsilon \to 0$.*

8.3.2 General multifractal random walk

We now look at the more general case when the change of time is a measure M for which the function $\varphi(q)$ is given by (8.6) and the structure exponent by $\zeta(q) = q - \psi(q)$ with $\psi(q) = \varphi(-iq)$.

We still have to make an assumption to avoid the issue of strong correlations. In this more general setting, Assumption (8.2) becomes:

Assumption 8.7. *Assume that the structure exponent of the MRM satisfies the condition:*

$$\zeta(2) > 5 - 4\zeta'(1). \tag{8.24}$$

and that there exists $\delta > 0$ such that:

$$\zeta(2+\delta) > 1. \tag{8.25}$$

As in the previous section, we conjecture that our results hold as soon as the measure M is non degenerated, i.e. (see [**?**]) $\zeta(1+\epsilon) > 1$ for some $\epsilon > 0$.

Theorems 8.3 and 8.4 remain unchanged for this more general context. Theorem 8.5 becomes:

Theorem 8.8. *The constant μ_z^2 and the bounded function $K_z(x)$ are uniquely determined for all $z \in \mathbb{C} \setminus \mathbb{R}$, by the following system of equations:*

$$\mu_z^2 = \mathbb{E}\left[\left(z - \int_0^1 K_z(t)M(dt)\right)^{-1}\right], \tag{8.26}$$

$$K_z(x) = \left(z - q\mathbb{E}\left[\left(z - \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\kappa} K_z(t)Q(dt)\right)^{-1}\right]\right)^{-1} \tag{8.27}$$

[1] The notation $(\cdot)_+$ is a shortcut for $\max(\cdot, 1)$.

with $\kappa = \psi(2)$ and where M is the MRM whose characteristic and structure exponent are respectively $\varphi(q), \zeta(q)$ and where the random Radon measure Q is defined, conditionally on M, as the almost sure weak limit as ϵ goes to 0 of the family of random measures $Q_\epsilon(dt) := e^{\overline{\omega}_\epsilon(t)} M(dt)$ where, for each $\epsilon > 0$, the random process $\overline{\omega}_\epsilon$ is independent of M and defined as $\overline{\omega}_\epsilon(t) = \overline{\mu}(A_\epsilon(t))$ where $\overline{\mu}$ is the independently scattered log infinitely divisible random measure associated to $(\bar{\varphi}, \theta(\cdot \cap A_0(x)))$ with:

$$\bar{\varphi}(p) = ip(\gamma^2 - \kappa) + \int_{\mathbb{R}} (e^{ipx} - 1)(e^x - 1)\nu(dx). \tag{8.28}$$

8.3.3 Lognormal random walk

Let us mention that one can easily adapt the methods used to prove the above theorems in the simpler case (lognormal case) where $X(t)$ is defined, for all $t \in [0;1]$, by:

$$X(t) = B\left(\int_0^t e^{2W(s)} ds\right), \tag{8.29}$$

where $(W(s))_{s\in[0;1]}$ is a stationary gaussian process with expectation m and stationary covariance kernel k. The normalization will be chosen such that: $m = -k(0)$.

In this context, the entries of X_N are given, for $1 \leqslant i \leqslant N, 1 \leqslant j \leqslant T$ by:

$$X_N(ij) = \frac{1}{\sqrt{T}} e^{W_i(\frac{j}{T})} B_j^i := r_i(j) \tag{8.30}$$

where the $(B_j^i)_{ij}$ are i.i.d standard centered Gaussian random variables and the W_i are i.i.d stationary Gaussian processes with expectation m and stationary covariance kernel k. Indeed, if one makes the following extra assumption:

Assumption 8.9. *Assume that for some constants $C > 0$ and $\beta > 0$, the covariance kernel k satisfies:*

$$\forall x \in \mathbb{R}, \quad |k(x) - k(0)| \leqslant C|x|^\beta.$$

With the same notations as in the previous section, we can now state the following theorem under assumption 8.9:

Theorem 8.10. *The system of equations for μ_z^2 and $K_z(x)$ becomes:*

$$\mu_z^2 = \mathbb{E}\left[\left(z - \int_0^1 K_z(t) e^{2W(t)}\, dt\right)^{-1}\right] \tag{8.31}$$

$$K_z(x) = \left(z - q\mathbb{E}\left[\left(z - \int_0^1 K_z(t) e^{4k(t-x)} e^{2W(t)}\, dt\right)^{-1}\right]\right)^{-1}. \tag{8.32}$$

where $(W(t))_{t\in[0;1]}$ is a stationary gaussian process with expectation m and stationary covariance kernel k.

8.4 Numerical results and computer simulations

In this section, we are interested in the case handled in sub-section 8.3.1, in which the price of an asset evolves as a lognormal multifractal random walk. We want to extract informations on the spectral density $\upsilon \circ (x^2)^{-1}$ 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 υ is entirely contained in its Stieltjes transform μ_z^2 which is the unique solution of the system of equations (8.21) and (8.22). Let us admit for clarity at this point that the measure υ admits a continuous density, at least on the set $\mathbb{R} \setminus \{0\}$. One should be able to show that this is indeed true using the two equations (8.21) and (8.22) that characterize the probability measure υ. Under this continuity assumption for $\upsilon(x)$, we can re-find the density $\upsilon(x)$ from μ_z^2 by the relation

$$\lim_{\epsilon\to 0} \frac{1}{\pi} \Im(\mu_{x-i\epsilon}^2) = \upsilon(x)\,. \tag{8.33}$$

Note that we just need to find the unique family of functions $(K_z(x))_{x\in[0;1]}$ for $z \in \mathbb{C}\setminus\mathbb{R}$ near the real line, that verifies the fixed point equation (8.22). Indeed, knowing $(K_z(x))_{x\in[0;1]}$, we can compute μ_z^2 by using equation (8.21), or even simpler, the additional relation that we stated above

$$\int_0^1 K_z(x)\,dx = q\mu_z^2 + \frac{1-q}{z}\,. \tag{8.34}$$

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

$$Tg(x) = \frac{1}{z - q\mathbb{E}\left[\left(z - \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} g(t)M(dt)\right)^{-1}\right]}\,. \tag{8.35}$$

It can easily be shown (see sub-section 8.5.6) that if $z \in \mathbb{C}\setminus\mathbb{R}$ 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 $K_z^{(n)} := T(K_z^{(n-1)})$ 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[2] and find the fixed point after a reasonable number of iterations of the operator T.

The multifractal lognormal random measure $M(dt)$ 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. [**?**, **?**]).

The results are as follows. In Fig. 8.2, we show the comparison between the theoretical value of the density $\upsilon \circ (x^2)^{-1}(x)$ (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 upward plot is done with an intermittency parameter $\gamma^2 = 1/4$ and an integral scale $\tau = 1/4$. The agreement is excellent as expected from Theorems 8.3, 8.4 and 8.5. The downward 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 8.6 which also covers the case $\gamma^2 \in [1,2[$).

In Fig. 8.3, we represent three curves (axis are in log-log) corresponding to the theoretical density $\upsilon \circ (x^2)^{-1}(x)$ for a parameter $q = 1$, an integral scale $\tau = 1/4$ and for three different values of γ^2. The black dashed curve corresponds to $\gamma^2 = 0$, which in fact is the Marcenko-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 γ^2 is, the heavier the tail of the distribution is.

In Fig. 8.4, we represent four curves corresponding to the thoeretical density $\upsilon \circ (x^2)^{-1}(x)$ but varying the integral scale τ instead of the intermittency parameter γ^2. We chose for this plot $q = 1$ and $\gamma^2 = 1/4$ and represented the density $\upsilon \circ (x^2)^{-1}(x)$ 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.

[2] Recall that, in view of equation (8.33), we are interested in the value of the Stieltjes transform near the real line.

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.

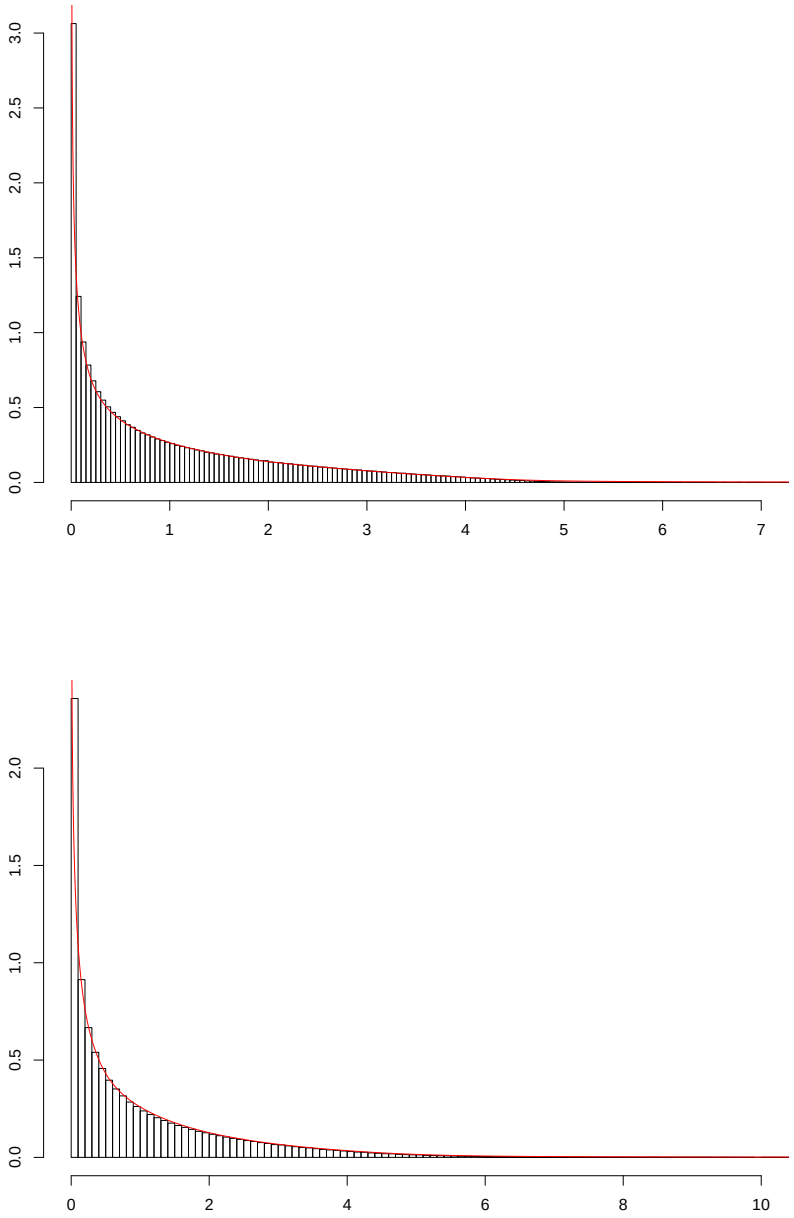


Figure 8.2: Comparison between the theoretical value of the density $\upsilon \circ (x^2)^{-1}(x)$ and the empirical histogram computed through a sample of simulated empirical covariance matrices R_N as defined in the introduction. For both plots, $q = 1$ but stock prices follow multifractal random walks with intermittency parameter $\gamma^2 = 1/4$ in the upward figure, $\gamma^2 = 1/2$ in the downward figure.

8.5 Proofs of the main results

In this section, we give the proofs of theorems 8.3, 8.4 and 8.5. The proof of Theorem 8.8 is very similar and we will not explain it in every detail, except for the final part where we establish the second equation of the system in Theorem 8.8 verified by K_z. We will give the details for this part of the proof in the appendix. The proof of theorem 8.10 is an easy adaptation of our

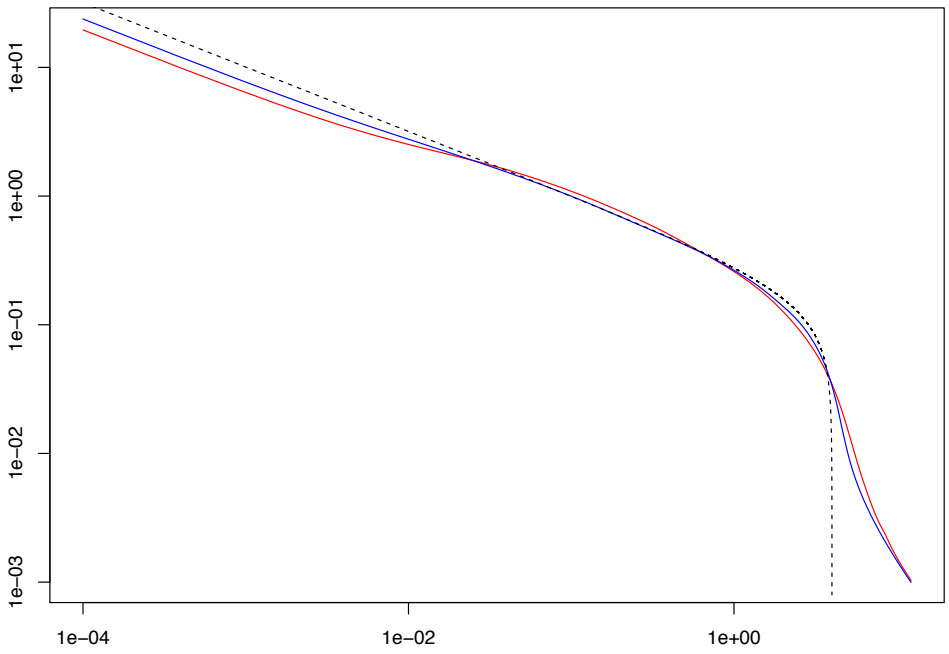


Figure 8.3: Log-log plot of the density $\upsilon \circ (x^2)^{-1}$ with $q = 1$, $\tau = 1/4$ for three different intermittency parameter: $\gamma^2 = 0$ (black dashed line), $\gamma^2 = 1/4$ (blue line) and $\gamma^2 = 1/2$ (red line).

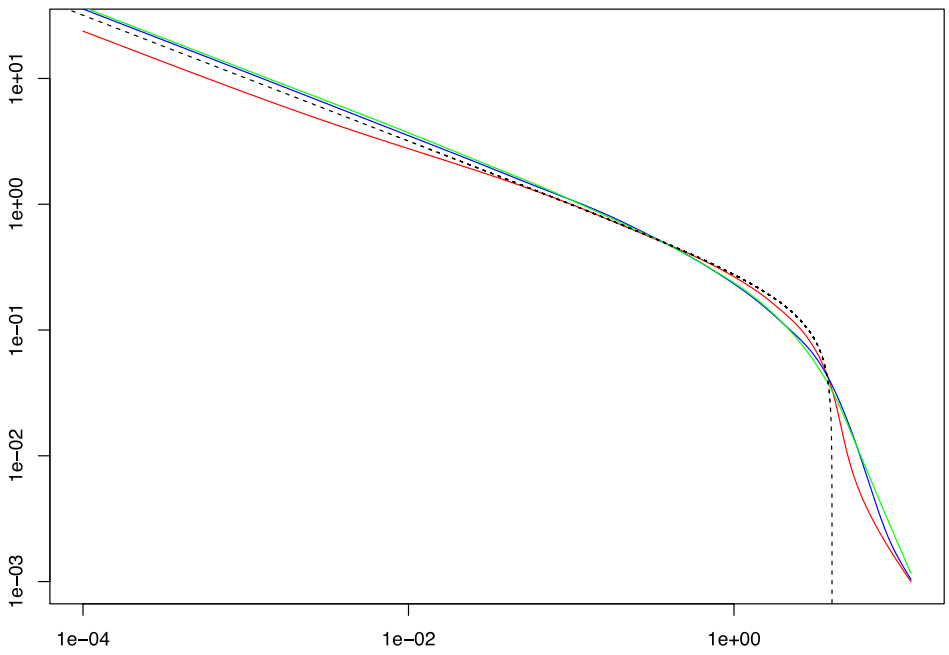


Figure 8.4: Log-log plot of the density $\upsilon \circ (x^2)^{-1}$ with $q = 1, \gamma^2 = 1/4$ for four different integral scales τ: $\tau = 0$ (black dashed line), $\tau = 1/4$ (red line), $\tau = 1$ (blue line) and $\tau = 2$ (green line).

proofs for theorems 8.3, 8.4 and 8.5; it is left to the reader. Furthermore, the proofs are very similar when $q = 1$ or when $q < 1$. For the sake of clarity, we assume $T = N$ and hence $q = 1$ in the proofs that follow.

Hence, in the following, we will suppose (unless otherwise stated) that:

$$\begin{aligned}\varphi(q) &= -iq\frac{\gamma^2}{2} + q^2\frac{\gamma^2}{2},\\ \psi(q) &= \varphi(-iq),\\ \zeta(q) &= (1+\frac{\gamma^2}{2})q + q^2\frac{\gamma^2}{2},\\ \gamma^2 &< \frac{1}{3},\end{aligned}$$

and M will be the MRM whose structure exponent is ζ (see section 8.2.1 for a reminder).

Our approach to show the convergence of $\mathbb{E}[L_N^{1,z}]$ and $\mathbb{E}[L_N^{2,z}]$ consists in proving tightness and characterizing uniquely the possible limit points. The classical Schur complement formula is our basic linear algebraic tool to study $\mathbb{E}[L_N^{1,z}]$ and $\mathbb{E}[L_N^{2,z}]$ 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 8.5 involving the limit point $K_z(x)$ of the measure $\mathbb{E}[L_N^{1,z}]$. We will also show that the limit points of the two complex measures $\mathbb{E}[L_N^{1,z}]$ and $\mathbb{E}[L_N^{2,z}]$ satisfy a fixed point system (written in theorem 8.5).

We begin by showing tightness.

8.5.1 Tightness of the complex measures $\mathbb{E}[L_N^{1,z}], \mathbb{E}[L_N^{2,z}]$ and limit points

Lemma 8.11. *The two families of complex measures $(\mathbb{E}[L_N^{i,z}])_{N\in\mathbb{N}}, i = 1, 2$ are tight and bounded in total variation.*

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

One has, for each $N \in \mathbb{N}$:

$$|\,\mathbb{E}[L_N^{1,z}]\,|\,[0,1] = \frac{1}{N}\sum_{k=1}^{N} |\,\mathbb{E}[G_N(z)_{kk}]\,| \leqslant \frac{1}{|\,\Im(z)\,|}, \tag{8.36}$$

and so the family of complex measures $(\mathbb{E}[L_N^{1,z}])_{N\in\mathbb{N}}$ 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 measure 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:

$$\mathbb{E}\left[L_N^{1,z}(f)\right] \to \int_0^1 f(x)\mu_z^1(dx). \tag{8.37}$$

Lemma 8.12. *The complex measure $\mu_z^1(dx)$ has Lebesgue density; more precisely, there exists a bounded measurable function $K_z(x)$ such that:*

$$\mu_z^1(dx) = K_z(x)dx. \tag{8.38}$$

Proof. One has:

$$\left|\mathbb{E}\left[L_N^{1,z}(f)\right]\right| \leqslant \frac{1}{N}\sum_{k=1}^{N} |f(k/N)|\mathbb{E}\left[G_N(z)_{kk}\right] \tag{8.39}$$

$$\leqslant \frac{1}{|\Im(z)|}\frac{1}{N}\sum_{k=1}^{N} |f(k/N)| \tag{8.40}$$

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

$$\left| \int_0^1 f(x) \mu_z^1(dx) \right| \leqslant \frac{1}{|\Im(z)|} \int_0^1 |f(x)| dx. \tag{8.41}$$

This proves the lemma. □

Thus, there exists a subsequence such that, as N tends to $+\infty$ along this subsequence:

$$\mathbb{E}\left[L_N^{1,z}(f) \right] \to \int_0^1 f(x) K_z(x) dx. \tag{8.42}$$

Lemma 8.13. *There exists a subsequence and a constant $\mu_z^2 \in \mathbb{C}$ such that, as N goes to $+\infty$ along this subsequence:*

$$\mathbb{E}\left[L_N^{2,z}(f) \right] \to \mu_z^2 \int_0^1 f(x) dx. \tag{8.43}$$

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

$$|\mu_z^2(N)| \leqslant \frac{1}{|\Im(z)|}. \tag{8.44}$$

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

$$\mathbb{E}\left[L_N^{2,z}(f) \right] \to \mu_z^2 \int_0^1 f(x) dx. \tag{8.45}$$

□

Following the classical method as in [33], [32], [93], we will show in the following that the limit point μ_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.

8.5.2 Preliminary results on resolvents

We first recall the following standard and general result; the next lemmas of this section are also standard but are applied to our particular case.

Lemma 8.14. *Let A be a symmetric real valued matrix of size N. For $z \in \mathbb{C} \setminus \mathbb{R}$, let us denote by $G(z)$ the matrix*

$$G(z) = (z - A)^{-1}. \tag{8.46}$$

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

$$\Im(z)\Im(G(z)_{kk}) < 0 \quad \text{and } |G(z)_{kk}| \leqslant \frac{1}{|\Im(z)|}. \tag{8.47}$$

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

$$\frac{\Im\left(z - \sum_{i \in F} a_i G(z)_{ii} \right)}{\Im(z)} \geqslant 1. \tag{8.48}$$

and we also have:

$$\frac{1}{\left| z - \sum_{i \in F} a_i G(z)_{ii} \right|} \leqslant \frac{1}{|\Im(z)|}. \tag{8.49}$$

Proof. Write $A = \bar{U}^t D U$ where D is a diagonal matrix with diagonal real entries $(\lambda_i)_{1 \leqslant i \leqslant N}$. Then

$$G(z)_{kk} = \sum_{i=1}^N |U_{ki}|^2 \frac{1}{z - \lambda_i}.$$

Since $\Re\left(\frac{1}{z-\lambda_i}\right) = \frac{\Re(z)-\lambda_i}{(\Re(z)-\lambda_i)^2+\Im(z)^2}$ and $\Im\left(\frac{1}{z-\lambda_i}\right) = \frac{-\Im(z)}{(\Re(z)-\lambda_i)^2+\Im(z)^2}$ the relation (8.47) follows. It is then straightforward to derive (8.48) from (8.47). □

For $i = 1, \dots, N$, let $X_N^{(i)} = (X_N(kl))_{k,l\neq i}$ be the matrix obtained from X_N by taking off the i-th column and row. Define, also for $i = 1, \dots, 2N$ the $(2N-1)\times(2N-1)$ matrix $A_N^{(i)}(z)$ obtained from $A_N(z)$ by taking off the i-th column and row. In particular, for $i = 1, \dots, N$,

$$A_N^{(N+i)}(z) = \begin{pmatrix} zI_N & -{}^tX_N^{(i)} \\ -X_N^{(i)} & zI_{N-1} \end{pmatrix},$$

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

$$G_N^{(i)}(z) = (A_N^{(i)}(z))^{-1}. \tag{8.50}$$

Let now $\hat{X}_N^{(i)}$ denote the matrix X_N with the i-th column and 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 $i = 1, \dots, N$:

$$\hat{A}_N^{(N+i)}(z) = \begin{pmatrix} zI_N & -{}^t\hat{X}_N^{(i)} \\ -\hat{X}_N^{(i)} & zI_N \end{pmatrix},$$

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

$$\hat{G}_N^{(i)}(z) = (\hat{A}_N^{(i)}(z))^{-1}. \tag{8.51}$$

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 8.15. *For all $k \in \{1, \dots, N\}$ and all $t \neq N+k$, one has:*

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(N+k)}(z)_{tt}\right|\right] \leqslant \frac{1}{\sqrt{N}|\Im(z)|^2}. \tag{8.52}$$

Proof. Multiply the identity:

$$\hat{A}_N^{(N+k)}(z) - A_N(z) = \hat{A}_N^{(N+k)}(0) - A_N(0) \tag{8.53}$$

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

$$G_N(z) - \hat{G}_N^{(N+k)}(z) = G_N(z)(\hat{A}_N^{(N+k)}(0) - A_N(0))\hat{G}_N^{(N+k)}(z). \tag{8.54}$$

Then one has:

$$G_N(z)_{tt} - \hat{G}_N^{(N+k)}(z)_{tt} = \left(G_N(z)(\hat{A}_N^{(N+k)}(0) - A_N(0))\hat{G}_N^{(N+k)}(z)\right)_{tt} \tag{8.55}$$

$$= \hat{G}_N^{(N+k)}(z)_{N+k,t} \sum_{i=1}^N G_N(z)_{ti} r_k(i) \tag{8.56}$$

$$+ G_N(z)_{t,N+k} \sum_{j=1}^N r_k(j) \hat{G}_N^{(N+k)}(z)_{jt} \tag{8.57}$$

$$= G_N(z)_{t,N+k} \sum_{j=1}^N r_k(j) \hat{G}_N^{(N+k)}(z)_{jt} \tag{8.58}$$

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:

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(N+k)}(z)_{tt}\right|\right] \leqslant \mathbb{E}\left[|G_N(z)_{t,N+k}|^2\right]^{1/2} \mathbb{E}\left[\left|\sum_{j=1}^N r_k(j)\hat{G}_N^{(N+k)}(z)_{jt}\right|^2\right]^{1/2} \tag{8.59}$$

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

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(N+k)}(z)_{tt}\right|\right] \leqslant \mathbb{E}\left[|G_N(z)_{t,N+k}|^2\right]^{1/2} \mathbb{E}\left[r_k(1)^2\right]^{1/2} \mathbb{E}\left[\sum_{j=1}^{N}\left|\hat{G}_N^{(N+k)}(z)_{jt}\right|^2\right]^{1/2}$$
$$\leqslant \frac{1}{\sqrt{N}|\Im(z)|^2}.$$

The proof is complete. □

Lemma 8.16. *There exists a constant $C > 0$ such that, for all $k \in \{1, \ldots, N\}$ and all $t \neq k$:*

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(k)}(z)_{tt}\right|\right] \leqslant \frac{C}{|\Im(z)|^2} \frac{1}{N^{\frac{1-\gamma^2}{4}}}. \tag{8.60}$$

Proof. Again, we start from the relation:

$$G_N(z) - \hat{G}_N^{(k)}(z) = G_N(z)(\hat{A}_N^{(k)}(0) - A_N(0))\hat{G}_N^{(k)}(z).$$

Thus we have

$$G_N(z)_{tt} - \hat{G}_N^{(k)}(z)_{tt} = \left(G_N(z)(\hat{A}_N^{(k)}(0) - A_N(0))\hat{G}_N^{(k)}(z)\right)_{tt} \tag{8.61}$$
$$= \hat{G}_N^{(k)}(z)_{k,t} \sum_{i=N+1}^{N} G_N(z)_{ti} r_i(k) \tag{8.62}$$
$$+ G_N(z)_{t,k} \sum_{j=1}^{N+1} r_j(k) \hat{G}_N^{(k)}(z)_{jt} \tag{8.63}$$
$$= G_N(z)_{t,k} \sum_{j=1}^{N+1} r_j(k) \hat{G}_N^{(k)}(z)_{jt} \tag{8.64}$$

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

Therefore, we find that:

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(k)}(z)_{tt}\right|\right] \leqslant \mathbb{E}\left[|G_N(z)_{t,k}|^2\right]^{1/2} \mathbb{E}\left[\left|\sum_{j=1}^{N} r_j(k) \hat{G}_N^{(k)}(z)_{jt}\right|^2\right]^{1/2} \tag{8.65}$$

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 $(r_j(k))_j$ are independent from $\hat{G}_N^{(k)}(z)$ and centered. Hence we have for $j \neq j'$,

$$\mathbb{E}\left[r_j(k) r_{j'}(k) \hat{G}_N^{(k)}(z)_{jt} \hat{G}_N^{(k)}(z)_{j't}\right] = 0.$$

Thus we get:

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(k)}(z)_{tt}\right|\right] \leqslant \mathbb{E}\left[|G_N(z)_{t,k}|^2\right]^{1/2} \left(\sum_{j=1}^{N+1} \mathbb{E}\left[r_j(k)^2 \left|\hat{G}_N^{(k)}(z)_{jt}\right|^2\right]\right)^{1/2}$$
$$\leqslant \mathbb{E}\left[|G_N(z)_{t,k}|^2\right]^{1/2} \left(\sum_{j=1}^{N+1} \mathbb{E}\left[r_j(k)^4\right]^{1/2} \mathbb{E}\left[\left|\hat{G}_N^{(k)}(z)_{jt}\right|^4\right]^{1/2}\right)^{1/2}$$
$$\leqslant \frac{\mathbb{E}[r_1(k)^4]^{1/4}}{|\Im(z)|} \left(\sum_{j=1}^{N+1} \mathbb{E}\left[\left|\hat{G}_N^{(k)}(z)_{jt}\right|^4\right]^{1/2}\right)^{1/2}$$
$$\leqslant \frac{\mathbb{E}[r_1(k)^4]^{1/4}}{|\Im(z)|} (N+1)^{1/4} \left(\sum_{j=1}^{N+1} \mathbb{E}\left[\left|\hat{G}_N^{(k)}(z)_{jt}\right|^4\right]\right)^{1/4}$$

Now we use the scaling properties of the MRM to obtain, for some positive constant C,

$$\mathbb{E}[r_j(k)^4] = 3\mathbb{E}[M(0, \frac{1}{N})^2] \leqslant CN^{-\zeta(2)}.$$

Furthermore, by using Lemma 8.26 which assures that, almost surely:

$$\sum_{j=1}^{N+1} \left|\hat{G}_N^{(k)}(z)_{jt}\right|^2 \leqslant \frac{1}{|\Im(z)|^2} \tag{8.66}$$

and the fact that:

$$\sum_{j=1}^{N+1} \left|\hat{G}_N^{(k)}(z)_{jt}\right|^4 \leqslant \left(\sum_{j=1}^{N+1} \left|\hat{G}_N^{(k)}(z)_{jt}\right|^2\right)^2, \tag{8.67}$$

we finally obtain

$$\mathbb{E}\left[\left|G_N(z)_{tt} - \hat{G}_N^{(k)}(z)_{tt}\right|\right] \leqslant \frac{C}{|\Im(z)|^2}\left(\frac{1}{N}\right)^{\frac{\zeta(2)-1}{4}}.$$

It just remains to check that $\zeta(2) = 2 - \gamma^2$. □

Lemma 8.17. *For each $k \in \{1, \ldots, 2N\}$, if $t \neq k$, then*

$$G_N^{(k)}(z)_{tt} = \hat{G}_N^{(k)}(z)_{tt}, \tag{8.68}$$

and if $t = k$, then $\hat{G}_N^{(k)}(z)_{k,k} = z^{-1}$.

Proof. It is straightforward to see that the two matrices $G_N^{(k)}(z)$ and $\hat{G}_N^{(k)}(z)$ have the same eigenvalues except that $\hat{G}_N^{(k)}(z)$ has one more zero eigenvalue. In addition, the eigenvectors look also very similar since you can obtain $2N$ eigenvectors of $\hat{G}_N^{(k)}(z)$ by adding a zero entry to the eigenvectors of $G_N^{(k)}(z)$ (between the entries $k-1$ and k). The last eigenvector of $\hat{G}_N^{(k)}(z)$ is the vector of $\mathbb{R}^N$ for which all entries are zero except the entry number k.

Now observe that with $G_N^{(k)}(z) = U\mathrm{diag}(z-\lambda)U^*$ and $\hat{G}_N^{(k)}(z) = V\mathrm{diag}(z-\widetilde{\lambda})V^*$,

$$G_N^{(k)}(z)_{tt} = \sum_{i=1}^{2N} |u_{ti}|^2 \frac{1}{z-\lambda_i} \tag{8.69}$$

$$\hat{G}_N^{(k)}(z)_{tt} = \sum_{i=1}^{N} |v_{ti}|^2 \frac{1}{z-\widetilde{\lambda}_i}. \tag{8.70}$$

The result follows since, for $t \neq k$,

$$\sum_{i=1}^{2N-1} |u_{ti}|^2 \frac{1}{z-\lambda_i} = \sum_{i=1}^{2N} |v_{ti}|^2 \frac{1}{z-\widetilde{\lambda}_i} \tag{8.71}$$

and, for $t = k, \hat{G}_N^{(k)}(z)_{k,k} = z^{-1}$. □

Lemma 8.18. *For all $z \in \mathbb{C}$ and Lebesgue almost every point $x \in [0,1]$, we have*

$$\Im(z)\Im(K_z(x)) \leqslant 0 \tag{8.72}$$

and

$$|\Im(K_z(x))| \leqslant \frac{1}{\Im(z)} \tag{8.73}$$

Proof. This is a straightforward consequence of Lemma 8.14. Indeed, we have for all positive continuous function f on $[0,1]$ and $N \in \mathbb{N}$:

$$\Im(z)\Im\Big(\int_0^1 f(x)\mathbb{E}[L_N^{1,z}](dx)\Big) \leqslant 0.$$

We pass to the limit as N goes to ∞ along some suitable subsequence and obtain:

$$\Im(z)\Im\Big(\int_0^1 f(x)K_z(x)\,dx\Big) \leqslant 0.$$

The result follows. □

8.5.3 Concentration inequalities

This lemma is adapted to our case from Lemma 5.4 in [33].

Lemma 8.19. *Let $f : [0,1] \to \mathbb{R}$ be a bounded measurable function. For each $i \in \{1,2\}$, we have the following concentration results:*

$$\mathbb{E}\left[\mid L_N^{i,z}(f) - \mathbb{E}[L_N^{i,z}(f)] \mid^2\right] \leqslant \frac{8}{N} \frac{\mid\mid f \mid\mid_\infty^2}{\mid \Im z \mid^4}. \tag{8.74}$$

Proof. Define two functions F_N^1 and F_N^2 such that:

$$F_N^1\left(\left(X_{ij}^{(N)}\right)_{1 \leqslant j \leqslant N+1}, 1 \leqslant i \leqslant N\right) = \frac{1}{N}\sum_{k=1}^{N} f\left(\frac{k}{N}\right) G_N(z)_{kk} \tag{8.75}$$

$$F_N^2\left(\left(X_{ij}^{(N)}\right)_{1 \leqslant j \leqslant N+1}, 1 \leqslant i \leqslant N\right) = \frac{1}{N}\sum_{k=1}^{N+1} f\left(\frac{k}{N+1}\right) G_N(z)_{k+N,k+N} \tag{8.76}$$

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

$$\mathcal{F}_k = \sigma\left(\left(X_{ij}^{(N)}\right)_{1 \leqslant j \leqslant N}, 1 \leqslant i \leqslant k\right) \tag{8.77}$$

If P denotes the law of the vector $\left(X_{1j}^{(N)}\right)_{1 \leqslant j \leqslant N}$,

$$\begin{aligned}
&\mathbb{E}\left[\mid F_N^1 - \mathbb{E}[F_N^1] \mid^2\right] \\
&= \sum_{i=0}^{N} \mathbb{E}\left[\mid \mathbb{E}[F_N^1 \mid \mathcal{F}_{i+1}] - \mathbb{E}[F_N^1 \mid \mathcal{F}_i] \mid^2\right] \\
&= \sum_{i=0}^{N} \int \mid \int \left(F_N(x_1,x_2,\dots,x_{i+1},y_{i+2},\dots,y_{N+1}) - F_N(x_1,x_2,\dots,x_i,y_{i+1},\dots,y_{N+1})\right) dP^{\otimes N+1}(y) \mid^2 \\
&dP^{\otimes i+1}(x) \\
&\leqslant \sum_{i=0}^{N} \int \mid \int \left(F_N(x_1,x_2,\dots,x_i,x_{i+1},x_{i+2},\dots,x_{N+1}) - F_N(x_1,x_2,\dots,x_i,y,x_{i+2},\dots,x_{N+1})\right) dP(y) \mid^2 \\
&dP^{\otimes N+1}(x) \\
&\leqslant \sum_{i=0}^{N} \sup_{\mathbb{R}^{(N+1)^2}} \mid\mid \nabla_{x_{i+1}} F_N \mid\mid^2 \int \mid\mid x-y \mid\mid^2 dP^{\otimes 2}(x,y).
\end{aligned}$$

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 $(\widetilde{B}^1, \widetilde{M}^1)$ independent from (B^1, M^1) with the same law, it is easy to see that:

$$\begin{aligned}
\int \mid\mid x-y \mid\mid^2 dP \otimes dP(x,y) &= \sum_{j=1}^{N} \mathbb{E}\Big[(B^1_{M^1(0,\frac{j}{N})} - B^1_{M^1(0,\frac{j-1}{N})} - \widetilde{B}^1_{\widetilde{M}^1(0,\frac{j}{N})} + \widetilde{B}^1_{\widetilde{M}^1(0,\frac{j-1}{N})})^2\Big]. \\
&= 2 - 2\sum_{j=1}^{N} \mathbb{E}\Big[(B^1_{M^1(0,\frac{j}{N})} - B^1_{M^1(0,\frac{j-1}{N})})(\widetilde{B}^1_{\widetilde{M}^1(0,\frac{j}{N})} \widetilde{B}^1_{\widetilde{M}^1(0,\frac{j-1}{N})})\Big] \\
&= 2.
\end{aligned}$$

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

$$\frac{\partial G_N(z)_{kk}}{\partial X_{ij}} = G_N(z)_{k,j} G_N(z)_{N+i,k} + G_N(z)_{k,N+i} G_N(z)_{j,k} \tag{8.78}$$

Thus,

$$\nabla_{x_{i+1}} F_N = \frac{1}{N} \sum_{k=1}^{N} f\left(\frac{k}{N}\right) \nabla_{x_{i+1}} G_N(z)_{kk} \tag{8.79}$$

It is now plain to compute:

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

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

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

One thus has:

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

where, in the last line, we used lemma 8.26 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,

$$\mathbb{E}\left[| F_N^1 - \mathbb{E}[F_N^1] |^2 \right] \leqslant \frac{8}{N} \frac{|| f ||_\infty^2}{| \Im z |^4}. \tag{8.80}$$

□

We also prove the following lemma:

Lemma 8.20. *For all $\alpha > 1$ such that $\zeta(2\alpha) > 1$, we have*

$$\mathbb{E}\left[\left| \sum_{t=1}^{N} r_k(t)^2 \left(\hat{G}_N^{(N+k)}(z)_{tt} - \mathbb{E}[\hat{G}_N^{(N+k)}(z)_{tt}] \right) \right| \right] \leqslant \frac{C (\ln N)^2}{N^{\frac{\zeta(2\alpha)-1}{\alpha}} |\Im(z)|^4} \tag{8.81}$$

for some positive constant C independent from N, z, k.

Proof. Notice that $(r_k(t))_t$ and $\hat{G}_N^{(N+k)}(z)$ are independent. Hence, by conditioning with respect to the process $(r_k(t))_t$, we can argue along the same lines as in the previous lemma with $r_k(t)$ instead of $\frac{1}{N} f(\frac{t}{N})$ and we get the formula:

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

We conclude with Proposition 8.29 in the appendix . □

In the following, we fix $\alpha > 1$ such that $\zeta(2\alpha) > 1$ (because of the expression of ζ and the inequality $\gamma^2 < 1/3$, it is clear that such a number exists).

8.5.4 The system verified by the limit point μ_z^2 and $K_z(x)$: first equation

From the Schur complement formula (see e.g. Lemma 4.2 in [33] for a reminder), one has for $k \in \{1, \ldots, N\}$:

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{s,t=1}^{N} r_k(s) r_k(t) G_N^{(N+k)}(z)_{st} \right]^{-1} \tag{8.82}$$

Using Lemma 8.27, one can write:

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{t=1}^{N} r_k(t)^2 G_N^{(N+k)}(z)_{tt} + \epsilon_{N,k}^1(z) \right]^{-1} \tag{8.83}$$

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

$$\mathbb{E}[|\epsilon_{N,k}^1(z)|^2] < \frac{C}{N^{1-\gamma^2}}. \tag{8.84}$$

By using Lemma 8.17, we can write:

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{t=1}^{N} r_k(t)^2 \hat{G}_N^{(N+k)}(z)_{tt} + \epsilon_{N,k}^1(z) \right]^{-1}. \tag{8.85}$$

Lemma 8.20 applied to $\alpha > 1$ such that $\zeta(2\alpha) > 1$ yields:

$$\mathbb{E}\left[\left| \sum_{t=1}^{N} r_k(t)^2 \left(\hat{G}_N^{(N+k)}(z)_{tt} - \mathbb{E}[\hat{G}_N^{(N+k)}(z)_{tt}] \right) \right|^2 \right] \leqslant \frac{C(\ln N)^2}{N^{\frac{\zeta(2\alpha)-1}{\alpha}} |\Im(z)|^4}. \tag{8.86}$$

Thus, one can write:

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{t=1}^{N} r_k(t)^2 \mathbb{E}\left[\hat{G}_N^{(N+k)}(z)_{tt} \right] + \epsilon_{N,k}^1(z) + \epsilon_{N,k}^2(z) \right]^{-1} \tag{8.87}$$

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

$$\mathbb{E}[|\epsilon_{N,k}^2(z)|^2] < \frac{C(\ln N)^2}{N^{\frac{\zeta(2\alpha)-1}{\alpha}} |\Im(z)|^4}. \tag{8.88}$$

In addition, using Lemma 8.15, we can show:

$$\mathbb{E}\left[\left| \sum_{t=1}^{N} r_k(t)^2 \left(\mathbb{E}\left[\hat{G}_N^{(N+k)}(z)_{tt} - G_N(z)_{tt} \right] \right) \right| \right] \tag{8.89}$$

$$\leqslant \sum_{t=1}^{N} \mathbb{E}[r_k(t)^2] \mathbb{E}\left[\left| \hat{G}_N^{(N+k)}(z)_{tt} - G_N(z)_{tt} \right| \right] \tag{8.90}$$

$$\leqslant \frac{1}{|\Im(z)|^2 \sqrt{N}}. \tag{8.91}$$

It follows:

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{t=1}^{N} r_k(t)^2 \mathbb{E}\left[G_N(z)_{tt} \right] + \epsilon_{N,k}^1(z) + \epsilon_{N,k}^2(z) + \epsilon_{N,k}^3(z) \right]^{-1} \tag{8.92}$$

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

$$\mathbb{E}\left[|\epsilon_{N,k}^3(z)| \right] < \frac{1}{|\Im(z)|^2 \sqrt{N}}. \tag{8.93}$$

Let us denote by I_N^t the interval $[\frac{t-1}{N}, \frac{t}{N}]$. Then we have:

Lemma 8.21. *The following inequality holds:*

$$\mathbb{E}\left[\left| \sum_{t=1}^{N} \left(r_k(t)^2 - M^k(I_N^t) \right) \mathbb{E}\left[G_N(z)_{tt} \right] \right|^2 \right] \leqslant \frac{C}{N^{1-\gamma^2} |\Im(z)|^2}$$

for some positive constant C.

Proof. We expand the square and, because $r_k(t)$ and $r_k(t')$ are independent for $t \neq t'$ conditionally to M^k, we have:

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\sum_{t=1}^{N}\Big(r_k(t)^2 - M^k(I_N^t)\Big)\mathbb{E}\left[G_N(z)_{tt}\right]\Big|^2\Big] \\
&\quad= \sum_{t,t'=1}^{N}\mathbb{E}\Big[\Big(r_k(t)^2 - M^k(I_N^t)\Big)\Big(r_k(t')^2 - M^k(I_N^{t'})\Big)\mathbb{E}\left[G_N(z)_{tt}\right]\mathbb{E}\left[G_N(z)_{t't'}\right]\Big] \\
&\quad= \sum_{t=1}^{N}\mathbb{E}\Big[\Big(r_k(t)^2 - M^k(I_N^t)\Big)^2\Big]\mathbb{E}\left[G_N(z)_{tt}\right]^2 \\
&\quad= 2\sum_{t=1}^{N}\mathbb{E}\left[\Big(M^k(I_N^t)\Big)^2\right]\mathbb{E}\left[G_N(z)_{tt}\right]^2 \\
&\quad\leqslant 2C\frac{N}{N^{\zeta(2)}|\Im(z)|^2}
\end{aligned}$$

□

We can thus write

$$G_N(z)_{N+k,N+k} = \Bigg[z - \sum_{t=1}^{N} M^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] \tag{8.94}$$

$$+ \epsilon_{N,k}^1(z) + \epsilon_{N,k}^2(z) + \epsilon_{N,k}^3(z) + \epsilon_{N,k}^4(z)\Bigg]^{-1} \tag{8.95}$$

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

$$\mathbb{E}\left[|\epsilon_{N,k}^4(z)|^2\right] \leqslant \frac{C}{N^{\zeta(2)-1}|\Im(z)|^2}. \tag{8.96}$$

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

$$G_N(z)_{N+k,N+k} = \Bigg[z - \sum_{t=1}^{N} M^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] + \epsilon_{N,k}(z)\Bigg]^{-1} \tag{8.97}$$

We now need to introduce the truncated Radon measure $M_\epsilon^k(dx)$ with Lebesgue density $e^{\omega_\epsilon^k(x)}$ which converges almost surely as ϵ goes to 0, in the sense of weak convergence in the space of Radon measure, to the measure M^k (see section 8.2.1).

Lemma 8.22. *For $\epsilon > 0$, the following uniform bound holds:*

$$\sup_N \mathbb{E}\Big[\Big|\sum_{t=1}^{N} M^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] - \sum_{t=1}^{N} M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]\Big|^2\Big] \leqslant \frac{C\epsilon^{1-\gamma^2}}{|\Im(z)|^2}.$$

Proof. We expand the square. Note that the covariance function ρ_ϵ of the process ω_ϵ increases as ϵ decreases to 0 and uniformly converges as $\epsilon \to 0$ towards $\ln_+ \frac{\tau}{|x|}$ over the complement of

any ball centered at 0. Thus we have:

$$\begin{aligned}
&\sup_N \mathbb{E}\big[|\sum_{t=1}^N M^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] - \sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]|^2\big] \\
&\qquad = \sup_N \sum_{t,t'=1}^N \mathbb{E}\big[(M^k(I_N^t) - M_\epsilon^k(I_N^t))(M^k(I_N^{t'}) - M_\epsilon^k(I_N^{t'}))\big]\mathbb{E}\left[G_N(z)_{tt}\right]\mathbb{E}\left[G_N(z)_{t't'}\right] \\
&\qquad = \sup_N \sum_{t,t'=1}^N \mathbb{E}\left[(M^k(I_N^t) - M_\epsilon^k(I_N^t))(M^k(I_N^{t'}) - M_\epsilon^k(I_N^{t'}))\right]\mathbb{E}\left[G_N(z)_{tt}\right]\mathbb{E}\left[G_N(z)_{t't'}\right] \\
&\qquad = \sup_N \sum_{t,t'=1}^N \left(\mathbb{E}\left[M^k(I_N^t)M^k(I_N^{t'})\right] - \mathbb{E}\left[M_\epsilon^k(I_N^t)M_\epsilon^k(I_N^{t'})\right]\right)\mathbb{E}\left[G_N(z)_{tt}\right]\mathbb{E}\left[G_N(z)_{t't'}\right] \\
&\qquad = \sup_N \sum_{t,t'=1}^N \mathbb{E}\left[G_N(z)_{tt}\right]\mathbb{E}\left[G_N(z)_{t't'}\right]\int_{I_N^t}\int_{I_N^{t'}} \left(e^{\psi(2)\ln_+\frac{\tau}{|r-u|}} - e^{\psi(2)\rho_\epsilon(r-u)}\right) drdu \\
&\qquad \leqslant \frac{1}{|\Im(z)|^2}\int_0^1\int_0^1 \left(e^{\psi(2)\ln_+\frac{\tau}{|r-u|}} - e^{\psi(2)\rho_\epsilon(r-u)}\right) drdu.
\end{aligned}$$

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

$$\rho_\epsilon(t) = \begin{cases} \ln\frac{\tau}{\epsilon} + 1 - \frac{|t|}{\epsilon} & \text{if } |t| \leqslant \epsilon \\ \ln\frac{\tau}{|t|} & \text{if } \epsilon \leqslant |t| \leqslant \tau \\ 0 & \text{if } \tau < |t| \end{cases} \tag{8.98}$$

By using the expression of ρ_ϵ, it is then plain to obtain the desired bound. □

We can thus write

$$G_N(z)_{N+k,N+k} = \left[z - \sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] + \epsilon_{N,k}(z) + \delta(\epsilon,N,z)\right]^{-1}, \tag{8.99}$$

where

$$\sup_N \mathbb{E}[|\delta(\epsilon,N,z)|^2] \to 0 \quad \text{as } \epsilon \to 0, \tag{8.100}$$

and also:

$$\mathbb{E}\left[G_N(z)_{N+k,N+k}\right] = \mathbb{E}\left[\left[z - \sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] + \epsilon_{N,k}(z) + \delta(\epsilon,N,z)\right]^{-1}\right]. \tag{8.101}$$

The next step is to study the convergence of the above quantity. Hence we prove (see the proof in the appendix):

Lemma 8.23. *The random variable $\sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]$ converges in probability as $N \to +\infty$ towards $\int_0^1 K_z(x)M_\epsilon^k(dx)$.*

We fix $\epsilon > 0$. For that ϵ, the family of random variables $(\delta(\epsilon,N,z))_N$ is bounded in L^2 so that it is tight. Even if it means extracting again a subsequence we assume that the couple $(\sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right], \delta(\epsilon,N,z))_N$ converges in law towards the couple $(\int_0^1 K_z(x)M_\epsilon^k(dx), Y_\epsilon)$. We remind the reader of (8.82) which implies that

$$\left|\left(z - \sum_{t=1}^N M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] + \epsilon_{N,k}(z) + \delta(\epsilon,N,z)\right)^{-1}\right| \leqslant \frac{1}{|\Im(z)|}.$$

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

$$\left(z - \int_0^1 K_z(x) M_\epsilon^k(dx) + Y_\epsilon\right)^{-1}.$$

We deduce that the expectation of the former quantity converges as $\epsilon \to 0$ towards the expectation of the latter quantity. From (8.101), we deduce that

$$\mu_z^2 = \mathbb{E}\left[\left(z - \int_0^1 K_z(x) M_\epsilon^k(dx) + Y_\epsilon\right)^{-1}\right]. \tag{8.102}$$

Clearly, standard arguments prove that $\int_0^1 K_z(x) M_\epsilon^k(dx)$ converges almost surely towards $\int_0^1 K_z(x) M^k(dx)$ as $\epsilon \to 0$ (K_z is deterministic (see lemma 8.19), measurable and bounded) and, because of (8.100), Y_ϵ converges almost surely towards 0 as $\epsilon \to 0$. Again, because the quantity $\left(z - \int_0^1 K_z(x) M_\epsilon^k(dx) + Y_\epsilon\right)^{-1}$ is bounded uniformly with respect to ϵ, we deduce that:

$$\mu_z^2 = \mathbb{E}\left[\left(z - \int_0^1 K_z(x) M^k(dx)\right)^{-1}\right]. \tag{8.103}$$

8.5.5 Second equation

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

$$G_N(z)_{kk} = \left[z - \sum_{i,j=1}^{N} r_i(k) r_j(k) G_N^{(k)}(z)_{N+i,N+j}\right]^{-1} \tag{8.104}$$

$$= \left[z - \sum_{i=1}^{N} r_i(k)^2 G_N^{(k)}(z)_{N+i,N+i} + \eta_{N,k}^1(z)\right]^{-1} \tag{8.105}$$

where, using Lemma 8.28, $\eta_{N,k}^1(z)$ is a complex valued random variable for which there exists $c > 0$ such that for all $N \in \mathbb{N}$ and $1 \leqslant k \leqslant N, \mathbb{E}[|\eta_{N,k}^1(z)|^2] < c/N$.

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

$$G_N(z)_{kk} = \left[z - \sum_{i=1}^{N} r_i(k)^2 \left[z - \sum_{s,t\neq k}^{N} r_i(s) r_i(t) G_N^{(k,N+i)}(z)_{st}\right]^{-1} + \eta_{N,k}^1(z)\right]^{-1} \tag{8.106}$$

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 $(r_i(t))_{t=1,\dots,N}$. Using the same arguments as in the derivation of the first equation (in particular Lemmas 8.27, 8.17, 8.20, 8.29, 8.16 and 8.15), one can show that:

$$G_N(z)_{kk} = \left[z - \sum_{i=1}^{N} \frac{r_i(k)^2}{z - \sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] + \delta_{N,k,i}(z)} + \eta_{N,k}^1(z)\right]^{-1} \tag{8.107}$$

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

$$\mathbb{E}[|\delta_{N,k,i}(z)|] \leqslant \frac{C}{N^{\min(\frac{1-\gamma^2}{4}, \frac{\zeta(2\alpha)-1}{\alpha})}} \tag{8.108}$$

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

Lemma 8.24. *One can write:*

$$G_N(z)_{kk} = \left[z - \sum_{i=1}^{N} \frac{r_i(k)^2}{z - \sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]} + \eta_{N,k}^1(z) + \eta_{N,k}^2(z)\right]^{-1} \tag{8.109}$$

where $\eta_{N,k}^2(z)$ is a random variable that tends to 0 in probability as N goes to ∞.

Proof. By using Lemma 8.14, we deduce that:

$$\sum_{i=1}^{N}\Big|\frac{r_i(k)^2}{z-\sum_{t=1}^{N}M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]+\delta_{N,k,i}(z)}-\frac{r_i(k)^2}{z-\sum_{t=1}^{N}M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]}\Big|$$
$$\leqslant\frac{1}{|\Im(z)|^2}\sum_{i=1}^{N}r_i(k)^2\min(|\delta_{N,k,i}(z)|,2). \tag{8.110}$$

We stress that the lemma is proved as soon as we can prove that the left-hand side in (9.19) converges in probability to 0. Hence it is enough to prove that

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

converges to 0 as N tends to ∞. By noticing that:

$$\delta_{N,k,i}(z)=\sum_{s,t\neq k}^{N}r_i(s)r_i(t)G_N^{(k,N+i)}(z)_{st}-\sum_{t=1}^{N}M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right], \tag{8.111}$$

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

$$\mathbb{E}\big[\sum_{i=1}^{N}r_i(k)^2\min(|\delta_{N,k,i}(z)|,2)\big]=N\mathbb{E}\big[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\big].$$

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

$$\begin{aligned}N\mathbb{E}\big[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\big]=&N\mathbb{E}\big[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\mathbf{1}_{\{Nr_1(k)^2\leqslant A\}}\big]\\&+N\mathbb{E}\big[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\mathbf{1}_{\{Nr_1(k)^2>A\}}\big]\\\leqslant&A\mathbb{E}[|\delta_{N,k,1}(z)|]+2\mathbb{E}\big[Nr_1(k)^2\mathbf{1}_{\{Nr_1(k)^2>A\}}\big]\\\leqslant&\frac{AC}{N^{\frac{\zeta(2)-1}{4}}}+\frac{2}{A^\alpha}\mathbb{E}\big[N^{1+\alpha}r_1(k)^{2(\alpha+1)}\big]\\=&\frac{AC}{N^{\frac{\zeta(2)-1}{4}}}+\frac{2N^{1+\alpha}}{A^\alpha}\mathbb{E}\big[M^1(0,\frac{1}{N})^{\alpha+1}\big]\end{aligned}$$

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

$$N\mathbb{E}\left[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\right]\leqslant\frac{AC}{N^{\frac{\zeta(2)-1}{4}}}+2\mathbb{E}\left[M^1(0,1)^{\alpha+1}\right]\frac{N^{\psi(1+\alpha)}}{A^\alpha}. \tag{8.112}$$

Since $\zeta(2)>5-4\zeta'(1)$ (this inequality is clear with $\zeta(q)=(1+\gamma^2/2)q+q^2\gamma^2/2$ and is due to our hypotheses of Assumption 8.24 in the more general case), we can choose $p>0$ such that

$$\frac{\zeta(2)-1}{4}>p>1-\zeta'(1)=\psi'(1). \tag{8.113}$$

The mapping $\alpha\in]0,+\infty[\mapsto p\alpha-\psi(1+\alpha)$ reduces to 0 for $\alpha=0$ and, because $p>\psi'(1)$, 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:

$$N\mathbb{E}\big[r_1(k)^2\min(|\delta_{N,k,1}(z)|,2)\big]\leqslant\frac{C}{N^{\frac{\zeta(2)-1}{4}-p}}+2^{2+\alpha}\mathbb{E}\big[M^1(0,T)^{\alpha+1}\big]\frac{1}{N^{\alpha p-\psi(1+\alpha)}}.$$

The result follows by letting $N\to\infty$ since $\min((\zeta(2)-1)/4-p,\alpha p-\psi(1+\alpha))>0$. □

Lemma 8.25. *There exists a constant $c > 0$, which does not depend on N, such that for each $N \in \mathbb{N}$:*

$$\mathbb{E}\left[\left|\sum_{i=1}^{N}\left(\frac{r_i(k)^2}{z-\sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]}-\mathbb{E}\left[\frac{r_i(k)^2}{z-\sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]}\right]\right)\right|^2\right] \leqslant \frac{c}{N^{1-\gamma^2}}.$$

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

$$\frac{r_i(k)^2}{z-\sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]} \tag{8.114}$$

are i.i.d. random variables and Lemma 8.14. □

Therefore we can write

$$G_N(z)_{kk} = \left[z-\sum_{i=1}^{N}\frac{r_i(k)^2}{z-\sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right]}+\eta_{N,k}^1(z)+\eta_{N,k}^2(z)+\eta_{N,k}^3(z)\right]^{-1} \tag{8.115}$$

with $\mathbb{E}[(\eta_{N,k}^3(z))^2] \leqslant \frac{c}{N^{1-\gamma^2}}$.

Now we can take the expectation in (8.115) to obtain

$$\begin{aligned}
&\mathbb{E}[L_N^{1,z}(f)]\\
&= \frac{1}{N}\sum_{k=1}^{N} f(k/N)\mathbb{E}[G_N(z)_{kk}]\\
&= \frac{1}{N}\sum_{k=1}^{N} f(k/N)\mathbb{E}\Big[\Big(z-\mathbb{E}\Big[\sum_{i=1}^{N}\frac{r_i(k)^2}{z-\sum_{t=1}^{N} M^i(I_N^t)\mathbb{E}\big[G_N(z)_{tt}\big]}\Big]+\eta_{N,k}(z)\Big)^{-1}\Big]\\
&= \frac{1}{N}\sum_{k=1}^{N} f(k/N)\mathbb{E}\Big[\Big(z-N\mathbb{E}\Big[\frac{M\left[\frac{k-1}{N};\frac{k}{N}\right]}{z-\sum_{t=1}^{N} M(I_N^t)\mathbb{E}\big[G_N(z)_{tt}\big]}\Big]+\eta_{N,k}(z)\Big)^{-1}\Big]
\end{aligned}$$

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_ϵ and by using the Girsanov formula, we can approximate (uniformly in N) this last expression by:

$$\frac{1}{N}\sum_{k=1}^{N} f(k/N)\mathbb{E}\Big[\Big(z-N\mathbb{E}\Big[\frac{M_\epsilon\left[\frac{k-1}{N};\frac{k}{N}\right]}{z-\sum_{t=1}^{N} M_\epsilon(I_N^t)\mathbb{E}\big[G_N(z)_{tt}\big]}\Big]\Big)^{-1}+\hat{\delta}(N,k,z,\epsilon)\Big] \tag{8.116}$$

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

$$\int_0^1 f(x)\mathbb{E}\left[\left(z-\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z-\int_0^1 K_z(r)\,M_\epsilon(dr)}\right]\right)^{-1}+Y^\epsilon\right]dx \tag{8.117}$$

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

$$\int_0^1 f(x)K_z(x)\,dx = \int_0^1 f(x)\mathbb{E}\left[\left(z-\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z-\int_0^1 K_z(r)\,M_\epsilon(dr)}\right]\right)^{-1}+Y^\epsilon\right]dx. \tag{8.118}$$

It remains to pass to the limit as $\epsilon \to 0$ in that expression. This job is carried out with the help of a Girsanov type transform in Appendix 8.8. □

8.5.6 Uniqueness of the solution to the system of equations

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

$$||f||_\infty = \sup_{x\in[0,1]} |f(x)|. \tag{8.119}$$

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

$$Tg(x) = \frac{1}{z - q\mathbb{E}\left[\left(z - \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} g(t)M(dt)\right)^{-1}\right]} \tag{8.120}$$

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

$$\begin{aligned}
|Tg(x) - Th(x)| &\leqslant \frac{q}{|\Im(z)|^4}\mathbb{E}\left[\int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} |g(t) - h(t)|M(dt)\right] \\
&\leqslant \frac{q}{|\Im(z)|^4}\mathbb{E}\left[\int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} M(dt)\right] ||g-h||_\infty \\
&\leqslant \frac{q}{|\Im(z)|^4}\int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} dt||g-h||_\infty.
\end{aligned}$$

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

$$\sup_{x\in[0,1]} \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} dt < +\infty \tag{8.121}$$

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

$$\sup_{x\in[0,1]} |Tg(x) - Th(x)| \leqslant \frac{C}{|\Im(z)|^4}||g-h||_\infty \tag{8.122}$$

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 $K_z : [0,1] \to \mathbb{C}$ such that for all $x \in [0,1]$:

$$K_z(x) = \frac{1}{z - q\mathbb{E}\left[\left(z - \int_0^1 \left(\frac{\tau}{|t-x|}\right)_+^{\gamma^2} K_z(t)M(dt)\right)^{-1}\right]}. \tag{8.123}$$

Using the first equation, it is now plain to see that, for z such that $C/|\Im(z)|^4 < 1$, the constant μ_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 μ_z^2 is uniquely defined for all $z \in \mathbb{C} \setminus \mathbb{R}$. It will be easy to see using analyticity arguments. Indeed, from the Montel theorem, every limit point μ_z^2 is holomorphic on the set $\mathbb{C}\setminus\mathbb{R}$ 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 $\mathbb{C} \setminus \mathbb{R}$ (see Lemma 8.14). Thus, μ_z^2 is uniquely defined for each $z \in \mathbb{C} \setminus \mathbb{R}$ by analytic extension (we have just seen that μ_z^2 is uniquely defined for a set of z with accumulation points).

The same argument holds for the unicity of the integral $\int_0^1 K_z(x)dx$. Indeed, every limit point $\int_0^1 K_z(x)dx$ is a holomorphic function on $\mathbb{C} \setminus \mathbb{R}$ that has some prescribed value on the set $\{z \in \mathbb{C} \setminus \mathbb{R} : C/|\Im(z)|^4 < 1\}$, which has accumulation points.

8.5.7 Proof of Theorem 8.3, 8.4 and 8.5

Let us gather the above arguments to prove the main theorems.

Proof of theorem 8.5: it is a direct consequence of sections 8.5.4, 8.5.5 and 8.5.6.

Proof of theorem 8.4 i): The limit points $K_z(x)dx$ and $\mu_z^2 dx$ of the two complex measures $\mathbb{E}[L_N^{1,z}]$ and $\mathbb{E}[L_N^{2,z}]$ are uniquely defined because μ_z^2 and $K_z(x)$ satisfy a fixed point system of equations (we have just seen this in theorem 8.5).

Proof of theorem 8.4 iii): We need to prove that μ_z^2 is the Stieltjes transform of a probability measure υ. From [81], it suffices to prove that μ_z^2 is holomorphic over $\mathbb{C}\setminus\mathbb{R}$, maps $\{z\in\mathbb{C}\setminus\mathbb{R};\Im(z)<0\}$ to $\{z\in\mathbb{C}\setminus\mathbb{R};\Im(z)>0\}$ and that $\lim_{y\to\infty} iy\mu_{iy}^2 = 1$ $(y\in\mathbb{R})$. Let us check those properties. We have already seen in section 8.5.6 that μ_z^2 is holomorphic. From Lemma 8.14, μ^2 maps $\{z\in\mathbb{C}\setminus\mathbb{R};\Im(z)<0\}$ to $\{z\in\mathbb{C}\setminus\mathbb{R};\Im(z)>0\}$. Finally, from Theorem 8.5, we have

$$z\mu_z^2 = \mathbb{E}\left[\frac{1}{1-z^{-1}\int_0^1 K_z(x)\,M(dx)}\right].$$

As $|K_z(x)|\leqslant|\Im(z)|^{-1}$, the term $\int_0^1 K_z(x)\,M(dx)/z$ converges pointwise towards 0 when $z=iy$ and $y\to\infty$. Furthermore, from Lemma 8.18, we have $\Im(z)\Im(K_z(x)\leqslant 0$ in such a way that $\left|z-\int_0^1 K_z(x)\,M(dx)\right|^{-1}\leqslant|\Im(z)|^{-1}$. Therefore

$$\left|\frac{z}{z-\int_0^1 K_z(x)\,M(dx)}\right|\leqslant 1$$

when z takes on the form $z=iy$ $(y\in\mathbb{R})$. The dominated convergence theorem then implies that $\lim_{y\to\infty} iy\mu_{iy}^2=1$ and we can conclude μ^2 is indeed the Stieltjes transform of a (unique) probability measure υ.

Proof of theorem 8.3 i) and 8.4 ii) We observe that, for $z\in\mathbb{C}\setminus\mathbb{R}$:

$$A_N(z)\begin{pmatrix} zI_T & 0\\ X_N & zI_N\end{pmatrix} = \begin{pmatrix} z^2I_T - {}^tX_NX_N & -z{}^tX_N\\ 0 & z^2I_N\end{pmatrix}. \tag{8.124}$$

Let us rewrite the matrix $G_N(z)=A_N(z)^{-1}$ under the form:

$$G_N(z) = \begin{pmatrix} G_1(z) & {}^tG_{1,2}(z)\\ G_{1,2}(z) & G_2(z)\end{pmatrix}, \tag{8.125}$$

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 (8.124), we obtain:

$$\begin{pmatrix} I_T/z & 0\\ -X_N/z^2 & I_N/z\end{pmatrix}\begin{pmatrix} G_1(z) & {}^tG_{1,2}(z)\\ G_{1,2}(z) & G_2(z)\end{pmatrix} = \begin{pmatrix} (z^2I_T - {}^tX_NX_N)^{-1} & B\\ 0 & I_N/z^2\end{pmatrix} \tag{8.126}$$

where $B=(z^2I_T-{}^tX_NX_N)^{-1}{}^tX_N/z$.

It can be rewritten, using the fact that $-X_NG_1(z)+zG_{1,2}(z)=0$ and $-X_N{}^tG_{1,2}(z)+zG_2(z)=I_N$, as:

$$\begin{pmatrix} G_1(z)/z & {}^tG_{1,2}(z)/z\\ 0 & I_N/z^2\end{pmatrix} = \begin{pmatrix} (z^2I_T - {}^tX_NX_N)^{-1} & B\\ 0 & I_N/z^2\end{pmatrix} \tag{8.127}$$

Therefore, taking the trace we get:

$$\frac{1}{Tz}\sum_{k=1}^{T} G_N(z)_{kk} = \frac{1}{T}\mathrm{tr}(z^2I_T - {}^tX_NX_N)^{-1}, \tag{8.128}$$

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

$$\frac{1}{Tz}\sum_{k=1}^{T} G_N(z)_{kk} = \frac{1}{T}\mathrm{tr}(z^2I_N - X_N{}^tX_N)^{-1} + \frac{T-N}{Tz^2}. \tag{8.129}$$

Now, taking expectation and using theorem 8.4, we deduce:

$$\int_0^1 K_z(x)dx = qz \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\mathrm{tr}(z^2 I_N - X_N {}^tX_N)^{-1}\right] + \frac{1-q}{z} \tag{8.130}$$

Using the fact that (by (8.12)) the spectrum of B_N contains $2N$ eigenvalues which are the positive and negative square-roots of the spectrum of $R_N = {}^tX_N X_N$ plus $T-N$ zero eigenvalues and the fact that $1/(z-\lambda) + 1/(z+\lambda) = 2z/(z^2-\lambda^2)$, we can see that:

$$\frac{1}{N+T}\sum_{k=1}^{N+T} G_N(z)_{kk} = \frac{2z}{N+T}\mathrm{tr}(z^2 I_N - X_N {}^tX_N)^{-1} + \frac{T-N}{T+N}\frac{1}{z} \tag{8.131}$$

Using the relation 8.17 and theorem 8.4, it is easy to see that:

$$\lim_{N\to+\infty} \frac{1}{N+T}\sum_{k=1}^{N+T} \mathbb{E}[G_N(z)_{kk}] = \frac{1}{1+q}\left(q\mu_z^2 + \int_0^1 K_z(x)dx\right) \tag{8.132}$$

Taking expectation in 8.131 and using (8.132), we get:

$$\frac{1}{1+q}\left(q\mu_z^2 + \int_0^1 K_z(x)dx\right) = \frac{2qz}{1+q}\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\mathrm{tr}(z^2 I_N - X_N {}^tX_N)^{-1}\right] \tag{8.133}$$

$$+ \frac{1-q}{1+q}\frac{1}{z}. \tag{8.134}$$

From equations (8.130) and (8.133), we get the following relation:

$$\int_0^1 K_z(x)dx = q\mu_z^2 + \frac{1-q}{z}. \tag{8.135}$$

and theorem 8.4 ii). is proved.

With (8.135), (8.132) becomes:

$$\lim_{N\to+\infty} \frac{1}{N+T}\sum_{k=1}^{N+T} \mathbb{E}[G_N(z)_{kk}] = \frac{1}{1+q}\left(2q\mu_z^2 + \frac{1-q}{z}\right) \tag{8.136}$$

and, we note that the right hand side of (8.136) is the Stieltjes transform of the measure $2q/(1+q)\upsilon(dx) + (1-q)/(1+q)\delta_0(dx)$. Thus, the mean spectral measure $\mathbb{E}[\mu_{B_N}]$ converges weakly to the measure $2q/(1+q)\upsilon(dx) + (1-q)/(1+q)\delta_0(dx)$.

We have also:

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\mathrm{tr}(z^2 I_N - X_N {}^tX_N)^{-1}\right] = \frac{\mu_z^2}{z} \tag{8.137}$$

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

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\mathrm{tr}(z^2 I_N - X_N {}^tX_N)^{-1}\right] = \frac{1}{z}\int_{\mathbb{R}}\frac{\upsilon(dx)}{z-x} \tag{8.138}$$

$$= \int_{\mathbb{R}}\frac{\upsilon\circ(x^2)^{-1}(dx)}{z^2-x}. \tag{8.139}$$

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

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left[\mathrm{tr}(zI_N - X_N {}^tX_N)^{-1}\right] = \int_{\mathbb{R}}\frac{\upsilon\circ(x^2)^{-1}(dx)}{z-x}. \tag{8.140}$$

and thus, the probability measure $\mathbb{E}[\mu_{R_N}]$ converges weakly to the measure $\upsilon\circ(x^2)^{-1}(dx)$.

Proof of theorem 8.3 ii): using relation (8.17) and lemma 8.19, it is plain to check that $\int_{\mathbb{R}}(z-x)^{-1}\mu_{B_N}(dx)$ converges in probability to the Stieltjes transform of the probability measure

$2q/(1+q)\upsilon(dx)+(1-q)/(1+q)\delta_0(dx)$. This convergence holds for finite dimensional vectors $(\int_{\mathbb{R}}(z_i-x)^{-1}\mu_{B_N}(dx)), i=1,\dots,d)$ as well. Using the fact that the set of functions $\{(z-x)^{-1}, z\in\mathbb{C}\setminus\mathbb{R}\}$ 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 $\int f(x)\mu_{B_N}(dx)$ converges in probability to $\int f(x)(2q/(1+q)\upsilon(dx)+(1-q)/(1+q)\delta_0(dx))$. But, since $\mu_{B_N}(\mathbb{R})=2q/(1+q)\upsilon(\mathbb{R})+(1-q)/(1+q)\delta_0(\mathbb{R})=1$, this vague convergence can be strengthened in a weak convergence. With the relations $\mu_{B_N^2}=2N/(N+T)\mu_{R_N}+(T-N)/(T+N)\delta_0$ and the fact that $\int f(x)\mu_{B_N^2}(dx)=\int f(x^2)\mu_{B_N}(dx)$, it is plain to conclude that μ_{R_N} converges weakly in probability to $\upsilon\circ(x^2)^{-1}(dx)$.

Proof of theorem 8.3 iii): again using relation (8.17) and lemma 8.19 together with Borel-Cantelli's lemma, one can show that the two spectral measures $\mu_{B_{N_k}}$ converges weakly almost surely to $2q/(1+q)\upsilon(dx)+(1-q)/(1+q)\delta_0(dx)$. It is then easy to deduce as before that $\mu_{R_{N_k}}$ converges weakly almost surely to $\upsilon\circ(x^2)^{-1}(dx)$.

8.6 Auxiliary lemmas

Lemma 8.26. *Let A be a $n\times n$ complex matrix such that the Hermitian matrix $M=A\bar{A}^T$ has spectral radius λ_{max}. Then, for all i, we have:*

$$\sum_{j=1}^{n} \mid A_{ij} \mid^2 \leqslant \lambda_{max}. \tag{8.141}$$

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

$$M_{ii}=\sum_{j=1}^{n} \mid A_{ij} \mid^2 .$$

and, thus:

$$\sum_{j=1}^{n} \mid A_{ij} \mid^2 \leqslant \lambda_{max}. \tag{8.142}$$

□

Lemma 8.27. *There exists $C>0$ such that for each $N\in\mathbb{N}$ and $k\in\{1,\dots,N\}$:*

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

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

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

Proof. We first expand the square and use the independence of $(r_k(s))_s$ from $G_N^{(N+k)}(z)$:

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

Now we compute

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

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

$$\mathbb{E}\left[r_k(s)^2 r_k(t)^2\right] \leqslant \frac{\tau^{\psi(2)}(2^{2-\psi(2)}-2)}{(1-\psi(2))(2-\psi(2))}\frac{1}{N^{2-\psi(2)}}. \tag{8.143}$$

Thus we have for some positive constant C

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

where we have used the fact that almost surely:

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

It just remains to see that $\psi(2)=\gamma^2$. To prove the second relation, we follow the same argument by noticing that $(r_i(t))_t$ and $G_N^{(k,N+i)}(z)$ are independent conditionally to M^i. □

Lemma 8.28. *There exists some constant $c>0$ such that for each $N\in\mathbb{N}$ and $k\in\{1,\dots,N\}$:*

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

Proof. Again we expand the square and we use the fact that, conditionally to the $(M^i)_i$, 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 $(M^i)_i$, 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

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

where we have used the fact that almost surely:

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

□

Proof of Lemma 8.23. We define the function $f_N^{k,\epsilon}$ on the interval $[0,1]$ by

$$f_N^{k,\epsilon}(x) = NM^{k,\epsilon}(I_N^t) \text{ if } x\in I_N^t.$$

Notice the relation:

$$\sum_{t=1}^{N} M_\epsilon^k(I_N^t)\mathbb{E}\left[G_N(z)_{tt}\right] = \int_0^1 f_N^{k,\epsilon}(r)\, d\mathbb{E}[L_N^{1,z}](dr).$$

Then, by stationarity, we have:

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^1 f_N^{k,\epsilon}(r)\, d\mathbb{E}[L_N^{1,z}](dr) - \int_0^1 e^{\omega_\epsilon^k(r)}\, d\mathbb{E}[L_N^{1,z}](dr)\right|\right] \\
&\leqslant \sum_{t=1}^N \mathbb{E}\left[\left|\int_{I_N^t} (f_N^{k,\epsilon}(r) - e^{\omega_\epsilon^k(r)})\, d\mathbb{E}[L_N^{1,z}](dr)\right|\right] \\
&\leqslant \frac{N}{|\Im(z)|} \sup_{r\in I_N^1} \mathbb{E}\left[\left|\int_{I_N^1} (e^{\omega_\epsilon^k(u)} - e^{\omega_\epsilon^k(r)})\, du\right|\right] \\
&\leqslant \frac{N}{|\Im(z)|} \sup_{r\in I_N^1} \int_{I_N^1} \mathbb{E}\left[\left|e^{\omega_\epsilon^k(u)} - e^{\omega_\epsilon^k(r)})\right|^2\right]^{1/2} du \\
&\leqslant \frac{N}{|\Im(z)|} \sup_{r\in I_N^1} \int_{I_N^1} \left(2e^{\psi(2)\rho_\epsilon(0)} - 2e^{\psi(2)\rho_\epsilon(r-u)}\right)^{1/2} du.
\end{aligned}$$

Because of the continuity of the function ρ_ϵ over $[0,1]$, we have

$$\mathbb{E}\Big[\Big|\int_0^1 f_N^{k,\epsilon}(r)\, d\mathbb{E}[L_N^{1,z}](dr) - \int_0^1 e^{\omega_\epsilon^k(r)}\, d\mathbb{E}[L_N^{1,z}](dr)\Big|\Big] \to 0 \quad \text{as } N\to\infty. \tag{8.144}$$

In a quite similar way, we can prove that

$$\mathbb{E}\Big[\Big|\int_0^1 e^{\omega_\epsilon^k} * \phi_p(r)\, d\mathbb{E}[L_N^{1,z}](dr) - \int_0^1 e^{\omega_\epsilon^k(r)}\, d\mathbb{E}[L_N^{1,z}](dr)\Big|\Big] \to 0 \quad \text{as } p\to\infty \text{ uniformly w.r.t. N} \tag{8.145}$$

and

$$\mathbb{E}\Big[\Big|\int_0^1 e^{\omega_\epsilon^k} * \phi_p(r) K_z(r)\, dr - \int_0^1 e^{\omega_\epsilon^k(r)} K_z(r)\, dr\Big|\Big] \to 0 \quad \text{as } p\to\infty \text{ uniformly w.r.t. N} \tag{8.146}$$

where $(\phi_p)_{p\in\mathbb{N}}$ is a regularizing sequence and $*$ stands for the convolution. Furthermore, for each fixed p and because of the weak convergence of $\mathbb{E}[L_N^{1,z}]$ towards $K_z(x)dx$, we have almost surely

$$\int_0^1 e^{\omega_\epsilon^k} * \phi_p(r)\, d\mathbb{E}[L_N^{1,z}](dr) \to \int_0^1 e^{\omega_\epsilon^k} * \phi_p(r) K_z(r)\, dr \quad \text{as } N\to\infty. \tag{8.147}$$

We prove the result by gathering (8.144) (8.145) (8.146) and (8.147). □

8.7 Sup of MRW

Here we prove

Proposition 8.29. *We have for all $k = 1,\dots,N+1$*

$$\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4\right] \leqslant C\frac{(\ln N)^2}{N^{\frac{\zeta(2\alpha)-1}{\alpha}}}.$$

for some positive constant C.

Proof. To prove the result, we first prove

Lemma 8.30. *There exists a constant C such that, if $(X_i)_{1\leqslant i\leqslant N}$ are iid centered Gaussian random variables then:*

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

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

$$\begin{aligned}
\mathbb{E}\left[\max_{1 \leqslant i \leqslant N} |X_i|^4\right] &\leqslant \delta + N \int_\delta^\infty \mathbb{P}(|X_1|^4 > t)dt \\
&\leqslant \delta + 2N \int_\delta^\infty \mathbb{P}(X_1 > t^{1/4})dt \\
&\leqslant \delta + \frac{2N}{\sqrt{2\pi}} \int_\delta^\infty e^{-\sqrt{t}} dt \\
&\leqslant \delta + \frac{4N}{\sqrt{2\pi}} \int_{\sqrt{\delta}}^\infty e^{-t} t dt \\
&\leqslant \delta + \frac{4N}{\sqrt{2\pi}} \left(\sqrt{\delta} e^{-\sqrt{\delta}} + e^{-\sqrt{\delta}}\right),
\end{aligned}$$

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:

$$\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4\right] = \mathbb{E}\left[\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4 | M^k\right]\right].$$

Notice then that, conditionally to $M^k(0, \frac{1}{N}) = x_1, \dots, M^k(\frac{N-1}{N}, 1) = x_N$, the vector $(r_k(1), \dots, r_k(N)$ has the same law as the increments of B: $(B_{x_1} - B_0, \dots, B_{x_N} - B_{x_{N-1}})$. By applying Lemma 8.30, we deduce that

$$\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4 | M^k\right] \leqslant C(\ln N)^2 \max_{t=1,\dots,N} M^k\left(\frac{t-1}{N}, \frac{t}{N}\right)^2.$$

Thus we deduce

$$\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4\right] \leqslant C(\ln N)^2 \mathbb{E}\left[\left(\max_{t=1,\dots,N} M^k\left(\frac{t-1}{N}, \frac{t}{N}\right)\right)^2\right]. \tag{8.148}$$

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

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

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

$$\mathbb{E}\left[\sup_{t=1,\dots,N} r_k(t)^4\right] \leqslant (1 + C) \frac{(\ln N)^2}{N^{\frac{\zeta(2\alpha)-1}{\alpha}}} \tag{8.149}$$

□

8.8 Girsanov transform

Lemma 8.31. *Let μ be an independently scattered infinitely divisible random measure associated to (ψ, θ), where*

$$\forall q \in \mathbb{R}, \quad \psi(q) = mq + \frac{1}{2}\sigma^2 q^2 + \int_{\mathbb{R}} (e^{qz} - 1)\nu(dz),$$

$\psi(2) < +\infty$ and $\psi(1) = 0$. Let B be a bounded Borelian set. We define a new probability measure $\mathbb{P}_B$ (with expectation $\mathbb{E}_B$) by:

$$\forall A \text{measurable set}, \quad \mathbb{P}_B(A) = \mathbb{E}[\mathbf{1}_A e^{\mu(B)}].$$

Then, under $\mathbb{P}_B$, μ has the same law as $\mu+\mu_B$ where μ_B is an independently scattered infinitely divisible random measures independent of μ and is associated to (ψ_B,θ_B) given by

$$\psi_B(q) = q\sigma^2 + \int_{\mathbb{R}} (e^{qx}-1)(e^x-1)\nu(dx)$$
$$\theta_B(\cdot) = \theta(\cdot\cap B).$$

Proof. It suffices to compute the joint distribution of p disjoint sets $A_1,\dots,A_p$. We have for any $\lambda_1,\dots,\lambda_p\in\mathbb{R}$:

$$\begin{aligned}
\mathbb{E}_B\left[e^{\lambda_1\mu(A_1)+\cdots+\lambda_p\mu(A_p)}\right] &= \mathbb{E}\left[e^{\lambda_1\mu(A_1)+\cdots+\lambda_p\mu(A_p)+\mu(B)}\right]\\
&= \mathbb{E}\left[e^{\lambda_1\mu(A_1\setminus B)+\cdots+\lambda_p\mu(A_p\setminus B)+\lambda_1\mu(A_1\cap B)+\cdots+\lambda_p\mu(A_p\cap B)+\mu(B)}\right]\\
&= \mathbb{E}\left[e^{\lambda_1\mu(A_1\setminus B)+\cdots+\lambda_p\mu(A_p\setminus B)+(\lambda_1+1)\mu(A_1\cap B)+\cdots+(\lambda_p+1)\mu(A_p\cap B)+\mu(B\setminus\bigcup_{i=1}^n A_i)}\right]\\
&= \mathbb{E}\left[e^{\lambda_1\mu(A_1\setminus B)+\cdots+\lambda_p\mu(A_p\setminus B)}\right]\mathbb{E}\left[e^{(\lambda_1+1)\mu(A_1\cap B)+\cdots+(\lambda_p+1)\mu(A_p\cap B)}\right]\\
&= e^{\psi(\lambda_1)\theta(A_1\setminus B)+\cdots+\psi(\lambda_p)\theta(A_p\setminus B)}e^{\psi(\lambda_1+1)\theta(A_1\cap B)+\cdots+\psi(\lambda_p+1)\theta(A_p\cap B)}\\
&= e^{\psi(\lambda_1)\theta(A_1)+\cdots+\psi(\lambda_p)\theta(A_p)}e^{(\psi(\lambda_1+1)-\psi(\lambda_1))\theta(A_1\cap B)+\cdots+(\psi(\lambda_p+1)-\psi(\lambda_p))\theta(A_p\cap B)}.
\end{aligned}$$

Then it suffices to notice that:

$$\psi(q+1)-\psi(q) = m+\sigma^2 q+\frac{1}{2}\sigma^2+\int_{\mathbb{R}}(e^{(q+1)z}-e^{qz})\nu(dz)$$

and $\psi(1)=0$. □

Lemma 8.32. *If the process ω_ϵ is defined as $\omega_\epsilon(x)=\mu(A_\epsilon(x))$ where μ is an independently scattered random measure associated to (φ,θ) with $\varphi(q)=-iq\gamma^2/2-q^2\gamma^2/2$ and θ given by 8.8, then:*

$$\lim_{\epsilon\to 0}\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z-\int_0^1 K_z(r)e^{\omega_\epsilon(r)}dr}\right] = \mathbb{E}\left[\left(z-\int_0^1\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}K_z(r)M(dr)\right)^{-1}\right]$$

where M is the lognormal MRM.

Proof. One can check that $(\omega_\epsilon(x))_{x\in[0;1]}$ is a stationary gaussian process with covariance given by $\gamma^2\rho_\epsilon(x-y)$. So, using Girsanov transform, we can write:

$$\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z-\int_0^1 K_z(r)e^{\omega_\epsilon(r)}dr}\right] = \mathbb{E}\left[\left(z-\int_0^1 K_z(r)e^{\gamma^2\rho_\epsilon(r-x)}e^{\omega_\epsilon(r)}dr\right)^{-1}\right]$$

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

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(z-\int_0^1 K_z(r)e^{\gamma^2\rho_\epsilon(r-x)}e^{\omega_\epsilon(r)}\,dr\right)^{-1}\right]\right.\\
&\left.-\mathbb{E}\left[\left(z-\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}e^{\omega_\epsilon(r)}\,dr\right)^{-1}\right]\right|\\
&\leqslant \frac{1}{|\Im(z)|^2}\mathbb{E}\left[\int_0^1|K_z(r)|e^{\omega_\epsilon(r)}\left|e^{\gamma^2\rho_\epsilon(r-x)}-\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}\right|dr\right]\\
&\leqslant \frac{1}{|\Im(z)|^3}\int_0^1\left|e^{\gamma^2\rho_\epsilon(r-x)}-\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}\right|dr \qquad (8.150)
\end{aligned}$$

where we have used Lemmas 8.14 and 8.18 and the normalization $\psi(1)=0$.

Because $\gamma^2<1$, the dominated convergence theorem implies that 8.150 converges to 0 when ϵ goes to 0.

We thus look at the limit when ϵ goes to 0 of the term:

$$\mathbb{E}\left[\left(z-\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} e^{\omega_\epsilon(r)}\,dr\right)^{-1}\right].$$

The random variable

$$\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} M(dr)$$

is well defined and is finite almost surely since:

$$\mathbb{E}\left[\left|\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} M(dr)\right|\right] \leqslant \int_0^1 |K_z(r)|\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} dr < +\infty.$$

And thus, we can compute:

$$\begin{aligned}&\left|\mathbb{E}\left[\left(z-\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} e^{\omega_\epsilon(r)}\,dr\right)^{-1}\right]\right.\\&\left.-\mathbb{E}\left[\left(z-\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2} M(dr)\right)^{-1}\right]\right|\\&\leqslant \frac{1}{|\Im(z)|^2}\mathbb{E}\left[\left|\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}(e^{\omega_\epsilon(r)}dr-M(dr))\right|\right],\end{aligned}$$

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

$$\mathbb{E}\left[\left|\int_0^1 K_z(r)\left[\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}-\min\left(\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2},n\right)\right]e^{\omega_\epsilon(r)}dr\right|\right] \tag{8.151}$$

$$+\mathbb{E}\left[\left|\int_0^1 K_z(r)\min\left(\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2},n\right)(e^{\omega_\epsilon(r)}dr-M(dr))\right|\right] \tag{8.152}$$

$$+\mathbb{E}\left[\left|\int_0^1 K_z(r)\left[\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}-\min\left(\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2},n\right)\right]M(dr)\right|\right]. \tag{8.153}$$

The two quantities 8.151 and 8.153 are smaller than

$$\int_0^1 |K_z(r)|\left[\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2}-\min\left(\left(\frac{\tau}{|r-x|}\right)_+^{\gamma^2},n\right)\right]dr \tag{8.154}$$

and thus converge to 0, uniformly in ϵ as n goes to infinity.

For a fixed n, the function $\min((\tau/|r-x|)_+^{\gamma^2},n)$ is measurable and bounded and thus it is plain to see that, for a fixed n, the term 8.152 goes to 0 when ϵ goes to 0.

The lemma follows gathering the above estimates. □

Lemma 8.33. *If the process ω_ϵ is defined as $\omega_\epsilon(x)=\mu(A_\epsilon(x))$ where μ is an independently scattered random measure associated to (φ,θ) where φ is given by (8.6),i.e.*

$$\varphi(q)=imq-\frac{\gamma^2}{2}q^2+\int_{\mathbb{R}}(e^{iqx}-1)\nu(dx)$$

and where θ given by (8.8), then:

$$\lim_{\epsilon\to 0}\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z-\int_0^1 K_z(r)e^{\omega_\epsilon(r)}dr}\right]=\mathbb{E}\left[\left(z-\int_0^1\left(\frac{\tau}{|r-x|}\right)_+^{\kappa}K_z(r)Q(dr)\right)^{-1}\right]$$

with $\kappa=\gamma^2+\int_{\mathbb{R}}(e^x-1)^2\nu(dx)$ and where the random Radon measure Q is defined, conditionally on a MRM denoted by M whose structure exponent is $\zeta(q):=q-\varphi(-iq)$, as the almost sure

weak limit as ϵ goes to 0 of the family of random measures $Q_\epsilon(dt) := e^{\overline{\omega}_\epsilon(t)}M(dt)$ where, for each $\epsilon > 0$, the random process $\overline{\omega}_\epsilon$ is independent of M and defined as $\overline{\omega}_\epsilon(t) = \overline{\mu}(A_\epsilon(t))$ where $\overline{\mu}$ is the independently scattered log infinitely divisible random measure associated to $(\bar{\varphi}, \theta(\cdot \cap A_0(x)))$ where

$$\bar{\varphi}(p) = ip(\gamma^2 - \kappa) + \int_{\mathbb{R}} (e^{ipx} - 1)(e^x - 1)\nu(dx). \tag{8.155}$$

Proof. We want to apply Lemma 8.31 to the process ω_ϵ. If we set $B = A_\epsilon(x)$, Lemma 8.31 tells us that, under $\mathbb{P}_B$, the process ω_ϵ possesses the same law as the process

$$\omega_\epsilon^{(1)}(r) + \omega_\epsilon^{(2)}(r) \quad \text{with } \omega_\epsilon^{(1)}(r) = \mu^{(1)}(A_\epsilon(r)) \text{ and } \omega_\epsilon^{(2)}(r) = \mu^{(2)}(A_\epsilon(r)),$$

where $\mu_\epsilon^{(1)}, \mu_\epsilon^{(2)}$ are independent independently scattered log infinitely divisible random measures respectively associated to (φ, θ) and $(\varphi^{(2)}, \theta^{(2)})$ with:

$$\varphi^{(2)}(q) = i\gamma^2 q + \int_{\mathbb{R}} (e^{iqx} - 1)(e^x - 1)\nu(dx) \text{ and } \theta^{(2)}(\cdot) = \theta(\cdot \cap A_\epsilon(x)). \tag{8.156}$$

Define:

$$\kappa = \gamma^2 + \int_{\mathbb{R}} (e^x - 1)^2 \nu(dx), \overline{\varphi}(q) = \varphi^{(2)}(q) - iq\kappa, \overline{\psi}(q) = \overline{\varphi}(-iq). \tag{8.157}$$

Notice that $\overline{\psi}$ is then normalized so as to make $\overline{\psi}(1) = \overline{\psi}(0) = 0$. Let us define the process $\overline{\omega}_\epsilon$ by:

$$\overline{\omega}_\epsilon(r) = \omega_\epsilon^{(2)}(r) - \kappa\theta(A_\epsilon(r) \cap A_\epsilon(x)) = \omega_\epsilon^{(2)}(r) - \kappa\rho_\epsilon(r - x), \tag{8.158}$$

and notice that $\mathbb{E}[e^{iq\overline{\omega}_\epsilon(r)}] = e^{\overline{\varphi}(q)\rho_\epsilon(r-x)}$.

We can now apply Lemma 8.31:

$$\mathbb{E}\left[\frac{e^{\omega_\epsilon(x)}}{z - \int_0^1 K_z(r)e^{\omega_\epsilon(r)}dr}\right] = \mathbb{E}\left[\left(z - \int_0^1 K_z(r)e^{\overline{\omega}_\epsilon(r)+\kappa\rho_\epsilon(r-x)+\omega_\epsilon(r)}dr\right)^{-1}\right]$$

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

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(z - \int_0^1 e^{\overline{\omega}_\epsilon(r)+\omega_\epsilon(r)+\kappa\rho_\epsilon(r-x)} K_z(r)\, dr\right)^{-1}\right]\right. \\
&\left. - \mathbb{E}\left[\left(z - \int_0^1 e^{\overline{\omega}_\epsilon(r)+\omega_\epsilon(r)} \left(\frac{\tau}{|r-x|}\right)_+^\kappa K_z(r)\, dr\right)^{-1}\right]\right| \\
&\leqslant \frac{1}{|\Im(z)|^2}\mathbb{E}\left[\int_0^1 e^{\overline{\omega}_\epsilon(r)+\omega_\epsilon(r)} \left|e^{\kappa\rho_\epsilon(r-x)} - \left(\frac{\tau}{|r-x|}\right)_+^\kappa\right| |K_z(r)|\, dr\right] \\
&\leqslant \frac{1}{|\Im(z)|^3}\int_0^1 \left|e^{\kappa\rho_\epsilon(r-x)} - \left(\frac{\tau}{|r-x|}\right)_+^\kappa\right| dr
\end{aligned} \tag{8.159}$$

where we have used Lemmas 8.14 and 8.18, the normalizations $\overline{\psi}(1) = 0, \psi(1) = 0$ and the independence between $\overline{\omega}_\epsilon$ and ω_ϵ.

Let us show that $\kappa < 1$. Indeed, we have:

$$\begin{aligned}
\kappa &= \gamma^2 + \int_{\mathbb{R}} (e^x - 1)^2 \nu(dx) \\
&= \gamma^2 + \int_{\mathbb{R}} (e^{2x} - 1)\nu(dx) - 2\int_{\mathbb{R}} (e^x - 1)\nu(dx) \\
&= \gamma^2 + \int_{\mathbb{R}} (e^{2x} - 1)\nu(dx) + 2(m + \frac{1}{2}\gamma^2) \\
&= 2m + 2\gamma^2 + \int_{\mathbb{R}} (e^{2x} - 1)\nu(dx) \\
&= \psi(2)
\end{aligned}$$

where, in the third line, we used the fact that $\psi(1) = 0$ (which implies the relation $\int_{\mathbb{R}}(e^x - 1)\nu(dx) = -(m + \gamma^2/2)$). We will now show that $\psi(2)$ is strictly less than 1. It suffices to show that $\zeta(2) > 1$. Using the concavity of the function ζ, we have the inequality:

$$\frac{\zeta(2+\epsilon) - \zeta(1)}{1+\epsilon} < \zeta(2) - \zeta(1) \tag{8.160}$$

and with assumption 8.24, we see that $\zeta(2) - \zeta(1) = \zeta(2) - 1 > 0$. We can thus conclude that $\kappa < 1$.

Because $\kappa < 1$, the dominated convergence theorem implies that 8.159 converges to 0 when ϵ goes to 0.

For each Borelian set A of $[0;1]$, the family $M_\epsilon(A) := \int_A e^{\omega_\epsilon(r)}dr, \epsilon > 0$ is a positive martingale with respect to ϵ and that it converges almost surely to $M(A)$. With the assumption 8.24 and in particular the condition $\zeta(2+\epsilon) > 1$, we can show (see [**?**] for a proof) that the family $(M_\epsilon(A))_{\epsilon>0}$ is in fact uniformly integrable. In particular, if we let $\mathcal{F}_\epsilon$ be the sigma field generated by the family of random variables $(\omega_\eta(r))_{\eta>\epsilon, r\in\mathbb{R}}$, we have the following almost sure equality:

$$\mathbb{E}\left[M(A)|\mathcal{F}_\epsilon\right] = M_\epsilon(A). \tag{8.161}$$

Conditionally to the random measure M, the family $P_\epsilon(A) := \int_A e^{\overline{\omega}_\epsilon(r)}M(dr), \epsilon > 0$ is also a positive martingale with respect to ϵ. Thus, $P_\epsilon(A)$ converges almost surely to a random variable that we will denote by $P(A)$. We know that this defines a random Radon measure P on $[0;1]$ and that the family of random Radon measures P_ϵ converges, when ϵ goes to 0, weakly almost surely to P in the space of Radon measures. Denote, conditionally to the random measure M, by $\mathbb{P}_M$ the law $\mathbb{P}[\cdot|M]$ and let us show that the family $(P_\epsilon([0;1]))_{\epsilon>0}$ is $\mathbb{P}_M$-uniformly integrable. Let $\overline{\delta}$ be such that $\overline{\psi}(1+\overline{\delta}) < +\infty$ (we can show, using the condition $\psi(2+\delta) < +\infty$, that that there exists such $\overline{\delta}$). We will show that the family $(P_\epsilon([0;1]))_{\epsilon>0}$ is uniformly bounded in $L^{1+\overline{\delta}}(\mathbb{P}_M)$. Indeed, conditionally to the random measure M:

$$\begin{aligned}
\mathbb{E}_M\left[\left(\int_0^1 e^{\overline{\omega}_\epsilon(r)}M(dr)\right)^{1+\overline{\delta}}\right] &\leqslant \mathbb{E}_M\left[\int_0^1 e^{(1+\overline{\delta})\overline{\omega}_\epsilon(r)}M(dr)\right] M[0;1]^{\overline{\delta}} \\
&\leqslant \int_0^1 e^{\overline{\psi}(1+\overline{\delta})\rho_\epsilon(r-x)}M(dr)M[0;1]^{\overline{\delta}} \\
&\leqslant M[0;1]^{\overline{\delta}}e^{\overline{\psi}(1+\overline{\delta})}\int_0^1 \left(\frac{\tau}{|r-x|}\right)_+^\kappa M(dr) < +\infty.
\end{aligned}$$

The family $(P_\epsilon([0;1]))_{\epsilon>0}$ is therefore $\mathbb{P}_M$-uniformly integrable, in particular, $P_\epsilon([0;1])$ converges to $P([0;1])$ also in L^1, which implies that P is a non degenerated random measure. Moreover, denoting by $\overline{\mathcal{F}}_\epsilon$ the sigma field generated by the family of random variables $(\overline{\omega}_\eta(r))_{\eta>\epsilon, r\in\mathbb{R}}$, we have, almost surely, conditionally to M, for all Borelian set A of $[0;1]$:

$$\mathbb{E}_M\left[P(A)|\mathcal{F}_\epsilon\right] = P_\epsilon(A).$$

Now, as before, it is easy to see that the family $Q_\epsilon(A) := \int_A e^{\omega_\epsilon(r)+\overline{\omega}_\epsilon(r)}dr, \epsilon > 0$ is also a positive martingale with respect to ϵ. Therefore, $Q_\epsilon(A)$ converges almost surely to a random variable that we will denote by $Q(A)$. This defines a random Radon measure Q and the family of random Radon measure Q_ϵ converges, as $\epsilon \to 0$, weakly almost surely to Q in the space of Radon measure. We want to show that the two random measures P and Q have the same law.

Gathering the above arguments, we can write, almost surely:

$$\begin{aligned}
\mathbb{E}\left[P(A)|\sigma(\mathcal{F}_\epsilon, \overline{\mathcal{F}}_\epsilon)\right] &= \mathbb{E}\left[\mathbb{E}[P(A)|\overline{\mathcal{F}}_\epsilon]\right] \\
&= \mathbb{E}\left[\int_A e^{\overline{\omega}_\epsilon(r)}M(dr)|\mathcal{F}_\epsilon\right] \\
&= \int_A e^{\omega_\epsilon(r)+\overline{\omega}_\epsilon(r)}dr,
\end{aligned}$$

and the latter quantity has the same law as $Q_\epsilon(A)$. Since the martingale $(\mathbb{E}[P(A)|\sigma(\mathcal{F}_\epsilon, \overline{\mathcal{F}}_\epsilon)])_{\epsilon>0}$ is uniformly integrable, we deduce that the family $(Q_\epsilon(A))_{\epsilon>0}$ is also uniformly integrable. Hence, both random variables $P(A)$ and $Q(A)$ have the same law. We can show easily that in fact the two random measures P and Q have the same law. In particular, Q is non degenerated.

It is now easy to see that, for all bounded and continuous function f, the two random variables $\int_{\mathbb{R}} f(r)P(dr)$ and $\int_{\mathbb{R}} f(r)Q(dr)$ have the same law. By regularizing the function $\left(\frac{\tau}{|r-x|}\right)_+^\kappa$ and with the dominated convergence theorem, we conclude as in the proof of lemma 8.32 using the fact that $\kappa < 1$ that:

$$\int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^\kappa Q(dr) \overset{(law)}{=} \int_0^1 K_z(r)\left(\frac{\tau}{|r-x|}\right)_+^\kappa P(dr). \tag{8.162}$$

Gathering the above argument and letting ϵ go to 0 concludes the proof. □

Chapter 9

Principal Regression Analysis and the index leverage effect

Résumé

Cet article est publié dans *Physica A* et est écrit en collaboration avec Jean-Philippe Bouchaud et Pierre-Alain Reigneron. Nous revisitons l'effet Levier pour l'indice, qui peut être décomposé en un effet sur la volatilité et un effet sur les corrélations. Nous nous intéressons à ce dernier en utilisant une régression linéaire matricielle, que nous appelons analyse en régression principale et pour lequel nous donnons un cadre de travail (grâce à la théorie des matrices aléatoires) théorique et numérique. Nous trouvons que les rendements négatifs passés ont pour effet d'augmenter la corrélation moyenne entre les prix des actions mais ont tendance à éloigner le mode marché du vecteur uniforme. Il y a deux échelles de temps associées à cet effet, une échelle courte de l'ordre d'un mois (correspondant à 20 jours d'échange sur les marchés) et une plus longue de l'ordre de un an. Nous trouvons aussi des traces d'un effet levier pour les secteurs, qui se révélent par les deuxième et troisième modes de la matrice issue de l'analyse en régression principale.

Abstract

We revisit the index leverage effect, that can be decomposed into a volatility effect and a correlation effect. We investigate the latter using a matrix regression analysis, that we call 'Principal Regression Analysis' (PRA) and for which we provide some analytical (using Random Matrix Theory) and numerical benchmarks. We find that downward index trends increase the average correlation between stocks (as measured by the most negative eigenvalue of the conditional correlation matrix), and makes the market mode more uniform. Upward trends, on the other hand, also increase the average correlation between stocks but rotates the corresponding market mode *away* from uniformity. There are two time scales associated to these effects, a short one on the order of a month (20 trading days), and a longer time scale on the order of a year. We also find indications of a leverage effect for sectorial correlations as well, which reveals itself in the second and third mode of the PRA.

9.1 Introduction

Among the best known stylized facts of financial markets lies the so-called "leverage effect" [82, 30, 118, 48, 119, 120], a name coined by Black to describe the negative correlation between past price returns and future realized volatilities in stock markets [38]. [1] It is indeed well documented that negative price returns induce increased future volatilities, an effect responsible for the observed skew on the implied volatility smile in stock option markets (see e.g. [35, 57]).

[1] While this effect holds for most markets in developed economies, Tenenbaum et al. [144] report that the situation appears to be different for markets in developing countries.

However, the association, made by Black, with a true leverage effect (i.e. that when the value of a stock goes down its debt to equity ratio increases, thereby making the company riskier and more volatile), is probably misleading. In particular, the amplitude of the leverage correlation for indices is noticeably stronger than for individual stocks, which even sounds paradoxical when the index return is by definition the average of individual stock returns! The volatility of an index in fact reflects both the volatility of underlying single stocks and the average correlation between these stocks. The increased leverage effect for indices must therefore mean that both these quantities are sensitive to a downward move of the market.

The aim of the present paper is to investigate more specifically this "correlation leverage effect", and make precise the common lore according to which correlations "jump to one" in crisis periods (see [70, 104, 141, 126] for early studies of the time evolution of the correlations in financial markets). Similar studies have appeared recently. In [27], a careful study of the average correlation between stock returns during contemporaneous upward/downward trends of the market index has confirmed that correlations are indeed stronger when the market goes down [40]. Our analyses confirm and make more precise these results, first by extending them to different markets, and second by devising and exploiting a new tool to investigate conditional correlations, that we call "principal regression analysis" (PRA). The idea here is to regress the instantaneous correlation matrix on the value of the index return (or any other conditioning variable). While the intercept of the regression gives the average correlation matrix, the regression slopes define a second symmetric (but not definite positive) matrix that can be diagonalized, leading to modes (eigenvectors) of sensitivity to the conditioning variable(s). The interpretation of these eigenvectors is particularly transparent when they coincide with those of the correlation matrix itself. The corresponding eigenvalues quantify how the whole correlation structure of stock returns is affected by the conditioning variable. The nice point about the PRA is that Random Matrix Theory (RMT) provides, as for standard PCA, a useful guide to decide whether or not these sensitivity modes are statistically meaningful (for a review on RMT, see [44]). When the conditioning variable is the past values of the index return, the conclusion of PRA is that the dominant mode is the market mode, associated to a negative eigenvalue, indeed corresponding to a correlation leverage effect. We characterize the temporal decay of this effect. Upon separating positive and negative index returns, we furthermore find that the correlation leverage effect is strongly asymmetric: whereas negative returns increase both the volatility of the underlying stocks and the average correlation between stocks, positive returns have weaker influence on these quantities (see Fig. 6 below). We furthermore find indications of a leverage effect for sectorial correlations as well, which reveals itself in the second and third modes of the PRA.

9.2 Data, notations and definitions

We have considered 6 pools of stocks corresponding to 6 major stock indices: SP500, BE500, Nikkei, FTSE, CAC 40 and DAX. We analyze the daily returns in a time period spanning from 01/01/2000 to 04/26/2010. Stocks are labelled by $\alpha = 1, \ldots, N$ (where N depends on the market), and days by $t = 1, \ldots, T$ (where $T = 2594$). Time average will be denoted by $\langle . \rangle$. The return of stock α between the close of day $t-1$ and the close of day t is denoted as $\eta_\alpha(t)$. We in fact understand $\eta_\alpha(t)$ as the *demeaned* return over the whole time period T. We define an inverse volatility weighted index return at time t as:

$$I(t) = \frac{1}{N} \sum_{\alpha=1}^{N} \widehat{\eta}_\alpha(t), \qquad \widehat{\eta}_\alpha(t) \equiv \frac{\eta_\alpha(t)}{\sigma_\alpha}, \tag{9.1}$$

where σ_α is the average volatility of the stock α over the whole time period:

$$\sigma_\alpha^2 := \frac{1}{T} \sum_{t=1}^{T} \eta_\alpha(t)^2. \tag{9.2}$$

We will further define the average instantaneous stock volatility $\sigma(t)$ at time t as:

$$\sigma(t)^2 := \frac{1}{N}\sum_{\alpha=1}^{N} \widehat{\eta}_\alpha(t)^2 \tag{9.3}$$

while the average instantaneous correlation between all pairs of stocks $\rho(t)$ is defined as:

$$\rho(t) := \frac{1}{N(N-1)} \sum_{\alpha\neq\beta=1}^{N} \frac{\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t)}{\sigma(t)^2}. \tag{9.4}$$

The average over time of the above two quantities will be denoted as σ_0^2 and ρ_0.

The squared index return $I(t)^2$ is a rough proxy for the instantaneous index volatility. Using the above definitions and the fact that N is large, it is easy to check that:

$$I(t)^2 \approx \rho(t)\sigma(t)^2 + O(\frac{1}{N}), \tag{9.5}$$

showing that both the average stock volatility and the average correlation contribute to the index volatility. It is therefore natural to decompose the full index leverage effect in two contributions: one coming from the dependence of the average stock volatility on the past returns of the index, and a second one describing the average correlation. We thus define a full leverage correlation function $\mathcal{L}_I(\tau)$:

$$\mathcal{L}_I(\tau) = \frac{\langle I(t-\tau)I(t)^2\rangle}{\langle I(t)^2\rangle}, \tag{9.6}$$

and two partial leverage correlation functions:

$$\mathcal{L}_\sigma(\tau) = \frac{\langle I(t-\tau)\sigma(t)^2\rangle}{\langle I(t)^2\rangle}, \qquad \mathcal{L}_\rho(\tau) = \frac{\langle I(t-\tau)\rho(t)\rangle}{\langle I(t)^2\rangle}. \tag{9.7}$$

All the above leverage correlation functions are normalized to be the regression slope of the corresponding observables on the past value of the index return, for example:

$$\rho(t) = \rho_0 + \mathcal{L}_\rho(\tau)I(t-\tau) + \varepsilon(t,\tau), \tag{9.8}$$

where $\varepsilon(t,\tau)$ is some noise. (Remember that by construction, $I(t)$ has zero mean.)

In the limit of weak correlations, the two effects are additive and one should find:

$$\mathcal{L}_I(\tau) \approx \rho_0\mathcal{L}_\sigma(\tau) + \sigma_0^2\mathcal{L}_\rho(\tau), \tag{9.9}$$

eliciting the contribution of the average stock volatility and of the average correlation to the full leverage correlation. The second term is responsible for the enhanced leverage effect for indices compared to single stocks.

9.3 Index leverage effect: A simple empirical analysis

As a first stab at understanding the index leverage effect, we plot in Fig. 9.1 the normalized partial leverage correlation functions, $\rho_0\mathcal{L}_\sigma(\tau)$, $\sigma_0^2\mathcal{L}_\rho(\tau)$, together with the full leverage $\mathcal{L}_I(\tau)$. In these plots, the data is averaged over the four indices, SP500, BE500, Nikkei and FTSE. From this figure, we draw the following conclusions:

- (a) the two contributions to the index leverage are of the same order of magnitude. In particular, the correlation leverage is significant and confirms the conclusions of Refs. [27, 40].
- (b) the correlation effect is stronger at short times but decays faster than the volatility effect; a two time scale exponential fit of these two contributions in the range $\tau \in [1, 250]$ (in days) indeed leads to

$$\sigma_0^2\mathcal{L}_\rho(\tau) \approx -0.053\exp(-\tau/18) - 0.005\exp(-\tau/350); \tag{9.10}$$
$$\rho_0\mathcal{L}_\sigma(\tau) \approx -0.02\exp(-\tau/14) - 0.02\exp(-\tau/280), \tag{9.11}$$

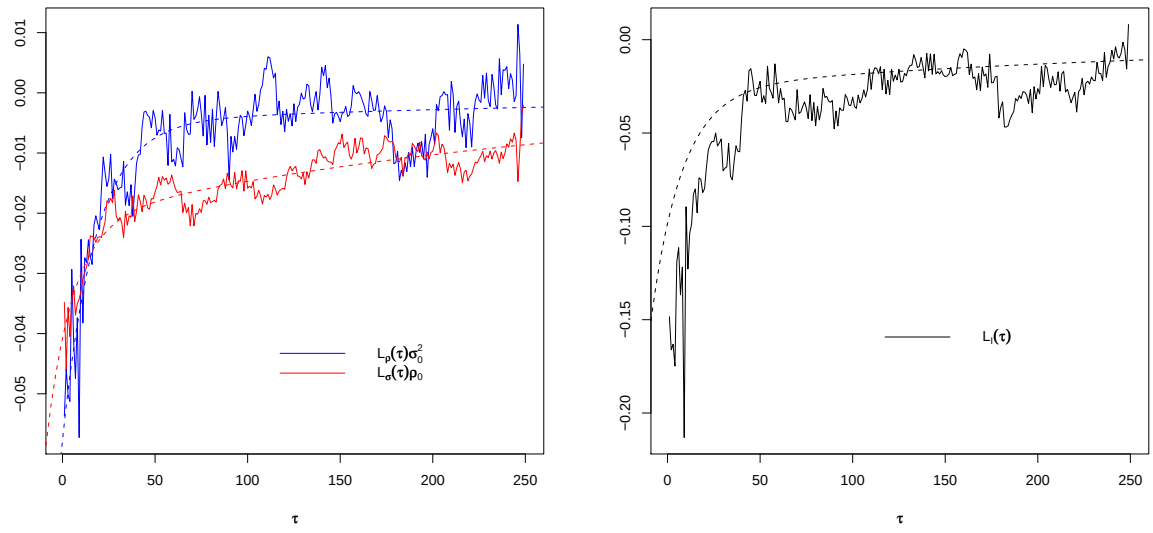


Figure 9.1: Left: normalized leverage correlation functions $\rho_0 \mathcal{L}_\sigma(\tau)$, $\sigma_0^2 \mathcal{L}_\rho(\tau)$, and an exponential fits with two scales (dotted lines). Right: Full leverage function $\mathcal{L}_I(\tau)$ and comparison with an additive model (dotted line).

- (c) a test of Eq. (9.9) with the sum of the above two fitted exponentials reproduces satisfactorily the full leverage effect, although the latter is underestimated at short times, when the correlations cease to be small enough for Eq. (9.9) to be accurate.

In fact, one can test directly whether linear regressions such as Eq. (9.8) above make sense or not, by averaging all values of $\rho(t)$ corresponding to a given value of $I(t-1)$ within some range. The resulting graphs are shown in Fig. 10.2, both for ρ and for σ^2. One sees that whereas a linear regression for ρ makes sense for $I(t-1) < 0$, there is in fact perhaps a small positive slope for $I(t-1) > 0$. For σ^2, the graph looks even more symmetric, reflecting the presence of volatility correlations on top of (asymmetric) leverage correlations.

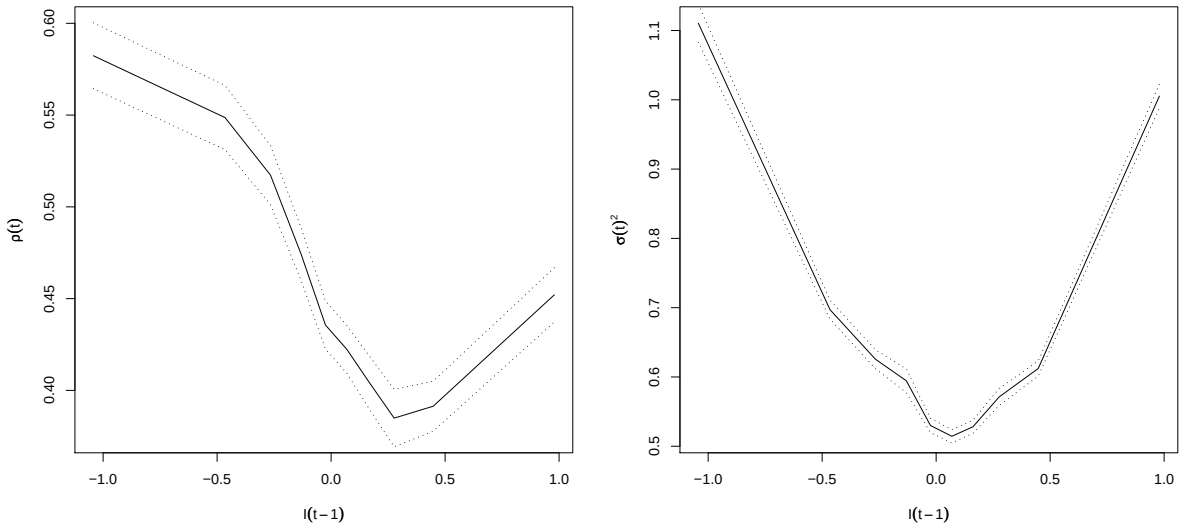


Figure 9.2: Dependence of the average correlation $\rho(t)$ and the average single stock volatility $\sigma^2(t)$ on the index return the previous day, $I(t-1)$. The result is obtained as an average over all 6 indices: SP500, BE500, Nikkei, FTSE, CAC40 and DAX, but the qualitative effects are robust and appear on each markets individually. These plots suggest that a quadratic $I^2(t-1)$ term should be included to the linear regressions. The printed error bars are the average of the error bars obtained for each of the 6 indices.

9.4 The "Principal Regression Analysis"

The above analysis, although interesting, is oversimplified, because the structure of inter-stock correlations is described by a full correlation matrix $\mathbf{C}$ and not by a single number ρ, that only captures the average correlations. In order to characterize the way the correlation matrix depends on the past value of the index (or on any other conditioning variable), we propose the following: consider a given pair of stocks, α, β, and regress the product of normalized returns $\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t)$ on the past value of the index return, i.e. write:

$$\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t) := C_{\alpha,\beta} + D_{\alpha,\beta}(\tau)I(t-\tau) + \varepsilon_{\alpha,\beta}(t,\tau). \tag{9.12}$$

Since $I(t)$ has zero mean, the intercept of the regression is exactly the empirical Pearson estimate of the correlation matrix. The regression *slopes* $D_{\alpha,\beta}(\tau)$ define another $N \times N$ symmetric matrix $\mathbf{D}(\tau)$, which encodes the full information about the dependence of the correlations on past returns. More precisely, the regression leads to the following empirical determination of $\mathbf{D}(\tau)$:

$$\langle I^2 \rangle D_{\alpha,\beta}(\tau) = \frac{1}{(T-\tau)} \sum_{t=\tau+1}^{T} \widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t)I(t-\tau). \tag{9.13}$$

The aim of this section is first to discuss the information contained in $\mathbf{D}(\tau)$, in particular its eigenvalues and eigenvectors, and second to use results from Random Matrix Theory to assess how meaningful this information is when the length of the sample, T, is not very large compared to the number of stocks N. Finally, we describe our empirical results on $\mathbf{D}(\tau)$, in particular its most negative eigenvalue and eigenvectors.

9.4.1 Interpretation

Define $\mathbf{C}(I)$ to be the correlation matrix conditioned to a certain past value of I, by:

$$\mathbf{C}(I) = \mathbf{C} + I\mathbf{D}. \tag{9.14}$$

The interpretation of the matrix $\mathbf{D}$ is particularly simple when it commutes with the correlation matrix $\mathbf{C}$, i.e. when the eigenvectors of $\mathbf{D}$ are the same as those of $\mathbf{C}$. In this case, the eigenvectors of $\mathbf{C}(I)$ are exactly the same as those of $\mathbf{C}$, whereas the eigenvalues $\lambda_k(I)$ are shifted as:[2]

$$\lambda_k(I) = \lambda_k(0) + I\langle v_k|\mathbf{D}|v_k\rangle, \tag{9.15}$$

where $\lambda_k(0)$ are the eigenvalues of $\mathbf{C}$ and $|v_k\rangle$ are the associated eigenvectors (in quantum mechanics notations). When $\mathbf{D}$ does not commute with $\mathbf{C}$, the structure of the eigenvectors themselves is impacted by the conditioning variable. If $\mathbf{D}I$ is small enough, standard first order perturbation theory gives back Eq. (9.15) for the eigenvalues and:

$$|v_k(I)\rangle = |v_k\rangle + I\sum_{\ell \neq k} \frac{\langle v_\ell|\mathbf{D}|v_k\rangle}{\lambda_k - \lambda_\ell}|v_\ell\rangle, \tag{9.16}$$

for the eigenvectors of the matrix $\mathbf{C}(I)$.

As we will find below, the eigenvector corresponding to the most negative eigenvalue of $\mathbf{D}$ turns out to be very close to the first eigenvector of $\mathbf{C}$ (i.e. the so-called market mode, $|v_1\rangle$), whereas all other eigenvalues are significantly smaller. In this case, the top eigenvalue of $\mathbf{C}$ is to a good approximation given by:

$$\lambda_1(I) \approx \lambda_1 + I\mu_1, \tag{9.17}$$

where μ_1 is the most negative eigenvalue of $\mathbf{D}$. Since λ_1 can be used to define the average correlation between stocks through $\lambda_1 := N\rho$, the meaning of μ_1 is similar to, but more precise than, the correlation leverage function $\mathcal{L}_\rho$ defined above.

[2] Note that the dependence on the lag τ is implied in the following formulas.

More generally, when $\mathbf{D}$ and $\mathbf{C}$ do not commute, one expects the "correlation leverage" to rotate the top eigenvector away from the market mode $|v_1\rangle$. The common lore is indeed that when markets go down, all stocks "move together", meaning that the top eigenvector should rotate towards the uniform vector $|e\rangle = (1/\sqrt{N}, 1/\sqrt{N}, \ldots, 1/\sqrt{N})$. The cosine of the angle between $|v_1\rangle$ and $|e\rangle$ is given by the scalar product $\langle e|v_1\rangle$, that one can compute using perturbation theory. Eq. (9.16) above. Assuming further that the top eigenvalue of $\mathbf{C}$ is much larger than all the others ($\lambda_1 \gg \lambda_{\ell\neq 1}$), one finds:

$$\langle e|v_1(I)\rangle \approx \langle e|v_1\rangle + \frac{I}{\lambda_1}\left[\langle e|\mathbf{D}|v_1\rangle - \langle v_1|\mathbf{D}|v_1\rangle\langle e|v_1\rangle\right]. \tag{9.18}$$

A measure of how strongly the top eigenvector moves towards $|e\rangle$ is therefore provided by the quantity Δ, defined as:

$$\Delta = \frac{1}{\lambda_1}\left[\langle e|\mathbf{D}|v_1\rangle - \langle v_1|\mathbf{D}|v_1\rangle\langle e|v_1\rangle\right]. \tag{9.19}$$

A negative Δ means that the instantaneous market mode is closer to the uniform mode $|e\rangle$ when the index goes down, since $\langle e|v_1(I)\rangle - \langle e|v_1\rangle = I\Delta > 0$.

9.4.2 Results from Random Matrix Theory

When N is large, the simultaneous determination – using Eq. (9.13) above – of the $N(N+1)/2$ different elements of $\mathbf{D}$ from the NT data points is problematic, exactly in the same way the correlation matrix $\mathbf{C}$ is hard to measure. We thus need to provide a benchmark to compare the empirical results obtained with the noise level of the benchmark case. This will enable to separate significant effect from noise level arising from the dimensionality problem. Let ξ be a random variable which will play the role of the conditioning variable (the past values of index returns in our context) and let $x_\alpha, \alpha = 1, \ldots, N$ be a gaussian vector of covariance matrix $\mathbf{C}$ which should be seen as instantaneous stock returns. The x_α will be supposed to have 0 mean and unit variance, so that $\mathbf{C}$ is the correlation matrix of the gaussian vector $(x_1, \ldots, x_N)$.

We begin by the case $\mathbf{C} = \mathbf{I}$. Suppose, in addition, that there is *no correlations whatsoever* between the conditioning variable ξ and the correlation $x_\alpha x_\beta$, and that one forms a matrix $\widetilde{\mathbf{D}}$ from:

$$\langle \xi^2 \rangle \widetilde{D}_{\alpha,\beta} = \frac{1}{T}\sum_{t=1}^{T} x_\alpha(t) x_\beta(t) \xi(t). \tag{9.20}$$

In the limit $T \to \infty$ for finite N one should find that all the elements of the matrix $\widetilde{\mathbf{D}}$ are zero, and therefore all its eigenvalues are zero as well. For finite T, however, the matrix $\widetilde{\mathbf{D}}$ will have a set of non trivial eigenvalues. Random Matrix Theory offers a way to compute the statistics of these eigenvalues when N and T are both large, with a fixed ratio $q = N/T$. The result depends both on the eigenvalue spectrum of the matrix $\mathbf{C}$ and, perhaps surprisingly, on the probability distribution of the conditioning variable, $P(\xi)$. The simplest, albeit unrealistic case for applications in finance, is when $\mathbf{C}$ is the identity matrix, i.e. there is no correlations between the $\widehat{\eta}$. In this case, using the theory of Free Random Matrices [145], one finds that the empirical eigenvalue spectrum of $\widetilde{\mathbf{D}}$, $\rho_1(\mu)$, is the solution of the following set of equations, in the limit where ϵ goes to zero: [44, 37]

$$\mu = \frac{G_R}{G_R^2 + \pi^2\rho_1^2} + \int \mathrm{d}\xi P(\xi) \frac{\xi(1 - q\xi G_R)}{(1 - q\xi G_R)^2 + (q\pi\xi\rho_1)^2} \tag{9.21}$$

$$\epsilon = \rho_1 \left(-\frac{1}{G_R^2 + \pi^2\rho_1^2} + \int \mathrm{d}\xi P(\xi) \frac{q\xi^2}{(1 - q\xi G_R)^2 + (q\pi\xi\rho_1)^2} \right), \tag{9.22}$$

where G_R is the real part of the resolvent. One can check that in the limit $q \to 0$, and using the fact that ξ has zero mean, the above equations boil down to:

$$\frac{1}{G_R - i\pi\rho_1} = \mu - i\epsilon \to \rho_1(\mu) = \delta(\mu), \tag{9.23}$$

i.e. all eigenvalues are zero, as they indeed should when $T \gg N$.

The case of an arbitrary correlation matrix $\mathbf{C}$ can also be solved completely using the above result on ρ_1 and the so-called S-transform of the eigenvalue spectrum [145], noting that the eigenvalues of $\widetilde{\mathbf{D}}$ are the same as those of the product $\mathbf{C} \times \mathbf{D}_1$, where $\mathbf{D}_1$ is a random matrix with eigenvalue spectrum $\rho_1(\mu)$. The resulting equation can in principle be solved numerically for any value of q and for an arbitrary correlation matrix $\mathbf{C}$. The resulting theoretical eigenvalue spectrum for the matrix $\widetilde{\mathbf{D}}$, assuming no correlation between the conditioning variable ξ and the instantaneous correlation $x_\alpha x_\beta$, can be compared to the empirical spectrum obtained from data using Eq. (9.13). Any difference between the two spectra can be interpreted as resulting from a true correlation with the conditioning variable.

In the null-hypothesis case, it is also clear that the quantity $\widetilde{\Delta}$ defined by:

$$\widetilde{\Delta} = \frac{1}{\lambda_1}\left[\langle e|\widetilde{\mathbf{D}}|v_1\rangle - \langle v_1|\widetilde{\mathbf{D}}|v_1\rangle\langle e|v_1\rangle\right]. \tag{9.24}$$

must be zero when averaged over ξ, x_α. One can compute its variance, which is found to be:

$$\langle\widetilde{\Delta}^2\rangle_{\xi,x_\alpha} = \frac{\langle e|\mathbf{C}|e\rangle - \lambda_1\langle e|v_1\rangle^2}{T\lambda_1}\langle\xi^2\rangle. \tag{9.25}$$

For large T, the central limit theorem ensures that $\widetilde{\Delta}$ becomes Gaussian with the above variance. This result will be used below to assess whether the empirical value of Δ (defined above) is meaningful or not.

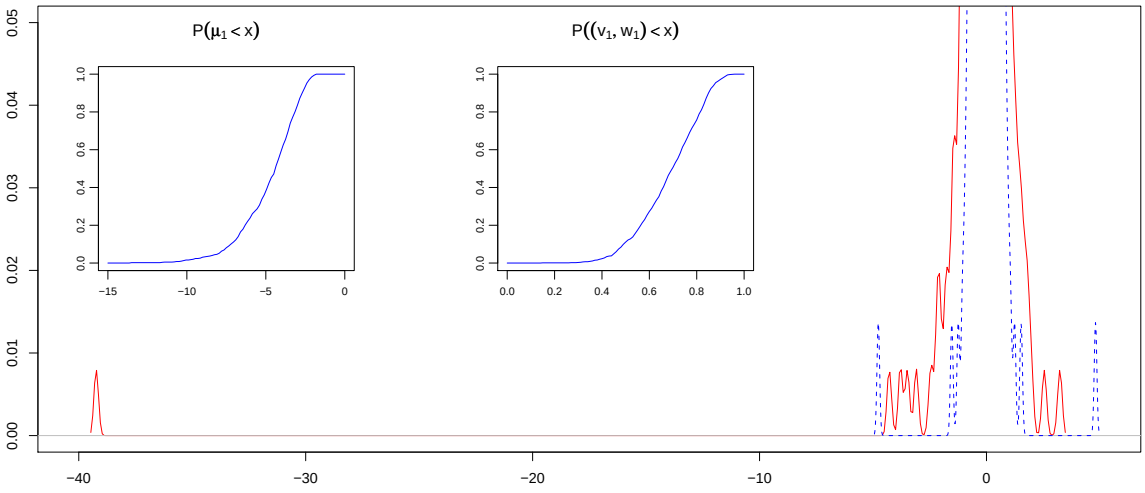


Figure 9.3: Main figure: empirical spectrum of $\mathbf{D}$ for the BE500 index (in red), compared to the null-hypothesis case (in blue). For the latter case, we have generated 1000 random samples, ranked the eigenvalues and averaged each of them separately. The leftmost blue peak therefore corresponds to the average value of the most negative eigenvalue. Insets: cumulative distributions of the most negative eigenvalue μ_1 and of the scalar product $S = \langle w_1|v_1\rangle$.

9.4.3 Numerical simulations

In practice, however, we found it more convenient to use direct numerical simulations rather than the above exact results. In principle, these results below could be obtained using the mathematical formalism above, but the effort required to solve numerically the equations above is larger than the one needed to make direct simulations. We measure the null-hypothesis spectrum of $\widetilde{\mathbf{D}}$ by choosing $\xi(t)$ to be a Gaussian random variable of zero mean and unit variance, completely independent of the true returns $\eta_\alpha(t)$, which we then diagonalize. The cumulative distribution of the largest negative eigenvalue in the null-hypothesis is shown in the inset. The average position of the most negative eigenvalue of $\widetilde{\mathbf{D}}$ in the null-hypothesis case is

found to be $\widetilde{\mu}_1 \approx -4.8$. The average position of the second and third most negative eigenvalues in the null-hypothesis case will be denoted by $\widetilde{\mu}_2$ and $\widetilde{\mu}_3$.

We have also measured the distribution of the scalar product $S = \langle w_1 | v_1 \rangle$ between the corresponding top eigenvector $|w_1\rangle$ and the top eigenvector of $\mathbf{C}$, $|v_1\rangle$. We find that even in the case where $\xi(t)$ is an independent random variable, the top eigenvector of $\mathbf{D}$ is in fact strongly correlated with $|v_1\rangle$, with an average scalar product equal to $S = 0.68$ for the correlation matrix of the returns of the BE500 index. We find numerically that $P(S \leqslant 0.5) \approx 0.11$ and $P(S \leqslant 0.65) \approx 0.38$ for the BE500 index – see Fig. 10.3. Results for the SP500 are very similar.

9.4.4 Comparison with empirical data

In order to reduce the measurement noise and compare with the above numerical simulations, we have estimated $\mathbf{D}(\tau)$ using Eq. (9.13) with "Gaussianized" empirical index returns, obtained by first ranking the true index return from most negative to most positive, defining the rank of day t, $k(t)$. The Gaussianized index return $I_G(t)$ is then obtained as $\Phi^{-1}(k(t)/T)$, where Φ is the error function.

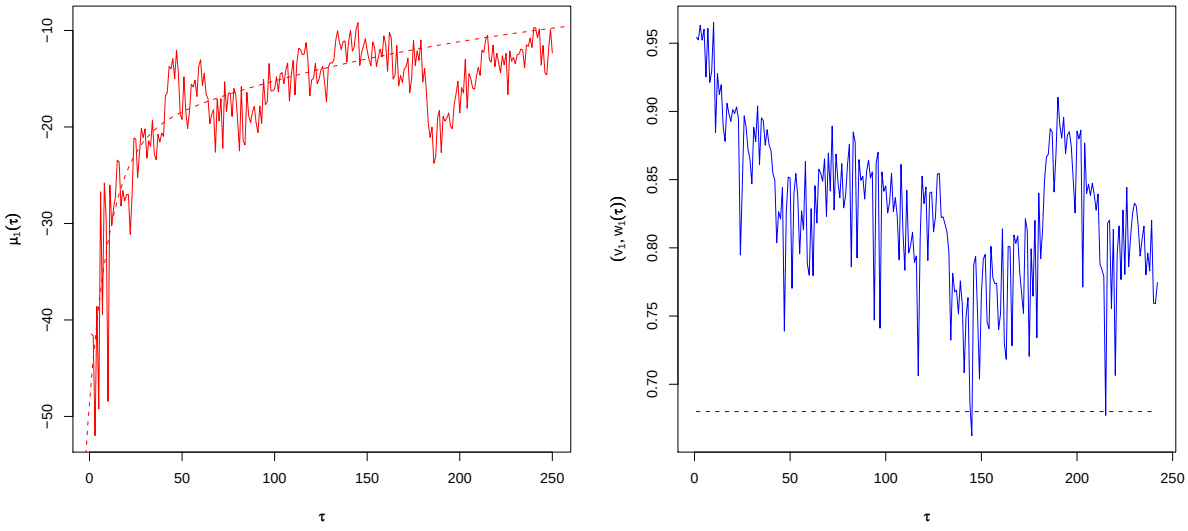


Figure 9.4: Left: Largest negative eigenvalue $\mu_1(\tau)$ of the lagged regression matrix $\mathbf{D}(\tau)$. The double exponential fit (dotted line) is given by : $\mu_1(\tau) = \mu_1^\infty - 26.6\exp(-\tau/11) - 17.1\exp(-\tau/200)$, where we fix the value of μ_1^∞ using the numerical results of the previous section: $\mu_1^\infty = \widetilde{\mu}_1 \approx -4.8$, since we expect that for large τ, all correlations are lost. Right: Evolution of the scalar product $S(\tau) = \langle v_1 | w_1(\tau) \rangle$ as a function of τ. The horizontal dashed line corresponds to the mean of the scalar product S in the null-hypothesis case. The data corresponds to the BE500 index, but the results for the SP500 are very similar.

We show in Fig. 10.4 the evolution of $\mu_1(\tau)$, the largest (in absolute value) eigenvalue of $\mathbf{D}(\tau)$ as a function of τ. We find that μ_1 is negative, corresponding to the correlation leverage effect (see Eq. (9.17)). Comparing with the null-hypothesis case, we find that $\mu_1(\tau)$ remains significant at the 1% confidence level up to $\tau \approx 240$. When fitting $\mu_1(\tau)$ with an exponential function with two scales that saturates at the noise level $\widetilde{\mu}_1$ determined above, we find $\mu_1(\tau) = \widetilde{\mu}_1 - 26.6\exp(-\tau/11) - 17.1\exp(-\tau/200)$. This reveals two time scales; a rather short one close to the one determined directly from $\mathcal{L}_\rho(\tau)$ above (see Fig. 9.1), and a much longer time scale on the order of a year, showing that the effect of market drops on the correlation is long lasting. The scalar product $S(\tau) = \langle w_1(\tau) | v_1 \rangle$ between the top eigenvectors of $\mathbf{D}(\tau)$ and $\mathbf{C}$ globally exceeds 0.8 in the whole range $\tau \in [1, 240]$, whereas the null-hypothesis average value is $S = 0.68$.

We have also studied the second ($\mu_2(\tau)$) and third ($\mu_3(\tau)$) eigenvalues of $\mathbf{D}(\tau)$ as a function of τ, which are both negative and clearly beyond the noise level, and are found to decay with very similar time scales: a month and a year (see Fig. 10.5). The corresponding eigenvectors are found to be mostly within the subspace spanned by the second and third eigenvectors of

C. The financial interpretation of these eigenvalues is of an increased sectorial correlation when the market drops on top of an increase of the market correlations. Therefore, all idiosyncratic effects disappear upon market drops, while global factors become dominant.

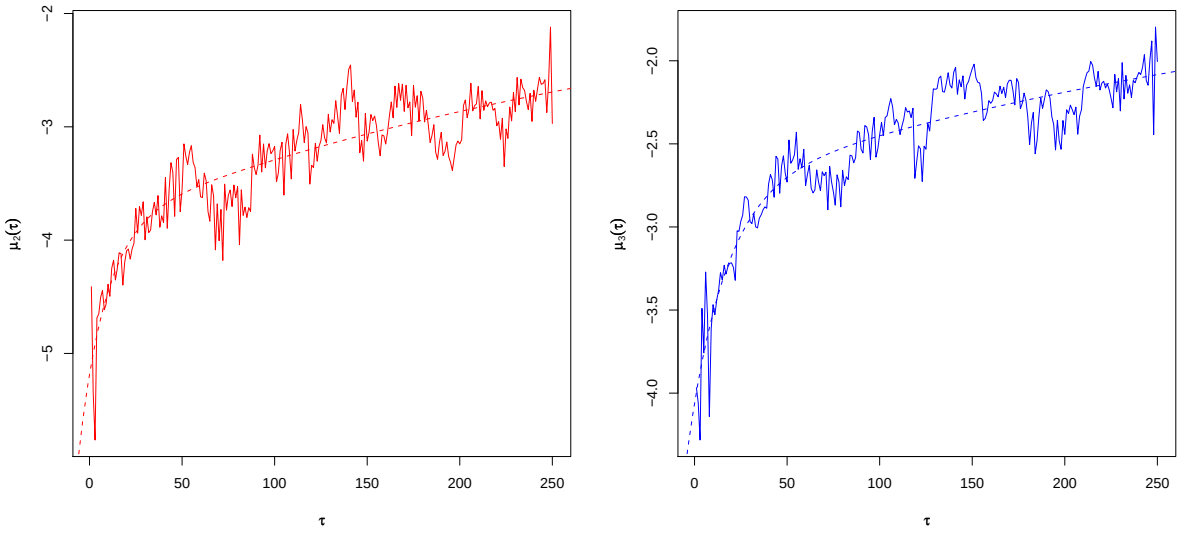


Figure 9.5: Left: Second eigenvalue $\mu_2(\tau)$ of the lagged regression matrix $\mathbf{D}(\tau)$. The exponential fit (dotted line) is given by: $\mu_2(\tau) = \widetilde{\mu}_2 - 1.3\exp(-\tau/14.4) - 2.3\exp(-\tau/364)$. Right: Third eigenvalue $\mu_3(\tau)$ of the lagged regression matrix $\mathbf{D}(\tau)$. The exponential fit (dotted line) is given by : $\mu_3(\tau) = \widetilde{\mu}_3 - 1.3\exp(-\tau/20) - 1.5\exp(-\tau/420)$. Direct numerical simulations of the random case lead to $\widetilde{\mu}_2 \approx -1.52$ and $\widetilde{\mu}_3 \approx -1.24$. The data corresponds to the BE500 index, but the results for the SP500 are again very similar.

9.4.5 Separating negative & positive returns

As Fig. 10.2 explicitely shows, the correlation depends on past index returns in a non-linear way. In fact, both negative and positive returns increase the correlations, although the effect is stronger for negative returns, which in turn leads to a non-zero linear term in the regression of $\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t)$ on $I(t-\tau)$. A way to capture the parabolic shape seen in Fig. 10.2 would be to extend the above model to:

$$\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t) := C_{\alpha,\beta} + D_{\alpha,\beta}(\tau)I(t-\tau) + E_{\alpha,\beta}(\tau)\left[I^2(t-\tau) - \langle I^2\rangle\right] + \varepsilon_{\alpha,\beta}(t), \tag{9.26}$$

defining a new matrix $\mathbf{E}$ that captures the symmetric effect of index returns on the correlation matrix. An alternative choice, that we adopt below, is to regress separately on negative returns and on positive returns:

$$\widehat{\eta}_\alpha(t)\widehat{\eta}_\beta(t) := C_{\alpha,\beta} + D^+_{\alpha,\beta}(\tau)\left[I^+(t-\tau) - \langle I^+\rangle\right]\delta_{\{I(t-\tau)>0\}} \tag{9.27}$$

$$+ D^-_{\alpha,\beta}(\tau)\left[I^-(t-\tau) - \langle I^-\rangle\right]\delta_{\{I(t-\tau)<0\}} + \varepsilon_{\alpha,\beta}(t), \tag{9.28}$$

where $I^+ = \max(I,0), I^- = \min(I,0)$ and δ is the Dirac function. With this definition, one can rewrite the correlation matrix conditioned to a certain past value of I more precisely, separating the effect of positive returns and negative returns, as follows:

$$\mathbf{C}(I) = \mathbf{C} + \mathbf{D}^-\left[I^- - \langle I^-\rangle\right]\delta_{\{I<0\}} + \mathbf{D}^+\left[I^+ - \langle I^+\rangle\right]\delta_{\{I>0\}}. \tag{9.29}$$

Again, in order to reduce the measurement noise, we used "Gaussianized" empirical index returns $I_G(t)$ instead of $I(t)$. We apply to $\mathbf{D}^\pm(\tau)$ the same analysis as above. As anticipated, the top eigenvalue μ_1^- of $\mathbf{D}^-$ is strongly negative, whereas the top eigenvalue μ_1^+ of $\mathbf{D}^+$ is positive, but with $\mu_1^+ < |\mu_1^-|$ — see Fig. 10.6. The projections of $|w_1^+\rangle$ and $|w_1^-\rangle$ onto $|v_1\rangle$ are both very

close to unity for small τ and gradually decay to the noise level as τ increases. To check the significancy of our effect, as before, we define a null-hypothesis case, introducing the matrix:

$$\langle \phi^2 \rangle \widetilde{D}^-_{\alpha,\beta} = \frac{1}{T} \sum_{t=1}^{T} x_\alpha(t) x_\beta(t) \phi(t) \tag{9.30}$$

where the conditioning variables ϕ is independent of the x_α (which are standard gaussian variables whose correlation matrix is $\mathbf{C}$ as above) and distributed as $\min(\xi, 0) - \langle \min(\xi, 0) \rangle$ where ξ is as before a standard gaussian variable. We define further the matrix $\widetilde{\mathbf{D}}^+$ exactly as $\widetilde{\mathbf{D}}^-$ except for the fact that the conditioning variable is now distributed as $\max(\xi, 0) - \langle \max(\xi, 0) \rangle$. As above, $\widetilde{\mu}^-_1, \widetilde{\mu}^-_2, \widetilde{\mu}^-_3$ will be the average positions of the first, second and third most negative eigenvalues of $\widetilde{\mathbf{D}}^-$ and $\widetilde{\mu}^+_1, \widetilde{\mu}^+_2, \widetilde{\mu}^+_3$ will be the average positions of the first, second and third most positive eigenvalues of $\widetilde{\mathbf{D}}^+$. Those values are all computed using numerical simulations.

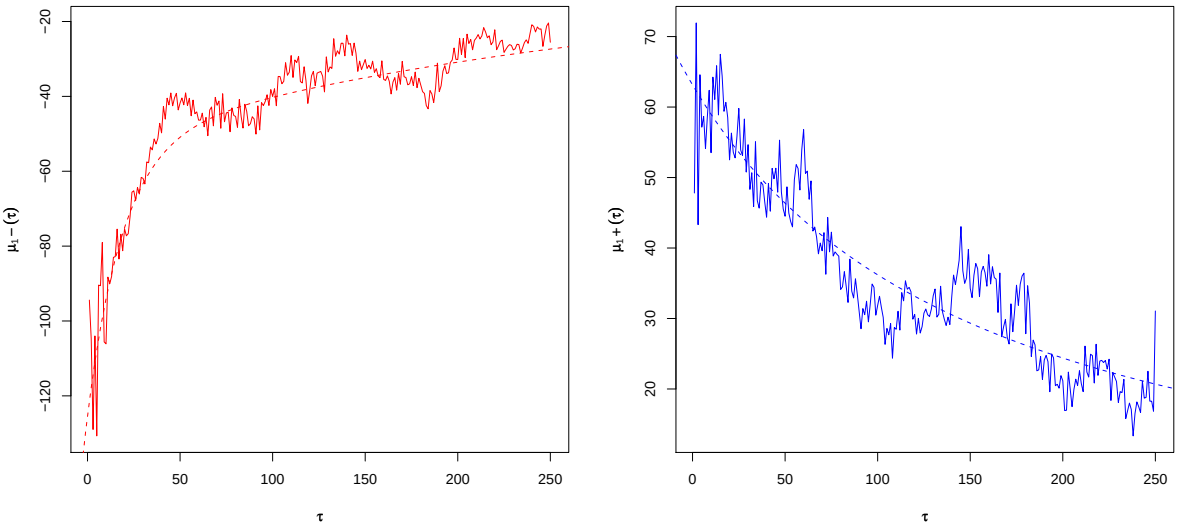


Figure 9.6: Left: $\mu^-_1(\tau)$ of the lagged regression matrix $\mathbf{D}^-(\tau)$. The exponential fit (dotted line) is given by : $\mu^-_1(\tau) = \widetilde{\mu}^-_1 - 73\exp(-\tau/19) - 41\exp(-\tau/300)$. Right: $\mu^+_1(\tau)$ of the lagged regression matrix $\mathbf{D}^+(\tau)$. The exponential fit (dotted line) is now given by : $\mu^+_1(\tau) = \widetilde{\mu}^+_1 + 10.6\exp(-\tau/49) + 44\exp(-\tau/200)$. Note again the presence of a long relaxation time on the order of a year. We have used direct numerical simulations to obtain $\widetilde{\mu}^-_1 = \widetilde{\mu}^+_1 \approx -8.3$. The data is for the returns of the BE500. Again, SP500 yields very similar results.

We have also studied the rotation parameter $\Delta^\pm$ for both matrices $\mathbf{D}^\pm(\tau)$ defined as:

$$\Delta^\pm = \frac{1}{\lambda_1} \left[\langle e|\mathbf{D}^\pm|v_1\rangle - \langle v_1|\mathbf{D}^\pm|v_1\rangle\langle e|v_1\rangle \right]. \tag{9.31}$$

The results are shown in Fig. 10.7. In agreement with the common lore, Δ^- is negative, indicating that strongly negative index returns (below $\langle I^- \rangle$) lead to a more uniform instantaneous market mode. On the other hand, Δ^+ is found to be negative as well, meaning that while strongly positive returns also tend to increase the average correlation between stocks, the instantaneous market mode rotates *away* from the uniform vector $|e\rangle$. The effects we are reporting are statistically significant since the root-mean square error on $\widetilde{\Delta}^\pm$ (defined as in Eq.(9.24)) in the null-hypothesis case is found to be $\sim 8 \cdot 10^{-4}$, a factor 3 to 4 smaller than the amplitude of the empirical values of $\widetilde{\Delta}^\pm$.

9.5 Summary & Conclusion

The aim of this paper was to revisit the index leverage effect, that can be decomposed into a volatility effect and a correlation effect. We investigated the latter in great detail using a

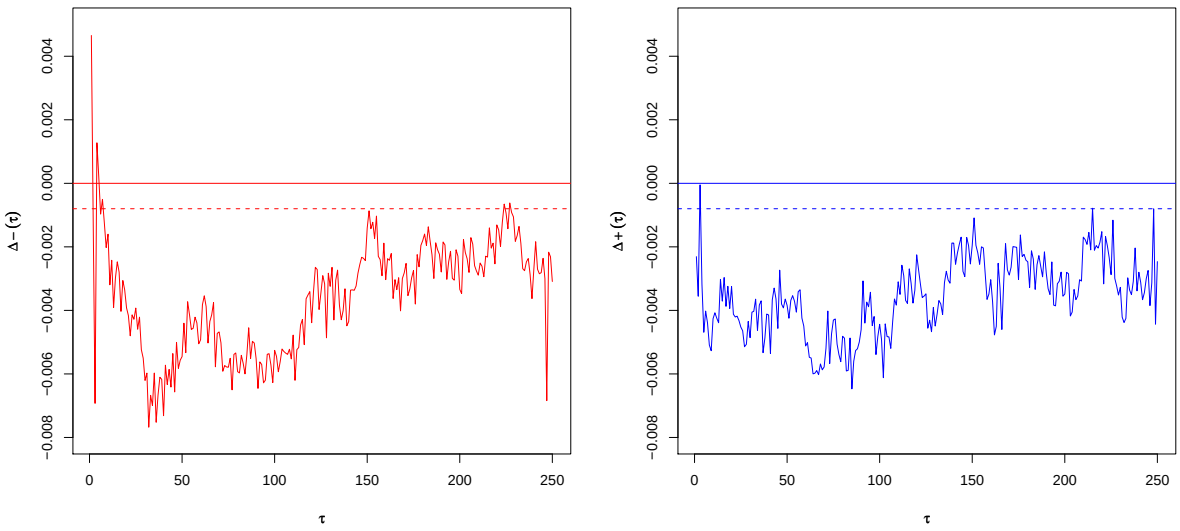


Figure 9.7: Plot of the rotation parameters Δ^- (left) and Δ^+ (right) as a function of τ. The horizontal dashed lines correspond to the root-mean square error on $\Delta^\pm$ in the null-hypothesis case. The data is for BE500; the amplitude of Δ^- and Δ^+ are found to be roughly a factor 2 larger for the SP500.

matrix regression analysis, that we called 'Principal Regression Analysis' (PRA) and for which we have provided, using Random Matrix Theory and simulations, some analytical and numerical benchmarks.

Using this refined analysis, we confirm that downward index trends increase the average correlation between stocks (as measured by the top eigenvalue of the conditional correlation matrix), which in turn explains why the index leverage effect is stronger than for single stocks. Compared to the null-hypothesis benchmark, this leverage correlation effect is highly significant (see Fig. 10.4 and Fig. 10.6). We also find that large downward trends implies a more uniform future market mode (see Fig. 10.7, left).

Upward trends, on the other hand, also increase the average correlation between stocks (see Fig. 10.6, right) but large upward trends rotate the future market mode *away* from uniformity (see Fig. 10.7, right). All these effects are characterized by two 'memory' time scales: a 'short' one on the order of a month and a longer one on the order of a year. The latter long time scale could be related to the fact that the market had long cycles of booms and busts within the studied time series, during which the average correlation went down and up again.

We have also studied the correlation leverage effect on intraday data, and we find (results not shown) that while the top eigenvalue of the 15 minutes correlation matrix is nearly insensitive to the sign of the previous 15 minutes index return, a significant effect emerges when the time scale reaches one hour.

Finally, we have found indications of a leverage effect for sectorial correlations as well, which reveals itself in the second and third modes of the PRA (see Fig. 10.5). It would be interesting to analyze other conditional correlation matrices using the tools developed in this paper, such as for example leader-lagger effects [124, 46, 123], or the role of other macro variables such as oil, currencies or interest rates.

Acknowledgements We have benefitted from insightful comments and suggestions by Giulio Biroli, Rémy Chicheportiche, Stefano Ciliberti, Marc Potters and Vincent Vargas.

Chapter 10

Individual and collective stock dynamics: intra-day seasonalities

Résumé

Cet article est publié dans le journal *New Journal of Physics* et est écrit en collaboration avec Jean-Philippe Bouchaud. Nous établissons plusieurs nouveaux faits stylisés concernant les saisonalités de la dynamique des prix des actions au cours de la journée. Au delà du célèbre effet U pour la volatilité au cours de la journée, nous trouvons que la corrélation moyenne entre les actions augmente au cours de la journée, ce qui conduit à une plus petite dispersion entre les stocks. D'une manière légèrement paradoxale, la kurtosis (une mesure des occurrences de volatilité inhabituelles) atteint son minimum à l'ouverture du marché, lorsque la volatilité est à son maximum. Nous confirmons que la kurtosis est une fonction décroissante du rendement de l'indice. Cela signifie que lors des grands mouvements du marché, la composante idiosyncratique devient moindre dans la dynamique des actions. Schématiquement, les heures du matin sont dominés par les mouvements idiosyncratiques ou sectorielle avec peu de surprise, tandis que l'influence du vecteur marché augmente au cours de la journée et que les surprises deviennent plus fréquentes.

Abstract

We establish several new stylized facts concerning the intra-day seasonalities of stock dynamics. Beyond the well known U-shaped pattern of the volatility, we find that the average correlation between stocks increases throughout the day, leading to a smaller relative dispersion between stocks. Somewhat paradoxically, the kurtosis (a measure of volatility surprises) reaches a minimum at the open of the market, when the volatility is at its peak. We confirm that the dispersion kurtosis is a markedly decreasing function of the index return. This means that during large market swings, the idiosyncratic component of the stock dynamics becomes subdominant. In a nutshell, early hours of trading are dominated by idiosyncratic or sector specific effects with little surprises, whereas the influence of the market factor increases throughout the day, and surprises become more frequent.

10.1 Introduction

Financial markets operate in sync with human activities. It is therefore no surprise that financial time series reveal a number of seasonalities related to human rhythms: markets open in the morning and close in the evening, remain closed during week-ends and during vacations; wages are paid and portfolios are re-balanced on a monthly basis, earnings are announced on a quarterly basis (in the US), etc. These periodicities leave a statistical trace on the time series of returns of many assets. Among the best known periodicities is the so-called U effect [2, 13], that describes the intra-day pattern of volatility of individual US stocks: the average volatility is observed to

be high after the market opens, then decreases as to reach a minimum around lunch time and increases again steadily until market close.[1]

In this short note we want to report on additional intra-day patterns concerning both individual and collective stock dynamics. First, we study the intra-day pattern of other moments of the individual stock dynamics, beyond the well known U-shaped volatility. Second, we characterise the cross-sectional distribution of returns and its typical evolution during the day. Finally, we study the correlation matrix between stock returns and find that the leading modes also have a very well defined intra-day pattern. Our study here is entirely empirical, but our results certainly beg for a detailed theoretical interpretation in terms of agent behaviour: strategies, information processing, risk aversion, etc. We provide some hints in that direction in the conclusion.

10.2 Data, notations and definitions

We have considered a set of $N = 126$ stocks of the New York Stock Exchange (which are among the 250 largest market capitalisations) that has been continuously traded during the period between 01/01/2000 and 12/31/2009 to form a statistical ensemble of 5 minutes stock returns. The total number of 5 minute bins is $186,498$, whereas the total number of days is $T = 2,391$ ($K = 78$ bins per day). Stocks will be labelled by $\alpha = 1, \ldots, N$, days by $t = 1, \ldots, T$ and bins by $k = 1, \ldots, K$. The return of stock α in bin k of day t will be denoted as $\eta_\alpha(k;t)$.

Different types of averages will be needed. Time averages for a given stock and a given bin are expressed with angled brackets: $\langle \ldots \rangle$, whereas averages over the ensemble of stocks for a given bin in a given day appear with square brackets: $[\ldots]$. For an arbitrary function $F(.)$ of these returns we therefore write:

$$\langle F \rangle(k;\alpha) := \frac{1}{T}\sum_{t=1}^{T} F(\eta_\alpha(k;t)); \quad [F](k;t) := \frac{1}{N}\sum_{\alpha=1}^{N} F(\eta_\alpha(k;t)); \quad [\langle F \rangle](k) = \frac{1}{T}\sum_{t=1}^{T}[F](k;t). \tag{10.1}$$

The first set of observables concerns single stock properties. We characterise the distribution of stock α in bin k by its four first moments: mean $\mu_\alpha(k)$, standard deviation (volatility) $\sigma_\alpha(k)$, skewness $\zeta_\alpha(k)$ and kurtosis $\kappa_\alpha(k)$. We will in fact use low moments, less noisy estimates of the last two quantities. We will define $m_\alpha(k)$ as the median of all returns of stock α in bin k, and define:

$$\mu_\alpha(k) = \langle \eta_\alpha(k;t) \rangle \tag{10.2a}$$

$$\sigma_\alpha^2(k) = \langle \eta_\alpha(k;t)^2 \rangle - \mu_\alpha^2(k) \tag{10.2b}$$

$$\zeta_\alpha(k) = \frac{6}{\sigma_\alpha(k)} \left(\mu_\alpha(k) - m_\alpha(k) \right) \tag{10.2c}$$

$$\kappa_\alpha(k) = 24 \left(1 - \sqrt{\frac{\pi}{2}} \frac{\langle |\eta_\alpha(k;t) - \mu_\alpha(k)| \rangle}{\sigma_\alpha(k)} \right) + \zeta_\alpha(k)^2. \tag{10.2d}$$

Within a cumulant expansion, the last two lines coincide with the usual definition of skewness and kurtosis, but no moments larger than two are needed to estimate them. Note that the correction term $\zeta_\alpha(k)^2$ to the kurtosis turns out to be negligible, and we have neglected it in the following. We will be interested below in the average over all stocks of the above quantities, as a way to characterize the typical intra-day evolution of the distribution of single stock returns.

One can also consider *cross sectional* distributions, i.e. the dispersion of the returns of the N stocks for a given bin k in a given day t, i.e. one distribution every five minutes. One can again characterize these distributions in terms of the first four moments. The median of all N

[1] This pattern is a little different in Europe or in the UK, with a second volatility spike at 2:30 pm GMT when the US market opens.

returns for a given $k;t$ is now $m_d(k;t)$ (d for "dispersion"), and we define:

$$\mu_d(k;t) = [\eta_\alpha(k;t)] \tag{10.3a}$$
$$\sigma_d^2(k;t) = [\eta_\alpha(k;t)^2] - \mu_d^2(k;t) \tag{10.3b}$$
$$\zeta_d(k;t) = \frac{6}{\sigma_d(k;t)} \left(\mu_d(k;t) - m_d(k;t)\right) \tag{10.3c}$$
$$\kappa_d(k) = 24 \left(1 - \sqrt{\frac{\pi}{2}} \frac{[|\eta_\alpha(k;t) - \mu_\alpha(k)|]}{\sigma_d(k)}\right) \tag{10.3d}$$

Note that $\mu_d(k;t)$ can be seen as the return of an index, equiweighted on all stocks. We will be interested below in the average over all days of the above quantities, as a way to characterize the typical intra-day evolution of the dispersion between stock returns.

Although the dispersion already captures part of the "Co-movements" of stocks, a more direct characterization is through the standard correlation of returns. In order to measure the correlation matrix of the returns, we first normalize each return by the dispersion of the corresponding bin. This factors in any "trivial" intra-day seasonality, and also accounts for the fact that the volatility fluctuates quite a bit during the 10 year time interval that we consider. Therefore, we introduce: $\widehat{\eta}_\alpha(k;t) = \eta_\alpha(k;t)/\sigma_d(k;t)$ and study the correlation matrix defined for a given bin k:

$$C_{\alpha,\beta}(k) := \frac{1}{\widehat{\sigma}_\alpha(k)\widehat{\sigma}_\beta(k)} \langle \widehat{\eta}_\alpha(k;t)\widehat{\eta}_\beta(k;t)\rangle_c, \tag{10.4}$$

where the subscript c means "connected part" (i.e. averages have been subtracted) and $\sigma_\alpha^2(k) := \langle \eta_\alpha^2(k;t)\rangle_c$. Of special interest are the largest eigenvalues and eigenvectors of $C_{\alpha,\beta}(k)$, which characterize the correlation structure of stock returns. This analysis has been performed in several papers ([45, 44, 146, 122]) using daily or high frequency returns, and it is well known that the structure of large eigenvectors reflects the existence of economic sectors of activity. The largest eigenvalue λ_1, in particular, corresponds to the market mode, and is associated to an eigenvector with all entries positive and close to $1/\sqrt{N}$. In fact, λ_1/N can be seen as a measure of the average correlation between stocks. We will be interested below in the k dependence of the largest eigenvalues and their associated eigenvectors, a study that, to the best of our knowledge, has not been reported in the literature before.

10.3 Single stock intra-day seasonalities

10.3.1 Odd moments

Odd moments tend to be small and noisy, so it is difficult to draw definite conclusions. The average return is on average over the whole period positive, but noisy and does not show any intra-day pattern. The average skewness of five minutes returns is also noisy and is compatible with zero, again without any identifiable intra-day pattern at all. This is at variance with the skewness of returns on a longer time interval, which is negative. The build up of negative skewness with time scale is a consequence of the leverage effect, i.e. negative returns tend to be followed by larger volatilities (see e.g. [45]).

10.3.2 Even moments

The average volatility, on the other hand, reveals a very clean U-shaped pattern that has been reported many times in the literature ([2, 13]). We show in Fig. 10.1 $\sigma(k) = [\sigma_\alpha(k)]$. Note that the overnight volatility $\approx 1.15\%$ is much larger than the typical five minute volatility, and is not shown in the graph. Interestingly, the average volatility is found to decay in the first two hours of trading as a power-law $k^{-\beta}$ with $\beta \approx 0.3$. This relaxation is reminiscent of the power-law decay of the volatility after large price swings [101, 142, 156, 90]. The overnight return is indeed usually quite large, and can be seen as a strong perturbation. The power-law relaxation suggests that some critical mechanism is involved in the way volatility reverts back to 'normal' after market jumps.

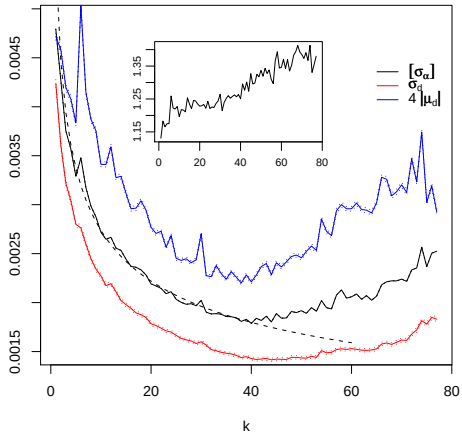


Figure 10.1: We show the average volatility of stocks $\sigma(k)$, the average cross sectional dispersion $\sigma_d(k)$ and the average absolute value of the index return $\langle|\mu_d(k,t)|\rangle$ (multiplied by 4 for clarity) as a function of k, with the corresponding statistical error bars. All display the well known U pattern. We also compare $\sigma(k)$ with a power-law decay $k^{-\beta}$ with $\beta \approx 0.3$ (dashed line), which is a good fit for the first half of the day. Inset: ratio $\sigma(k)/\sigma_d(k)$ as a function of k, showing that dispersion effects diminish throughout the day.

Turning now to the kurtosis $\kappa(k) = [\kappa_\alpha(k)]$, we find, perhaps surprisingly, that there is a clear intra-day *growth* of the kurtosis from $\kappa \approx 3.5$ at the beginning of the day to $\kappa \approx 5$ around 1 p.m., and stays approximately constant (but noisy) until the end of the day – see Fig. 10.2, left. The overnight kurtosis remains around 5. This finding is counter-intuitive because one would naively associate the large volatility in the morning with huge swings, symptomatic of the market uncertainty at the open. But this is not the case: the maximum of the intra-day volatility corresponds to a *minimum* in kurtosis. We will report similar counter-intuititive results below. Possible mechanisms are discussed in the conclusion.

10.4 Cross-sectional intra-day seasonalities

10.4.1 Odd moments

Noting that the average over stocks of $\mu_\alpha(k)$ is identical to the average over time of $\mu_d(k;t)$, the discussion of the first moment of the cross-sectional distribution is redundant. The average of $|\mu_d(k;t)|$ is a proxy for the index volatility, and is displayed in Fig. 10.1 : it shows a U-shaped pattern similar to that of $\sigma(k)$, with however a stronger end-of-day surge. This is due to the correlation pattern discussed in section 5 below: the average correlation between stock indeed increases as the day proceeds, leading to an increased index volatility.

As far as the average skewness $\zeta_d(k) = \langle\zeta_d(k;t)\rangle$ is concerned, we again find a very noisy quantity with no particular intra-day pattern. The only notable feature is that this time, the skewness is significantly positive, albeit small: the average over k of $\zeta_d(k)$ is found to be ≈ 0.025.

10.4.2 Even moments

As above, the even moments show clear patterns. The average dispersion $\sigma_d(k) = \langle\sigma_d(k;t)\rangle$ exhibits a U-shaped pattern very similar to that of $\sigma(k)$ — see Fig. 10.1. In fact, the ratio $\sigma(k)/\sigma_d(k)$ is plotted in the inset of Fig. 10.1 as a function of k and increases from ≈ 1.15 at the open to 1.45 at the close. In relative terms, the dispersion is thus stronger in the morning, and decreases as the day proceeds.

The dispersion kurtosis $\kappa_d(k) = \langle\kappa_d(k;t)\rangle$, on the other hand, has an *inverted* U shape, and reaches a minimum at the open and at the close of the market, i.e. when the dispersion and the volatility are locally maximum. So even when the dispersion of returns is at its peak, with stocks all over the place (so to say), the cross-sectional distribution of returns is on average

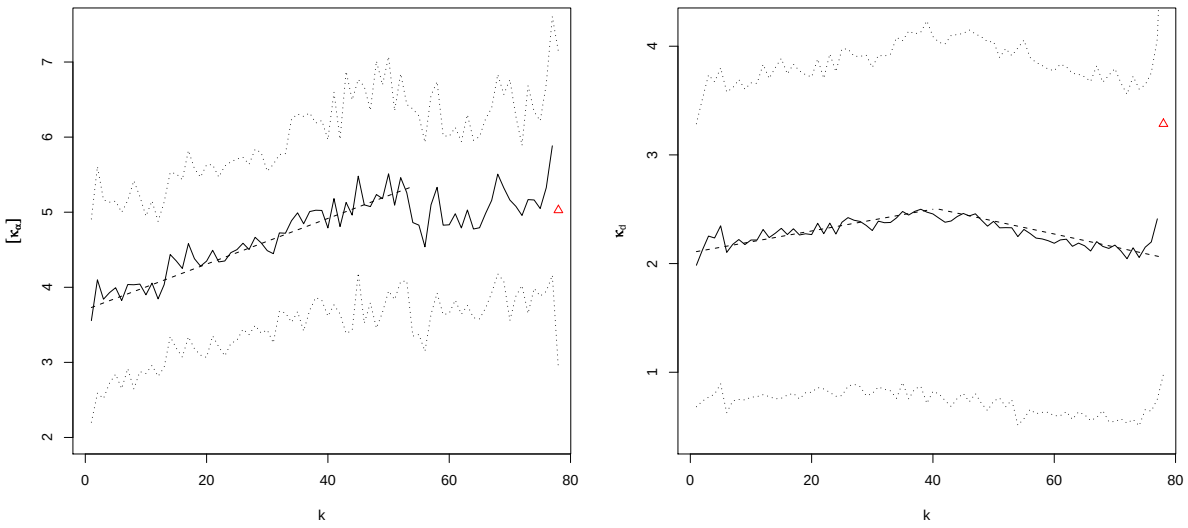


Figure 10.2: Left: average kurtosis of individual stocks $\kappa(k)$ as a function of k. Right: average cross sectional kurtosis $\kappa_d(k)$ as a function of k. In both cases, we show the 1-σ dispersion around the mean (i.e. not the error bar). The red triangles correspond to the overnight values.

closer to a Gaussian! Note however that the variation of the kurtosis is not large, from $\kappa_d = 2$ to $\kappa_d = 2.4$. The overnight dispersion kurtosis, on the other hand, is much stronger: $\kappa_d \approx 3.3$.

10.4.3 Conditioning on the index return

As noted above, the quantity $\mu_d(k;t)$ is the return of an equiweighted index. It is interesting to condition the value of the moments of the cross-sectional dispersion on this quantity. Such a study was performed on daily returns in [103] and more recently by L. Borland [39]. In agreement with the results of [103, 58], we find that the average dispersion σ_d is an increasing function of the amplitude of the index return, see Fig. 10.3. As noted in [58], this observation shows that the volatility of the stock residuals in a one-factor model must depend on the volatility of the market mode. Fig. 10.3 furthermore suggests that this dependence is sub-linear (see [56] for some elaborations on this observation.)

As first established in [102] on daily data, we find that the skewness ζ_d is an odd function of μ_d, as shown in Fig. 10.4. Note that the skewness increases very abruptly for small μ_d and saturates for larger values of the index return. Pictorially, a positive index return can be thought of as resulting from a few "winners" running ahead of the pack, contributing both to the mean μ_d and to the skewness. The slope of $\zeta_d(\mu_d)$ around the origin, together with the fact that the index has made on average positive daily gains in the period 2000 – 2009, are enough to explain the average value of the dispersion skewness $\zeta_d(k) \approx 0.025$ reported above.

Finally, the dispersion kurtosis κ_d shows again a non-intuitive *decreasing* behaviour as a function of $|\mu_d|$, see Fig. 10.5. The average kurtosis conditioned to a value of $|\mu_d|$ decreases from ≈ 2.8 for small index returns to ≈ 1.8 for index returns larger than 2% in absolute value. This was first noticed *en passant* in [58] on daily data and recently emphasised by Borland [39]. Here, we confirm on five minute returns this strange stylised fact: the cross-sectional distribution of returns appears to be more Gaussian when its mean is off-centred.

However, if we now condition κ_d on the dispersion σ_d (which, as we found above, is *positively* correlated with $|\mu_d|$), we find (see Fig. 10.5) the opposite behaviour, i.e. the larger the dispersion, the larger the kurtosis κ_d! We will offer a discussion of these confusing effects in the discussion section below.

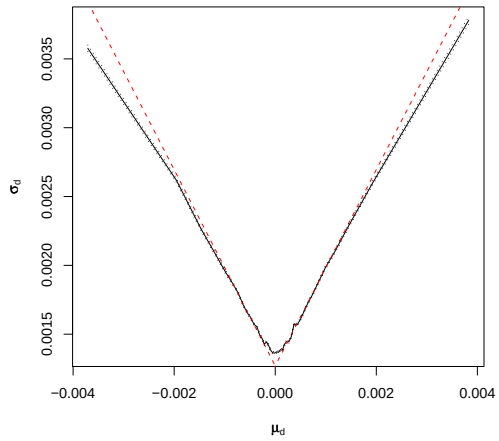


Figure 10.3: Cross sectional dispersion σ_d as a function of the index return (equiweighted on all stocks) μ_d. We added error bars, and linear branches that fit the small $|\mu_d|$ slopes, that emphasise the sub-linear behaviour of σ_d.

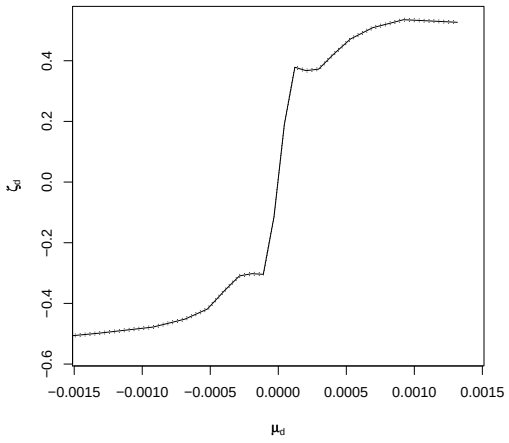


Figure 10.4: Cross sectional skewness ζ_d as a function of the index return (equiweighted on all stocks) μ_d. We added error bars (dotted lines), that are actually difficult to see near the origin.

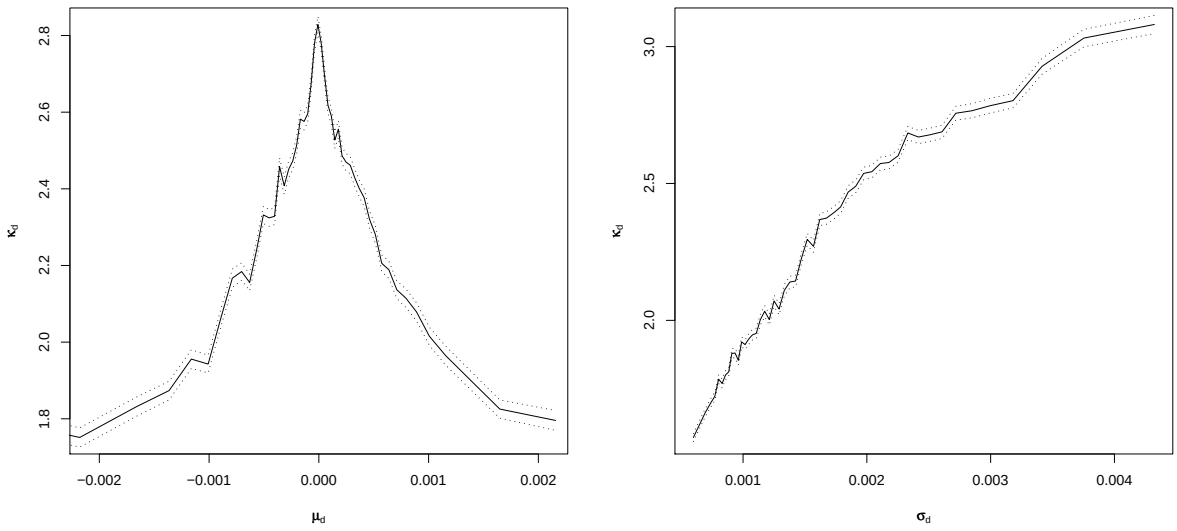


Figure 10.5: Left: Cross sectional kurtosis κ_d as a function of the index return (equiweighted on all stocks) μ_d. Note that the dependence is nearly the same for positive and negative market returns. Right: Cross sectional kurtosis κ_d as a function of the cross sectional dispersion σ_d. We added error bars (dashed lines) on both figures.

10.5 Intra-day seasonalities in the inter-stock correlations

Let us now turn to the properties of the eigenvalues and eigenvectors of the $N \times N$ correlation matrix $C_{\alpha,\beta}(k)$ defined by Eq. (10.4) above.

10.5.1 The top eigenvalue

The largest eigenvalue λ_1 of the correlation matrix of stock returns is well known to be associated with the "market mode", i.e. all stocks moving more or less in sync. As recalled above, the quantity λ_1/N can be used to define the average correlation between stocks.

We show in Fig. 10.6 (left side) the magnitude of λ_1/N as a function of k. Interestingly, the average correlation clearly increases as time elapses, from a rather small value ≈ 0.12 when the market opens to ≈ 0.3 near market close. This is in agreement with the fact that the dispersion $\sigma_d(k)$ is, in relative terms, smaller at the end of the day (see Fig. 10.1, inset). The value of λ_1/N for the correlation of overnight returns is also around 0.3, in continuity with the value at the end of the trading day.

In agreement with the idea that the stock dynamics become more and more uniform as the day proceeds, we find a substantial increase of the scalar product of the largest eigenvector $\vec{v}_1(k)$ with the uniform normalised vector $\vec{e} = (1/\sqrt{N}, 1/\sqrt{N}, \ldots, 1/\sqrt{N})$ — see Fig. 10.6 right. This scalar product is always close to unity, confirming the market mode interpretation of the top eigenvalue, but starts the day around 0.97 and ends the day at 0.995, before dropping again in the last bins of the day and during the overnight, when it is equal to 0.985 (i.e. larger than the open value).

10.5.2 Smaller eigenvalues

The evolution of the next six eigenvalues $\lambda_i(k)$, $i = 2, \ldots, 7$ is shown in Fig. 10.7. We see that the amplitude of this risk factors now *decreases* with time, before shooting back up during the overnight (see the last point of the graphs). Although by construction the trace of the correlation matrix, and therefore the sum of all N eigenvalues is constant (and equal to N), this decrease is not a trivial consequence of the increase of λ_1, since the sum of the first five eigenvalues is ~ 50, still small compared to $\mathrm{Tr}(C) = N = 126$. What we see here is that as the day proceeds, more and more risk is carried by the market factor, while the amplitude of sectorial moves shrivels in

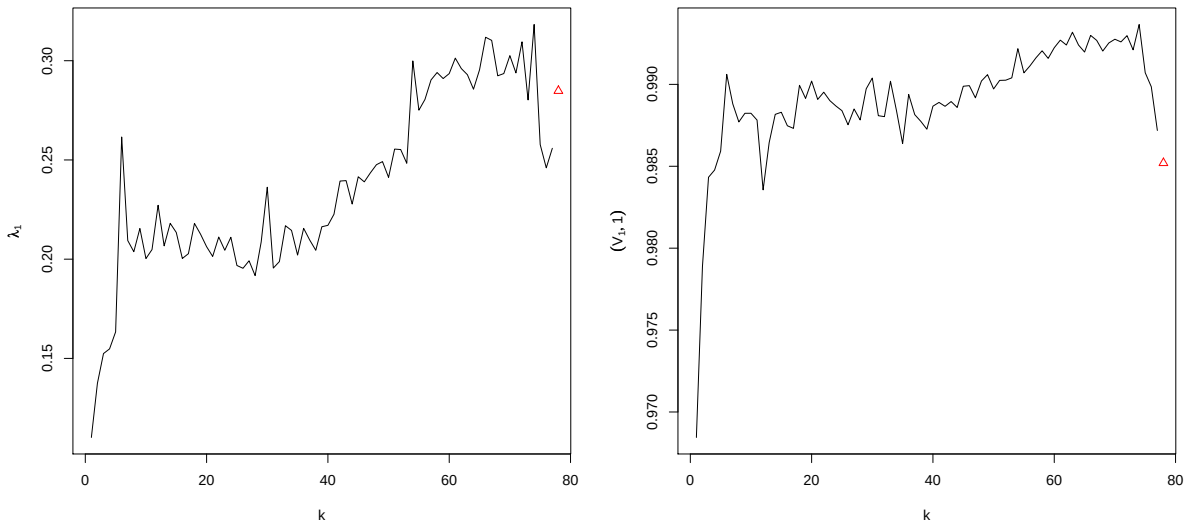


Figure 10.6: Left: Dependence of the top eigenvalue of the correlation matrix $C(k)$, $\lambda_1(k)/N$, as a function of time of day. Right: Evolution of the scalar product between top eigenvector $\vec{v}_1(k)$ and the uniform vector $\vec{e}$. The red triangles correspond to the overnight values.

relative terms (but remember that the correlation matrix is defined after normalising the returns by the local volatility, which increases in the last hours of the day).

It is more difficult to visualise the evolution of the corresponding eigenvectors, since there is no natural vectors to compare them with. Furthermore, eigenvalues can "collide" and cross, resulting in an interchange between two consecutive eigenvectors. We have therefore chosen to take as a reference the eigenvectors $\vec{v}_i(1)$ in the opening bin $k = 1$, corresponding to the largest values of $\lambda_i(k)$, $i = 2, \ldots, 7$. We then form the 6×6 matrix of scalar products $W_{ij}(k) = \vec{v}_i(1) \cdot \vec{v}_j(k)$. The singular values $s_\ell(k)$ of this matrix (equal to the square-root of the eigenvalues of $W^T W$) give a measure of the overlap between the eigenspace spanned by the $\vec{v}_i(1)$ and that spanned by the $\vec{v}_j(k)$. If the $\vec{v}_j(k)$ are a permutation of the $\vec{v}_i(1)$, all the s_i's are equal to unity, indicating maximum overlap. In particular, $s_\ell(1) \equiv 1$ trivially. The evolution of the $s_\ell(k)$, $\ell = 2, \ldots, 7$ is shown in Fig. 10.7 right. Using the results of [46], we conclude that all $s_\ell(k)$ are meaningful, since in the absence of any true correlations between the $\vec{v}_i(1)$ and the $\vec{v}_j(k)$, one would expect all singular values to lie in the interval $[0, 0.12]$. Therefore, although the structure of correlations clearly evolves between the opening hours and the closing hours, there is as expected a strong overlap between the principal components throughout the day.

10.6 Discussion & Conclusion

Let us present a synthetic account of the above empirical results, for which we only propose an interpretation stub. We have seen that during the opening hours of the market, the volatility and the dispersion of returns are high, whereas kurtosis effects are relatively low. These two quantities are different measures of the *heterogeneity* of stock returns, and quite paradoxically they are found to behave in opposite ways. But while the volatility and dispersion are dimensional measures of heterogeneity (measuring the spread of returns in %), the kurtosis is a relative, a-dimensional measure of *surprise*. What our results mean in intuitive terms is that although the typical amplitudes of stock returns are high in the morning, outliers are relatively rare, both over time and over stocks. In a sense, agitation is the norm during these early hours of trading, stocks move in different directions in such a way that the average correlation is weaker than average, and the top eigenvector of the correlation matrix is farther away from the uniform mode $\vec{e} = (1/\sqrt{N}, 1/\sqrt{N}, \ldots, 1/\sqrt{N})$. But anomalously large jumps rarely take place in the morning — as expected, these jumps are more likely overnight (and are to be related to arrival of corporate specific or market-wide information), where kurtosis effects are strongest,

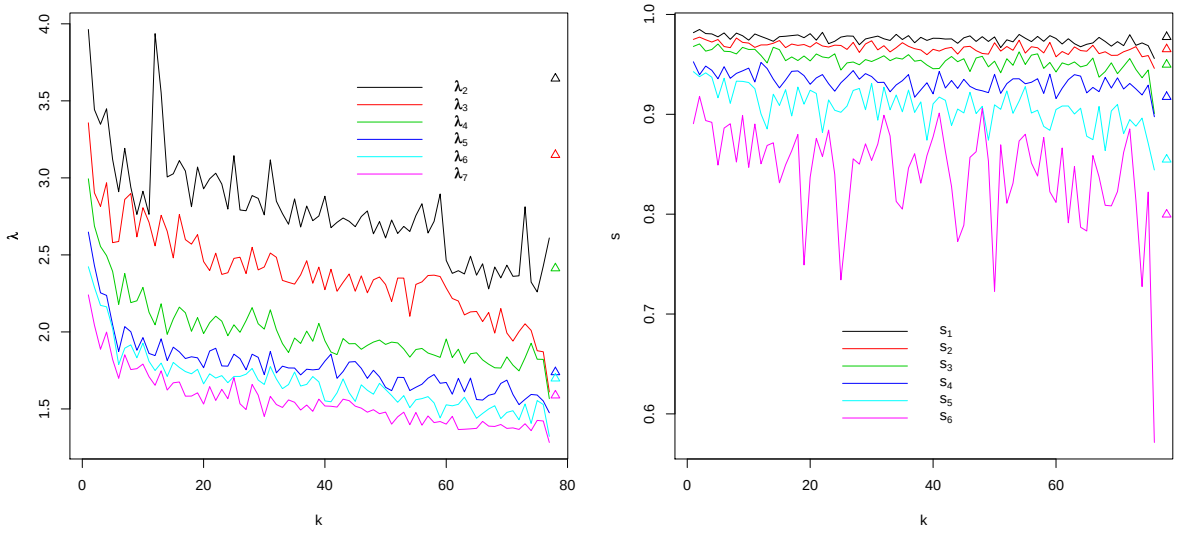


Figure 10.7: Left: Smaller eigenvalues $\lambda_i(k)$, $i = 2, \ldots, 7$) as a function of k. Right: The 6 singular values $s_\ell(k)$ ($\ell = 2, \ldots, 7$) of the matrix $W_{ij}(k)$. In the absence of any true persistence, one would expect all singular value to lie in the interval $[0, 0.12]$, much below the smallest singular value $s_7 \sim 0.8$. The triangles correspond to the overnight values.

both for single stock and cross-sectional returns. As the day proceeds, correlations increase and dispersion decreases, but *unexpectedly large jumps become more probable*, thereby increasing the kurtosis.

The second somewhat paradoxical effect is the dependence of the kurtosis on the index return, which was recently interpreted by L. Borland as a signature of collective behaviour during crises [39]. Again, days when the market as a whole moves a lot are also large dispersion days where all stocks move a lot in different directions, but with little outliers, i.e. one or a handful of stocks that would jump up or down. In this sense, these days are more homogeneous. Should one deduce from this that there is a stronger "synchronisation", or collective dynamics, during these periods, as suggested by Borland? While it is true that the average correlation between stocks depends on the index return, this dependence is in fact signed: correlations are stronger for negative index returns and weaker for positive returns, see [27, 8]. This is in contrast with the kurtosis effect discussed here, which is surprisingly symmetrical (see Fig. 10.5). A quantitative model for this behaviour is missing at this stage. Qualitatively, however, we believe that the mechanism is the following [56]: when the index return is large, the dominant source of dispersion becomes the market exposure (the 'β's') of the different stocks, rather than the idiosyncratic residuals. Since the distribution of the β's is roughly Gaussian, kurtosis effects do indeed decrease for large index returns. This interpretation however requires that the volatility of the residuals increases sub-linearly with the index volatility, as indeed suggested by the data shown in Fig. 10.3. The fact that during large swings of the index, the market exposure of stocks becomes the dominant factor is probably a result of index/futures arbitrage.

Finally, although large index return days are large dispersion days, the converse is not true. A typical large dispersion day is in fact a day when one or a handful of stocks gyrate wildly, contributing both to the dispersion and to the kurtosis, and explaining the *positive* correlation between σ_d and κ_d. If this interpretation is correct, this positive correlation should diminish when one uses the mean-absolute deviation and not the variance to compute the dispersion, since the former is less sensitive to outliers. We have checked that this is indeed the case.

To summarise, we have established several new stylised facts concerning the intra-day seasonalities of stock dynamics. Beyond the well known U-shaped pattern of the volatility, we have found that the average correlation between stocks increases throughout the day, leading to a smaller dispersion between stocks (in relative terms). However, the kurtosis, which is a measure of volatility surprises, is in fact minimum at the open of the market, when the volatility is at its peak. We have also confirmed that the dispersion kurtosis is a symmetric, markedly decreasing

function of the index return. This means that during large market swings, the idiosyncratic component of the stock dynamics becomes sub-dominant, an effect that we have confirmed directly. Finally, while the market mode component of the dynamics becomes stronger as the day proceeds, the sectorial components recede. In a nutshell, early hours of trading are dominated by idiosyncratic or sector specific effects with little surprises, whereas the influence of macro, market factor increases throughout the day, and surprises become more frequent. A detailed quantitative interpretation of our results, for example of the power-law decay of the volatility in the morning, is at this stage lacking. We believe that, when available, such an interpretation will shed light on the relative importance of behavioural and informational effects on price formation and volatility.

Acknowledgements We have benefitted from insightful comments and suggestions by Rémy Chicheportiche, Stefano Ciliberti, Benoit Crespin, Laurent Laloux, Marc Potters, Pierre-Alain Reigneron and Vincent Vargas.

Bibliography

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

[2] A. Admati and P. Pfleiderer. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, 1:3 ? 40, (1988).

[3] G. Akemann, J. Baik, Ph. Di Francesco, *The Oxford Handbook of Random Matrix Theory*, Oxford University Press (2011).

[4] G. Akemann and P. Vivo, J. Stat. Mech. P09002 (2008).

[5] R. Allez, R. Rhodes, V. Vargas, *Lognormal $\star$-scale invariant random measures*, Probability Theory and Related Fields, DOI 10.1007/s00440-012-0412-9, arXiv:1102.1895v2.

[6] R. Allez, J.-P. Bouchaud and A. Guionnet, *Invariant β-ensembles and the Gauss-Wigner crossover*, Phys. Rev. Lett. **109**, 094102 (2012).

[7] R. Allez, J.-P. Bouchaud, *Eigenvector dynamics: general theory and some applications*, Phys. Rev. E **86**, 046202 (2012).

[8] P.-A. Reigneron, R. Allez and J.-P. Bouchaud, *Principal regression analysis and the index leverage effect*, Physica A **390**, 3026 (2011).

[9] R. Allez and J.-P. Bouchaud, *Individual and collective stock dynamics: intra-day seasonalities*, New Journal of Physics, **13**, 025010 (2011).

[10] R. Allez and A. Guionnet, *A diffusive matrix model for invariant β-ensembles*, http://arxiv.org/abs/1206.1460

[11] R. Allez, R. Rhodes and V. Vargas, *Marchenko Pastur type theorem for independent MRW processes: convergence of the empirical spectral measure*, http://arxiv.org/abs/1106.5891

[12] R. Allez, J.-P. Bouchaud, S. N. Majumdar, P. Vivo, *Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marčenko-Pastur law*, to appear in J. Phys. A.

[13] T. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance, **4**, 115:158, (1997).

[14] P. W. Anderson, *Infrared Catastrophe in Fermi Gases with Local Scattering Potentials*, Phys. Rev. Lett. **18**, 1049 (1967).

[15] G.W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics, Cambridge University Press (2009).

[16] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge studies in advanced mathematics **93**, Cambridge University Press, Cambridge, (2004).

[17] R. Askey and J. Wimp. *Associated Laguerre and Hermite polynomials*. Proc. Roy. Soc. Edinburgh A, **96** 15-37, (1984), see [Theorem 8.2.2].

[18] Z. D. Bai, B. Q. Miao and G. M. Pan,*On asymptotics of eigenvectors of large sample covariance matrix*, The Annals of Probability, 2007, **35** (4), 1532-1572.

[19] J. Baik, J. Silverstein, *Eigenvalues of Large Sample Covariance Matrices of Spiked Population Models*, Journal of Multivariate Analysis **97**, 1382 (2006).

[20] Z. Bai, J. Silverstein, *Spectral Analysis of Large dimensional random matrices*, Second Edition, Springer (2010), see [Theorem 9.2].

[21] E. Bacry, A. Kozhemyak and J.F. Muzy, *Log-normal continuous cascade model of asset returns: aggregation properties and estimation*, Quantitative Finance, DOI:10.1080/14697688.2011.647411.

[22] E. Bacry and J.F. Muzy, *Log-infinitely divisible multifractal processes*, Comm. Math. Phys., **236**, 449–475, (2003).

[23] E. Bacry, J. Delour and J.-F. Muzy, *Multifractal random walks*, Phys. Rev. E, **64** (2001).

[24] E. Bacry, J.-F. Muzy, *Log-infinitely divisible multifractal processes*, Comm. Math. Phys. **236** (3), 449-475, (2003).

[25] E. Bacry, A. Kozhemyak, J.-F. Muzy, *Continuous cascade models for asset returns*, Journal of Economic Dynamics and Control, **32** (2008) **1**, 156-199.

[26] J. Barral, B.B. Mandelbrot: *Multifractal products of cylindrical pulses*, *Probab. Theory Relat. Fields* **124**, 409-430, (2002).

[27] E. Balogh, I. Simonsen, B. Nagy, and Z. Neda. *Persistent collective trend in stock markets.* Phys. Rev. E **82**, 066113 (2010).

[28] L. Bauwens, S. Laurent, J. K. Rombouts, *Multivariate GARCH models: A survey*, J. Appl. Econ. **21** 79-109 (2006).

[29] J. Baik, G. Ben Arous and S. Péché, *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrix*, The Annals of Probability, **33**, 1643-1697, (2005).

[30] G. Bekaert and G. Wu. Asymmetric volatility and risk in equity markets. Rev. Fin. Stud., 13:1, 2000.

[31] S. T. Belinschi, M. Bozejko, F. Lehner, and R. Speicher, *The classical normal distribution is $\boxplus$-infinitely divisible*, Advances in Mathematics, **226**, 3677-3698 (2011).

[32] S. T. Belinschi, A. Dembo, and A. Guionnet, *Spectral measure of heavy tailed band and covariance random matrices*, Comm. Math. Phys., **289**, 1023–1055, (2009).

[33] G. Ben Arous and A. Guionnet, *The spectrum of heavy-tailed random matrices*, Comm. Math. Phys., **278**, 715–751, (2008).

[34] I. Benjamini and O. Schramm, *KPZ in one dimensional random geometry of multiplicative cascades*, Communications in mathematical physics, **2** , 653-662, (2009).

[35] L. Bergomi. Smile Dynamics II. Risk, page 67, (2005); L. Bergomi. Smile Dynamics III. Risk, page 94, (2008).

[36] P. Billingsley, *Convergence of Probability Measures*, (Wiley Series in Probability and Statistics).

[37] G. Biroli, J.-P. Bouchaud, and M. Potters. *The Student ensemble of correlation matrices: eigenvalue spectrum and Kullback-Leibler entropy.* Acta Phys. Pol. B, **38**:4009, (2007).

[38] F. Black. Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section, (1976).

[39] L. Borland. *Statistical Signatures in Times of Panic: Markets as a Self-Organizing System.* ArXiv e-prints, 0908.0111, August 2009.

[40] L. Borland and Y. Hassid. *Market panic on different time-scales.* ArXiv e-prints,1010.4917, October 2010.

[41] G. Borot, C. Nadal, *Right tail expansion of Tracy-Widom beta laws*, arXiv:1111.2761.

[42] G. Borot, B. Eynard, S. N. Majumdar et C. Nadal, *Large deviations of the maximal eigenvalue of random matrices*, J. Stat. Mech. (2011) P11024

[43] J.-P. Bouchaud and P. Cizeau, *Theory of Lévy matrices*, Phys. Rev. E, **3**, 1810-1822, (1994).

[44] J.-P. Bouchaud, M. Potters, *Financial Applications of Random Matrix Theory: a short review*, in [3].

[45] J.-P. Bouchaud, M. Potters, *Theory of Financial Risk and Derivative Pricing From Statistical Physics to Risk Management*, Cambridge University Press.

[46] J.-P. Bouchaud, L. Laloux, M. A. Miceli and M. Potters, *Large dimension forecasting models and random singular value spectra*, Eur. Phys. J. B **55** (2007) 201.

[47] J.-P. Bouchaud, *The Endogenous Dynamics of Markets: Price Impact, Feedback Loops and Instabilities*, in "Lessons from the 2008 Crisis", edited by A. Berd (Risk Books, Incisive Media, London, 2011).

[48] J.-P. Bouchaud, A. Matacz and M. Potters, *Leverage Effect in Financial Markets: The Retarded Volatility Model*, Phys. Rev. Lett., **87**, 228701, (2001).

[49] M. F. Bru, *Wishart Processes*, Journal of Theoretical Probability, **4** (4), (1991).

[50] M. F. Bru, *Diffusions of Perturbed Principal Component Analysis*, Journal of multivariate analysis, **29**, 127-136 (1989).

[51] Z. Burda, A.T. Gorlich and B. Waclaw, Phys. Rev. E **74**, 041129, (2006).

[52] B. Castaing, Y. Gagne, E. J. Hopfinger, *Velocity probability density-functions of high Reynolds-number turbulence*, Physica D, **46** 2, 177-200, (1990).

[53] B. Castaing, Y. Gagne, M. Marchand, *Conditional velocity pdf in 3-D turbulence*, J. Phys. II France **4**, 1-8, (1994).

[54] E. Cépa and D. Lépingle, *Diffusing particles with electrostatic repulsion*, Probability Theory and Related Fields, (1997).

[55] E. Cépa and D. Lépingle, *No multiple collisions for mutually repelling Brownian particles*, Séminaire de Probabilités **40** (2007) 241-246.

[56] R. Chicheportiche, J.-P. Bouchaud, *Goodness-of-Fit tests with Dependent Observations*, J. Stat. Mech. (2011).

[57] S. Ciliberti, J.-P. Bouchaud, and M. Potters. *Smile Dynamics: a Theory of the Implied Leverage Effect.* Wilmott Journal, 1:8794, 2009.

[58] P. Cizeau, M. Potters, and J.-P. Bouchaud. Correlation structure of extreme stock returns. Quantitative Finance,1:217, 2001.

[59] R. Cont, *Empirical properties of asset returns: Stylized facts and statistical issues*, Quantitative Finance, **1**, 223–236, (2001).

[60] R. Cont, L. Wagalath, *Running for the exit: distressed selling and endogeneous correlations in financial markets*, Mathematical Finance (in press).

[61] D.J. Daley, D. Vere-Jones, *An introduction to the theory of point processes*, **2**, Probability and its applications, Springer, 2nd edition, (2007).

[62] D. S. Dean, S. Majumdar, *Extreme value statistics of eigenvalues of Gaussian random matrices*, Phys. Rev. E **77**, 041108 (2008)

[63] P. Desrosier, P. Forrester, Nucl. Phys.B **743** 307-332 (2006).

[64] S.-H. Dong, A. Gonzalez-Cisneros. *Energy spectra of the hyperbolic and second Poschl-Teller like potentials solved by new exact quantization rule*, Annals of Physics 323 (2008) 1136-1149; S. Odake, R. Sasaki, *Infinitely many shape invariant potentials and new orthogonal polynomials*, Physics Letters B, **679** 414 (2009).

[65] J. Duchon, R. Robert and V. Vargas, *Forecasting volatility with the multifractal random walk model*, Mathematical Finance.

[66] B. Duplantier, S. Sheffield, *Liouville Quantum Gravity and KPZ*, to appear in *Inventiones Mathematicae*, available on arxiv at the URL http://arxiv.org/abs/0808.1560.

[67] L. Dumaz and B. Virág, *The right tail exponent of the Tracy-Widom-beta distribution*, arXiv:1102.4818, Annales de l'IHP.

[68] I. Dumitriu, A. Edelman, *Matrix Models for Beta Ensembles*, Journal of Mathematical Physics **43**, 5830–5847 (2002).

[69] F. J. Dyson, *A Brownian-Motion Model for the Eigenvalues of a Random Matrix*, J. Math. Phys. **3**, 1191 (1962).

[70] C.B. Erb, C.R. Harvey, and T.E. Viskanta. *Forecasting International Equity Correlations. Financial Analysts Journal*, **50**:32?45, (1994).

[71] see e.g.: H. Kohler, H.J. Sommers, S. Aberg, T. Guhr, *Exact Fidelity and Full Fidelity Statistics in Regular and Chaotic Surroundings*, Phys. Rev. E, **81**, 050103(R) (2010) and refs. therein.

[72] P. J. Forrester, *Log gases and Random Matrices*, Princeton University Press (2010).

[73] P.J. Forrester, *Spectral density asymptotics for Gaussian and Laguerre β-ensembles in the exponentially small region*, J. Phys. A.: Math. Theo. **45**, 075206 (2012).

[74] P. J. Forrester, E. M. Rains, *Interpretations of some parameter dependent generalizations of classical matrix ensembles*, Prob. Th. Rel. Fields **131** 1-61, (2005).

[75] P. J. Forrester, N. Frankel and T. M. Garoni, Asymptotic form of the density profile for Gaussian and Laguerre random matrix ensembles with orthogonal and symplectic symmetry, Journal of Mathematical Physics, 47, 1-26, (2006).

[76] U. Frisch, Turbulence, Cambridge University Press (1995).

[77] Y. V. Fyodorov, P. Le Doussal, A. Rosso, *Freezing transition in decaying Burgers turbulence*, Europhys. Lett. **90** (2010) 60004.

[78] Y. V. Fyodorov and H.-J. Sommers, J. Math. Phys. **38**, 1918 (1997); Y. V. Fyodorov and B. A. Khoruzhenko, Phys. Rev. Lett. **83**, 66 (1999).

[79] Y. V. Fyodorov and H.-J. Sommers, Z.Phys.B **99** 123 (1995).

[80] I.M. Gelfand, N. Y. Vilenkin, *Generalized Functions*, **4**, Academic Press, New York (1964).

[81] J. S. Geronimo and T. P. Hill, *Necessary and sufficient condition that the limit of Stieltjes transforms is a Stieltjes transform*, J. Approx. Theory, **121**, 54–60, (2003).

[82] L. Glosten, R. Jagannathan, and D. Runkle. *Relationship between the expected value and the volatility of nominal excess return.* J. Finance, **48**:1779?1801, (1993).

[83] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, San Diego, (1980).

[84] M. Hentschel, D. Ullmo, H. Baranger, *Fermi edge singularities in the mesoscopic regime: Anderson orthogonality catastrophe*, Phys. Rev. B **72**, 035310 (2005).

[85] M. E. H. Ismail and D. H. Kelkar, Siam J. Math. Anal. **10**, 884-901 (1979).

[86] J. Jacod and A.N. Shiryaev, *Limit theorems for stochastic processes*, Springer, Berlin Heidelberg New York 1987.

[87] A. T. James, Ann. Math. Statistics, **35**, 475, (1964).

[88] K. Johansson, Comm. Math. Phys. **209**, 437 (2000).

[89] K. Johansson, Duke Math. J. **91** (1998).

[90] A. Joulin, A. Lefevre, D. Grunberg, and J.-P. Bouchaud. *Stock price jumps: news and volume play a minor role*. Wilmott Magazine, Sept/Oct, 2008.

[91] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Graduate Texts in Mathematics, Springer.

[92] S. Kerov. *Interlacing measures*, in Kirillov's seminar on representation theory, Amer. Math. Soc. Transl. **181**, 35-83. Providence, RI, (1998).

[93] 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, page 73, (1992).

[94] J.-P. Kahane, Sur le chaos multiplicatif, Ann. Sci. Math. Québec, **9** (2), 105-150, (1985).

[95] J.-P. Kahane and J. Peyrière, *Sur certaines martingales de Benoit Mandelbrot*, Advances in Math., **22**(2), 131–145, (1976).

[96] F. Kalisch, D. Braak, J. Phys. A **35** (2002), 9957-9969

[97] V.G. Knizhnik, A.M. Polyakov, A.B. Zamolodchikov, *Fractal structure of 2D-quantum gravity* , Modern Phys. Lett A, **3**(8), 819-826, (1988).

[98] A. N. Kolmogorov, *The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers*, Proc. Roy. Soc. London Ser. A, 434 (1890), 913, (1991). Translated from the Russian by V. Levin, Turbulence and stochastic processes: Kolmogorov's ideas 50 years on.

[99] O. Ledoit and S. Péché, *Eigenvectors of some large sample covariance matrix ensembles*, Probability Theory Related Fields, **151** 233 (2011).

[100] Y. Li and X. Zheng, *On the estimation of integrated covariance matrices of high dimensional diffusion processes*, Annals of Statistics, **39** (6), 3121–3151, (2011).

[101] F. Lillo and R. Mantegna. *Power law relaxation in a complex system: Omori Law After a Financial Market Crash*. Physical Review E, 68:016119, 2006.

[102] F. Lillo and R. N. Mantegna. *Symmetry alteration of ensemble return distribution in crash and rally days of financial markets*. Eur. Phys. J. B, 15:603?606, 2000.

[103] F. Lillo and R. N. Mantegna. Variety and volatility in financial markets. Physical Review E, 62:6126?6134, 2000.

[104] F. Longin and B. Solnik. *Is the correlation in international equity returns constant: 1960-1990*. Journal of International Money and Finance, **14**:3?26, (1995).

[105] A. Lytova, L. Pastur, *Central limit theorem for linear eigenvalue statistics of random matrices with independent entries*, http://arxiv.org/abs/0809.4698/

[106] S. N. Majumdar, private communication, and in preparation.

[107] S. N. Majumdar and P. Vivo, Phys. Rev. Lett. **108**, 200601 (2012).

[108] S. N. Majumdar and M. Vergassola, Phys. Rev. Lett. **102**, 060601 (2009).

[109] B.B. Mandelbrot, *Multiplications aléatoires itérées et distributions invariantes par moyenne pondérée aléatoire, I and II. Comptes Rendus* (Paris): 278A, 289-292 and 355-358.

[110] B.B. Mandelbrot, *Intermittent turbulence in self-similar cascades, divergence of high moments and dimension of the carrier*, *J. Fluid. Mech.* **62**, 331-358, (1974).

[111] B. B. Mandelbrot, *Intermittent turbulence in self-similar cascades, divergence of high moments and dimension of the carrier*, J. Fluid. Mech. **62** , 331- 358, (1974).

[112] B.B. Mandelbrot, *Possible refinement of the lognormal hypothesis concerning the distribution of energy in intermittent turbulence*, Statistical Models and Turbulence. In Rosenblatt, M. and Atta, C.V. ed. *Lectures Notes in Physics*. **12**, 333–351. Springer–Verlag, New York, (1972).

[113] V.A. Marchenko and L.A. Pastur, *Distribution of eigenvalues for some sets of random matrices*, Math. USSR-Sb., **1**, 457–483, (1967).

[114] M.L. Mehta, *Random matrices* (3rd edition), Pure and Applied Mathematics Series 142, Elsevier (London - 2004).

[115] C. Monthus, A. Comtet. *On the flux distribution in a one dimensional disordered system*, J. Phys. 1 France **4** 635-653, (1994).

[116] M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, H. E. Stanley, *Identifying States of a Financial Market*, arXiv 1202.1623

[117] C. Nadal and S. N. Majumdar, Phys. Rev. E **79**, 061117 (2009).

[118] D. B. Nelson. *Conditional Heteroskedasticity in Asset Returns: A New Approach*. Econometrica.

[119] J. Perelló and J. Masoliver. *Random diffusion and leverage effect in financial markets*. Phys. Rev. E, **67**(3):037102, (2003).

[120] J. Perelló, J. Masoliver, and J.-P. Bouchaud. *Multiple time scales in volatility and leverage correlations: a stochastic volatility model*. Applied Mathematical Finance, **11**:27?50, (2004).

[121] O. Pfaffel and E. Schlemm, *Eigenvalue distribution of large sample covariance matrices of linear processes*, Arxiv e-print 1201.4134.

[122] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Amaral, T. Guhr, and H. E. Stanley. *Random matrix approach to cross correlations in financial data*. Physical Review E, **65**(6):066126, (2002).

[123] B. Podobnik, D. Wang, D. Horvatic, I. Grosse, and H. E. Stanley. *Time-lag cross-correlations in collective phenomena*. Europhysics Letters, 90, (2010).

[124] M. Potters, J.-P. Bouchaud, L. Laloux, *Financial Applications of Random Matrix Theory: Old laces and new pieces*, Acta Phys. Pol. B **36**, 2767 (2005).

[125] J. Rambeau and G. Schehr, Europhys. Lett. **91**, 60006 (2010); Phys. Rev. E, **83**, 061146 (2011).

[126] L. Ramchand and R. Susmel. *Volatility and cross-correlation across major stock markets*. Journal of Empirical Finance, **5** 397-416, (1998).

[127] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Third edition, Springer.

[128] C. Robert and M. Rosenbaum, *On the limiting spectral distribution of the covariance matrices of time-lagged processes*, Journal of multivariate analysis, **101** (31), 2434–2451, (2010).

[129] R. Robert,V. Vargas, *Gaussian Multiplicative Chaos revisited*, Annals of Probability, **38** (2), 605-631, (2010).

[130] L. C. G. Rogers and Z. Shi, *Interacting Brownian particles and the Wigner law*, Probab. Theory Relat. Fields **95**, 555-570 (1993).

[131] R. Rhodes, V. Vargas, *KPZ formula for log-infinitely divisible multifractal random measures*, to appear in ESAIM, available on arxiv at the URL http://arxiv.org/abs/0807.1036.

[132] R. Rhodes, V. Vargas, *Multidimensional multifractal random measures*, Electronic Journal of Probability, **15**, 241-258, (2010).

[133] R. Rhodes, J. Sohier, V. Vargas, Star-scale invariant random measures, http://arxiv.org/abs/1201.5219 .

[134] M. Sadek, A. Tarighat and A.H. Sayed, IEEE Trans. Signal Pro- cessing **55**, 1498 (2007).

[135] G. Schehr, S. N. Majumdar, A. Comtet and J. Randon-Furling, Phys. Rev. Lett. **101**, 150601 (2008).

[136] A. Schenzle, H. Brand, *Multiplicative stochastic processes in statistical physics*, Phys. Rev. A **20** 1629, (1979).

[137] F. Schmitt, D. Lavallee, D. Schertzer, S. Lovejoy, *Empirical determination of universal multifractal exponents in turbulent velocity fields*, Phys. Rev. Lett. **68** , 305-308, (1992).

[138] Schmitt, F., Lavallee, D., Schertzer, D., S. Lovejoy, *Empirical determination of universal multifractal exponents in turbulent velocity fields*, *Phys. Rev. Lett.* **68**, 305-308, (1992).

[139] Schwartz, L.: *Théorie des distributions*, Hermann (1997).

[140] B. D. Simon, P. A. Lee and B. A. Altshuler, *Exact Description of Spectral Correlations by a Quantum One-dimensional Model with Inverse-square Interaction*, Phys. Rev. Lett. **70** 4122, (1993).

[141] B. Solnik, C. Boucrelle, and Y. Le Fur. *International Market Correlation and Volatility.* Financial Analysts Journal, **52**:17?34, (1996).

[142] D. Sornette, Y. Malevergne, and J.F. Muzy. *What causes crashes?* Risk, Feb., 2003.

[143] G. Stolovitzky, P. Kailasnath, K.R. Sreenivasan, *Kolmogorov's Refined Similarity Hypotheses*, Phys. Rev. Lett. **69**(8), 1178-1181, (1992).

[144] J. Tenenbaum, D. Horvatic, S.C. Bajic, B. Pehlivanovic, B. Podobnik, and H.E. Stanley. *Comparison between response dynamics in transition economies and developed economies.* Physical Review E, **82**:397?416, (2010).

[145] A. Tulino, S. Verdù, *Random Matrix Theory and Wireless Communications*, Foundations and Trends in Communication and Information Theory, **1**, 1-182 (2004).

[146] A. Utsugi, K. Ino, and M. Oshikawa. *Random matrix theory analysis of cross correlations in financial markets.* Physical Review E, 70(2):026110, 2004.

[147] R. O. Vallejos, C. H. Lewenkopf, Y. Gefen, *Orthogonality catastrophe in parametric random matrices*, Phys. Rev. B, **65**, 085309.

[148] J. J. M. Verbaarschot, Phys. Rev. Lett. **72**, 2531 (1994).

[149] for an attempt in the context of 2×2 matrices, see: P. Vivo and S. N. Majumdar, *Tunable correlations in a* 2×2 *hermitian random matrix model* Physica A **387**, 4839 (2008).

[150] F. von Oppen, Phys. Rev. Lett. **73** 798 (1994).

[151] E. P. Wigner, *On the Distribution of the Roots of Certain Symmetric Matrices*, Ann. of Math., **67**, 325-327, (1957).

[152] E.P. Wigner. *On the statistical distribution of the widths and spacings of nuclear resonance levels*, Math. Proc. Cambridge Philos. Soc., 47, 790-798 (1951). xiii, 3.

[153] M. Wilkinson, P. Walker, *A Brownian motion model for the parameter dependence of matrix elements*, J. Phys. A: Math. Gen. **28**, 6143, (1995).

[154] M. Wilkinson, *Statistical aspects of dissipation by Landau-Zener transitions*, J. Phys. A: Math. Gen. 21 (1988) 4021-4037; *Diffusion and dissipation in complex quantum systems*, Phys. Rev. A 41, 4645-4652 (1990).

[155] J. Wishart, Biometrika **20**, 32 (1928).

[156] A. Zawadowski, J. Kertesz, and G. Andor. *Short-term market reaction after extreme price changes of liquid stocks.* Quantitative Finance, **6**:283?295, 2006.

[157] G. Zumbach. *The empirical properties of large covariance matrices.* RiskMetrics Journal, **9**, No. 1, Winter 2009; ArXiv e-prints, 0903.1525.