



HAL
open science

Random Matrices and Quantum Information Theory

Félix Parraud

► **To cite this version:**

Félix Parraud. Random Matrices and Quantum Information Theory. Probability [math.PR]. Université de Lyon; Kyōto daigaku, 2021. English. NNT : 2021LYSEN031 . tel-03366206

HAL Id: tel-03366206

<https://theses.hal.science/tel-03366206>

Submitted on 5 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Numéro National de Thèse : 2021LYSEN031

THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée par

l'École Normale Supérieure de Lyon

en cotutelle avec

Kyoto University

École Doctorale N° 512

École Doctorale en Informatique et Mathématiques de Lyon

Discipline : Mathématiques

Soutenue publiquement le 07/07/2021, par :

Félix PARRAUD

Matrices aléatoires et théorie de l'information quantique

Après l'avis de :

Philippe BIANE	D.R.	Université Paris-Est
Dimitri SHLYAKHTENKO	Prof.	UCLA

Devant le jury composé de :

Philippe BIANE	D.R.	Université Paris-Est	Rapporteur
Mireille CAPITAINE	C.R.	Université Paul Sabatier	Examinatrice
Mikael DE LA SALLE	C.R.	E.N.S. de Lyon	Examineur
Masanori HINO	Prof.	Kyoto University	Examineur
Masaki IZUMI	Prof.	Kyoto University	Examineur
Mylène MAÏDA	Prof. Univ.	Université de Lille	Examinatrice
Alice GUIONNET	D.R.	E.N.S. de Lyon	Directrice de thèse
Benoît COLLINS	Prof.	Kyoto University	Co-tuteur de thèse

Summary:

This PhD lies at the intersection of Random Matrix Theory and Free Probability Theory. The study of random matrices originated from outside of the field of mathematics, first in statistics and data analysis with Wishart [1], and then during the second part of the twentieth century in physics with Wigner [2]. As time went on, more and more connections with other fields of mathematics and physics were made. In particular, Free Probability emerged as a new field between Operator Algebra and Random Matrix Theory. It was created in the 1980s by Dan Voiculescu as a non-commutative probability theory where the notion of freeness plays the role of independence. This notion of freeness is intimately related with the notion of freeness in groups. In the seminal paper [3] published in 1991, he realized that the asymptotic behavior of many random matrices can be described within the framework of Free Probability. He studied GUE random matrix, which are Hermitian random matrices with Gaussian coefficients of variance 1, and whose law is invariant by the action by conjugation of the unitary group. He showed that a family of independent GUE matrices is asymptotically free in the sense that the trace of their joint moments converges towards those of free semicircular variables. The later can be described in terms of non-crossing pair partitions. A consequence of Voiculescu's result is that the empirical measure of the eigenvalues of a polynomial in independent GUE matrices, provided that it is a self-adjoint matrix, converges weakly towards the spectral measure of the same polynomial evaluated in free semicircular variables. Free Probability has been widely used since then in Random Matrix Theory, see for example chapter 22 of [4] for a more detailed survey. The link between Random matrix Theory and Operator Algebra is extremely fruitful in both ways and applies beyond these two fields (see for example [5] for a survey of the different techniques used in Quantum Information Theory). One of the spectacular application of the uses of Random matrix Theory in Operator Algebra was provided by Haagerup and Thorbjørnsen in [6] who proved by using random matrices that the Ext-invariant for the reduced C^* -algebra of the free group on 2 generators is not a group but only a semi-group.

The first chapter of this manuscript is a brief introduction to the fields of Random Matrix Theory and Free Probability, as well as the papers that we present in this manuscript. In the second chapter, which is based on [7], we got, through new methods, an estimate on the non renormalized trace of some smooth functions in GUE matrices which in turns gives us a different proof of the main result of Haagerup and Thorbjørnsen in [6]. That is, the convergence of the norm of any polynomials in independent GUE random matrices. Our result also allows to consider polynomials in GUE matrices, deterministic matrices and tensor matrices. In the third chapter, based on [8], we prove similar results but for Haar unitary matrices instead of GUE matrices by adapting our method to the unitary case. Although some of those results were already proved in [9], this new method has the advantage of giving us quantitative estimates. In the fourth chapter, which is based on [10], we refine the method used in the first paper to prove a finite Taylor expansion around the dimension N of the non renormalized trace of some smooth functions in GUE matrices of size N . This also gives several interesting corollary on the local spectrum of polynomials in independent GUE matrices. In the fifth chapter, based on [11], we use the estimates obtained in [8] to get measure concentration estimates on a random set. This set is defined as the output of a very specific quantum channel which is related to an important problem in Quantum Information Theory. The concentration estimates that we obtained give us explicit parameters of the size of the solutions. Finally in the last chapter, which is based on the paper [12], we expand Voiculescu's result of convergence in law of the empirical measure

of non-commutative polynomials in independent GUE matrices to non-commutative rational function.

Résumé:

Ce doctorat est au croisement de la théorie des matrices aléatoires et de celle des probabilités libres. Les matrices aléatoires ont d'abord été étudiées en dehors du cadre des mathématiques théoriques, notamment en statistiques et en analyse de données grâce aux travaux de Wishart (voir [1]), puis durant la seconde moitié du vingtième siècle, en physique avec Wigner (voir [2]). Au fur et à mesure du temps, de plus en plus de connexions se sont faites entre la théorie des matrices aléatoires et d'autres domaines des mathématiques et de la physique. Notamment, les probabilités libres sont un domaine récent des mathématiques à mi-chemin entre la théorie des algèbres d'opérateurs et celle des matrices aléatoires. Sa naissance remonte aux années 1980 via les travaux de Dan Voiculescu qui inventa cette théorie des probabilités non commutatives où la notion de liberté remplace celle d'indépendance dans la théorie classique des probabilités. Comme son nom l'indique, la notion de liberté en probabilité libre est fortement reliée à celle que l'on définit dans la théorie des groupes. Dans son article fondateur [3] publié en 1991, il démontra que le comportement asymptotique de nombre de modèles de matrice aléatoire peut être décrit dans le cadre des probabilités libres. Il étudia la matrice du GUE, qui est l'unique matrice aléatoire Hermitienne dont les coefficients sont des Gaussiennes de variance 1, et dont la loi est invariante par conjugaison par des matrices unitaires. Il démontra qu'une famille de matrices du GUE indépendantes est asymptotiquement libre au sens où la trace d'un monôme en ces matrices converge vers la trace du même monôme mais évalué en des semicirculaires libres. Cette dernière peut être définie à l'aide d'outils combinatoires. Une conséquence du résultat de Voiculescu est que la mesure empirique définie à partir des valeurs propres d'un polynôme évalué en des matrices du GUE indépendantes, sous l'hypothèse que ce polynôme de matrices est lui-même une matrice auto-adjointe, converge en loi vers la mesure spectral du même polynôme évalué en des semicirculaires libres. Les probabilités libres ont aujourd'hui de nombreuses applications en théorie des matrices aléatoires, voir par exemple le chapitre 22 de [4] pour une étude détaillée. Le lien entre la théorie des algèbres d'opérateurs et celle des matrices aléatoires a généré de nombreux résultats qui ont eux-mêmes des applications dans d'autres domaines (voir par exemple [5] pour une étude exhaustive des différentes méthodes en lien avec les deux précédents domaines qui sont utilisées en théorie de l'information quantique). Une application particulièrement spectaculaire de l'utilisation des matrices aléatoires en théorie des algèbres d'opérateurs nous est donnée par Haagerup et Thorbjørnsen dans leur article [6] dans lequel ils démontrent que l'invariant extérieur de la C^* -algèbre réduite du groupe libre à deux générateurs n'est pas un groupe mais seulement un semi-groupe.

Le premier chapitre de ce manuscrit est une courte introduction à la théorie des matrices aléatoires et de celle des probabilités libres, ainsi que des articles écrits pendant ma thèse. Dans le second chapitre, qui se base sur l'article [7], nous calculons grâce à de nouvelles méthodes une estimée de la trace non renormalisée de certaines fonctions lisses évaluées en des matrices du GUE indépendantes. Nous en déduisons une nouvelle preuve du résultat principal de Haagerup et Thorbjørnsen dans leur article [6]. C'est-à-dire la convergence de la norme d'un polynôme évalué en des matrices du GUE indépendantes. Notre preuve nous permet également de prouver un résultat similaire pour des polynômes évalués en des matrices du GUE indépendantes, des matrices déterministes, ainsi que des tenseurs de matrices. Dans le troisième chapitre, qui se base sur [8], nous démontrons des résultats similaires mais pour des matrices unitaires de Haar au lieu des matrices du GUE. Notre approche consista à adapter notre méthode au cas unitaire. Bien que certains de ces résultats aient déjà été démontrés dans [9], cette nouvelle méthode à

l'avantage de donner des estimées quantitatives. Dans le quatrième chapitre, qui se base sur l'article [10], nous avons amélioré la méthode utilisée dans notre premier article [7] pour prouver un développement limité (en la dimension N de nos matrices aléatoires) de la trace de certaines fonctions lisses évaluées en des matrices du GUE. Nous en déduisons plusieurs corollaires sur le spectre local d'un polynôme en des matrices du GUE indépendantes. Dans le cinquième chapitre, qui se base sur l'article [11], nous utilisons les résultats obtenus précédemment dans [8] pour en déduire des estimées de concentration de la mesure d'un ensemble aléatoire. Cet ensemble correspond aux données de sortie d'un canal quantique qui a été utilisé comme contre-exemple pour résoudre un problème important en théorie de l'information quantique. Cette estimée nous permet d'obtenir des paramètres explicites sur la taille des solutions à ce problème. Finalement dans le dernier chapitre, qui se base sur [12], nous élargissons le résultat de Voiculescu de convergence en loi de la mesure empirique d'un polynôme non commutatif évalué en des matrices du GUE indépendantes à l'ensemble des expressions rationnelles.

Acknowledgments:

Naturally I would like to thank first and foremost my PhD supervisors, Benoît Collins and Alice Guionnet without whom I would never have been able to go anywhere with my research. Their guidance when I was stuck was always really helpful, and I mean it. I would like to point out their patience and ability to make time for me even when busy. Especially since handling the administrative work and the logistic of a cosupervised PhD between France and Japan has proved to be sometimes challenging and time consuming.

Similarly I would also like to thank Mikael de la Salle for numerous fruitful conversations during the time I spent at the UMPA. He gave me several good lead for some lemma I was struggling with. I am also grateful towards him as well as Grégory Miermont for writing letters of recommendation for my applications to postdocs. Grégory Miermont was also a professor I did some teaching duty for, and I thank him for making it as easy as possible by providing a lot of teaching materials and organizing weekly meetings to talk about it. Among the people that helped me quite a bit with this PhD, there are obviously my co-authors. First my supervisors, Benoît Collins and Alice Guionnet, but also Tobias Mai, Akihiro Miyagawa and Sheng Yin. I would also like to thank Philippe Biane and Dima Shlyakhtenko for accepting to review this manuscript. Their feedback was especially useful to perfect this manuscript. As well as Mireille Capitaine, Mikael de la Salle, Masaki Izumi, Mylene Maida, Masanori Hino for being part of the dissertation committee along with Philippe Biane and my supervisors.

More generally I would like to thank the UMPA and the lab of Kyoto university as a whole. The administration department was always on point when I had questions. And I had many. This definitely made my travel much easier and the regular paperwork way less tiresome. The atmosphere at my lab was always very good and I will definitely miss having lunch with the different PhD or Postdoc students from the UMPA, as well as the regular cake, beer or dinner coupled with board games and not so serious conversations. I would like to thank in particular Mickaël Maazoun and Benoit Loisel with whom I shared an office for most of my PhD, but there are many many more people that I intend on continuing to see regularly.

There were obviously many people outside of mathematics which were of great help during my PhD. Christian A in particular, with whom I have been friend for now more than twenty years and that has the impressive ability to always lift up my mood when we meet. Song Hui K, the first friend that I made in Japan, thank you for helping me to have such a great first stay and motivating me to invest myself in the student organization to help international students when I came back to France. By the way, I have to thank the people who shared with me this memorable experience that the ASSET was, Laura B, Elodie H, Ophélie S, Blanche T. Just by yourselves, you made the time invested in this association more than worth it. Through the ASSET, I also met Denis T with whom I always have a great time when we manage to meet. Anatole E, with whom I made the entirety of my studies in Lyon, including my PhD. Antoine H, with whom I will definitely meet again since we still have to go on a hike together.

I also have to thank my flatmates with whom I spent so much time during my PhD. First, Joachim C, Thibaut B and Isabel M from my last year of master, hopefully we can meet again to share a bottle of cider as usual. During those last three years, I spent most of my time in France at the 231 avenue Jean Jaurès, thank you to my past, present or almost flatmates, Quentin B, Mathieu D, Nassima M, Marie C and Léna D. Thank you for always being motivated to spend time together in the common room, and the memorable party that we had. In Japan, I would like to thank the many people of HdB for making my life as a PhD student so much

fun. Thank you in particular to Mathieu F for introducing me to HdB, Rutvika M for trying so hard to correct my english accent, Richard W for being himself.

Et enfin, merci à ma famille, merci de toujours m'avoir motivé à faire ce que je fais.

Contents

1	Introduction	1
1.1	Random Matrix Theory and Free Probability	2
1.1.1	A short introduction to Random Matrix Theory	2
1.1.2	A short introduction to Free Probability	3
1.1.3	Seminal works preceding our research	5
1.1.4	The Stein’s method for free probability	6
1.2	Summary of research results	9
1.2.1	On the operator norm of noncommutative polynomials in deterministic matrices and iid GUE matrices	9
1.2.2	On the operator norm of noncommutative polynomials in deterministic matrices and iid Haar unitary matrices	12
1.2.3	Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices	13
1.2.4	Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory	16
1.2.5	Convergence for noncommutative rational functions evaluated in random matrices	18
2	On the operator norm of noncommutative polynomials in deterministic matrices and iid GUE matrices	21
2.1	Introduction	22
2.2	Framework and standard properties	28
2.2.1	Usual definitions in free probability	28
2.2.2	Non-commutative polynomials and derivatives	30
2.2.3	GUE random matrices	32
2.3	Proof of Theorem 2.1.1	33
2.3.1	Overview of the proof	33
2.3.2	Proof of Theorem 2.1.1	35
2.4	Consequences of the main result	53
2.4.1	Proof of Corollary 2.1.3	53
2.4.2	Proof of Corollary 2.1.4	53
2.4.3	Proof of Theorem 2.1.2	54
2.4.4	Proof of Theorem 2.1.5	60
3	On the operator norm of noncommutative polynomials in deterministic matrices and iid Haar unitary matrices	69
3.1	Introduction	70

3.2	Framework and standard properties	75
3.2.1	Usual definitions in free probability	75
3.2.2	Non-commutative polynomials and derivatives	76
3.2.3	Free stochastic calculus	78
3.2.4	Notations	82
3.2.5	Random matrix model	83
3.3	Preliminaries	84
3.3.1	A matricial inequality	84
3.3.2	A Poincaré type equality	85
3.3.3	Convergence of the free unitary Brownian motion	90
3.3.4	Free stochastic calculus and free unitary Brownian motion	92
3.4	Proof of Theorem 3.1.1, the main result	95
3.4.1	Overview of the proof	95
3.4.2	Proof of Theorem 3.4.1	97
3.5	Proof of Corollaries	109
3.5.1	Proof of Corollary 3.1.3	109
3.5.2	Proof of Corollary 3.1.4	110
3.5.3	Proof of Theorem 3.1.2	111
4	Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices	115
4.1	Introduction	116
4.2	Framework and standard properties	120
4.2.1	Usual definitions in free probability	120
4.2.2	Non-commutative polynomials and derivatives	121
4.2.3	GUE random matrices	125
4.3	Proof of Theorem 4.1.1	126
4.3.1	A Poincaré type equality	126
4.3.2	A first rough formulation of the coefficients	128
4.3.3	Proof of Theorem 4.1.1	135
4.4	Consequences of Theorem 4.3.4	145
4.4.1	Proof of corollary 4.1.2	145
4.4.2	Proof of Corollary 4.1.3	145
5	Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory	149
5.1	Introduction	150
5.2	Notations and main theorem	150
5.3	Proof of main theorem	153
5.4	Application to Quantum Information Theory	161
5.4.1	Preliminaries on entropy	161
5.4.2	Corollary of the main result	162
5.4.3	Application to violation of the Minimum Output Entropy of Quantum Channels	162
6	Convergence for noncommutative rational functions evaluated in random matrices	165

6.1	Introduction	166
6.2	Preliminaries	168
6.2.1	Noncommutative rational functions and expressions	168
6.2.2	Linearization	171
6.2.3	Self-adjointness for matrix-valued noncommutative rational expressions	175
6.2.4	Unbounded random variables	178
6.2.5	The quantity Δ	179
6.3	Evaluations of non-degenerate matrix-valued noncommutative rational expressions	180
6.3.1	Evaluations in random matrices	180
6.3.2	Evaluation in operators with maximal Δ	183
6.4	Convergence in law of the spectral measure	184
6.4.1	Estimate on the cumulative distribution function of the spectral measure of self-adjoint operators	184
6.4.2	Main result	185

Bibliography

Chapter 1

Introduction

This chapter is a brief introduction to the fields of Random Matrix Theory and Free Probability, as well as a presentation of the papers that are contained in this manuscript. We first introduce usual definitions and notations that arise regularly in this manuscript. Then we summarize important results of older papers on which this PhD is built on. Finally we present the research results that I and my co-authors obtained and how they connect to each others.

1.1 Random Matrix Theory and Free Probability

1.1.1 A short introduction to Random Matrix Theory

One of the most studied model of random matrices is the Gaussian Unitary Ensemble (GUE), one can view it as the natural generalization of the Gaussian random variable to the vector space of Hermitian matrices. Indeed if dA denotes the Lebesgue measure on the set of $N \times N$ Hermitian matrices, then a $N \times N$ self-adjoint random matrix is said to be a GUE matrix if its distribution is proportional to the measure $\exp(-N \operatorname{Tr}(A^2)/2) dA$. Another way to define a GUE matrix is as the unique random Hermitian matrix whose coefficients are all independent Gaussian and whose law is invariant by conjugation by any deterministic unitary matrix. This random matrix model has been studied thoroughly, in particular we have an explicit expression for the probability density of the distribution of the eigenvalues of a GUE random matrix.

Proposition 1.1.1. *Let X^N be a GUE random matrix, if $\lambda_1 \leq \dots \leq \lambda_N$ are its eigenvalues, then with Z_N a renormalization constant, for any $A \in \mathbb{R}^N$,*

$$\mathbb{P}(\lambda_1, \dots, \lambda_N \in A) = \frac{1}{Z_N} \int_A \mathbf{1}_{x_1 \leq \dots \leq x_N} \prod_{1 \leq i < j \leq N} |x_j - x_i|^2 \prod_{1 \leq i \leq N} e^{-\frac{x_i^2}{2}} dx_1 \dots dx_N. \quad (1.1)$$

For a proof, see Theorem 2.5.2 from [13]. GUE random matrices are at the heart of my research. More precisely, I study polynomials in several independent copies of these matrices, which depend not only on the eigenvalues but also on the eigenvectors. Similarly, I also study Haar unitary matrices, which in fact describe the joint distribution of the eigenvectors of GUE matrices. A matrix is said to be a Haar unitary matrix of size N if its law is simply the Haar measure on the group of unitary matrices of size N . Like the GUE, this random matrix has been extensively studied and we have an explicit expression of its eigenvalue probability density. However in the world of Random Matrix Theory, it is very rare to actually have such exact formula. For most random matrix models, in order to study their eigenvalues we have to rely on different methods. To do so, we study instead the so-called empirical measure.

Definition 1.1.2. *Let A be a self-adjoint matrix of size N , let $\lambda_1 \leq \dots \leq \lambda_N$ be its eigenvalues, then one defines the empirical measure associated to A by*

$$\mu_A = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i},$$

where δ_λ is the Dirac measure in λ .

The link between the empirical measure and the eigenvalues is rather straightforward, if $\sigma(A)$ is the spectrum of A , then for any interval $[a, b]$,

$$\mu_A([a, b]) = \frac{\#\{\lambda \in \sigma(A) \mid a \leq \lambda \leq b\}}{N}.$$

Thus if A_N is a sequence of random matrices, one gets a sequence of random measures μ_{A_N} . Then one usually proceeds to show that almost surely this sequence converges in law towards some deterministic measure μ . Thus, assuming that the limiting measure has no atom, by

Portmanteau theorem this implies that the proportion of eigenvalues in the segment $[a, b]$ converges towards $\mu([a, b])$. The natural way to show the convergence of the empirical measure is to exploit the connection with traces, indeed if f is a continuous function, then

$$\mu_A(f) = \frac{1}{N} \sum_{\lambda \in \sigma(A)} f(\lambda) = \frac{1}{N} \text{Tr}(f(A)). \quad (1.2)$$

This means that to prove the convergence in law of the empirical measure, we simply need to prove the convergence of traces of smooth functions. Besides, assuming that the limiting measure has a compact support, which is the case for the overwhelming majority of models of random matrix studied, since one can approximate smooth functions by polynomials, it is sufficient to prove the convergence of the traces of the moments of our random matrices. This is where the connection with free probability starts.

1.1.2 A short introduction to Free Probability

In a nutshell, Free Probability stems from applying definitions and strategies from probability in operator algebra by viewing operators as random variables. In particular, if the Gaussian random variable is one of the most important random variable in classical probability, then the semicircular variable is just as important in Free Probability. It is defined as follows,

Definition 1.1.3. *If \mathcal{A} is a C^* -algebra endowed with a trace τ , then $x \in \mathcal{A}$ is said to be a semicircular variable if it is self-adjoint and for any $k \in \mathbb{N}$,*

$$\tau(x^k) = \frac{1}{2\pi} \int_{-2}^2 t^k \sqrt{4-t^2} dt. \quad (1.3)$$

It has been known for a very long time that almost surely the trace of the moments of a GUE random matrix converges towards the trace of the moments of a semicircular variable. This prompts us to define the notion of convergence in distribution.

Definition 1.1.4. *let $A = (a_1, \dots, a_k)$ be a k -tuple of elements of a C^* -algebra endowed with a trace τ , we denote those elements as noncommutative random variables. The joint distribution of the family A is the linear form*

$$P \in \mathbb{C}\langle X_1, \dots, X_{2k} \rangle \mapsto \tau[P(A, A^*)] \in \mathbb{C}$$

on the set of polynomials in $2k$ noncommutative indeterminates.

Definition 1.1.5. *Let $(A_N)_{N \geq 1} = (a_1^N, \dots, a_k^N)_{N \geq 1}$ be a sequence of families of noncommutative variables in C^* -algebras $\mathcal{A}_N, *$ endowed with a trace τ_N , we say that A_N converges in distribution towards A if the map*

$$P \mapsto \tau_N[P(A_N, A_N^*)]$$

converges pointwise towards the map

$$P \mapsto \tau_N[P(A, A^*)].$$

In Random Matrix Theory, one usually has $\mathcal{A}_N = \mathbb{M}_N(\mathbb{C})$ and $\tau_N = N^{-1} \text{Tr}$. Thus almost surely a GUE random matrix converges in distribution towards a semicircular variable. This definition also applies to the Haar unitary matrix. Indeed $u \in \mathcal{A}$ is said to be a free Haar unitary if $uu^* = u^*u = 1$ and for any $k \in \mathbb{Z}$, $\tau(u^k) = \mathbf{1}_{k=0}$. And a Haar unitary matrix almost surely converges in distribution towards a free Haar unitary. However, both of those results are more of a reformulation of long-known properties of GUE and Haar unitary matrices and do not justify by themselves the introduction of the notion of convergence in distribution. It is only when we start working with families of several random matrices that this definition fully makes sense.

Dan Voiculescu paved the way to tackle this kind of problem with the concept of freeness. In order to explain it, let us draw a parallel between classical and free independence. Indeed one can define classical independence as such, if \mathcal{A} is a σ -algebra endowed with an expectation \mathbb{E} , the sub- σ -algebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ are said to be independent if for any $I \subset [1, n]$ and $A_i \in \mathcal{A}_i$,

$$\mathbb{E} \left[\prod_{i \in I} (\mathbf{1}_{A_i} - \mathbb{E}[\mathbf{1}_{A_i}]) \right] = 0.$$

This prompts us to define the following notion of free independence.

Definition 1.1.6. *If \mathcal{A} is a \mathcal{C}^* -algebra endowed with a trace τ , the sub-algebras $\mathcal{A}_1, \dots, \mathcal{A}_n$ are said to be freely independent if for any $k \geq 1$, for any i_1, \dots, i_k such that for any j , $i_j \neq i_{j+1}$ and $a_j \in \mathcal{A}_{i_j}$,*

$$\tau \left(\prod_{1 \leq j \leq k} (a_j - \tau(a_j)) \right) = 0.$$

The fact that we need to take an unlimited number k of indices, unlike the classical case where $k = n$ is enough, is due to non-commutativity. While given several \mathcal{C}^* -algebra \mathcal{A}_i one can build an explicit \mathcal{C}^* -algebra which contains free copies of the \mathcal{A}_i , in Random Matrix Theory we usually do not use this construction. Indeed, in classical probability we can build an explicit probability space which contains copies of independent random variables but we never use it beyond the fact that it exists, we only work with the distribution of the random variables. Similarly, in Random Matrix Theory we are mainly interested by the existence of free noncommutative random variables, and the previous formula gives a way to compute the trace of their mixed moments. If (x_1, \dots, x_d) are free noncommutative random variables, and P is a self-adjoint polynomial, that is such that $P(x_1, \dots, x_d, x_1^*, \dots, x_d^*)$ is self-adjoint, then the trace of its moments define a unique measure μ_P such that

$$\tau(P(x_1, \dots, x_d, x_1^*, \dots, x_d^*)^k) = \int t^k d\mu_P(t).$$

More rigorously, one defines μ_P by using usual notion of operator algebra. Notably if y is a self-adjoint element of of a \mathcal{C}^* -algebra, then by functional calculus, for every continuous function f , one can define the operator $f(y)$ such that the map $f \in \mathcal{C}^0(\mathbb{R}) \mapsto f(y)$ is a morphism of algebra.

Definition 1.1.7. *Let y be a self-adjoint element of a \mathcal{C}^* -algebra endowed with a trace τ . Then by the Riez theorem, there exist a unique probability measure ν_y , whose support is bounded by $\|y\|$, such that for any $f \in \mathcal{C}^0(\mathbb{R})$,*

$$\tau(f(y)) = \int f d\nu_y.$$

We say that ν_y is the spectral measure of y . Then we set $\mu_P = \nu_{P(x_1, \dots, x_d, x_1^*, \dots, x_d^*)}$.

It turns out that thanks to equation (1.2), approximation by polynomials coupled with the fact that the spectral measure of an element of a \mathcal{C}^* -algebra is always compactly supported, the convergence in distribution of a family of matrices $X^N = (X_1^N, \dots, X_d^N)$ towards $x = (x_1, \dots, x_d)$ implies that for any self-adjoint polynomial P , the empirical measure of $P(X^N, X^{N*})$ converges in law towards the spectral measure of $P(x, x^*)$. As previously mentioned this gives an estimate on the proportion of eigenvalues of $P(X^N, X^{N*})$ in a given segment. It was only until 1991 that the first result of convergence in distribution linking Free Probability and Random Matrix Theory was proved by Dan Voiculescu in the seminal work [3]:

Theorem 1.1.8. *Let $X^N = (X_1^N, \dots, X_d^N)$ be a family of independent GUE matrices, then almost surely it converges in distribution towards a family $x = (x_1, \dots, x_d)$ of free semicircular variables.*

Later in [14], he proved a similar result for independent Haar unitary matrices and free Haar unitaries. While this result is quite important, it leaves a few questions open. For example, it does not give an approximation of the number of eigenvalues in an interval whose size converges towards 0. It also does not tell anything about the existence of outliers, that is whether the support of the empirical measure converges towards the support of the limiting measure, or in other words, whether there are eigenvalues outside of the support of a neighborhood of the limiting measure.

1.1.3 Seminal works preceding our research

In a few words, the convergence of the empirical measure of the eigenvalues of a matrix does not say much about the local properties of its spectrum. When dealing with a single matrix, very precise results are known. For example it is well-known that the largest eigenvalue of a GUE random matrix converges almost surely towards 2. More precisely, if X_N is a GUE random matrix of size N , then almost surely

$$\lim_{N \rightarrow \infty} \|X_N\| = 2.$$

The proof, for the more general case of a Wigner matrix with entries with finite moments, was given in [15]. This result was later obtained under the optimal assumption that their fourth moment is finite in [16]. Concerning the GUE, much more precise results were obtained by Tracy and Widom in the early nineties in [17]. The main result of their paper is the existence of a continuous decreasing function F_2 from \mathbb{R} to $[0, 1]$ such that if $\lambda_1(X^N)$ denotes the largest eigenvalue of X^N ,

$$\lim_{N \rightarrow \infty} P\left(N^{2/3}(\lambda_1(X^N) - 2) \geq s\right) = F_2(s).$$

This was recently generalized to Wigner matrices [18, 19, 20, 21] up to optimal hypotheses. One can as well study the localization of the eigenvalues in the bulk as well as their fluctuations [22, 19].

On the other hand, there are much less results available when one deals with a polynomial in several random matrices. In fact, up to today, the only local fluctuations results concern perturbative polynomials [23] or local laws [24] under some assumptions which are shown to hold for homogeneous polynomials of degree two. However, a beautiful breakthrough was made

in 2005 by Haagerup and Thorbjørnsen [6]: they proved the almost sure convergence of the norm of a polynomial evaluated in independent GUE matrices. For P a self-adjoint polynomial, they proved that almost surely, for N large enough,

$$\sigma\left(P(X_1^N, \dots, X_d^N)\right) \subset \text{Supp } \mu_P + (-\varepsilon, \varepsilon), \quad (1.4)$$

where $\sigma(H)$ is the spectrum of H and $\text{Supp } \mu_P$ the support of the measure μ_P . This is equivalent to saying that for any polynomial P , $\|P(X_1^N, \dots, X_d^N)\|$ converges almost surely towards $\sup\{|x| \mid x \in \text{Supp } \mu_P\}$. The result (1.4) was a major progress in free probability and was refined in multiple ways. In [25], Schultz used the method of [6] to prove the same result with Gaussian orthogonal or symplectic matrices instead of Gaussian unitary matrices. In [26], Capitaine and Donati-Martin proved it for Wigner matrices under some technical hypothesis on the law of the entries. This result itself was then extended by Anderson in [27] to remove most of the technical assumptions. In [28], Male made a conceptual improvement to the result of Haagerup and Thorbjørnsen, by allowing to work both with GUE and deterministic matrices. Finally, Belinschi and Capitaine proved in [29] that one could even work with Wigner and deterministic matrices, while keeping the same assumptions on the Wigner matrices as Anderson. As for the unitary case, Collins and Male proved in [9] the same result with unitary Haar matrices instead of GUE matrices by using Male's former paper.

With the exception of [9], all of these results are essentially based on the method introduced by Haagerup and Thorbjørnsen. Their first tool is called the linearization trick: it allows to relate the spectrum of a polynomial of degree d with coefficients in \mathbb{C} by a polynomial of degree 1 with coefficients in $\mathbb{M}_{k(d)}(\mathbb{C})$ where $k(d)$ is an integer which only depends on d . The second idea to understand the spectrum of this larger matrix is to study its Stieltjes transform close to the real axis by using the Dyson-Schwinger equations. An issue of this method is that it does not allow to retrieve easily sharp quantitative estimates and it is hard to generalize to other models such as Haar unitary matrices. Solving this issue ended up being the starting point of my PhD. In order to do so, we had to find an alternative proof of Haagerup and Thorbjørnsen's result of convergence of the norm.

1.1.4 The Stein's method for free probability

In this subsection we present the heuristic behind many of the proofs in my PhD. The Stein's method is an old trick from classical probability introduced by Charles Stein in [30] to estimate the difference between the distribution of a random variable Y and a Gaussian distribution. Although this is not the approach we initially took in the beginning of my PhD, through a discussion with Martin Venker, it turned out that one could view the method that we were developing and using as an adaptation of the Stein's method to a free probability setting. More specifically this explains why we can expect interesting results when we interpolate random matrices with their free limit through an Ornstein-Uhlenbeck process. Stein's method can be summarized as such, it is known that Z is a centered Gaussian variable of variance 1 if and only if for any smooth function f ,

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[f'(Z)]. \quad (1.5)$$

Thus if we set $\mathcal{A} : f \in \mathcal{C}^\infty(\mathbb{R}) \mapsto (x \rightarrow f'(x) - xf(x))$, the Gaussian distribution is the unique solution to the system of equations $\{\forall f \text{ smooth}, \mathbb{E}[\mathcal{A}f(Z)] = 0\}$. One can wonder

if there is some sort of continuity, that is if Y is such that for any function f , $\mathbb{E}[\mathcal{A}f(Y)]$ is small, then how close from a Gaussian distribution is the one of Y ? Indeed, this intuition is confirmed by stochastic calculus. If we consider $(B_t)_{t \geq 0}$ a Brownian motion, then $(X_t)_{t \geq 0}$ is an Ornstein-Uhlenbeck process started in Y if it satisfies

$$\forall t \geq 0, \quad X_t = Y - \frac{1}{2} \int_0^t X_s ds + B_t.$$

It is known that independently of the distribution of Y , as t goes to infinity, X_t converges in law towards a Gaussian random variable Z . So basically $X_0 = Y$, and $X_\infty = Z$. Besides an Ornstein-Uhlenbeck process is a Markov process whose generator is \mathcal{A} composed with the differentiation, without going into detail for the computations, thanks to stochastic calculus,

$$\begin{aligned} \mathbb{E}[f(Z)] - \mathbb{E}[f(Y)] &= \frac{1}{2} \int_0^\infty \mathbb{E}[\mathcal{A}f'(X_t)] dt \\ &= \frac{1}{2} \int_0^\infty e^{-t/2} \mathbb{E}[\mathcal{A}g_t(Y)] dt, \end{aligned}$$

where $g_t : x \rightarrow \mathbb{E}[f'(e^{-t/2}x + (1 - e^{-t})^{1/2}Z)]$. Hence if $\mathbb{E}[\mathcal{A}f(Y)]$ is small for any smooth f , we do get that the distribution of Y is close from a Gaussian one.

It is natural to generalize Stein's method to the matricial case. Instead of a Gaussian random variable we will take a GUE random matrix X of size N . By definition, its coefficients are all independent Gaussian random variable (except to respect the self-adjointness) of variance $N^{-1/2}$, real on the diagonal and complex off-diagonal. Then thanks to equation (1.5), with $\tau_N = N^{-1} \text{Tr}$ and $E_{r,s}$ the matrix whom every coefficient is 0 but the one at line r and column s which is 1, for any $p \geq 0$,

$$\begin{aligned} \mathbb{E}[\tau_N(X^p)] &= \frac{1}{N} \sum_{r,s} \mathbb{E}[X_{r,s} \text{Tr}(E_{r,s} X^{p-1})] \\ &= \frac{1}{N^2} \sum_{r,s} \sum_{i=1}^{p-1} \mathbb{E}[\text{Tr}(E_{r,s} X^{i-1} E_{s,r} X^{p-i-1})] \\ &= \mathbb{E} \left[\sum_{i=1}^{p-1} \tau_N(X^{i-1}) \tau_N(X^{p-i-1}) \right]. \end{aligned}$$

Hence a GUE matrix satisfies the system of equations

$$\forall p \geq 0, \quad \mathbb{E} \left[\tau_N(X^p) - \sum_{i=1}^{p-1} \tau_N(X^{i-1}) \tau_N(X^{p-i-1}) \right] = 0. \quad (1.6)$$

Wigner matrices are defined similarly to GUE matrices but without the assumption that the coefficients are Gaussian. It turns out that if Y is a centered random variable of variance 1, with all moments finite, one can show that for any smooth functions f ,

$$\mathbb{E}[Y f(Y)] = \mathbb{E}[f'(Y)] + \mathcal{O}(\|f^{(2)}\|).$$

Thus it turns out that with the same kind of computation, if W is a Wigner matrix then

$$\forall p \geq 0, \quad \mathbb{E} \left[\tau_N(W^p) - \sum_{i=1}^{p-1} \tau_N(W^{i-1}) \tau_N(W^{p-i-1}) \right] = \mathcal{O}(N^{-1/2}). \quad (1.7)$$

Once again, similarly to the scalar case, one can define a Hermitian Brownian motion $(H_t)_{t \geq 0}$ which is defined similarly to GUE random matrices but with Brownian motions instead of Gaussian random variables. Then we can define a Hermitian Ornstein-Uhlenbeck process started in W by

$$X_t = W - \frac{1}{2} \int_0^t X_s ds + H_t.$$

Naturally, the distribution of a Hermitian Ornstein-Uhlenbeck process converges towards the one of a GUE random matrix since every coefficients are scalar Ornstein-Uhlenbeck process. Then finally through stochastics calculus,

$$\mathbb{E} [\tau_N(W^p)] - \mathbb{E} [\tau_N(X^p)] = \frac{p}{2} \int_0^\infty \mathbb{E} \left[\tau_N(X_t^p) - \sum_{i=1}^{p-1} \tau_N(X_t^{i-1}) \tau_N(X_t^{p-1-i}) \right] dt.$$

Hence with the help of equations (1.6) and (1.7), we get that the expectation of the moments of X and W are $N^{-1/2}$ -close. As it turns out, it has been known for a long time (see [3]) that the moments of GUE random matrices converge towards the moments of the semicircle distribution. Hence this generalizes to Wigner matrices too. This result is nothing new and the assumption on the moments of the Wigner matrices are far from optimal, however this gives a good strategy on how to tackle the question of estimating the difference between the trace of random matrices and free operators.

In free probability, the semicircular variable is a well-known object. As we said previously in equation (1.3), it is an operator whose trace of its moments matches those of the semicircle distribution. We also stated that it is known that if P is a polynomial, $\mathbb{E} \left[\frac{1}{N} \text{Tr} (P(X^N)) \right]$ converges towards $\tau(P(x))$. But how can we estimate the difference between those two quantities? The starting idea of my PhD was precisely to create a free version of Stein's method to do so. It is known that the semicircular variable satisfies the following equations,

$$\forall p \geq 0, \tau(x^p) - \sum_{i=1}^{p-1} \tau(x^i) \tau(x^{p-1-i}) = 0. \quad (1.8)$$

Besides we know that a GUE random matrix will satisfy (1.6), so by using well-known measure concentration estimate, we get that

$$\forall p \geq 0, \mathbb{E} [\tau_N(X^p)] - \sum_{i=1}^{p-1} \mathbb{E} [\tau_N(X^{i-1})] \mathbb{E} [\tau_N(X^{p-i-1})] = \mathcal{O}(N^{-2}). \quad (1.9)$$

And once again there exists a free equivalent to the usual Brownian motion, the free Brownian motion, with whom we can build a free Ornstein-Uhlenbeck process started in X . There also exists a theory of free stochastic calculus and free Markov processes. Besides the generator of the free Ornstein-Uhlenbeck process matches with equation (1.8). However we are not simply considering moments of matrices and operators, that is the polynomial case, but as we said previously, we need to consider smooth functions. This makes the computations way more difficult since we cannot proceed by induction on the degree of the polynomial. We also need to get sharp enough asymptotic estimates since we are mostly interested by the non-renormalized trace. Finally, unlike the simplified case which we exposed in this subsection, we will be working with several independent matrices.

1.2 Summary of research results

This is the summary of the research results obtained during my PhD, each subsection corresponds to one paper. Each of the following chapter in this manuscript will also be dedicated to one paper. All of them are available on arXiv and the first four were submitted to different journals. Among them, the first, fourth and fifth papers were written in collaboration with other researchers, whereas the second and third one are single authored. By chronological order we have,

1. *On the operator norm of noncommutative polynomials in deterministic matrices and iid GUE matrices*, Benoît Collins, Alice Guionnet and Félix Parraud, [arXiv:1912.04588](https://arxiv.org/abs/1912.04588), 2019, submitted to *Cambridge Journal of Mathematics*.
2. *On the operator norm of noncommutative polynomials in deterministic matrices and iid Haar unitary matrices*, Félix Parraud, [arXiv:2005.13834](https://arxiv.org/abs/2005.13834), 2020, submitted to *Probability Theory and Related Fields*.
3. *Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices*, Félix Parraud, [arXiv:2011.04146](https://arxiv.org/abs/2011.04146), 2020, submitted to *Communications in Mathematical Physics*.
4. *Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory*, Benoît Collins and Félix Parraud, [arXiv:2012.00159](https://arxiv.org/abs/2012.00159), 2020, submitted to *Probability and Mathematical Physics*.
5. *Convergence for noncommutative rational functions evaluated in random matrices*, Benoît Collins, Tobias Mai, Akihiro Miyagawa, Félix Parraud and Sheng Yin, [arXiv:2103.05962](https://arxiv.org/abs/2103.05962), 2021.

1.2.1 On the operator norm of noncommutative polynomials in deterministic matrices and iid GUE matrices

As mentioned previously, my goal was to find an alternative proof of Haagerup and Thorbjørnsen's result of convergence of the norm of a polynomial in several independent GUE matrices. In Random Matrix Theory, we say that a sequence of family of matrices $(A_N)_{N \geq 1} = (a_1^N, \dots, a_k^N)_{N \geq 1}$ converges strongly if for any polynomials in those matrices, both the trace and the operator norm of this polynomial converge. More precisely if A is a family of operator in a C^* -algebras with a trace τ such that for any polynomial P ,

$$\frac{1}{N} \text{Tr}(P(A^N)) \xrightarrow{N \rightarrow \infty} \tau(P(A)),$$

$$\|P(A^N)\| \xrightarrow{N \rightarrow \infty} \|P(A)\|,$$

then we say that A^N converges strongly towards A . As mentioned in the first subsection, the strong convergence of a family of independent GUE matrices was proved in [6]. In [28] Male proved the strong convergence of a family of independent GUE matrices and deterministic matrices under some assumption. Finally in [31], the author also considered tensor matrices of size $N^{1/4}$ where N is the size of the GUE matrices. A tensor matrix is defined as follows: if

$E_{i,j} \in \mathbb{M}_N(\mathbb{C})$ is the matrix whose coefficients are all 0 but the coefficient (i, j) which is 1, $E_{k,l} \in \mathbb{M}_M(\mathbb{C})$ is similarly defined, then $E_{i,j} \otimes E_{k,l} \in \mathbb{M}_{NM}(\mathbb{C})$ is defined by $E_{i,j} \otimes E_{k,l} = E_{kN+i, lN+j}$. In [7], we worked with the following random matrices. First we naturally have independent GUE matrices $X^N = (X_1^N, \dots, X_d^N)$, we also consider deterministic matrices $Z^N = (Z_1^N, \dots, Z_q^N)$ and $Y^M = (Y_1^M, \dots, Y_r^M)$ with M an integer which depends on N such that $M = o(N^{1/3})$. We then proved the strong convergence of the family $(X_i^N \otimes I_M, Z_j^N \otimes I_M, I_N \otimes Y_l^M)_{i,j,l}$ assuming that the family Z^N and Y^M converges strongly themselves. Compared to [6] and [28], our result gives better quantitative bounds but also allows to consider tensors in higher dimension.

Before stating theorems, let us explain the heuristic behind the proof. To minimize notations, we focus on the case where we do not have any matrices Y^M or Z^N , but part of why this proof is interesting is that the general case is not much more difficult or different. Thanks to classical concentration of measure estimates for Gaussian random variables, one only need to show the convergence of the expectation of the norm. Then the first and most important steps was to get a quantitative estimate of the difference between

$$\mathbb{E} \left[\frac{1}{N} \operatorname{Tr} \left(f(P(X^N)) \right) \right] \quad (1.10)$$

and its deterministic limit for general sufficiently smooth function f . Indeed if $\kappa(P)$ is the limit of $\|P(X^N)\|$, then let f_ε be a non-negative function equal to 1 on the interval $[\kappa(P) + \varepsilon, \infty)$, 0 on $(-\infty, \kappa(P)]$, then for any $\alpha > 0$,

$$\begin{aligned} \mathbb{E} \left[\|P(X^N)\| \right] - \kappa(P) &\leq \alpha + \int_\alpha^\infty \mathbb{P} \left(\|P(X^N)\| \geq \kappa(P) + \varepsilon \right) d\varepsilon \\ &\leq \alpha + \int_\alpha^\infty \mathbb{P} \left(\operatorname{Tr} \left(f_\varepsilon(P(X^N)) \right) \geq 1 \right) d\varepsilon \\ &\leq \alpha + \int_\alpha^\infty \mathbb{E} \left[\operatorname{Tr} \left(f_\varepsilon(P(X^N)) \right) \right] d\varepsilon. \end{aligned} \quad (1.11)$$

Hence getting an estimate on (1.10) lets us prove that the right-hand side of (1.11) is converging towards 0, which means that $\lim_{N \rightarrow \infty} \mathbb{E} \left[\|P(X^N)\| \right] \leq \kappa(P)$. Since the other inequality is an easy consequence of Voiculescu's result, the conclusion follows. This leads us to the first important theorem of the paper, the estimation of the expectation of the trace.

Theorem 1.2.1. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices in $\mathbb{M}_N(\mathbb{C})$,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variables,
- $Z^{NM} = (Z_1^{NM}, \dots, Z_q^{NM})$ deterministic matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$,
- P a self-adjoint polynomial,
- $f \in \mathcal{C}^6(\mathbb{R})$.

Then there exists a constant l_P which only depends on P such that for any N, M ,

$$\begin{aligned} &\left| \mathbb{E} \left[\tau_{NM} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau_{NM} \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| \\ &\leq \frac{M^2}{N^2} \|f\|_{\mathcal{C}^6} \left(1 + \|Z^{NM}\| \right)^{l_P}, \end{aligned}$$

where $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$ and $\|f\|_{\mathcal{C}^6}$ is the sum of the supremum on \mathbb{R} of the first six derivatives.

To better understand this theorem, it is important to realize that the variables $x \otimes I_M$ and Z^{NM} live in the same space. Namely the tensor of $\mathbb{M}_M(\mathbb{C})$ with the free product of $\mathbb{M}_N(\mathbb{C})$ and the \mathcal{C}^* -algebra in which our semicircular variables x are. Besides τ_{NM} is the trace on this space. In particular, when restricted to $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$, as it is the case under the expectation, this trace is simply $(NM)^{-1} \text{Tr}_{\mathbb{M}_N(\mathbb{C})} \otimes \text{Tr}_{\mathbb{M}_M(\mathbb{C})}$. It is worth noting that although Haagerup and Thorbjørnsen could probably have retrieved a similar result through the Stieljes transform and the linearization trick, this would yield an estimate with a worse upper bound with respect to M and f , which then has an impact on the corollaries that one can deduce from Theorem 1.2.1. Indeed, from 1.2.1 we get an estimate on the Stieljes transform of any polynomials and another one on the bounded Lipschitz metric between the empirical measure of a random polynomial and the spectral measure towards which it converges. It also gives us a concentration estimate of the norm. We refer to Corollary 1.3 and 1.4, as well as Theorem 1.5 from [7] for a full statement of those results. The main corollary of Theorem 1.2.1 remains the following result of strong convergence.

Theorem 1.2.2. *Let the following objects be given:*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variables,
- $Y^M = (Y_1^M, \dots, Y_p^M)$ random matrices of size M a function of N ,
- $Z^N = (Z_1^N, \dots, Z_q^N)$ random matrices of size N .

Assuming that Y^M and Z^N almost surely converges strongly in distribution towards y and z families of noncommutative random variables free from x , then the following holds true:

- *If X^N and Z^N are independent, almost surely, (X^N, Z^N) converges strongly in distribution towards (x, z) .*
- *If X^N and Y^M are independent and $M = o(N^{1/3})$, almost surely, $(X^N \otimes I_M, I_N \otimes Y^M)$ converges strongly in distribution towards $(x \otimes 1, 1 \otimes y)$.*

The M^2 that appears in Theorem 1.2.1 is especially interesting since if we assume that the matrices in the family Y^M commutes it disappears. Not being able to get rid of this term has for consequences in Theorem 1.2.2 that we need to assume that $M_N = o(N^{1/3})$. Up to this day we have not found a way to improve this bound in all generality, however once again if we assume that the matrices in the family Y^M commutes, then one could even prove that strong convergence still holds if M is exponentially large relatively to N thanks to computations similar to those of [10]. The case where Y^M are independent GUE matrices and $M = N$ is of special interest since Ben Hayes proved in [32] that it is equivalent to the Peterson-Thom conjecture. While our result does not allow us to answer this question yet, we have been having interesting result by changing our approach. Namely we have been getting partial results by dropping the analytical approach, that is studying the trace of a smooth function evaluated in a polynomial $P(X^N \otimes I_M, I_N \otimes Y^M)$, and instead using the moment method, that is studying the trace of $P(X^N \otimes I_M, I_N \otimes Y^M)^k$ for $k \gg \log(N)$. However this is still ongoing work.

1.2.2 On the operator norm of noncommutative polynomials in deterministic matrices and iid Haar unitary matrices

In [8], we proved similar results but for Haar unitary matrices instead of GUE matrices. While those two models of random matrices are quite different, which means that the computations have not much in common, the heuristic of the proofs remains the same. To understand the heuristic on how to adapt this method to the unitary case, one can simply make the following remark. In the self-adjoint setting, we interpolate GUE matrices and semicircular variables by a free Ornstein Uhlenbeck process. To do so, in the proof we also make use of the Hermitian Ornstein Uhlenbeck process. One of the reason we work with those processes is that those are Markov processes whose invariant law are respectively the semicircular variable for the free Ornstein Uhlenbeck process, and the law of a GUE random matrix for the Hermitian one. In the unitary case, it turns out that the Markov process whose invariant law is the law of a Haar unitary matrix is the unitary Brownian motion, and the free Markov process whose invariant law is a free Haar unitary is unsurprisingly the free unitary Brownian motion. From there on, the overarching structure of the proof remains the same although many technical lemmas and most of the actual computations differ greatly. This yields a theorem very similar to 1.2.1.

Theorem 1.2.3. *We define*

- $u = (u_1, \dots, u_p, u_1^*, \dots, u_p^*)$ a family of p free Haar unitaries and their adjoints,
- $U^N = (U_1^N, \dots, U_p^N, (U_1^N)^*, \dots, (U_p^N)^*)$ random i.i.d. Haar unitary matrices of size N and their adjoints.
- $Z^{NM} = (Z_1^{NM}, \dots, Z_q^{NM})$ deterministic matrices and their adjoints,
- P a self-adjoint polynomial,
- $f : \mathbb{R} \rightarrow \mathbb{R}$ a smooth enough function.

Then there exists a constant l_P which only depends on P such that for any N, M ,

$$\left| \mathbb{E} \left[\tau_{MN} \left(f \left(P \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] - \tau_{NM} \left(f \left(P \left(u \otimes I_M, Z^{NM} \right) \right) \right) \right| \leq \frac{M^2}{N^2} \left(1 + \|Z^{NM}\| \right)^{l_P} \times \|f\|_{C^7}.$$

where $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$ and $\|f\|_{C^7}$ is the sum of the supremum on \mathbb{R} of the first seven derivatives.

Once again, we have similar limitations and deduce similar corollaries. The term M^2 disappears when we assume that the family Y^M does commute but we do not have a lead on how to improve this result. We also get an estimate on the Stieljes transform of any polynomials and another one on the bounded Lipschitz metric between the empirical measure of a random polynomial and the spectral measure towards which it converges. We refer to Corollary 1.1 and 1.2 from [8] for a full statement of those results. We do also get a result of strong convergence.

Theorem 1.2.4. *Let the following objects be given:*

- $U^N = (U_1^N, \dots, U_d^N)$ independent unitary Haar matrices of size N ,
- $u = (u_1, \dots, u_d)$ a system of free Haar unitaries,
- $Y^M = (Y_1^M, \dots, Y_p^M)$ random matrices of size M , a function of N ,
- $Z^N = (Z_1^N, \dots, Z_q^N)$ random matrices of size N .

Assuming that Y^M and Z^N almost surely converges strongly in distribution towards y and z families of noncommutative random variables free from u , then the following holds true:

- *If U^N and Z^N are independent, almost surely, (U^N, Z^N) converges strongly in distribution towards (u, z) .*
- *If U^N and Y^M are independent and $M = o(N^{1/3})$, almost surely, $(U^N \otimes I_M, I_N \otimes Y^M)$ converges strongly in distribution towards $(u \otimes 1, 1 \otimes y)$.*

The first point of the previous theorem is the same as Theorem 1.4 from [9], the second one however is entirely new. That being said, the most interesting part of this paper remains Theorem 1.2.3 and its corollaries, indeed the strategy in [9] was to view a Haar unitary matrix as a random function evaluated in a GUE random matrix, then to prove that almost surely this random function converges towards a deterministic function f . Then naturally if x is a semicircular, $f(x)$ is a free Haar unitary, and to conclude one simply need to apply the main result of Camille Male in [28], which is an improvement of the result of strong convergence of Haagerup and Thorbjørnsen to include deterministic matrices. The main issue of this method is that one has to approximate the random function by a random polynomial obtained by applying Weierstrass theorem almost surely. As a consequence, getting any kind of explicit estimate on the speed of convergence out of this proof seems impossible. On the contrary, our method does not have this drawback.

1.2.3 Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices

Topological expansion in Random Matrix Theory is a topic which has received a lot of attention. It creates bridges between different worlds, including topology, statistical mechanics, and quantum field theory. In mathematics, a breakthrough was made in 1986 by Harer and Zagier in [33] who used the large dimension expansion of the moments of Gaussian matrices to compute the Euler characteristic of the moduli space of curves. A good introduction to this topic is given in the survey [34] by Zvonkin. In physics, the seminal works of t'Hooft [35] and Brézin, Parisi, Itzykson and Zuber [36] related matrix models with the enumeration of maps of any genus, hence providing a purely analytical tool to solve these hard combinatorial problems. Considering matrices in interaction via a potential, the so-called matrix models, indeed allows to consider the enumeration of maps with several vertices, including a possible coloring of the edges when the matrix model contains several matrices. This relation allowed to associate matrix models to statistical models on random graphs [37, 38, 39, 40, 41], as well as in [42] and [43] for the unitary case. This was also extended to the so-called β -ensembles in [44, 45, 46, 47, 48, 49]. Among other objects, these works study correlation functions and the so-called free energy and

show that they expand as power series in the inverse of the dimension, and the coefficients of these expansions enumerate maps sorted by their genus. To compute asymptotic expansions, often referred to in the literature as topological expansions, one of the most successful methods is the loop equations method, see [50] and [51]. Depending on the model of random matrix, those are Tutte's equations, Schwinger-Dyson equations, Ward identities, Virasoro constraints, W-algebra or simply integration by part. This method was refined and used repeatedly in physics, see for example the work of Eynard and his collaborators, [52, 53, 54, 55]. At first those equations were only solved for the first few orders, however in 2004, in [53] and later [56] and [57], this method was refined to push the expansion to any orders recursively in [58].

While the paper [10], whom this subsection focuses on, does fit in the continuity of the previously mentioned papers. It is also very much in the continuity of the work of Dan Voiculescu, Haagerup and Thorbjørnsen. Indeed in [3], Dan Voiculescu proved that given a smooth function f and a self-adjoint polynomial P , then

$$\mathbb{E} \left[\frac{1}{N} \left(P(X^N) \right) \right] = \tau(P(x)) + o(1).$$

Then through the papers [6, 28, 7], it was proved that actually the last term was of order N^{-2} , the question that then arises is what happens at the next order? Among previous works, there is notably the paper [59] from Haagerup and Thorbjørnsen which does precisely that by computing a finite Taylor expansion for the previous quantity but only in the case where the family X^N consists of a single random GUE matrix. To do so, they use repeatedly the explicit formula (1.1) of the law of the probability distribution of the eigenvalues of a GUE matrix. Since there are no equivalent of such a formula for polynomials in several independent GUE matrices, it was not possible to adapt this proof. This is where free probability comes into play, the idea was to reuse the strategy from [7] but instead of using upper bound we would do exact computations. At first we kept the tensor matrices and got an explicit expression for the term of order $N^{-2}M^2$. However we realized that if we dropped the tensor matrices (i.e. $M = 1$), then we could apply the very same strategy to the term of order N^{-2} , thus getting a deterministic term of order N^{-2} and a term of order N^{-4} . And thus by induction we ended up proving a finite order Taylor expansion around N .

Theorem 1.2.5. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices in $\mathbb{M}_N(\mathbb{C})$,
- $Z^N = (Z_1^N, \dots, Z_q^N)$ deterministic matrices in $\mathbb{M}_N(\mathbb{C})$ whose norm is uniformly bounded over \mathbb{N} ,
- P a self-adjoint polynomial,
- $f \in \mathcal{C}^{4k+6}(\mathbb{R})$.

Then there exist deterministic constants $(\alpha_i^P(f, Z^N))_{i \in \mathbb{N}}$ such that,

$$\mathbb{E} \left[\frac{1}{N} \operatorname{Tr} \left(f(P(X^N, Z^N)) \right) \right] = \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(f, Z^N) + \mathcal{O}(N^{-2(k+1)}).$$

Besides, if the support of f and the spectrum of $P(x, Z^N)$ are disjoint, then for any i , $\alpha_i^P(f, Z^N) = 0$.

In order to make theorem 1.2.5 easier to read, we give an abridged version of the theorem we proved, see Theorem 1.1 from [10] for the full statement. In particular it is worth noting that one has an estimate of the term $\mathcal{O}(N^{-2(k+1)})$ in function of every parameters, such as P, f or k . Which means that one can assume that those parameters have a dependency in N . This has for consequence that if the support of f is disjoint from the spectrum of $P(x, Z^N)$, then since the coefficients of the topological expansion are all 0, $\mathbb{E} \left[\frac{1}{N} \text{Tr} \left(f(P(X^N, Z^N)) \right) \right]$ is actually exponentially small with respect to N . This also gives rise to several interesting corollaries, the first shows that almost surely the spectrum of $P(X^N, Z^N)$ is very close from the one of its free limit. If E and F are subset of \mathbb{R} , we say that $E \subset F + \varepsilon$ if every element of E is at most ε -close from an element of F .

Corollary 1.2.6. *Let X^N be independent GUE matrices of size N , Z^N a family of deterministic matrices whose norm is uniformly bounded over N , x be a free semicircular system and P a self-adjoint polynomial. Given $\alpha < 1/2$, almost surely for N large enough,*

$$\sigma \left(P(X^N, Z^N) \right) \subset \sigma \left(P(x, Z^N) \right) + N^{-\alpha},$$

where $\sigma(X)$ is the spectrum of X , and x is free from $\mathbb{M}_N(\mathbb{C})$.

One can show that a family of random matrices converges strongly if and only if the spectrum of any self-adjoint polynomial converges for the Hausdorff distance towards its free limit, see Proposition 2.2 from [7] for example. In a sense, Corollary 1.2.6 is a quantitative estimate of the speed of convergence.

We also derived another corollary from Theorem 1.2.5, it is a concentration inequality of the norm of a polynomial around the norm of its free limit. The aim was to try to prove a version of Ledoux's bound (see Lemma 3.3.2 from [13]) for general polynomial in i.i.d GUE random matrices. Ledoux's bound main interest is in being the first hint that when properly centered and rescaled by $N^{2/3}$, the tail of distribution of the largest eigenvalue converges towards the Tracy-Widom distribution. However we could not prove a result as strong for general polynomials and in the following we only rescale by $N^{1/2} \ln^{-4} N$. While we are still trying to improve this result, it is not entirely sure that it is possible to do so, since studying the norm isn't exactly the same as studying the largest eigenvalue, even for self-adjoint polynomials.

Corollary 1.2.7. *Let X^N be a tuple of independent GUE matrices of size N , A^N a family of deterministic matrices whose norm is uniformly bounded over N , x be a free semicircular system and P a polynomial. Then there exists a constant C such that for N large enough,*

$$\mathbb{P} \left(\frac{\sqrt{N}}{\ln^4 N} \left(\|P(X^N, A^N)\| - \|P(x, A^N)\| \right) \geq C(\delta + 1) \right) \leq e^{-N} + e^{-\delta^2 \ln^8 N}.$$

One of the unexpected corollary of Theorem 1.2.5, is that it can very well be applied to polynomials. However the polynomial has already been extensively studied through other methods. Indeed in 1986, Harer and Zagier proved in [33] an asymptotic expansion for the trace of moments of a GUE matrix. There is also a generalization to the trace of the mixed moments of independent GUE matrices which can be found in [60], chapter 22. The coefficients of this expansion are defined with the help of map enumeration. More precisely we say that a graph on a surface is a map if it is connected and its faces are homeomorphic to discs. It is of genus g if it can be embedded in a surface of genus g but not $g - 1$. For an edge-colored graph

on an orientated surface we say that a vertex is of type $q = X_{i_1} \dots X_{i_p}$ if it has degree p and when we look at the half-edges going out of it, starting from a distinguished one and going in the clockwise order the first half-edge is of color i_1 , the second i_2 , and so on. If $\mathcal{M}_g(X_{i_1} \dots X_{i_p})$ is the number of such maps of genus g with a single vertex, then given X_i^N independent GUE matrices

$$\mathbb{E} \left[\frac{1}{N} \text{Tr} \left(X_{i_1}^N \dots X_{i_p}^N \right) \right] = \sum_{g \in \mathbb{N}} \frac{1}{N^{2g}} \mathcal{M}_g(X_{i_1} \dots X_{i_p}). \quad (1.12)$$

And thus, thanks to my own expansion methods, we were able to get an alternative formula for $\mathcal{M}_g(X_{i_1} \dots X_{i_p})$ defined through Free Probability. We refer to remark 3.8 of [10] for more details.

Finally, it is worth noting that even though it could not be done in time for this manuscript, one can prove a similar result for unitary Haar matrices instead of GUE matrices.

1.2.4 Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory

One of the first problems that we worked on at the beginning of my PhD was the question of the additivity of the so-called Minimum Output Entropy (MOE). To summarize the problem, the main goal was to determine whether a certain map defined on quantum channels was additive or not. This problem was solved in [61] with important preliminary work from [62]. However all proofs available so far are not constructive in the sense that constructions rely on the probabilistic method. After the initial construction of [61], the probabilistic tools involved in the proof have been found to have deep relation with random matrix theory in many respects, including large deviation principle [63], Free probability [64], convex geometry [65] and Operator Algebra [66]. The last two probably give the most conceptual proofs, and in particular convex geometry gives explicit numbers. Free probability gives the best numbers for the output dimension [64] but was unable to give estimates for the input dimension so far. Thus our aim was to find explicit and as optimal as possible parameters for both the input and output dimension.

For X a matrix, its von Neumann entropy is defined by functional calculus by $H(X) = -\text{Tr}(X \ln X)$, where $0 \ln 0$ is assumed by continuity to be zero. In other words, $H(X) = -\sum_{\lambda \in \sigma(X)} \lambda \ln \lambda$ where the sum is counted with multiplicity. A quantum channel $\Phi : \mathbb{M}_d(\mathbb{C}) \rightarrow \mathbb{M}_k(\mathbb{C})$ is a completely positive trace preserving linear map. The Minimum Output Entropy of Φ is defined by

$$H_{min}(\Phi) = \min_{X \in \mathcal{D}_d} H(\Phi(X)), \quad (1.13)$$

where \mathcal{D}_d is the set of non-negative self-adjoint matrices of non-renormalized trace 1 in $\mathbb{M}_d(\mathbb{C})$. During the last decade, a crucial problem in Quantum Information Theory was to determine whether one can find two quantum channels Φ_1 and Φ_2 such that

$$H_{min}(\Phi_1 \otimes \Phi_2) < H_{min}(\Phi_1) + H_{min}(\Phi_2). \quad (1.14)$$

Indeed it is easy to prove that $H_{min}(\Phi_1 \otimes \Phi_2) \leq H_{min}(\Phi_1) + H_{min}(\Phi_2)$. But if this inequality had turned out to be an equality, then H_{min} would have been said to be additive for the tensor operation. This would have implied the additivity of the so-called Holevo capacity, which is in turn related to another very important quantity in Quantum Information Theory, the classical capacity of a quantum channel. It describes the transmission rate of classical information which

can be transmitted through repeated use of the quantum channel. If ϕ is a quantum channel and $\chi(\phi)$ is its Holevo capacity, then $C(\phi)$ the classical capacity of ϕ is defined by

$$C(\phi) = \lim_{r \rightarrow \infty} \frac{1}{r} \chi(\phi^{\otimes r}).$$

Thus, should the Holevo capacity be additive, then we would have $C(\phi) = \chi(\phi)$. This is especially strong since in general computing the Holevo capacity is much easier than computing the classical capacity. However, as said previously, in all generality it was proved that the MOE is not additive, but none of the proof provided until now gave a constructive example of quantum channels. With $\Phi_i : \mathbb{M}_{d_i}(\mathbb{C}) \rightarrow \mathbb{M}_{k_i}(\mathbb{C})$ for $i = 1, 2$, the aim of our paper was to remedy to this issue by giving explicit parameter d_i, k_i for whom one could find quantum channels Φ_1 and Φ_2 such that (1.14) is true.

The strategy that we used relied on [67], a previous work of Belinschi, Collins and Nechita in which they obtained a law of large number on the output of a random quantum channel, but since the speed of convergence was not explicit it could not be used directly. Indeed the method they used relied heavily on the strong convergence of a family of Haar unitary and deterministic matrices, proved in [9], but as we said in subsection 1.2.2, those do not provide any quantitative estimates. However thanks to my paper [8], we now have such estimates. Before giving the result that we got, let me explain the general strategy.

Let $d, k, n \in \mathbb{N}$, let U be distributed according to the Haar measure on the unitary group of $\mathbb{M}_{kn}(\mathbb{C})$, let P_n be the canonical projection from \mathbb{C}^d to \mathbb{C}^{kn} , that is the matrix with kn lines and d columns with 1 on the diagonal and 0 elsewhere. With Tr_n the trace on $\mathbb{M}_n(\mathbb{C})$, we define the following random linear map,

$$\Phi_n : X \in \mathbb{M}_d(\mathbb{C}) \mapsto id_k \otimes \text{Tr}_n(UP_n X P_n^* U^*) \in \mathbb{M}_k(\mathbb{C}). \quad (1.15)$$

This map is trace preserving, linear and completely positive and as such, it is a quantum channel. Let $t \in (0, 1)$. We fix d an integer sequence (depending on n) such that $d \sim tkn$, and define

$$K_{n,k,t} = \Phi_n(\mathcal{D}_d). \quad (1.16)$$

We also define the conjugate channel $\bar{\Phi}_n$ similarly to Φ_n but with \bar{U} instead of U . Then we have

$$H_{min}(\Phi_n) = H_{min}(\bar{\Phi}_n) = \min_{X \in K_{n,k,t}} H(X).$$

In [67], they proved that almost surely, $K_{n,k,t}$ converges towards a deterministic set $K_{k,t}$. Thus $H_{min}(\Phi_n) + H_{min}(\bar{\Phi}_n)$ converges towards $c_t = 2 \min_{X \in K_{k,t}} H(X)$. Then in [64] they proved that $H_{min}(\Phi_n \otimes \bar{\Phi}_n)$ has a deterministic upper bound which is strictly smaller than c_t . Thus almost surely, for n large enough

$$H_{min}(\Phi_n \otimes \bar{\Phi}_n) < H_{min}(\Phi_n) + H_{min}(\bar{\Phi}_n).$$

To get explicit parameters, we get measure concentration estimate on $K_{n,k,t}$ around $K_{k,t}$ for the Hausdorff distance, which in turn gives us measure concentration estimate on $H_{min}(\Phi_n)$ and let us conclude. To get those concentration estimates, we use first use a geometric description of $K_{n,k,t}$, indeed one can show that

$$K_{n,k,t} = \{X \in \mathcal{D}_k \mid \forall A \in \mathcal{D}_k, \text{Tr}_k(XA) \leq \|P_n^* U^* A \otimes I_n U P_n\|\}.$$

Then, to summarize, we use Theorem 1.2.3 with measure concentration properties of unitary Haar matrices to conclude. Eventually we get the following theorem.

Theorem 1.2.8. *If we endow $\mathbb{M}_k(\mathbb{C})$ with the norm $M \mapsto \sqrt{\text{Tr}_k(M^*M)}$, and that we assume $d \leq tkn$, then for $n \geq 3^4 \times 2^{29} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$,*

$$\mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

Which in turn gives us after some computations the following explicit parameters.

Theorem 1.2.9. *For the following values $(k, t, n) = (184, 1/10, 10^{52}), (185, 1/10, 2 \times 10^{51}), (200, 1/10, 10^{47}), (500, 1/10, 4 \times 10^{45}), (500, 1/2, 6 \times 10^{44})$ violation of additivity is achieved with probability at least $1 - \exp(-10^{20})$.*

1.2.5 Convergence for noncommutative rational functions evaluated in random matrices

This paper is a bit different from the first four since it does not rely on this idea of interpolating random matrices and free operator through free stochastic calculus. However it is still very much in the continuity of the work of Dan Voiculescu. Indeed as mentioned in the introduction, he proved that almost surely the trace of the moments of independent GUE matrices converges towards the trace of the moments of free semicircular variables. And if we define a self-adjoint polynomial P as a polynomial such that for any self-adjoint matrices X_i , $P(X_1, \dots, X_d)$ is self-adjoint, then Voiculescu's result is equivalent to the convergence in law of the empirical measure of any self-adjoint polynomial P evaluated in GUE matrices towards the spectral measure of the same polynomial evaluated in free semicirculars. In this paper we prove a similar result for more general functions. Indeed one can define the set of noncommutative polynomials by induction as the functions we can build with $\mathbb{C}, X_1, \dots, X_d$ and the operation $+$ and \times . If we add the operation $(\cdot)^{-1}$, then this is the set of noncommutative rational expression.

In their recent works [68] and [69], Mai, Speicher and Yin proved that it made sense to evaluate rational expressions in different free operators by viewing them as unbounded operators. Thus our aim is to show the convergence of the spectrum of self-adjoint rational expressions evaluated in independent random matrices towards the spectrum of the same rational expressions evaluated in free operators. Surprisingly one of the most technical and difficult part of the proof was simply to show that if we evaluate a rational expression in independent random matrices, it is indeed almost surely well-defined. Eventually though we did manage to prove that we could do so for some type of random matrices, such as Haar unitary matrices, or self-adjoint random matrix with a density with respect to the Lebesgue measure. Then we proved that as long as we could evaluate our rational expression r in both the matrices X^N and the limit operators x , the convergence in law of the empirical measure of $r(X^N)$ towards the spectral measure of $r(x)$ was simply a corollary of the convergence in distribution of the family X^N towards x .

Theorem 1.2.10. *Let $X^N = (X_1^N, \dots, X_{d_1}^N)$ be a d_1 -tuple of self-adjoint random matrices and let $U^N = (U_1^N, \dots, U_{d_2}^N)$ be a d_2 -tuple of unitary random matrices. Further, let R be a non-degenerate square matrix-valued noncommutative rational expression in $d = d_1 + d_2$ variables which is self-adjoint of type (d_1, d_2) ; see Definition 6.2.9. Suppose that the following conditions are satisfied:*

1. (X^N, U^N) converges almost surely in $*$ -distribution towards a d -tuple of noncommutative random variables (x, u) in some tracial W^* -probability space (\mathbb{M}, τ) satisfying the regularity condition $\Delta(x, u) = d$; see Sections 6.2.4 and 6.2.5.

2. For N large enough $R(X^N, U^N)$ is well-defined almost surely.

Then $R(x, u)$ is well-defined, and the empirical measure of $R(X^N, U^N)$ converges almost surely in law towards the analytic distribution of $R(x, u)$.

The assumption 2 is satisfied for random matrix models (X^N, U^N) whose law on $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ is absolutely continuous with respect to the product measure of the Lebesgue measure on $M_N(\mathbb{C})_{\text{sa}}$ and the Haar measure on $U_N(\mathbb{C})$.

In particular, the assumptions 1 and 2 are satisfied for random matrix models satisfying the following conditions:

- (X^N, U^N) are almost surely asymptotically free.
- The law of each X_j^N has a density with respect to the Lebesgue measure on $M_N(\mathbb{C})_{\text{sa}}$ and its eigenvalue distribution almost surely converges weakly to some compactly supported probability measure on \mathbb{R} that is non-atomic.
- U^N are i.i.d. Haar distributed.

Let us provide an outline of the proof. It is known that given a self-adjoint rational expression r , we can find a self-adjoint linearization $\rho = (Q, w)$. That is a matrix $Q \in \mathbb{M}_k(\langle X_1, \dots, X_d \rangle)$ and a vector $w \in \mathbb{C}^k$ such that $r(X) = w^* Q(X)^{-1} w$. Thanks to the Portman-teau theorem, we only need to show the convergence of the cumulative distribution function to conclude. Then we prove that the cumulative distribution function of the spectral measure of an operator (not necessarily bounded) is in a way continuous with respect to the rank. This implies that we can ignore the singularity in 0 of $Q(x, u, u^*)^{-1}$ as long as the spectral measure of $Q(x, u, u^*)$ has no atom in 0. Thus it turns out that the cumulative distribution function of $w^* Q(X^N, U^N, U^{N*})^{-1} w$ is close from the one of $w^* f_\varepsilon(Q(X^N, U^N, U^{N*})) w$ where f_ε is a continuous function which is equal to $t \mapsto t^{-1}$ outside of a neighborhood of 0 of size ε . Then we can use the convergence in $*$ -distribution of the family (X^N, U^N) to show that the cumulative distribution function of $w^* f_\varepsilon(Q(X^N, U^N, U^{N*})) w$ converges towards the correct limit when we let N go to infinity and ε go to 0.

On the operator norm of noncommutative polynomials in deterministic matrices and iid GUE matrices

Let $X^N = (X_1^N, \dots, X_d^N)$ be a d -tuple of $N \times N$ independent GUE random matrices and Z^{NM} be any family of deterministic matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$. Let P be a self-adjoint noncommutative polynomial. A seminal work of Voiculescu shows that the empirical measure of the eigenvalues of $P(X^N)$ converges towards a deterministic measure defined thanks to free probability theory. Let now f be a smooth function, the main technical result of this paper is a precise bound of the difference between the expectation of

$$\frac{1}{MN} \operatorname{Tr}_{\mathbb{M}_N(\mathbb{C})} \otimes \operatorname{Tr}_{\mathbb{M}_M(\mathbb{C})} \left(f(P(X^N \otimes I_M, Z^{NM})) \right),$$

and its limit when N goes to infinity. If f is six times differentiable, we show that it is bounded by $M^2 \|f\|_{\mathcal{C}^6} N^{-2}$. As a corollary we obtain a new proof and slightly improve a result of Haagerup and Thorbjørnsen, later developed by Male, which gives sufficient conditions for the operator norm of a polynomial evaluated in (X^N, Z^{NM}, Z^{NM*}) to converge almost surely towards its free limit.

This chapter is adapted from [7], which is a joint work with Benoît Collins and Alice Guionnet.

2.1 Introduction

Given several deterministic matrices whose spectra are known, the spectra of a noncommutative polynomial evaluated in these matrices is not well defined since it depends as well on the eigenvectors of these matrices. If one takes these vectors at random, it is possible to get some surprisingly good results, in particular when the dimension of these matrices goes to infinity. Indeed, the limit can then be computed thanks to free probability. This theory was introduced by Voiculescu in the early nineties as a noncommutative probability theory equipped with a notion of freeness analogous to independence in classical probability theory. Voiculescu showed that this theory was closely related with Random Matrix Theory in a seminal paper [3]. He considered independent matrices taken from the Gaussian Unitary Ensemble (GUE), which is a random matrix is an $N \times N$ self-adjoint random matrix whose distribution is proportional to the measure $\exp(-N/2 \operatorname{Tr}_N(A^2)) dA$, where dA denotes the Lebesgue measure on the set of $N \times N$ Hermitian matrices. We refer to Definition 2.2.8 for a more precise statement. Voiculescu proved that given X_1^N, \dots, X_d^N independent GUE matrices, the renormalized trace of a polynomial P evaluated in these matrices converges towards a deterministic limit $\alpha(P)$. Specifically, the following holds true almost surely:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{Tr}_N \left(P(X_1^N, \dots, X_d^N) \right) = \alpha(P). \quad (2.1)$$

Voiculescu computed the limit $\alpha(P)$ with the help of free probability. If A_N is a self-adjoint matrix of size N , then one can define the empirical measure of its (real) eigenvalues by

$$\mu_{A_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i},$$

where δ_λ is the Dirac mass in λ and $\lambda_1, \dots, \lambda_N$ are the eigenvalue of A_N . In particular, if P is a self-adjoint polynomial, that is such that for any self adjoint matrices A_1, \dots, A_d , $P(A_1, \dots, A_d)$ is a self-adjoint matrix, then one can define the random measure $\mu_{P(X_1^N, \dots, X_d^N)}$. In this case, Voiculescu's result (2.1) implies that there exists a measure μ_P with compact support such that almost surely $\mu_{P(X_1^N, \dots, X_d^N)}$ converges weakly towards μ_P : it is given by $\mu_P(x^k) = \alpha(P^k)$ for all integer numbers k .

However, the convergence of the empirical measure of the eigenvalues of a matrix does not say anything about the local properties of its spectrum, in particular about the convergence of the norm of this matrix, or the local fluctuations of its spectrum. When dealing with a single matrix, incredibly precise results are known. For exemple it is well-known that the largest eigenvalue of a GUE random matrix converges almost surely towards 2. More precisely, if X_N is a GUE random matrix of size N , then almost surely

$$\lim_{N \rightarrow \infty} \|X_N\| = 2.$$

The proof, for the more general case of a Wigner matrix with entries with finite moments, was given in [15]. This result was later obtained under the optimal assumption that their fourth moment is finite [16]. Concerning the GUE, much more precise results were obtained by Tracy and Widom in the early nineties in [17]. The main result of their paper is the existence of a continuous decreasing function F_2 from \mathbb{R} to $[0, 1]$ such that if $\lambda_1(X^N)$ denotes the largest eigenvalue of X^N ,

$$\lim_{N \rightarrow \infty} P\left(N^{2/3}(\lambda_1(X^N) - 2) \geq s\right) = F_2(s).$$

This was recently generalized to Wigner matrices [18, 19, 20, 21] up to optimal hypotheses. One can as well study the localization of the eigenvalues in the bulk as well as their fluctuations [22, 19].

On the other hand, there are much less results available when one deals with a polynomial in several random matrices. In fact, up to today, the only local fluctuations results concern perturbative polynomials [23] or local laws [24] under some assumptions which are shown to hold for homogeneous polynomials of degree two. However, a beautiful breakthrough was made in 2005 by Haagerup and Thorbjørnsen [6]: they proved the almost sure convergence of the norm of a polynomial evaluated in independent GUE matrices. For P a self-adjoint polynomial, they proved that almost surely, for N large enough,

$$\sigma\left(P(X_1^N, \dots, X_d^N)\right) \subset \text{Supp } \mu_P + (-\varepsilon, \varepsilon), \quad (2.2)$$

where $\sigma(H)$ is the spectrum of H and $\text{Supp } \mu_P$ the support of the measure μ_P . This is equivalent to saying that for any polynomial P , $\|P(X_1^N, \dots, X_d^N)\|$ converges almost surely towards $\sup\{|x| \mid x \in \text{Supp } \mu_P\}$ (see proposition 2.2.2). The result (2.2) was a major progress in free probability. It was refined in multiple ways. In [25], Schultz used the method of [6] to prove the same result with Gaussian orthogonal or symplectic matrices instead of Gaussian unitary matrices. In [26], Capitaine and Donati-Martin proved it for Wigner matrices under some technical hypothesis on the law of the entries. This result itself was then extended by Anderson in [27] to remove most of the technical assumptions. In [28], Male made a conceptual improvement to the result of Haagerup and Thorbjørnsen, by allowing to work both with GUE and deterministic matrices. Finally, Belinschi and Capitaine proved in [29] that one could even work with Wigner and deterministic matrices, while keeping the same assumptions on the Wigner matrices as Anderson. It is also worth noting that Collins and Male proved in [9] the same result with unitary Haar matrices instead of GUE matrices by using Male's former paper.

With the exception of [9], all of these results are essentially based on the method introduced by Haagerup and Thorbjørnsen. Their first tool is called the linearization trick: it allows to relate the spectrum of a polynomial of degree d with coefficients in \mathbb{C} by a polynomial of degree 1 with coefficients in $\mathbb{M}_{k(d)}(\mathbb{C})$. The second idea to understand the spectrum of the spectral measure of this larger matrix is to study its Stieltjes transform close to the real axis by using the Dyson-Schwinger equations. An issue of this method is that it does not give easily good quantitative estimates. One aim of this paper is to remedy to this problem. We develop a new method that allows us to give a new proof of the main theorem of Male in [28], and thus a new proof of the result of Haagerup and Thorbjørnsen. Our approach requires neither the linearization trick, nor the study of the Stieltjes transform and attacks the problem directly. In this sense the proof is more direct and less algebraic. We will apply it to a generalization of GUE matrices by tackling the case of GUE random matrices tensorized with deterministic matrices.

A usual strategy to study outliers, that are the eigenvalues going away from the spectrum, is to study the *non-renormalized* trace of smooth non-polynomial functions evaluated in independent GUE matrices i.e. if P is self-adjoint:

$$\text{Tr}_N\left(f(P(X_1^N, \dots, X_d^N))\right).$$

This strategy was also used by Haagerup, Thorbjørnsen and Male. Indeed it is easy to see that if f is a function which takes value 0 on $(-\infty, C - \varepsilon]$, 1 on $[C, \infty)$ and in $[0, 1]$ elsewhere, then

$$\mathbb{P}\left(\lambda_1(P(X_1^N, \dots, X_d^N)) \geq C\right) \leq \mathbb{P}\left(\mathrm{Tr}_N\left(f(P(X_1^N, \dots, X_d^N))\right) \geq 1\right)$$

Hence, if we can prove that $\mathrm{Tr}_N\left(f(P(X_1^N, \dots, X_d^N))\right)$ converges towards 0 in probability, this would already yield expected results. The case where f is a polynomial function has already been studied a long time ago, starting with the pioneering works [70, 71], and later formalized by the concept of second order freeness [72]. However here we have to deal with a function f which is at best C^∞ . This makes things considerably more difficult and forces us to adopt a completely different approach. The main result is the following Theorem. For the notations, we refer to Section 2.2 – for now, let us specify that $\frac{1}{N} \mathrm{Tr}_N$ denotes the usual renormalized trace on $N \times N$ matrices whereas τ denotes its free limit.

Theorem 2.1.1. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices in $\mathbb{M}_N(\mathbb{C})$,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variables,
- $Z^{NM} = (Z_1^{NM}, \dots, Z_q^{NM})$ deterministic matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$,
- $P \in \mathbb{C}\langle X_1, \dots, X_{d+2q} \rangle_{sa}$ a self-adjoint polynomial,
- $f \in \mathcal{C}^6(\mathbb{R})$.

Then there exists a polynomial $L_P \in \mathbb{R}^+[X]$ which only depends on P such that with $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$, for any N, M ,

$$\left| \mathbb{E} \left[\frac{1}{MN} \mathrm{Tr}_{MN} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau_N \otimes \tau_M \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| \leq \frac{M^2}{N^2} \|f\|_{\mathcal{C}^6} L_P \left(\|Z^{NM}\| \right),$$

where $\|f\|_{\mathcal{C}^6}$ is the sum of the supremum on \mathbb{R} of the first six derivatives. Besides if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, then we have the same inequality without the M^2 .

This theorem is a consequence of the slightly sharper, but less explicit, Theorem 2.3.1. It is essentially the same statement, but instead of having the norm C^6 of f , we have the fourth moment of the Fourier transform of f . The above Theorem calls for a few remarks.

- We assumed that the matrices Z^{NM} were deterministic, but thanks to Fubini's Theorem we can assume that they are random matrices as long as they are independent from X^N . In this situation though, $L_P \left(\|Z^{NM}\| \right)$ in the right side of the inequality is a random variable (and thus we need some additional assumptions if we want its expectation to be finite for instance).

- In Theorems 2.1.1 and 2.3.1 we have $X^N \otimes I_M$ and $x \otimes I_M$, however it is very easy to replace them by $X^N \otimes Y^M$ and $x \otimes Y^M$ for some matrices $Y_i^M \in M_M(\mathbb{C})$. Indeed we just need to apply Theorem 2.1.1 or 2.3.1 with $Z^{NM} = I_N \otimes Y^M$. Besides, in this situation, $L_P(\|Z^{NM}\|) = L_P(\|Y^M\|)$ does not depend on N . What this means is that if we have a matrix whose coefficients are polynomial in X^N , and that we replace X^N by x , we only change the spectra of this matrix by M^2N^{-2} in average.
- Unfortunately we cannot get rid of the M^2 in all generality. The specific case where we can is when $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$, where the Y_i^M commute: this indicates that the M^2 term is really a noncommutative feature.

A detailed overview of the proof is given in Subsection 2.3.1. The main idea of the proof is to use a free version of Stein's method by interpolating GUE matrices with a free semicircular system with the help of a free Ornstein-Uhlenbeck process. For a reference, see [73]. When using this process, the Schwinger-Dyson equations, which can be seen as an integration by part formula, appear in the computation. We refer to Proposition 2.2.10 for more information which will play a major role in this paper. Theorem 2.1.1 is the crux of the paper and allows us to deduce many corollaries. Firstly we rederive a new proof of the following theorem. The first statement is basically Theorem 1.6 from [28]. The second one is an improvement of Theorem 7.8 from [31] on the size of the tensor from $N^{1/4}$ to $N^{1/3}$. This theorem is about strong convergence of random matrices, that is the convergence of the norm of polynomials in these matrices, see definition 2.2.1.

Theorem 2.1.2. *Let the following objects be given:*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variable,
- $Y^M = (Y_1^M, \dots, Y_p^M)$ random matrices of size M , which almost surely, as M goes to infinity, converge strongly in distribution towards a p -tuple y of noncommutative random variables in a C^* -probability space \mathcal{B} with a faithful trace $\tau_{\mathcal{B}}$,
- $Z^N = (Z_1^N, \dots, Z_q^N)$ random matrices of size N , which almost surely, as N goes to infinity, converges strongly in distribution towards a q -tuple z of noncommutative random variables in a C^* -probability space with a faithful trace.

Then, the following holds true:

- If X^N and Z^N are independent, almost surely, (X^N, Z^N) converges strongly in distribution towards $\mathcal{F} = (x, z)$, where \mathcal{F} belongs to a C^* -probability space $(\mathcal{A}, *, \tau_{\mathcal{A}}, \|\cdot\|)$ in which x and z are free.
- If $(M_N)_{N \geq 1}$ is a sequence of integers such that $M_N = o(N^{1/3})$, X^N and Y^{M_N} are independent, then almost surely $(X^N \otimes I_{M_N}, I_N \otimes Y^{M_N})$ converges strongly in distribution towards $\mathcal{F} = (x \otimes 1, 1 \otimes y)$ when N goes to infinity. The family \mathcal{F} thus belongs to $\mathcal{A} \otimes_{\min} \mathcal{B}$ (see definition 2.4.1). Besides if the matrices Y^{M_N} commute, then we can weaken the assumption on M_N by only assuming that $M_N = o(N)$.

As we mentioned earlier, understanding the Stieljes transform of a matrix gives a lot of information about its spectrum. This was actually a very important point in the proof of Haagerup and Thorbjørnsen's Theorem. Our proof does not use this tool, however our final result, Theorem 2.3.1, allows us to deduce the following estimate with sharper constant than what has previously been done. Being given a self-adjoint $NM \times NM$ matrix, we denote by G_A its Stieltjes transform:

$$G_A(z) = \frac{1}{NM} \text{Tr}_{NM} \left(\frac{1}{z - A} \right).$$

This definition extends to the tensor product of free semi-circular variables by replacing $(NM)^{-1} \text{Tr}_{NM}$ by $\tau_N \otimes \tau_M$.

Corollary 2.1.3. *Given*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variable,
- $Y^M = (Y_1^M, \dots, Y_p^M, Y_1^{M*}, \dots, Y_p^{M*})$ deterministic matrices of size M and their adjoints,
- $P \in \mathbb{C}\langle X_1, \dots, X_d, Y_1, \dots, Y_{2p} \rangle_{sa}$ a self-adjoint polynomial,

there exists a polynomial $L_P \in \mathbb{R}^+[X]$ such that for every Y^M , $z \in \mathbb{C} \setminus \mathbb{R}$, $M, N \in \mathbb{N}$,

$$\left| \mathbb{E} \left[G_{P(X^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(x \otimes I_M, I_N \otimes Y^M)}(z) \right| \leq L_P \left(\|Y^M\| \right) \frac{M^2}{N^2} \left(\frac{1}{|\Im(z)|^5} + \frac{1}{|\Im(z)|^2} \right),$$

where $\|Y^M\| = \sup_{1 \leq i \leq p} \|Y_i^M\|$.

One of the limitation of Theorem 2.1.1 is that we need to pick f regular enough. Actually by approximating f , we can afford to take f less regular at the cost of a slower speed of convergence. In other words, we trade some degree of regularity on f for a smaller exponent in N . The best that we can achieve is to take f Lipschitz. Thus it makes sense to introduce the Lipschitz-bounded metric. This metric is compatible with the topology of the convergence in law of measure. Let \mathcal{F}_{LU} be the set of Lipschitz functions from \mathbb{R} to \mathbb{R} , uniformly bounded by 1 and with Lipschitz constant at most 1, then

$$d_{LU}(\mu, \nu) = \sup_{f \in \mathcal{F}_{LU}} \left| \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \right|.$$

For more information about this metric we refer to Annex C.2 of [13]. In this paper, we get the following result:

Corollary 2.1.4. *Under the same notations as in Corollary 2.1.3, there exists a polynomial $L_P \in \mathbb{R}^+[X]$ such that for every matrices Y^M and $M, N \in \mathbb{N}$,*

$$d_{LU} \left(\mathbb{E}[\mu_{P(X^N \otimes I_M, I_N \otimes Y^M)}], \mu_{P(x \otimes I_M, I_N \otimes Y^M)} \right) \leq L_P \left(\|Y^M\| \right) \frac{M^2}{N^{1/3}}.$$

One of the advantage of Theorem 2.1.1 over the original proof of Haagerup and Thorbjørnsen is that if we take f which depends on N , we get sharper estimates in N . For exemple if we assume that g is a C^∞ function with bounded support, as we will see later in this paper we like to work with $f : x \mapsto g(N^\alpha x)$ for some constant α . Then its n -th derivative will be of order $N^{n\alpha}$. In the original work of Haagerup, Thorbjørnsen (see [6], Theorem 6.2) the eighth derivative appears for the easiest case where our polynomial P is of degree 1, and the order is even higher in the general case. But in Theorem 2.1.1 the sixth derivative appears in the general case. Actually if we look at the sharper Theorem 2.3.1, the fourth moment of the Fourier transform appears, which is roughly equivalent to the fourth derivative for our computations. This allows us to compute an estimate of the difference between $\mathbb{E} \left[\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right]$ and its limit. To do that, we use Proposition 2.4.8 from [74, Theorem 1.1] which implies that if we denote by $\mu_{P(x \otimes I_M, 1 \otimes Y^M)}$ the spectral measure of $P(x \otimes I_M, 1 \otimes Y^M)$, then there exists $\beta \in \mathbb{R}^+$ such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^{-\beta} \mu_{P(x \otimes I_M, 1 \otimes Y^M)} \left(\left(\left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| - \varepsilon, \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right) \right) > 0. \quad (2.3)$$

With the help of standard measure concentration estimates, we then get the following Theorem:

Theorem 2.1.5. *We consider*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variable,
- $Y^M = (Y_1^M, \dots, Y_p^M)$ deterministic matrices of size M a fixed integer and their adjoints.

For any polynomial $P \in \mathbb{C}\langle X_1, \dots, X_d, Y_1, \dots, Y_p \rangle$, there exists constants K and C such that for any $\delta > 0$ and $N \in \mathbb{N}$,

$$\begin{aligned} \mathbb{P} \left(N^{1/4} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right) \geq \delta + C \right) \\ \leq e^{-K\delta^2\sqrt{N}} + de^{-N}, \end{aligned} \quad (2.4)$$

$$\begin{aligned} \mathbb{P} \left(N^{1/(3+\beta)} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right) \leq -\delta - C \right) \\ \leq e^{-K\delta^2 N^{\frac{1+\beta}{3+\beta}}} + de^{-N}. \end{aligned} \quad (2.5)$$

This theorem is interesting because of its similarity with Tracy and Widom's result about the tail of the law of the largest eingenvale of a GUE matrix. We have smaller exponent in N , and thus we can only show the convergence towards 0 with exponential speed, however we are not restricted to a single GUE matrix, we can chose any polynomial evaluated in GUE matrices. Besides by applying Borel-Cantelli's Lemma, we immediately get:

Theorem 2.1.6. *We consider*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variable,

- $Y^M = (Y_1^M, \dots, Y_p^M)$ deterministic matrices of size M a fixed integer and their adjoints.

Then almost surely, for any polynomial $P \in \mathbb{C}\langle X_1, \dots, X_d, Y_1, \dots, Y_p \rangle$, there exists a constant $c(P) > 0$ such that for any $c(P) > \alpha > 0$,

$$\lim_{N \rightarrow \infty} N^\alpha \left| \left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right| = 0.$$

Moreover, if β satisfies (2.3), then almost surely for any $\alpha < (3 + \beta)^{-1}$ and $\varepsilon < 1/4$, for N large enough,

$$-N^{-\alpha} \leq \left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \leq N^{-\varepsilon}.$$

In order to conclude this introduction, we would like to say that while it is not always easy to compute the constant β in all generality, it is possible for some polynomials. In particular, if our polynomial is evaluated in a single GUE matrix, then the computation is heavily simplified by the fact that we know the distribution of a single semicircular variable. Finally, the constant $(3 + \beta)^{-1}$ is clearly a worst case scenario and can be easily improved if β is explicit.

This paper is organised as follows. In Section 2.2, we recall the definitions and properties of free probability, noncommutative calculus and Random Matrix Theory needed for this paper. Section 2.3 contains the proof of Theorem 2.1.1. And finally in Section 2.4 we give the proof of the remaining Theorem and Corollaries.

2.2 Framework and standard properties

2.2.1 Usual definitions in free probability

In order to be self-contained, we begin by reminding the following definitions from free probability.

Definition 2.2.1. • A \mathcal{C}^* -probability space $(\mathcal{A}, *, \tau, \|\cdot\|)$ is a unital \mathcal{C}^* -algebra $(\mathcal{A}, *, \|\cdot\|)$ endowed with a state τ , i.e. a linear map $\tau : \mathcal{A} \rightarrow \mathbb{C}$ satisfying $\tau(1_{\mathcal{A}}) = 1$ and $\tau(a^*a) \geq 0$ for all $a \in \mathcal{A}$. In this paper we always assume that τ is a trace, i.e. that it satisfies $\tau(ab) = \tau(ba)$ for any $a, b \in \mathcal{A}$. An element of \mathcal{A} is called a (noncommutative) random variable. We will always work with a faithful trace, namely, for $a \in \mathcal{A}$, $\tau(a^*a) = 0$ if and only if $a = 0$. In this case the norm is determined by τ thanks to the formula:

$$\|a\| = \lim_{k \rightarrow \infty} \left(\tau((a^*a)^{2k}) \right)^{1/2k}.$$

- Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be $*$ -subalgebras of \mathcal{A} , having the same unit as \mathcal{A} . They are said to be free if for all k , for all $a_i \in \mathcal{A}_{j_i}$ such that $j_1 \neq j_2, j_2 \neq j_3, \dots, j_{k-1} \neq j_k$:

$$\tau\left((a_1 - \tau(a_1))(a_2 - \tau(a_2)) \dots (a_k - \tau(a_k))\right) = 0.$$

Families of noncommutative random variables are said to be free if the $*$ -subalgebras they generate are free.

- Let $A = (a_1, \dots, a_k)$ be a k -tuple of random variables. The joint distribution of the family A is the linear form $\mu_A : P \mapsto \tau[P(A, A^*)]$ on the set of polynomials in $2k$ noncommutative indeterminates. By convergence in distribution, for a sequence of families of variables $(A_N)_{N \geq 1} = (a_1^N, \dots, a_k^N)_{N \geq 1}$ in \mathcal{C}^* -algebras $(\mathcal{A}_N, *, \tau_N, \|\cdot\|)$, we mean the pointwise convergence of the map

$$\mu_{A_N} : P \mapsto \tau_N[P(A_N, A_N^*)],$$

and by strong convergence in distribution, we mean convergence in distribution, and pointwise convergence of the map

$$P \mapsto \|P(A_N, A_N^*)\|.$$

- A family of noncommutative random variables $x = (x_1, \dots, x_p)$ is called a free semicircular system when the noncommutative random variables are free, selfadjoint ($x_i = x_i^*$, $i = 1 \dots p$), and for all k in \mathbb{N} and $i = 1, \dots, p$, one has

$$\tau(x_i^k) = \int t^k d\sigma(t),$$

with $d\sigma(t) = \frac{1}{2\pi} \sqrt{4-t^2} \mathbf{1}_{|t| \leq 2} dt$ the semicircle distribution.

The strong convergence of noncommutative random variables is actually equivalent to the convergence of the spectrum of their polynomials for the Hausdorff distance. More precisely we have the following proposition whose proof can be found in [9, Proposition 2.1] :

Proposition 2.2.2. *Let $\mathbf{x}_N = (x_1^N, \dots, x_p^N)$ and $\mathbf{x} = (x_1, \dots, x_p)$ be p -tuples of variables in \mathcal{C}^* -probability spaces, $(\mathcal{A}_N, *, \tau_N, \|\cdot\|)$ and $(\mathcal{A}, *, \tau, \|\cdot\|)$, with faithful states. Then, the following assertions are equivalent.*

- \mathbf{x}_N converges strongly in distribution to \mathbf{x} .
- For any self-adjoint variable $h_N = P(\mathbf{x}_N)$, where P is a fixed polynomial, μ_{h_N} converges in weak-* topology to μ_h where $h = P(\mathbf{x})$. Weak-* topology means relatively to continuous functions on \mathbb{C} . Moreover, the spectrum of h_N converges in Hausdorff distance to the spectrum of h , that is, for any $\varepsilon > 0$, there exists N_0 such that for any $N \geq N_0$,

$$\sigma(h_N) \subset \sigma(h) + (-\varepsilon, \varepsilon). \quad (2.6)$$

In particular, the strong convergence in distribution of a single self-adjoint variable is equivalent to its convergence in distribution together with the Hausdorff convergence of its spectrum.

It is important to note that thanks to [60, Theorem 7.9], that we recall below, one can consider free version of any random variable.

Theorem 2.2.3. *Let $(\mathcal{A}_i, \phi_i)_{i \in I}$ be a family of \mathcal{C}^* -probability spaces such that the functionals $\phi_i : \mathcal{A}_i \rightarrow \mathbb{C}$, $i \in I$, are faithful traces. Then there exist a \mathcal{C}^* -probability space (\mathcal{A}, ϕ) with ϕ a faithful trace, and a family of norm- preserving unital $*$ -homomorphism $W_i : \mathcal{A}_i \rightarrow \mathcal{A}$, $i \in I$, such that:*

- $\phi \circ W_i = \phi_i$, $\forall i \in I$.

- The unital C^* -subalgebras form a free family in (\mathcal{A}, ϕ) .

Let us finally fix a few notations concerning the spaces and traces that we use in this paper.

Definition 2.2.4. • (\mathcal{A}_N, τ_N) is the free sum of $\mathbb{M}_N(\mathbb{C})$ with a system of d free semicircular variable, this is the C^* -probability space built in Theorem 2.2.3. Note that when restricted to $\mathbb{M}_N(\mathbb{C})$, τ_N is just the regular renormalized trace on matrices. The restriction of τ_N to the C^* -algebra generated by the free semicircular system x is denoted as τ .

- Tr_N is the non-renormalized trace on $\mathbb{M}_N(\mathbb{C})$.
- $\mathbb{M}_N(\mathbb{C})_{sa}$ is the set of self adjoint matrix of $\mathbb{M}_N(\mathbb{C})$. We denote $E_{r,s}$ the matrix with coefficients equal to 0 except in (r, s) where it is equal to one.
- We regularly identify $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$ with $\mathbb{M}_{kN}(\mathbb{C})$ through the isomorphism $E_{i,j} \otimes E_{r,s} \mapsto E_{i+rN, j+sN}$, similarly we identify $\text{Tr}_N \otimes \text{Tr}_k$ with Tr_{kN} .
- If $A^N = (A_1^N, \dots, A_d^N)$ and $B^M = (B_1^M, \dots, B_d^M)$ are two families of matrices, then we denote $A^N \otimes B^M = (A_1^N \otimes B_1^M, \dots, A_d^N \otimes B_d^M)$. We typically use the notation $X^N \otimes I_M$ for the family $(X_1^N \otimes I_M, \dots, X_d^N \otimes I_M)$.

2.2.2 Non-commutative polynomials and derivatives

We set $\mathcal{A}_{d,q} = \mathbb{C}\langle X_1, \dots, X_d, Y_1, \dots, Y_q, Y_1^*, \dots, Y_q^* \rangle$ the set of noncommutative polynomial in $d + 2q$ indeterminates. We endow this vector space with the norm

$$\|P\|_A = \sum_{M \text{ monomial}} |c_M(P)| A^{\deg M}, \quad (2.7)$$

where $c_M(P)$ is the coefficient of P for the monomial M and $\deg M$ the total degree of M (that is the sum of its degree in each letter $X_1, \dots, X_d, Y_1, \dots, Y_q, Y_1^*, \dots, Y_q^*$). Let us define several maps which we use frequently in the sequel First, for $A, B, C \in \mathcal{A}_{d,q}$, let

$$A \otimes B \# C = ACB,$$

$$A \otimes B \tilde{\#} C = BCA,$$

$$m(A \otimes B) = BA.$$

Definition 2.2.5. If $1 \leq i \leq d$, one defines the noncommutative derivative $\partial_i : \mathcal{A}_{d,q} \longrightarrow \mathcal{A}_{d,q} \otimes \mathcal{A}_{d,q}$ by its value on a monomial $M \in \mathcal{A}_{d,q}$ given by

$$\partial_i M = \sum_{M=AX_iB} A \otimes B,$$

and then extend it by linearity to all polynomials. Similarly one defines the cyclic derivative $D_i : \mathcal{A}_{d,q} \longrightarrow \mathcal{A}_{d,q}$ for $P \in \mathcal{A}_{d,q}$ by

$$D_i P = m \circ \partial_i P .$$

The map ∂_i is called the noncommutative derivative. It is related to Schwinger-Dyson equation on semicircular variable thanks to the following property 2.2.6. One can find a proof of the first part in [13], Lemma 5.4.7. As for the second part it is a direct consequence of the first one which can easily be verified by taking P monomial and then concluding by linearity.

Proposition 2.2.6. *Let $x = (x_1, \dots, x_p)$ be a free semicircular system, $y = (y_1, \dots, y_q)$ be noncommutative random variables free from x , if the family (x, y) belongs to the \mathcal{C}^* -probability space $(\mathcal{A}, *, \tau, \|\cdot\|)$, then for any $P \in \mathcal{A}_{d,q}$,*

$$\tau(P(x, y, y^*) x_i) = \tau \otimes \tau(\partial_i P(x, y, y^*)) .$$

Moreover, one can deduce that if Z^{NM} are matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$ that we view as a subspace of $\mathcal{A}_N \otimes \mathbb{M}_M(\mathbb{C})$, then for any $P \in \mathcal{A}_{d,q}$,

$$\begin{aligned} & \tau_N \otimes \tau_M \left(P(x \otimes I_M, Z^{NM}, Z^{NM*}) x_i \otimes I_M \right) \\ &= \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(\partial_i P(x \otimes I_M, Z^{NM}, Z^{NM*}) \right) \right) . \end{aligned}$$

We define an involution $*$ on $\mathcal{A}_{d,q}$ such that

$$(X_i)^* = X_i, \quad (Y_i)^* = Y_i^*, \quad (Y_i^*)^* = Y_i$$

and then we extend it to $\mathcal{A}_{d,q}$ by the formula $(\alpha PQ)^* = \bar{\alpha} Q^* P^*$. $P \in \mathcal{A}_{d,q}$ is said to be self-adjoint if $P^* = P$. Self-adjoint polynomials have the property that if $x_1, \dots, x_d, z_1, \dots, z_q$ are elements of a \mathcal{C}^* -algebra such as x_1, \dots, x_d are self-adjoint, then so is $P(x_1, \dots, x_d, z_1, \dots, z_q, z_1^*, \dots, z_q^*)$. Now that we have defined the notion of self-adjoint polynomial we remark for later use that

Proposition 2.2.7. *Let the following objects be given,*

- $x = (x_1, \dots, x_p)$ a free semicircular system ,
- $X^N = (X_1^N, \dots, X_d^N)$ self-ajoint matrices of size N ,
- $X_t^N = e^{-t/2} X^N + (1 - e^{-t})^{1/2} x$ elements of \mathcal{A}_N ,
- Z^{NM} matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$,
- $f \in \mathcal{C}^0(\mathbb{R})$,
- P a self-adjoint polynomial.

Then the following map is measurable:

$$(X^N, Z^{NM}) \mapsto \tau_N \otimes \tau_M \left(f \left(P(X_t^N \otimes I_M, Z^{NM}, Z^{NM*}) \right) \right) .$$

Proof. This is obvious if f is a polynomial and the general case is obtained by approximation. \square

Actually we could easily prove that this map is continuous, however we do not need it. The only reason we need this property is to justify that if X^N is a d -tuple of independent GUE matrices, then the random variable $\tau_N \otimes \tau_M \left(f \left(P(X_t^N \otimes I_M, Z^{NM}, Z^{NM*}) \right) \right)$ is well-defined and measurable.

2.2.3 GUE random matrices

We conclude this section by reminding the definition of Gaussian random matrices and stating a few useful properties about them.

Definition 2.2.8. *A GUE random matrix X^N of size N is a self adjoint matrix whose coefficients are random variables with the following laws:*

- For $1 \leq i \leq N$, the random variables $\sqrt{N}X_{i,i}^N$ are independent centered Gaussian random variables of variance 1.
- For $1 \leq i < j \leq N$, the random variables $\sqrt{2N} \Re X_{i,j}^N$ and $\sqrt{2N} \Im X_{i,j}^N$ are independent centered Gaussian random variables of variance 1, independent of $(X_{i,i}^N)_i$.

We now present two of the most useful tools when it comes to computation with Gaussian variable, the Poincaré inequality and Gaussian integration by part. Firstly, the Poincaré inequality:

Proposition 2.2.9. *Let (x_1, \dots, x_n) be i.i.d. centered Gaussian random variable with variance 1, let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^1 , then*

$$\text{Var} \left(f(x_1, \dots, x_n) \right) \leq \mathbb{E} \left[\|\nabla f(x_1, \dots, x_n)\|_2^2 \right].$$

For more details about the Poincaré inequality, we refer to Definition 4.4.2 in [13]. As for Gaussian integration by part, it comes from the following formula, if Z is a centered Gaussian variable with variance 1 and f a \mathcal{C}^1 function, then

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[\partial_Z f(Z)] . \quad (2.8)$$

A direct consequence of this, is that if x and y are centered Gaussian variable with variance 1, and $Z = \frac{x+iy}{\sqrt{2}}$, then

$$\mathbb{E}[Zf(x, y)] = \mathbb{E}[\partial_Z f(x, y)] \quad \text{and} \quad \mathbb{E}[\bar{Z}f(x, y)] = \mathbb{E}[\partial_{\bar{Z}} f(x, y)] , \quad (2.9)$$

where $\partial_Z = \frac{1}{2}(\partial_x + i\partial_y)$ and $\partial_{\bar{Z}} = \frac{1}{2}(\partial_x - i\partial_y)$. When working with GUE matrices, an important consequence of this are the so-called Schwinger-Dyson equation, which we summarize in the following proposition. For more information about these equations and their applications, we refer to [13], Lemma 5.4.7.

Proposition 2.2.10. *Let X^N be GUE matrices of size N , $Q \in \mathcal{A}_{d,q}$, then for any i ,*

$$\mathbb{E} \left[\frac{1}{N} \text{Tr}_N (X_i^N Q(X^N)) \right] = \mathbb{E} \left[\left(\frac{1}{N} \text{Tr}_N \right)^{\otimes 2} (\partial_i Q(X^N)) \right].$$

Proof. One can write $X_i^N = \frac{1}{\sqrt{N}}(x_{r,s}^i)_{1 \leq r,s \leq N}$ and thus

$$\begin{aligned}
 \mathbb{E} \left[\frac{1}{N} \operatorname{Tr}_N(X_i^N Q(X^N)) \right] &= \frac{1}{N^{3/2}} \sum_{r,s} \mathbb{E} \left[x_{r,s}^i \operatorname{Tr}_N(E_{r,s} Q(X^N)) \right] \\
 &= \frac{1}{N^{3/2}} \sum_{r,s} \mathbb{E} \left[\operatorname{Tr}_N(E_{r,s} \partial_{x_{r,s}^i} Q(X^N)) \right] \\
 &= \frac{1}{N^2} \sum_{r,s} \mathbb{E} \left[\operatorname{Tr}_N(E_{r,s} \partial_i Q(X^N) \# E_{s,r}) \right] \\
 &= \mathbb{E} \left[\left(\frac{1}{N} \operatorname{Tr}_N \right)^{\otimes 2} (\partial_i Q(X^N)) \right].
 \end{aligned}$$

□

Now to finish this section we state a property that we use several times in this paper:

Proposition 2.2.11. *There exist constants C, D and α such that for any $N \in \mathbb{N}$, if X^N is a GUE random matrix of size N , then for any $u \geq 0$,*

$$\mathbb{P} \left(\|X^N\| \geq u + D \right) \leq e^{-\alpha u N}.$$

Consequently, for any $k \leq \alpha N/2$,

$$\mathbb{E} \left[\|X^N\|^k \right] \leq C^k.$$

Proof. The first part is a direct consequence of Lemma 2.2 from [40] in the specific case of the GUE. As for the second part, if $k \leq \alpha N/2$, then we have,

$$\begin{aligned}
 \mathbb{E} \left[\|X^N\|^k \right] &= k \int_0^\infty \mathbb{P} \left(\|X^N\| \geq u \right) u^{k-1} du \\
 &\leq kD^k + k \int_D^\infty e^{-N\alpha(u-D)} u^{k-1} du \\
 &\leq kD^k + ke^{DN\alpha} \int_D^\infty e^{(k-N\alpha)u} du \\
 &\leq kD^k + \frac{2k}{\alpha N} e^{kD} \leq C^k
 \end{aligned}$$

for some C independent of N and k . In the third line we used that $\ln |u| \leq u$ for all positive real numbers,

□

2.3 Proof of Theorem 2.1.1

2.3.1 Overview of the proof

Given two families of non-commutative random variables, $(X^N \otimes I_M, Z^{NM})$ and $(x \otimes I_M, Z^{NM})$, we want to study the difference between their distributions. As mentioned in the introduction, the main idea of the proof is to interpolate these two families with the help of d free Ornstein-Uhlenbeck processes $X^{t,N}$ started in the matrices X^N . However, as we shall

explain just below, we are only interested into the law of the marginals at time t of this process, and thus since there exists a simplified expression of this law, we do not need to define it as a stochastic process. We refer to [73] for the reader interested by the global definition. Many properties of the free Ornstein-Uhlenbeck process are similar to the classical case. For example, if $(S_t)_t$ is an Ornstein-Uhlenbeck process, then it is well-known that for any function f and $t \geq 0$,

$$\mathbb{E}[f(S_t)] = \mathbb{E}[f(e^{-t/2}S_0 + (1 - e^{-t})^{1/2}X)]$$

where X is a centered Gaussian random variable of variance 1 independent of S_0 . Likewise, if μ is the trace on the C^* -algebra which contains $(X_t^N)_{t \geq 0}$, we have for any function f such that this is well-defined and $t \geq 0$,

$$\mu(f(X_t^N)) = \tau_N \left(f(e^{-t/2}X^N + (1 - e^{-t})^{1/2}x) \right) \quad (2.10)$$

where x is a system of free semicircular variables, free from $\mathbb{M}_N(\mathbb{C})$. Thus a free Ornstein-Uhlenbeck process started at time t has the same distribution in the sense of Definition 2.2.1 as the family

$$e^{-t/2}X^N + (1 - e^{-t})^{1/2}x .$$

Consequently, from now on, we write $X_t^N = e^{-t/2}X^N + (1 - e^{-t})^{1/2}x$. Since our aim in this subsection is not to give a proof but to outline the strategy used in subsection 2.3.2, we also assume that we have no matrix Z^{NM} and that $M = 1$. Now under the assumption that this is well-defined, if $Q \in \mathcal{A}_{d,0} = \mathbb{C}\langle X_1, \dots, X_d \rangle$,

$$\mathbb{E} \left[\frac{1}{N} \text{Tr}_N \left(Q(X^N) \right) \right] - \tau \left(Q(x) \right) = - \int_0^\infty \mathbb{E} \left[\frac{d}{dt} \left(\tau_N \left(Q(X_t^N) \right) \right) \right] dt .$$

On the other hand, it is known that by using the free Markov property of the free Brownian motion, we have for $Q \in \mathcal{A}_{d,0}$

$$\frac{d}{dt} \tau_N(Q(X_t^N)) = -\frac{1}{2} \sum_i \left\{ \tau_N \left((X_t^N)_i (D_i Q)(X_t^N) \right) - \tau_N \otimes \tau_N \left((\partial_i D_i Q)(X_t^N) \right) \right\} .$$

One can already recognize the Schwinger-Dyson equation. Indeed thanks to Proposition 2.2.10, one can see that

$$\mathbb{E} \left[\frac{d}{dt} \tau_N(Q(X_t^N)) \right] \Big|_{t=0} = -\frac{1}{2} \sum_i \mathbb{E} \left[\tau_N \left(X_i^N (D_i Q)(X^N) \right) - \tau_N \otimes \tau_N \left((\partial_i D_i Q)(X^N) \right) \right] = 0 .$$

And then, thanks to Proposition 2.2.6,

$$\mathbb{E} \left[\frac{d}{dt} \tau_N(Q(X_t^N)) \right] \Big|_{t=\infty} = -\frac{1}{2} \sum_i \left\{ \tau \left(x_i (D_i Q)(x) \right) - \tau \otimes \tau \left((\partial_i D_i Q)(x) \right) \right\} = 0 .$$

However what happens at time t is much harder to estimate and is the core of the proof. The main idea to deal with this issue is to view the family (X^N, x) as the asymptotic limit when k goes to infinity of the family $(X^N \otimes I_k, R^{kN})$ where R^{kN} are independent GUE matrices of size kN and independent of X^N .

Another issue is that to prove Theorem 2.1.1, we would like to set $Q = f(P)$ but since f is not polynomial this means that we need to extend the definition of operators such as ∂_i . In order to do so we assume that there exist μ a measure on \mathbb{R} such that,

$$\forall x \in \mathbb{R}, \quad f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y) .$$

While we have to assume that the support of μ is indeed on the real line, μ can be a complex measure. However we will usually work with measure such that $|\mu|(\mathbb{R})$ is finite. Indeed under this assumption we can use Fubini's Theorem, and we get

$$\mathbb{E} \left[\frac{1}{M} \text{Tr}_N \left(f \left(P(X^N) \right) \right) \right] - \tau \left(f \left(P(x) \right) \right) = \int_{\mathbb{R}} \left\{ \mathbb{E} \left[\frac{1}{N} \text{Tr}_N \left(e^{iyP(X^N)} \right) \right] - \tau \left(e^{iyP(x)} \right) \right\} d\mu(y) .$$

We can then set $Q = e^{iyP}$. And even though this is not a polynomial function, since it is a power series, most of the properties associated to polynomials remain true with some assumption on the convergence. The main difficulty with this method is that we need to find a bound which does not depend on too high moments of y . Indeed terms of the form

$$\int_{\mathbb{R}} |y|^l d|\mu|(y)$$

appear in our estimates. Thanks to Fourier integration we can relate the exponent l to the regularity of the function f , thus we want to find a bound with l as small as possible. It turns out that with our proof $l = 4$.

2.3.2 Proof of Theorem 2.1.1

In this section we focus on proving Theorem 2.1.1 from which we deduce all of the important corollaries. It will be a consequence of the following Theorem :

Theorem 2.3.1. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $x = (x_1, \dots, x_d)$ a system of free semicircular variables,
- $Z^{NM} = (Z_1^{NM}, \dots, Z_q^{NM})$ deterministic matrices,
- $P \in \mathcal{A}_{d,q}$ a polynomial that we assume to be self-adjoint,
- $f : \mathbb{R} \mapsto \mathbb{R}$ such that there exists a measure on the real line μ with $\int (1+y^4) d|\mu|(y) < +\infty$ and for any $x \in \mathbb{R}$,

$$f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y) .$$

Then, there exists a polynomial $L_P \in \mathbb{R}^+[X]$ which only depends on P such that with $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$, for any N, M ,

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{MN} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau_N \otimes \tau_M \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| \\ & \leq \frac{M^2}{N^2} L_P \left(\|Z^{NM}\| \right) \int_{\mathbb{R}} (|y| + y^4) d|\mu|(y) . \end{aligned}$$

The proof is a direct corollary of Lemmas 2.3.3 and 2.3.4 below. The first one shows that the crux of the proof lies in understanding the following quantity:

Definition 2.3.2. *Let the following objects be given,*

- $\alpha, \beta \in [0, 1]$,
- $A, B, C, D \in \mathcal{A}_{d,q}$ monomials,
- $X_t^N = e^{-t/2} X^N + (1 - e^{-t})^{1/2} x$
- $Z_t^N = (X_t^N \otimes I_M, Z^{NM}, Z^{NM*})$,
- $S_t = (Ae^{i\beta y^P} B)(Z_t^N)$,
- $V_t = (Ce^{i\alpha y^P} D)(Z_t^N)$.

Then we define:

$$\begin{aligned} \mathcal{S}_{N,t}^{\alpha,\beta}(A, B, C, D) &= \mathbb{E} \left[\frac{1}{N} \sum_{1 \leq s, r \leq N} \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M \times S_t \times E_{r,s} \otimes I_M \times V_t \right) \right] \\ &\quad - \mathbb{E} \left[\tau_M \left((\tau_N \otimes I_M)(S_t) (\tau_N \otimes I_M)(V_t) \right) \right]. \end{aligned}$$

We can now state the next lemma which explains why this object appears:

Lemma 2.3.3. *Let f be a function such that there exists a measure μ such that for any $x \in \mathbb{R}$,*

$$f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y)$$

We also assume that $\int_{\mathbb{R}} (1 + y^4) d|\mu|(y) < \infty$. Then one can write

$$\mathbb{E} \left[\frac{1}{MN} \text{Tr} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau_N \otimes \tau_M \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right)$$

as a finite linear combination of terms of the following kinds :

$$\int_0^\infty e^{-t} \int y^2 \int_0^1 \mathcal{S}_{N,t}^{\alpha, 1-\alpha}(A, B, C, D) d\alpha d\mu(y) dt, \quad (2.11)$$

and

$$\int_0^\infty e^{-t} \int y \mathcal{S}_{N,t}^{1,0}(A, B, C, D) d\mu(y) dt \quad (2.12)$$

where the monomials $A, B, C, D \in \mathcal{A}_{d,q}$ and the coefficients of the linear combination are uniquely determined by P .

Proof. First, we define the natural interpolation between the trace of matrices at size N and the trace of semicircular variables,

$$s(t, y) = \mathbb{E} \left[\tau_N \otimes \tau_M \left(e^{iyP(Z_i^N)} \right) \right].$$

By definition of f we have

$$\begin{aligned} \int_{\mathbb{R}} s(0, y) d\mu(y) &= \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{MN} \left(f(P(X^N \otimes I_M, Z^{NM}, Z^{NM*})) \right) \right], \\ \int_{\mathbb{R}} s(\infty, y) d\mu(y) &= \tau_N \otimes \tau_M \left(f(P(x \otimes I_M, Z^{NM}, Z^{NM*})) \right). \end{aligned}$$

Thus under the assumption that this is well-defined, we have

$$\begin{aligned} &\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{MN} \left(f(P(X^N \otimes I_M, Z^{NM}, Z^{NM*})) \right) \right] - \tau_N \otimes \tau_M \left(f(P(x \otimes I_M, Z^{NM}, Z^{NM*})) \right) \\ &= - \int_0^\infty \int_{\mathbb{R}} \partial_t s(t, y) d\mu(y) dt. \end{aligned} \quad (2.13)$$

We compute

$$\partial_t s(t, y) = iy \frac{e^{-t}}{2} \mathbb{E} \left[\tau_N \otimes \tau_M \left(e^{iyP(Z_i^N)} \sum_i \partial_i P(Z_i^N) \# \left(\left(\frac{x_i}{(1-e^{-t})^{1/2}} - e^{t/2} X_i^N \right) \otimes I_M \right) \right) \right]. \quad (2.14)$$

Since we assumed that μ is such that $\int f(1+y^4)d\mu(y) < +\infty$ and that since X_i^N and x_i have all moments uniformly bounded by Lemma 2.2.11, we can find a constant C independent from y and t such that

$$|\partial_t s(t, y)| \leq C y e^{-t/2},$$

we can deduce that (2.13) is well-defined. Besides, writing $P = \sum c_V(P)V$ with monomials $V \in \mathcal{A}_{d,q}$, we get

$$\begin{aligned} \partial_t s(t, y) &= iy \frac{e^{-t}}{2} \sum c_V(P) \sum_{V=BX_iA} \mathbb{E} \left[\tau_N \otimes \tau_M \left(A(Z_i^N) e^{iyP(Z_i^N)} B(Z_i^N) \right. \right. \\ &\quad \left. \left. \times \left(\frac{x_i}{(1-e^{-t})^{1/2}} - e^{t/2} X_i^N \right) \otimes I_M \right) \right]. \end{aligned} \quad (2.15)$$

Hence, $\partial_t s$ is a finite linear combination of terms of the form

$$y e^{-t} S_t(A, B) = y e^{-t} S_t^1(A, B) - y e^{-t} S_t^2(A, B) \quad (2.16)$$

with

$$S_t^1(A, B) = S_t(A, B, (1-e^{-t})^{-1/2} x_i) \text{ and } S_t^2(A, B) = S_t(A, B, e^{t/2} X_i^N)$$

where

$$S_t(A, B, G) = \mathbb{E} \left[\tau_N \otimes \tau_M \left(A(Z_i^N) e^{iyP(Z_i^N)} B(Z_i^N) \times G \otimes I_M \right) \right]. \quad (2.17)$$

We first study $S_t^2(A, B)$. We denote by $Q = Ae^{iyP}B$. We want to use Gaussian integration by part: if we set $\sqrt{N}X_i^N = (x_{s,r}^i)_{1 \leq s,r \leq N}$, then with $\partial_{x_{s,r}^i}$ as in equations (2.8) and (2.9), thanks to Duhamel formula

$$\begin{aligned} \sqrt{N}e^{t/2} \partial_{x_{s,r}^i} Q(Z_t^N) &= \partial_i A(Z_t^N) \#(E_{r,s} \otimes I_M) e^{iyP(Z_t^N)} B(Z_t^N) \\ &\quad + iy \int_0^1 A(Z_t^N) e^{i(1-\alpha)yP(Z_t^N)} \partial_i P(Z_t^N) \#(E_{r,s} \otimes I_M) e^{i\alpha yP(Z_t^N)} B(Z_t^N) d\alpha \\ &\quad + A(Z_t^N) e^{iyP(Z_t^N)} \partial_i B(Z_t^N) \#(E_{r,s} \otimes I_M). \end{aligned} \quad (2.18)$$

Consequently, expanding in $S_t^2(A, B)$ the product by X_i^N in terms of its entries, we have

$$\begin{aligned} S_t^2(A, B) &= e^{t/2} \mathbb{E} \left[\tau_N \otimes \tau_M \left((Ae^{iyP}B)(Z_t^N) X_i^N \otimes I_M \right) \right] \\ &= N^{-1/2} e^{t/2} \sum_{1 \leq s,r \leq N} \mathbb{E} \left[x_{s,r}^i \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M (Ae^{iyP}B)(Z_t^N) \right) \right] \\ &= \frac{1}{N} \sum_{1 \leq s,r \leq N} \mathbb{E} \left[\tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M e^{t/2} \partial_{x_{s,r}^i} Q(Z_t^N) \right) \right] \\ &= \mathbb{E} \left[\frac{1}{N} \sum_{1 \leq s,r \leq N} \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M \partial_i A \#(E_{r,s} \otimes I_M) e^{iyP} B \right) \right] \\ &\quad + iy \int_0^1 \mathbb{E} \left[\frac{1}{N} \sum_{1 \leq s,r \leq N} \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M A e^{i(1-\alpha)yP} \partial_i P \#(E_{r,s} \otimes I_M) e^{i\alpha yP} B \right) \right] d\alpha \\ &\quad + \mathbb{E} \left[\frac{1}{N} \sum_{1 \leq s,r \leq N} \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M A e^{iyP} \partial_i B \#(E_{r,s} \otimes I_M) \right) \right] \end{aligned} \quad (2.19)$$

where A, B, P are evaluated at Z_t^N . To deal with $S_t^1(A, B)$, since a priori we defined free integration by parts only for polynomials, we expand the exponential as a power series,

$$\begin{aligned} &\tau_N \otimes \tau_M \left(A(Z_t^N) e^{iyP(Z_t^N)} B(Z_t^N) \frac{x_i \otimes I_M}{(1 - e^{-t})^{1/2}} \right) \\ &= \sum_{k \geq 0} \frac{1}{k!} \tau_N \otimes \tau_M \left(A(Z_t^N) (iyP(Z_t^N))^k B(Z_t^N) \frac{x_i \otimes I_M}{(1 - e^{-t})^{1/2}} \right). \end{aligned}$$

We define $(\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) : (\mathcal{A}_N \otimes \mathbb{M}_M(\mathbb{C}))^{\otimes 2} \rightarrow M_M(\mathbb{C})$ the linear map which is defined on simple tensor by $(\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M)(A \otimes B) = (\tau_N \otimes I_M)(A) \times (\tau_N \otimes I_M)(B)$. Hence, thanks to Proposition 2.2.6, with the convention that $A \times (B \otimes C) \times D = (AB) \otimes (CD)$, we have

$$\begin{aligned}
 & \tau_N \otimes \tau_M \left(A(Z_t^N) (\mathbf{i}yP(Z_t^N))^k B(Z_t^N) \frac{x_i \otimes I_M}{(1 - e^{-t})^{1/2}} \right) \\
 &= \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(\partial_i A(Z_t^N) (\mathbf{i}yP(Z_t^N))^k B(Z_t^N) \right) \right) \\
 & \quad + \mathbf{i}y \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A(Z_t^N) (\mathbf{i}y)^{k-1} \sum_{1 \leq l \leq k} P(Z_t^N)^{l-1} \partial_i P(Z_t^N) P(Z_t^N)^{k-l} B(Z_t^N) \right) \right) \\
 & \quad + \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A(Z_t^N) (\mathbf{i}yP(Z_t^N))^k \partial_i B(Z_t^N) \right) \right).
 \end{aligned}$$

Now we can use the fact that

$$\frac{1}{k!} = \int_0^1 \frac{\alpha^{l-1} (1 - \alpha)^{k-l}}{(l-1)!(k-l)!} d\alpha,$$

to deduce that

$$\begin{aligned}
 & \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A(Z_t^N) \sum_{k \geq 1} \frac{(\mathbf{i}y)^{k-1}}{k!} \sum_{l=1}^k P(Z_t^N)^{l-1} \partial_i P(Z_t^N) P(Z_t^N)^{k-l} B(Z_t^N) \right) \right) \\
 &= \int_0^1 \sum_{k \geq 1} \sum_{l=1}^k \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A(Z_t^N) \frac{(\mathbf{i}y\alpha P(Z_t^N))^{l-1}}{(l-1)!} \partial_i P(Z_t^N) \right. \right. \\
 & \quad \left. \left. \frac{(\mathbf{i}y(1-\alpha)P(Z_t^N))^{k-l}}{(k-l)!} B(Z_t^N) \right) \right) d\alpha \\
 &= \int_0^1 \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A(Z_t^N) e^{\mathbf{i}(1-\alpha)yP(Z_t^N)} \partial_i P(Z_t^N) e^{\mathbf{i}\alpha yP(Z_t^N)} B(Z_t^N) \right) \right) d\alpha.
 \end{aligned}$$

And thus, after summation, we obtain

$$\begin{aligned}
 S_t^1(A, B) &= \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(\partial_i A e^{\mathbf{i}yP} B \right) \right) \\
 & \quad + \mathbf{i}y \int_0^1 \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A e^{\mathbf{i}(1-\alpha)yP} \partial_i P e^{\mathbf{i}\alpha yP} B \right) \right) d\alpha \\
 & \quad + \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(A e^{\mathbf{i}yP} \partial_i B \right) \right).
 \end{aligned}$$

Therefore, after making the difference (2.16) to compute $S_t(A, B)$, we conclude that the difference we wish to estimate in (2.13) is a linear combination of terms, whose coefficients only depend on P , of the form (2.11) and (2.12). □

Thus the next step is to study the quantity $\mathcal{S}_{N,t}^{\alpha,\beta}(A, B, C, D)$. More precisely we show:

Lemma 2.3.4. *There is a polynomial $L \in \mathbb{R}^+[X]$ which only depends on A, B, C, D and P such that with $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$, for any $\alpha, \beta \in [0, 1]$, $N \in \mathbb{N}$, $t \in \mathbb{R}^+$ and $y \in \mathbb{R}$,*

$$\left| \mathcal{S}_{N,t}^{\alpha,\beta} \left(A, B, C, D \right) \right| \leq \frac{(1+y^2)M^2}{N^2} L \left(\|Z^{NM}\| \right).$$

This lemma is a direct consequence of Lemmas 2.3.6 and 2.3.7. We first show that the family $(X^N \otimes I_M, x \otimes I_M, Z^{NM})$ is actually the asymptotic distribution (in the sense of Definition 2.2.1) as k goes to infinity of the family $(X^N \otimes I_{kM}, R^{kN} \otimes I_M, Z^{NM} \otimes I_k)$ where R^{kN} is a family of independent GUE random matrices of size kN . The advantage of this representation is that it allows us to use classical analysis, and to treat the GUE variables and the semi-circle variables in a more symmetric way. A direct proof using semi-circular variables should however be possible.

Proposition 2.3.5. *If R^{kN} is a family of independent GUE random matrices of size kN , independent of X^N , we set*

$$U_t^k = \left(\left(e^{-t/2} X^N \otimes I_k + (1 - e^{-t})^{1/2} R^{kN} \right) \otimes I_M, Z^{NM} \otimes I_k, Z^{NM*} \otimes I_k \right).$$

Then if $q = Ae^{i\beta y P} B$, we have that \mathbb{P}_{X^N} -almost surely for any t ,

$$(\tau_N \otimes I_M)(q(Z_t^N)) = \lim_{k \rightarrow \infty} \mathbb{E}_R \left[(\tau_{kN} \otimes I_M)(q(U_t^k)) \right],$$

where \mathbb{E}_R is the expectation with respect to R^{kN} . Here M, N are kept fixed.

Proof. This proposition is mostly a corollary of Theorem 5.4.5 of [13]. Indeed this theorem states that if R^{kN} are GUE matrices and D^{kN} are deterministic matrices such that

$$\sup_{l \in \mathbb{N}} \max_i \sup_{k \in \mathbb{N}} \left(\frac{1}{N} \text{Tr}(|D_i^{kN}|^l) \right)^{1/l} < \infty,$$

and if D^{kN} converges in distribution towards a family of noncommutative random variables d , then the family (R^{kN}, D^{kN}) in the noncommutative probability space $(\mathbb{M}_{kN}(\mathbb{C}), *, \mathbb{E}[\frac{1}{kN} \text{Tr}])$ converges in distribution towards the family (x, d) where x is a system of free semicircular variables free from d . In our situation we can write for every i ,

$$Z_i^{NM} = \sum_{1 \leq r, s \leq N} E_{r,s} \otimes A_{r,s,i}^M.$$

Thus, if $E^N = (E_{r,s})_{1 \leq r, s \leq N}$, we fix $D^{kN} = (X^N \otimes I_k, E^N \otimes I_k)$, $d = (X^N, E^N)$, and we can apply Theorem 5.4.5 from [13] to get that for any non-commutative polynomial P ,

$$\lim_{k \rightarrow \infty} \mathbb{E}_R \left[\tau_{kN}(P(R^{kN}, X^N \otimes I_k, E^N \otimes I_k)) \right] = \tau_N \left(P(x, X^N, E^N) \right).$$

Consequently, for any non-commutative polynomial P , we also have

$$\begin{aligned} & \lim_{k \rightarrow \infty} \mathbb{E}_R \left[\tau_{kN} \otimes I_M \left(P(R^{kN} \otimes I_M, X^N \otimes I_{kM}, E^N \otimes I_{kM}, I_{kN} \otimes A^M, I_{kN} \otimes (A^M)^*) \right) \right] \\ &= \tau_N \left(P(x \otimes I_M, X^N \otimes I_M, E^N \otimes I_M, I_N \otimes A^M, I_N \otimes (A^M)^*) \right). \end{aligned}$$

Hence, for any $P \in \mathcal{A}_{d,q}$,

$$\lim_{k \rightarrow \infty} \mathbb{E}_R \left[\tau_{kN} \otimes I_M(P(U_t^k)) \right] = \tau_N \otimes I_M \left(P(Z_t^N) \right). \quad (2.20)$$

Thanks to Property 2.2.11, we know that there exist $\alpha > 0$ and $D < \infty$ such that for all $u \geq D$, for N large enough, $\forall i \in [1, d]$,

$$\mathbb{P} \left(\left\| R_i^{kN} \right\| \geq u \right) \leq e^{-\alpha u kN}. \quad (2.21)$$

Since if $c_M(P)$ is the coefficient of P associated with the monomial M , one has

$$\left\| P(U_t^k) \right\| \leq \sum_{M \text{ monomials}} |c_M(P)| \left\| M(U_t^k) \right\|,$$

there exist constants L and C which do depend on $\left\| Z_j^{NM} \right\|$ and $\left\| X_i^N \right\|$ such that for N large enough

$$\mathbb{P} \left(\left\| P(U_t^k) \right\| \geq C \right) \leq e^{-LkN}. \quad (2.22)$$

Knowing this, let $f_\varepsilon \in \mathbb{C}[X]$ be a polynomial which is ε -close from $x \mapsto e^{i\beta yx}$ on the interval $[-1 - C, C + 1]$. Since one can always assume that $C > \left\| P(Z_t^N) \right\|$, we have, with $q = Ae^{i\beta yP} B$:

$$\left\| (\tau_N \otimes I_M)(q(Z_t^N)) - (\tau_N \otimes I_M)((Af_\varepsilon(P)B)(Z_t^N)) \right\| \leq D\varepsilon,$$

where D is some constant which can depend on the dimensions N, M but not on k . Thus

$$\begin{aligned} & \left\| (\tau_N \otimes I_M)(q(Z_t^N)) - \mathbb{E}_R \left[(\tau_{kN} \otimes I_M)(q(U_t^k)) \right] \right\| \\ & \leq D\varepsilon + D\mathbb{E}_R \left[\left\| (q - Af_\varepsilon(P)B)(U_t^k) \right\| \mathbf{1}_{\left\| P(U_t^k) \right\| \geq C+1} \right] \\ & \quad + \left\| (\tau_N \otimes I_M)((Af_\varepsilon(P)B)(Z_t^N)) - \mathbb{E}_R \left[(\tau_{kN} \otimes I_M)((Af_\varepsilon(P)B)(U_t^k)) \right] \right\| \end{aligned}$$

The last term goes to zero as k goes to infinity by (2.20). Besides

$$\begin{aligned} & \mathbb{E}_R \left[\left\| (q - Af_\varepsilon(P)B)(U_t^k) \right\| \mathbf{1}_{\left\| P(U_t^k) \right\| \geq C+1} \right] \\ & \leq \mathbb{E}_R \left[\left(\left\| A(U_t^k) \right\| \left\| B(U_t^k) \right\| + \left\| (Af_\varepsilon(P)B)(U_t^k) \right\| \right)^2 \right]^{1/2} \mathbb{P}(\left\| P(U_t^k) \right\| \geq C + 1)^{1/2}. \end{aligned}$$

The first term is bounded independently of k thanks to (2.21) and the second converges exponentially fast towards 0 thanks to (2.22). Consequently

$$\limsup_{k \rightarrow \infty} \left\| (\tau_N \otimes I_M)(q(Z_t^N)) - \mathbb{E}_R \left[(\tau_{kN} \otimes I_M)(q(U_t^k)) \right] \right\| \leq D\varepsilon.$$

Hence the conclusion follows since the left hand side does not depend on ε . □

Recall that by definition

$$\mathcal{S}_{N,t}^{\alpha,\beta}(A, B, C, D) := \mathbb{E}[\Lambda_{N,t}^{\alpha,\beta}(A, B, C, D)] \quad (2.23)$$

with, following the notations of Definition 2.3.2 :

$$\begin{aligned} \Lambda_{N,t}^{\alpha,\beta}(A, B, C, D) &= \frac{1}{N} \sum_{1 \leq s, r \leq N} \tau_N \otimes \tau_M \left(E_{s,r} \otimes I_M \times S_t \times E_{r,s} \otimes I_M \times V_t \right) \\ &\quad - \tau_M \left((\tau_N \otimes I_M)(S_t) (\tau_N \otimes I_M)(V_t) \right). \end{aligned}$$

By Proposition 2.3.5, we deduce that

$$\Lambda_{N,t}^{\alpha,\beta}(A, B, C, D) = \lim_{k \rightarrow \infty} \Lambda_{k,N,t}^{\alpha,\beta}(A, B, C, D) \quad (2.24)$$

where $\Lambda_{k,N,t}^{\alpha,\beta}(A, B, C, D)$ equals

$$\begin{aligned} &\mathbb{E}_R \left[\frac{1}{N} \sum_{1 \leq s, r \leq N} \tau_{kN} \otimes \tau_M \left(E_{s,r} \otimes I_k \otimes I_M (Ae^{i\beta y P} B)(U_t^k) E_{r,s} \otimes I_k \otimes I_M (Ce^{i\alpha y P} D)(U_t^k) \right) \right] \\ &\quad - \tau_M \left(\mathbb{E}_R \left[\tau_{kN} \otimes I_M (Ae^{i\beta y P} B)(U_t^k) \right] \mathbb{E}_R \left[\tau_{kN} \otimes I_M (Ce^{i\alpha y P} D)(U_t^k) \right] \right) \end{aligned} \quad (2.25)$$

We can now prove the following intermediary lemma in view of deriving Lemma 2.3.4.

Lemma 2.3.6. *Define U_t^k as in Proposition 2.3.5, and let*

- $P_{1,2} = I_N \otimes E_{1,2} \otimes I_M,$
- $Q = (Ae^{i\beta y P} B)(U_t^k),$
- $T = (Ce^{i\alpha y P} D)(U_t^k).$

Then there is a constant C and a polynomial $L \in \mathbb{R}^+[X, Y]$ which only depend on A, B, C, D and P such that with $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$ and $\|X^N\| = \sup_{1 \leq i \leq d} \|X_i^N\|$, for any $\alpha, \beta \in [0, 1]$, $M, N \in \mathbb{N}$, $t \in \mathbb{R}^+$ and $y \in \mathbb{R}$,

$$\begin{aligned} |\Lambda_{k,N,t}^{\alpha,\beta}(A, B, C, D)| &\leq \frac{(1+y^2)M^2}{N^2} L \left(\|Z^{NM}\|, \|X^N\| \right) \\ &\quad + k^3 |\tau_M (\mathbb{E}_R [(\tau_{kN} \otimes I_M)(QP_{1,2})] \mathbb{E}_R [(\tau_{kN} \otimes I_M)(TP_{1,2})])|. \end{aligned} \quad (2.26)$$

Proof. We denote in short $\Lambda_{k,N,t}^{\alpha,\beta}(A, B, C, D) = \Lambda_{k,N,M} = \tau_M (\mathbb{E}_R[\Gamma_{k,N,M}]) - \Theta_{k,N,M}$ with

$$\begin{aligned} \Gamma_{k,N,M} &= \frac{1}{N} \sum_{1 \leq s, r \leq N} \tau_{kN} \otimes I_M (E_{s,r} \otimes I_k \otimes I_M Q E_{r,s} \otimes I_k \otimes I_M T) \\ \Theta_{k,N,M} &= \tau_M \left(\mathbb{E}_R [\tau_{kN} \otimes I_M Q] \mathbb{E}_R [\tau_{kN} \otimes I_M (T)] \right) \end{aligned} \quad (2.27)$$

Let $(g_i)_{i \in [1, N]}$ and $(f_i)_{i \in [1, k]}$ be the canonical basis of \mathbb{C}^N and \mathbb{C}^k , $E_{i,j}$ is the matrix whose only non-zero coefficient is (i, j) and this coefficient has value 1, the size of the matrix $E_{i,j}$ will

depend on the context. We use the fact that $E_{r,s} = g_r g_s^*$ and $I_k = \sum_l E_{l,l}$ with $E_{l,l} = f_l^* f_l$ to deduce that

$$\begin{aligned}
 \Gamma_{k,N,M} &= \frac{1}{N} \sum_{1 \leq s, r \leq N} \sum_{1 \leq l, l' \leq k} \tau_{kN} \otimes I_M (E_{s,r} \otimes E_{l,l} \otimes I_M Q E_{r,s} \otimes E_{l',l'} \otimes I_M T) \\
 &= \frac{1}{N^2 k} \sum_{1 \leq l, l' \leq k} \sum_{1 \leq r \leq N} g_r^* \otimes f_l^* \otimes I_M Q g_r \otimes f_{l'} \otimes I_M \sum_{1 \leq s \leq N} g_s^* \otimes f_{l'}^* \otimes I_M T g_s \otimes f_l \otimes I_M \\
 &= \frac{1}{k} \sum_{1 \leq l, l' \leq k} (\tau_N \otimes I_M) (I_N \otimes f_l^* \otimes I_M Q I_N \otimes f_{l'} \otimes I_M) \\
 &\quad \times (\tau_N \otimes I_M) (I_N \otimes f_{l'}^* \otimes I_M T I_N \otimes f_l \otimes I_M) \\
 &= k \sum_{1 \leq l, l' \leq k} (\tau_{kN} \otimes I_M) (Q I_N \otimes E_{l',l} \otimes I_M) (\tau_{kN} \otimes I_M) (T I_N \otimes E_{l,l'} \otimes I_M). \tag{2.28}
 \end{aligned}$$

The last line of the above equation prompts us to set $P_{l',l} = I_N \otimes E_{l',l} \otimes I_M$. If $(e_i)_{i \in [1,M]}$ is the canonical basis of \mathbb{C}^M , we set

$$F_{l,l',u,v}^q(R^{kN}) = e_u^* (\tau_{kN} \otimes I_M) \left(q \left((e^{-t/2} X^N \otimes I_k + \left(\frac{1 - e^{-t}}{Nk} \right)^{1/2} R^{kN}) \otimes I_M, Z^{NM}, Z^{NM*} \right) P_{l',l} \right) e_v$$

with $q = Q = A e^{i\beta y P} B$ or $q = T = C e^{i\alpha y P} D$. We thus have with (2.28)

$$\begin{aligned}
 \tau_M (\mathbb{E}_R [\Gamma_{k,N,M}]) &= k \sum_{1 \leq l, l' \leq k} \tau_M \left(\mathbb{E}_R \left[(\tau_{kN} \otimes I_M) (Q P_{l',l}) (\tau_{kN} \otimes I_M) (T P_{l,l'}) \right] \right) \tag{2.29} \\
 &= \frac{k}{M} \sum_{\substack{1 \leq l, l' \leq k \\ 1 \leq u, v \leq M}} \text{Cov}_R \left(F_{l,l',u,v}^Q(R^{kN}), F_{l',l,u,v}^T(R^{kN}) \right) \\
 &\quad + k \sum_{1 \leq l, l' \leq k} \tau_M \left(\mathbb{E}_R \left[(\tau_{kN} \otimes I_M) (Q P_{l',l}) \right] \mathbb{E}_R \left[(\tau_{kN} \otimes I_M) (T P_{l,l'}) \right] \right).
 \end{aligned}$$

However, the law of U_t^k is invariant under conjugation by $I_N \otimes U \otimes I_M$, where $U \in M_k(\mathbb{C})$ is a permutation matrix. Therefore, if $l = l'$, $\mathbb{E}_R[\tau_{kN}(Q P_{l,l})] = \mathbb{E}_R[\tau_{kN}(Q P_{1,1})]$, and if $l \neq l'$, $\mathbb{E}_R[\tau_{kN}(Q P_{l',l})] = \mathbb{E}_R[\tau_{kN}(Q P_{1,2})]$. We get the same equation when replacing Q by T . Consequently, we get

$$\begin{aligned}
 &k \sum_{1 \leq l, l' \leq k} \mathbb{E}_R \left[(\tau_{kN} \otimes I_M) (Q P_{l',l}) \right] \mathbb{E}_R \left[(\tau_{kN} \otimes I_M) (T P_{l,l'}) \right] \\
 &= k^2 \mathbb{E}_R[(\tau_{kN} \otimes I_M)(Q P_{1,1})] \mathbb{E}_R[(\tau_{kN} \otimes I_M)(T P_{1,1})] \\
 &\quad + (k-1)k^2 \mathbb{E}_R[(\tau_{kN} \otimes I_M)(Q P_{1,2})] \mathbb{E}_R[(\tau_{kN} \otimes I_M)(T P_{1,2})].
 \end{aligned}$$

where the trace τ_M of the first term in the right hand side equals $\Theta_{k,N,M} = \tau_M(\mathbb{E}_R[(\tau_{kN} \otimes I_M)(Q)] \mathbb{E}_R[(\tau_{kN} \otimes I_M)(T)])$ because $I_M = \sum_l P_{l,l}$. Thus equation (2.29) yields

$$\begin{aligned}
 |\Lambda_{k,N,M}| &\leq \frac{k}{M} \sum_{\substack{1 \leq l, l' \leq k \\ 1 \leq u, v \leq M}} \left| \text{Cov}_R \left(F_{l,l',u,v}^Q(R^{kN}), F_{l',l,u,v}^T(R^{kN}) \right) \right| \tag{2.30} \\
 &\quad + \left| k^3 \tau_M \left(\mathbb{E}_R[(\tau_{kN} \otimes I_M)(Q P_{1,2})] \mathbb{E}_R[(\tau_{kN} \otimes I_M)(T P_{1,2})] \right) \right|.
 \end{aligned}$$

Hence, we only need to bound the first term to complete the proof of the lemma. Thanks to Cauchy-Schwartz's inequality, it is enough to bound the covariance of $F_{l,l',u,v}^q(R^{kN})$, for $q = Q$ and T . To study these variances, we shall use the Poincaré inequality, see Proposition 2.2.9. If we set $x_{r,s}^i$ and $y_{r,s}^i$ the real and imaginary part of $\sqrt{2kN}(R_i^{kN})_{r,s}$ for $r < s$ and $x_{r,r}^i = \sqrt{kN}(R_i^{kN})_{r,r}$, then these are real centered Gaussian random variables of variance 1 and one can view $F_{l,l',u,v}^q$ as a function on $(x_{r,s}^i)_{r \leq s, i}$ and $(y_{r,s}^i)_{r < s, i}$. By a computation similar to (2.18), we find

$$\begin{aligned} \frac{kN}{1 - e^{-t}} \left\| \nabla F_{l,l',u,v}^q \right\|_2^2 &= \sum_i \sum_{1 \leq r, s \leq kN} e_u^* (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right) e_v \\ &\quad \times e_v^* (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right)^* e_u. \end{aligned} \quad (2.31)$$

It is worth noting that here the matrices $E_{r,s}$ have size kN in this formula. Thanks to Poincaré inequality (see Proposition 2.2.9), we deduce

$$\begin{aligned} \frac{k}{M} \sum_{1 \leq u, v \leq M} \text{Var}_R(F_{l,l',u,v}^q(R_{kN})) &\leq \frac{k}{M} \sum_{1 \leq u, v \leq M} \mathbb{E} \left[\left\| \nabla F_{l,l',u,v}^q \right\|_2^2 \right] \\ &\leq \frac{1}{N} \sum_i \sum_{1 \leq r, s \leq kN} \mathbb{E}_R \left[\frac{1}{M} \sum_{1 \leq u, v \leq M} e_u^* (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right) e_v e_v^* \right. \\ &\quad \left. \times (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right)^* e_u \right] \\ &\leq \frac{1}{N} \sum_i \sum_{1 \leq r, s \leq kN} \mathbb{E}_R \left[\tau_M \left((\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right) \right. \right. \\ &\quad \left. \left. \times (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right)^* \right) \right]. \end{aligned} \quad (2.32)$$

Moreover we have, if e_l is an orthonormal basis of \mathbb{C}^k ,

$$\begin{aligned} &\sum_{1 \leq l, l' \leq k} \tau_M \left((\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right) (\tau_{kN} \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M P_{l',l} \right)^* \right) \\ &= \frac{1}{k^2} \sum_{1 \leq l, l' \leq k} \tau_M \left(e_l^* \otimes I_M (\tau_N \otimes I_k \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M \right) e_{l'} e_{l'}^* \otimes I_M \right. \\ &\quad \left. (\tau_N \otimes I_k \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M \right)^* e_l \otimes I_M \right) \\ &= \frac{1}{k} \tau_k \otimes \tau_M \left((\tau_N \otimes I_k \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M \right) (\tau_N \otimes I_k \otimes I_M) \left(\partial_i q \# E_{r,s} \otimes I_M \right)^* \right). \end{aligned} \quad (2.33)$$

Hence by combining equations (2.32) and (2.33) we have proved that

$$\begin{aligned}
 & \frac{k}{M} \sum_{\substack{1 \leq l, l' \leq k \\ 1 \leq u, v \leq M}} \text{Var}_R \left(F_{l, l', u, v}^q (R^{kN}) \right) \\
 & \leq \frac{1}{kN} \sum_i \sum_{1 \leq r, s \leq kN} \mathbb{E}_R \left[\tau_k \otimes \tau_M \left((\tau_N \otimes I_{kM}) \left(\partial_i q \# E_{r, s} \otimes I_M \right) (\tau_N \otimes I_{kM}) \left(\partial_i q \# E_{r, s} \otimes I_M \right)^* \right) \right]
 \end{aligned} \tag{2.34}$$

Moreover, let us remind that, with the convention $A \times (B \otimes C) \times D = (AB) \otimes (CD)$, we have (for $q = Q = Ae^{i\beta y P} B$ but with obvious changes for $q = T$)

$$\partial_i q = \partial_i A e^{i\beta y P} B + i\beta y A \int_0^1 e^{i(1-u)\beta y P} \partial_i P e^{iu\beta y P} B du + A e^{i\beta y P} \partial_i B.$$

Consequently, (2.34) is a finite linear combination of terms of the three following types $Q_N^i = \mathbb{E}_R[q_N^i]$, $1 \leq i \leq 3$, with

$$\begin{aligned}
 q_N^1 &= \frac{1}{kN} \sum_{1 \leq r, s \leq kN} \tau_k \otimes \tau_M \left((\tau_N \otimes I_{kM}) \left(A_1 E_{r, s} \otimes I_M A_2 e^{i\beta y P} A_3 \right) \right. \\
 & \quad \left. (\tau_N \otimes I_{kM}) \left(B_3 E_{s, r} \otimes I_M B_2 e^{-i\beta y P} B_1 \right) \right), \\
 q_N^2 &= \frac{\beta y}{kN} \int_0^1 \sum_{1 \leq r, s \leq kN} \tau_k \otimes \tau_M \left((\tau_N \otimes I_{kM}) \left(A_1 e^{i(1-u)\beta y P} A_2 E_{r, s} \otimes I_M A_3 e^{iu\beta y P} A_4 \right) \right. \\
 & \quad \left. (\tau_N \otimes I_{kM}) \left(B_3 E_{s, r} \otimes I_M B_2 e^{-i\beta y P} B_1 \right) \right) du, \\
 q_N^3 &= \frac{(\beta y)^2}{kN} \int_0^1 \int_0^1 \sum_{1 \leq r, s \leq kN} \tau_k \otimes \tau_M \left((\tau_N \otimes I_{kM}) \left(A_1 e^{i(1-u)\beta y P} A_2 E_{r, s} \otimes I_M A_3 e^{iu\beta y P} A_4 \right) \right. \\
 & \quad \left. (\tau_N \otimes I_{kM}) \left(B_4 e^{-iv\beta y P} B_3 E_{s, r} \otimes I_M B_2 e^{-i(1-v)\beta y P} B_1 \right) \right) du dv,
 \end{aligned} \tag{2.35}$$

where the A_i and B_i are monomials in U_t^k . Besides the coefficients of this linear combination only depend on A, B and P . We first show how to estimate q_N^3 . Let us recall that we set $(e_i)_{1 \leq i \leq N}$, $(f_i)_{1 \leq i \leq k}$ and $(g_i)_{1 \leq i \leq M}$ as the canonical basis of \mathbb{C}^M , \mathbb{C}^k and \mathbb{C}^N . Then, for any matrices $A, B, C, D \in \mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$, we have

$$\begin{aligned}
& \sum_{1 \leq r, s \leq kN} \text{Tr}_{kM} \left(\text{Tr}_N \otimes I_{kM} \left(A E_{r,s} \otimes I_M B \right) \times \text{Tr}_N \otimes I_{kM} \left(C E_{s,r} \otimes I_M D \right) \right) \quad (2.36) \\
&= \sum_{\substack{1 \leq a, b, r_1, s_1 \leq N \\ 1 \leq c, d, r_2, s_2 \leq k \\ 1 \leq e, f, g, h \leq M}} g_a^* \otimes f_c^* \otimes e_e^* A g_{r_1} \otimes f_{r_2} \otimes e_f \times g_{s_1}^* \otimes f_{s_2}^* \otimes e_f^* B g_a \otimes f_d \otimes e_g \\
&\quad \times g_b^* \otimes f_d^* \otimes e_g^* C g_{s_1} \otimes f_{s_2} \otimes e_h \times g_{r_1}^* \otimes f_{r_2}^* \otimes e_h^* D g_b \otimes f_c \otimes e_e \\
&= \sum_{\substack{1 \leq a \leq N \\ 1 \leq c, d \leq k \\ 1 \leq e, f, g, h \leq M}} g_a^* \otimes f_c^* \otimes e_e^* A I_N \otimes I_k \otimes (e_f e_h^*) D I_N \otimes (f_c f_d^*) \otimes (e_e e_g^*) \\
&\quad \times C I_N \otimes I_k \otimes (e_h e_f^*) B g_a \otimes f_d \otimes e_g \\
&= \sum_{1 \leq u, v \leq M} \text{Tr}_N \left(I_N \otimes \text{Tr}_{kM} (A I_{kN} \otimes e_u e_v^* D) I_N \otimes \text{Tr}_{kM} (C I_{kN} \otimes e_v e_u^* B) \right).
\end{aligned}$$

Let K_M be a GUE matrix of size M , independent of everything else. Performing a Gaussian integration by part, we get

$$\begin{aligned}
& \frac{1}{M} \sum_{1 \leq u, v \leq M} \text{Tr}_N \left(I_N \otimes \text{Tr}_{kM} (A I_{kN} \otimes e_u e_v^* D) I_N \otimes \text{Tr}_{kM} (C I_{kN} \otimes e_v e_u^* B) \right) \quad (2.37) \\
&= \mathbb{E}_K \left[\text{Tr}_N \left(I_N \otimes \text{Tr}_{kM} \left(A I_{kN} \otimes K_M D \right) I_N \otimes \text{Tr}_{kM} \left(C I_{kN} \otimes K_M B \right) \right) \right].
\end{aligned}$$

Consequently by combining equations (2.36) and (2.37), we have

$$\begin{aligned}
q_N^3 &= \left(\frac{\beta y M}{N} \right)^2 \int_0^1 \int_0^1 \mathbb{E}_K \left[\tau_N \left((I_N \otimes \tau_{kM}) \left(A_1 e^{i(1-u)\beta y P} A_2 I_{kN} \otimes K_M B_2 e^{-i(1-v)\beta y P} B_1 \right) \right. \right. \\
&\quad \left. \left. \times (I_N \otimes \tau_{kM}) \left(B_4 e^{-iv\beta y P} B_3 I_{kN} \otimes K_M A_3 e^{iu\beta y P} A_4 \right) \right) \right] du dv.
\end{aligned}$$

Since P is self-adjoint, we know that for any real r , $\|e^{i r P(U_t^k)}\| = 1$. Besides $\|I_N \otimes \tau_{kM}(A)\| \leq \|A\|$, thus we can bound q_N^3 in (2.35) by

$$|q_N^3| \leq \left(\frac{yM}{N} \right)^2 \|A_1\| \|A_2\| \|A_3\| \|A_4\| \|B_1\| \|B_2\| \|B_3\| \|B_4\| \mathbb{E}_K [\|K_M\|^2]. \quad (2.38)$$

Finally, by [15], $\mathbb{E}_K [\|K_M\|^2]$ is bounded by 3. One can bound similarly q_N^1 and q_N^2 , the only difference on the final result is that we would have 1 or y instead of y^2 . Finally after taking the expectation with respect to R^{kN} in equation (2.38) and using Proposition 2.2.11, we deduce that there exists S which only depends on A, B and P , hence is independent of N, M, y, t, α or β , such that the covariance in (2.34) is bounded by

$$\frac{k}{M} \sum_{\substack{1 \leq l, l' \leq k \\ 1 \leq u, v \leq M}} \text{Var}_R \left(F_{l, l', u, v}^q(R^{kN}) \right) \leq \frac{(1+y^2)M^2}{N^2} S \left(\|X^N\|, \|Z^{NM}\| \right).$$

Thus, we deduce that there exists a polynomial H which only depends on A, B, C, D and P such that the first term in the right hand side of (2.30) is bounded by

$$\frac{k}{M} \sum_{\substack{1 \leq l, l' \leq k \\ 1 \leq u, v \leq M}} \left| \text{Cov}_R \left(F_{l, l', u, v}^Q(R^{kN}), F_{l', l, u, v}^T(R^{kN}) \right) \right| \leq \frac{(1+y^2)M^2}{N^2} H \left(\|X^N\|, \|Z^{NM}\| \right). \quad (2.39)$$

This completes the proof of the Lemma in the general case. For the specific case where $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, we can get better estimate in equation (2.38) thanks to a refinement of equation (2.37). Indeed if A, B, C, D are monomials in U_t^k , then we can write $A = A_1 \otimes A_2$ in $\mathbb{M}_{kN}(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$ and likewise for B, C, D such that A_2, B_2, C_2, D_2 commute. Thus,

$$\begin{aligned} & \frac{1}{M} \sum_{1 \leq u, v \leq M} \text{Tr}_N \left(I_N \otimes \text{Tr}_{kM} (A I_{kN} \otimes e_u e_v^* D) I_N \otimes \text{Tr}_{kM} (C I_{kN} \otimes e_v e_u^* B) \right) \\ &= \frac{1}{M} \text{Tr}_N \left(I_N \otimes \text{Tr}_k (A_1 D_1) I_N \otimes \text{Tr}_k (C_1 B_1) \right) \sum_{1 \leq u, v \leq M} \text{Tr}_M (A_2 e_u e_v^* D_2) \text{Tr}_M (C_2 e_v e_u^* B_2) \\ &= \frac{1}{M} \text{Tr}_N \left(I_N \otimes \text{Tr}_k (A_1 D_1) I_N \otimes \text{Tr}_k (C_1 B_1) \right) \text{Tr}_M (D_2 A_2 B_2 C_2) \\ &= \frac{1}{M} \text{Tr}_N \left(I_N \otimes \text{Tr}_k (A_1 D_1) I_N \otimes \text{Tr}_k (C_1 B_1) \right) \text{Tr}_M (A_2 D_2 C_2 B_2) \\ &= \frac{1}{M} \text{Tr}_{NM} \left(I_{NM} \otimes \text{Tr}_k (AD) I_{NM} \otimes \text{Tr}_k (CB) \right). \end{aligned}$$

By linearity and density this equality is true if we assume that A, B, C, D are power series in U_t^k . Thus combining this equality with equation (2.36), we get that in this case

$$|q_N^3| \leq \left(\frac{y}{N} \right)^2 \|A_1\| \|A_2\| \|A_3\| \|A_4\| \|B_1\| \|B_2\| \|B_3\| \|B_4\|.$$

The same argument as in the general case applies and the proof follows. \square

In order to prove Lemma 2.3.4, we show in the following lemma that the term appearing in the second line of equation (2.26) vanishes.

Lemma 2.3.7. *Let $U_t^k, P_{1,2}, Q$ and T be defined as in Lemma 2.3.6, then \mathbb{P}_{X^N} -almost surely,*

$$\lim_{k \rightarrow \infty} k^3 \tau_M \left(\mathbb{E}_R \left[(\tau_{kN} \otimes I_M)(QP_{1,2}) \right] \mathbb{E}_R \left[(\tau_{kN} \otimes I_M)(TP_{1,2}) \right] \right) = 0.$$

Proof. It is enough to show that given $y \in \mathbb{R}$ and monomial A and B , we have

$$\lim_{k \rightarrow \infty} k^{3/2} \mathbb{E}_R \left[(\tau_{kN} \otimes I_M) \left((A e^{iyP} B)(U_t^k) P_{1,2} \right) \right] = 0.$$

For this purpose, let us define for monomials A, B and $y \geq 0$

$$f_{A,B}(y) = \mathbb{E}_R \left[(\text{Tr}_{kN} \otimes I_M) ((A e^{iyP} B)(U_t^k) P_{1,2}) \right].$$

We want to show that $f_{A,B}$ goes to zero faster than $k^{-1/2}$. We first show that we can reduce the problem to the case $y = 0$. To this end, we also define

$$d_n(y) = \sup_{\deg(A) + \deg(B) \leq n} \|f_{A,B}(y)\|.$$

We know thanks to Proposition 2.2.11 that there exist constants α and C such that for any i and $n \leq \alpha k N / 2$,

$$\mathbb{E} \left[\|R_i^{kN}\|^n \right] \leq C^n.$$

Besides, with as previously $(f_i)_{i \in [1,k]}$ the canonical basis of \mathbb{C}^k ,

$$\begin{aligned} \left\| (\text{Tr}_{kN} \otimes I_M) ((A e^{iyP} B)(U_t^k) P_{1,2}) \right\| &= \left\| (\text{Tr}_N \otimes I_M) (f_2^* \otimes I_{MN} (A e^{iyP} B)(U_t^k) f_1 \otimes I_{MN}) \right\| \\ &\leq N \times \left\| f_2^* \otimes I_{MN} (A e^{iyP} B)(U_t^k) f_1 \otimes I_{MN} \right\| \\ &\leq N \times \left\| A(U_t^k) \right\| \left\| B(U_t^k) \right\|. \end{aligned}$$

Consequently, P_{X^N} -almost surely, there exist constants γ and D (which do depend on, N , $\|X^N\|$ and $\|Z^{NM}\|$) such that for any $n \leq \gamma k$,

$$d_n(y) \leq D^n. \quad (2.40)$$

It is important to point out that this constant D can be very large when N is, it does not matter since, in the end, we will show that this quantity will go towards 0 when k goes to infinity and the other parameters such as N, M or y are fixed. Next, we define

$$g_{k,a}(y) = \sum_{0 \leq n \leq \gamma k} d_n(y) a^n.$$

But we have

$$\frac{df_{A,B}(y)}{dy} = \mathbf{i} \mathbb{E}_R \left[(\text{Tr}_{kN} \otimes I_M) ((A P e^{iyP} B)(U_t^k) P_{1,2}) \right]$$

so that if we set $c_L(P)$ to be the coefficient associated to the monomial L in P , $P = \sum c_L(P)L$,

$$\left| \frac{df_{A,B}(y)}{dy} \right| \leq \sum_{L \text{ monomials}} |c_L(P)| d_{\deg(A) + \deg(B) + \deg(L)}(y).$$

Thus, for any $y \geq 0$, any monomials A, B with $\deg(A) + \deg(B) = n$,

$$f_{A,B}(y) \leq f_{A,B}(0) + \sum_{L \text{ monomials}} |c_L(P)| \int_0^y d_{n+\deg(L)}(u) du.$$

Therefore, we have for $y \geq 0$ and any $n \geq 0$,

$$a^n d_n(y) \leq a^n d_n(0) + \sum_{L \text{ monomials}} |c_L(P)| a^{-\deg(L)} \int_0^y d_{n+\deg(L)}(u) a^{n+\deg(L)} du.$$

And with $\|\cdot\|_{a^{-1}}$ defined as in (2.7), thanks to (2.40), we find a finite constant c_a independent of k such that

$$g_{k,a}(y) \leq g_{k,a}(0) + c_a(aD)^{\gamma k} + \|P\|_{a^{-1}} \int_0^y g_{k,a}(u) du,$$

As a consequence of Gronwall's inequality, we deduce that for $y \geq 0$,

$$g_{k,a}(y) \leq \left(g_{k,a}(0) + c_a(aD)^{\gamma k}\right) e^{y\|P\|_{a^{-1}}}. \quad (2.41)$$

Next we will prove that for a small enough, $g_{k,a}(0) = O(1/k)$. Then at the cost of by replacing P by $-P$, thanks to equation (2.41) we have that for a small enough and any $y \in \mathbb{R}$, $g_{k,a}(y) = O(1/k)$. This completes the proof since with A, B, C, D as in Lemma 2.3.6, for a small enough and k large enough

$$\begin{aligned} & \left| k^3 \tau_M \left(\mathbb{E}_R [(\tau_{kN} \otimes I_M)(QP_{1,2})] \mathbb{E}_R [(\tau_{kN} \otimes I_M)(TP_{1,2})] \right) \right| \\ & \leq k \times d_{\deg(A)+\deg(B)}(\beta y) d_{\deg(C)+\deg(D)}(\alpha y) \\ & \leq k \times a^{-\deg(A)-\deg(B)-\deg(C)-\deg(D)} g_{k,a}(\beta y) g_{k,a}(\alpha y) \end{aligned}$$

Hence, it is enough to find an estimate on $g_{k,a}(0)$. First for any j , one can write $Z_j^{NM} = \sum_{1 \leq u, v \leq N} E_{u,v} \otimes I_k \otimes A_{u,v}^j$ for some matrices $A_{u,v}^j$, then we define

$$U_{N,k} = \left(R^{kN}, X^N \otimes I_k, (E_{u,v} \otimes I_k)_{u,v} \right), \quad c_n = \sup_{\deg(L) \leq n, L \text{ monomial}} \left| \mathbb{E}_R [\text{Tr}_{kN}(L(U_{N,k}) P_{1,2})] \right|.$$

Note that since we are taking the trace of $L(U_{N,k})P_{1,2}$ with $P_{1,2} = I_N \otimes f_1 f_2^* \otimes I_M$, we have $c_0 = c_1 = 0$. We consider K the supremum over u, v, j of $\|A_{u,v}^j\|$, we also assume without loss of generality that $K \geq 1$. Thus, since

$$Z_j^{NM} = \sum_{1 \leq u, v \leq N} E_{u,v} \otimes I_k \otimes A_{u,v}^j, \quad X_t^N = e^{-t/2} X^N \otimes I_k + (1 - e^{-t})^{1/2} R^{kN},$$

if L is a monomial in $U_t^k = (X_t^N \otimes I_M, Z^{NM} \otimes I_k, Z^{NM*} \otimes I_k)$ of degree n , then we can view $L(U_t^k)$ as a sum of at most $2^n N^{2n}$ monomials in $e^{-t/2} X^N \otimes I_k, (1 - e^{-t})^{1/2} R^{kN}, E_{u,v} \otimes I_k \otimes A_{u,v}^j, E_{v,u} \otimes I_k \otimes A_{u,v}^{j*}$. Consequently, since $\sup_{u,v,j} \|A_{u,v}^j\| \leq K$, we have

$$\left\| \mathbb{E}_R \left[\text{Tr}_{kN} \otimes I_M (L(U_t^k) P_{1,2}) \right] \right\| \leq 2^n N^{2n} K^n c_n.$$

Thus, if we set

$$f_p(a) = \sum_{0 \leq n \leq p} c_n a^n,$$

we have

$$g_{k,a}(0) \leq f_{\gamma k}(2N^2 K a). \quad (2.42)$$

Now we need to study the behaviour of $f_k(a)$ when k goes to infinity for a small enough. In order to do so, let us consider a monomial L in $U_{N,k}$. Since $X^N \otimes I_k$ and $E_{u,v} \otimes I_k$ commute with $P_{1,2}$, one can assume that $L = R_i^{kN} S$ for some i (unless L is a monomial in $X^N \otimes I_k$ and $E_{u,v} \otimes I_k$ in which case $\text{Tr}_{kN}(LP_{1,2}) = 0$), thus thanks to Schwinger-Dyson equation (see Proposition 2.2.10),

$$\mathbb{E}_R [\mathrm{Tr}_{kN}(LP_{1,2})] = \frac{1}{Nk} \mathbb{E}_R [\mathrm{Tr}_{kN} \otimes \mathrm{Tr}_{kN}(\partial_i(SP_{1,2}))] = \frac{1}{Nk} \sum_{S=UR_iV} \mathbb{E}[\mathrm{Tr}_{Nk}(U) \mathrm{Tr}_{Nk}(VP_{1,2})]. \quad (2.43)$$

To use this Schwinger-Dyson equation as an inductive bound we shall use Poincaré inequality to bound the covariance in the above right hand side. We hence compute for any monomial V ,

$$\begin{aligned} \|\nabla \mathrm{Tr}_{kN}(VP_{1,2})\|_2^2 &= \frac{1}{Nk} \sum_i \sum_{r,s} \mathrm{Tr}_{kN}(\partial_s V \# E_{r,s} P_{1,2}) \mathrm{Tr}_{kN}(\partial_s V \# E_{s,r} P_{1,2})^* \\ &= \sum_i \sum_{V=AR_iB, V=CR_iD} \frac{1}{Nk} \mathrm{Tr}_{kN}(BP_{1,2}AC^*P_{1,2}D^*) \end{aligned} \quad (2.44)$$

Thus with $\Theta = \max\{C, \|X^N\|, 1\}$, since $P_{1,2}$ is of rank N , we get

$$\mathrm{Var}_R(\mathrm{Tr}_{kN}(VP_{1,2})) \leq \frac{1}{k} (\deg V)^2 \Theta^{2 \deg V}.$$

Likewise, for any monomial U , we find

$$\mathrm{Var}_R(\mathrm{Tr}_{kN}(U)) \leq (\deg U)^2 \Theta^{2 \deg U}.$$

Therefore, if n is the degree of L , we deduce from (2.44), (2.43) and Poincaré inequality that

$$\begin{aligned} |\mathbb{E}_R [\mathrm{Tr}_{kN}(LP_{1,2})]| &\leq \frac{1}{k^{3/2}N} \sum_{i=0}^{n-2} i(n-2-i) \Theta^n + \sum_{S=UR_iV} \left| \frac{1}{Nk} \mathbb{E}_R[\mathrm{Tr}_{kN}(U)] \mathbb{E}_R[\mathrm{Tr}_{kN}(VP_{1,2})] \right| \\ &\leq \frac{n^3 \Theta^n}{k^{3/2}N} + \sum_{S=UR_iV} |\mathbb{E}_R[\mathrm{Tr}_{kN}(VP_{1,2})]| \Theta^{\deg U}. \end{aligned}$$

By replacing D by $\max\{D, \Theta\}$, we can always assume that $\Theta < D$. We also bound N^{-1} by 1, thus for $n \geq 2$,

$$c_n \leq \frac{n^3 D^n}{k^{3/2}} + \sum_{i=0}^{n-2} c_i D^{n-2-i}.$$

We use this estimate to bound $f_g(a)$ with g such that $g^3 D^g \leq \sqrt{k}$. Since $c_0 = c_1 = 0$ and for any $n \leq g$, $n^3 D^n k^{-3/2} \leq k^{-1}$, we have for $aD < 1$

$$f_g(a) = \sum_{n=2}^g c_n a^n \leq \frac{1}{k} \times \frac{a^2 - a^{g+1}}{1-a} + a^2 \sum_{m=0}^{g-2} \sum_{n=0}^m c_i D^{n-i} a^n \leq \frac{1}{k} \times \frac{a^2}{1-a} + a^2 \frac{f_g(a)}{1-Da}.$$

Thus, for a small enough,

$$f_g(a) \leq \frac{(1-Da)a^2}{(1-a)(1-Da-a^2)} \times \frac{1}{k}.$$

Besides, we want g such that $g^3 D^g \leq \sqrt{k}$, hence we can take g the integer part of $\frac{\ln k}{2(\ln D+3)}$. Since by definition we have $c_n \leq \Theta^n$, this also means that $c_n \leq D^n$, thus

$$\sum_{g < n \leq \gamma k} c_n a^n \leq \sum_{n > g} (Da)^n \leq \frac{(Da)^{g+1}}{1 - Da} \leq k^{\frac{\ln(Da)}{2(\ln D+3)}} \times \frac{1}{1 - Da}.$$

Thus, if we fix a small enough, $f_{\gamma k}(a) = O(1/k)$. Hence, we deduce from (2.42) that for a small enough (depending on N, K but not k) there exists a finite constant C independent of k such that

$$g_{k,a}(0) \leq f_k(2N^2 K a) \leq \frac{C}{k}.$$

□

We can now prove Theorem 2.1.1.

Proof of Theorem 2.1.1. It is based on Theorem 2.3.1. To use it, we would like to take the Fourier transform of f and use Fourier inversion formula. However we did not assume that f is integrable. Thus the first step of the proof is to show that up to a term of order e^{-N} , we can assume that f has compact support. Thanks to Proposition 2.2.11, there exist constants D and α such that for any N and i , for any $u \geq 0$,

$$\mathbb{P}\left(\|X_i^N\| \geq u + D\right) \leq e^{-\alpha u N}.$$

Thus, there exist constants C and K , independent of M, N, P and f , such that

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \mathbf{1}_{\{\exists i, \|X_i^N\| > D+1\}} \right) \right] \right| \\ & \leq \mathbb{E} \left[\left\| f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right\| \mathbf{1}_{\{\exists i, \|X_i^N\| > D+1\}} \right] \\ & \leq \|f\|_\infty \mathbb{P} \left(\exists i, \|X_i^N\| > D + 1 \right) \\ & \leq C \|f\|_\infty e^{-KN}. \end{aligned}$$

There exists a polynomial $H \in \mathbb{R}^+[X]$ which only depends on P such that

$$\left\| P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right\| \mathbf{1}_{\{\forall i, \|X_i^N\| \leq D+1\}} \leq H \left(\|Z^{NM}\| \right).$$

We can also assume that $\left\| P(x \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \leq H \left(\|Z^{NM}\| \right)$. We take g a \mathcal{C}^∞ -function which takes value 1 on $[-H \left(\|Z^{NM}\| \right), H \left(\|Z^{NM}\| \right)]$, 0 on $[-H \left(\|Z^{NM}\| \right) - 1, H \left(\|Z^{NM}\| \right) + 1]^c$ and belongs to $[0, 1]$ elsewhere. From the bound above, we deduce

$$\begin{aligned}
& \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| \\
& \leq \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \mathbf{1}_{\{\forall i, \|X_i^N\| \leq D+1\}} \right] \right. \\
& \quad \left. - \tau \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| + C \|f\|_\infty e^{-KN} \\
& \leq \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr} \left((fg) \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] \right. \\
& \quad \left. - \tau \left((fg) \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| + 2C \|f\|_\infty e^{-KN}.
\end{aligned} \tag{2.45}$$

Since fg has compact support and can be differentiated six times, we can take its Fourier transform and then invert it so that with the convention $\hat{h}(y) = \frac{1}{2\pi} \int_{\mathbb{R}} h(x) e^{-ixy} dx$, we have

$$\forall x \in \mathbb{R}, \quad (fg)(x) = \int_{\mathbb{R}} e^{ixy} \widehat{fg}(y) dy.$$

Besides, since if h has compact support bounded by K then $\|\hat{h}\|_\infty \leq 2K \|h\|_\infty$, we have

$$\begin{aligned}
\int_{\mathbb{R}} (|y| + y^4) |\widehat{fg}(y)| dy & \leq \int_{\mathbb{R}} \frac{|y| + |y|^3 + y^4 + y^6}{1 + y^2} |\widehat{fg}(y)| dy \\
& \leq \int_{\mathbb{R}} \frac{|\widehat{(fg)^{(1)}}(y)| + |\widehat{(fg)^{(3)}}(y)| + |\widehat{(fg)^{(4)}}(y)| + |\widehat{(fg)^{(6)}}(y)|}{1 + y^2} dy \\
& \leq 2 \left(H(\|Z^{NM}\|) + 1 \right) \|fg\|_{C^6} \int_{\mathbb{R}} \frac{1}{1 + y^2} dy \\
& \leq C \left(H(\|Z^{NM}\|) + 1 \right) \|f\|_{C^6},
\end{aligned}$$

for some absolute constant C . Hence fg satisfies the hypothesis of Theorem 2.3.1 with $\mu(dy) = \widehat{fg}(y)dy$. Therefore, combining with equation (2.45), we conclude that

$$\begin{aligned}
& \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr} \left(f \left(P \left(X^N \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right] - \tau \left(f \left(P \left(x \otimes I_M, Z^{NM}, Z^{NM*} \right) \right) \right) \right| \\
& \leq \|f\|_\infty e^{-KN} + \frac{M^2}{N^2} L_P(\|Z^{NM}\|) \int_{\mathbb{R}} (|y| + y^4) |\widehat{fg}(y)| dy \\
& \leq \frac{M^2}{N^2} \left(CL_P(\|Z^{NM}\|) \left(H(\|Z^{NM}\|) + 1 \right) + e^{-KN} \right) \|f\|_{C^6}.
\end{aligned}$$

□

2.4 Consequences of the main result

In this section, we deduce Corollaries 2.1.3 and 2.1.4, as well as Theorems 2.1.2 and 2.1.5.

2.4.1 Proof of Corollary 2.1.3

We could directly apply Theorem 2.1.1 to $f_z : x \rightarrow (z - x)^{-1}$, however we have $\|f\|_{\mathbb{C}^6} = O(|\Im z|^7)$ when we want an exponent 5. Since $\overline{G_{P(x)}(z)} = G_{P(x)}(\bar{z})$ we can assume that $\Im z < 0$, but then

$$f_z(x) = \int_0^\infty e^{ixy} (\mathbf{i}e^{-iyz}) dy.$$

Consequently, with $\mu_z(dy) = \mathbf{i}e^{-iyz} dy$, we have

$$\int_0^\infty (y + y^4) d|\mu_z|(y) = \frac{1}{|\Im z|^2} + \frac{24}{|\Im z|^5}.$$

Thus, by applying Theorem 2.3.1 with $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_p^M)$, P and f_z , we have

$$\left| \mathbb{E} \left[G_{P(X^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(x \otimes I_M, I_N \otimes Y^M)}(z) \right| \leq \frac{M^2}{N^2} L_P \left(\|Z^{NM}\| \right) \int_{\mathbb{R}} (1 + y^4) d|\mu_z|(y).$$

Now since

$$\|Z^{NM}\| = \sup_{1 \leq i \leq p} \|Y_i^M\| = \|Y^M\|$$

which does not depend on N , we get the desired estimate

$$\left| \mathbb{E} \left[G_{P(X^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(x \otimes I_M, I_N \otimes Y^M)}(z) \right| \leq \frac{M^2}{N^2} L_P \left(\|Y^M\| \right) \left(\frac{1}{|\Im z|^2} + \frac{24}{|\Im z|^5} \right).$$

2.4.2 Proof of Corollary 2.1.4

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz function uniformly bounded by 1 and with Lipschitz constant at most 1. We want to bound from above the quantity

$$\Delta_{N,M}(f) = \left| \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} \left(f \left(P \left(X^N \otimes I_M, I_N \otimes Y_M \right) \right) \right) \right] - \tau \otimes \tau_M \left(f \left(P \left(x \otimes I_M, I_N \otimes Y_M \right) \right) \right) \right| \quad (2.46)$$

Firstly, one can see that with the same argument as in the proof of Theorem 2.1.1 (in particular equation (2.45)), we can assume that the support of f is bounded by a constant $S = H(\|Y^M\|)$ for some polynomial H independent of everything. However, we cannot apply directly Theorem 2.1.1 since f is not regular enough. In order to deal with this issue we use the convolution with Gaussian random variable, thus let G be a centered Gaussian random variable, we set

$$f_\varepsilon : x \rightarrow \mathbb{E}[f(x + \varepsilon G)].$$

Since f has Lipschitz constant 1, we have for any $x \in \mathbb{R}$,

$$|\mathbb{E}[f(x + \varepsilon G)] - f(x)| \leq \varepsilon.$$

Since f_ε is regular enough we could now apply Theorem 2.1.1, however we get a better result by using Theorem 2.3.1. Indeed we have

$$\begin{aligned} f_\varepsilon(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x + \varepsilon y) e^{-y^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) \frac{e^{-\frac{(x-y)^2}{2\varepsilon^2}}}{\varepsilon} dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} e^{i(x-y)u} e^{-(u\varepsilon)^2/2} du dy. \end{aligned}$$

Since the support of f is bounded, we can apply Fubini's Theorem:

$$f_\varepsilon(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \int_{\mathbb{R}} f(y) e^{-iyu} dy e^{-(u\varepsilon)^2/2} du.$$

And so with the convention $\hat{h}(u) = \frac{1}{2\pi} \int_{\mathbb{R}} h(y) e^{-iyu} dy$, we have

$$f_\varepsilon(x) = \int_{\mathbb{R}} e^{iux} \hat{f}(u) e^{-(u\varepsilon)^2/2} du.$$

Thus, if we set $\mu_\varepsilon(dy) = \hat{f}(y) e^{-(y\varepsilon)^2/2} dy$, then, since $\|f\|_\infty \leq 1$,

$$\int_{\mathbb{R}} (1 + y^4) d|\mu_\varepsilon|(y) \leq 2S \int_{\mathbb{R}} (1 + y^4) e^{-y^2/2} dy \varepsilon^{-5}.$$

Consequently, we can apply Theorem 2.3.1 with f_ε and since $\|f - f_\varepsilon\|_\infty \leq \varepsilon$, there exists a polynomial R_P such that the difference in (2.46) can be bounded by:

$$\Delta_{N,M}(f) \leq 2\varepsilon + R_P(\|Y^M\|) \frac{M^2}{N^2\varepsilon^5}.$$

We finally choose $\varepsilon = N^{-1/3}$ to get the desired bound

$$\Delta_{N,M}(f) \leq 2R_P(\|Y^M\|) \frac{M^2}{N^{1/3}}.$$

2.4.3 Proof of Theorem 2.1.2

Firstly, we need to define properly the operator norm of tensor of \mathcal{C}^* -algebras. When writing the proof it appears that we should work with the minimal tensor product.

Definition 2.4.1. *Let \mathcal{A} and \mathcal{B} be \mathcal{C}^* -algebras with faithful representations $(H_{\mathcal{A}}, \phi_{\mathcal{A}})$ and $(H_{\mathcal{B}}, \phi_{\mathcal{B}})$, then if \otimes_2 is the tensor product of Hilbert spaces, $\mathcal{A} \otimes_{\min} \mathcal{B}$ is the completion of the image of $\phi_{\mathcal{A}} \otimes \phi_{\mathcal{B}}$ in $B(H_{\mathcal{A}} \otimes_2 H_{\mathcal{B}})$ for the operator norm in this space. This definition is independent of the representations that we fixed.*

The following two lemmas are well known facts in operator algebra. The first one is Lemma 4.1.8 from [75]:

Lemma 2.4.2. *Let $(\mathcal{A}, \tau_{\mathcal{A}})$ and $(\mathcal{B}, \tau_{\mathcal{B}})$ be \mathcal{C}^* -algebra with faithful traces, then $\tau_{\mathcal{A}} \otimes \tau_{\mathcal{B}}$ extends uniquely to a faithful trace $\tau_{\mathcal{A} \otimes_{\min} \mathcal{B}}$ on $\mathcal{A} \otimes_{\min} \mathcal{B}$.*

We did not find a reference with an explicit proof for the following Lemma, so we give our own. In order to learn more about this second lemma, especially how to weaken the hypothesis, we refer to [31].

Lemma 2.4.3. *Let \mathcal{C} be an exact \mathcal{C}^* -algebra endowed with a faithful state $\tau_{\mathcal{C}}$, let $Y^N \in \mathcal{A}_N$ be a sequence of families of noncommutative random variables in a \mathcal{C}^* -algebra \mathcal{A}_N which converges strongly towards a family Y in a \mathcal{C}^* -algebra \mathcal{A} endowed with a faithful state $\tau_{\mathcal{A}}$. Let $S \in \mathcal{C}$ be a family of noncommutative random variables, then the family $(S \otimes 1, 1 \otimes Y^N)$ converges strongly in distribution towards the family $(S \otimes 1, 1 \otimes Y)$.*

Proof. The following sets

$$\mathcal{M} = \left\{ (x_N)_{N \in \mathbb{N}} \mid x_N \in \mathcal{A}_N, \sup_{N \geq 0} \|x_N\| < \infty \right\},$$

$$\mathcal{I} = \left\{ (x_N)_{N \in \mathbb{N}} \in \mathcal{M} \mid \lim_{N \rightarrow \infty} \|x_N\| = 0 \right\},$$

are \mathcal{C}^* -algebras for the norm $\|x\| = \sup_{N \geq 0} \|x_N\|$. We also define

$$\mathcal{B} = \mathcal{C}^*((Y_N)_{N \in \mathbb{N}}, \mathcal{I}),$$

the \mathcal{C}^* -algebra generated by \mathcal{I} and the family $(Y_N)_{N \in \mathbb{N}}$. Since \mathcal{I} is a closed ideal of \mathcal{B} , by Theorem 3.1.4 of [76], \mathcal{B}/\mathcal{I} is a \mathcal{C}^* -algebra for the quotient norm. We naturally have the following exact sequence

$$0 \rightarrow \mathcal{I} \rightarrow \mathcal{B} \rightarrow \mathcal{B}/\mathcal{I} \rightarrow 0.$$

And by hypothesis, since \mathcal{C} is exact, we have the following exact sequence

$$0 \rightarrow \mathcal{C} \otimes_{\min} \mathcal{I} \rightarrow \mathcal{C} \otimes_{\min} \mathcal{B} \rightarrow \mathcal{C} \otimes_{\min} (\mathcal{B}/\mathcal{I}) \rightarrow 0.$$

By definition, this means that $(\mathcal{C} \otimes_{\min} \mathcal{B})/(\mathcal{C} \otimes_{\min} \mathcal{I}) \simeq \mathcal{C} \otimes_{\min} (\mathcal{B}/\mathcal{I})$. If $\pi_{\mathcal{I}}$ is the quotient map from \mathcal{B} to \mathcal{B}/\mathcal{I} , the isomorphism between these two spaces is

$$f : x + \mathcal{C} \otimes_{\min} \mathcal{I} \mapsto \text{id}_{\mathcal{C}} \otimes_{\min} \pi_{\mathcal{I}}(x).$$

Hence

$$f(P(1 \otimes (Y_N)_{N \in \mathbb{N}}, S \otimes 1) + \mathcal{C} \otimes_{\min} \mathcal{I}) = P(1 \otimes ((Y_N)_{N \in \mathbb{N}} + \mathcal{I}), S \otimes 1). \quad (2.47)$$

Let (H, φ) be a faithful representation of \mathcal{C} , and (H_N, φ_N) a faithful representation of \mathcal{A}_N . The direct sum $(\bigoplus_{N \in \mathbb{N}} H_N, \bigoplus_{N \in \mathbb{N}} \varphi_N)$ is a faithful representation of \mathcal{M} and consequently of \mathcal{B} too. More precisely, it is defined by

$$\bigoplus_{N \in \mathbb{N}} H_N = \left\{ (x_N)_{N \in \mathbb{N}} \mid x_N \in H_N, \sum_N \|x_N\|_2^2 < \infty \right\}.$$

Consequently, by definition of the spatial tensor product, it is the completion of the algebraic tensor $\mathcal{C} \otimes \mathcal{B}$ in the operator space $B(H \otimes_2 (\bigoplus_N H_N))$ endowed with the operator norm. The

notation \otimes_2 means that we completed the algebraic tensor $H \overline{\otimes} (\oplus_N H_N)$ to make it a Hilbert space. It is important to see that this space is isomorphic to $\oplus_N (H \otimes_2 H_N)$, indeed it means that if P is a noncommutative polynomial, then

$$\left\| P\left(1 \otimes (Y_N)_{N \in \mathbb{N}}, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{B}} = \sup_{N \geq 0} \left\| P\left(1 \otimes Y_N, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}_N}.$$

Consequently by using the definition of the quotient norm, we have

$$\left\| P\left(1 \otimes (Y_N)_{N \in \mathbb{N}}, S \otimes 1\right) + \mathcal{C} \otimes_{\min} \mathcal{I} \right\|_{(\mathcal{C} \otimes_{\min} \mathcal{B}) / (\mathcal{C} \otimes_{\min} \mathcal{I})} = \limsup_{N \rightarrow \infty} \left\| P\left(1 \otimes Y_N, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}_N}. \quad (2.48)$$

Since f is a \mathcal{C}^* -algebra isomorphism, thanks to (2.47), we have

$$\begin{aligned} & \left\| P\left(1 \otimes (Y_N)_{N \in \mathbb{N}}, S \otimes 1\right) + \mathcal{C} \otimes_{\min} \mathcal{I} \right\|_{(\mathcal{C} \otimes_{\min} \mathcal{B}) / (\mathcal{C} \otimes_{\min} \mathcal{I})} \\ &= \left\| P\left(1 \otimes ((Y_N)_{N \in \mathbb{N}} + \mathcal{I}), S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} (\mathcal{B}/\mathcal{I})}. \end{aligned}$$

By definition of \mathcal{I} , if P is a noncommutative polynomial, we have

$$\left\| P((Y_N)_{N \in \mathbb{N}} + \mathcal{I}) \right\|_{\mathcal{B}/\mathcal{I}} = \left\| P(Y) \right\|_{\mathcal{A}}.$$

For our purposes, we can assume that $\mathcal{A} = \mathcal{C}^*(Y)$. Therefore the map

$$P((Y_N)_{N \in \mathbb{N}} + \mathcal{I}) \in \mathbb{C}\langle (Y_N)_{N \in \mathbb{N}} + \mathcal{I} \rangle \mapsto P(Y) \in \mathbb{C}\langle Y \rangle$$

is well-defined and is an isometry. Thus since $\mathbb{C}\langle (Y_N)_{N \in \mathbb{N}} + \mathcal{I} \rangle$ is dense in \mathcal{B}/\mathcal{I} and $\mathbb{C}\langle Y \rangle$ is dense in \mathcal{A} , this isometry extends into an isomorphism between \mathcal{B}/\mathcal{I} and \mathcal{A} . Consequently

$$\left\| P\left(1 \otimes ((Y_N)_{N \in \mathbb{N}} + \mathcal{I}), S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} (\mathcal{B}/\mathcal{I})} = \left\| P\left(1 \otimes Y, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}}.$$

Thus, combined with (2.48), we have

$$\limsup_{N \rightarrow \infty} \left\| P\left(1 \otimes Y_N, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}_N} = \left\| P\left(1 \otimes Y, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}}. \quad (2.49)$$

Finally let f be a function which takes value 0 on $(-\infty, \left\| P(1 \otimes Y, S \otimes 1) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}} - \varepsilon]$ and positive value on $(\left\| P(1 \otimes Y, S \otimes 1) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}} - \varepsilon, \infty)$. Since the family $(S \otimes 1, 1 \otimes Y^N)$ converges clearly in distribution towards the family $(S \otimes 1, 1 \otimes Y)$, we have

$$\lim_{N \rightarrow \infty} \tau_{\mathcal{C}} \otimes_{\min} \tau_{\mathcal{A}_N} \left(f(P(1 \otimes Y_N, S \otimes 1)) \right) = \tau_{\mathcal{C}} \otimes_{\min} \tau_{\mathcal{A}} \left(f(P(1 \otimes Y, S \otimes 1)) \right).$$

Thanks to Lemma 2.4.2, we know that $\tau_{\mathcal{C}} \otimes_{\min} \tau_{\mathcal{A}}$ is faithful, consequently

$$\tau_{\mathcal{C}} \otimes_{\min} \tau_{\mathcal{A}} \left(f(P(1 \otimes Y, S \otimes 1)) \right) > 0.$$

This means that for N large enough, $\tau_{\mathcal{C}} \otimes_{\min} \tau_{\mathcal{A}_N} \left(f(P(1 \otimes Y_N, S \otimes 1)) \right) > 0$, thus for any $\varepsilon > 0$,

$$\liminf_{N \rightarrow \infty} \left\| P\left(1 \otimes Y_N, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}_N} \geq \left\| P\left(1 \otimes Y, S \otimes 1\right) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}} - \varepsilon.$$

This allows to conclude with (2.49) that

$$\lim_{N \rightarrow \infty} \left\| P(1 \otimes Y_N, S \otimes 1) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}_N} = \left\| P(1 \otimes Y, S \otimes 1) \right\|_{\mathcal{C} \otimes_{\min} \mathcal{A}}.$$

□

In order to prove Theorem 2.1.2 we use well-known concentration properties of Gaussian random variable coupled with an estimation of the expectation, let us begin by stating the concentration properties (see [13] Lemma 2.3.3).

Proposition 2.4.4. *Let G be a Lipschitz function on \mathbb{R}^n with Lipschitz constant K for the ℓ^2 -norm $\|\gamma\|_2 = (\sum_i \gamma_i^2)^{1/2}$, $\gamma = (\gamma_1, \dots, \gamma_n)$ independent centered Gaussian random variable of variance 1. Then for all $\delta > 0$,*

$$\mathbb{P}(G(\gamma) - \mathbb{E}[G(\gamma)] \geq \delta) \leq e^{-\frac{\delta^2}{2K^2}}.$$

In our situation, we have p independent GUE matrices $(X^{N,i})_s$ of size N , hence we fix γ the random vector of size dN^2 which consists of the union of $(\sqrt{N}X_{s,s}^{N,i})_{i,s}$, $(\sqrt{2N}\Re X_{s,r}^{N,i})_{s < r, i}$ and $(\sqrt{2N}\Im X_{s,r}^{N,i})_{s < r, i}$ which are indeed centered Gaussian random variable of variance 1 as stated in Definition 2.2.8. We would like to apply Proposition 2.4.4 to

$$G_N(\gamma) = \left\| P^* P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\|.$$

However G_N is not Lipschitz on \mathbb{R}^{dN^2} because of its polynomial behaviour at infinity. Hence we cannot use directly Proposition 2.4.4. The following lemma is a well-known tool for this kind of situation, the proof can be found in [77, Lemma 5.9].

Lemma 2.4.5. *Let (X, d) be a metric space and μ a probability measure on (X, d) which satisfies a concentration inequality, i.e. for all $f : X \rightarrow \mathbb{R}$ with Lipschitz constant $|f|_{\mathcal{L}}$, for all $\delta > 0$,*

$$\mu\left(|f - \mu(f)| \geq \delta\right) \leq e^{-g\left(\frac{\delta}{|f|_{\mathcal{L}}}\right)}$$

for some increasing function g on \mathbb{R}^+ . Let B be a subset of X and $|f|_{\mathcal{L}}^B$ be the Lipschitz constant of f as a function from B to \mathbb{R} . Let $\delta(f) = \mu(\mathbf{1}_{x \in B^c} (|f(x)| + \sup_{u \in B} |f(u)| + |f|_{\mathcal{L}}^B d(x, B)))$, then

$$\mu\left(|f - \mu(f)| \geq \delta + \delta(f)\right) \leq \mu(B^c) + e^{-g\left(\frac{\delta}{|f|_{\mathcal{L}}^B}\right)}.$$

We can now prove the concentration inequality that we will use in the rest of this paper. To simplify notations we will write M instead of M_N . We also set $Z^{NM} = (Z^N \otimes I_M, I_N \otimes Y^M)$ and $Z = (z \otimes 1, 1 \otimes y)$.

Proposition 2.4.6. *Let $P \in \mathcal{A}_{d,p+q}$, there are some polynomials $H_P, K_P \in \mathbb{R}^+[X]$ which only depends on P such that with $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$, for any N, M ,*

$$\begin{aligned} & \mathbb{P}\left(\left| \left\| P^* P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| - \mathbb{E}\left[\left\| P^* P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \right] \right| \right. \\ & \quad \left. \geq \delta + K_P\left(\|Z^{NM}\|\right) e^{-N}\right) \leq d e^{-2N} + e^{-\frac{\delta^2 N}{H_P(\|Z^{NM}\|)}}. \end{aligned}$$

Proof. We want to use Lemma 2.4.5 and Proposition 2.4.4. The metric space we will work with is \mathbb{R}^n endowed with the Euclidian norm, and we can take the function g to be $g : x \mapsto x^2/2$ by Lemma 2.4.4. Thus we get that for any $B \subset \mathbb{R}^n$, for any $G : \mathbb{R}^n \mapsto \mathbb{R}$, if $\gamma = (\gamma_1, \dots, \gamma_n)$ is a vector of independent centered Gaussian random variables of variance 1, then for all $\delta > 0$,

$$\mathbb{P}(G(\gamma) - \mathbb{E}[G(\gamma)] \geq \delta + \delta(G)) \leq e^{-\frac{\delta^2}{2(\|G\|_{\mathcal{L}}^B)^2}}. \quad (2.50)$$

If $0 \in B$ as it will be the case later on, we have $\delta(G) \leq \mathbb{E}[\mathbf{1}_{\gamma \notin B}(|G(\gamma)| + \sup_{u \in B} |G(u)| + |f|_{\mathcal{L}}^B \|\gamma\|_2)]$. We set $B_N = \{\forall i, \|X_i^N\| \leq D\}$ where D was chosen thanks to 2.2.11 such that for any N and i ,

$$\mathbb{P}(\|X_i^N\| \geq D) \leq e^{-2N}. \quad (2.51)$$

Thus we have $\mathbb{P}(B_N^c) \leq d e^{-2N}$. With γ the vector of size dN^2 which consists of the union of $(\sqrt{N}X_{s,s}^{N,i})_{i,s}$, $(\sqrt{2N} \Re X_{s,r}^{N,i})_{s < r, i}$ and $(\sqrt{2N} \Im X_{s,r}^{N,i})_{s < r, i}$, we set

$$G_N(\gamma) = \left\| P^* P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\|.$$

One can see that on B_N we can find a polynomial H'_P such that for any N and Z^{NM} ,

$$|G_N(\gamma) - G_N(\tilde{\gamma})| \leq H'_P(\|Z^{NM}\|) \sum_i \|X_i^N - \tilde{X}_i^N\|,$$

where $\|\cdot\|$ is the operator norm. Besides

$$\sum_i \|X_i^N - \tilde{X}_i^N\| \leq \sum_i \text{Tr}_N \left((X_i^N - \tilde{X}_i^N)^* (X_i^N - \tilde{X}_i^N) \right)^{1/2} \leq \frac{2^d}{\sqrt{N}} \|\gamma - \tilde{\gamma}\|_2.$$

Thus, on B_N , G_N has Lipschitz constant $2^d H'_P(\|Z^{NM}\|) N^{-1/2}$. Consequently with (2.50), we get that

$$\mathbb{P}(G_N(\gamma) - \mathbb{E}[G_N(\gamma)] \geq \delta + \delta(G_N)) \leq e^{-\frac{\delta^2 N}{2^{d+1} H'_P(\|Z^{NM}\|)^2}}.$$

Therefore, we set $H_P = 2^{d+1} H'_P$, we also have that $\|\gamma\|_2^2 = N \sum_i \text{Tr}_N((X_i^N)^2)$. Consequently we have some polynomial K'_P such that,

$$\delta(G) \leq \mathbb{E} \left[\mathbf{1}_{\{\exists i, \|X_i^N\| > D\}} \left(|G_N(\gamma)| + K'_P(\|Z^{NM}\|) + 2^d H'_P(\|Z^{NM}\|) N^{1/2} \sqrt{\sum_i \|X_i^N\|^2} \right) \right]$$

Hence the conclusion thanks to Proposition 2.2.11 and our choice of D in equation (2.51). \square

We can now prove Theorem 2.1.2. Firstly, we can assume that Z^N and Y^M are deterministic matrices by Fubini's Theorem. The convergence in distribution is a well-known theorem, we refer to [13], Theorem 5.4.5. We set g a C^∞ function which takes value 0 on $(-\infty, 1/2]$ and value 1 on $[1, \infty)$, and belongs to $[0, 1]$ otherwise. Let us define $f_\varepsilon : t \mapsto g(\varepsilon^{-1}(t - \|PP^*(x \otimes 1, Z, Z^*)\|))$. By Theorem 2.1.1, there exists a constant C which only depends on P , $\sup_M \|Y^M\|$ and $\sup_N \|Z^N\|$ (which is finite thanks to the strong convergence assumption on Z^N) such that,

$$\begin{aligned}
 & \left| \mathbb{E} \left[\text{Tr}_{MN} \left(f_\varepsilon \left(PP^* \left(X^N \otimes I_M, Z^{NM}, Z^{NM^*} \right) \right) \right) \right] \right. \\
 & \quad \left. - MN \tau_N \otimes \tau_M \left(f_\varepsilon \left(PP^* \left(x \otimes I_M, Z^{NM}, Z^{NM^*} \right) \right) \right) \right| \\
 & \leq C \varepsilon^{-6} \frac{M^3}{N}.
 \end{aligned}$$

According to Theorem A.1 from [28], $(x, Z^N)_{N \geq 1}$ converges strongly in distribution towards (x, z) . Besides thanks to Lemma 2.4.3 and Corollary 17.10 from [78], we have that $(x \otimes I_M, 1 \otimes Y^M)_{M \geq 1}$ converges strongly in distribution towards $(x \otimes 1, 1 \otimes y)$. In Theorem 2.1.2, we are interested in the situation where $Z^{NM} = Z^N \otimes I_M$ or $Z^{NM} = I_N \otimes Y^M$. So, without loss of generality, we restrict ourselves to this kind of Z^{NM} . We know that $(x \otimes I_M, Z^{NM})$ converges strongly towards $(x \otimes 1, Z)$, but since the support of f_ε is disjoint from the spectrum of $PP^*(x \otimes 1, Z, Z^*)$, thanks to Proposition 2.2.2, for N large enough, $\tau_N \otimes \tau_M \left(f_\varepsilon(PP^*(x \otimes I_M, Z^{NM}, Z^{NM^*})) \right) = 0$ and therefore,

$$\mathbb{E} \left[\text{Tr}_{MN} \left(f_\varepsilon \left(PP^* \left(X^N \otimes I_M, Z^{NM}, Z^{NM^*} \right) \right) \right) \right] \leq C \varepsilon^{-6} \frac{M^3}{N}. \quad (2.52)$$

Hence, using Proposition 2.2.11, we deduce for N large enough,

$$\begin{aligned}
 & \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM^*}) \right\| \right] - \left\| PP^*(x \otimes I_M, Z, Z^*) \right\| \\
 & \leq \varepsilon + \int_\varepsilon^\infty \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM^*}) \right\| \geq \left\| PP^*(x \otimes I_M, Z, Z^*) \right\| + \alpha \right) d\alpha \\
 & \leq \varepsilon + \int_\varepsilon^K \mathbb{P} \left(\text{Tr}_{NM} \left(f_\alpha(PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM^*})) \right) \geq 1 \right) d\alpha + C e^{-N} \\
 & \leq \varepsilon + C' \varepsilon^{-5} \frac{M^3}{N}.
 \end{aligned}$$

Finally we get that,

$$\limsup_{N \rightarrow \infty} \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM^*}) \right\| \right] \leq \left\| PP^*(x \otimes I_M, Z, Z^*) \right\|.$$

Besides, we know thanks to Theorem 5.4.5 of [13] that if h is a continuous function taking positive values on $(\left\| PP^*(x \otimes 1, Z, Z^*) \right\| - \varepsilon, \infty)$ and taking value 0 elsewhere. Then

$$\frac{1}{MN} \text{Tr}_{MN} (h(PP^*(X^N \otimes I_M, Z, Z^*)))$$

converges almost surely towards $\tau_A \otimes_{\min} \tau_B (h(PP^*(x \otimes 1, Z, Z^*)))$. If this quantity is positive for any h , then for any $\varepsilon > 0$, for N large enough,

$$\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM^*}) \right\| \geq \left\| PP^*(x \otimes 1, Z, Z^*) \right\| - \varepsilon.$$

Since h is non-negative and the intersection of the support of h with the spectrum of $PP^*(x \otimes 1, Z, Z^*)$ is non-empty, we have that $h(PP^*(x \otimes 1, Z, Z^*)) \geq 0$ and is not 0. Besides, we know that the trace on the space where z is defined is faithful, and so is the trace on the \mathcal{C}^* -algebra generated by a single semicircular variable, hence by Theorem 2.2.3, so is $\tau_{\mathcal{A}}$. Thus, since both $\tau_{\mathcal{A}}$ and $\tau_{\mathcal{B}}$ are faithful, by Lemma 2.4.2, so is $\tau_{\mathcal{A} \otimes_{\min} \tau_{\mathcal{B}}}$ and $\tau_{\mathcal{A} \otimes_{\min} \tau_{\mathcal{B}}}(h(PP^*(x \otimes 1, Z, Z^*))) > 0$. As a consequence, almost surely,

$$\liminf_{N \rightarrow \infty} \left\| P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \geq \|P(x \otimes 1, Z, Z^*)\|. \quad (2.53)$$

Thanks to Fatou's Lemma, we deduce

$$\liminf_{N \rightarrow \infty} \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \right] \geq \|PP^*(x \otimes I_M, Z, Z^*)\|.$$

We conclude that

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \right] = \|PP^*(x \otimes I_M, Z, Z^*)\|. \quad (2.54)$$

Let us define the following objects,

$$\varepsilon_N = \left| \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| \right] - \|PP^*(x \otimes I_M, Z, Z^*)\| \right|,$$

$$K = \sup_{N, M \geq 0} K_P \left(\|Z^{NM}\| \right) + H_P \left(\|Z^{NM}\| \right).$$

K is finite thanks to the strong convergence of the families Z^N and Y^M . Then thanks to Proposition 2.4.6, we have that for any $\delta > 0$,

$$\begin{aligned} & \mathbb{P} \left(\left| \left\| P^*P(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| - \|PP^*(x \otimes I_M, Z, Z^*)\| \right| \geq \delta + Ke^{-N} + \varepsilon_N \right) \\ & \leq d e^{-2N} + e^{-\frac{\delta^2 N}{K}}. \end{aligned}$$

Since this is true for any $\delta > 0$, by Borel-Cantelli's Lemma, almost surely,

$$\lim_{N \rightarrow \infty} \left\| PP^*(X^N \otimes I_M, Z^{NM}, Z^{NM*}) \right\| = \|PP^*(x \otimes 1, Z, Z^*)\|.$$

We finally conclude thanks to the fact that for any y in a \mathcal{C}^* -algebra, $\|yy^*\| = \|y\|^2$.

2.4.4 Proof of Theorem 2.1.5

We first prove the following estimate that we use multiple times during the proofs.

Lemma 2.4.7. *Let g be a \mathcal{C}^∞ function which takes value 0 on $(-\infty, 1/2]$ and value 1 on $[1, \infty)$, and in $[0, 1]$ otherwise. We set $f_\varepsilon : t \mapsto g(\varepsilon^{-1}(t - \alpha))$ with $\alpha = \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\|$, then there exists a constant C such that for any $\varepsilon > 0$ and N ,*

$$\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} \left(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right] \leq C \frac{\varepsilon^{-4}}{N^2}.$$

Proof. To estimate the above expectation we use the same method as in the proof of Theorem 2.1.2 with a few refinements to have an optimal estimate with respect to ε . We set $f_\varepsilon^\kappa : t \mapsto g(\varepsilon^{-1}(t - \alpha))g(\varepsilon^{-1}(\kappa - t) + 1)$ with $\alpha = \|PP^*(x \otimes I_M, 1 \otimes Y_M)\|$ and $\kappa > \alpha$. Since g has compact support and is sufficiently smooth we can apply Theorem 2.3.1. Setting $h : t \mapsto g(t - \varepsilon^{-1}\alpha)g(\varepsilon^{-1}\kappa + 1 - t) = \hat{f}_\varepsilon^\kappa(\varepsilon t)$, we have

$$\begin{aligned} 2\pi \int y^4 |\hat{f}_\varepsilon^\kappa(y)| dy &= \int y^4 \left| \int g(\varepsilon^{-1}(t - \alpha))g(\varepsilon^{-1}(\kappa - t) + 1)e^{-iyt} dt \right| dy \\ &= \int y^4 \left| \int h(t)e^{-iy\varepsilon t} \varepsilon dt \right| dy \\ &= \varepsilon^{-4} \int y^4 \left| \int h(t)e^{-iyt} dt \right| dy \\ &\leq \varepsilon^{-4} \int \frac{1}{1 + y^2} dy \int (|h^{(4)}(t)| + |h^{(6)}(t)|) dt. \end{aligned}$$

The derivatives $h^{(4)}$ and $h^{(6)}$ are uniformly bounded independently of t or ε . Since the support of these functions is included in $[\varepsilon^{-1}\alpha, \varepsilon^{-1}\alpha + 1] \cup [\varepsilon^{-1}\kappa, \varepsilon^{-1}\kappa + 1]$, there is a universal constant C such that for any ε and κ ,

$$\int y^4 |\hat{f}_\varepsilon^\kappa(y)| dy \leq C\varepsilon^{-4}.$$

With similar computations we can find a constant C such that for any ε and κ ,

$$\int (|y| + y^4) |\hat{f}_\varepsilon^\kappa(y)| dy \leq C\varepsilon^{-4}. \quad (2.55)$$

Since the support of f_ε^κ is disjoint from the spectrum of $PP^*(x \otimes I_M, 1 \otimes Y^M)$, for any ε and N one have $\tau \otimes \tau_M \left(f_\varepsilon^\kappa(PP^*(x \otimes I_M, 1 \otimes Y^M)) \right) = 0$. Consequently thanks to Theorem 2.3.1, we have a constant C such that for any N , ε and κ ,

$$\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} \left(f_\varepsilon^\kappa(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right] \leq C \frac{\varepsilon^{-4}}{N^2}.$$

Then by the monotone convergence Theorem, we deduce

$$\mathbb{E} \left[\text{Tr}_{NM} \left(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right] = \lim_{\kappa \rightarrow \infty} \mathbb{E} \left[\text{Tr}_{NM} \left(f_\varepsilon^\kappa(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right].$$

Hence we have a constant C such that for any N and $\varepsilon > 0$,

$$\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} \left(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right] \leq C \frac{\varepsilon^{-4}}{N^2}.$$

□

We finally complete the proof of Theorem 2.1.5. One can view $X^N = (X_1^N, \dots, X_d^N)$ as the random vector of size dN^2 which consists of the union of $(\sqrt{N}X_{s,s}^{N,i})_{i,s}$, $(\sqrt{2N} \Re X_{s,r}^{N,i})_{s < r, i}$ and $(\sqrt{2N} \Im X_{s,r}^{N,i})_{s < r, i}$ which are indeed centered Gaussian random variable of variance 1 as stated in Definition 2.2.8. Thus we can apply the Poincaré inequality (see Proposition 2.2.9) to the function

$$\varphi : X^N \mapsto \frac{1}{MN} \operatorname{Tr}_{MN} \left(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right),$$

and we get

$$\operatorname{Var} \left(\varphi(X^N) \right) \leq \frac{1}{(MN)^2} \mathbb{E} [\| \nabla \varphi(X^N) \|_2^2]$$

Besides, as in the proof of Lemma 2.3.6, if $Q \in \mathcal{A}_{d,p+q}$,

$$\begin{aligned} & N \left\| \nabla \operatorname{Tr}_{MN} \left(Q(X^N \otimes I_M, I_N \otimes Y^M) \right) \right\|_2^2 \\ &= \sum_s \sum_{i,j} \operatorname{Tr}_{MN} \left(D_s Q E_{i,j} \otimes I_M \right) \operatorname{Tr}_{MN} \left(D_s Q E_{i,j} \otimes I_M \right)^*. \end{aligned}$$

Besides, if f_k is a polynomial with a single variable, then $D_s f_k(PP^*) = \partial_s(PP^*) \# f'_k(PP^*)$. Thus, taking f_k such that f'_k converges towards f'_ε for the sup norm on the spectrum of $PP^*(X^N \otimes I_M, I_N \otimes Y^M)$, we deduce that

$$\begin{aligned} \operatorname{Var} \left(\varphi(X^N) \right) &\leq \frac{1}{M^2 N^3} \sum_{s,i,j} \mathbb{E} \left[\operatorname{Tr}_{MN} \left(\partial_s(PP^*) \# f'_\varepsilon(PP^*) E_{i,j} \otimes I_M \right) \right. \\ &\quad \left. \times \operatorname{Tr}_{MN} \left(\partial_s(PP^*) \# f'_\varepsilon(PP^*) E_{i,j} \otimes I_M \right)^* \right]. \end{aligned}$$

Now with $A = \partial_s(PP^*) \# f'_\varepsilon(PP^*)$,

$$\begin{aligned} & \sum_{i,j} \operatorname{Tr}_{MN} \left(A E_{i,j} \otimes I_M \right) \operatorname{Tr}_{MN} \left(A E_{i,j} \otimes I_M \right)^* \\ &= \sum_{i,j,k,l} g_j^* \otimes e_k^* A g_i \otimes e_k g_i^* \otimes e_l^* A^* g_j \otimes f_l \\ &= \sum_{j,k,l} g_j^* (I_N \otimes e_k^* A I_N \otimes e_k I_N \otimes e_l^* A^* I_N \otimes e_l) g_j \\ &= \operatorname{Tr}_N (I_N \otimes \operatorname{Tr}_M(A) I_N \otimes \operatorname{Tr}_M(A^*)) \\ &= \operatorname{Tr}_N (I_N \otimes \operatorname{Tr}_M(A) (I_N \otimes \operatorname{Tr}_M(A))^*). \end{aligned}$$

So by contractivity of the conditional expectation over $\mathbb{M}_N(\mathbb{C}) \otimes I_M$, that is $I_N \otimes \frac{1}{M} \operatorname{Tr}_M$, we have

$$\sum_{i,j} \operatorname{Tr}_{MN} \left(A E_{i,j} \otimes I_M \right) \operatorname{Tr}_{MN} \left(A E_{i,j} \otimes I_M \right)^* \leq \operatorname{Tr}_{MN}(AA^*) M.$$

As a consequence, we find that

$$\operatorname{Var} \left(\varphi(X^N) \right) \leq \frac{1}{N^3 M} \sum_s \mathbb{E} \left[\operatorname{Tr}_{MN} \left(\partial_s(PP^*) \# f'_\varepsilon(PP^*) (\partial_s(PP^*) \# f'_\varepsilon(PP^*))^* \right) \right].$$

Besides, if U, V and W are monomials,

$$\begin{aligned} |\mathrm{Tr}_{MN}(U f'_\varepsilon(PP^*) V f'_\varepsilon(PP^*) W)| &\leq \sqrt{\mathrm{Tr}_{MN}(U f'_\varepsilon{}^2(PP^*) U^*) \mathrm{Tr}_{MN}(V f'_\varepsilon(PP^*) W W^* f'_\varepsilon(PP^*) V^*)} \\ &\leq \mathrm{Tr}_{MN}(f'_\varepsilon{}^2(PP^*)) \|U\| \|V\| \|W\|. \end{aligned}$$

Therefore there is a constant C depending only on P and $\sup_i \|Y_i^M\|$ such that

$$\mathrm{Var}(\varphi(X^N)) \leq \frac{C}{N^2} \mathbb{E} \left[\prod_s \left(\|X_s^N\|^{2 \deg P} + 1 \right) \frac{1}{MN} \mathrm{Tr}_{NM} \left(\left| f'_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right|^2 \right) \right].$$

Thanks to Proposition 2.2.11, we can find w and α such that for any s and $u \geq 0$,

$$\mathbb{P}(\|X_s^N\| \geq w + u) \leq e^{-\alpha u N}.$$

There is a constant C independent of N and ε such that

$$\mathrm{Var}(\varphi(X^N)) \leq \frac{C}{N^2} \left(\mathbb{E} \left[\frac{1}{MN} \mathrm{Tr}_{NM} \left((f'_\varepsilon)^2(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \right] + \varepsilon^{-2} e^{-N} \right). \quad (2.56)$$

We can now apply Theorem 2.3.1 to the right hand side of the above equation, noticing that (2.55) still holds if we replace f'_ε by $(\varepsilon f'_\varepsilon)^2$. As a consequence, we find an inequality similar the one of Lemma 2.4.7 and thus a constant C such that for any N or ε ,

$$\mathrm{Var} \left(\frac{1}{MN} \mathrm{Tr}_{NM}(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M))) \right) \leq C \left(\frac{\varepsilon^{-6}}{N^4} + \varepsilon^{-2} e^{-N} \right).$$

Therefore, thanks to Lemma 2.4.7 there exists a constant C such that for any $N \in \mathbb{N}$ and ε such that $\varepsilon^4 > C \frac{M}{N}$,

$$\begin{aligned} &\mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \geq \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| + \varepsilon \right) \\ &\leq \mathbb{P} \left(\frac{1}{MN} \mathrm{Tr}_{NM} \left(f_\varepsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M)) \right) \geq \frac{1}{MN} \right) \\ &\leq \mathbb{P} \left(\left| \frac{1}{MN} \mathrm{Tr}_{NM} (f_\varepsilon(PP^*)) - \mathbb{E} \left[\frac{1}{MN} \mathrm{Tr}_{NM} (f_\varepsilon(PP^*)) \right] \right| \geq \frac{1}{MN} - \frac{C}{N^2 \varepsilon^4} \right) \\ &\leq C \left(\frac{\varepsilon^{-6}}{N^4} + \varepsilon^{-2} e^{-N} \right) \left(\frac{1}{MN} - \frac{C}{N^2 \varepsilon^4} \right)^{-2}. \end{aligned}$$

Let us now set $s = cN^{-1/4}$ with c a constant such that for any N ,

$$\frac{1}{MN} - \frac{C}{N^2 s^4} \geq \frac{1}{2MN}.$$

Therefore, if $x_+ = \max(x, 0)$, we have for some constant C ,

$$\begin{aligned}
& \mathbb{E} \left[\left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \right)_+ \right] \\
&= \int_{\mathbb{R}^+} \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \geq \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| + \varepsilon \right) d\varepsilon \\
&\leq s + 4CM^2N^2 \int_s^\infty \frac{\varepsilon^{-6}}{N^4} + \varepsilon^{-2}e^{-N} d\varepsilon \leq s + 4CM^2N^2(s^{-5}N^{-4} + s^{-1}e^{-N}) \\
&\leq CN^{-1/4}.
\end{aligned}$$

On one side, we have

$$\begin{aligned}
& \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right] \geq \delta + K_P \left(\left\| Y^M \right\| \right) e^{-N} \right) \\
&\geq \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \right. \\
&\quad \left. \geq \delta + K_P \left(\left\| Y^M \right\| \right) e^{-N} + \mathbb{E} \left[\left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \right)_+ \right] \right) \\
&\geq \mathbb{P} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \geq \frac{\delta + CN^{-1/4}}{\left\| P(x \otimes I_M, 1 \otimes Y^M) \right\|} \right).
\end{aligned}$$

On the other side, thanks to Proposition 2.4.6, we have

$$\begin{aligned}
& \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right] \geq \delta + K_P \left(\left\| Y^M \right\| \right) e^{-N} \right) \\
&\leq e^{-\frac{\delta^2}{HP(\|Y^M\|)}N} + de^{-2N}.
\end{aligned}$$

Hence we can find constants K and C such that for any $N \in \mathbb{N}$ and $\delta > 0$,

$$\mathbb{P} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \geq \delta + CN^{-1/4} \right) \leq e^{-K\delta^2N} + de^{-2N}.$$

And we get (2.4) by replacing δ by $N^{-1/4}\delta$.

The other inequality is trickier because we need to study the spectral measure of $PP^*(x \otimes I_M, 1 \otimes Y^M)$, which is far from easy. We mainly rely on the Theorem 1.1 from [74]. We summarize the part of this theorem which is interesting for us in the proposition below.

Proposition 2.4.8. *Let $x = (x_1, \dots, x_d)$ be a system of free semicircular variable, $p_{i,j} \in \mathbb{C}\langle X_1, \dots, X_d \rangle$ be such that $S = (p_{i,j}(x))_{i,j}$ is self-adjoint with spectral measure ρ with support K . Then there exists a finite subset $A \subset \mathbb{R}$ such that if I is a connected component of $\mathbb{R} \setminus A$, then either $\rho|_I = 0$, or $I \subset K$. In the second situation there exists an analytic function g defined for some $\delta > 0$ on*

$$W := \{z \in \mathbb{C} \mid |\Im z| < \delta\} \setminus \bigcup_{a \in A} \{a - it \mid t \in \mathbb{R}^+\}$$

such that for each $a \in A$, there exist $N \in \mathbb{N}$ and $\epsilon > 0$ such that $(z - a)^N g(z)$ admits an expansion on $W \cap \{z \in \mathbb{C} \mid |z - a| < \epsilon\}$ as a convergent powerseries in $r_N(z - a)$ where $r_N(z)$ is the analytic N^{th} -root of z defined with branch $C \setminus \{-it \mid t \in \mathbb{R}^+\}$. Then $\Im g_I$ is the probability density function of ρ_I .

What this means for us is that at the edge of the spectrum of $PP^*(x \otimes I_M, 1 \otimes Y^M)$, either we have an atom or the density of the spectral measure decays like $\frac{1}{|x-a|^r}$ with $r \in \mathbb{Q}$ when approaching a . Consequently we can find $\beta \geq 0$ such that if ρ is the spectral measure of $PP^*(x \otimes I_M, 1 \otimes Y^M)$ then for $\epsilon > 0$ small enough,

$$\rho \left(\left[\left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \epsilon, \infty \right] \right) \geq \epsilon^\beta.$$

Consequently if once again g is a \mathcal{C}^∞ function which takes value 0 on $(-\infty, 0]$, 1 on $[1/2, \infty)$, and belongs to $(0, 1]$ otherwise. We then take $f_\epsilon : t \mapsto g(\epsilon^{-1}(t - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| + \epsilon))$ for some $\epsilon \geq 0$. Then

$$\begin{aligned} \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \leq \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \epsilon \right) \\ = \mathbb{P} \left(\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*(X^N \otimes I_M, I_N \otimes Y^M))) = 0 \right) \\ \leq \mathbb{P} \left(\left| \frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) - \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) \right] \right| \geq \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) \right] \right) \\ \leq \frac{\text{Var} \left(\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) \right)}{\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) \right]^2}. \end{aligned}$$

Thanks to (2.56), we have

$$\begin{aligned} \text{Var} \left(\frac{1}{MN} \text{Tr}_N (f_\epsilon(PP^*)) \right) &\leq \frac{C}{N^2} \left(\mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} ((f'_\epsilon)^2(PP^*)) \right] + \epsilon^{-2} e^{-N} \right) \\ &\leq \frac{C}{N^2} (\|f'_\epsilon\|^2 + \epsilon^{-2}) \leq \frac{C'}{N^2} \epsilon^{-2}. \end{aligned}$$

On the contrary, with the same kind of computations which let us get Lemma 2.4.7, we can find constants C and K such that

$$\begin{aligned} \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} (f_\epsilon(PP^*)) \right] &\geq \tau \otimes \tau_M (f_\epsilon(PP^*)) - C \frac{\epsilon^{-4}}{N^2} \\ &\geq \rho \left(\left[\left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \epsilon/2, \infty \right] \right) - C \frac{\epsilon^{-4}}{N^2} \\ &\geq K \min(1, \epsilon)^\beta - C \frac{\epsilon^{-4}}{N^2}. \end{aligned}$$

Therefore we find finite constants C and K such that

$$\begin{aligned} \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \leq \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \epsilon \right) \\ \leq \frac{K}{N^2 \epsilon^2} \left(\min(1, \epsilon)^\beta - C \frac{\epsilon^{-4}}{N^2} \right)^{-2}. \end{aligned}$$

Now we fix $r = cN^{-1/(3+\beta)}$, with c constant such that for any N ,

$$\min(1, r)^\beta - \frac{C}{N^2 r^4} \geq \frac{\min(1, r)^\beta}{2}.$$

Then, we have

$$\begin{aligned} & \mathbb{E} \left[\left(\left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right)_+ \right] \\ &= \int_{\mathbb{R}^+} \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \leq \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \varepsilon \right) d\varepsilon \\ &\leq r + 4KN^{-2} \int_r^\infty \varepsilon^{-2} \min(1, \varepsilon)^{-2\beta} d\varepsilon \leq r + 4KN^{-2}(r^{-1-2\beta} + 1) \\ &\leq CN^{-1/(3+\beta)}. \end{aligned}$$

We deduce the following bound

$$\begin{aligned} & \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right] \leq -\delta - K_P \left(\left\| Y^M \right\| \right) e^{-N} \right) \\ &\geq \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \right. \\ &\quad \left. \leq -\delta - K_P \left(\left\| Y^M \right\| \right) e^{-N} - \mathbb{E} \left[\left(\left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| - \left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right)_+ \right] \right) \\ &\geq \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \leq -\delta - CN^{-1/(3+\beta)} \right). \end{aligned}$$

Since on the event $\{\forall i, \|X_i^N\| \leq D\}$ with D as in (2.51), we have

$$\begin{aligned} & \left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \\ &\leq \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right) \left(J_P \left(\left\| Y^M \right\| \right) + \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \right), \end{aligned}$$

we deduce that

$$\begin{aligned} & \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right] \leq -\delta - K_P \left(\left\| Y^M \right\| \right) e^{-N} \right) \\ &\geq \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| PP^*(x \otimes I_M, 1 \otimes Y^M) \right\| \leq -\delta - CN^{-1/(3+\beta)} \right. \\ &\quad \left. \text{and } \forall i, \|X_i^N\| \leq D \right) \\ &\geq \mathbb{P} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \leq \frac{-\delta - CN^{-1/(3+\beta)}}{J_P \left(\left\| Y^M \right\| \right) + \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\|} \right) \\ &\quad - \mathbb{P}(\exists i, \|X_i^N\| \geq D) \\ &\geq \mathbb{P} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \leq \frac{-\delta - CN^{-1/(3+\beta)}}{J_P \left(\left\| Y^M \right\| \right) + \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\|} \right) \\ &\quad - de^{-2N}. \end{aligned}$$

On the other side thanks to Proposition 2.4.6, we have

$$\begin{aligned} & \mathbb{P} \left(\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \mathbb{E} \left[\left\| PP^*(X^N \otimes I_M, I_N \otimes Y^M) \right\| \right] \leq -\delta - K_P \left(\left\| Y^M \right\| \right) e^{-N} \right) \\ & \leq d e^{-2N} + e^{-\frac{\delta^2 N}{H_P(\|Y^M\|)}}. \end{aligned}$$

Hence we can find constants K and C such that for any $N \in \mathbb{N}$ and $\delta > 0$,

$$\mathbb{P} \left(\left\| P(X^N \otimes I_M, I_N \otimes Y^M) \right\| - \left\| P(x \otimes I_M, 1 \otimes Y^M) \right\| \leq -\delta - CN^{-1/(3+\beta)} \right) \leq e^{-K\delta^2 N} + 2d e^{-2N}.$$

And we get (2.5) by replacing δ by $N^{-1/(3+\beta)}\delta$.

Acknowledgements

B. C. was partially funded by JSPS KAKENHI 17K18734, 17H04823, 15KK0162. F. P. benefited also from the aforementioned Kakenhi grants and a MEXT JASSO fellowship. A. G. and F. P. were partially supported by Labex Milyon (ANR-10-LABX-0070) of Université de Lyon. The authors would like to thank Narutaka Ozawa for supplying reference [75] for Lemma 2.4.2.

On the operator norm of noncommutative polynomials in deterministic matrices and iid Haar unitary matrices

Let $U^N = (U_1^N, \dots, U_p^N)$ be a p -tuple of $N \times N$ independent Haar unitary matrices and Z^{NM} be any family of deterministic matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$. Let P be a self-adjoint non-commutative polynomial. In [14], Voiculescu showed that the empirical measure of the eigenvalues of this polynomial evaluated in Haar unitary matrices and deterministic matrices converges towards a deterministic measure defined thanks to free probability theory. Now, let f be a smooth function. The main technical result of this paper is a precise bound of the difference between the expectation of

$$\frac{1}{MN} \operatorname{Tr}_{\mathbb{M}_N(\mathbb{C})} \otimes \operatorname{Tr}_{\mathbb{M}_M(\mathbb{C})} \left(f(P(U^N \otimes I_M, Z^{NM})) \right),$$

and its limit when N goes to infinity. If f is seven times differentiable, we show that it is bounded by $M^2 \|f\|_{\mathcal{C}^6} \ln^2(N) \times N^{-2}$. As a corollary we obtain a new proof with quantitative bounds of a result of Collins and Male which gives sufficient conditions for the operator norm of a polynomial evaluated in Haar unitary matrices and deterministic matrices to converge almost surely towards its free limit. Our result also holds in much greater generality. For instance, it allows to prove that if U^N and Y^{M_N} are independent and $M_N = o(N^{1/3} \ln^{-2/3}(N))$, then the norm of any polynomial in $(U^N \otimes I_{M_N}, I_N \otimes Y^{M_N})$ converges almost surely towards its free limit. Previous results required that $M = M_N$ is constant.

This chapter is adapted from [8].

3.1 Introduction

Understanding the behaviour of random matrices in large dimension is the core of random matrix theory. In the early nineties Voiculescu showed that one could get very accurate results with the help of non-commutative probability theory. This theory is equipped with a notion of freeness, analogous to independence in classical probability theory, which often describes accurately the asymptotic behaviour of random matrices. In [3] he studied the asymptotic behaviour of independent matrices taken from the Gaussian Unitary Ensemble (GUE). In a later paper he proved a similar theorem for Haar unitary matrices, which are random matrices whose law is the Haar measure on the unitary group \mathbb{U}_N . In a nutshell, Voiculescu proved in [14] that given U_1^N, \dots, U_p^N independent Haar unitary matrices, the renormalized trace of a polynomial P evaluated in these matrices converges towards a deterministic limit $\alpha(P)$. Specifically, the following holds true almost surely:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{Tr}_N \left(P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*}) \right) = \alpha(P). \quad (3.1)$$

Voiculescu computed the limit $\alpha(P)$ with the help of free probability. To give more detail, let B_N be a self-adjoint matrix of size N , then one can define the empirical measure of its (real) eigenvalues by

$$\mu_{B_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i},$$

where δ_λ is the Dirac mass in λ and $\lambda_1, \dots, \lambda_N$ are the eigenvalue of B_N . In particular, if P is a self-adjoint polynomial, that is such that for any matrices $A_1, \dots, A_d, P(A_1, \dots, A_d, A_1^*, \dots, A_d^*)$ is a self-adjoint matrix, then one can define the random measure $\mu_{P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*})}$. In this case, Voiculescu's result (3.1) implies that there exists a measure μ_P with compact support such that almost surely $\mu_{P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*})}$ converges weakly towards μ_P : its moments are given by $\mu_P(x^k) = \alpha(P^k)$ for all integer numbers k .

However, the convergence of the empirical measure of the eigenvalues of a matrix does not say much about the local properties of its spectrum, in particular about the convergence of the norm of this matrix, or the local fluctuations of its spectrum. For a comprehensive survey of important milestones related to these questions, we refer to the introduction of our previous paper [7]. In a nutshell, when dealing with a single matrix, incredibly precise results are known. Typically, concerning the GUE, very precise results were obtained by Tracy and Widom in the early nineties in [17]. On the other hand, there are much less results available when one deals with a polynomial in several random matrices. One of the most notable result was found by Haagerup and Thorbjørnsen in 2005 in [6]: they proved the almost sure convergence of the norm of a polynomial evaluated in independent GUE matrices. Equivalently, for P a self-adjoint polynomial, they proved that almost surely, for N large enough,

$$\sigma \left(P(X_1^N, \dots, X_d^N) \right) \subset \operatorname{Supp} \mu_P + (-\varepsilon, \varepsilon), \quad (3.2)$$

where $\sigma(H)$ is the spectrum of H and $\operatorname{Supp} \mu_P$ the support of the measure μ_P . The result (3.2) was a major breakthrough in the context of free probability and was refined in multiple ways, see [25, 26, 27, 29, 31, 7]. Those results all have in common that the basic random matrix is always self-adjoint. Much less is known in the non self-adjoint case. However Collins and

Male proved in [9] the same result as in [28] but with unitary Haar matrices instead of GUE matrices by using Male's former paper. With the exception of [9] and [7], all of these results are essentially based on the method introduced by Haagerup and Thorbjørnsen who relies on the so-called linearization trick. The main idea of this tool is that given a polynomial P , the spectrum of $P(X_1^N, \dots, X_d^N)$ is closely related to the spectrum of

$$L_N = a_0 \otimes I_N + \sum_{i=1}^d a_i \otimes X_i^N,$$

where a_0, \dots, a_d are matrices of size k depending only on P . Thus we trade a polynomial of degree d with coefficient in \mathbb{C} by a polynomial of degree 1 with coefficient in $\mathbb{M}_{k(d)}(\mathbb{C})$. In [9], the main idea was to view Haar unitary matrices as a random function of a GUE random matrix. Then the authors showed that almost surely this random function converges uniformly and they concluded by using the main result of [28]. An issue of this method is that it does not give any quantitative estimate. An important aim of this paper is to remedy to this problem. Our approach requires neither the linearization trick, nor the study of the Stieljes transform and attacks the problem directly without using previous results about the strong convergence of GUE random matrices. In this sense the proof is more direct and less algebraic. We will apply it to a generalization of Haar unitary matrices by tackling the case of Haar unitary matrices tensorized with deterministic matrices.

A usual strategy to study outliers, that are the eigenvalues going away from the spectrum, is to study the *non-renormalized* trace of smooth non-polynomial functions evaluated in independent Haar matrices i.e. if P is self-adjoint:

$$\mathrm{Tr}_N \left(f \left(P \left(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*} \right) \right) \right).$$

This strategy was also used by Haagerup, Thorbjørnsen and Male. Indeed it is easy to see that if f is a function which takes value 0 on $(-\infty, C - \varepsilon]$, 1 on $[C, \infty)$ and in $[0, 1]$ elsewhere, then with $\lambda_1(X)$ the largest eigenvalue of X ,

$$\mathbb{P} \left(\lambda_1(P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*})) \geq C \right) \leq \mathbb{P} \left(\mathrm{Tr}_N \left(f(P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*})) \right) \geq 1 \right).$$

Hence, if we can prove that $\mathrm{Tr}_N \left(f(P(U_1^N, \dots, U_p^N, U_1^{N*}, \dots, U_p^{N*})) \right)$ converges towards 0 in probability, this would already yield expected results. The above is just a well-known example, but one can get much more out of this strategy. Therefore, we need to study the non-renormalized trace. The case where f is a polynomial function has already been studied a long time ago, starting with the pioneering works [70, 71], and later formalized by the concept of second order freeness [72, 79]. However here we have to deal with a function f which is at best C^∞ . This makes things considerably more difficult and forces us to adopt a completely different approach. The main result is the following theorem (for the notations, we refer to Section 3.2 – for now, let us specify that $\frac{1}{N} \mathrm{Tr}_N$ denotes the usual renormalized trace on $N \times N$ matrices whereas τ denotes its free limit):

Theorem 3.1.1. *We define*

- $u = (u_1, \dots, u_p, u_1^*, \dots, u_p^*)$ a family of p free Haar unitaries and their adjoints,
- $U^N = (U_1^N, \dots, U_p^N, (U_1^N)^*, \dots, (U_p^N)^*)$ i.i.d. Haar unitary matrices of size N , and their adjoints.

- $Z^{NM} = (Z_1^{NM}, \dots, Z_q^{NM}) \in \mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$ deterministic matrices and their adjoint,
- P a self-adjoint polynomial,
- $f : \mathbb{R} \mapsto \mathbb{R}$ a function at least six times differentiable.

Then there exists a polynomial $L_P \in \mathbb{R}^+[X]$ which only depends on P such that for any N, M ,

$$\left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr}_{MN} \left(f \left(P \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] - \tau_N \otimes \tau_M \left(f \left(P \left(u \otimes I_M, Z^{NM} \right) \right) \right) \right| \leq \frac{\ln^2(N)M^2}{N^2} L_P \left(\|Z^{NM}\| \right) \times \|f\|_{C^6}.$$

where $\|Z^{NM}\| = \sup_{1 \leq i \leq q} \|Z_i^{NM}\|$ and $\|f\|_{C^k}$ is the sum of the supremum on \mathbb{R} of the first k derivatives. If $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, then we have the same inequality without the M^2 .

This theorem is a consequence of the slightly sharper, but less explicit, Theorem 3.4.1. Those two theorems are essentially the same, but in Theorem 3.4.1, instead of having the norm C^6 of f , we have the fourth moment of the Fourier transform of f . The above theorem calls for a few remarks.

- We assumed that the matrices Z^{NM} were deterministic, but thanks to Fubini's theorem we can assume that they are random matrices as long as they are independent from U^N . In this situation though, $L_P \left(\|Z^{NM}\| \right)$ in the right side of the inequality is a random variable (and thus we need some additional assumptions on the law of Z^{NM} if we want its expectation to be finite for instance).
- In Theorems 3.1.1 and 3.4.1 we have $U^N \otimes I_M$ and $u \otimes I_M$, however it is very easy to replace them by $U^N \otimes Y^M$ and $u \otimes Y^M$ for some matrices $Y_i^M \in M_M(\mathbb{C})$. Indeed we just need to apply Theorem 3.1.1 or 3.4.1 with $Z^{NM} = I_N \otimes Y^M$. Besides, in this situation, $L_P \left(\|Z^{NM}\| \right) = L_P \left(\|Y^M\| \right)$ does not depend on N . What this means is that if we have a matrix whose coefficients are polynomial in U^N , and that we replace U^N by u , we only change the spectra of this matrix by M^2N^{-2} in average.
- In the specific case where $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and the Y_i^M commute, as we stated in Theorem 3.1.1, we have the same inequality without the M^2 . Lowering the exponent in all generality would yield a direct improvement to Theorem 3.1.2 but we currently do not know whether it is actually possible. A lead to do so would be to prove a sharper version of Lemma 3.3.1. While this seems unrealistic for deterministic matrices, it might be possible to get some results when the matrices Y_i^M are random.

A detailed overview of the proof is given in Subsection 3.4.1. Similarly to [7], we interpolate Haar unitary matrices and free Haar unitaries with the help of a free Ornstein-Uhlenbeck process on the unitary group, i.e. the free unitary Brownian motion. For a reference, see Definition 3.2.11. However in [7] this idea was only to understand the intuition of the proof. In this paper the computations involved were quite different, indeed since we were considering the usual free

Ornstein-Uhlenbeck process, we could use a computation trick to replace this process by a well-chosen interpolation between GUE matrices and free semicirculars. This means that we did not need to use free stochastic calculus. There is no such trick for the free unitary Brownian motion, hence the computations use much more advanced tools.

When using this process, the Schwinger-Dyson equations, which can be seen as an integration by part, appear in the computation. For more information about these equations we refer to [80] to find numerous applications. In the specific case of the unitary group it is worth checking the proof of Theorem 5.4.10 from [13]. Even though those equations only come into play in the proof of Lemma 3.4.6, they play a major role in the proof since we could get a theorem similar to Theorem 3.1.1 for any random matrices which satisfies those equations.

Theorem 3.1.1 is the crux of the paper and allows us to deduce many corollaries. Firstly we get the following result. The first statement is basically Theorem 1.4 from [9]. The second one is entirely new and let us tensorize by matrices whose size goes to infinity when until now we could only work with tensor of finite size. This theorem is about strong convergence of random matrices, that is the convergence of the norm of polynomials in these matrices, see Definition 3.2.1.

Theorem 3.1.2. *Let the following objects be given:*

- $U^N = (U_1^N, \dots, U_p^N)$ independent unitary Haar matrices of size N ,
- $u = (u_1, \dots, u_p)$ a system of free Haar unitaries,
- $Y^M = (Y_1^M, \dots, Y_q^M)$ random matrices of size M , which almost surely, as M goes to infinity, converges strongly in distribution towards a q -tuple y of non-commutative random variables in a C^* -probability space \mathcal{B} with a faithful trace $\tau_{\mathcal{B}}$.
- $Z^N = (Z_1^N, \dots, Z_q^N)$ random matrices of size N , which almost surely, as N goes to infinity, converges strongly in distribution towards a q -tuple z of non-commutative random variables in a C^* -probability space with a faithful trace,

then the following holds true.

- If U^N and Z^N are independent, almost surely, (U^N, Z^N) converges strongly in distribution towards $\mathcal{F} = (u, z)$, where \mathcal{F} belongs to a C^* -probability space $(\mathcal{A}, *, \tau_{\mathcal{A}}, \|\cdot\|)$ in which u and z are free.
- If $(M_N)_{N \geq 0}$ is a sequence of integers such that $M_N = o(N^{1/3} \ln^{-2/3}(N))$, U^N and Y^{M_N} are independent, then almost surely $(U^N \otimes I_{M_N}, I_N \otimes Y^{M_N})$ converges strongly in distribution towards $\mathcal{F} = (u \otimes 1, 1 \otimes y)$ when N goes to infinity. The family \mathcal{F} thus belongs to $\mathcal{A} \otimes_{\min} \mathcal{B}$ (see Definition 3.2.6). Besides if the matrices Y^{M_N} commute, then we can weaken the assumption on M_N by only assuming that $M_N = o(N \ln^{-2}(N))$.

Understanding the Stieljes transform of a matrix gives a lot of information about its spectrum. This was actually a very important point in the proof of Haagerup and Thorbjørnsen's theorem. Our proof does not use this tool, however our final result, Theorem 3.4.1, allows us to deduce the following estimate. Being given a self-adjoint $NM \times NM$ matrix, we denote by G_A its Stieltjes transform:

$$G_A(z) = \frac{1}{NM} \operatorname{Tr}_{NM} \left(\frac{1}{z - A} \right).$$

This definition extends to the tensor product of free Haar unitaries with deterministic matrices by replacing $(NM)^{-1} \text{Tr}_{NM}$ by $\tau_N \otimes \tau_M$.

Corollary 3.1.3. *Given*

- $u = (u_1, \dots, u_p, u_1^*, \dots, u_p^*)$ a family of p free Haar unitaries and their adjoints,
- $U^N = (U_1^N, \dots, U_p^N, (U_1^N)^*, \dots, (U_p^N)^*)$ i.i.d. Haar unitary matrices of size N , and their adjoints.
- $Y^M = (Y_1^M, \dots, Y_q^M, Y_1^{M*}, \dots, Y_q^{M*})$ deterministic matrices of size M and their adjoints,
- P a self-adjoint polynomial,

there exists a polynomial L_P such that for every Y^M , $z \in \mathbb{C} \setminus \mathbb{R}$, $M, N \in \mathbb{N}$,

$$\left| \mathbb{E} \left[G_{P(U^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(u \otimes I_M, 1 \otimes Y^M)}(z) \right| \leq L_P \left(\|Y^M\| \right) \frac{M^2 \ln^2(N)}{N^2} \left(\frac{1}{|\Im(z)|^5} + \frac{1}{|\Im(z)|^2} \right).$$

where $\|Y^M\| = \sup_{1 \leq i \leq q} \|Y_i^M\|$.

One of the limitation of Theorem 3.1.1 is that we need to pick f regular enough. Actually by approximating f , we can afford to take f less regular at the cost of a slower speed of convergence. In other words, we trade some degree of regularity on f for a smaller exponent in N . The best that we can achieve is to take f Lipschitz. Thus it makes sense to introduce the Lipschitz-bounded metric which is the standard metric to metrize the topology of the weak convergence of probability measures on \mathbb{R} . Let \mathcal{F}_{LU} be the set of Lipschitz function from \mathbb{R} to \mathbb{R} , uniformly bounded by 1 and with Lipschitz constant at most 1, then

$$d_{LU}(\mu, \nu) = \sup_{f \in \mathcal{F}_{LU}} \left| \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \right|.$$

This metric is a slight weakening of the Wasserstein-1 distance which is defined similarly but without the assumption of boundedness on the functions f . For more information about this metric we refer to Appendix C.2 of [13]. In this paper, we get the following result:

Corollary 3.1.4. *Under the same notations as in Corollary 3.1.3, there exists a polynomial L_P such that for every matrices Y^M and $M, N \in \mathbb{N}$,*

$$d_{LU} \left(\mathbb{E}[\mu_{P(U^N \otimes I_M, I_N \otimes Y^M)}], \mu_{P(u \otimes I_M, 1 \otimes Y^M)} \right) \leq L_P \left(\|Y^M\| \right) M^2 \left(\frac{\ln N}{N} \right)^{1/3}.$$

This paper is organized as follows. In section 3.2 we give many usual definitions and notations in free probability, commutative and non-commutative stochastic calculus. Section 3.3 contains the proof of many important properties which we will need later on. Section 3.4 contains the proof of Theorem 3.1.1. Finally in section 3.5 we prove all of the corollaries.

3.2 Framework and standard properties

3.2.1 Usual definitions in free probability

In order to be self-contained, we begin by reminding the following definitions of free probability.

Definition 3.2.1. • A \mathcal{C}^* -probability space $(\mathcal{A}, *, \tau, \|\cdot\|)$ is a unital \mathcal{C}^* -algebra $(\mathcal{A}, *, \|\cdot\|)$ endowed with a state τ , i.e. a linear map $\tau: \mathcal{A} \rightarrow \mathbb{C}$ satisfying $\tau(1_{\mathcal{A}}) = 1$ and $\tau(a^*a) \geq 0$ for all $a \in \mathcal{A}$. In this paper we always assume that τ is a trace, i.e. that it satisfies $\tau(ab) = \tau(ba)$ for any $a, b \in \mathcal{A}$. An element of \mathcal{A} is called a (non commutative) random variable. We will always work with faithful trace, that is such that if $a \in \mathcal{A}$, $\tau(a^*a) = 0$ if and only if $a = 0$, in which case the norm is determined by τ thanks to the formula:

$$\|a\| = \lim_{k \rightarrow \infty} \left(\tau((a^*a)^{2k}) \right)^{1/2k}.$$

- Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be $*$ -subalgebras of \mathcal{A} , having the same unit as \mathcal{A} . They are said to be free if for all k , for all $a_i \in \mathcal{A}_{j_i}$ such that $j_1 \neq j_2, j_2 \neq j_3, \dots, j_{k-1} \neq j_k$:

$$\tau\left((a_1 - \tau(a_1))(a_2 - \tau(a_2)) \dots (a_k - \tau(a_k))\right) = 0.$$

Families of non-commutative random variable are said to be free if the $*$ -subalgebras they generate are free.

- Let $A = (a_1, \dots, a_k)$ be a k -tuple of non-commutative random variables. The joint distribution of the family A is the linear form $\mu_A: P \mapsto \tau[P(A, A^*)]$ on the set of polynomials in $2k$ non commutative indeterminates. By convergence in distribution, for a sequence of families of variables $(A_N)_{N \geq 1} = (a_1^N, \dots, a_k^N)_{N \geq 1}$ in \mathcal{C}^* -algebras $(\mathcal{A}_N, *, \tau_N, \|\cdot\|)$, we mean the pointwise convergence of the map

$$\mu_{A_N}: P \mapsto \tau_N[P(A_N, A_N^*)],$$

and by strong convergence in distribution, we mean convergence in distribution, and pointwise convergence of the map

$$P \mapsto \|P(A_N, A_N^*)\|.$$

- A non commutative random variable u is called a Haar unitary when it is unitary, that is $uu^* = u^*u = 1_{\mathcal{A}}$, and for all $n \in \mathbb{Z}$, one has

$$\tau(u^n) = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{else.} \end{cases}$$

The strong convergence of non-commutative random variable is actually equivalent to the convergence of its spectrum for the Hausdorff distance. More precisely we have the following proposition whose proof can be found in [9] (see Proposition 2.1):

Proposition 3.2.2. Let $\mathbf{x}_N = (x_1^N, \dots, x_p^N)$ and $\mathbf{x} = (x_1, \dots, x_p)$ be p -tuples of variables in \mathcal{C}^* -probability spaces, $(\mathcal{A}_N, *, \tau_N, \|\cdot\|)$ and $(\mathcal{A}, *, \tau, \|\cdot\|)$, with faithful states. Then, the following assertions are equivalent.

- \mathbf{x}_N converges strongly in distribution to \mathbf{x} .
- For any self-adjoint variable $h_N = P(\mathbf{x}_N, \mathbf{x}_N^*)$, where P is a fixed polynomial, μ_{h_N} converges in weak-* topology to μ_h where $h = P(\mathbf{x}, \mathbf{x}^*)$. Weak-* topology means relatively to continuous functions on \mathbb{C} . Moreover, the spectrum of h_N converges in Hausdorff distance to the spectrum of h , that is, for any $\varepsilon > 0$, there exists N_0 such that for any $N \geq N_0$,

$$\sigma(h_N) \subset \sigma(h) + (-\varepsilon, \varepsilon). \quad (3.3)$$

In particular, the strong convergence in distribution of a single self-adjoint variable is equivalent to its convergence in distribution together with the Hausdorff convergence of its spectrum.

It is important to note that thanks to Theorem 7.9 from [60], that we recall below, one can consider free copy of any random variable.

Theorem 3.2.3. *Let $(\mathcal{A}_i, \phi_i)_{i \in I}$ be a family of \mathcal{C}^* -probability spaces such that the functionals $\phi_i : \mathcal{A}_i \rightarrow \mathbb{C}$, $i \in I$, are faithful traces. Then there exist a \mathcal{C}^* -probability space (\mathcal{A}, ϕ) with ϕ a faithful trace, and a family of norm-preserving unital $*$ -homomorphism $W_i : \mathcal{A}_i \rightarrow \mathcal{A}$, $i \in I$, such that:*

- $\phi \circ W_i = \phi_i$, $\forall i \in I$.
- The unital \mathcal{C}^* -subalgebras $(W_i(\mathcal{A}_i))_{i \in I}$ form a free family in (\mathcal{A}, ϕ) .

3.2.2 Non-commutative polynomials and derivatives

We set $\mathbb{C}\langle Y_1, \dots, Y_d \rangle$ the set of non-commutative polynomials in d indeterminates and in particular we fix $\mathcal{P}_d = \mathbb{C}\langle Y_1, \dots, Y_{2d} \rangle$. Given a constant $A \in \mathbb{R}$, we can endow this vector space with the norm

$$\|P\|_A = \sum_{M \text{ monomial}} |c_M(P)| A^{\deg M}, \quad (3.4)$$

where $c_M(P)$ is the coefficient of P for the monomial M . In this subsection we define several maps on \mathcal{P}_d which we use multiple times in the rest of the paper, but first let us set a few notations. For A, B, C non-commutative polynomials,

$$(A \otimes B) \# C = ACB,$$

$$(A \otimes B) \tilde{\#} C = BCA,$$

$$m(A \otimes B) = BA.$$

We define an involution $*$ on \mathcal{P}_d by fixing for all $i \in [1, d]$, $(Y_i)^* = Y_{i+d}$, $(Y_{i+d})^* = Y_i$ and then extending it to \mathcal{P}_d with the formula $(\alpha PQ)^* = \bar{\alpha} Q^* P^*$. We then define the following maps.

Definition 3.2.4. If $1 \leq i \leq d$, one set $\partial_i : \mathcal{P}_d \longrightarrow \mathcal{P}_d \otimes \mathcal{P}_d$ such that for $P, Q \in \mathcal{P}_d$,

$$\begin{aligned}\partial_i(PQ) &= \partial_i P \times 1 \otimes Q + P \otimes 1 \times \partial_i Q, \\ \partial_i Y_j &= \mathbf{1}_{i=j} 1 \otimes 1.\end{aligned}$$

We also define $D_i : \mathcal{P}_d \longrightarrow \mathcal{P}_d$ by $D_i P = m \circ \partial_i P$. We similarly define ∂_i^* and D_i^* with the difference that for any j , $\partial_i^* Y_j = \mathbf{1}_{i+d=j} 1 \otimes 1$.

Because they satisfy the Leibniz's rule, the maps ∂_i and ∂_i^* are called non-commutative derivatives. It is related to Schwinger-Dyson equations on semicircular variable, for more information see [13], Lemma 5.4.7. While we do not use those equations in this paper, we use those associated with Haar unitary matrices. To do so, we define the following non-commutative derivative.

Definition 3.2.5. If $1 \leq i \leq d$, one set $\delta_i : \mathcal{P}_d \longrightarrow \mathcal{P}_d \otimes \mathcal{P}_d$ such that for $P, Q \in \mathcal{P}_d$,

$$\begin{aligned}\delta_i(PQ) &= \delta_i P \times 1 \otimes Q + P \otimes 1 \times \delta_i Q, \\ \forall j \in [1, d], \quad \delta_i Y_j &= \mathbf{1}_{i=j} Y_i \otimes 1, \quad \delta_i Y_{j+d} = -\mathbf{1}_{i=j} 1 \otimes Y_{i+d}.\end{aligned}$$

We also define $\mathcal{D}_i : \mathcal{P}_d \longrightarrow \mathcal{P}_d$ by $\mathcal{D}_i P = m \circ \delta_i P$.

We would like to apply the map δ_i to power series, more precisely the exponential of a polynomial, however since this is not well-defined in all generality we will need a few more definitions. Firstly, we need to define properly the operator norm of tensor of C^* -algebras. Since we use it later in this paper, we work with the minimal tensor product also named the spatial tensor product. For more information we refer to chapter 6 of [76].

Definition 3.2.6. Let \mathcal{A} and \mathcal{B} be C^* -algebra with faithful representations $(H_{\mathcal{A}}, \phi_{\mathcal{A}})$ and $(H_{\mathcal{B}}, \phi_{\mathcal{B}})$, then if \otimes_2 is the tensor product of Hilbert spaces, $\mathcal{A} \otimes_{\min} \mathcal{B}$ is the completion of the image of $\phi_{\mathcal{A}} \otimes \phi_{\mathcal{B}}$ in $B(H_{\mathcal{A}} \otimes_2 H_{\mathcal{B}})$ for the operator norm in this space. This definition is independent of the representations that we fixed.

Consequently if $P \in \mathcal{P}_d$, $z = (z_1, \dots, z_d)$ belongs to a C^* -algebra \mathcal{A} , then $(\delta_i P^k)(z, z^*)$ belongs to $\mathcal{A} \otimes_{\min} \mathcal{A}$, and $\|(\delta_i P^k)(z, z^*)\| \leq C_P k \|P(z, z^*)\|^{k-1}$ for some constant C_P independent of k . Thus we can define

$$(\delta_i e^P)(z, z^*) = \sum_{k \in \mathbb{N}} \frac{1}{k!} (\delta_i P^k)(z, z^*). \quad (3.5)$$

While we will not always be in this situation during this paper, it is important to note that if $\mathcal{A} = \mathbb{M}_N(\mathbb{C})$, then up to isomorphism $\mathcal{A} \otimes_{\min} \mathcal{A}$ is simply $\mathbb{M}_{N^2}(\mathbb{C})$ with the usual operator norm. Now we prove the following property whose proof is inspired of the one of Duhamel's formula which states that given two operators a and b ,

$$e^a - e^b = \int_0^1 e^{\alpha a} (a - b) e^{(1-\alpha)b} d\alpha. \quad (3.6)$$

Proposition 3.2.7. Let $P \in \mathcal{P}_d$, $z = (z_1, \dots, z_d)$ elements of a C^* -algebra \mathcal{A} , then

$$\left(\delta_i e^P \right) (z, z^*) = \int_0^1 \left(e^{\alpha P} \delta_i P e^{(1-\alpha)P} \right) (z, z^*) d\alpha,$$

with convention

$$A \times (B \otimes C) \times D = (AB) \otimes (CD).$$

Proof. One has,

$$\begin{aligned} \int_0^1 \left(e^{\alpha P} \delta_i P e^{(1-\alpha)P} \right) (z, z^*) d\alpha &= \sum_{n,m} \int_0^1 \frac{\alpha^n (1-\alpha)^m}{n!m!} d\alpha (P^n \delta_i P P^m) (z, z^*) \\ &= \sum_k \sum_{n+m=k} \int_0^1 \frac{\alpha^n (1-\alpha)^m}{n!m!} d\alpha (P^n \delta_i P P^m) (z, z^*). \end{aligned}$$

But if $m > 0$, by integration by part,

$$\int_0^1 \alpha^n (1-\alpha)^m d\alpha = \frac{m}{n+1} \int_0^1 \alpha^{n+1} (1-\alpha)^{m-1} d\alpha.$$

Thus for any n, m ,

$$\int_0^1 \frac{\alpha^n (1-\alpha)^m}{n!m!} d\alpha = \int_0^1 \frac{\alpha^{n+m}}{(n+m)!} d\alpha = \frac{1}{(m+n+1)!}.$$

Hence,

$$\int_0^1 \left(e^{\alpha P} \delta_i P e^{(1-\alpha)P} \right) (z, z^*) d\alpha = \sum_k \frac{1}{(k+1)!} \sum_{n+m=k} (P^n \delta_i P P^m) (z, z^*) = \left(\delta_i e^P \right) (z, z^*).$$

□

3.2.3 Free stochastic calculus

The main idea of this paper is to use an interpolation between Haar unitary matrices and their free limit. In order to do so, we will need some notion of free stochastic calculus. The main reference in this field is the paper [81] of Biane and Speicher to which we refer for most of the proofs in this subsection. For the sake of completeness we had to introduce notations and objects that we will not necessarily use outside of this subsection. For the reader not familiar with free probability, we would suggest to focus on understanding Theorem 3.2.10 and Definition 3.2.11.

From now on, (\mathcal{A}, τ) is a W^* -non-commutative probability space, that is \mathcal{A} is a von Neumann algebra, and τ is a faithful, normal (i.e. continuous for the ultraweak topology), tracial state on \mathcal{A} . We take \mathcal{A} filtered, that is there exists a family $(\mathcal{A}_t)_{t \in \mathbb{R}^+}$ of unital, weakly closed $*$ -subalgebras of \mathcal{A} , such that $\mathcal{A}_s \subset \mathcal{A}_t$ for all $s \leq t$. Besides we also assume that there exist p freely independent $(\mathcal{A}_t)_{t \in \mathbb{R}^+}$ -free Brownian motions $((S_t^i)_{t \in \mathbb{R}^+})_{1 \leq i \leq p}$. That is $t \mapsto S_t^i$ is weakly continuous, S_t^i is a self-adjoint element of \mathcal{A}_t with semi-circular distribution of mean 0 and variance t , and for all $s \leq t$, $S_t^i - S_s^i$ is free with \mathcal{A}_s , and has semi-circular distribution of mean 0 and variance $t - s$. Besides since the state τ is tracial, for any unital, weakly closed $*$ -subalgebra \mathcal{B} of \mathcal{A} , there exists a unique conditional expectation onto \mathcal{B} . We shall denote it by $\tau(\cdot | \mathcal{B})$. A map $t \in \mathbb{R}^+ \mapsto M_t \in \mathcal{A}$ will be called a martingale with respect to the filtration $(\mathcal{A}_t)_{t \in \mathbb{R}^+}$ if for every $s \leq t$ one has $\tau(M_t | \mathcal{A}_s) = M_s$.

We define the opposite algebra \mathcal{A}^{op} as the algebra \mathcal{A} endowed of the same addition, norm and involution, but with the product $a \times b = b \cdot a$ where \cdot is the product in \mathcal{A} . We can endow

\mathcal{A}^{op} with a faithful normal tracial state τ^{op} , which as a linear map on \mathcal{A} is actually τ . Similarly to the minimal tensor product, we will denote $L^\infty(\tau \otimes \tau^{\text{op}})$ the von Neuman algebra generated by $\mathcal{A} \otimes \mathcal{A}^{\text{op}}$ in $B(L^2(\mathcal{A}, \tau) \otimes_2 L^2(\mathcal{A}^{\text{op}}, \tau^{\text{op}}))$ where \otimes_2 is the usual tensor product for Hilbert spaces. Similarly to classical stochastic calculus, we now introduce piecewise constant maps.

Definition 3.2.8. *A simple biprocess is a piecewise constant map $t \mapsto U_t$ from \mathbb{R}^+ into the algebraic tensor product $\mathcal{A} \otimes \mathcal{A}^{\text{op}}$, such that $U_t = 0$ for t large enough. Besides it is called adapted if for any $t \geq 0$, $U_t \in \mathcal{A}_t \otimes \mathcal{A}_t$.*

The space of simple biprocesses form a complex vector space that we can endow with the norm

$$\|U\|_{\mathcal{B}^\infty}^2 = \int_0^\infty \|U_s\|_{L^\infty(\tau \otimes \tau^{\text{op}})}^2 ds. \quad (3.7)$$

We will denote by \mathcal{B}_a^∞ the completion of the vector space of adapted simple biprocesses for this norm. Now that we have defined the notion of simple process, we can define its stochastic integral that we will later extend to the space \mathcal{B}_a^∞ .

Definition 3.2.9. *Let $(S_t)_{t \geq 0}$ be a free Brownian motion, U be a simple adapted biprocess, we can find a decomposition $U = \sum_{j=1}^n A^j \otimes B^j$ and $0 = t_0 \leq t_1 \leq \dots \leq t_m$ such that for $t \in [t_i, t_{i+1})$, $A_t^j = A_{t_i}^j \in \mathcal{A}_{t_i}$ and $B_t^j = B_{t_i}^j \in \mathcal{A}_{t_i}^{\text{op}}$. We define its stochastic integral by*

$$\int_0^\infty U_s \# dS_s = \sum_{i=0}^{m-1} U_{t_i} \# (S_{t_{i+1}} - S_{t_i}) = \sum_{j=1}^n \sum_{i=0}^{m-1} A_{t_i}^j (S_{t_{i+1}} - S_{t_i}) B_{t_i}^j.$$

This definition is independent of the decomposition chosen. Besides $t \mapsto \int_0^t U_s \# dX_s$ is a martingale.

Thanks to Burkholder-Gundy inequality, that is Theorem 3.2.1 of [81], if we see the stochastic integral as a linear map from the space of adapted simple biprocesses endowed with the norm $\|\cdot\|_{\mathcal{B}^\infty}$ to \mathcal{A} , then this map is continuous. Hence we can extend it to \mathcal{B}_a^∞ and the martingale property remains true. Before talking about Itô's formula, as in the classical case, we need to introduce the quadratic variation. We will not develop the idea, but by studying random matrices, in the case of simple tensors, we are prompted to define

$$\langle\langle a \otimes b, c \otimes d \rangle\rangle = a \tau(bc) d.$$

We denote by $\#$ the product law in $\mathcal{A} \otimes \mathcal{A}^{\text{op}}$. If by contrast we want to use the usual product in $\mathcal{A} \otimes \mathcal{A}$, we will not put any sign. Let \dagger be the linear map such that on simple tensors, $(a \otimes b)^\dagger = b \otimes a$. In all generality for any $Z, Y \in \mathcal{A} \otimes \mathcal{A}^{\text{op}}$,

$$\langle\langle Z, Y \rangle\rangle = (\mathbf{1}_{\mathcal{A}} \otimes \tau^{\text{op}}) (Z \# (Y^\dagger)).$$

Since $\|\langle\langle Z, Y \rangle\rangle\| \leq \|Z\|_{L^\infty(\tau \otimes \tau^{\text{op}})} \|Y\|_{L^\infty(\tau \otimes \tau^{\text{op}})}$, we can extend this bilinear map to $Z, Y \in L^\infty(\tau \otimes \tau^{\text{op}})$. Besides by Cauchy-Schwarz, for $U, V \in \mathcal{B}_a^\infty$, $\langle\langle U, V \rangle\rangle$ is integrable.

Now that we have defined all of the necessary object to do stochastic calculus, we can state Itô's formula. We will need to handle polynomials in several processes, however Biane and Speicher only stated Itô's formula for a product of two processes, that is if $X_0, Y_0 \in \mathcal{A}$, $U^i, V^i \in \mathcal{B}_a^\infty$ and $K, L \in L^1(\mathbb{R}^+, \mathcal{A})$, we set

$$Y_t = Y_0 + \int_0^t K_s ds + \sum_{1 \leq i \leq p} \int_0^t U_s^i \# dS_s^i,$$

$$Z_t = Z_0 + \int_0^t L_s ds + \sum_{1 \leq i \leq p} \int_0^t V_s^i \# dS_s^i,$$

then for any $t \geq 0$,

$$\begin{aligned} Y_t Z_t &= Y_0 Z_0 + \int_0^t \left(Y_s L_s + K_s Z_s + \sum_{1 \leq i \leq p} \langle \langle U_s^i, V_s^i \rangle \rangle \right) ds \\ &\quad + \sum_{1 \leq i \leq p} \int_0^t \left((Y_s \otimes 1_{\mathcal{A}}) V_s^i + U_s^i (1_{\mathcal{A}} \otimes Z_s) \right) \# dS_s^i. \end{aligned} \quad (3.8)$$

To find a proof of (3.8), see Theorem 4.1.2 in [81]. While this theorem only proves the case where there is a single Brownian motion and $L = K = 0$, deducing equation (3.8) does not require much more work. We can then deduce from equation (3.8), the general Itô's formula. Even though this formula is used without a proof by Dabrowski in [82], we do not know of any satisfying reference. Hence we include a proof for self-containedness. Let us first fix a few notations.

- If $P \in \mathbb{C}\langle X_1, \dots, X_d \rangle$, $X \in (L^\infty(\mathbb{R}^+, \mathcal{A}))^d$ and $K \in (L^1(\mathbb{R}^+, \mathcal{A}))^d$, then

$$\partial P(X) \# K = \sum_{1 \leq j \leq d} \partial_j P(X) \# K_j.$$

- Similarly if $U \in (\mathcal{B}_a^\infty)^d$, then $\partial P(X) \# U = \sum_{1 \leq j \leq d} \partial_j P(X) \# U_j$.
- Finally if $U, V \in \mathcal{B}_a^\infty$, $A, B, C \in L^\infty(\mathbb{R}^+, \mathcal{A})$, then $(A \otimes B \otimes C) \# (U, V) = ((A \otimes B) \# U, (1 \otimes C) \# V)$.

Theorem 3.2.10. *Let $X_0 \in \mathcal{A}^d$, P be a non-commutative polynomial in d indeterminates, for any $t \geq 0$, $K \in (L^1([0, t], \mathcal{A}))^d$ and for every $i \in [1, p]$, $(\mathbf{1}_{s \leq t} U_s^i)_{s \in \mathbb{R}^+} \in (\mathcal{B}_a^\infty)^d$. With I the identity map on \mathcal{P}_d , we define*

$$X_t = X_0 + \int_0^t K_s ds + \sum_{1 \leq i \leq p} \int_0^t U_s^i \# dS_s^i,$$

$$\Delta_U(P)(X) = \sum_{1 \leq i \leq p} \sum_{1 \leq j, k \leq d} \langle \langle ((\partial_j \otimes I) \circ \partial_k P(X)) \# (U^{i,j}, U^{i,k}) \rangle \rangle.$$

Then for any $t \geq 0$, $\partial P(X) \# K$ and $\Delta_U(P)(X) \in L^1([0, t], \mathcal{A})$, and $(\mathbf{1}_{s \leq t} \partial P(X_s) \# U_s)_{s \in \mathbb{R}^+} \in \mathcal{B}_a^\infty$. Finally for any $t \geq 0$,

$$P(X_t) = P(X_0) + \int_0^t \partial P(X_s) \# K_s ds + \sum_{1 \leq i \leq p} \int_0^t (\partial P(X) \# U_s^i) \# dS_s^i + \int_0^t \Delta_U(P)(X_s) ds.$$

Proof. Thanks to Burkholder-Gundy inequality, that is Theorem 3.2.1 of [81], we know that

$$\sup_{0 \leq s \leq t} \|X_s^j\| \leq \|X_0^j\| + \|K_s^j\|_{L^1([0, t], \mathcal{A})} + \sum_{1 \leq i \leq p} \|U^{i,j} \mathbf{1}_{[0, t]}\|_{\mathcal{B}_a^\infty}.$$

Thus for any $t \in \mathbb{R}^+$, $(X_s)_{s \in [0,t]} \in L^\infty([0,t], \mathcal{A})^d$, hence for any polynomial P , $\partial P(X) \# K \in L^\infty([0,t], \mathcal{A})$, and $(\mathbf{1}_{s \leq t} \partial P(X_s) \# U_s)_{s \in \mathbb{R}^+} \in \mathcal{B}_a^\infty$. Given that

$$\|\langle \langle Z, Y \rangle \rangle\| \leq \|Z\|_{L^\infty(\tau \otimes \tau^{\circ p})} \|Y\|_{L^\infty(\tau \otimes \tau^{\circ p})},$$

we also have that $\Delta_U(P)(X) \in L^1([0,t], \mathcal{A})$. Finally to prove the formula, we proceed recurrently. If P is of degree 1, there is nothing to prove. For larger degree, by linearity we only need to deal with the case where P is a monomial. Thus we can write $P = QR$ with Q and R monomials of smaller degree for which the formula is verified. Thus thanks to equation (3.8), we have that

$$\begin{aligned} P(X_t) &= Q(X_0)R(X_0) + \int_0^t Q(X_s) \times (\partial R(X_s) \# K_s) + (\partial Q(X_s) \# K_s) \times R(X_s) ds \\ &\quad + \int_0^t Q(X_s) \Delta_U(R)(X_s) + \Delta_U(Q)(X_s) R(X_s) + \sum_{1 \leq i \leq p} \langle \langle \partial Q(X_s) \# U_s^i, \partial R(X_s) \# U_s^i \rangle \rangle ds \\ &\quad + \sum_{1 \leq i \leq p} \int_0^t \left((Q(X_s) \otimes \mathbf{1}_{\mathcal{A}}) \times (\partial R(X_s) \# U_s^i) + (\partial Q(X_s) \# U_s^i) \times (\mathbf{1}_{\mathcal{A}} \otimes R(X_s)) \right) \# dS_s^i. \end{aligned}$$

It is clear that,

$$\begin{aligned} \partial(QR)(X_s) \# K_s &= Q(X_s) \times (\partial R(X_s) \# K_s) + (\partial Q(X_s) \# K_s) \times R(X_s), \\ \partial(QR)(X_s) \# U_s^i &= (Q(X_s) \otimes \mathbf{1}_{\mathcal{A}}) \times (\partial R(X_s) \# U_s^i) + (\partial Q(X_s) \# U_s^i) \times (\mathbf{1}_{\mathcal{A}} \otimes R(X_s)). \end{aligned}$$

And finally,

$$\begin{aligned} \Delta_U(QR)(X) &= \sum_{1 \leq i \leq p} \sum_{1 \leq j, k \leq d} \langle \langle (\partial_j \otimes I) \circ \partial_k(QR)(X) \rangle \rangle \# (U^{i,j}, U^{i,k}) \\ &= \sum_{1 \leq i \leq p} \sum_{1 \leq j, k \leq d} \sum_{Q=AX_j B X_k C} \langle \langle (A(X) \otimes B(X) \otimes C(X)) \# (U^{i,j}, U^{i,k}) \rangle \rangle R(X) \\ &\quad + \sum_{R=AX_j B X_k C} Q(X) \langle \langle (A(X) \otimes B(X) \otimes C(X)) \# (U^{i,j}, U^{i,k}) \rangle \rangle \\ &\quad + \sum_{Q=AX_j B, R=C X_k D} \langle \langle (A(X) \otimes (BC)(X) \otimes D(X)) \# (U^{i,j}, U^{i,k}) \rangle \rangle \\ &= Q(X) \Delta_U(R)(X) + \Delta_U(Q)(X) R(X) + \sum_{1 \leq i \leq p} \langle \langle \partial Q(X) \# U^i, \partial R(X) \# U^i \rangle \rangle. \end{aligned}$$

□

Finally, one of the fundamental tool that we use in this paper is the free unitary Brownian motion, a good reference on the matter is [83]. In particular one can find a proof of its existence.

Definition 3.2.11. *Let $(S_t)_{t \geq 0}$ be a free Brownian motion adapted to a filtered W^* -probability space $(\mathcal{A}, (\mathcal{A}_t)_{t \geq 0}, \tau)$, the free unitary Brownian motion $(u_t)_{t \geq 0}$ is the unique solution to the equation*

$$\forall t \geq 0, \quad u_t = \mathbf{1}_{\mathcal{A}} - \int_0^t \frac{u_s}{2} ds + \mathbf{i} \int_0^t (u_s \otimes \mathbf{1}_{\mathcal{A}}) \# dS_s. \quad (3.9)$$

In particular, for any $t \geq 0$, u_t is unitary, that is $u_t u_t^* = u_t^* u_t = \mathbf{1}_{\mathcal{A}}$.

Although we do not use this notation in this paper, similarly to the classical case, it is usual to write equation (3.9) as

$$u_0 = \mathbf{1}_{\mathcal{A}}, \quad du_t = -\frac{u_t}{2} dt + \mathbf{i}(u_t \otimes \mathbf{1}_{\mathcal{A}}) \# dS_t.$$

3.2.4 Notations

Let us now fix a few notations concerning the spaces and traces that we use in this paper.

- Definition 3.2.12.**
- (\mathcal{A}_N, τ_N) is the free product of $\mathbb{M}_N(\mathbb{C})$ with the von Neuman algebra \mathcal{A} from the former subsection. To build \mathcal{A}_N we use Theorem 3.2.3 and we get a \mathcal{C}^* -probability space C with a faithful trace φ . Since we want (\mathcal{A}_N, τ_N) to be a von Neuman algebra, we set $L^2(C, \varphi)$ as the completion of C for the norm $a \mapsto \phi(a^*a)^{1/2}$, we have an injective \mathcal{C}^* -algebra morphism from C to $B(L^2(C, \varphi))$. We then proceed to take \mathcal{A}_N the closure of the image of C in this space for the weak topology. As for τ_N , since we can extend $(x, y) \mapsto \varphi(x^*y)$ to a scalar product $\langle \cdot, \cdot \rangle_\varphi$ on $L^2(C, \varphi)$, we set for $a \in B(L^2(C, \varphi))$, $\tau_N(a) = \langle a(1), 1 \rangle_\varphi$.
 - Note that when restricted to $\mathbb{M}_N(\mathbb{C})$, τ_N is just the regular renormalized trace on matrices. Similarly we will usually denote τ_M and τ_k the renormalized trace on $M_M(\mathbb{C})$ and $M_k(\mathbb{C})$. As in the former subsection, the restriction of τ_N to the \mathcal{C}^* -algebra \mathcal{A} is denoted as τ .
 - Tr_N is the non-renormalized trace on $\mathbb{M}_N(\mathbb{C})$.
 - $E_{i,j}$ is the matrix whose only non-zero coefficient is (i, j) and this coefficient has value 1, the size of the matrix $E_{i,j}$ will depend on the context.
 - In general we identify $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$ with $\mathbb{M}_{kN}(\mathbb{C})$ through the isomorphism $E_{i,j} \otimes E_{r,s} \mapsto E_{i+(r-1)N, j+(s-1)N}$, similarly we identify $\text{Tr}_N \otimes \text{Tr}_k$ with Tr_{kN} .
 - If $A^N = (A_1^N, \dots, A_d^N)$ and $B^M = (B_1^M, \dots, B_d^M)$ are two families of matrices, then we denote $A^N \otimes B^M = (A_1^N \otimes B_1^M, \dots, A_d^N \otimes B_d^M)$ and if $M = N$, $A^N B^N = (A_1^N B_1^N, \dots, A_d^N B_d^N)$. We typically use the notation $X^N \otimes I_M$ for the family $(X_1^N \otimes I_M, \dots, X_d^N \otimes I_M)$.
 - If $P \in \mathcal{P}_d$, in order to avoid cumbersome notations when evaluating P in (X, X^*) , instead of denoting $P(X, X^*)$ we will write $\tilde{P}(X)$.
 - We define $(e_i)_{i \in [1, M]}$, $(g_i)_{i \in [1, N]}$ and $(f_i)_{i \in [1, k]}$ the canonical basis of \mathbb{C}^M , \mathbb{C}^N and \mathbb{C}^k .

A polynomial $P \in \mathcal{P}_d$ is said to be self-adjoint if $P^* = P$. Self-adjoint polynomials have the property that if z_1, \dots, z_d are elements of a \mathcal{C}^* -algebra, then $P(z_1, \dots, z_d, z_1^*, \dots, z_d^*)$ is self-adjoint. Now that we have defined the notion of self-adjoint polynomial we give a property which justifies computations that we will do later on:

Proposition 3.2.13. *Let the following objects be given,*

- $u = (u_t^1, \dots, u_t^p)_{t \geq 0}$ a family of p freely independent free unitary Brownian motions,
- $f \in \mathcal{C}(\mathbb{R})$ the set of continuous function on \mathbb{R} ,
- P a self-adjoint polynomial.

Then with \mathbb{U}_N the group of unitary matrices of size N , the following map is measurable:

$$(U^N, Z^{NM}) \in \mathbb{U}_N^p \times \mathbb{M}_{NM}(\mathbb{C})^{d-p} \mapsto \tau_N \otimes \tau_M \left(f \left(\tilde{P} \left((U^N u_t) \otimes I_M, Z^{NM} \right) \right) \right).$$

For a full proof we refer to [7], Proposition 2.7. But in a few words, it is easy to see the measurability when f is a polynomial since then this map is also polynomial in the coefficient of U^N and Z^{NM} , and we conclude by density. Actually we could easily prove that this map is continuous, however we do not need it. The only reason we need this property is to justify that if U^N is a family of random matrices, then the random variable $\tau_N \otimes \tau_M \left(f \left(\tilde{P}(u_t^N \otimes I_M, Z^{NM}) \right) \right)$ is well-defined. To conclude this subsection we introduce different notations related to maps defined on tensor spaces.

Definition 3.2.14. Let $n : A \otimes B \in \mathbb{M}_M(\mathbb{C})^{\otimes 2} \mapsto AB \in \mathbb{M}_M(\mathbb{C})$, we define the linear map $(\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) : (\mathcal{A}_N \otimes \mathbb{M}_M(\mathbb{C}))^{\otimes 2} \rightarrow \mathbb{M}_M(\mathbb{C})$ as

$$(\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) = n \circ (\tau_N \otimes I_M)^{\otimes 2}.$$

We will also use the shorter notation $(\tau_N \otimes I_M) \otimes^2$.

3.2.5 Random matrix model

We conclude this section by giving the definition and a few properties on the models of random matrices that we will study.

Definition 3.2.15. A Haar unitary matrix of size N is a random matrix distributed according to the Haar measure on the group of unitary matrices of size N .

Definition 3.2.16. A Hermitian Brownian motion $(X_t^N)_{t \in \mathbb{R}^+}$ of size N is a self-adjoint matrix whose coefficients are random variables with the following laws:

- For $1 \leq i \leq N$, the random variables $\sqrt{N}((X_t^N)_{i,i})_{t \in \mathbb{R}^+}$ are independent Brownian motions.
- For $1 \leq i < j \leq N$, the random variables $(\sqrt{2N} \Re(X_t^N)_{i,j})_{t \in \mathbb{R}^+}$ and $(\sqrt{2N} \Im(X_t^N)_{i,j})_{t \in \mathbb{R}^+}$ are independent Brownian motions, independent of $\sqrt{N}((X_t^N)_{i,i})_{t \in \mathbb{R}^+}$.

To study the free unitary Brownian motion, we will need to study its finite dimensional equivalent, the unitary Brownian motion. Typically it is defined as the Markov process whose infinitesimal generator is the Laplacian operator on the unitary group. However given the upcoming computations in this paper, it is better to use an equivalent definition as the solution of a stochastic differential equation. We refer to subsection 2.1 of [84] for a short summary on the different definitions.

Definition 3.2.17. Let X^N be a Hermitian Brownian motion, then the unitary Brownian motion $(U_t^N)_{t \geq 0}$ is the solution of the following stochastic differential equation:

$$dU_t^N = \mathbf{i}U_t^N dX_t^N - \frac{1}{2}U_t^N dt, \quad U_0^N = I_N, \quad (3.10)$$

where we formally define $U_t^N dX_t^N$ by simply taking the matrix product

$$(U_t^N dX_t^N)_{i,j} = \sum_k (U_t^N)_{i,k} d(X_t^N)_{k,j}.$$

In particular, almost surely, for any t , U_t^N is a unitary matrix of size N .

The following property is typical of the kind of computation that we can do with unitary Brownian motion with classical stochastic calculus, see [84] for example.

Proposition 3.2.18. *Let U_1^N, \dots, U_p^N be independent unitary Brownian motions of size N , A_{p+1}^N, \dots, A_d^N be deterministic matrices, $Q \in \mathcal{P}_d$ be a monomial, we set Q_s the monomial evaluated in $(U_{1,s}^N, \dots, U_{p,s}^N, A_{p+1}^N, \dots, A_d^N)$ and their adjoints, $|Q|_B$ the degree of Q with respect to $(U_1, \dots, U_p, U_1^*, \dots, U_p^*)$. Then there exists a martingale J such that,*

$$\begin{aligned} d\mathrm{Tr}_N Q_s &= dJ_s - \frac{|Q|_B}{2} \mathrm{Tr}_N Q_s ds - \frac{1}{N} \sum_{i \leq p, Q=AU_i B U_i^* C} \mathrm{Tr}_N(A_s U_{i,s}^N C_s) \mathrm{Tr}_N(B_s U_{i,s}^N) ds \\ &\quad - \frac{1}{N} \sum_{i \leq p, Q=AU_i^* B U_i^* C} \mathrm{Tr}_N(A_s U_{i,s}^{N*} C_s) \mathrm{Tr}_N(B_s U_{i,s}^{N*}) ds \\ &\quad + \frac{1}{N} \sum_{i \leq p, Q=AU_i B U_i^* C} \mathrm{Tr}(A_s C_s) \mathrm{Tr}(B_s) ds \\ &\quad + \frac{1}{N} \sum_{i \leq p, Q=AU_i^* B U_i^* C} \mathrm{Tr}(A_s C_s) \mathrm{Tr}(B_s) ds. \end{aligned}$$

3.3 Preliminaries

3.3.1 A matricial inequality

We are indebted to Mikael de la Salle for supplying us with the proof of the following lemma.

Lemma 3.3.1 (de la Salle). *Let \mathcal{A} be a C^* -algebra, $A_1, A_2 \in \mathcal{A}$, $B_1, B_2 \in \mathbb{M}_M(\mathbb{C})$, as in subsection 3.2.3 we define $(A_1 \otimes B_1) \sharp (A_2 \otimes B_2) = (A_1 A_2) \otimes (B_2 B_1)$. Then if $x, y \in \mathcal{A} \otimes \mathbb{M}_M(\mathbb{C})$, with the operator norm in their respective space,*

$$\|x \sharp y\| \leq M \|x\| \|y\|.$$

Proof. We write $x = \sum_{1 \leq i, j \leq M} x_{i,j} \otimes E_{i,j}$, $y = \sum_{1 \leq k, l \leq M} y_{k,l} \otimes E_{k,l}$, then

$$x \sharp y = \sum_{i,j,k} x_{k,j} y_{i,k} \otimes E_{i,j}.$$

We define $A_k = \sum_{i,j} x_{k,j} y_{i,k} \otimes E_{i,j}$, $X_k = \sum_j x_{k,j} \otimes E_{k,j} \otimes I_M$, $Y_k = \sum_i y_{i,k} \otimes I_M \otimes E_{i,k}$. Then by using the fact that X_k and Y_k are band matrices, we have $\|X_k\| \leq \|x\|$ and $\|Y_k\| \leq \|y\|$. Besides $\|x \sharp y\| \leq \sum_{1 \leq k \leq M} \|A_k\|$. Finally we have for any k ,

$$\begin{aligned}
 \|X_k Y_k\|^2 &= \|X_k Y_k (X_k Y_k)^*\| \\
 &= \left\| \left(\sum_{i,j} x_{k,j} y_{i,k} \otimes E_{k,j} \otimes E_{i,k} \right) \left(\sum_{i,j} x_{k,j} y_{i,k} \otimes E_{k,j} \otimes E_{i,k} \right)^* \right\| \\
 &= \left\| \sum_{i,j,u,v} x_{k,j} y_{i,k} y_{u,k}^* x_{k,v}^* \otimes E_{k,j} E_{v,k} \otimes E_{i,k} E_{k,u} \right\| \\
 &= \left\| \sum_{i,j,u} x_{k,j} y_{i,k} y_{u,k}^* x_{k,v}^* \otimes E_{k,k} \otimes E_{i,j} E_{v,u} \right\| \\
 &= \|A_k A_k^* \otimes E_{k,k}\| \\
 &= \|A_k\|^2.
 \end{aligned}$$

Thus $\|x \sharp y\| \leq \sum_{1 \leq k \leq M} \|A_k\| \leq M \|x\| \|y\|$.

□

3.3.2 A Poincaré type equality

One of the main tool when dealing with GUE random matrices is the Poincaré inequality (see Definition 4.4.2 from [13]), which gives us a sharp majoration of the variance of a function in these matrices. Typically this inequality shows that the variance of the renormalized trace of a polynomial in GUE random matrices, which a priori is of order $\mathcal{O}(1)$, is of order $\mathcal{O}(N^{-2})$. In this paper we need a similar type of inequality but instead of working with independent GUE random matrices, we work with marginal of independent unitary Brownian motions at times t . We will follow the approach of [85], Proposition 6.1. I would like to thank one of the reviewer for pointing out that an alternative approach could be to use the more general Theorem 4.3 from [86]. This theorem deals with Brownian motions defined with the help of two parameters r and s , and the unitary Brownian motion matches with the case where $(r, s) = (1, 0)$. However the proof of Proposition 3.3.2 that we give below is a good introduction to the kind of computations that we deal with in this paper unlike the approach taken in [86] and needs less prerequisite to be understood.

Proposition 3.3.2. *Let $Q \in \mathcal{P}_d$, $(U_t^N)_{t \in \mathbb{R}^+}$, $(V_t^N)_{t \in \mathbb{R}^+}$, $(W_t^N)_{t \in \mathbb{R}^+}$ be independent families of p unitary Brownian motions of size N . Let A^N be a family of $d - p$ deterministic matrices, with notations as in Definition 3.2.12, one has for any $T \geq 0$,*

$$\text{Var} \left(\text{Tr}_N \left(\tilde{Q}(U_T^N, A^N) \right) \right) = \frac{1}{N} \sum_{1 \leq k \leq p} \int_0^T \mathbb{E} \left[\text{Tr}_N \left(\widetilde{\mathcal{D}}_k Q(V_{T-t}^N U_t^N, A^N) \times \widetilde{\mathcal{D}}_k Q(W_{T-t}^N U_t^N, A^N)^* \right) \right] dt.$$

Proof. To simplify notations, we will not write the index N in U_t^N, V_t^N, W_t^N and A^N . For $U \in \mathbb{M}_N(\mathbb{C})^p$, we set $f : (U, U^*) \mapsto \text{Tr}_N(Q(U, A, U^*, A^*))$. We can view f as a polynomial in the coefficients of the matrices U and their conjugates, since those are complex variables we use the notion of complex differential. That is if $g : (x, y) \in \mathbb{R}^2 \rightarrow g(x, y) \in \mathbb{C}$ is a differentiable function, we define $\partial_z g = \frac{1}{2}(\partial_x g - \mathbf{i} \partial_y g)$ and $\partial_{\bar{z}} g = \frac{1}{2}(\partial_x g + \mathbf{i} \partial_y g)$. If $u_k^{i,j}$ is the (i, j) -coefficient

of the k -th matrix in U , we denote the differential of f with respect to $u_k^{i,j}$ by $\partial_{u_k^{i,j}} f$, and the differential of f with respect to the conjuguate of this coefficient by $\partial_{u_k^{j,i}}^* f$. In particular,

$$\begin{aligned} \partial_{u_k^{i,j}}((U_k)_{i,j}) &= 1, & \partial_{u_k^{i,j}}^*((U_k^*)_{i,j}) &= 1, \\ \partial_{u_k^{i,j}}((U_k)_{a,b}) &= 0, & \partial_{u_k^{i,j}}^*((U_k^*)_{a,b}) &= 0, \text{ for all } (a,b) \neq (i,j), \\ \partial_{u_k^{i,j}}((U_k^*)_{a,b}) &= 0, & \partial_{u_k^{i,j}}((U_k)_{a,b}) &= 0, \text{ for any } (a,b). \end{aligned}$$

Next we introduce

$$M_t = P_{T-t} f(U_t, U_t^*),$$

where $P_{T-t} f(U, U^*) = \mathbb{E}_V[f(V_{T-t}U, (V_{T-t}U)^*)]$ with $(V_t)_{t \geq 0}$, p independent unitary Brownian motions of size N and E_V the expectation with respect to $(V_t)_{t \geq 0}$. We will follow the approach of [85], Proposition 6.1, and show that $(M_t)_{0 \leq t \leq T}$ is a martingale. It will follow that,

$$\begin{aligned} \text{Var} \left(\text{Tr}_N \left(\tilde{Q}(U_T^N, A^N) \right) \right) &= \mathbb{E}[|f(U_T, U_T^*)|^2 - |\mathbb{E}[f(U_T, U_T^*)]|^2] \\ &= \mathbb{E}[M_T \overline{M_T} - M_0 \overline{M_0}] \\ &= \mathbb{E} \left[\langle M_T, \overline{M_T} \rangle \right]. \end{aligned} \tag{3.11}$$

If we set $(X_t)_{t \geq 0}$, p independent Hermitian Brownian motions of size N , and $f_t = P_{T-t} f$, then

$$\begin{aligned} dM_t &= (\partial_t f_t)(U_t, U_t^*) dt + \sum_{1 \leq k \leq p, 1 \leq i, j \leq N} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) d(U_{k,t})_{i,j} + (\partial_{u_k^{j,i}}^* f_t)(U_t, U_t^*) d(U_{k,t}^*)_{i,j} \\ &\quad + \frac{1}{2} \sum_{\substack{1 \leq k \leq p, 1 \leq i, j, r, s \leq N \\ \varepsilon_1, \varepsilon_2 \in \{1, *\}}} (\partial_{u_k^{i,j}}^{\varepsilon_1} \partial_{u_k^{r,s}}^{\varepsilon_2} f_t)(U_t, U_t^*) d\langle (U_{k,t}^{\varepsilon_1})_{i,j}, (U_{k,t}^{\varepsilon_2})_{r,s} \rangle_t. \end{aligned}$$

By using equation (3.10), we can isolate the martingale term in the previous equation. As for the term associated to dt , as long as we show that $(M_t)_{t \geq 0}$ is a martingale it will be 0. To do so, we set $\mathcal{F}_t = \sigma((U_s)_{0 \leq s \leq t})$, then \mathcal{F}_t is a filtration adapted to $(M_t)_{t \geq 0}$. Besides if we set $N_t = \mathbb{E}[f(U_T, U_T^*) \mid \mathcal{F}_t]$, then since if we redefine $V_{T-t} = U_T U_t^*$, it is still a family of p independent unitary Brownian motions of size N , independent of \mathcal{F}_t . This implies that

$$N_t = \mathbb{E}[f(U_T, U_T^*) \mid \mathcal{F}_t] = \mathbb{E}[f(V_{T-t}U_t, (V_{T-t}U_t)^*) \mid \mathcal{F}_t] = \mathbb{E}_V[f(V_{T-t}U_t, (V_{T-t}U_t)^*)] = M_t.$$

Hence $(M_t)_{t \geq 0}$ is a martingale and

$$dM_t = \mathbf{i} \sum_{1 \leq k \leq p, 1 \leq i, j \leq N} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) (U_{k,t} dX_{k,t}^N)_{i,j} - (\partial_{u_k^{j,i}}^* f_t)(U_t, U_t^*) ((U_{k,t} dX_{k,t}^N)^*)_{i,j}, \tag{3.12}$$

Therefore, as we saw in equation (3.11), we are left with computing the bracket of M_t . To begin with we have,

$$\begin{aligned} \langle (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{(U_{k,t} dX_{k,t}^N)_{r,s}} \rangle &= \mathbf{1}_{i=r, j=s} \frac{dt}{N}, \\ \langle ((U_{k,t} dX_{k,t}^N)^*)_{i,j}, \overline{((U_{k,t} dX_{k,t}^N)^*)_{r,s}} \rangle &= \mathbf{1}_{i=r, j=s} \frac{dt}{N}, \end{aligned}$$

$$\langle (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{((U_{k,t} dX_{k,t}^N)^*)_{r,s}} \rangle = (U_{k,t})_{i,r} (U_{k,t})_{s,j} \frac{dt}{N},$$

$$\langle ((U_{k,t} dX_{k,t}^N)^*)_{i,j}, \overline{(U_{k,t} dX_{k,t}^N)_{r,s}} \rangle = (U_{k,t}^*)_{s,j} (U_{k,t}^*)_{i,r} \frac{dt}{N}.$$

Let us remind that $f : (U, U^*) \mapsto \text{Tr}_N(Q(U, A, U^*, A^*))$, thus one has

$$(\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) = \mathbb{E}_V \left[\text{Tr}_N(\widetilde{D}_k Q(V_{T-t} U_t, A) V_{k,T-t} E_{i,j}) \right],$$

$$(\partial_{u_k^{i,j}}^* f_t)(U_t, U_t^*) = \mathbb{E}_V \left[\text{Tr}_N(V_{k,T-t}^* \widetilde{D}_k^* Q(V_{T-t} U_t, A) E_{i,j}) \right].$$

We will now compute four different brackets, and by summing them we will get the bracket of M_t (see equation (3.12)). First,

$$\begin{aligned} & \left\langle \sum_{i,j,k} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{\sum_{i,j,k} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) (U_{k,t} dX_{k,t}^N)_{i,j}} \right\rangle \quad (3.13) \\ &= \sum_k \sum_{i,j,r,s} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) \overline{(\partial_{u_k^{r,s}} f_t)(U_t, U_t^*)} \left\langle (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{(U_{k,t} dX_{k,t}^N)_{r,s}} \right\rangle \\ &= \frac{1}{N} \sum_k \sum_{i,j} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) \overline{(\partial_{u_k^{i,j}} f_t)(U_t, U_t^*)} dt \\ &= \frac{1}{N} \sum_k \sum_{i,j} \mathbb{E}_V \left[\text{Tr}_N(\widetilde{D}_k Q(V_{T-t} U_t, A) V_{k,T-t} E_{i,j}) \right] \mathbb{E}_V \left[\text{Tr}_N(E_{j,i} (V_{k,T-t})^* \widetilde{D}_k^* Q(V_{T-t} U_t, A)^*) \right] dt \\ &= \frac{1}{N} \sum_k \mathbb{E}_{V,W} \left[\text{Tr}_N(\widetilde{D}_k Q(V_{T-t} U_t, A) V_{k,T-t} W_{k,T-t}^* \widetilde{D}_k^* Q(W_{T-t} U_t, A)^*) \right] dt. \end{aligned}$$

Similarly one has,

$$\begin{aligned} & \left\langle \sum_{i,j,k} (\partial_{u_k^{i,j}}^* f_t)(U_t, U_t^*) ((U_{k,t} dX_{k,t}^N)^*)_{i,j}, \overline{\sum_{i,j,k} (\partial_{u_k^{i,j}}^* f_t)(U_t, U_t^*) ((U_{k,t} dX_{k,t}^N)^*)_{i,j}} \right\rangle \quad (3.14) \\ &= \frac{1}{N} \sum_k \mathbb{E}_{V,W} \left[\text{Tr}_N \left(V_{k,T-t}^* \widetilde{D}_k^* Q(V_{T-t} U_t, A) \widetilde{D}_k^* Q(W_{T-t} U_t, A)^* W_{k,T-t} \right) \right] dt. \end{aligned}$$

Next we have,

$$\left\langle \sum_{i,j,k} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{\sum_{i,j,k} (\partial_{u_k^{i,j}}^* f_t)(U_t, U_t^*) ((U_{k,t} dX_{k,t}^N)^*)_{i,j}} \right\rangle \quad (3.15)$$

$$\begin{aligned} &= \sum_k \sum_{i,j,r,s} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) \overline{(\partial_{u_k^{r,s}}^* f_t)(U_t, U_t^*)} \left\langle (U_{k,t} dX_{k,t}^N)_{i,j}, \overline{((U_{k,t} dX_{k,t}^N)^*)_{r,s}} \right\rangle \\ &= \frac{1}{N} \sum_k \sum_{i,j,r,s} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) \overline{(\partial_{u_k^{r,s}}^* f_t)(U_t, U_t^*)} (U_{k,t})_{i,r} (U_{k,t})_{s,j} dt \\ &= \frac{1}{N} \sum_k \sum_{i,j,r,s} \mathbb{E}_V \left[\text{Tr}_N (\widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} E_{i,j}) \right] \\ &\quad \times \mathbb{E}_W \left[\text{Tr}_N (E_{s,r} \widetilde{D}_k^* \widetilde{Q}(W_{T-t} U_t, A)^* W_{k,T-t}) \right] (U_{k,t})_{i,r} (U_{k,t})_{s,j} dt \\ &= \frac{1}{N} \sum_k \mathbb{E}_{V,W} \left[\sum_{i,j,r,s} \left(\widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} \right)_{j,i} (U_{k,t})_{i,r} \right. \\ &\quad \left. \times \left(\widetilde{D}_k^* \widetilde{Q}(W_{T-t} U_t, A)^* W_{k,T-t} \right)_{r,s} (U_{k,t})_{s,j} \right] dt \\ &= \frac{1}{N} \sum_k \mathbb{E}_{V,W} \left[\text{Tr}_N \left(\widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} U_{k,t} \widetilde{D}_k^* \widetilde{Q}(W_{T-t} U_t, A)^* W_{k,T-t} U_{k,t} \right) \right] dt. \end{aligned} \quad (3.16)$$

And similarly,

$$\begin{aligned} &\left\langle \sum_{i,j,k} (\partial_{u_k^{i,j}}^* f_t)(U_t, U_t^*) ((U_{k,t} dX_{k,t}^N)_{i,j}), \overline{\sum_{i,j,k} (\partial_{u_k^{i,j}} f_t)(U_t, U_t^*) (U_{k,t} dX_{k,t}^N)_{i,j}} \right\rangle \quad (3.17) \\ &= \frac{1}{N} \sum_k \mathbb{E}_{V,W} \left[\text{Tr}_N \left((V_{k,T-t} U_t^k)^* \widetilde{D}_k^* \widetilde{Q}(V_{T-t} U_t, A) (W_{k,T-t} U_t^k)^* \widetilde{D}_k \widetilde{Q}(W_{T-t} U_t, A)^* \right) \right] dt. \end{aligned}$$

We sum equations (3.13) to (3.17).

$$\begin{aligned} &\text{Var} \left(\text{Tr}_N (\widetilde{Q}(U_T^N, A^N)) \right) \\ &= \frac{1}{N} \sum_k \int_0^T \mathbb{E} \left[\text{Tr}_N \left(\widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} U_{k,t} (W_{k,T-t} U_{k,t})^* \widetilde{D}_k \widetilde{Q}(W_{T-t} U_t, A)^* \right. \right. \\ &\quad \left. \left. + (V_{k,T-t} U_{k,t})^* \widetilde{D}_k^* \widetilde{Q}(V_{T-t} U_t, A) \widetilde{D}_k^* \widetilde{Q}(W_{T-t} U_t, A)^* W_{k,T-t} U_{k,t} \right. \right. \\ &\quad \left. \left. - \widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} U_{k,t} \widetilde{D}_k \widetilde{Q}(W_{T-t} U_t, A)^* W_{k,T-t} U_{k,t} \right. \right. \\ &\quad \left. \left. - (V_{k,T-t} U_t^k)^* \widetilde{D}_k^* \widetilde{Q}(V_{T-t} U_t, A) (W_{k,T-t} U_t^k)^* \widetilde{D}_k \widetilde{Q}(W_{T-t} U_t, A)^* \right) \right] dt \\ &= \frac{1}{N} \sum_k \int_0^T \mathbb{E} \left[\text{Tr}_N \left(\left(\widetilde{D}_k \widetilde{Q}(V_{T-t} U_t, A) V_{k,T-t} U_{k,t} - (V_{k,T-t} U_{k,t})^* \widetilde{D}_k^* \widetilde{Q}(V_{T-t} U_t, A) \right) \times \right. \right. \\ &\quad \left. \left. \left(\widetilde{D}_k \widetilde{Q}(W_{T-t} U_t, A) W_{k,T-t} U_{k,t} - (W_{k,T-t} U_{k,t})^* \widetilde{D}_k^* \widetilde{Q}(W_{T-t} U_t, A)^* \right) \right) \right] dt. \end{aligned}$$

Hence the conclusion. \square

Corollary 3.3.3. *Let $P, Q \in \mathcal{P}_d$, $(U_t^N)_{t \in \mathbb{R}^+}$, $(V_t^N)_{t \in \mathbb{R}^+}$, $(W_t^N)_{t \in \mathbb{R}^+}$ be independent families of p unitary Brownian motions of size N . Let Z^{NM} be a family of deterministic matrices. With*

$$h : x \otimes y \in (\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C}))^{\otimes 2} \mapsto y \sharp x \in \mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C}),$$

and notations as in subsection 3.2.4, one has for any $T \geq 0$,

$$\begin{aligned} & \mathbb{E} \left[(\mathrm{Tr}_N \otimes I_M)^{\otimes 2} \left(\tilde{P}(U_T^N \otimes I_M, Z^{NM}) \otimes \tilde{Q}(U_T^N \otimes I_M, Z^{NM}) \right) \right] \\ & - \mathbb{E} \left[\mathrm{Tr}_N \otimes I_M \right]^{\otimes 2} \left(\tilde{P}(U_T^N \otimes I_M, Z^{NM}) \otimes \tilde{Q}(U_T^N \otimes I_M, Z^{NM}) \right) \\ & = -\frac{1}{N} \sum_{1 \leq k \leq p} \int_0^T \mathbb{E} \left[\mathrm{Tr}_N \otimes I_M \left(h \circ \delta_k \tilde{P}(V_{T-t}^N U_t^N \otimes I_M, Z^{NM}) h \circ \delta_k \tilde{Q}(W_{T-t}^N U_t^N \otimes I_M, Z^{NM}) \right) \right] dt. \end{aligned}$$

Besides if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, then we have the same equality but with \mathcal{D}_k instead of $h \circ \delta_k$.

Proof. Let A^N be a family of deterministic matrices, by polarization and the fact that $\mathcal{D}_k(Q^*)^* = -\mathcal{D}_k Q$, we have

$$\begin{aligned} & \mathbb{E} \left[\mathrm{Tr}_N \left(\tilde{P}(U_T^N, A^N) \right) \mathrm{Tr}_N \left(\tilde{Q}(U_T^N, A^N) \right) \right] - \mathbb{E} \left[\mathrm{Tr}_N \left(\tilde{P}(U_T^N, A^N) \right) \right] \mathbb{E} \left[\mathrm{Tr}_N \left(\tilde{Q}(U_T^N, A^N) \right) \right] \\ & = \mathbb{E} \left[\left(\mathrm{Tr}_N \left(\tilde{P}(U_T^N, A^N) \right) - \mathbb{E} \left[\mathrm{Tr}_N \left(\tilde{P}(U_T^N, A^N) \right) \right] \right) \right. \\ & \quad \left. \times \overline{\left(\mathrm{Tr}_N \left(\tilde{Q}^*(U_T^N, A^N) \right) - \mathbb{E} \left[\mathrm{Tr}_N \left(\tilde{Q}^*(U_T^N, A^N) \right) \right] \right)} \right] \\ & = \frac{1}{N} \sum_{k \leq p} \int_0^T \mathbb{E} \left[\mathrm{Tr}_N \left(\widetilde{\mathcal{D}_k P}(V_{T-t}^N U_t^N, A^N) \times \widetilde{\mathcal{D}_k Q^*}(W_{T-t}^N U_t^N, A^N)^* \right) \right] dt \\ & = -\frac{1}{N} \sum_{k \leq p} \int_0^T \mathbb{E} \left[\mathrm{Tr}_N \left(m \circ \delta_k \tilde{P}(V_{T-t}^N U_t^N, A^N) \times m \circ \delta_k \tilde{Q}(W_{T-t}^N U_t^N, A^N) \right) \right] dt. \end{aligned}$$

Now we want to study a polynomial in $(U_T^N \otimes I_M, Z^{NM})$ and their adjoints. By linearity we can assume that P is a monomial. One can also assume that Z_i^{NM} is a simple tensor, i.e. that $Z_i^{NM} = A_i \otimes B_i$ where $A_i \in \mathbb{M}_N(\mathbb{C})$ and $B_i \in \mathbb{M}_M(\mathbb{C})$. Indeed if $Z_i^{NM} = \sum_l A_{i,l} \otimes B_{i,l}$, then a polynomial in $(U_T^N \otimes I_M, Z^{NM})$ is a linear combination in monomials in $(U_T^N \otimes I_M, (A_{i,l} \otimes B_{i,l})_{i,l})$. Thus

$$\tilde{P}(U_T^N \otimes I_M, Z^{NM}) = \tilde{P}(U_T^N, A) \otimes \tilde{P}(I_M, B).$$

Thus assuming that P and Q are monomials, we have

$$\begin{aligned}
& \mathbb{E} \left[(\mathrm{Tr}_N \otimes I_M) \otimes^2 \left(\tilde{P}(U_T^N \otimes I_M, Z^{NM}) \otimes \tilde{Q}(U_T^N \otimes I_M, Z^{NM}) \right) \right] \\
& - \mathbb{E} [\mathrm{Tr}_N \otimes I_M] \otimes^2 \left(\tilde{P}(U_T^N \otimes I_M, Z^{NM}) \otimes \tilde{Q}(U_T^N \otimes I_M, Z^{NM}) \right) \\
& = \left(\mathbb{E} [\mathrm{Tr}_N(\tilde{P}(U_T^N, A)) \mathrm{Tr}_N(\tilde{Q}(U_T^N, A))] - \mathbb{E} [\mathrm{Tr}_N(\tilde{P}(U_T^N, A))] \mathbb{E} [\mathrm{Tr}_N(\tilde{Q}(U_T^N, A))] \right) \\
& \quad \otimes \tilde{P}(I_M, B) \tilde{Q}(I_M, B) \\
& = -\frac{1}{N} \int_0^T \mathbb{E} \left[\mathrm{Tr}_N \left(m \circ \delta_k \tilde{P}(V_{T-t}^N U_t^N, A) \times m \circ \delta_k \tilde{Q}(W_{T-t}^N U_t^N, A) \right) \right] \otimes \tilde{P}(I_M, B) \tilde{Q}(I_M, B) dt \\
& = -\frac{1}{N} \sum_{k \leq p} \int_0^T \mathbb{E} \left[\mathrm{Tr}_N \otimes I_M \left(h \circ \delta_k \tilde{P}(V_{T-t}^N U_t^N \otimes I_M, Z^{NM}) \times h \circ \delta_k \tilde{Q}(W_{T-t}^N U_t^N \otimes I_M, Z^{NM}) \right) \right] dt.
\end{aligned}$$

Hence the conclusion by linearity. Besides if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$, and that these matrices commute, then for any p -tuple of unitary matrices U , with $Z = (Y_i^M)_i$,

$$\begin{aligned}
h \circ \delta_k \tilde{P}(U \otimes I_M, Z^{NM}) &= \sum_{P=SY_k T} \tilde{T}(U, I_N) \tilde{S}(U, I_N) U_k \otimes \tilde{S}(I_M, Z) \tilde{T}(I_M, Z) \\
&\quad - \sum_{P=SY_{k+d} T} U_k^* \tilde{T}(U, I_N) \tilde{S}(U, I_N) \otimes \tilde{S}(I_M, Z) \tilde{T}(I_M, Z) \\
&= \sum_{P=SY_k T} (\tilde{T}(U, I_N) \otimes \tilde{T}(I_M, Z)) (\tilde{S}(U, I_N) \otimes \tilde{S}(I_M, Z)) \times U_k \otimes I_M \\
&\quad - \sum_{P=SY_{k+d} T} U_k^* \otimes I_M \times (\tilde{T}(U, I_N) \otimes \tilde{T}(I_M, Z)) (\tilde{S}(U, I_N) \otimes \tilde{S}(I_M, Z)) \\
&= \mathcal{D}_k \tilde{P}(U \otimes I_M, Z^{NM})
\end{aligned}$$

Hence once again the conclusion by linearity. □

3.3.3 Convergence of the free unitary Brownian motion

If u_t is a free unitary Brownian motion at time t , one can define μ_{u_t} as in Definition 3.2.1. Then thanks to Riesz theorem, there is a measure ν_t such that for any polynomial P in two commuting variables,

$$\tau(P(u_t, u_t^*)) = \int_{\mathbb{C}} P(z, z^*) d\nu_t(z).$$

The measure ν_t is well-known albeit not explicit. The proof of the following theorem can be found in [83].

Theorem 3.3.4. *For every $t > 0$, the measure ν_t is absolutely continuous with respect to the Haar measure on $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$. For $t > 4$, the support of ν_t is equal to \mathbb{T} , and its density is positive on \mathbb{T} . We set $\kappa(t, \omega)$ the density of ν_t at the point $\omega \in \mathbb{T}$. Then for $t > 4$, $\kappa(t, \omega)$ is the real part of the only solution with positive real part of the equation,*

$$\frac{z-1}{z+1} e^{\frac{t}{2}z} = \omega. \tag{3.18}$$

The following theorem states that given a free unitary Brownian motion $(u_t)_{t \geq 0}$, there exists a family of Haar unitaries $(\tilde{u}_t)_{t \geq 0}$ such that when t goes to infinity, u_t and \tilde{u}_t are exponentially close for the operator norm topology. And more importantly it gives explicit estimates.

Proposition 3.3.5. *There exists a \mathcal{C}^* -algebra \mathcal{C} which contains u_t^1, \dots, u_t^p freely independent free unitary Brownian motions at time $t \geq 5$ and $\tilde{u}_t^1, \dots, \tilde{u}_t^p$ freely independent Haar unitaries such that for any i , $\|u_t^i - \tilde{u}_t^i\| \leq 4e^2\pi e^{-\frac{t}{2}}$.*

Proof. We view $B(L^2([0, 2\pi]))$ as the \mathcal{C}^* -algebra endowed with the state

$$\tau(u) = \langle u(\mathbf{1}_{[0, 2\pi]}), \mathbf{1}_{[0, 2\pi]} \rangle_{L^2([0, 2\pi])}.$$

The endomorphism $x : f \mapsto (t \rightarrow tf(t))$ is self-adjoint and has distribution (as defined in 3.2.1) $\mu_x(f) = \int_{[0, 2\pi]} f$. Consequently we set $g : s \rightarrow \kappa(t, e^{is})$ and $G : s \rightarrow \int_0^s g(u) du$. Since g is positive, we can define $u_t = e^{iG^{-1}(x)}$ which has the distribution of a free unitary Brownian motion at time t , indeed for any polynomial P in two commuting indeterminates,

$$\tau(P(u_t, u_t^*)) = \int_0^{2\pi} P(e^{iG^{-1}(s)}, e^{-iG^{-1}(s)}) ds = \int_0^{2\pi} P(e^{is}, e^{-is})g(s)ds = \int_{\mathbb{C}} f(z, z^*) d\nu_t(z).$$

And similarly, $u = e^{ix}$ is a Haar unitary. Besides, since

$$u_t - u = \int_0^1 e^{i\alpha G^{-1}(x)} (G^{-1}(x) - x) e^{i(1-\alpha)x} d\alpha,$$

thanks to the fact that G is a diffeomorphism of $[0, 2\pi]$,

$$\|u_t - u\| \leq \|G^{-1}(x) - x\| = \sup_{s \in [0, 2\pi]} |G^{-1}(s) - s| = \sup_{s \in [0, 2\pi]} |s - G(s)| \leq 2\pi \sup_{s \in [0, 2\pi]} |1 - g(s)|.$$

We set $y(s)$ the imaginary part of the only solution with positive real part of the equation (3.18). Then we have for any s ,

$$\frac{(g(s) - 1)^2 + y(s)^2}{(g(s) + 1)^2 + y(s)^2} \leq e^{-tg(s)}.$$

However since $(g(s) - 1)^2 \leq (g(s) + 1)^2$, we have,

$$\frac{(g(s) - 1)^2}{(g(s) + 1)^2} \leq \frac{(g(s) - 1)^2 + y(s)^2}{(g(s) + 1)^2 + y(s)^2} \leq e^{-tg(s)}.$$

First in the case where $g(s) \geq 1$, then since we assumed $t \geq 4$, $|g(s) - 1| \leq (|g(s) - 1| + 2)e^{-2|g(s)-1|}e^{-\frac{t}{2}}$, and since the function $u \rightarrow (u+2)e^{-2u}$ is decreasing, we have, $|g(s) - 1| \leq 2e^{-\frac{t}{2}}$. If $g(s) \leq 1$, then after studying the graph of the function $h : g \mapsto e^{-tg/2} - \frac{1-g}{1+g}$, we have that this function takes value 0 in in 0, is negative on $(0, c_t)$ for some $c_t \in (0, 1)$, and finally is positive for $g > c_t$. Since we know that $g(s)$ is positive for $t > 4$ and $h(g(s)) \geq 0$, if we find g such that $h(g) \leq 0$, then $g(s) \geq g$. Besides for $t \geq 5$, we have that $h\left(\ln(t/2)\frac{2}{t}\right) \leq 0$. Thus necessarily $g(s) \geq \ln(t/2)\frac{2}{t}$, consequently since $g(s) \leq 1$, we know that $1 - g(s) \leq 2e^{-\frac{t}{2}g(s)}$. Hence,

$$1 - g(s) \leq 2e^{-\frac{t}{2} \times \ln(t/2)\frac{2}{t}} = \frac{4}{t}.$$

Thus by bootstrapping, for any s ,

$$1 - g(s) \leq 2e^{-\frac{t}{2}(1-\frac{4}{t})} = 2e^2e^{-\frac{t}{2}}.$$

Consequently $\|u_t - u\| \leq 4e^2\pi e^{-t/2}$, and thanks to Theorem 3.2.3, to conclude we just need to take the free product of p copies of $B(L^2([0, 2\pi]))$. □

3.3.4 Free stochastic calculus and free unitary Brownian motion

In this subsection, we consider $u_t^N = (U_1^N u_t^1, \dots, U_p^N u_t^p)$, i.e. free unitary Brownian started in U^N . As we will see later, thanks to Proposition 3.3.5, this will let us interpolate between $U^N = (U_1^N, \dots, U_p^N)$ random Haar unitary matrices and $u = (u^1, \dots, u^p)$ free Haar unitaries. Concretely if $P \in \mathcal{P}_d$, we set

$$H(t) = \tau_N \otimes \tau_M \left(\tilde{P}(u_t^N \otimes I_M, Z^{NM}) \right).$$

Then,

$$H(0) = \frac{1}{NM} \text{Tr}_{MN} \left(\tilde{P}(U^N \otimes I_M, Z^{NM}) \right),$$

$$H(\infty) = \lim_{t \rightarrow \infty} H(t) = \tau_N \otimes \tau_M \left(\tilde{P}(u \otimes I_M, Z^{NM}) \right).$$

To prove the second line, that is to prove that H converges towards $H(\infty)$, we first use Proposition 3.3.5 to prove that H converges towards $\tau_N \otimes \tau_M \left(\tilde{P}((U^N u) \otimes I_M, Z^{NM}) \right)$. Then thanks to the invariance of the distribution of a Haar unitary by multiplication by a unitary operator, almost surely this quantity is equal to $H(\infty)$. The invariance can easily be proved by using the fact (see Theorem 5.4.10 of [13]) that if $V^{kN} = (V_1^{kN}, \dots, V_p^{kN})$ are independent Haar unitary matrices of size kN , then for any polynomial P ,

$$\tau(\tilde{P}(U^N u)) = \lim_{k \rightarrow \infty} \mathbb{E} \left[\tau_N \otimes \tau_k(\tilde{P}(U^N \otimes I_k \times V^{kN})) \right] = \lim_{k \rightarrow \infty} \mathbb{E} \left[\tau_{kN}(\tilde{P}(V^{kN})) \right] = \tau(\tilde{P}(u)).$$

Consequently as long as the integral is well-defined, we can write

$$\frac{1}{NM} \text{Tr}_{MN} \left(\tilde{P}(U^N \otimes I_M, Z^{NM}) \right) - \tau_N \otimes \tau_M \left(\tilde{P}(u \otimes I_M, Z^{NM}) \right) = \int_0^\infty \frac{dH}{dt}(t) dt.$$

Hence we need to compute the differential of H with respect to t , which we do in the following proposition.

Proposition 3.3.6. *Let the following objects be given,*

- $u = (u_t^1, \dots, u_t^p)_{t \geq 0}$ a family of p free unitary Brownian motions,
- $U^N = (U_1^N, \dots, U_p^N)$ matrices of size N ,
- $u_t^N = (U_1^N u_t^1, \dots, U_p^N u_t^p)$ elements of \mathcal{A}_N ,
- $Z^{NM} = (Z_{p+1}^{NM}, \dots, Z_d^{NM})$ matrices in $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$,
- $P \in \mathcal{P}_d$.

With notation as in subsection 3.2.4, the map $H : t \mapsto \tau_N \otimes \tau_M \left(\tilde{P}(u_t^N \otimes I_M, Z^{NM}) \right)$ is differentiable on \mathbb{R}^+ and,

$$\frac{dH}{dt}(t) = -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(\delta_i \mathcal{D}_i \tilde{P}(u_t^N \otimes I_M, Z^{NM}) \right) \right).$$

Proof. We want to use Theorem 3.2.10 to write H as an integral which we can then easily differentiate. We need to define $X_0 \in \mathcal{A}^d$, K such that for any $t \geq 0$, $K \in (L^1([0, t], \mathcal{A}))^d$, U such that for any $t \geq 0$, $(\mathbf{1}_{s \leq t} U_s^i)_{t \in \mathbb{R}^+} \in (\mathcal{B}_a^\infty)^d$, and then,

$$X_t = X_0 + \int_0^t K_s ds + \sum_i \int_0^t U_s^i \# dS_s^i.$$

By using the linearity of the trace and the non-commutative differential, we can assume that $Z_i^{NM} = A_i \otimes B_i$ where $A_i \in \mathbb{M}_N(\mathbb{C})$ and $B_i \in \mathbb{M}_M(\mathbb{C})$. We then set $X_t = (u_t^N, u_t^{N*}, A, A^*)$. Since (A, A^*) is free from \mathcal{A} , the processes K and U associated to (A, A^*) are zero. As for (u_t^N, u_t^{N*}) , by definition of a free unitary Brownian motion, we have

$$\forall t \geq 0, \quad u_t^N = U^N - \int_0^t \frac{u_s^N}{2} ds + \mathbf{i} \int_0^t (u_s^N \otimes \mathbf{1}_A) \# dS_s,$$

$$\forall t \geq 0, \quad (u_t^N)^* = (U^N)^* - \int_0^t \frac{(u_s^N)^*}{2} ds - \mathbf{i} \int_0^t (\mathbf{1}_A \otimes (u_s^N)^*) \# dS_s.$$

To minimize cumbersome notations, for the rest of this proof we will forget about the N in u_t^N , and assimilate u_t^N with u_t . Consequently we set for any $s \geq 0$,

$$\begin{aligned} \forall i \in [1, p], \forall j \in [1, p], \quad K_s^j &= -u_{j,s}/2, \quad U_s^{i,j} = \mathbf{i} \mathbf{1}_{i=j} u_{j,s} \otimes \mathbf{1}_A, \\ \forall i \in [1, p], \forall j \in [p+1, 2p], \quad K_s^j &= -u_{j,s}^*/2, \quad U_s^{i,j} = -\mathbf{i} \mathbf{1}_{i=j} \mathbf{1}_A \otimes u_{j,s}^*, \\ \forall i \in [1, p], \forall j > 2p, \quad K_s^j &= 0, \quad U_s^{i,j} = 0 \otimes 0. \end{aligned}$$

Thus we have for any monomial Q ,

$$\partial Q(X) \# K = -\frac{1}{2} \sum_{1 \leq i \leq p} \partial_i Q(X) \# u_i + \partial_i^* Q(X) \# (u_i)^*,$$

$$\begin{aligned} \Delta_U(Q)(X) &= - \sum_{1 \leq i \leq p} \langle \langle (\partial_i \otimes I) \circ \partial_i Q(X) \rangle \# (u_i \otimes \mathbf{1}_A, u_i \otimes \mathbf{1}_A) \rangle \rangle \\ &\quad - \langle \langle (\partial_i \otimes I) \circ \partial_i^* Q(X) \rangle \# (u_i \otimes \mathbf{1}_A, \mathbf{1}_A \otimes (u_i)^*) \rangle \rangle \\ &\quad - \langle \langle (\partial_i^* \otimes I) \circ \partial_i Q(X) \rangle \# (\mathbf{1}_A \otimes (u_i)^*, u_i \otimes \mathbf{1}_A) \rangle \rangle \\ &\quad + \langle \langle (\partial_i^* \otimes I) \circ \partial_i^* Q(X) \rangle \# (\mathbf{1}_A \otimes (u_i)^*, \mathbf{1}_A \otimes (u_i)^*) \rangle \rangle. \end{aligned}$$

And thanks to Theorem 3.2.10, we have for any $t \geq 0$,

$$Q(X_t) = Q(X_0) + \int_0^t \partial Q(X_s) \# K_s ds + \sum_{1 \leq i \leq p} \int_0^t (\partial Q(X) \# U_s^i) \# dS_s^i + \int_0^t \Delta_U(Q)(X_s) ds.$$

Thus if we fix $t \in \mathbb{R}^+$, then for any $\varepsilon \geq -t$,

$$Q(X_{t+\varepsilon}) - Q(X_t) = \int_t^{t+\varepsilon} \partial Q(X_s) \# K_s ds + \sum_{1 \leq i \leq p} \int_t^{t+\varepsilon} (\partial Q(X) \# U_s^i) \# dS_s^i + \int_t^{t+\varepsilon} \Delta_U(Q)(X_s) ds.$$

As we said in section 3.2.3, $(\sum_i \int_0^t (\partial Q(X) \# U_s^i) \# dS_s^i)_{t \geq 0}$ is a martingale, thus

$$\tau_N(Q(X_{t+\varepsilon})) - \tau_N(Q(X_t)) = \int_t^{t+\varepsilon} \tau_N(\partial Q(X_s) \# K_s) ds + \int_t^{t+\varepsilon} \tau_N(\Delta_U(Q)(X_s)) ds.$$

Finally we have,

$$\frac{d\tau_N(Q(X_t))}{dt} = \tau_N(\partial Q(X_t) \# K_t) + \tau_N(\Delta_U(Q)(X_t)). \quad (3.19)$$

Besides,

$$\tau_N(\partial Q(X_t) \# K_t) = -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N(D_i Q(X_t) u_{i,t}) + \tau_N(u_{i,t}^* D_i^* P(X_t)),$$

and,

$$\tau_N(\langle\langle (\partial_i \otimes I) \circ \partial_i Q(X) \# (u_i \otimes \mathbf{1}_A, u_i \otimes \mathbf{1}_A) \rangle\rangle) = \sum_{Q=AY_i BY_i C} \tau_N(C(X)A(X)u_i) \tau_N(u_i B(X)),$$

$$\tau_N(\langle\langle (\partial_i \otimes I) \circ \partial_i^* Q(X) \# (u_i \otimes \mathbf{1}_A, \mathbf{1}_A \otimes u_i^*) \rangle\rangle) = \sum_{Q=AY_i BY_i^* C} \tau_N(A(X)u_i u_i^* C(X)) \tau_N(B(X)),$$

$$\tau_N(\langle\langle (\partial_i^* \otimes I) \circ \partial_i Q(X) \# (\mathbf{1}_A \otimes u_i^*, u_i \otimes \mathbf{1}_A) \rangle\rangle) = \sum_{Q=AY_i^* BY_i C} \tau_N(A(X)C(X)) \tau_N(u_i u_i^* B(X)),$$

$$\tau_N(\langle\langle (\partial_i^* \otimes I) \circ \partial_i^* Q(X) \# (\mathbf{1}_A \otimes u_i^*, \mathbf{1}_A \otimes u_i^*) \rangle\rangle) = \sum_{Q=AY_i^* BY_i^* C} \tau_N(C(X)A(X)u_i^*) \tau_N(u_i^* B(X)).$$

Besides we also have,

$$\begin{aligned} \tau_N \otimes \tau_N (\delta_i D_i Q(X) \times 1 \otimes u_i) &= 2 \sum_{Q=AY_i BY_i C} \tau(B(X)u_i) \tau(C(X)A(X)u_i) \\ &\quad - \sum_{Q=AY_i^* BY_i C} \tau(C(X)A(X)) \tau(B(X)u_i u_i^*) \\ &\quad - \sum_{Q=AY_i BY_i^* C} \tau(B(X)) \tau(C(X)u_i u_i^* A(X)), \end{aligned}$$

$$\begin{aligned} \tau_N \otimes \tau_N (\delta_i D_i^* Q(X) \times u_i^* \otimes 1) &= -2 \sum_{Q=AY_i^* BY_i^* C} \tau(u_i^* B(X)) \tau(C(X)A(X)u_i^*) \\ &\quad + \sum_{Q=AY_i^* BY_i C} \tau(B(X)u_i u_i^*) \tau(C(X)A(X)) \\ &\quad + \sum_{Q=AY_i BY_i^* C} \tau(C(X)u_i u_i^* A(X)) \tau(B(X)). \end{aligned}$$

Which means that

$$\tau_N(\Delta_U(Q)(X)) = -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N \otimes \tau_N (\delta_i D_i Q(X) \times 1 \otimes u_i) - \tau_N \otimes \tau_N (\delta_i D_i^* Q(X) \times u_i^* \otimes 1).$$

And thus when combined with equation (3.19), we get that

$$\begin{aligned}
 \frac{d\tau_N(Q(X_t))}{dt} &= -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N(D_i Q(X_t) u_{i,t}) + \tau_N \otimes \tau_N (\delta_i D_i Q(X_t) \times 1 \otimes u_{i,t}) \\
 &\quad + \tau_N(u_{i,t}^* D_i^* Q(X_t)) - \tau_N \otimes \tau_N (\delta_i D_i^* Q(X_t) \times u_{i,t}^* \otimes 1) \\
 &= -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N \otimes \tau_N (\delta_i (D_i Q(X_t) u_{i,t})) - \tau_N \otimes \tau_N (\delta_i (u_{i,t}^* D_i^* Q(X_t))) \\
 &= -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N \otimes \tau_N (\delta_i \mathcal{D}_i Q(X_t)).
 \end{aligned}$$

Now we want to study a polynomial in (u_t^N, Z^{NM}) and their adjoints. If P is a monomial, we have,

$$\tilde{P}(u_t^N \otimes I_M, Z^{NM}) = \tilde{P}(u_t^N, A) \otimes \tilde{P}(I_M, B).$$

Therefore,

$$\frac{dH}{dt}(t) = -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_N \otimes \tau_N (\delta_i \mathcal{D}_i \tilde{P}(u_t^N, A)) \times \tau_M (\tilde{P}(I_M, B)).$$

And since for any $S, T \in \mathcal{P}_d$,

$$\begin{aligned}
 &\tau_N \otimes \tau_N (\delta_i \widetilde{TS}(u_t^N, A)) \times \tau_M (\widetilde{ST}(I_M, B)) \\
 &= \tau_M ((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) (\delta_i \widetilde{TS}(u_t^N \otimes I_M, A \otimes B))).
 \end{aligned}$$

Hence after summing,

$$\frac{d}{dt} \tau_N \otimes \tau_M (\tilde{P}(u_t^N, Z^{NM})) = -\frac{1}{2} \sum_{1 \leq i \leq p} \tau_M ((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) (\delta_i \mathcal{D}_i \tilde{P}(u_t^N, Z^{NM}))),$$

and we conclude by linearity. □

3.4 Proof of Theorem 3.1.1, the main result

3.4.1 Overview of the proof

If we take the point of view of free probability – for details we refer to the third point of Definition 3.2.1 – we have two families of non-commutative random variables, $(U^N \otimes I_M, Z^{NM})$ and $(u \otimes I_M, Z^{NM})$, and we want to study the difference between their distributions. As mentioned in the introduction the main idea of the proof is to interpolate those two families with the help of p free unitary Brownian motions $(u_t^N)_{t \geq 0}$ started in the Haar unitary matrices U^N . A big difference with the case of the GUE which was treated in [7] is that we do not have an explicit expression of the law of u_t^N in function of U^N and u , which is why we had to introduce notions of free stochastic calculus.

This idea of interpolating random matrices is not entirely new. Indeed in [84], the authors proved in theorem 1.3 that given $(U_t^N)_{t \geq 0}$ a unitary Brownian motion of size N and $(u_t)_{t \geq 0}$ a free unitary Brownian motion, for any $n \in \mathbb{N}$,

$$\left| \mathbb{E} [\tau_N((U_t^N)^n)] - \tau(u_t^n) \right| \leq \frac{t^2 n^4}{N^2}. \quad (3.20)$$

The main idea of their proof was to interpolate $\mathbb{E} \left[\tau_N((U_t^N)^n) \right]$ and $\mathbb{E} \left[\tau_N((U_t^{2N})^n) \right]$ through a stochastic process defined with a unitary Brownian motion. Thus they get an estimate of the difference between those two expectations, and by iterating this method they get an estimate of

$$\mathbb{E} \left[\tau_N((U_t^{2^i N})^n) \right] - \mathbb{E} \left[\tau_N((U_t^{2^{i+1} N})^n) \right]$$

for any i . Then by summing over i they get equation (3.20). In this paper, while the method used are very different, we do keep this idea of interpolation. But instead of doing it step by step from $2^i N$ to $2^{i+1} N$, we directly interpolate between N and ∞ .

Since our aim in this subsection is not to give a proof but to outline the strategy used in subsection 3.4.2, we assume that we have no matrix Z^{NM} and that $M = 1$. Now under the assumption that this is well-defined, if Q is a non-commutative polynomial,

$$\mathbb{E} \left[\frac{1}{N} \text{Tr}_N \left(Q(U^N) \right) \right] - \tau \left(Q(u) \right) = - \int_0^\infty \mathbb{E} \left[\frac{d}{dt} \left(\tau_N \left(Q(u_t^N) \right) \right) \right] dt.$$

In the classical case, if $(S_t)_{t \geq 0}$ is a Markov process with generator θ , then under the appropriate assumption we have

$$\frac{d}{dt} \mathbb{E}[f(S_t)] = \mathbb{E}[(\theta f)(S_t)].$$

And if the law of the process at time 0 is invariant for this Markov process we have that for any t , $\mathbb{E}[(\theta f)(S_t)] = 0$. Since $(u_t^N)_{t \geq 0}$ is a free Markov process, we expect to get similarly that

$$\frac{d}{dt} \left(\tau_N \left(Q(u_t^N) \right) \right) = \tau_N \left((\Theta Q)(u_t^N) \right),$$

for some generator Θ which we will compute with the help of Proposition 3.3.6. Besides the invariant law of a free Brownian motion is the law of free Haar unitaries. Thus if $(u_t)_{t \geq 0}$ is a free Brownian motion started in free Haar unitaries, we have that $\tau((\Theta Q)(u_t)) = 0$. Since unitary Haar matrices converges in distribution towards free Haar unitaries (see [13], Theorem 5.4.10), we have that $\tau_N((\Theta Q)(u_t^N))$ converges towards $\tau((\Theta Q)(u_t)) = 0$. As we will see in this proof, the convergence happens at a speed of N^{-2} . To prove this, the main idea is to view free unitary Brownian motions started in U^N as the asymptotic limit when k goes to infinity of a unitary Brownian motion of size kN started in $U^N \otimes I_k$ (see Proposition 3.4.4).

Another issue is that to prove Theorem 3.1.1, we would like to set $Q = f(P)$ but since f is not polynomial this means that we need to extend the definition of operators such as δ_i . In order to do so we assume that there exists μ a measure on \mathbb{R} such that,

$$\forall x \in \mathbb{R}, \quad f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y).$$

While we have to assume that the support of μ is indeed on the real line, μ can be a complex measure. However we will usually work with measure such that $|\mu|(\mathbb{R})$ is finite. Indeed under this assumption we can use Fubini's theorem, and we get

$$\mathbb{E} \left[\frac{1}{M} \text{Tr}_N \left(f \left(P(U^N) \right) \right) \right] - \tau \left(f \left(P(u) \right) \right) = \int_{\mathbb{R}} \left\{ \mathbb{E} \left[\frac{1}{N} \text{Tr}_N \left(e^{iyP(U^N)} \right) \right] - \tau \left(e^{iyP(u)} \right) \right\} d\mu(y).$$

We can then set $Q = e^{iyP}$. And even though this is not polynomial, since it is a power series, most of the properties associated to polynomials remain true with some assumption on the

convergence. The main difficulty with this method is that we need to find a bound uniform in y , indeed we have terms of the form

$$\int_{\mathbb{R}} |y|^l d|\mu|(y)$$

which appear. Thanks to Fourier integration we can relate the exponent l to the regularity of the function f , thus we want to find a bound with l as small as possible. It turns out that with our proof $l = 4$.

3.4.2 Proof of Theorem 3.4.1

In this section we focus on proving theorem 3.4.1 from which we deduce all of the important corollaries, and notably Theorem 3.1.1. Since this subsection is dedicated to proving only one theorem but is the longest of the paper, in the next few paragraphs we explain the structure of the proof. Unlike subsection 3.4.1 where we gave ideas on the method of the proof, here we mainly focus on which specific purpose serves every lemma. We begin by applying Proposition 3.3.6 to directly obtain Lemma 3.4.2 which states that the difference between the trace at time N , that is a trace in Haar unitary matrices, and at infinity, that is a trace in free Haar unitaries, can be written as an integral with respect to t of a trace in free unitary Brownian motions at time t . We then proceed to study the term under the integral that we name $S_{t,y}^N(A, B)$ in Definition 3.4.3.

The next two lemmas are technical lemmas that we need to justify further computations. First Proposition 3.4.4 shows that one can see a trace of a power series in free unitary Brownian motions and matrices U^N of size N as the limit of the trace of the same power series but evaluated in independent unitary Brownian motions of size kN and $U^N \otimes I_k$. The proof of Proposition 3.4.4 can be summarized as using well-known theorems on the convergence in distribution of random matrices. The proof of the second one, Lemma 3.4.5, is much more subtle. It gives an estimate in k of the non diagonal coefficients of our power series in independent unitary Brownian motions of size kN and $U^N \otimes I_k$. The proof relies on Gronwall's inequality to reduce the problem to the polynomial case and classical stochastic calculus to deal with the former.

Lemma 3.4.6 let us write $S_{t,y}^N(A, B)$ as a linear combination of covariance terms. Indeed, thanks to Proposition 3.4.4 one can write $S_{t,y}^N(A, B)$ as the limit when k goes to infinity of a linear combination of products of expectations of traces in power series in independent unitary Brownian motions of size kN started in $U^N \otimes I_k$ where U^N are Haar unitary matrices. It turns out that the expectation of the product of those traces converges towards 0 when k goes to infinity thanks to Lemma 3.4.5 and the properties of Haar unitary matrices. Hence every product of expectations can be viewed as a covariance term. Interestingly enough, this is the only part of the proof where we use that U^N are Haar unitary matrices. More precisely we use that the law of such a random matrix is invariant by multiplication by a unitary matrix. In every other part of the proof one only need to assume at most that U^N are unitary matrices.

Finally, in Lemma 3.4.7, we use Corollary 3.3.3 to get an upper bound independent of k of those covariance terms. Thus we can let k go to infinity to get an upper bound of $S_{t,y}^N(A, B)$. As usual, the covariance of renormalized traces on $\mathbb{M}_N(\mathbb{C})$ is of order N^{-2} . And even though we are not exactly working with a trace on $\mathbb{M}_N(\mathbb{C})$ here, when using Corollary 3.3.3 the upper bound that we get is of order N^{-2} . Which means that so is the difference between the trace at time N and at infinity. Finally this upper bound immediately yields Theorem 3.4.1. Then we

conclude this section by a proof of Theorem 3.1.1 which, up to a trick to assume that they have compact support, mainly consist in checking that the functions that we consider in Theorem 3.1.1 satisfies the hypothesis of Theorem 3.4.1.

Theorem 3.4.1. *We define*

- $u = (u^1, \dots, u^p)$ a family of p free Haar unitaries,
- $U^N = (U_1^N, \dots, U_p^N)$ i.i.d. Haar unitary matrices of size N .
- $Z^{NM} = (Z_{p+1}^{NM}, \dots, Z_d^{NM})$ deterministic matrices,
- $P \in \mathcal{P}_d$ a self-adjoint polynomial,
- $f : \mathbb{R} \mapsto \mathbb{R}$ such that there exists a measure on the real line μ with $\int (1+y^4) d|\mu|(y) < +\infty$ and for any $x \in \mathbb{R}$,

$$f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y).$$

Then there exists a polynomial $L_P \in \mathbb{R}^+[X]$ which only depends on P such that for any N, M ,

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{MN} \left(f \left(\tilde{P} \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] - \tau_N \otimes \tau_M \left(f \left(\tilde{P} \left(u \otimes I_M, Z^{NM} \right) \right) \right) \right| \\ & \leq \frac{\ln^2(N)M^2}{N^2} L_P \left(\|Z^{NM}\| \right) \times \int_{\mathbb{R}} (|y| + y^4) d|\mu|(y). \end{aligned}$$

where $\|Z^{NM}\| = \sup_{p+1 \leq i \leq d} \|Z_i^{NM}\|$.

Even though we do not give an explicit expression for L_P , it is possible to compute it rather easily by following the proof of Lemma 3.4.7. In particular given a set of polynomials whose degree and coefficients are uniformly bounded, we can find a polynomial R such that for any P in this set and any matrices Z^{NM} , $L_P \left(\|Z^{NM}\| \right) \leq R \left(\|Z^{NM}\| \right)$. Besides, if we replace P by αP where $\alpha \in \mathbb{C}$, then up to a constant independent from α , we can bound $L_{\alpha P}$ by $(|\alpha| + |\alpha|^5)L_P$, or even $(|\alpha| + |\alpha|^4)L_P$ if one picks the first expression in the minimum. It is also worth noting that the set of function $f : \mathbb{R} \mapsto \mathbb{C}$ such that there exists a Borel complex measure on the real line μ with $\int y^4 d|\mu|(y) < +\infty$ and for any $x \in \mathbb{R}$, $f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y)$, is the so-called 4th Wiener space $W_4(\mathbb{R})$. We refer to [87], section 4.3 for a brief introduction on the matter.

The first step to prove this theorem is the following lemma, who is a direct consequence of Proposition 3.3.6 and equation (3.5),

Lemma 3.4.2. *With the same notation as in Theorem 3.4.1, we define*

- $u = (u_t^1, \dots, u_t^p)_{t \geq 0}$ a family of p free unitary Brownian motions,
- $u_t^N = (U_1^N u_t^1, \dots, U_p^N u_t^p)$ elements of \mathcal{A}_N .

Then with notation as in subsection 3.2.4, almost surely

$$\begin{aligned} & \frac{1}{MN} \text{Tr}_{MN} \left(f \left(\tilde{P} \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) - \tau_N \otimes \tau_M \left(f \left(\tilde{P} \left(u_T^N \otimes I_M, Z^{NM} \right) \right) \right) \\ & = \frac{1}{2} \sum_{i \leq p} \int \int_0^T \tau_M \left(\left(\tau_N \otimes I_M \right) \otimes \left(\tau_N \otimes I_M \right) \left(\delta_i \left(\mathcal{D}_i e^{iy\tilde{P}} \right) \left(u_t^N \otimes I_M, Z^{NM} \right) \right) \right) dt d\mu(y). \end{aligned}$$

Since $\mathcal{D}_i e^{iyP} = iy \delta_i P \# e^{iyP}$, this prompts us to define the following quantity.

Definition 3.4.3. Let $A, B \in \mathcal{P}_d$, we set

$$S_{t,y}^N(A, B) = \mathbb{E} \left[\tau_M \left((\tau_N \otimes I_M) \otimes (\tau_N \otimes I_M) \left(\delta_i \left(A e^{iyP} B \right) \left(Z_t^N \right) \right) \right) \right],$$

where $Z_t^N = \left(u_t^N \otimes I_M, Z^{NM}, (u_t^N)^* \otimes I_M, (Z^{NM})^* \right)$.

The following proposition justifies why the family $(U^N \otimes I_M, u_t \otimes I_M, Z^{NM})$ has in the large k limit the distribution – in the sense of Definition 3.2.1 – of the family $(U^N \otimes I_{kM}, U_t^{kN} \otimes I_M, Z^{NM} \otimes I_k)$ where U_t^{kN} are independent unitary Brownian motions of size kN at time t .

Proposition 3.4.4. If U_t^{kN} are unitary Brownian motions of size kN at time t , independent of U^N , we set

$$Y_t^k = \left((U^N \otimes I_k U_t^{kN}) \otimes I_M, Z^{NM} \otimes I_k, (U^N \otimes I_k U_t^{kN})^* \otimes I_M, (Z^{NM})^* \otimes I_k \right).$$

Then if $q = A e^{iyP} B$, we have that for any t , almost surely with respect to U^N ,

$$(\tau_N \otimes I_M)(q(Z_t^N)) = \lim_{k \rightarrow \infty} \mathbb{E}_k \left[(\tau_{kN} \otimes I_M)(q(Y_t^k)) \right],$$

where \mathbb{E}_k is the expectation with respect to $(U_t^{kN})_{t \geq 0}$.

Proof. It has been known for a long time that the unitary Brownian motion converges in distribution towards the free unitary Brownian motion, see [88]. However since we also have to consider deterministic matrices we will use Theorem 1.4 of [84]. This theorem states that if $(U_t^{kN})_{t \geq 0}$ are independent unitary Brownian motions and D^{kN} is a family of deterministic matrices which converges strongly in distribution towards a family of non-commutative random variables d , the family (U_t^{kN}, D^{kN}) in the non-commutative probability space $(\mathbb{M}_{kN}(\mathbb{C}), *, \mathbb{E}_k[\frac{1}{kN} \text{Tr}])$ converges strongly in distribution towards the family (u_t, d) where u_t are freely independent free unitary Brownian motions at time t free from d . That being said, we do not use the convergence of the norm, we only need the convergence in distribution which is way easier to prove through induction and stochastic calculus. In our situation we can write for every i ,

$$Z_i^{NM} = \sum_{1 \leq r, s \leq N} E_{r,s} \otimes A_{r,s,i}^M.$$

Thus if $E^N = (E_{r,s})_{1 \leq r, s \leq N}$, we fix $D^{kN} = (U^N \otimes I_k, E^N \otimes I_k)$, and we can apply Theorem 1.4 of [84] to get that for any non-commutative polynomial P ,

$$\lim_{k \rightarrow \infty} \mathbb{E}_k \left[\tau_{kN}(\tilde{P}(U_t^{kN}, D^{kN})) \right] = \tau_N \left(\tilde{P}(u_t, U^N, E^N) \right).$$

Consequently, for any non-commutative polynomial P , we also have

$$\lim_{k \rightarrow \infty} \mathbb{E}_k \left[\tau_{kN} \otimes I_M(\tilde{P}(U_t^{kN}, D^{kN}, A^M)) \right] = \tau_N \otimes I_M \left(\tilde{P}(u_t, U^N, E^N, A^M) \right).$$

Hence for any $P \in \mathcal{P}_d$,

$$\lim_{k \rightarrow \infty} \mathbb{E}_k \left[\tau_{kN} \otimes I_M(P(Y_t^k)) \right] = \tau_N \otimes I_M \left(P(Z_t^N) \right).$$

Now since U_t^{kN} are unitary matrices, we can find a polynomial L such that for any k , $\|P(Y_t^k)\| \leq C = L(\|U^N\|, \|Z^{NM}\|)$. Knowing this, let $f_\varepsilon \in \mathbb{C}[X]$ be a polynomial which is ε -close to $x \mapsto e^{iyx}$ on the interval $[-C, C]$. Since one can always assume that $C > \|P(Z_t^N)\|$, we have a constant K such that

$$\begin{aligned} & \|(\tau_N \otimes I_M)(q(Z_t^N)) - (\tau_N \otimes I_M)((Af_\varepsilon(P)B)(Z_t^N))\| \leq K\varepsilon, \\ & \|(\tau_N \otimes I_M)(q(Y_t^k)) - (\tau_N \otimes I_M)((Af_\varepsilon(P)B)(Y_t^k))\| \leq K\varepsilon. \end{aligned}$$

Thus

$$\begin{aligned} & \|(\tau_N \otimes I_M)(q(Z_t^N)) - \mathbb{E}_k [(\tau_{kN} \otimes I_M)(q(Y_t^k))]\| \\ & \leq \|(\tau_N \otimes I_M)((Af_\varepsilon(P)B)(Z_t^N)) - \mathbb{E}_k [(\tau_{kN} \otimes I_M)((Af_\varepsilon(P)B)(Y_t^k))]\| + 2K\varepsilon. \end{aligned}$$

Consequently

$$\limsup_{k \rightarrow \infty} \|(\tau_N \otimes I_M)(q(Z_t^N)) - \mathbb{E}_k [(\tau_{kN} \otimes I_M)(q(Y_t^k))]\| \leq 2K\varepsilon.$$

This completes the proof. □

The next lemma shows that the non-diagonal coefficients can actually be neglected.

Lemma 3.4.5. *We define Y_t^k as in Proposition 3.4.4, $P_{1,2} = I_N \otimes E_{1,2} \otimes I_M$, $q = Ae^{iyP}B$, then*

$$\lim_{k \rightarrow \infty} k^{1/2} \mathbb{E}_k [(\text{Tr}_{kN} \otimes I_M)(q(Y_t^k)P_{1,2})] = 0.$$

Proof. Let us first define for $A, B \in \mathcal{P}_d$,

$$f_{A,B}^t(y) = \mathbb{E}_k \left[(\text{Tr}_{kN} \otimes I_M) \left((\tilde{A} e^{iy\tilde{P}} \tilde{B})(U^N \otimes I_{kM}, U_t^{kN} \otimes I_M, Z^{NM} \otimes I_k) P_{1,2} \right) \right],$$

$$d_n^t(y) = \sup_{\substack{A, B \in \mathcal{P}_d \text{ monomials,} \\ \deg(AB) \leq n \\ 0 \leq s \leq t}} \|f_{A,B}^t(y)\|.$$

Since given a matrix $Z \in \mathbb{M}_{NkM}(\mathbb{C})$, we have

$$\begin{aligned} \|(\text{Tr}_{kN} \otimes I_M)(ZP_{1,2})\| &= \|(\text{Tr}_N \otimes I_M)(I_{NM} \otimes f_2^* \times Z \times I_{NM} \otimes f_1)\| \\ &\leq N \|I_{NM} \otimes f_2^* \times Z \times I_{NM} \otimes f_1\| \\ &\leq N \|Z\|. \end{aligned}$$

Consequently, we can find a constant D such that for any n , $d_n^t(y) \leq D^n$. Note that this constant D can be exponentially large in N or M , indeed it does not matter since in the end we will show that this quantity tends to 0 when k goes to infinity and the other parameters such as N, M or y are fixed. This implies that for a small enough,

$$g_{k,a}^t(y) = \sum_{n \geq 0} d_n^t(y) a^n,$$

is well-defined. But if we set $c_R(P)$ the coefficient associated to the monomial R in P , we have for any $s \leq t$,

$$\left| \frac{df_{A,B}^s(y)}{dy} \right| \leq \sum_{R \text{ monomials}} |c_R(P)| d_{\deg(AB)+\deg(R)}^t(y).$$

Thus if $\deg(AB) \leq n$, we have for any $y \geq 0$,

$$f_{A,B}^s(y) \leq f_{A,B}^s(0) + \sum_{R \text{ monomials}} |c_R(P)| \int_0^y d_{n+\deg(R)}^t(u) du.$$

Thus we have for $y \geq 0$ and any $n \geq 0$,

$$a^n d_n^t(y) \leq a^n d_n^t(0) + \sum_{L \text{ monomials}} |c_R(P)| a^{-\deg(L)} \int_0^y d_{n+\deg(R)}^t(u) a^{n+\deg(L)} du.$$

And with $\|\cdot\|_{a^{-1}}$ defined as in (3.4), we have

$$g_{k,a}^t(y) \leq g_{k,a}^t(0) + \|P\|_{a^{-1}} \int_0^y g_{k,a}^t(u) du.$$

Thanks to Gronwall's inequality, we have for $y \geq 0$,

$$g_{k,a}^t(y) \leq g_{k,a}^t(0) e^{y\|P\|_{a^{-1}}}. \quad (3.21)$$

In order to conclude the proof, we are going to show that

$$g_{k,a}^t(0) = \mathcal{O}(k^{-2}). \quad (3.22)$$

In combination with equation (3.21), it will yields

$$\left| k^{3/2} \mathbb{E}_k \left[(\tau_{kN} \otimes I_M)(q(Y_t^k) P_{1,2}) \right] \right| \leq k^{1/2} a^{-\deg(AB)} g_{k,a}^t(y) = \mathcal{O}(k^{-3/2}).$$

Hence the conclusion. To show equation (3.22), first one can find deterministic matrices $L_j^{u,v} \in \mathbb{M}_N(\mathbb{C})$ such that for every $j \in [p+1, d]$, $Z_j^{NM} = \sum_{1 \leq u,v \leq M} L_j^{u,v} \otimes E_{u,v}$ where $E_{u,v} \in \mathbb{M}_M(\mathbb{C})$ is the matrix whose every coefficient is 0 but the (u,v) coefficients. Thus with $L = (L_j^{u,v})_{j,u,v}$, we proceed by defining

$$V_{N,k}^t = \left(U_t^{kN}, (U_t^{kN})^*, U^N \otimes I_k, (U^N)^* \otimes I_k, L \otimes I_k, L^* \otimes I_k \right), \quad (3.23)$$

$$c_n^t = \sup_{\substack{\deg(Q) \leq n, \\ 0 \leq s \leq t}} \sup_{Q \text{ monomial}} \left| \mathbb{E}_k \left[\text{Tr}_{kN}(Q(V_{N,k}^s) P_{1,2}) \right] \right|.$$

where for the rest of the proof $P_{1,2} = I_N \otimes E_{1,2}$. Then for any $A \in \mathcal{P}_d$ monomials,

$$(\text{Tr}_{kN} \otimes I_M) \left(\tilde{A}(U^N \otimes I_{kM}, U_t^{kN} \otimes I_M, Z^{NM} \otimes I_k) P_{1,2} \right) \quad (3.24)$$

is a linear combination of at most M^{2n} terms of the form

$$\text{Tr}_{kN} \left(A_r(V_{N,k}^t) P_{1,2} \right) \times E_{u,v}, \quad (3.25)$$

where $(A_r)_r$ are monomials. Consequently we have

$$d_n^t(0) \leq M^{2n} c_n^t.$$

Thus if we set

$$f_k^t(a) = \sum_{n \geq 0} c_n^t a^n,$$

we have

$$g_{k,a}^t(0) \leq f_k^t(M^2 a).$$

So all we need to do is to prove that for a small enough, $f_k^t(a) = \mathcal{O}(k^{-2})$. Let Q be a monomial, we define Q_t as the monomial Q evaluated in $V_{N,k}^t$. Thanks to Proposition 3.2.18,

$$\begin{aligned} \frac{d}{dt} \mathbb{E}_k [\mathrm{Tr}_{kN} (Q_t P_{1,2})] &= -\frac{|Q|_B}{2} \mathbb{E}_k [\mathrm{Tr}_{kN} (Q_t P_{1,2})] \\ &\quad - \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i B U_i C} \mathbb{E}_k \left[\mathrm{Tr}_{kN} (A_t U_{i,t}^{kN} C_t P_{1,2}) \mathrm{Tr}_{kN} (B_t U_{i,t}^{kN}) \right] \\ &\quad - \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i^* B U_i^* C} \mathbb{E}_k \left[\mathrm{Tr}_{kN} (A_t U_{i,t}^{kN*} C_t P_{1,2}) \mathrm{Tr}_{kN} (B_t U_{i,t}^{kN*}) \right] \\ &\quad + \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i B U_i^* C} \mathbb{E}_k [\mathrm{Tr}_{kN} (A_t C_t P_{1,2}) \mathrm{Tr}_{kN} (B_t)] \\ &\quad + \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i^* B U_i C} \mathbb{E}_k [\mathrm{Tr}_{kN} (A_t C_t P_{1,2}) \mathrm{Tr}_{kN} (B_t)]. \end{aligned}$$

Since $\mathbb{E}_k [\mathrm{Tr}_{kN} (Q_0 P_{1,2})] = 0$, we have for any t ,

$$\begin{aligned} &\mathbb{E}_k [\mathrm{Tr}_{kN} (Q_t P_{1,2})] \\ &= \int_0^t e^{-\frac{|Q|_B}{2}(t-s)} \left(-\frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i B U_i C} \mathbb{E}_k \left[\mathrm{Tr}_{kN} (A_s U_{i,s}^{kN} C_s P_{1,2}) \mathrm{Tr}_{kN} (B_s U_{i,s}^{kN}) \right] \right. \\ &\quad - \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i^* B U_i^* C} \mathbb{E}_k \left[\mathrm{Tr}_{kN} (A_s U_{i,s}^{kN*} C_s P_{1,2}) \mathrm{Tr}_{kN} (B_s U_{i,s}^{kN*}) \right] \\ &\quad + \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i B U_i^* C} \mathbb{E}_k [\mathrm{Tr}_{kN} (A_s C_s P_{1,2}) \mathrm{Tr}_{kN} (B_s)] \\ &\quad \left. + \frac{1}{kN} \sum_{1 \leq i \leq p, Q=AU_i^* B U_i C} \mathbb{E}_k [\mathrm{Tr}_{kN} (A_s C_s P_{1,2}) \mathrm{Tr}_{kN} (B_s)] \right) ds. \end{aligned}$$

As in Proposition 3.3.2, we consider $(V_t^{kN})_{t \in \mathbb{R}^+}$ and $(W_t^{kN})_{t \in \mathbb{R}^+}$ independent families of p unitary Brownian motions of size kN , independent of $(U_t^{kN})_{t \in \mathbb{R}^+}$. We define $V_{N,k}^{r,1}$ and $V_{N,k}^{r,2}$ as $V_{N,k}^r$ (see (3.23)) but with $V_{s-r}^{kN} U_r^{kN}$ and $W_{s-r}^{kN} U_r^{kN}$ instead of U_r^{kN} . Thanks to Proposition 3.3.2, by polarization and the fact that $(\mathcal{D}_i Q^*)^* = -\mathcal{D}_i Q$, we have with $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$,

$$\begin{aligned} &\mathrm{Cov}_k (\mathrm{Tr}_{kN} (A_s C_s P_{1,2}), \mathrm{Tr}_{kN} (B_s)) \\ &= -\frac{1}{kN} \sum_{i \leq p} \int_0^s \mathbb{E}_k \left[\mathrm{Tr}_{kN} \left((\delta_i(AC) \# P_{1,2}) (V_{N,k}^{r,1}) (\mathcal{D}_i B) (V_{N,k}^{r,2}) \right) \right] dr. \end{aligned}$$

Since $P_{1,2}$ is a matrix of rank N , we now fix $D = \max(1, \sup_{u,v} \|L_j^{u,v}\|)$, we have

$$|\text{Cov}_k(\text{Tr}_{kN}(A_s C_s P_{1,2}), \text{Tr}_{kN}(B_s))| \leq \frac{s}{k} \deg(AC) \deg(B) D^{\deg(ABC)}.$$

We now assume that Q has degree at most n , then $|\text{Cov}_k(\text{Tr}_{kN}(A_s C_s P_{1,2}), \text{Tr}_{kN}(B_s))| \leq \frac{s}{k} n^2 D^n$. Thus we have,

$$\begin{aligned} |\mathbb{E}_k[\text{Tr}_{kN}(Q_t P_{1,2})]| &\leq \frac{n^4 t^2 D^n}{k^2 N} \\ &+ \frac{1}{kN} \int_0^t e^{-\frac{|Q|_B}{2}(t-s)} \left(\sum_{i \leq p, Q=AU_i B U_i C} |\mathbb{E}_k[\text{Tr}_{kN}(A_s U_{i,s}^{kN} C_s P_{1,2})] \mathbb{E}_k[\text{Tr}_{kN}(B_s U_{i,s}^{kN})]| \right. \\ &\quad \sum_{i \leq p, Q=AU_i^* B U_i^* C} |\mathbb{E}_k[\text{Tr}_{kN}(A_s U_{i,s}^{kN*} C_t P_{1,2})] \mathbb{E}_k[\text{Tr}_{kN}(B_s U_{i,s}^{kN*})]| \\ &\quad \sum_{i \leq p, Q=AU_i B U_i^* C} |\mathbb{E}_k[\text{Tr}_{kN}(A_s C_s P_{1,2})] \mathbb{E}_k[\text{Tr}_{kN}(B_s)]| \\ &\quad \left. \sum_{i \leq p, Q=AU_i^* B U_i C} |\mathbb{E}_k[\text{Tr}_{kN}(A_s C_s P_{1,2})] \mathbb{E}_k[\text{Tr}_{kN}(B_s)]| \right) ds. \end{aligned}$$

This means that,

$$\begin{aligned} |\mathbb{E}_k[\text{Tr}_{kN}(Q_t P_{1,2})]| &\leq \frac{n^4 t^2 D^n}{k^2 N} + \int_0^t e^{-\frac{|Q|_B}{2}(t-s)} \left(\sum_{1 \leq i \leq p, Q=AU_i B U_i C} c_{\deg(AC)+1}^t D^{\deg(B)+1} \right. \\ &\quad + \sum_{1 \leq i \leq p, Q=AU_i^* B U_i^* C} c_{\deg(AC)+1}^t D^{\deg(B)+1} \\ &\quad + \sum_{1 \leq i \leq p, Q=AU_i B U_i^* C} c_{\deg(AC)}^t D^{\deg(B)} \\ &\quad \left. + \sum_{1 \leq i \leq p, Q=AU_i^* B U_i C} c_{\deg(AC)}^t D^{\deg(B)} \right) ds \\ &\leq \frac{n^4 t^2 D^n}{k^2 N} + \int_0^t |Q|_B e^{-\frac{|Q|_B}{2}s} ds \sum_{0 \leq d \leq n-1} D^d c_{n-1-d}^t \\ &\leq \frac{n^4 t^2 D^n}{k^2 N} + 2 \sum_{0 \leq d \leq n-1} D^d c_{n-1-d}^t. \end{aligned}$$

Hence, for any $n \geq 1$,

$$c_n^t \leq \frac{n^4 t^2 D^n}{k^2 N} + 2 \sum_{0 \leq d \leq n-1} D^d c_{n-1-d}^t.$$

Since we are taking the trace of $L(V_{N,k}^s) P_{1,2}$ with $P_{1,2} = I_N \otimes E_{1,2}$, we have $c_0 = 0$. We fix $s : a \mapsto \sum_{n \geq 0} \frac{n^4 t^2 (aD)^n}{N}$, thus for a small enough,

$$\begin{aligned} f_k^t(a) &\leq \frac{s(a)}{k^2} + 2 \sum_{n \geq 1} \left(\sum_{0 \leq d \leq n-1} D^d c_{n-1-d}^t \right) a^n \\ &\leq \frac{s(a)}{k^2} + \frac{2a}{1-aD} f_k^t(a) \end{aligned}$$

Thus for a small enough, $f_k^t(a) \leq 2s(a)k^{-2}$. Which means that $f_k^t(a) = \mathcal{O}(k^{-2})$, hence the conclusion. \square

We can now prove the following intermediary lemma that will allow us to derive Lemma 3.4.7. This lemma is the only one where the law of U^N actually plays an important part. To be more precise, we use the invariance of the law of a Haar unitary matrix by multiplication by a deterministic unitary matrix.

Lemma 3.4.6. *We define Y_t^k as in Proposition 3.4.4, we set*

- $P_{l,l'} = I_N \otimes E_{l,l'} \otimes I_M$,
- $q = A e^{iyP} B$.

Then for every $M, N \in \mathbb{N}$, $t \in \mathbb{R}^+$ and $y \in \mathbb{R}$,

$$S_{t,y}^N(A, B) = \lim_{k \rightarrow \infty} -\frac{1}{kN^2} \mathbb{E} \left[\tau_M \left(\sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[(\text{Tr}_{kN} \otimes I_M) \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] - \mathbb{E}_k [\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right) \right]$$

where thanks to Proposition 3.2.7, we set

$$\delta_i q = \delta_i A e^{iyP} B + iyA \int_0^1 e^{i\alpha yP} \delta_i P e^{i(1-\alpha)yP} B d\alpha + A e^{iyP} \delta_i B. \quad (3.26)$$

Proof. Since all of our random variables are unitary matrices, thanks to Proposition 3.4.4 and the dominated convergence theorem,

$$S_{t,y}^N(A, B) = \lim_{k \rightarrow \infty} \mathbb{E} \left[\tau_M \left(\mathbb{E}_k [\tau_{kN} \otimes I_M] \otimes \mathbb{E}_k [\tau_{kN} \otimes I_M] (\delta_i (A e^{iyP} B) (Y_t^k)) \right) \right], \quad (3.27)$$

where $\mathbb{E}_k [\tau_{kN} \otimes I_M] \otimes \mathbb{E}_k [\tau_{kN} \otimes I_M] (A \otimes B (Y_t^k)) = \mathbb{E}_k [\tau_{kN} \otimes I_M (A(Y_t^k))] \mathbb{E}_k [\tau_{kN} \otimes I_M (B(Y_t^k))]$. Since given $V \in \mathbb{U}_N$, $(U_{t,1}^{kN}, U_1^N \otimes I_k, \dots, U_{t,p}^{kN}, U_p^N \otimes I_k)$ has the same law as $((V^* \otimes I_k) U_{t,1}^{kN} (V \otimes I_k), (U_1^N V) \otimes I_k, U_{t,2}^{kN}, U_2^N \otimes I_k, \dots, U_{t,p}^{kN}, U_p^N \otimes I_k)$, we have

$$\mathbb{E}[q(Y_t^k)] = \mathbb{E} \left[\tilde{q} \left((U_1^N \otimes I_k U_{t,1}^{kN}) \otimes I_M (V \otimes I_{kM}), (U_2^N \otimes I_k U_{t,2}^{kN}) \otimes I_M, \dots, (U_p^N \otimes I_k U_{t,p}^{kN}) \otimes I_M, Z^{NM} \otimes I_k \right) \right].$$

Hence let H be an skew-Hermitian matrix, then for any $s \in \mathbb{R}$, $e^{sH} \in \mathbb{U}_N$, thus by taking V this matrix and differentiating with respect to s we get that, $\mathbb{E} \left[\delta_1 q(Y_t^k) \# (H \otimes I_{kM}) \right] = 0$. And similarly we get that for any i ,

$$\mathbb{E} \left[\delta_i q(Y_t^k) \# (H \otimes I_{kM}) \right] = 0.$$

Since every matrix is a linear combination of skew-Hermitian matrices (indeed if $A \in \mathbb{M}_N(\mathbb{C})$, then $2A = (A - A^*) + \mathbf{i} \times (-\mathbf{i})(A^* + A)$), this is true for any matrix $H \in \mathbb{M}_N(\mathbb{C})$, and thus for any i ,

$$\mathbb{E} \left[(\mathrm{Tr}_N \otimes I_{kM})^{\otimes 2} (\delta_i q(Y_t^k)) \right] = \sum_{1 \leq r, s \leq N} g_r^* \otimes I_{kM} \mathbb{E} \left[\delta_i q(Y_t^k) \# (E_{r,s} \otimes I_{kM}) \right] g_s \otimes I_{kM} = 0 \quad (3.28)$$

Let $S, T \in \mathbb{M}_{NkM}(\mathbb{C})$ be deterministic matrices, then

$$\begin{aligned} & \mathrm{Tr}_k \otimes I_M \left((\mathrm{Tr}_N \otimes I_{kM})^{\otimes 2} (S \otimes T) \right) \\ &= \sum_{1 \leq m, n \leq N} \mathrm{Tr}_{Nk} \otimes I_M (S E_{m,n} \otimes I_{kM} T E_{n,m} \otimes I_{kM}) \\ &= \sum_{1 \leq l, l' \leq k} \sum_{1 \leq m \leq N} g_m^* \otimes f_l^* \otimes I_M S g_m \otimes f_{l'} \otimes I_M \sum_{1 \leq n \leq N} g_n^* \otimes f_{l'}^* \otimes I_M T g_n \otimes f_l \otimes I_M \\ &= \sum_{1 \leq l, l' \leq k} \mathrm{Tr}_N \otimes I_M (I_N \otimes f_l^* \otimes I_M S I_N \otimes f_{l'} \otimes I_M) \mathrm{Tr}_N \otimes I_M (I_N \otimes f_{l'}^* \otimes I_M T I_N \otimes f_l \otimes I_M) \\ &= \sum_{1 \leq l, l' \leq k} \mathrm{Tr}_{kN} \otimes I_M (S I_N \otimes E_{l',l} \otimes I_M) \mathrm{Tr}_{kN} \otimes I_M (T I_N \otimes E_{l,l'} \otimes I_M). \end{aligned}$$

Thus by using equation (3.28), we have for any i ,

$$\sum_{1 \leq l, l' \leq k} \mathbb{E} \left[(\mathrm{Tr}_{kN} \otimes I_M)^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) \right] = 0.$$

And consequently,

$$\begin{aligned} & \sum_{1 \leq l, l' \leq k} \mathbb{E} \left[(\mathrm{Tr}_{kN} \otimes I_M)^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) \right] \\ & \quad - \mathbb{E} \left[\mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) \right] \\ &= - \sum_{1 \leq l, l' \leq k} \mathbb{E} \left[\mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) \right]. \end{aligned} \quad (3.29)$$

Let $V, W \in \mathbb{M}_k(\mathbb{C})$ be permutation matrices. Since $I_{NM} \otimes V$ commutes with $Z^{NM} \otimes I_k$ and $U^N \otimes I_{kM}$, and that the law of U_t^{kN} is invariant by conjugation by a unitary matrix, it follows that the law of every matrix of Y_t^k is invariant by conjugation by $I_{NM} \otimes V$ or $I_{NM} \otimes W$. Thus,

$$\mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) = \mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times V P_{l',l} V^* \otimes W P_{l,l'} W^*).$$

Thus by using well-chosen matrices, we get

- if $l = l'$,

$$\mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l,l} \otimes P_{l,l}) = \mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{1,1} \otimes P_{1,1}),$$

- if $l \neq l'$,

$$\mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{l',l} \otimes P_{l,l'}) = \mathbb{E}_k [\mathrm{Tr}_{kN} \otimes I_M]^{\otimes 2} (\delta_i q(Y_t^k) \times P_{1,2} \otimes P_{1,2}).$$

Consequently, we have that

- equation (3.29) is equal to

$$\begin{aligned} & \sum_{1 \leq l, l' \leq k} \mathbb{E} \left[(\text{Tr}_{kN} \otimes I_M) \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] \\ & \quad - \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] \\ & = -k \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{1,1} \otimes P_{1,1}) \right] \\ & \quad - k(k-1) \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{1,2} \otimes P_{1,2}) \right]. \end{aligned}$$

- Whereas the quantity inside the trace τ_M in equation (3.27) is equal to

$$\begin{aligned} & \mathbb{E} \left[\mathbb{E}_k[\tau_{kN} \otimes I_M] \otimes \mathbb{E}_k[\tau_{kN} \otimes I_M] (\delta_i q(Y_t^k)) \right] \\ & = \frac{1}{(kN)^2} \sum_{1 \leq l, l' \leq k} \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] \\ & = \frac{1}{N^2} \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{1,1} \otimes P_{1,1}) \right]. \end{aligned}$$

Thus, we have

$$\begin{aligned} S_{t,y}^N(A, B) & = \lim_{k \rightarrow \infty} -\frac{1}{kN^2} \tau_M \left(\sum_{1 \leq l, l' \leq k} \mathbb{E} \left[(\text{Tr}_{kN} \otimes I_M) \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] \right. \\ & \quad \left. - \mathbb{E} \left[\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{l,l'} \otimes P_{l,l'}) \right] \right) \\ & \quad - \frac{k-1}{N^2} \mathbb{E} \left[\tau_M \left(\mathbb{E}_k[\text{Tr}_{kN} \otimes I_M] \otimes^2 (\delta_i q(Y_t^k) \times P_{1,2} \otimes P_{1,2}) \right) \right]. \end{aligned}$$

Thanks to Lemma 3.4.5 and Proposition 3.2.7, the last term converges towards 0, which gives the expected formula. \square

Lemma 3.4.6 makes a covariance appears. Thus it is natural to want to use Corollary 3.3.3 to get an upper bound of $S_{t,y}^N(A, B)$, explicit in all of its parameters.

Lemma 3.4.7. *There exists a polynomial $L_P \in \mathbb{R}^+[X]$ such that for any t, y, N, M, Z^{NM} ,*

$$|S_{t,y}^N(A, B)| \leq L_P \left(\|Z^{NM}\| \right) \frac{M^2}{N^2} (1 + |y|^3) t. \quad (3.30)$$

Besides if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, then we have the same inequality without the M^2 .

Proof. As mentioned in equation (3.26), we have

$$\delta_i q = \delta_i A e^{iyP} B + iyA \int_0^1 e^{i\alpha yP} \delta_i P e^{i(1-\alpha)yP} B d\alpha + A e^{iyP} \delta_i B.$$

Consequently, we set $q_1 = A_1 e^{i\alpha y^P} B_1$ and $q_2 = A_2 e^{i(1-\alpha)y^P} B_2$ where $A_1, B_1, A_2, B_2 \in \mathcal{P}_d$ are monomials, then thanks to equation (3.5) and Proposition 3.2.7, we can use Lemma 3.3.3 even though q_1 and q_2 are not exactly polynomials and we get that

$$\begin{aligned} & \tau_M \left(\mathbb{E}_k \left[(\text{Tr}_{kN} \otimes I_M)^{\otimes 2} \left(q_1(Y_t^k) P_{\nu,l} \otimes q_2(Y_t^k) P_{l,\nu} \right) \right] \right. \\ & \quad \left. - \mathbb{E}_k [\text{Tr}_{kN} \otimes I_M]^{\otimes 2} \left(q_1(Y_t^k) P_{\nu,l} \otimes q_2(Y_t^k) P_{l,\nu} \right) \right) \\ &= \sum_{1 \leq j \leq p} \int_0^t \mathbb{E}_k \left[\tau_{kN} \otimes \tau_M \left(h \circ \delta_j(\tilde{q}_1 P_{\nu,l})(U^N \otimes I_k V_{t-s}^{kN} U_s^{kN}, Z^{NM} \otimes I_M) \right. \right. \\ & \quad \left. \left. \times h \circ \delta_j(\tilde{q}_2 P_{\nu,l})(U^N \otimes I_k W_{t-s}^{kN} U_s^{kN}, Z^{NM} \otimes I_M) \right) \right] ds, \end{aligned}$$

where $(V_s^{kN})_{s \in \mathbb{R}}$ and $(W_s^{kN})_{s \in \mathbb{R}}$ are p -tuples of independent unitary Brownian motions of size kN . Thus thanks to Lemma 3.3.1, we get that there exist a polynomial $L_{A_1, B_1, A_2, B_2, P}$ such that

$$\begin{aligned} & \left| \tau_M \left(\mathbb{E}_k \left[(\text{Tr}_{kN} \otimes I_M)^{\otimes 2} \left(q_1(Y_t^k) P_{\nu,l} \otimes q_2(Y_t^k) P_{l,\nu} \right) \right] \right. \right. \\ & \quad \left. \left. - \mathbb{E}_k [\text{Tr}_{kN} \otimes I_M]^{\otimes 2} \left(q_1(Y_t^k) P_{\nu,l} \otimes q_2(Y_t^k) P_{l,\nu} \right) \right) \right| \\ & \leq L_{A_1, B_1, A_2, B_2, P} \left(\|Z^{NM}\| \right) \times (1 + y^2) \frac{M^2 t}{k}, \end{aligned}$$

where we used the fact that $P_{\nu,l}$ has rank NM and that the renormalized trace of a matrix of rank NM in $\mathbb{M}_{kN}(\mathbb{C}) \otimes \mathbb{M}_M(\mathbb{C})$ is smaller than its norm renormalized by k .

Since this upper bound does not depend on α , it remains true if we integrate with respect to α from 0 to 1. But then $\delta_i q$ is a finite linear combination of such terms. Consequently, one gets that there exists a polynomial L_P such that for any k ,

$$\begin{aligned} & \left| \frac{1}{kN^2} \tau_M \left(\sum_{1 \leq l, l' \leq k} \mathbb{E} \left[(\text{Tr}_{kN} \otimes I_M)^{\otimes 2} \left(\delta_i q(Y_t^k) \times P_{\nu,l} \otimes P_{l,\nu} \right) \right] \right. \right. \\ & \quad \left. \left. - \mathbb{E} \left[\mathbb{E}_k [\text{Tr}_{kN} \otimes I_M]^{\otimes 2} \left(\delta_i q(Y_t^k) \times P_{\nu,l} \otimes P_{l,\nu} \right) \right] \right) \right| \\ & \leq L_P \left(\|Z^{NM}\| \right) \times (1 + |y|^3) \frac{M^2 t}{N^2}. \end{aligned}$$

Finally, when the matrices Z^{NM} commute, as specified in Lemma 3.3.3, we have the same proof where we replaced $h \circ \delta_j$ by \mathcal{D}_j , hence we do not need to use Lemma 3.3.1 and hence we have the same inequality without the M^2 . Finally we get the conclusion thanks to Lemma 3.4.6. \square

We now have the tools to prove Theorem 3.4.1.

Proof of Theorem 3.4.1. Thanks to Lemma 3.4.2, and since $\mathcal{D}_i e^{iy^P} = iy \delta_i P \tilde{\#} e^{iy^P}$, there exist a family of monomials $(A_k, B_k)_k$ and a constant C which only depends on P such that,

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{MN} \left(f \left(\tilde{P} \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) - \tau_N \otimes \tau_M \left(f \left(\tilde{P} \left(U^N u_T \otimes I_M, Z^{NM} \right) \right) \right) \right] \right| \\ & \leq C \sum_k \int |y| \int_0^T |S_{t,y}^N(A_k, B_k)| dt d|\mu|(y). \end{aligned}$$

Thanks to equation (3.30), we get that for some polynomial L_P ,

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{MN} \operatorname{Tr}_{MN} \left(f \left(\tilde{P} \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) - \tau_N \otimes \tau_M \left(f \left(\tilde{P} \left(U^N u_T \otimes I_M, Z^{NM} \right) \right) \right) \right] \right| \\ & \leq T^2 \frac{M^2}{N^2} L_P \left(\|Z^{NM}\| \right) \times \int_{\mathbb{R}} (|y| + y^4) d|\mu|(y). \end{aligned}$$

And besides if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, we have the same inequality without the M^2 . Finally, thanks to Proposition 3.3.5, thanks to Duhamel's formula (3.6) we can find a polynomial L'_P such that

$$\begin{aligned} & \left| \tau_N \otimes \tau_M \left(e^{\mathbf{i}y \tilde{P}(u \otimes I_M, Z^{NM})} \right) - \tau_N \otimes \tau_M \left(e^{\mathbf{i}y \tilde{P}(U^N u_T \otimes I_M, Z^{NM})} \right) \right| \\ & = \left| \tau_N \otimes \tau_M \left(e^{\mathbf{i}y \tilde{P}(U^N u \otimes I_M, Z^{NM})} \right) - \tau_N \otimes \tau_M \left(e^{\mathbf{i}y \tilde{P}(U^N u_T \otimes I_M, Z^{NM})} \right) \right| \\ & \leq e^{-T/2} L'_P \left(\|Z^{NM}\| \right) \times |y|. \end{aligned}$$

Hence the conclusion by fixing $T = 4 \ln(N)$. □

We can finally prove Theorem 3.1.1.

Proof of Theorem 3.1.1. We want to use Theorem 3.4.1. To do so we would like to take the Fourier transform of f and use Fourier inversion formula. However we did not assume that f is integrable. Thus the first step of the proof is to show that we can assume that f has compact support. Since U^N and u are unitaries, there exists a polynomial $H \in \mathbb{R}^+[X]$ which only depends on P such that $\|\tilde{P}(U^N \otimes I_M, Z^{NM})\| \leq H(\|Z^{NM}\|)$. Consequently since we also have that $\|\tilde{P}(u \otimes I_M, Z^{NM})\| \leq H(\|Z^{NM}\|)$, we can replace f by fg where g is a \mathcal{C}^∞ function which takes value in $[0, 1]$, takes value 1 in $[-H(\|Z^{NM}\|), H(\|Z^{NM}\|)]$ and 0 outside of $[-H(\|Z^{NM}\|) - 1, H(\|Z^{NM}\|) + 1]$. Since f can be differentiated six times, we can take its Fourier transform and then invert it so that with the convention $\hat{f}(y) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x) e^{-ixy} dx$, we have

$$\forall x \in \mathbb{R}, \quad f(x) = \int_{\mathbb{R}} e^{ixy} \hat{f}(y) dy.$$

Besides since if f has compact support bounded by $H(\|Z^{NM}\|) + 1$, then

$$\|\hat{f}\|_\infty \leq 2 \left(H(\|Z^{NM}\|) + 1 \right) \|f\|_\infty,$$

we get that

$$\begin{aligned}
 \int_{\mathbb{R}} (|y| + y^4) |\widehat{f}(y)| dy &\leq \int_{\mathbb{R}} \frac{|y| + |y|^3 + y^4 + y^6}{1 + y^2} |\widehat{f}(y)| dy \\
 &\leq \int_{\mathbb{R}} \frac{|\widehat{(f)^{(1)}}(y)| + |\widehat{(f)^{(3)}}(y)| + |\widehat{(f)^{(4)}}(y)| + |\widehat{(f)^{(6)}}(y)|}{1 + y^2} dy \\
 &\leq 2 \left(H(\|Z^{NM}\|) + 1 \right) \|f\|_{C^6} \int_{\mathbb{R}} \frac{1}{1 + y^2} dy \\
 &\leq 2\pi \left(H(\|Z^{NM}\|) + 1 \right) \|f\|_{C^6},
 \end{aligned}$$

Hence it satisfies the hypothesis of Theorem 3.4.1 with $\mu(dy) = \widehat{f}(y)dy$, thus we have

$$\begin{aligned}
 &\left| \mathbb{E} \left[\frac{1}{MN} \text{Tr} \left(f \left(\tilde{P} \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] - \tau \left(f \left(\tilde{P} \left(u \otimes I_M, Z^{NM} \right) \right) \right) \right| \\
 &\leq \frac{M^2 \ln^2(N)}{N^2} L_P \left(\|Z^{NM}\| \right) \int_{\mathbb{R}} (|y| + y^4) |\widehat{f}(y)| dy \\
 &\leq \frac{M^2 \ln^2(N)}{N^2} \times 2\pi L_P \left(\|Z^{NM}\| \right) \left(H(\|Z^{NM}\|) + 1 \right) \|f\|_{C^6}.
 \end{aligned}$$

And finally, if $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_q^M)$ and that these matrices commute, then we have the same inequality without the M^2 as specified in Theorem 3.4.1. \square

3.5 Proof of Corollaries

3.5.1 Proof of Corollary 3.1.3

We could directly apply Theorem 3.1.1 to $f_z : x \rightarrow (z - x)^{-1}$, however for z such that $\Im z$ is small, we have $\|f\|_{C^6} = O((\Im z)^{-7})$ when we want $O((\Im z)^{-5})$ instead. Since P is self-adjoint, $\overline{G_P(z)} = G_P(\bar{z})$, thus we can assume that $\Im z < 0$, but then

$$f_z(x) = \int_0^\infty e^{ixy} (\mathbf{i}e^{-iyz}) dy.$$

Consequently with $\mu_z(dy) = \mathbf{i}e^{-iyz} dy$, we have

$$\int_0^\infty (y + y^4) d|\mu_z|(y) = \frac{1}{|\Im z|^2} + \frac{24}{|\Im z|^5}.$$

Thus by applying Theorem 3.4.1 with $Z^{NM} = (I_N \otimes Y_1^M, \dots, I_N \otimes Y_p^M)$, P and f_z , we have

$$\left| \mathbb{E} \left[G_{P(U^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(u \otimes I_M, 1 \otimes Y^M)}(z) \right| \leq \frac{M^2 \ln^2(N)}{N^2} L_P \left(\|Z^{NM}\| \right) \int_{\mathbb{R}} (1+y^4) d|\mu_z|(y).$$

Now since $\|Z^{NM}\| = \|Y^M\|$ which does not depend on N , we finally have

$$\left| \mathbb{E} \left[G_{P(U^N \otimes I_M, I_N \otimes Y^M)}(z) \right] - G_{P(u \otimes I_M, 1 \otimes Y^M)}(z) \right| \leq \frac{M^2 \ln^2(N)}{N^2} L_P \left(\|Y^M\| \right) \left(\frac{1}{|\Im z|^2} + \frac{24}{|\Im z|^5} \right).$$

3.5.2 Proof of Corollary 3.1.4

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz function uniformly bounded by 1 and with Lipschitz constant at most 1, we want to find an upper bound on

$$\left| \mathbb{E} \left[\frac{1}{MN} \text{Tr}_{NM} \left(f \left(P \left(U^N \otimes I_M, I_N \otimes Y_M \right) \right) \right) \right] - \tau \otimes \tau_M \left(f \left(P \left(u \otimes I_M, 1 \otimes Y_M \right) \right) \right) \right|. \quad (3.31)$$

Firstly, since U^N are unitary matrices, we can assume that the support of f is bounded by a constant $S = H(\|Y^M\|)$ for some polynomial $H \in \mathbb{R}^+[X]$ independent of everything. However we cannot apply directly Theorem 3.1.1 since f is not regular enough. In order to deal with this issue we use the convolution with gaussian random variable, thus let G be a centered gaussian random variable, we set

$$f_\varepsilon : x \rightarrow \mathbb{E}[f(x + \varepsilon G)].$$

Since f has Lipschitz constant 1, we have for any $x \in \mathbb{R}$,

$$|\mathbb{E}[f(x + \varepsilon G)] - f(x)| \leq \varepsilon.$$

Since f_ε is regular enough we could now apply Theorem 3.1.1, however we get a better result by using Theorem 3.4.1. Indeed we have

$$\begin{aligned} f_\varepsilon(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x + \varepsilon y) e^{-y^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) \frac{e^{-\frac{(x-y)^2}{2\varepsilon^2}}}{\varepsilon} dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} f(y) \int_{\mathbb{R}} e^{i(x-y)u} e^{-(u\varepsilon)^2/2} du dy. \end{aligned}$$

Since the support of f is bounded, we can apply Fubini's theorem:

$$f_\varepsilon(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iux} \int_{\mathbb{R}} f(y) e^{-iyu} dy e^{-(u\varepsilon)^2/2} du.$$

And so with the convention $\hat{h}(u) = \frac{1}{2\pi} \int_{\mathbb{R}} h(y) e^{-iyu} dy$, we have

$$f_\varepsilon(x) = \int_{\mathbb{R}} e^{iux} \hat{f}(u) e^{-(u\varepsilon)^2/2} du.$$

Thus if we set $\mu_\varepsilon(dy) = \hat{f}(y) e^{-(y\varepsilon)^2/2} dy$, then, since $\|f\|_\infty \leq 1$,

$$\int_{\mathbb{R}} (1 + y^4) d|\mu_\varepsilon|(y) \leq 2S \int_{\mathbb{R}} (1 + y^4) e^{-y^2/2} dy \varepsilon^{-5}.$$

Consequently we can apply Theorem 3.4.1 with f_ε and since $\|f - f_\varepsilon\|_\infty \leq \varepsilon$, there exists a polynomial R_P such that (3.31) can be bounded by

$$2\varepsilon + R_P(\|Y^M\|) \frac{M^2 \ln^2(N)}{N^2 \varepsilon^5}.$$

Thus we can now fix $\varepsilon = (N^{-1} \ln(N))^{1/3}$ and we get that for any f Lipschitz function uniformly bounded by 1 and with Lipschitz constant at most 1, (3.31) can be bounded by

$$2R_P(\|Y^M\|) M^2 \left(\frac{\ln N}{N} \right)^{1/3}.$$

3.5.3 Proof of Theorem 3.1.2

Firstly, we need to set the operator norm of tensor of \mathcal{C}^* -algebras we will work with. When writing the proof it appears that it is the minimal tensor product as defined in 3.2.6. The following two lemmas were used in [7], see Lemma 4.1.8 from [75] for a proof of the first one and Lemma 4.3 from [7] for the second one. In order to learn more about the second lemma, especially how to weaken the hypothesis, we refer to [31].

Lemma 3.5.1. *Let $(\mathcal{A}, \tau_{\mathcal{A}})$ and $(\mathcal{B}, \tau_{\mathcal{B}})$ be \mathcal{C}^* -algebras with faithful traces, then $\tau_{\mathcal{A}} \otimes \tau_{\mathcal{B}}$ extends uniquely to a faithful trace $\tau_{\mathcal{A}} \otimes_{\min} \tau_{\mathcal{B}}$ on $\mathcal{A} \otimes_{\min} \mathcal{B}$.*

Lemma 3.5.2. *Let \mathcal{C} be an exact \mathcal{C}^* -algebra endowed with a faithful state $\tau_{\mathcal{C}}$, let $Y^N \in \mathcal{A}_N$ be a sequence of family of noncommutative random variable in a \mathcal{C}^* -algebra \mathcal{A}_N which converges strongly towards a family Y in a \mathcal{C}^* -algebra \mathcal{A} endowed with a faithful state $\tau_{\mathcal{A}}$. Let $S \in \mathcal{C}$ be a family of noncommutative random variable, then the family $(S \otimes 1, 1 \otimes Y^N)$ converges strongly in distribution towards the family $(S \otimes 1, 1 \otimes Y)$.*

In order to prove Theorem 3.1.2 we use well-known concentration properties of unitary Haar matrices coupled with an estimation of the expectation, let us begin by stating the concentration properties that we will use.

Proposition 3.5.3. *Let f be a continuous function on \mathbb{U}_N^p , such that for any $X, Y \in \mathbb{U}_N^p$,*

$$|f(X) - f(Y)| \leq C \sum_i \text{Tr}_N((X_i - Y_i)(X_i - Y_i)^*)^{1/2}.$$

Then if W is a family of p independent random matrices distributed according to the Haar measure on SU_N , and U a family of p independent unitary Haar matrices of size N independent from W , we have,

$$\mathbb{P}(|f(U) - \mathbb{E}_W[f(WU)]| \geq \delta) \leq 4p e^{-\left(\frac{\delta}{2pC}\right)^2 N}.$$

Proof. We want to use Corollary 4.4.28 from [13], in order to do so let us first assume that f takes real values. We then set,

$$f_{U_{i+1}, \dots, U_p}^i : U_i \rightarrow \mathbb{E}_{W_1, \dots, W_{i-1}} [f(W_1 U_1, \dots, W_{i-1} U_{i-1}, U_i, U_{i+1}, \dots, U_p)].$$

Thus,

$$f(U) - \mathbb{E}_W [f(WU)] = \sum_{1 \leq i \leq p} f_{U_{i+1}, \dots, U_p}^i(U_i) - \mathbb{E}_{W_i} [f_{U_{i+1}, \dots, U_p}^i(W_i U_i)].$$

Besides for any U_i, V_i , we have that

$$|f_{U_{i+1}, \dots, U_p}^i(U_i) - f_{U_{i+1}, \dots, U_p}^i(V_i)| \leq C \operatorname{Tr}_N ((U_i - V_i)(U_i - V_i)^*)^{1/2}.$$

Thus thanks to Corollary 4.4.28 from [13] we have that,

$$\begin{aligned} \mathbb{P} (|f(U) - \mathbb{E}_W [f(WU)]| \geq \delta) &\leq \sum_i \mathbb{P} \left(\left| f_{U_{i+1}, \dots, U_p}^i(U_i) - \mathbb{E}_{Y_i} [f_{U_{i+1}, \dots, U_p}^i(W_i U_i)] \right| \geq \frac{\delta}{p} \right) \\ &\leq 2p e^{-\left(\frac{\delta}{pC}\right)^2 N}. \end{aligned}$$

Finally we conclude by taking the real and imaginary part of f . □

We can now prove the concentration inequality that we will use in the rest of this paper. To simplify notations we will write M instead of M_N . We also set $Z^{NM} = (Z^N \otimes I_M, I_N \otimes Y^M)$ and $Z = (z \otimes 1, 1 \otimes y)$.

Proposition 3.5.4. *Let $P \in \mathcal{P}_d$, there are polynomials $H_P, K_P \in \mathbb{R}^+[X]$ which only depends on P such that for any N, M ,*

$$\mathbb{P} \left(\left| \left\| \tilde{P}(U^N \otimes I_M, Z^{NM}) \right\| - \mathbb{E} \left[\left\| \tilde{P}(U^N \otimes I_M, Z^{NM}) \right\| \right] \right| \geq \delta + \frac{K_P(\|Z^{NM}\|)}{N} \right) \leq e^{-\frac{\delta^2 N}{H_P(\|Z^{NM}\|)}},$$

where $\|Z^{NM}\| = \sup_i \|Z_i^{NM}\|$.

Proof. We set $G_N : X \mapsto \left\| \tilde{P}(X \otimes I_M, Z^{NM}) \right\|$. One can find a polynomial $L_P \in \mathbb{R}^+[X]$ such that for any N and Z^{NM} ,

$$|G_N(X) - G_N(Y)| \leq L_P \left(\|Z^{NM}\| \right) \sum_i \|X_i - Y_i\|,$$

where $\|\cdot\|$ is the operator norm. Besides

$$\sum_i \|X_i - Y_i\| \leq \sum_i \operatorname{Tr}_N ((X_i - Y_i)^*(X_i - Y_i))^{1/2}.$$

Hence with Proposition 3.5.3, there is a polynomial $H_P \in \mathbb{R}^+[X]$ which only depends on P such that for any N, M ,

$$\mathbb{P} \left(\left| \left\| \tilde{P}(U^N \otimes I_M, Z^{NM}) \right\| - \mathbb{E}_W \left[\left\| \tilde{P}(WU^N \otimes I_M, Z^{NM}) \right\| \right] \right| \geq \delta \right) \leq e^{-\frac{\delta^2 N}{H_P(\|Z^{NM}\|)}}.$$

Besides for any matrix $U \in \mathbb{U}_N$, there exist $S \in SU_N$ and $\theta \in [0, 2\pi]$ such that $U = e^{i\frac{\theta}{N}}S$. Indeed we just have to pick θ such that $e^{i\theta} = \det(U)$. Thus there is a polynomial K_P such that

$$\left| \mathbb{E}_W \left[\left\| \tilde{P}(WU^N \otimes I_M, Z^{NM}) \right\| \right] - \mathbb{E} \left[\left\| \tilde{P}(U^N \otimes I_M, Z^{NM}) \right\| \right] \right| \leq \frac{K_P(\|Z^{NM}\|)}{N}.$$

This concludes the proof. \square

We can now prove Theorem 3.1.2. Firstly, we can assume that Z^N and Y^M are deterministic matrices by Fubini's theorem. The convergence in distribution is a well-known theorem, we refer to [13], Theorem 5.4.10. We set g a function of class \mathcal{C}^∞ which takes value 0 on $(-\infty, 1/2]$ and value 1 on $[1, \infty)$, and belongs to $[0, 1]$ otherwise. Let us define $f_\varepsilon : t \mapsto g\left(\varepsilon^{-1}\left(t - \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\| \right)\right)$. By Theorem 3.1.1, there exists a constant C which only depends on P , $\sup_M \|Y^M\|$ and $\sup_N \|Z^N\|$ (which is finite thanks to the strong convergence assumption on Y^M and Z^N) such that,

$$\left| \mathbb{E} \left[\text{Tr}_{MN} \left(f_\varepsilon \left(\tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] - MN \tau_N \otimes \tau_M \left(f_\varepsilon \left(\tilde{P}\tilde{P}^* \left(u \otimes I_M, Z^{NM} \right) \right) \right) \right| \leq C \varepsilon^{-6} \frac{M^3 \ln^2(N)}{N}.$$

According to Theorem A.1 from [28], $(u, Z^N)_{N \geq 1}$ converges strongly in distribution towards (u, z) since, given a system of free semi-circular variable, we can write $u_i = f(x_i)$ for a specific function f built with the help of Lemma 3.1 of [9]. Besides thanks to Lemma 3.5.2 and Corollary 17.10 from [78], we have that $(u \otimes I_M, 1 \otimes Y^M)_{M \geq 1}$ converges strongly in distribution towards $(u \otimes 1, 1 \otimes y)$. In Theorem 3.1.2, we are interested in the situation where $Z^{NM} = Z^N \otimes I_M$ or $Z^{NM} = I_N \otimes Y^M$. So, without loss of generality, we restrict ourselves to this kind of Z^{NM} . We know that $(u \otimes I_M, Z^{NM})$ converges strongly towards $(u \otimes 1, Z)$, but since the support of f_ε is disjoint from the spectrum of $\tilde{P}\tilde{P}^*(u \otimes 1, Z)$, thanks to Proposition 3.2.2, for N large enough, $\tau_N \otimes \tau_M \left(f_\varepsilon \left(\tilde{P}\tilde{P}^* \left(u \otimes I_M, Z^{NM} \right) \right) \right) = 0$ and therefore,

$$\mathbb{E} \left[\text{Tr}_{MN} \left(f_\varepsilon \left(\tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \right] \leq C \varepsilon^{-6} \frac{M^3 \ln^2(N)}{N}. \quad (3.32)$$

Hence, we deduce for N large enough,

$$\begin{aligned} & \mathbb{E} \left[\left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \right] - \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\| \\ & \leq \varepsilon + \int_\varepsilon^\infty \mathbb{P} \left(\left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \geq \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\| + \alpha \right) d\alpha \\ & \leq \varepsilon + \int_\varepsilon^K \mathbb{P} \left(\text{Tr}_{NM} \left(f_\alpha \left(\tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right) \right) \geq 1 \right) d\alpha \\ & \leq \varepsilon + C' \varepsilon^{-6} \frac{M^3 \ln^2(N)}{N}. \end{aligned}$$

Finally we get that,

$$\limsup_{N \rightarrow \infty} \mathbb{E} \left[\left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \right] \leq \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\|.$$

Thanks to Proposition 3.5.4, by taking $\delta_N = N^{-1/4}$ and using Borel-Cantelli lemma, we get that almost surely,

$$\lim_{N \rightarrow \infty} \left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| - \mathbb{E} \left[\left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \right] = 0$$

And consequently almost surely,

$$\limsup_{N \rightarrow \infty} \left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \leq \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\|.$$

Besides, we know thanks to Theorem 5.4.10 of [13] that if h is a continuous function taking positive values on $\left(\left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\| - \varepsilon, \infty \right)$ and taking value 0 elsewhere. Then

$$\frac{1}{MN} \text{Tr}_{MN}(h(\tilde{P}\tilde{P}^*(U^N \otimes I_M, Z^{NM})))$$

converges almost surely towards $\tau_{\mathcal{A}} \otimes_{\min} \tau_{\mathcal{B}}(h(\tilde{P}\tilde{P}^*(u \otimes 1, Z)))$. If this quantity is positive, then almost surely for N large enough so is $\frac{1}{MN} \text{Tr}_{MN}(h(\tilde{P}\tilde{P}^*(U^N \otimes I_M, Z^{NM})))$, thus

$$\left\| \tilde{P}\tilde{P}^*(U^N \otimes I_M, Z^{NM}) \right\| \geq \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\| - \varepsilon.$$

Since h is non-negative and the intersection of the support of h with the spectrum of $\tilde{P}\tilde{P}^*(u \otimes 1, Z)$ is non-empty, we have that $h(\tilde{P}\tilde{P}^*(u \otimes 1, Z)) \geq 0$ and is not 0. Besides, we know that the trace on the space where z is defined is faithful, and so is the trace on the \mathcal{C}^* -algebra generated by a single semicircular variable, hence by Theorem 3.2.3, so is $\tau_{\mathcal{A}}$. Thus, since both $\tau_{\mathcal{A}}$ and $\tau_{\mathcal{B}}$ are faithful, by Lemma 3.5.1, so is $\tau_{\mathcal{A}} \otimes_{\min} \tau_{\mathcal{B}}$ and $\tau_{\mathcal{A}} \otimes_{\min} \tau_{\mathcal{B}}(h(\tilde{P}\tilde{P}^*(u \otimes 1, Z))) > 0$. As a consequence, almost surely,

$$\liminf_{N \rightarrow \infty} \left\| \tilde{P}\tilde{P}^* \left(U^N \otimes I_M, Z^{NM} \right) \right\| \geq \left\| \tilde{P}\tilde{P}^*(u \otimes 1, Z) \right\|.$$

We finally conclude thanks to the fact that for any y in a \mathcal{C}^* -algebra, $\|yy^*\| = \|y\|^2$.

Acknowledgements

The author would like to thanks his PhD supervisors Benoît Collins and Alice Guionnet for proofreading this paper and their continuous help, as well as Mikael de la Salle for its advices and the proof of Lemma 3.3.1. The author was partially supported by a MEXT JASSO fellowship and Labex Milyon (ANR-10-LABX-0070) of Université de Lyon.

Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid GUE matrices

Let X^N be a family of $N \times N$ independent GUE random matrices, Z^N a family of deterministic matrices, P a self-adjoint noncommutative polynomial, that is for any N , $P(X^N, Z^N)$ is self-adjoint, f a smooth function. We prove that for any k , if f is smooth enough, there exist deterministic constants $\alpha_i^P(f, Z^N)$ such that

$$\mathbb{E} \left[\frac{1}{N} \text{Tr} \left(f(P(X^N, Z^N)) \right) \right] = \sum_{i=0}^k \frac{\alpha_i^P(f, Z^N)}{N^{2i}} + \mathcal{O}(N^{-2k-2}).$$

Besides the constants $\alpha_i^P(f, Z^N)$ are built explicitly with the help of free probability. In particular, if x is a free semicircular system, then when the support of f and the spectrum of $P(x, Z^N)$ are disjoint, for any i , $\alpha_i^P(f, Z^N) = 0$. As a corollary, we prove that given $\alpha < 1/2$, for N large enough, every eigenvalue of $P(X^N, Z^N)$ is $N^{-\alpha}$ -close from the spectrum of $P(x, Z^N)$.

This chapter is adapted from [10].

4.1 Introduction

Asymptotic expansions in Random Matrix Theory created bridges between different worlds, including topology, statistical mechanics, and quantum field theory. In mathematics, a breakthrough was made in 1986 in [33] by Harer and Zagier who used the large dimension expansion of the moments of Gaussian matrices to compute the Euler characteristic of the moduli space of curves. A good introduction to this topic is given in the survey [34] by Zvonkin. In physics, the seminal works of t'Hooft [35] and Brézin, Parisi, Itzykson and Zuber [36] related matrix models with the enumeration of maps of any genus, hence providing a purely analytical tool to solve these hard combinatorial problems. Considering matrices in interaction via a potential, the so-called matrix models, indeed allows to consider the enumeration of maps with several vertices, including a possible coloring of the edges when the matrix model contains several matrices. This relation allowed to associate matrix models to statistical models on random graphs [37, 38, 39, 40, 41], as well as in [42] and [43] for the unitary case. This was also extended to the so-called β -ensembles in [44, 45, 46, 47, 48, 49]. Among other objects, these works study correlation functions and the so-called free energy and show that they expand as power series in the inverse of the dimension, and the coefficients of these expansions enumerate maps sorted by their genus. To compute asymptotic expansions, often referred to in the literature as topological expansions, one of the most successful methods is the loop equations method, see [50] and [51]. Depending on the model of random matrix, those are Tutte's equations, Schwinger-Dyson equations, Ward identities, Virasoro constraints, W-algebra or simply integration by part. This method was refined and used repeatedly in physics, see for example the work of Eynard and his collaborators, [52, 53, 54, 55]. At first those equations were only solved for the first few orders, however in 2004, in [53] and later [56] and [57], this method was refined to push the expansion to any orders recursively [58].

In this paper we want to generalize Harer-Zagier expansion for the moments of Gaussian matrices to more general smooth functions. Instead of a single GUE matrix, we will consider several independent matrices and deterministic matrices. We repeatedly use Schwinger-Dyson equations associated to GUE matrices to carry out our estimates. While we do not use the link between the coefficients of our expansion and map enumeration, as a corollary we get a new expression of these combinatorial objects. We show that the number of colored maps of genus g with a single specific vertex can be expressed as an integral, see remark 4.3.8 for a precise statement.

Most papers quoted above have in common that they deal with polynomials or exponentials of polynomial evaluated in random matrices. With a few exceptions, such as [37] and [59], smooth functions have not been considered. However being able to work with such functions is important for the applications. In particular we need to be able to work with functions with compact support to prove strong convergence results, that is proving the convergence of the spectrum for the Hausdorff distance. In this paper we establish a finite expansion of any orders around the dimension of the random matrix for the trace of smooth functions evaluated in polynomials in independent GUE random matrices. We refer to Definition 4.2.16 for a definition of those objects. The link between maps and topological expansion is a good motivation to prove such kind of theorem. Another motivation is to study the spectrum of polynomials in these random matrices: because we consider general smooth functions, our expansion will for instance allow to study the spectrum outside of the limiting bulk. In the case of a single GUE matrix, we have an explicit formula for the distribution of the eigenvalues of those random

matrices, see Theorem 2.5.2 of [13]. However, if we consider polynomials in independent GUE matrices, we have no such result. The first result in this direction dates back to 1991 when Voiculescu proved in [3] that the renormalized trace of such polynomials converges towards a deterministic limit $\alpha(P)$. In particular given X_1^N, \dots, X_d^N independent GUE matrices, the following holds true almost surely:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{Tr}_N \left(P(X_1^N, \dots, X_d^N) \right) = \alpha(P). \quad (4.1)$$

Voiculescu computed the limit $\alpha(P)$ with the help of free probability. Besides if A_N is a self-adjoint matrix of size N , then one can define the empirical measure of its (real) eigenvalues by

$$\mu_{A_N} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i},$$

where δ_λ is the Dirac mass in λ and $\lambda_1, \dots, \lambda_N$ are the eigenvalue of A_N . In particular, if P is a self-adjoint polynomial, that is such that for any self adjoint matrices A_1, \dots, A_d , $P(A_1, \dots, A_d)$ is a self-adjoint matrix, then one can define the random measure $\mu_{P(X_1^N, \dots, X_d^N)}$. In this case, Voiculescu's result (4.1) implies that there exists a measure μ_P with compact support such that almost surely $\mu_{P(X_1^N, \dots, X_d^N)}$ converges weakly towards μ_P : it is given by $\mu_P(x^k) = \alpha(P^k)$ for all integer numbers k . Consequently, assuming we can apply Portmanteau theorem, the proportion of eigenvalues of $A_N = P(X_1^N, \dots, X_d^N)$ in the interval $[a, b]$, that is $\mu_{A_N}([a, b])$, converges towards $\mu_P([a, b])$.

Therefore in order to study the eigenvalues of a random matrix one has to study the renormalized trace of its moments. However if instead of studying the renormalized trace of polynomials in A_N , we study the non-renormalized trace of smooth function in A_N , then we can get precise information on the location of the eigenvalues. It all comes from the following remark, let f be a non-negative function such that f is equal to 1 on the interval $[a, b]$, then if $\sigma(A_N)$ is the spectrum of A_N ,

$$\mathbb{P} \left(\sigma(A_N) \cap [a, b] \neq \emptyset \right) \leq \mathbb{P} \left(\operatorname{Tr}_N (f(A_N)) \geq 1 \right) \leq \mathbb{E} \left[\operatorname{Tr}_N (f(A_N)) \right].$$

Thus if one can show that the right-hand side of this inequality converges towards zero when N goes to infinity, then asymptotically there is no eigenvalue in the segment $[a, b]$. In the case of the random matrices that we study in this paper, that is polynomials in independent GUE matrices, a breakthrough was made in 2005 by Haagerup and Thorbjørnsen in [6]. They proved the almost sure convergence of the norm of those matrices. More precisely, they proved that for P a self-adjoint polynomial, almost surely, for any $\varepsilon > 0$ and N large enough,

$$\sigma \left(P(X_1^N, \dots, X_d^N) \right) \subset \operatorname{Supp} \mu_P + (-\varepsilon, \varepsilon), \quad (4.2)$$

where $\operatorname{Supp} \mu_P$ is the support of the measure μ_P . In order to do so, they showed that given a smooth function f , there is a constant $\alpha_0^P(f)$, which can be computed explicitly with the help of free probability, such that

$$\mathbb{E} \left[\frac{1}{N} \operatorname{Tr}_N (f(A_N)) \right] = \alpha_0^P(f) + \mathcal{O}(N^{-2}).$$

A similar equality was proved in [7] with a better estimation of the dependency in the parameters such as f and Z^N in the $\mathcal{O}(N^{-2})$. Given the important consequences that studying the first

two orders had, one can wonder what happens at the next order. More precisely, could we write this expectation as a finite order Taylor expansion, and what consequences would it have on the eigenvalues? That is, can we prove that for any k , if f is smooth enough, there exist deterministic constants $\alpha_i^P(f)$ such that

$$\mathbb{E} \left[\frac{1}{N} \text{Tr} \left(f(P(X_1^N, \dots, X_d^N)) \right) \right] = \sum_{i=0}^k \frac{\alpha_i^P(f)}{N^{2i}} + \mathcal{O}(N^{-2k-2})?$$

In 2002, by using Riemann-Hilbert techniques, Ercolani and McLaughlin gave in [37] a positive answer for the case of a single random matrix (but not necessarily a GUE random matrix), that is $d = 1$. Haagerup and Thorbjørnsen gave a simplified proof in 2010 (see [59]) for the specific case of a single GUE matrix. However the method of the proof relied heavily on the explicit formula of the law of the eigenvalues of a GUE matrix and since there is no equivalent for polynomials in GUE matrices we cannot adapt this proof. Instead, we developed a proof whose main tool is free probability. The main idea of the proof is to interpolate independent GUE matrices and free semicirculars with free Ornstein-Uhlenbeck processes. It is similar to the method used in [7]. The main result is the following Theorem.

Theorem 4.1.1. *We define,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- $Z^N = (Z_1^N, \dots, Z_r^N, Z_1^{N*}, \dots, Z_r^{N*})$ deterministic matrices whose norm is uniformly bounded over N ,
- P a self-adjoint polynomial which can be written as a linear combination of \mathbf{m} monomials of degree at most n and coefficients at most c_{\max} ,
- $f : \mathbb{R} \mapsto \mathbb{R}$ a function of class $\mathcal{C}^{4(k+1)+2}$. We define $\|f\|_{\mathcal{C}^i}$ the sum of the supremum on \mathbb{R} of the first i -th derivatives of f .

Then there exist deterministic coefficients $(\alpha_i^P(f, Z^N))_{1 \leq i \leq k}$ and constants C, K and c independent of P , such that with $K_N = \max\{\|Z_1^N\|, \dots, \|Z_q^N\|, K\}$, $C_{\max}(P) = \max\{1, c_{\max}\}$, for any N , if $k \leq cNn^{-1}$,

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{N} \text{Tr}_N \left(f(P(X^N, Z^N)) \right) \right] - \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(f, Z^N) \right| \\ & \leq \frac{1}{N^{2(k+1)}} \|f\|_{\mathcal{C}^{4(k+1)+2}} \times \left(C \times n^2 K_N^n C_{\max} \mathbf{m} \right)^{4(k+1)+1} \times k^{12k}. \end{aligned} \quad (4.3)$$

Besides if we define \widehat{K}_N like K_N but with 2 instead of K , then we have that for any i ,

$$\left| \alpha_i^P(f, Z^N) \right| \leq \|f\|_{\mathcal{C}^{4i+2}} \times \left(C \times n^2 \widehat{K}_N^n C_{\max} \mathbf{m} \right)^{4i+1} \times i^{12i}. \quad (4.4)$$

Finally if f and g are functions of class $\mathcal{C}^{4(k+1)}$ equal on a neighborhood of the spectrum of $P(x, Z^N)$, where x is a free semicircular system free from $\mathbb{M}_N(\mathbb{C})$, then for any $i \leq k$, $\alpha_i^P(f, Z^N) = \alpha_i^P(g, Z^N)$. In particular if the support of f and the spectrum of $P(x, Z^N)$ are disjoint, then for any i , $\alpha_i^P(f, Z^N) = 0$.

This theorem is a consequence of the slightly sharper, but less explicit, Theorem 4.3.4. It is essentially the same statement, but instead of having the norm C^k of f , we make the moment of the Fourier transform of f appears. We also give an explicit expression for the coefficients α_i^P . The above Theorem calls for a few remarks.

- We assumed that the matrices Z^N are deterministic, but thanks to Fubini's theorem we can assume that they are random matrices as long as they are independent from X^N . In this situation though, K_N^n in the right side of the inequality is a random variable (and thus we need some additional assumptions if we want its expectation to be finite for instance).
- We assumed that the matrices Z^N were uniformly bounded over N . This is a technical assumption which is necessary to make sure that the coefficients α_i^P are well-defined. However as we can see in Theorem 4.3.4, one can relax this assumption. That being said, in order for equation (4.3) to be meaningful one has to be careful that the term $K_N^{n^4}$ is not compensating the term N^{-2} .
- The exponent 12 in the term k^{12k} is very suboptimal and could easily be optimized a bit more in the proof of Theorem 4.1.1. For a better bound we refer to Theorem 4.3.4, where the term k^{12k} is replaced by k^{3k} . However in order to work with the norm C^k instead of the moments of the Fourier transform, we were forced to increase this term.
- Although we cannot take $k = \infty$, hence only getting a finite Taylor expansion, we can still take k which depends on N . However to keep the last term under control we need to estimate the k -th derivative of f .
- Since the probability that there is an eigenvalue of $P(X^N, Z^N)$ outside of a neighborhood of $P(x, Z^N)$ is exponentially small as N goes to infinity. The hypothesis of smoothness on f only need to be verified on a neighborhood of $P(X^N, Z^N)$ for an asymptotic expansion to exist.

As we said earlier in the introduction, by studying the trace of a smooth function evaluated in $P(X_1^N, \dots, X_d^N)$, Haagerup and Thorbjørnsen were able to show in [6] that the spectrum of $P(X_1^N, \dots, X_d^N)$ converges for the Hausdorff distance towards an explicit subset of \mathbb{R} . We summarized this result in equation (4.2). With the full finite order Taylor expansion, by taking $f : x \rightarrow g(N^\alpha x)$ where g is a well-chosen smooth function, one can show the following proposition.

Corollary 4.1.2. *Let X^N be independent GUE matrices of size N , A^N a family of deterministic matrices whose norm is uniformly bounded over N , x be a free semicircular system and P a self-adjoint polynomial. Given $\alpha < 1/2$, almost surely for N large enough,*

$$\sigma\left(P(X^N, A^N)\right) \subset \sigma\left(P(x, A^N)\right) + N^{-\alpha},$$

where $\sigma(X)$ is the spectrum of X , and x is free from $\mathbb{M}_N(\mathbb{C})$.

In the case of a single GUE matrix, much more precise results were obtained by Tracy and Widom in [17]. They proved the existence of a continuous decreasing function F_2 from \mathbb{R} to $[0, 1]$ such that if $\lambda_1(X^N)$ denotes the largest eigenvalue of X^N ,

$$\lim_{N \rightarrow \infty} P\left(N^{2/3}(\lambda_1(X^N) - 2) \geq s\right) = F_2(s).$$

This was generalized to β -matrix model in [89] and to polynomials in independent GUE matrices which are close to the identity in [23]. But there is no such result for general polynomials in independent GUE matrices. However with Theorem 4.1.1 we managed to get an estimate on the tail of the distribution of $\sqrt{N} \|P(X^N, A^N)\|$.

Corollary 4.1.3. *Let X^N be a family of independent GUE matrices of size N , A^N a family of deterministic matrices whose norm is uniformly bounded over N , x be a free semicircular system and P a polynomial. Then there exists a constant C such that for N large enough,*

$$\mathbb{P} \left(\frac{\sqrt{N}}{\ln^4 N} \left(\|P(X^N, A^N)\| - \|P(x, A^N)\| \right) \geq C(\delta + 1) \right) \leq e^{-N} + e^{-\delta^2 \ln^8 N}.$$

This corollary is similar to Theorem 1.5 obtained in [7], but with a substantial improvement on the exponent since in this theorem, instead of $1/2$, we only had $1/4$. Theorem 1.5 of [7] also gave a similar bound on the probability that $\|P(X^N)\|$ be smaller than its deterministic limit, but Theorem 4.3.4 does not yield any improvement on this inequality. The proof of this corollary can be summarized in two steps: first use measure concentration to get an estimate on the probability that $\|P(X^N, A^N)\|$ is far from its expectation, and secondly use Theorem 4.1.1 to estimate the difference between the expectation and the deterministic limit. Finally it is worth noting that the exponent $1/2$ comes from the fact that for every N^2 that we gain in equation (4.3), we also have to differentiate our function f four more times. Thus if we take $f : x \rightarrow g(N^\alpha x)$ where g is smooth, then in order for N^{-2} to compensate the differential, we have to take $\alpha = 1/2$. If we only had to differentiate our function three more times, then we could take $\alpha = 2/3$ which is the same exponent as in Tracy-Widom.

4.2 Framework and standard properties

4.2.1 Usual definitions in free probability

In order to be self-contained, we begin by recalling the following definitions from free probability.

Definition 4.2.1. • *A \mathcal{C}^* -probability space $(\mathcal{A}, *, \tau, \|\cdot\|)$ is a unital \mathcal{C}^* -algebra $(\mathcal{A}, *, \|\cdot\|)$ endowed with a state τ , i.e. a linear map $\tau : \mathcal{A} \rightarrow \mathbb{C}$ satisfying $\tau(1_{\mathcal{A}}) = 1$ and $\tau(a^*a) \geq 0$ for all $a \in \mathcal{A}$. In this paper we always assume that τ is a trace, i.e. that it satisfies $\tau(ab) = \tau(ba)$ for any $a, b \in \mathcal{A}$. An element of \mathcal{A} is called a (noncommutative) random variable. We will always work with a faithful trace, namely, for $a \in \mathcal{A}$, $\tau(a^*a) = 0$ if and only if $a = 0$.*

- *Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be $*$ -subalgebras of \mathcal{A} , having the same unit as \mathcal{A} . They are said to be free if for all k , for all $a_i \in \mathcal{A}_{j_i}$ such that $j_1 \neq j_2, j_2 \neq j_3, \dots, j_{k-1} \neq j_k$:*

$$\tau \left((a_1 - \tau(a_1))(a_2 - \tau(a_2)) \dots (a_k - \tau(a_k)) \right) = 0.$$

Families of noncommutative random variables are said to be free if the $$ -subalgebras they generate are free.*

- Let $A = (a_1, \dots, a_k)$ be a k -tuple of random variables. The joint distribution of the family A is the linear form $\mu_A : P \mapsto \tau[P(A, A^*)]$ on the set of polynomials in $2k$ noncommutative variables.
- A family of noncommutative random variables $x = (x_1, \dots, x_d)$ is called a free semicircular system when the noncommutative random variables are free, self-adjoint ($x_i = x_i^*$), and for all k in \mathbb{N} and i , one has

$$\tau(x_i^k) = \int t^k d\sigma(t),$$

with $d\sigma(t) = \frac{1}{2\pi} \sqrt{4 - t^2} \mathbf{1}_{|t| \leq 2} dt$ the semicircle distribution.

It is important to note that thanks to [60, Theorem 7.9], that we recall below, one can consider free copies of any noncommutative random variable.

Theorem 4.2.2. *Let $(\mathcal{A}_i, \phi_i)_{i \in I}$ be a family of \mathcal{C}^* -probability spaces such that the functionals $\phi_i : \mathcal{A}_i \rightarrow \mathbb{C}$, $i \in I$, are faithful traces. Then there exist a \mathcal{C}^* -probability space (\mathcal{A}, ϕ) with ϕ a faithful trace, and a family of norm-preserving unital $*$ -homomorphism $W_i : \mathcal{A}_i \rightarrow \mathcal{A}$, $i \in I$, such that:*

- $\phi \circ W_i = \phi_i$, $\forall i \in I$.
- The unital \mathcal{C}^* -subalgebras form a free family in (\mathcal{A}, ϕ) .

Let us finally fix a few notations concerning the spaces and traces that we use in this paper.

Definition 4.2.3. • (\mathcal{A}_N, τ_N) is the free sum of $\mathbb{M}_N(\mathbb{C})$ with a system of d free semicircular variable, this is the \mathcal{C}^* -probability space built in Theorem 4.2.2. Note that when restricted to $\mathbb{M}_N(\mathbb{C})$, τ_N is just the regular renormalized trace on matrices. The restriction of τ_N to the \mathcal{C}^* -algebra generated by the free semicircular system x is denoted as τ . Note that one can view this space as the limit of a matrix space, we refer to Proposition 3.5 from [7].

- Tr_N is the non-renormalized trace on $\mathbb{M}_N(\mathbb{C})$.
- We denote $E_{r,s}$ the matrix with coefficients equal to 0 except in (r, s) where it is equal to one.
- We regularly identify $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$ with $\mathbb{M}_{kN}(\mathbb{C})$ through the isomorphism $E_{i,j} \otimes E_{r,s} \mapsto E_{i+rN, j+sN}$, similarly we identify $\text{Tr}_N \otimes \text{Tr}_k$ with Tr_{kN} .
- If $A^N = (A_1^N, \dots, A_d^N)$ and $B^k = (B_1^k, \dots, B_d^k)$ are two families of random matrices, then we denote $A^N \otimes B^k = (A_1^N \otimes B_1^k, \dots, A_d^N \otimes B_d^k)$. We typically use the notation $X^N \otimes I_k$ for the family $(X_1^N \otimes I_k, \dots, X_d^N \otimes I_k)$.

4.2.2 Non-commutative polynomials and derivatives

Let $\mathcal{A}_{d,2r} = \mathbb{C}\langle X_1, \dots, X_d, Y_1, \dots, Y_{2r} \rangle$ be the set of noncommutative polynomial in $p + 2r$ variables. We set $q = 2r$ to simplify notations. We endow this vector space with the norm

$$\|P\|_A = \sum_{M \text{ monomial}} |c_M(P)| A^{\deg M}, \quad (4.5)$$

where $c_M(P)$ is the coefficient of P for the monomial M and $\deg M$ the total degree of M (that is the sum of its degree in each letter $X_1, \dots, X_d, Y_1, \dots, Y_{2r}$). Let us define several maps which we use frequently in the sequel. First, for $A, B, C \in \mathcal{A}_{d,q}$, let

$$A \otimes B \# C = ACB, \quad A \otimes B \tilde{\#} C = BCA, \quad m(A \otimes B) = BA. \quad (4.6)$$

We define an involution $*$ on $\mathcal{A}_{d,q}$ by $X_i^* = X_i, Y_i^* = Y_{i+r}$ if $i \leq d+r, Y_i^* = Y_{i-r}$ else, and then we extend it to $\mathcal{A}_{d,q}$ by linearity and the formula $(\alpha PQ)^* = \bar{\alpha} Q^* P^*$. $P \in \mathcal{A}_{d,q}$ is said to be self-adjoint if $P^* = P$. Self-adjoint polynomials have the property that if $x_1, \dots, x_d, z_1, \dots, z_r$ are elements of a C^* -algebra such as x_1, \dots, x_d are self-adjoint, then so is $P(x_1, \dots, x_d, z_1, \dots, z_r, z_1^*, \dots, z_r^*)$.

Definition 4.2.4. *If $1 \leq i \leq d$, one defines the noncommutative derivative $\partial_i : \mathcal{A}_{d,q} \rightarrow \mathcal{A}_{d,q} \otimes \mathcal{A}_{d,q}$ by its value on a monomial $M \in \mathcal{A}_{d,q}$ given by*

$$\partial_i M = \sum_{M=AX_iB} A \otimes B,$$

and then extend it by linearity to all polynomials. We can also define ∂_i by induction with the formulas,

$$\forall P, Q \in \mathcal{A}_{d,q}, \quad \partial_i(PQ) = \partial_i P \times 1 \otimes Q + P \otimes 1 \times \partial_i Q, \quad (4.7)$$

$$\forall i, j, \quad \partial_i X_i = \partial_{i,j} 1 \otimes 1.$$

Similarly, with m as in (4.6), one defines the cyclic derivative $D_i : \mathcal{A}_{d,q} \rightarrow \mathcal{A}_{d,q}$ for $P \in \mathcal{A}_{d,q}$ by

$$D_i P = m \circ \partial_i P.$$

Definition 4.2.5. *We define $\mathcal{F}_{d,q}$ to be the $*$ -algebra generated by $\mathcal{A}_{d,q}$ and the family*

$$\{e^{iQ} \mid Q \in \mathcal{A}_{d,q} \text{ self-adjoint}\}.$$

Then, as we will see in the next proposition, a natural way to extend the definition of ∂_i (and D_i) to $\mathcal{F}_{d,q}$ is by setting

$$\partial_i e^{iQ} = \mathbf{i} \int_0^1 e^{i\alpha Q} \otimes 1 \partial_i Q 1 \otimes e^{i(1-\alpha)Q} d\alpha. \quad (4.8)$$

However we cannot define the integral properly on $\mathcal{F}_{d,q} \otimes \mathcal{F}_{d,q}$. After evaluating our polynomials in C^* -algebras, the integral will be well-defined as we will see. Firstly, we need to define properly the operator norm of tensor of C^* -algebras. We work with the minimal tensor product also named the spatial tensor product. For more information we refer to chapter 6 of [76].

Definition 4.2.6. *Let \mathcal{A} and \mathcal{B} be C^* -algebra with faithful representations $(H_{\mathcal{A}}, \phi_{\mathcal{A}})$ and $(H_{\mathcal{B}}, \phi_{\mathcal{B}})$, then if \otimes_2 is the tensor product of Hilbert spaces, $\mathcal{A} \otimes_{\min} \mathcal{B}$ is the completion of the image of $\phi_{\mathcal{A}} \otimes \phi_{\mathcal{B}}$ in $B(H_{\mathcal{A}} \otimes_2 H_{\mathcal{B}})$ for the operator norm in this space. This definition is independent of the representations that we fixed.*

While we will not always be in this situation during this paper, it is important to note that if $\mathcal{A} = \mathbb{M}_N(\mathbb{C})$, then up to isomorphism $\mathcal{A} \otimes_{\min} \mathcal{A}$ is simply $\mathbb{M}_{N^2}(\mathbb{C})$ with the usual operator norm. If $P \in \mathcal{A}_{d,q}$, $z = (z_1, \dots, z_{d+q})$ belongs to a C^* -algebra \mathcal{A} , then $(\partial_i P^k)(z)$ belongs to

$\mathcal{A} \otimes_{\min} \mathcal{A}$, and $\|(\partial_i P^k)(z)\| \leq C_P k \|P(z)\|^{k-1}$ for some constant C_P independent of k . Thus we can define

$$(\partial_i e^P)(z) = \sum_{k \in \mathbb{N}} \frac{1}{k!} (\partial_i P^k)(z). \quad (4.9)$$

We have now defined the noncommutative differential of the exponential of a polynomial twice, in (4.8) and (4.9). However those two definitions are compatible thanks to the following proposition (see [8], Proposition 2.2 for the proof).

Proposition 4.2.7. *Let $P \in \mathcal{A}_{d,q}$, $z = (z_1, \dots, z_{d+q})$ elements of a \mathcal{C}^* -algebra \mathcal{A} , then with $(\partial_i e^P)(z)$ defined as in (4.9),*

$$(\partial_i e^P)(z) = \int_0^1 e^{\alpha P(z)} \otimes 1 \partial_i P(z) 1 \otimes e^{(1-\alpha)P(z)} d\alpha.$$

We explained why (4.8) was well-defined when we evaluate our polynomials in a \mathcal{C}^* -algebra. However in order to be perfectly rigorous we need to give the following definition for the noncommutative differential that we use in the rest of this paper.

Definition 4.2.8. *For $\alpha \in [0, 1]$, let $\partial_{\alpha,i} : \mathcal{F}_{d,q} \rightarrow \mathcal{F}_{d,q} \otimes \mathcal{F}_{d,q}$ which satisfies (4.7) and such that for any $P \in \mathcal{A}_{d,q}$ self-adjoint,*

$$\partial_{\alpha,i} e^{iP} = i e^{i\alpha P} \otimes 1 \partial_i P 1 \otimes e^{i(1-\alpha)P}.$$

And then, given $z = (z_1, \dots, z_{d+q})$ elements of a \mathcal{C}^* -algebra, we define for any $Q \in \mathcal{F}_{d,q}$,

$$\partial_i Q(z) = \int_0^1 \partial_{\alpha,i} Q(z) d\alpha.$$

In particular, it means that we can define rigorously the composition of those maps. Since the map $\partial_{\alpha,i}$ goes from $\mathcal{F}_{d,q}$ to $\mathcal{F}_{d,q} \otimes \mathcal{F}_{d,q}$ it is very easy to do so. For example one can define the following operator (see Definition 4.2.10 for the notation ∂_i^1 , ∂_i^2 and \boxtimes). We will use a similar one later on.

Definition 4.2.9. *Let $Q \in \mathcal{F}_{d,q}$, given $z = (z_1, \dots, z_{d+q})$ elements of a \mathcal{C}^* -algebra, let $i, j \in [1, d]$, we define*

$$\partial_j D_i Q(z) = \int_{[0,1]^2} \partial_{\alpha_2,j} \circ (m \circ \partial_{\alpha_1,i}) Q(z) d\alpha.$$

For the sake of clarity, we introduce the following notation which is close to Sweedler's convention. Its interest will be clear in section 4.3.

Definition 4.2.10. *Let $Q \in \mathcal{F}_{d,q}$, \mathcal{C} be a \mathcal{C}^* -algebra. Then let $\alpha : \mathcal{F}_{d,q} \rightarrow \mathcal{C}$ and $\beta : \mathcal{F}_{d,q} \rightarrow \mathcal{C}$ be morphisms. We also set $n : A \otimes B \in \mathcal{C} \otimes \mathcal{C} \mapsto AB \in \mathcal{C}$. Then we use the following notation,*

$$\alpha(\partial_i^1 P) \boxtimes \beta(\partial_i^2 P) = n \circ (\alpha \otimes \beta(\partial_i P)).$$

This notation is especially useful when our maps α and β are simply evaluation of P as it is the case in section 4.3. Indeed we typically denote, $\partial_i^1 P(X) \boxtimes \partial_i^2 P(Y)$, rather than define $h_X : P \rightarrow P(X)$ and use the more cumbersome and abstract notation, $n \circ (h_X \otimes h_Y(\partial_i P))$.

The map ∂_i is related to the so-called Schwinger-Dyson equations on semicircular variable thanks to the following proposition. One can find a proof for polynomials in [13], Lemma 5.4.7, and then extend it to $\mathcal{F}_{d,q}$ thanks to Definition (4.9).

Proposition 4.2.11. *Let $x = (x_1, \dots, x_d)$ be a free semicircular system, $y = (y_1, \dots, y_r)$ be noncommutative random variables free from x , if the family (x, y) belongs to the \mathcal{C}^* -probability space $(\mathcal{A}, *, \tau, \|\cdot\|)$, then for any $Q \in \mathcal{F}_{d,q}$,*

$$\tau(Q(x, y, y^*) x_i) = \tau \otimes \tau(\partial_i Q(x, y, y^*)) .$$

Now that we have defined the usual noncommutative polynomial spaces, we build a very specific one which we need to define properly the coefficients of the topological expansion.

Definition 4.2.12. *Let $(c_n)_n$ be the sequence such that $c_0 = 0$, $c_{n+1} = 6c_n + 6$. We define by induction, $J_0 = \{\emptyset\}$ and for $n \geq 0$, $j \in [1, 2n]$,*

$$\begin{aligned} J_{n+1}^{j,1} &= \left\{ \{s_1 + c_n, \dots, s_{j-1} + c_n, s_j + c_n, s_j, \dots, s_{2n}, 3c_n + 1\} \mid I = \{s_1, \dots, s_{2n}\} \in J_n \right\}, \\ J_{n+1}^{2n+1,1} &= \left\{ \{s_1 + c_n, \dots, s_{2n} + c_n, 3c_n + 2, 3c_n + 1\} \mid I = \{s_1, \dots, s_{2n}\} \in J_n \right\}, \\ J_{n+1}^{j,2} &= \left\{ \{s_1 + 2c_n, \dots, s_{j-1} + 2c_n, s_j + 2c_n, s_j, \dots, s_{2n}, 3c_n + 1\} \mid I = \{s_1, \dots, s_{2n}\} \in J_n \right\}, \\ J_{n+1}^{2n+1,2} &= \left\{ \{s_1 + 2c_n, \dots, s_{2n} + 2c_n, 3c_n + 3, 3c_n + 1\} \mid I = \{s_1, \dots, s_{2n}\} \in J_n \right\}. \end{aligned}$$

We similarly define $\tilde{J}_{n+1}^{j,1}$ and $\tilde{J}_{n+1}^{j,2}$ by adding $3c_n + 3$ to every integer in every set. Finally we fix

$$J_{n+1} = \bigcup_{1 \leq j \leq 2n+1} J_{n+1}^{j,1} \cup J_{n+1}^{j,2} \cup \tilde{J}_{n+1}^{j,1} \cup \tilde{J}_{n+1}^{j,2}.$$

Then we define $\mathcal{A}_{d,q}^n = \mathbb{C}\langle X_{i,I}, 1 \leq i \leq d, I \in J_n, Y_1, \dots, Y_{2r} \rangle$. We also define $\mathcal{F}_{d,q}^n$ as the $*$ -algebra generated by $\mathcal{A}_{d,q}^n$ and the family $\{e^{iQ} \mid Q \in \mathcal{A}_{d,q}^n \text{ self-adjoint}\}$.

Definition 4.2.13. *Similarly to Definition 4.2.8, we define ∂_i and $\partial_{i,I}$ on $\mathcal{F}_{d,q}^n$ which satisfies (4.7) and (4.8) and*

$$\forall i, j \in [1, p], I, K \in J_n, \quad \partial_{i,I} X_{j,K} = \delta_{i,j} \delta_{I,K} 1 \otimes 1, \quad \partial_i X_{j,K} = \delta_{i,j} 1 \otimes 1.$$

We then define $D_i = m \circ \partial_i$ and $D_{i,I} = m \circ \partial_{i,I}$ on $\mathcal{F}_{d,q}^n$.

In particular, $\mathcal{F}_{d,q}^0 = \mathcal{F}_{d,q}$ and the two definitions of ∂_i coincide. The following lemma will be important for a better estimation of the remainder term in the expansion.

Lemma 4.2.14. *Given $s \in [1, c_n]$, there exists a unique $l \in [1, n]$ such that for any $I = \{s_1, \dots, s_{2n}\} \in J_n$, either $s_l = s$ or $s \notin I$. We refer to l as the depth of s in J_n , and will denote it $\text{depth}^n(s)$.*

Proof. Let's proceed by induction. If this is true for J_n , then let $s \in [1, c_{n+1}]$. If $s \leq 3c_n + 3$ then it will only appear in the elements of $J_{n+1}^{j,1}$ and $J_{n+1}^{j,2}$ for j from 1 to $n + 1$. If $s > 3c_n + 3$ then on the contrary it will only appear in the elements of $\tilde{J}_{n+1}^{j,1}$ and $\tilde{J}_{n+1}^{j,2}$ for j from 1 to $n + 1$. Since both cases are identical, one can assume that $s \leq 3c_n + 3$. If s is equal to $3c_n + 3, 3c_n + 2$ or $3c_n + 1$, then it is straightforward. Thus there remains three possibilities:

- Either $s \in [1, c_n]$, then let l be the depth of s in J_n , then by construction the depth of s in J_{n+1} will be $l + 1$.
- Either $s \in [c_n + 1, 2c_n]$, then with l the depth of $s - c_n$ in J_n , the depth of s in J_{n+1} will also be l .
- Either $s \in [2c_n + 1, 3c_n]$, then with $l - 2c_n$ the depth of s in J_n , the depth of s in J_{n+1} will also be l .

□

This prompts us to define the following non-commutative derivative on $\mathcal{F}_{d,q}^n$.

Definition 4.2.15. *Given $s \in [0, c_n]$ with depth l in J_n , then we set*

$$\partial_{i,s,n} = \sum_{I \in J_n \text{ such that } s \in I} \partial_{i,I} = \sum_{I \in J_n \text{ such that } s_I = s} \partial_{i,I}.$$

And finally we define $D_{i,s,n} = m \circ \partial_{i,s,n}$.

4.2.3 GUE random matrices

We conclude this section by reminding the definition of Gaussian random matrices and stating a few useful properties about them.

Definition 4.2.16. *A GUE random matrix X^N of size N is a self-adjoint matrix whose coefficients are random variables with the following laws:*

- For $1 \leq i \leq N$, the random variables $\sqrt{N}X_{i,i}^N$ are independent centered Gaussian random variables of variance 1.
- For $1 \leq i < j \leq N$, the random variables $\sqrt{2N} \Re X_{i,j}^N$ and $\sqrt{2N} \Im X_{i,j}^N$ are independent centered Gaussian random variables of variance 1, independent of $(X_{i,i}^N)_i$.

When doing computations with Gaussian variables, the main tool that we use is Gaussian integration by part. It can be summarized into the following formula, if Z is a centered Gaussian variable with variance 1 and f a \mathcal{C}^1 function, then

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[\partial_Z f(Z)] . \quad (4.10)$$

A direct consequence of this, is that if x and y are centered Gaussian variable with variance 1, and $Z = \frac{x+iy}{\sqrt{2}}$, then

$$\mathbb{E}[Zf(x, y)] = \mathbb{E}[\partial_Z f(x, y)] \quad \text{and} \quad \mathbb{E}[\bar{Z}f(x, y)] = \mathbb{E}[\partial_{\bar{Z}} f(x, y)] , \quad (4.11)$$

where $\partial_Z = \frac{1}{2}(\partial_x + i\partial_y)$ and $\partial_{\bar{Z}} = \frac{1}{2}(\partial_x - i\partial_y)$. When working with GUE matrices, an important consequence of this are the so-called Schwinger-Dyson equations, which we summarize in the following proposition. For more information about these equations and their applications, we refer to [13], Lemma 5.4.7.

Proposition 4.2.17. *Let X^N be GUE matrices of size N , $Q \in \mathcal{F}_{d,q}$, then for any i ,*

$$\mathbb{E} \left[\frac{1}{N} \text{Tr}_N(X_i^N Q(X^N)) \right] = \mathbb{E} \left[\left(\frac{1}{N} \text{Tr}_N \right)^{\otimes 2} (\partial_i Q(X^N)) \right].$$

Proof. Let us first assume that $Q \in \mathcal{A}_{d,q}$. One can write $X_i^N = \frac{1}{\sqrt{N}}(x_{r,s}^i)_{1 \leq r,s \leq N}$ and thus

$$\begin{aligned} \mathbb{E} \left[\frac{1}{N} \text{Tr}_N(X_i^N Q(X^N)) \right] &= \frac{1}{N^{3/2}} \sum_{r,s} \mathbb{E} \left[x_{r,s}^i \text{Tr}_N(E_{r,s} Q(X^N)) \right] \\ &= \frac{1}{N^{3/2}} \sum_{r,s} \mathbb{E} \left[\text{Tr}_N(E_{r,s} \partial_{x_{r,s}^i} Q(X^N)) \right] \\ &= \frac{1}{N^2} \sum_{r,s} \mathbb{E} \left[\text{Tr}_N(E_{r,s} \partial_i Q(X^N) \# E_{s,r}) \right] \\ &= \mathbb{E} \left[\left(\frac{1}{N} \text{Tr}_N \right)^{\otimes 2} (\partial_i Q(X^N)) \right]. \end{aligned}$$

If $Q \in \mathcal{F}_{d,q}$, then the proof is pretty much the same but we need to use Duhamel's formula which states that for any matrices A and B ,

$$e^B - e^A = \int_0^1 e^{\alpha B} (B - A) e^{(1-\alpha)A} d\alpha. \quad (4.12)$$

Thus this let us prove that

$$\partial_{x_{r,s}^i} e^{iP(X^N)} = \mathbf{i} \int_0^1 e^{i\alpha P(X^N)} \partial_i P(X^N) \# E_{s,r} e^{i(1-\alpha)P(X^N)} d\alpha.$$

And the conclusion follows. □

Now to finish this section we state a property that we use several times in this paper. For the proof we refer to Proposition 2.11 in [7].

Proposition 4.2.18. *There exist constants C, D and α such that for any $N \in \mathbb{N}$, if X^N is a GUE random matrix of size N , then for any $u \geq 0$,*

$$\mathbb{P} \left(\|X^N\| \geq u + D \right) \leq e^{-\alpha u N}.$$

Consequently, for any $k \leq \alpha N/2$,

$$\mathbb{E} \left[\|X^N\|^k \right] \leq C^k.$$

4.3 Proof of Theorem 4.1.1

4.3.1 A Poincaré type equality

One of the main tool when dealing with GUE random matrices is the Poincaré inequality (see Definition 4.4.2 from [13]), which gives us a sharp upper bound of the variance of a function

in these matrices. Typically this inequality shows that the variance of a trace of a polynomial in GUE random matrices, which a priori is of order $\mathcal{O}(1)$, is of order $\mathcal{O}(N^{-2})$. In this paper we use the same kind of argument which are used to prove the Poincaré inequality to get an exact formula for the variances we are interested in.

Proposition 4.3.1. *Let $P, Q \in \mathcal{F}_{d,q}$, R^N, S^N, T^N be independent families of d independent GUE matrices of size N . Let A^N be a family of deterministic matrices and their adjoints. With convention $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, for any $t \geq 0$, we have:*

$$\begin{aligned} & \text{Cov} \left(\text{Tr}_N \left(P \left((1 - e^{-t})^{1/2} R^N, A^N \right) \right), \text{Tr}_N \left(Q \left((1 - e^{-t})^{1/2} R^N, A^N \right) \right) \right) \\ &= \frac{1}{N} \sum_i \int_0^t \mathbb{E} \left[\text{Tr}_N \left(D_i P \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} S^N, A^N \right) \right. \right. \\ & \quad \left. \left. \times D_i Q \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} T^N, A^N \right) \right) \right] ds. \end{aligned}$$

Proof. We define the following function,

$$\begin{aligned} h(s) = \mathbb{E} \left[\text{Tr}_N \left(P \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} S^N, A^N \right) \right. \right. \\ \left. \left. \text{Tr}_N \left(Q \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} T^N, A^N \right) \right) \right) \right]. \end{aligned}$$

To simplify notations, we set

$$S_s^N = \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} S^N, A^N \right),$$

$$T_s^N = \left((e^{-s} - e^{-t})^{1/2} R^N + (1 - e^{-s})^{1/2} T^N, A^N \right).$$

Then we have,

$$\text{Cov} \left(\text{Tr}_N \left(P \left((1 - e^{-t})^{1/2} R^N, A^N \right) \right), \text{Tr}_N \left(Q \left((1 - e^{-t})^{1/2} R^N, A^N \right) \right) \right) = - \int_0^t \frac{dh}{ds}(s) ds.$$

Thanks to Duhamel's formula (see (4.12)) we find

$$\frac{dP(S_s^N, A^N)}{ds} = -\frac{e^{-s}}{2} \sum_{i=1}^d \partial_i P(S_s^N, A^N) \# \left(\frac{R_i^N}{(e^{-s} - e^{-t})^{1/2}} - \frac{S_i^N}{(1 - e^{-s})^{1/2}} \right).$$

Since $\text{Tr}_N(\partial_i P \# B) = \text{Tr}_N(D_i P \times B)$, we compute,

$$\begin{aligned} \frac{dh}{ds}(s) = -\frac{e^{-s}}{2} \sum_i \mathbb{E} \left[\text{Tr}_N \left(D_i P(S_s^N) \left(\frac{R_i^N}{(e^{-s} - e^{-t})^{1/2}} - \frac{S_i^N}{(1 - e^{-s})^{1/2}} \right) \right) \text{Tr}_N(Q(T_s^N)) \right. \\ \left. + \text{Tr}_N(P(S_s^N)) \text{Tr}_N \left(D_i Q(S_s^N) \left(\frac{R_i^N}{(e^{-s} - e^{-t})^{1/2}} - \frac{T_i^N}{(1 - e^{-s})^{1/2}} \right) \right) \right]. \end{aligned}$$

But by using integration by part formula (4.11), we get that

$$\begin{aligned} & \mathbb{E} \left[\mathrm{Tr}_N \left(D_i P \left(S_s^N \right) \frac{R_i^N}{(e^{-s} - e^{-t})^{1/2}} \right) \mathrm{Tr}_N \left(Q \left(T_s^N \right) \right) \right] \\ &= \frac{1}{N} \sum_{1 \leq a, b \leq N} \mathbb{E} \left[\mathrm{Tr}_N \left(E_{a,b} \partial_i D_i P \left(S_s^N \right) \# E_{b,a} \right) \times \mathrm{Tr}_N \left(Q \left(T_s^N \right) \right) \right. \\ & \quad \left. + \mathrm{Tr}_N \left(D_i P \left(S_s^N \right) E_{a,b} \right) \times \mathrm{Tr}_N \left(D_i Q \left(T_s^N \right) E_{b,a} \right) \right]. \end{aligned}$$

And similarly

$$\begin{aligned} & \mathbb{E} \left[\mathrm{Tr}_N \left(D_i P \left(S_s^N \right) \frac{S_i^N}{(1 - e^{-s})^{1/2}} \right) \mathrm{Tr}_N \left(Q \left(T_s^N \right) \right) \right] \\ &= \frac{1}{N} \sum_{1 \leq a, b \leq N} \mathbb{E} \left[\mathrm{Tr}_N \left(E_{a,b} \partial_i D_i P \left(S_s^N \right) \# E_{b,a} \right) \times \mathrm{Tr}_N \left(Q \left(T_s^N \right) \right) \right]. \end{aligned}$$

Therefore with similar computations we conclude,

$$\frac{dh}{ds}(s) = -\frac{1}{N} e^{-s} \sum_i \mathbb{E} \left[\mathrm{Tr}_N \left(D_i P \left(S_s^N \right) D_i P \left(S_s^N \right) \right) \right].$$

Hence the conclusion. □

4.3.2 A first rough formulation of the coefficients

In this subsection we prove the following lemma which will be the backbone of the proof of the topological expansion. The heuristic behind this lemma is that if $Q \in \mathcal{F}_{d(p+1),q}$, X^N is a family of d independent GUE matrices, $(y_i)_{i \geq 1}$ are systems of d free semicircular variables free between each other. Then we can find $R \in \mathcal{F}_{6d(p+1)+1,q}$ such that

$$\begin{aligned} & \mathbb{E} \left[\tau_N \left(Q \left(X^N, (y_i)_{1 \leq i \leq p} \right), Z^N \right) \right] - \tau_N \left(Q \left(x, (y_i)_{1 \leq i \leq p} \right), Z^N \right) \\ &= \frac{1}{N^2} \mathbb{E} \left[\tau_N \left(R \left(X^N, (y_i)_{1 \leq i \leq 6p+6} \right), Z^N \right) \right]. \end{aligned}$$

Then we will only need to apply this lemma recursively to build the topological expansion. Note that thanks to the definition of \mathcal{A}_N in Definition 4.2.3, it makes sense to consider matrices and free semicirculars in the same space. One can also assume that those matrices are random thanks to Proposition 2.7 of [7].

Lemma 4.3.2. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,
- x, z^1, z^2 free families of d free semi-circular variables,

- $y_s = (y_{s,1}, \dots, y_{s,d_s})$ for s from 1 to n , systems of free semicircular variables, free between each other and from x ,
- v_s, w_s free copies of y_s , free between each other,
- $Z^N = (Z_1^N, \dots, Z_q^N)$ deterministic matrices and their adjoints,
- $Y^N = \left((1 - e^{-t_1})^{1/2} y_1, \dots, (1 - e^{-t_n})^{1/2} y_n, X^N, Z^N \right)$,
- $Y = \left((1 - e^{-t_1})^{1/2} y_1, \dots, (1 - e^{-t_n})^{1/2} y_n, x, Z^N \right)$,
- $z_r^1 = \left((1 - e^{-t_1})^{1/2} v_1, \dots, (1 - e^{-t_n})^{1/2} v_n, (1 - e^{-r})^{1/2} z^1 + (e^{-r} - e^{-t})^{1/2} x + e^{-t} X^N, Z^N \right)$,
- for s from 1 to n ,

$$z_r^{1,s} = \left((1 - e^{-t_1})^{1/2} v_1, \dots, (1 - e^{-r})^{1/2} v_s + (e^{-r} - e^{-t_s})^{1/2} y_s, \right. \\ \left. (1 - e^{-t_{s+1}})^{1/2} y_{s+1}, \dots, (1 - e^{-t_n})^{1/2} y_n, (1 - e^{-t})^{1/2} x + e^{-t} X^N, Z^N \right),$$

- z_r^1 and $z_r^{2,s}$, defined similarly but with w and z^2 instead of v and z^1 ,
- $\tilde{z}_r^1, \tilde{z}_r^2, \tilde{z}_r^{1,s}$ and $\tilde{z}_r^{2,s}$ defined similarly but where we replaced $v_s, w_s, y_s, z^1, z^2, x$ by free copies,
- $Q \in \mathcal{F}_{d_1 + \dots + d_n + d, q}$.

Then, for any N , with $\partial_{s,j}$ the noncommutative differential as defined in 4.2.4 but with respect to $(1 - e^{-t_s})^{1/2} y_{s,j}$

$$\begin{aligned} & \mathbb{E} \left[\tau_N \left(Q \left(Y^N \right) \right) \right] - \tau_N \left(Q \left(Y \right) \right) \\ &= \frac{1}{2N^2} \int_0^\infty e^{-t} \sum_{\substack{1 \leq i \leq d \\ 1 \leq s \leq n \\ 1 \leq j \leq d_s}} \int_0^{t_s} \mathbb{E} \left[\tau_N \left(\left(\partial_{s,j}^2 \left(\partial_i^1 D_i Q \right) \left(z_r^{1,s} \right) \boxtimes \partial_{s,j}^1 \left(\partial_i^1 D_i Q \right) \left(\tilde{z}_r^{1,s} \right) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_{s,j}^2 \left(\partial_i^2 D_i Q \right) \left(\tilde{z}_r^{2,s} \right) \boxtimes \partial_{s,j}^1 \left(\partial_i^2 D_i Q \right) \left(z_r^{2,s} \right) \right) \right) \right] dr dt \\ &+ \frac{1}{2N^2} \int_0^\infty e^{-t} \sum_{1 \leq i, j \leq d} \int_0^t \mathbb{E} \left[\tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i Q \right) \left(z_r^1 \right) \boxtimes \partial_j^1 \left(\partial_i^1 D_i Q \right) \left(\tilde{z}_r^1 \right) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i Q \right) \left(\tilde{z}_r^2 \right) \boxtimes \partial_j^1 \left(\partial_i^2 D_i Q \right) \left(z_r^2 \right) \right) \right) \right] dr dt. \end{aligned}$$

First we need to prove the following technical lemma.

Lemma 4.3.3. *If Y^{kN} is a family of l independent GUE matrices of size kN , then let*

$$S_k = \left(Y^{kN}, X^N \otimes I_k, Z^N \otimes I_k \right)$$

With $P_{1,2} = I_N \otimes E_{1,2}$, \mathbb{E}_k the expectation with respect to Y^{kN} , given $Q \in \mathcal{F}_{l+d,q}$, we have that

$$\lim_{k \rightarrow \infty} k^{3/2} \mathbb{E}_k \left[\tau_{kN} \left(Q(S_k) P_{1,2} \right) \right] = 0$$

Proof. Given $A_1, \dots, A_l, B_1, \dots, B_l \in \mathcal{A}_{l+d,q}$, we define the following quantity,

$$f_A(y) = \mathbb{E}_k \left[\tau_{kN} \left((A_1 e^{iyB_1} \dots A_l e^{iyB_l})(S_k) P_{1,2} \right) \right],$$

$$d_n(y) = \sup_{\sum_i \deg A_i \leq n, A_i \text{ monomials}} |f_A(y)|.$$

Thanks to Proposition 4.2.18, we know that there exists constants γ and D (depending on $N, \|X^N\|$ and $\|Z^N\|$) such that for any $n \leq \gamma k$, $|d_n(y)| \leq D^n$. Consequently we define

$$g(a, y) = \sum_{n \leq \gamma k/2} d_n(y) a^n.$$

Let $m = \sup_i \deg B_i$ and A be such that $\sum_i \deg A_i \leq n$, there exists a constant C_B which only depends on the coefficients of the B_i such that

$$\left| \frac{df_A(y)}{dy} \right| \leq C_B d_{n+m}(y).$$

Naturally we get that for any $y \in [0, 1]$

$$|f_A(y)| \leq |f_A(0)| + C_B \int_0^y d_{n+m}(y) dy.$$

And by taking the supremum over A , we get that

$$d_n(y) \leq d_n(0) + C_B \int_0^y d_{n+m}(y) dy.$$

Hence by summing over n , we have for a small enough,

$$g(a, y) \leq g(a, 0) + C_B a^{-m} \left(\sum_{1 \leq i \leq m} (aD)^{\gamma k/2+i-1} + \int_0^y g(a, y) dy \right).$$

Thanks to Grönwall's inequality we get that there exist a constant κ such that for any $y \in [0, 1]$,

$$g(a, y) \leq \left(g(a, 0) + C_B a^{-m} \sum_{1 \leq i \leq m} (aD)^{\gamma k/2+i-1} \right) e^{y C_B a^{-m}}.$$

Thus for $a < 1/D$, we have

$$\limsup_{k \rightarrow \infty} k^{3/2} g(a, y) \leq e^{y C_B a^{-m}} \limsup_{k \rightarrow \infty} k^{3/2} g(a, 0).$$

However, we have

$$g(a, 0) = \sum_{n \leq \gamma k/2} a^n \sup_{A \text{ monomial}, \deg A \leq n} |\mathbb{E}_k [\tau_{kN} (A(S_k) P_{1,2})]|.$$

We refer to the proof of Lemma 3.7 of [7] to prove that $\limsup_{k \rightarrow \infty} k^{3/2} g(a, 0) = 0$ (with the notations of [7], it is the same thing as to show that $k^{3/2} f_{\gamma k/2}(a)$ converges towards 0). Hence for any A, B ,

$$\limsup_{k \rightarrow \infty} k^{3/2} \left| \mathbb{E}_k \left[\tau_{kN} \left((A_1 e^{iyB_1} \dots A_l e^{iyB_l})(S_k) P_{1,2} \right) \right] \right| \leq a^{-\sum_i \deg A_i} \limsup_{k \rightarrow \infty} k^{3/2} g(a, 1) = 0.$$

Hence the conclusion. □

Proof of Lemma 4.3.2. With

$$Y_t = \left((1 - e^{-t_1})^{1/2} y_1, \dots, (1 - e^{-t_n})^{1/2} y_n, (1 - e^{-t})^{1/2} x + e^{-t/2} X^N, Z^N \right),$$

we have,

$$\mathbb{E} \left[\tau_N \left(Q \left(Y^N \right) \right) \right] - \tau_N \left(Q \left(Y \right) \right) = - \int_0^\infty \mathbb{E} \left[\frac{d}{dt} \tau_N \left(Q \left(Y_t^N \right) \right) \right] dt.$$

We can compute

$$\frac{d}{dt} \tau_N \left(Q \left(Y_t^N \right) \right) = \frac{e^{-t}}{2} \sum_i \tau_N \left(D_i Q \left(Y_t^N \right) \left(\frac{x_i}{(1 - e^{-t})^{1/2}} - e^{t/2} X_i^N \right) \right)$$

Thus thanks to Gaussian integration by part (see (4.11)) and Schwinger-Dyson equations (see Proposition 4.2.11), we get that

$$\begin{aligned} & \mathbb{E} \left[\frac{d}{dt} \tau_N \left(Q \left(Y_t^N \right) \right) \right] \\ &= \mathbb{E} \left[\frac{e^{-t}}{2} \sum_i \left(\tau_N \otimes \tau_N \left(\partial_i D_i Q \left(Y_t^N \right) \right) - \frac{1}{N} \sum_{u,v} \tau_N \left(E_{u,v} \partial_i D_i Q \left(Y_t^N \right) \# E_{v,u} \right) \right) \right]. \end{aligned} \quad (4.13)$$

Let

$$\Lambda_{N,t} = \tau_N \otimes \tau_N \left(\partial_i D_i Q \left(Y_t^N \right) \right) - \frac{1}{N} \sum_{u,v} \tau_N \left(E_{u,v} \partial_i D_i Q \left(Y_t^N \right) \# E_{v,u} \right).$$

Thanks to Theorem 5.4.5 of [13], we have that if

$$Z_k = \left((1 - e^{-t_1})^{1/2} Y_1^{kN}, \dots, (1 - e^{-t_n})^{1/2} Y_n^{kN}, (1 - e^{-t})^{1/2} Y_{n+1}^{kN} + e^{-t} X^N \otimes I_k, Z^N \otimes I_k \right)$$

with Y_s^{kN} being independent families of d_s independent GUE matrices (with $d_{n+1} = d$). Then with \mathbb{E}_k the expectation with respect to Y_s^{kN} for every s ,

$$\begin{aligned} \Lambda_{N,t} &= \lim_{k \rightarrow \infty} E_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q \left(Z_k \right) \right) \\ &\quad - \mathbb{E}_k \left[\frac{1}{N} \sum_{1 \leq u,v \leq N} \tau_{kN} \left(E_{u,v} \otimes I_k \partial_i D_i Q \left(Z_k \right) \# E_{v,u} \otimes I_k \right) \right]. \end{aligned} \quad (4.14)$$

For more information, we refer to Proposition 3.5 of [7]. See also the definition of \mathcal{A}_N in Definition 4.2.3. Let A, B be matrices of $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$, with $(g_a)_{1 \leq a \leq N}$ the canonical basis of \mathbb{C}^N and $(f_b)_{1 \leq b \leq k}$ the one of \mathbb{C}^k , since $I_k = \sum_l E_{l,l}$ and $\tau_{kN}(M) = \sum_{a,b} g_a^* \otimes f_b^* M g_a \otimes f_b$,

$$\begin{aligned} & \frac{1}{N} \sum_{1 \leq u,v \leq N} \tau_{kN} \left(E_{u,v} \otimes I_k A E_{v,u} \otimes I_k B \right) \\ &= \frac{1}{N} \sum_{1 \leq u,v \leq N} \sum_{1 \leq l,l' \leq k} \tau_{kN} \left(E_{u,v} \otimes E_{l,l'} A E_{v,u} \otimes E_{l',l} B \right) \\ &= \frac{1}{N^2 k} \sum_{1 \leq l,l' \leq k} \sum_{1 \leq v \leq N} g_v^* \otimes f_l^* A g_v \otimes f_{l'} \sum_{1 \leq u \leq N} g_u^* \otimes f_{l'}^* B g_u \otimes f_l \\ &= \frac{1}{k} \sum_{1 \leq l,l' \leq k} \tau_N \left(I_N \otimes f_l^* A I_N \otimes f_{l'} \right) \tau_N \left(I_N \otimes f_{l'}^* B I_N \otimes f_l \right) \\ &= k \sum_{1 \leq l,l' \leq k} \tau_{kN} \left(A I_N \otimes E_{l',l} \right) \tau_{kN} \left(B I_N \otimes E_{l,l'} \right). \end{aligned}$$

Hence with convention $P_{l,l'} = I_N \otimes E_{l,l'}$, we have

$$\frac{1}{N} \sum_{1 \leq u, v \leq N} \tau_{kN} \left(E_{u,v} \otimes I_k \partial_i D_i Q(Z_k) \# E_{v,u} \otimes I_k \right) = k \sum_{1 \leq l, l' \leq k} \tau_{kN} \otimes \tau_{kN} \left(\partial_i D_i Q(Z_k) \times P_{l,l'} \otimes P_{l,l'} \right) \quad (4.15)$$

Consequently, we get that

$$\begin{aligned} \Lambda_{N,t} &= \lim_{k \rightarrow \infty} E_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) \right) \\ &\quad - \mathbb{E}_k \left[k \sum_{1 \leq l, l' \leq k} \tau_{kN} \otimes \tau_{kN} \left(\partial_i D_i Q(Z_k) \times P_{l,l'} \otimes P_{l,l'} \right) \right]. \end{aligned} \quad (4.16)$$

Let $U \in \mathbb{M}_k(\mathbb{C})$ be a unitary matrix, then since for any i ,

$$\begin{aligned} I_N \otimes U X_i^N \otimes I_k I_N \otimes U^* &= X_i^N \otimes I_k, \\ I_N \otimes U Z_i^N \otimes I_k I_N \otimes U^* &= Z_i^N \otimes I_k, \end{aligned}$$

and that the law of $Y_{s,j}^{kN}$ is invariant by conjugation by a unitary matrix, we get that for any unitary matrices U and V ,

$$\begin{aligned} &\mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) \times P_{l,l'} \otimes P_{l,l'} \right) \\ &= \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) \times (I_N \otimes U^* P_{l,l'} I_N \otimes U) \otimes (I_N \otimes V^* P_{l,l'} I_N \otimes V) \right). \end{aligned}$$

Thus if $l = l'$, we can pick U such that $U^* E_{l,l} U = E_{1,1}$, and if $l \neq l'$, we can pick U such that $U^* E_{l,l} U = E_{1,2}$. By doing the same for V , we have

$$\begin{aligned} &k \sum_{1 \leq l, l' \leq k} \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \\ &= k^2 \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{1,1} \otimes P_{1,1} \right) \\ &\quad + k^2 (k-1) \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{1,2} \otimes P_{1,2} \right). \end{aligned} \quad (4.17)$$

Similarly we also have,

$$\begin{aligned} &\mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) \right) \\ &= \sum_{1 \leq l, l' \leq k} \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{l,l} \otimes P_{l,l'} \right) \\ &= k^2 \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{1,1} \otimes P_{1,1} \right). \end{aligned} \quad (4.18)$$

By combining equations (4.16), (4.17) and (4.18), we get that

$$\begin{aligned} \Lambda_{N,t} &= \lim_{k \rightarrow \infty} - \left\{ k \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\tau_{kN} \otimes \tau_{kN} \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \right] \right. \\ &\quad \left. - \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \right. \\ &\quad \left. + k^2 (k-1) \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{1,2} \otimes P_{1,2} \right) \right\}. \end{aligned}$$

Thanks to Lemma 4.3.3, the last term converges towards 0. Consequently,

$$\Lambda_{N,t} = \lim_{k \rightarrow \infty} -k \left\{ \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\tau_{kN} \otimes \tau_{kN} \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \right] - \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \right\}. \quad (4.19)$$

Let R_s^{kN}, S_s^{kN} for s from 1 to $n+1$ be independent families of d_s independent GUE random matrices. We set the following notations,

$$\begin{aligned} Z_{k,r}^1 &= \left((1 - e^{-t_1})^{1/2} S_1^{kN}, \dots, (1 - e^{-t_n})^{1/2} S_n^{kN}, \right. \\ &\quad \left. (1 - e^{-r})^{1/2} S_{n+1}^{kN} + (e^{-r} - e^{-t})^{1/2} Y_{n+1}^{kN} + e^{-t} X^N \otimes I_k, Z^N \otimes I_k \right), \\ Z_{k,r}^{1,s} &= \left((1 - e^{-t_1})^{1/2} S_1^{kN}, \dots, (1 - e^{-r})^{1/2} S_s^{kN} + (e^{-r} - e^{-t_s})^{1/2} Y_s^{kN}, (1 - e^{-t_{s+1}})^{1/2} Y_{s+1}^{kN}, \right. \\ &\quad \left. \dots, (1 - e^{-t_n})^{1/2} Y_n^{kN}, (1 - e^{-t})^{1/2} Y_{n+1}^{kN} + e^{-t} X^N \otimes I_k, Z^N \otimes I_k \right), \end{aligned}$$

And similarly we define $Z_{k,r}^2$ and $Z_{k,r}^{2,s}$ but with $R_1^{kN}, \dots, R_s^{kN}$ instead of $S_1^{kN}, \dots, S_s^{kN}$. Thanks to Proposition 4.3.1, we get that

$$\begin{aligned} &k \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\tau_{kN} \otimes \tau_{kN} \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \right] - \mathbb{E}_k[\tau_{kN}] \otimes \mathbb{E}_k[\tau_{kN}] \left(\partial_i D_i Q(Z_k) P_{l,l'} \otimes P_{l,l'} \right) \\ &= k \sum_{1 \leq l, l' \leq k} \sum_{1 \leq s \leq n} \mathbb{E}_k \left[\tau_{kN} \left(\partial_i^1 D_i Q(Z_{k,0}^{1,s}) P_{l,l'} \boxtimes \partial_i^1 D_i Q(Z_{k,0}^{2,s}) P_{l,l'} \right) \right] \\ &\quad - \mathbb{E}_k \left[\tau_{kN} \left(\partial_i^1 D_i Q(Z_{k,t_s}^{1,s}) P_{l,l'} \boxtimes \partial_i^1 D_i Q(Z_{k,t_s}^{2,s}) P_{l,l'} \right) \right] \\ &\quad + \mathbb{E}_k \left[\tau_{kN} \left(\partial_i^1 D_i Q(Z_{k,0}^1) P_{l,l'} \boxtimes \partial_i^1 D_i Q(Z_{k,0}^2) P_{l,l'} \right) \right] \\ &\quad - \mathbb{E}_k \left[\tau_{kN} \left(\partial_i^1 D_i Q(Z_{k,t}^1) P_{l,l'} \boxtimes \partial_i^1 D_i Q(Z_{k,t}^2) P_{l,l'} \right) \right] \\ &= \frac{1}{k^2 N^3} \sum_{1 \leq l, l' \leq k} \sum_{\substack{1 \leq s \leq n, \\ 1 \leq j \leq d_s}} \int_0^{t_s} \mathbb{E}_k \left[\text{Tr}_{kN} \left(\partial_{s,j} \left(\partial_i^1 D_i Q \right) (Z_{k,r}^{1,s}) \tilde{\#} P_{l,l'} \boxtimes \partial_{s,j} \left(\partial_i^2 D_i Q \right) (Z_{k,r}^{2,s}) \tilde{\#} P_{l,l'} \right) \right] dr \\ &\quad + \sum_{1 \leq j \leq d} \int_0^t \mathbb{E}_k \left[\text{Tr}_{kN} \left(\partial_j \left(\partial_i^1 D_i Q \right) (Z_{k,r}^1) \tilde{\#} P_{l,l'} \boxtimes \partial_j \left(\partial_i^2 D_i Q \right) (Z_{k,r}^2) \tilde{\#} P_{l,l'} \right) \right] dr, \end{aligned}$$

where $\partial_{s,j}$ is the noncommutative differential as defined in 4.2.4 but with respect to $(1 - e^{-t_s})^{1/2} Y_{s,j}^{kN}$. Besides if U, V are matrices of $\mathbb{M}_N(\mathbb{C}) \otimes \mathbb{M}_k(\mathbb{C})$, then

$$\sum_{1 \leq l, l' \leq k} \text{Tr}_{kN}(U P_{l,l'} V P_{l,l'}) = \text{Tr}_N(I_N \otimes \text{Tr}_k(U) I_N \otimes \text{Tr}_k(V)).$$

Hence given $A, B, C, D \in \mathcal{F}_{d_1 + \dots + d_n + d, q}$,

$$\begin{aligned} &\frac{1}{k^2 N} \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\text{Tr}_{kN} \left(B(Z_{k,r}^{1,s}) P_{l,l'} A(Z_{k,r}^{1,s}) D(Z_{k,r}^{2,s}) P_{l,l'} C(Z_{k,r}^{2,s}) \right) \right] \\ &= \mathbb{E}_k \left[\text{Tr}_N \left(I_N \otimes \tau_k \left(A(Z_{k,r}^{1,s}) D(Z_{k,r}^{2,s}) \right) I_N \otimes \tau_k \left(C(Z_{k,r}^{2,s}) B(Z_{k,r}^{1,s}) \right) \right) \right]. \end{aligned}$$

If $Q = A(Z_{k,r}^{1,s})D(Z_{k,r}^{2,s})$ and $T = C(Z_{k,r}^{2,s})B(Z_{k,r}^{1,s})$, then thanks to Proposition 4.3.1,

$$\begin{aligned}
& \tau_N (\mathbb{E}_k [I_N \otimes \tau_k(Q)I_N \otimes \tau_k(T)]) \\
&= \frac{1}{Nk^2} \sum_{\substack{1 \leq i, j \leq N \\ 1 \leq l, l' \leq k}} \mathbb{E}_k [g_i^* \otimes f_l^* Q g_j \otimes f_l \times g_j^* \otimes f_{l'}^* T g_i^* \otimes f_{l'}^*] \\
&= \frac{1}{Nk^2} \sum_{\substack{1 \leq i, j, m, m' \leq N \\ 1 \leq l, l' \leq k}} \mathbb{E}_k [g_m^* \otimes f_l^* Q \times E_{j,i} \otimes I_k g_m \otimes f_l \times g_{m'}^* \otimes f_{l'}^* T \times E_{i,j} \otimes I_k g_{m'} \otimes f_{l'}] \\
&= \frac{1}{Nk^2} \sum_{1 \leq i, j \leq N} \mathbb{E}_k [\text{Tr}_{kN}(Q I_k \otimes E_{j,i}) \text{Tr}_{kN}(T I_k \otimes E_{i,j})] \\
&= \mathcal{O}(k^{-2}) + \frac{1}{Nk^2} \sum_{1 \leq i, j \leq N} \mathbb{E}_k [\text{Tr}_{kN}(Q I_k \otimes E_{j,i})] [\text{Tr}_{kN}(T I_k \otimes E_{i,j})] \\
&= \mathcal{O}(k^{-2}) + \tau_N(\mathbb{E}_k [I_N \otimes \tau_k(Q)] \mathbb{E}_k [I_N \otimes \tau_k(T)]).
\end{aligned}$$

We can view a GUE matrix of size kN as a matrix of size N with matrix coefficients. The diagonal coefficients are independent GUE matrices of size k multiplied by $N^{-1/2}$. The upper non-diagonal coefficients are independent random matrices of size k which have the same law as $(2N)^{-1/2}(X + \mathbf{i}Y)$ where X and Y are independent GUE matrices of size k , and the lower non-diagonal coefficients are the adjoints of the upper coefficients. Thus if U^{kN} is a family of l independent GUE matrices of size kN , u a family of l free semicircular variables, we then define \mathbf{u}^N as a family of l matrices of size N whose diagonal coefficients are free semicirculars multiplied by $N^{-1/2}$, and the upper non-diagonal coefficients are free between each other, free from the diagonal one, and they are of the form $(2N)^{-1/2}(a + \mathbf{i}b)$ where a and b are free semicirculars. Finally the lower non-diagonal coefficients are the adjoints of the upper coefficients. We also assume that semicirculars from different matrices are free and that all of those semicirculars live in a \mathcal{C}^* -algebra endowed with a trace τ . Then with \tilde{U}^{kN} an independent copy of U^{kN} , \tilde{u} a free copy of u and $\tilde{\mathbf{u}}^N$ a free copy of \mathbf{u}^N , for $L, K \in \mathcal{F}_{l,q}$,

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \tau_N \left(\mathbb{E}_k [I_N \otimes \tau_k(L(U^{kN}, Z^N \otimes I_k))] \mathbb{E}_k [I_N \otimes \tau_k(K(U^{kN}, Z^N \otimes I_k))] \right) \\
&= \tau_N \left(I_N \otimes \tau(L(\mathbf{u}^N, Z^N)) \times I_N \otimes \tau(K(\mathbf{u}^N, Z^N)) \right) \\
&= \tau_N \otimes \tau \left(L(\mathbf{u}^N, Z^N) \times K(\tilde{\mathbf{u}}^N, Z^N) \right) \\
&= \lim_{k \rightarrow \infty} \mathbb{E}_k \left[\tau_{kN} \left(L(U^{kN}, Z^N \otimes I_k) \times K(\tilde{U}^{kN}, Z^N \otimes I_k) \right) \right] \\
&= \tau_N \left(L(u, Z^N) \times K(\tilde{u}, Z^N) \right)
\end{aligned}$$

Consequently, we have that

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \frac{1}{k^2 N} \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\text{Tr}_{kN} \left(B(Z_{k,r}^{1,s}) P_{l',l} A(Z_{k,r}^{1,s}) D(Z_{k,r}^{2,s}) P_{l,l'} C(Z_{k,r}^{2,s}) \right) \right] \\
&= \tau_N \left(B(z_r^{1,s}) A(\tilde{z}_r^{1,s}) D(\tilde{z}_r^{2,s}) C(z_r^{2,s}) \right).
\end{aligned}$$

This means that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \frac{1}{k^2 N} \sum_{1 \leq l, l' \leq k} \mathbb{E}_k \left[\text{Tr}_{kN} \left(\partial_{s,j} \left(\partial_i^1 D_i Q \right) (Z_{k,r}^{1,s}) \tilde{\#} P_{l,l'} \boxtimes \partial_{s,j} \left(\partial_i^2 D_i Q \right) (Z_{k,r}^{2,s}) \tilde{\#} P_{l,l'} \right) \right] \\ &= \tau_N \left(\left(\partial_{s,j}^2 \left(\partial_i^1 D_i Q \right) (z_r^{1,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^1 D_i Q \right) (\tilde{z}_r^{1,s}) \right) \right. \\ & \quad \left. \boxtimes \left(\partial_{s,j}^2 \left(\partial_i^2 D_i Q \right) (\tilde{z}_r^{2,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^2 D_i Q \right) (z_r^{2,s}) \right) \right). \end{aligned}$$

Which in turn means that $\Lambda_{N,t}$ is equal to

$$\begin{aligned} & -\frac{1}{N^2} \sum_{\substack{1 \leq s \leq n, \\ 1 \leq j \leq d_s}} \int_0^{t_s} \tau_N \left(\left(\partial_{s,j}^2 \left(\partial_i^1 D_i Q \right) (z_r^{1,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^1 D_i Q \right) (\tilde{z}_r^{1,s}) \right) \right. \\ & \quad \left. \boxtimes \left(\partial_{s,j}^2 \left(\partial_i^2 D_i Q \right) (\tilde{z}_r^{2,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^2 D_i Q \right) (z_r^{2,s}) \right) \right) dr \\ & -\frac{1}{N^2} \sum_{1 \leq j \leq d} \int_0^t \tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i Q \right) (z_r^1) \boxtimes \partial_j^1 \left(\partial_i^1 D_i Q \right) (\tilde{z}_r^1) \right) \right. \\ & \quad \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i Q \right) (\tilde{z}_r^2) \boxtimes \partial_j^1 \left(\partial_i^2 D_i Q \right) (z_r^2) \right) \right) dr. \end{aligned}$$

Thus by using this result in equation (4.13), we have in conclusion

$$\begin{aligned} & \mathbb{E} \left[\frac{d}{dt} \tau_N \left(Q \left(Y_t^N \right) \right) \right] \\ &= -\frac{e^{-t}}{2N^2} \sum_{\substack{1 \leq i \leq d \\ 1 \leq s \leq n, \\ 1 \leq j \leq d_s}} \int_0^{t_s} \mathbb{E} \left[\tau_N \left(\left(\partial_{s,j}^2 \left(\partial_i^1 D_i Q \right) (z_r^{1,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^1 D_i Q \right) (\tilde{z}_r^{1,s}) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_{s,j}^2 \left(\partial_i^2 D_i Q \right) (\tilde{z}_r^{2,s}) \boxtimes \partial_{s,j}^1 \left(\partial_i^2 D_i Q \right) (z_r^{2,s}) \right) \right) \right] dr \\ & -\frac{e^{-t}}{2N^2} \sum_{1 \leq i, j \leq d} \int_0^t \mathbb{E} \left[\tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i Q \right) (z_r^1) \boxtimes \partial_j^1 \left(\partial_i^1 D_i Q \right) (\tilde{z}_r^1) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i Q \right) (\tilde{z}_r^2) \boxtimes \partial_j^1 \left(\partial_i^2 D_i Q \right) (z_r^2) \right) \right) \right] dr. \end{aligned}$$

□

4.3.3 Proof of Theorem 4.1.1

In this section we focus on proving Theorem 4.1.1 from which we deduce all of the important corollaries. It will mainly be a corollary of the following theorem, which is slightly stronger but less explicit. We refer to Definition 4.3.6 for the definition of L^{T_i} and Lemma 4.3.5 for the one of x^{T_i} .

Theorem 4.3.4. *Let the following objects be given,*

- $X^N = (X_1^N, \dots, X_d^N)$ independent GUE matrices of size N ,

- $Z^N = (Z_1^N, \dots, Z_q^N)$ deterministic matrices and their adjoints,
- $P \in \mathcal{A}_{d,q}$ a polynomial that we assume to be self-adjoint,
- $f : \mathbb{R} \mapsto \mathbb{R}$ such that there exists a measure on the real line μ with $\int (1+y^{4(k+1)}) d|\mu|(y) < +\infty$ and for any $x \in \mathbb{R}$,

$$f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y). \quad (4.20)$$

Then with notations as in Lemma 4.3.5 and Proposition 4.3.7 if we set,

$$\alpha_i^P(f, Z^N) = \int_{\mathbb{R}} \int_{A_i} \tau_N \left(\left(L^{\tilde{T}_1} \dots L^{\tilde{T}_i} \right) (e^{iyP})(x^{\tilde{T}_i}, Z^N) \right) dt_1 \dots dt_{2i} d\mu(y),$$

and that we write $P = \sum_{1 \leq i \leq Nb(P)} c_i M_i$ where the M_i are monomials and $c_i \in \mathbb{C}$, if we set $C_{\max}(P) = \max\{1, \max_i |c_i|\}$, then there exist constants C, K and c independent of P such that with $K_N = \max\{\|Z_1^N\|, \dots, \|Z_q^N\|, K\}$, for any N and $k \leq cN(\deg P)^{-1}$,

$$\begin{aligned} & \left| \mathbb{E} \left[\tau_N \left(f(P(X^N, Z^N)) \right) \right] - \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(f, Z^N) \right| \\ & \leq \frac{1}{N^{2k+2}} \int_{\mathbb{R}} (|y| + y^{4(k+1)}) d|\mu|(y) \times \left(C \times K_N^{\deg P} C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)} \times k^{3k}. \end{aligned} \quad (4.21)$$

Besides if we define \widehat{K}_N like K_N but with 2 instead of K , then we have that for any $j \in \mathbb{N}^*$,

$$|\alpha_j^P(f, Z^N)| \leq \int_{\mathbb{R}} (|y| + y^{4j}) d|\mu|(y) \times \left(C \times \widehat{K}_N^{\deg P} C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4j} \times j^{3j}. \quad (4.22)$$

Finally if f and g satisfies (4.20) and are bounded functions equal on a neighborhood of the spectrum of $P(x, Z^N)$, where x is a free semicircular system free from $\mathbb{M}_N(\mathbb{C})$, then for any i , $\alpha_i^P(f, Z^N) = \alpha_i^P(g, Z^N)$. In particular if f is a bounded function such that its support and the spectrum of $P(x, Z^N)$ are disjoint, then for any i , $\alpha_i^P(f, Z^N) = 0$.

The following lemma allows us to define the coefficients of the topological expansion by induction. It is basically a reformulation of Lemma 4.3.2 with the notations of Definitions 4.2.12 and 4.2.13. Although the notations in this formula are a bit heavy, such a formulation is necessary in order to get a better upper bound on the remainder term. It is the first step of the proof of Theorem 4.3.4.

Lemma 4.3.5. *Let x, y^1, \dots, y^{c_n} be free semicircular systems of d variables. Then with $T_n = \{t_1, \dots, t_{2n}\}$ an increasing sequence of non-negative number and $I = \{s_1, \dots, s_{2n}\} \in J_n$, with $t_0 = 0$, we set*

$$\begin{aligned} X_{i,I}^{N, T_n} &= \sum_{l=1}^{2n} (e^{-t_{l-1}} - e^{-t_l})^{1/2} y_i^{s_l} + e^{-t_n/2} X_i^N, \\ x_{i,I}^{T_n} &= \sum_{l=1}^{2n} (e^{-t_{l-1}} - e^{-t_l})^{1/2} y_i^{s_l} + e^{-t_n/2} x_i. \end{aligned}$$

We define the following subfamily of $(X_{i,I})_{i \in [1,d], I \in J_{n+1}}$,

$$X_{l,1} = (X_{i,I})_{i \in [1,d], I \in J_{n+1}^{l,1}}, X_{l,2} = (X_{i,I})_{i \in [1,d], I \in J_{n+1}^{l,2}},$$

$$\widetilde{X}_{l,1} = (X_{i,I})_{i \in [1,d], I \in \widetilde{J}_{n+1}^{l,1}}, \widetilde{X}_{l,2} = (\widetilde{X}_{i,I})_{i \in [1,d], I \in \widetilde{J}_{n+1}^{l,2}}.$$

Since there is a natural bijection between J_n and $J_{n+1}^{l,1}$, one can evaluate an element of $\mathcal{F}_{d,q}^n$ in $(X_{l,1}, Z)$ where $Z = (Y_1, \dots, Y_{2r})$ as in Definition 4.2.12, and similarly for $X_{l,2}, \widetilde{X}_{l,1}$ and $\widetilde{X}_{l,2}$. Then given a set T_{n+1} of $2n+2$ numbers, we define the following operators from $\mathcal{F}_{d,q}^n$ to $\mathcal{F}_{d,q}^{n+1}$, for l from 1 to $2n$,

$$L_l^{T_{n+1}}(Q) := \frac{1}{2} \sum_{\substack{1 \leq i, j \leq d \\ s \in [1, c_n], \\ \text{depth}(s)=l}} e^{-t_{2n+2}-t_{i-1}} \tau_N \left(\left(\partial_{j,s,n}^2 (\partial_i^1 D_i Q) (X_{l,1}, Z) \boxtimes \partial_{j,s,n}^1 (\partial_i^1 D_i Q) (\widetilde{X}_{l,1}, Z) \right) \right. \\ \left. \boxtimes \left(\partial_{j,s,n}^2 (\partial_i^2 D_i Q) (\widetilde{X}_{l,2}, Z) \boxtimes \partial_{j,s,n}^1 (\partial_i^2 D_i Q) (X_{l,2}, Z) \right) \right).$$

We also define

$$L_{2n+1}^{T_{n+1}}(Q) := \frac{1}{2} \sum_{1 \leq i, j \leq d} e^{-t_{2n+2}-t_{2n}} \tau_N \left(\left(\partial_j^2 (\partial_i^1 D_i Q) (X_{2n+1,1}, Z) \boxtimes \partial_j^1 (\partial_i^1 D_i Q) (\widetilde{X}_{2n+1,1}, Z) \right) \right. \\ \left. \boxtimes \left(\partial_j^2 (\partial_i^2 D_i Q) (\widetilde{X}_{2n+1,2}, Z) \boxtimes \partial_j^1 (\partial_i^2 D_i Q) (X_{2n+1,2}, Z) \right) \right).$$

And finally we set

$$L^{T_{n+1}}(Q) := \sum_{1 \leq l \leq 2n} \mathbf{1}_{[t_{l-1}, t_l]}(t_{2n+1}) L_l^{T_{n+1}}(Q) + \mathbf{1}_{[t_{2n}, t_{2n+2}]}(t_{2n+1}) L_{2n+1}^{T_{n+1}}(Q). \quad (4.23)$$

Given $Q \in \mathcal{F}_{d,q}^n$, if T_{n+1} a set of $2n+2$ numbers, then \widetilde{T}_{n+1} is the same set but whose elements have been sorted by increasing order, then ,

$$\mathbb{E} \left[\tau_N \left(Q(X^{N, T_n}, Z^N) \right) \right] - \tau_N \left(Q(x^{T_n}, Z^N) \right) \\ = \int_{t_{2n}}^{\infty} \int_0^{t_{2n+2}} \tau_N \left(L^{\widetilde{T}_{n+1}}(Q) \left(X^{N, \widetilde{T}_{n+1}}, Z^N \right) \right) dt_{2n+1} dt_{2n+2}.$$

Proof. With the notations of Lemma 4.3.2, let $S \in \mathcal{F}_{d(c_n+1), q}$ be such that

$$S \left((1 - e^{-(t_1-t_0)})(y^s)_{\{s | \text{depth}^n(s)=1\}}, \dots, (1 - e^{-(t_{2n}-t_{2n-1})})(y^s)_{\{s | \text{depth}^n(s)=2n\}}, X^N, Z^N \right) \\ = Q \left(\left(\sum_{l=1}^{2n} e^{-t_{l-1}/2} (1 - e^{-(t_l-t_{l-1})})^{1/2} y_i^{s_l} + e^{-t_{2n}/2} X_i^N \right)_{I \in J_n}, Z^N \right)$$

Consequently $\partial_i D_i S = e^{-t_{2n}} \partial_i D_i Q$ where on the left side we used the non-commutative differential defined in 4.2.4 which is also the one used in Lemma 4.3.2, whereas on the right side we used the non-commutative differential defined in Definition 4.2.13. Thus with the convention

of Definition 4.2.15, we set d_l to be d times the number of $s \in J_n$ which have depth l , then

$$\begin{aligned} & \mathbb{E} \left[\tau_N \left(S \left(Y^N \right) \right) \right] - \tau_N \left(S \left(Y \right) \right) \\ &= \frac{1}{2N^2} \int_0^\infty e^{-t} \sum_{\substack{1 \leq i \leq d \\ 1 \leq l \leq 2n, \\ 1 \leq j \leq d_i}} \int_0^{t_l - t_{l-1}} \mathbb{E} \left[\tau_N \left(\left(\partial_{l,j}^2 \left(\partial_i^1 D_i S \right) \left(z_r^{1,l} \right) \boxtimes \partial_{l,j}^1 \left(\partial_i^1 D_i S \right) \left(\tilde{z}_r^{1,l} \right) \right. \right. \right. \\ & \quad \left. \left. \left. \boxtimes \left(\partial_{l,j}^2 \left(\partial_i^2 D_i S \right) \left(\tilde{z}_r^{2,l} \right) \boxtimes \partial_{l,j}^1 \left(\partial_i^2 D_i S \right) \left(z_r^{2,l} \right) \right) \right) \right] dr dt \\ &+ \frac{1}{2N^2} \int_0^\infty e^{-t} \sum_{1 \leq i, j \leq d} \int_0^t \mathbb{E} \left[\tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i S \right) \left(z_r^1 \right) \boxtimes \partial_j^1 \left(\partial_i^1 D_i S \right) \left(\tilde{z}_r^1 \right) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i S \right) \left(\tilde{z}_r^2 \right) \boxtimes \partial_j^1 \left(\partial_i^2 D_i S \right) \left(z_r^2 \right) \right) \right) \right] dr dt. \end{aligned}$$

Consequently, since $S(Y^N) = Q(X^{N, T_n}, Z^N)$ and $S(Y) = Q(x^{T_n}, Z^N)$, if we set

$$\begin{aligned} X_{l,1}^{N, T_{n+1}} &= \left(X_I^{N, T_{n+1}} \right)_{I \in J_{n+1}^{l,1}}, X_{l,2}^{N, T_{n+1}} = \left(X_I^{N, T_{n+1}} \right)_{I \in J_{n+1}^{l,2}}, \\ \tilde{X}_{l,1}^{N, T_{n+1}} &= \left(X_I^{N, T_{n+1}} \right)_{I \in \tilde{J}_{n+1}^{l,1}}, \tilde{X}_{l,2}^{N, T_{n+1}} = \left(\tilde{X}_I^{N, T_{n+1}} \right)_{I \in \tilde{J}_{n+1}^{l,2}}, \end{aligned}$$

we have that

$$\begin{aligned} & \mathbb{E} \left[\tau_N \left(Q \left(X^{N, T_n}, Z^N \right) \right) \right] - \tau_N \left(Q \left(x^{T_n}, Z^N \right) \right) \\ &= \frac{1}{2N^2} \int_0^\infty e^{-t-2t_n} \sum_{\substack{1 \leq i, j \leq d \\ 1 \leq l \leq 2n}} \sum_{\substack{s \in [1, c_n], \\ \text{depth}(s)=l}} e^{-t_{l-1}} \int_0^{t_l - t_{l-1}} \\ & \quad \mathbb{E} \left[\tau_N \left(\left(\partial_{j,s,n}^2 \left(\partial_i^1 D_i Q \right) \left(X_{l,1}^{N, \{T_n, r+t_{l-1}, t+t_{2n}\}} \right) \boxtimes \partial_{j,s,n}^1 \left(\partial_i^1 D_i Q \right) \left(\tilde{X}_{l,1}^{N, \{T_n, r+t_{l-1}, t+t_{2n}\}} \right) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_{j,s,n}^2 \left(\partial_i^2 D_i Q \right) \left(\tilde{X}_{l,2}^{N, \{T_n, r+t_{l-1}, t+t_{2n}\}} \right) \boxtimes \partial_{j,s,n}^1 \left(\partial_i^2 D_i Q \right) \left(X_{l,2}^{N, \{T_n, r+t_{l-1}, t+t_{2n}\}} \right) \right) \right) \right] dr dt \\ &+ \frac{1}{2N^2} \int_0^\infty e^{-t-2t_{2n}} \sum_{1 \leq i, j \leq d} \int_0^t \\ & \quad \mathbb{E} \left[\tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i Q \right) \left(X_{n+1,1}^{N, \{T_n, r+t_{2n}, t+t_{2n}\}} \right) \boxtimes \partial_j^1 \left(\partial_i^1 D_i Q \right) \left(\tilde{X}_{n+1,1}^{N, \{T_n, r+t_{2n}, t+t_{2n}\}} \right) \right) \right. \right. \\ & \quad \left. \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i Q \right) \left(\tilde{X}_{n+1,2}^{N, \{T_n, r+t_{2n}, t+t_{2n}\}} \right) \boxtimes \partial_j^1 \left(\partial_i^2 D_i Q \right) \left(X_{n+1,2}^{N, \{T_n, r+t_{2n}, t+t_{2n}\}} \right) \right) \right) \right] dr dt. \end{aligned}$$

Thus if we set $T_{n+1} = \{T_n, r, t\}$, after a change of variable, we get that

$$\begin{aligned}
 & \mathbb{E} \left[\tau_N \left(Q(X^{N, T_n}, Z^N) \right) \right] - \tau_N \left(Q(x^{T_n}, Z^N) \right) \\
 &= \frac{1}{2N^2} \sum_{\substack{1 \leq i, j \leq d \\ 1 \leq l \leq 2n \\ s \in [1, c_n], \\ \text{depth}(s)=l}} \int_{t_{2n}}^{\infty} e^{-t-t_{l-1}} \int_{t_{l-1}}^{t_l} \mathbb{E} \left[\tau_N \left(\left(\partial_{j,s,n}^2 \left(\partial_i^1 D_i Q \right) \left(X_{l,1}^{N, T_{n+1}} \right) \boxtimes \partial_{j,s,n}^1 \left(\partial_i^1 D_i Q \right) \left(\widetilde{X}_{l,1}^{N, T_{n+1}} \right) \right) \right. \right. \\
 & \quad \left. \left. \boxtimes \left(\partial_{j,s,n}^2 \left(\partial_i^2 D_i Q \right) \left(\widetilde{X}_{l,2}^{N, T_{n+1}} \right) \boxtimes \partial_{j,s,n}^1 \left(\partial_i^2 D_i Q \right) \left(X_{l,2}^{N, T_{n+1}} \right) \right) \right) \right] dr dt \\
 &+ \frac{1}{2N^2} \int_{t_{2n}}^{\infty} e^{-t-t_{2n}} \sum_{1 \leq i, j \leq d} \int_{t_{2n}}^t \mathbb{E} \left[\tau_N \left(\left(\partial_j^2 \left(\partial_i^1 D_i Q \right) \left(X_{n+1,1}^{N, T_{n+1}} \right) \boxtimes \partial_j^1 \left(\partial_i^1 D_i Q \right) \left(\widetilde{X}_{n+1,1}^{N, T_{n+1}} \right) \right) \right. \right. \\
 & \quad \left. \left. \boxtimes \left(\partial_j^2 \left(\partial_i^2 D_i Q \right) \left(\widetilde{X}_{n+1,2}^{N, T_{n+1}} \right) \boxtimes \partial_j^1 \left(\partial_i^2 D_i Q \right) \left(X_{n+1,2}^{N, T_{n+1}} \right) \right) \right) \right] dr dt.
 \end{aligned}$$

Hence the conclusion by renaming r in t_{2n+1} and t in t_{2n+2} . \square

This prompts us to define the following operator:

Definition 4.3.6. *In order to be perfectly rigorous, as pointed out in Definition 4.2.8, we have to define $\widetilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{T_{n+1}} : \mathcal{F}_{d,q}^n \rightarrow \mathcal{F}_{d,q}^{n+1}$ as in (4.23) but where we replaced every non-commutative differential by their equivalent defined in 4.2.8. Then given $x = (x_I)_{I \in J_{n+1}}$ and z elements of a C^* -algebra, we define $L^{T_{n+1}}(Q)(x, z)$ as the integral of $\widetilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{T_{n+1}}(Q)(x, z)$ over $\alpha_n, \beta_n, \gamma_n$ and δ_n .*

Thus we get directly the following proposition.

Proposition 4.3.7. *Let x be a free semicircular system, $(y^i)_{i \geq 1}$ be free semicircular systems free from x , and X^N be independent GUE matrices. We define X^{N, T_n} and x^{T_n} as in 4.3.5, $A_i = \{t_{2i} \geq t_{2i-2} \geq \dots \geq t_2 \geq 0\} \cap \{\forall s \in [1, i], t_{2s} \geq t_{2s-1} \geq 0\} \subset \mathbb{R}^{2i}$, then for any $Q \in \mathcal{F}_{d,q}$,*

$$\begin{aligned}
 & \mathbb{E} \left[\tau_N \left(Q(X^N, Z^N) \right) \right] \\
 &= \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \int_{A_i} \tau_N \left(\left(L^{\widetilde{T}_i} \dots L^{\widetilde{T}_1} \right) (Q)(x^{\widetilde{T}_i}, Z^N) \right) dt_1 \dots dt_{2i} \\
 &+ \frac{1}{N^{2(k+1)}} \int_{A_{k+1}} \mathbb{E} \left[\tau_N \left(\left(L^{\widetilde{T}_{k+1}} \dots L^{\widetilde{T}_1} \right) (Q)(X^{N, \widetilde{T}_{k+1}}, Z^N) \right) \right] dt_1 \dots dt_{2(k+1)}.
 \end{aligned}$$

Before giving the proof of Theorem 4.3.4, as mentioned in the introduction, the former proposition gives some insight in map enumeration.

Remark 4.3.8. *We say that a graph on a surface is a map if it is connected and its faces are homeomorphic to discs. It is of genus g if it can be embedded in a surface of genus g but not $g-1$. For an edge-colored graph on an orientated surface we say that a vertex is of type $q = X_{i_1} \dots X_{i_p}$ if it has degree p and when we look at the half-edges going out of it, starting from a distinguished one and going in the clockwise order the first half-edge is of color i_1 , the*

second i_2 , and so on. If $\mathcal{M}_g(X_{i_1} \dots X_{i_p})$ is the number of such maps of genus g with a single vertex, then given X_i^N independent GUE matrices

$$\mathbb{E} \left[\frac{1}{N} \operatorname{Tr}_N \left(X_{i_1}^N \dots X_{i_p}^N \right) \right] = \sum_{g \in \mathbb{N}} \frac{1}{N^{2g}} \mathcal{M}_g(X_{i_1} \dots X_{i_p}).$$

For a proof we refer to [33] for the one matrix case and [60], chapter 22, for the multimatrix case. Thanks to Proposition 4.3.7, we immediately get that

$$\mathcal{M}_g(X_{i_1} \dots X_{i_p}) = \int_{A_g} \tau \left(L^{\tilde{T}_g} \dots L^{\tilde{T}_1} \left(X_{i_1} \dots X_{i_p} \right) (x^{\tilde{T}_g}) \right) dt_1 \dots dt_{2g}.$$

We can now prove Theorem 4.3.4.

Proof of Theorem 4.3.4. Thanks to Proposition 4.3.7, we immediately get that

$$\begin{aligned} \mathbb{E} \left[\tau_N \left(f(P(X^N, Z^N)) \right) \right] &= \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(f, Z^N) \\ &+ \frac{1}{N^{2(k+1)}} \int_{\mathbb{R}} \int_{A_{k+1}} \mathbb{E} \left[\tau_N \left(\left(L^{\tilde{T}_{k+1}} \dots L^{\tilde{T}_1} \right) (e^{iyP})(X^{N, \tilde{T}_{k+1}}, Z^N) \right) \right] dt_1 \dots dt_{2(k+1)} d\mu(y). \end{aligned}$$

All we need to do from now on is to get an estimate on the last line. To do so we use the following remark. Let $Q \in \mathcal{F}_{d,q}^n$, then we can write

$$Q = \sum_{1 \leq i \leq Nb(Q)} c_i M_i$$

where $c_i \in \mathbb{C}$ and $M_i \in \mathcal{F}_{d,q}^n$ are monomials (not necessarily distinct). We also define $C_{\max}(Q) = \max\{1, \sup_i |c_i|\}$. Since for any $I \in J_n$, $\|X_{i,I}^{N, T_n}\| \leq 2 + \|X_i^N\|$, given

$$\mathcal{B} = \left\{ 2 + \|X_i^N\| \right\}_{1 \leq i \leq p} \cup \left\{ \|Z_j^N\| \right\}_{1 \leq j \leq q},$$

and D_N the maximum of this family, we get that

$$\|Q(X^{N, T_n}, Z^N)\| \leq Nb(Q) \times C_{\max}(Q) \times D_N^{\deg(Q)}. \quad (4.24)$$

It is worth noting that this upper bound is not optimal at all and heavily dependent on the decomposition chosen. Now let us consider $\tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}}$ defined as in Definition 4.3.6. We also consider $\tilde{\mathcal{F}}_{d,q}^n$ the $*$ -algebra generated by $\mathcal{A}_{d,q}^n$ and the family

$$\left\{ e^{i\lambda y P((X_{i,I})_{1 \leq i \leq d, Z})} \mid I \in J_n, \lambda \in [0, 1] \right\}.$$

Then $\tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}}$ send $\tilde{\mathcal{F}}_{d,q}^n$ to $\tilde{\mathcal{F}}_{d,q}^{n+1}$. Let $Q \in \tilde{\mathcal{F}}_{d,q}^n$, then we get that

$$\deg \left(\tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}}(Q) \right) \leq \deg Q + 4 \deg P,$$

$$C_{\max} \left(\tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}}(Q) \right) \leq \frac{e^{-t_{2n+2}}}{2} (1 + |y|)^4 C_{\max}(P)^4 C_{\max}(Q),$$

$$\begin{aligned} Nb\left(\tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}}(Q)\right) &\leq \deg(Q)(\deg Q + \deg P)(\deg Q + 2 \deg P) \\ &\quad \times (\deg Q + 3 \deg P) \times (Nb(P) \deg P)^4 \times Nb(Q). \end{aligned}$$

Thus if we define by induction $Q_0 = e^{iyP}$, and $Q_{n+1} = \tilde{L}_{\alpha_n, \beta_n, \gamma_n, \delta_n}^{\tilde{T}_{n+1}} Q_n$, by a straightforward induction we get that

$$\deg Q_n \leq 4n \deg P \quad (4.25)$$

$$C_{\max}(Q_n) \leq \frac{\prod_{r=1}^n e^{-t_{2r}}}{2^n} (1 + |y|)^{4n} C_{\max}(P)^{4n} \quad (4.26)$$

$$Nb(Q_n) \leq \left(Nb(P)(\deg P)^2\right)^{4n} (4n)! \quad (4.27)$$

Actually since we have $D_{\delta_{1,i}} e^{iyP} = iy \partial_{\delta_{1,i}} P \# e^{iyP}$, one can replace $(1 + |y|)^{4n}$ in equation (4.26) by $|y|(1 + |y|)^{4n-1}$. Thus thanks to (4.24), we get that

$$\begin{aligned} &\left\| \tilde{L}_{\alpha_{k+1}, \beta_{k+1}, \gamma_{k+1}, \delta_{k+1}}^{\tilde{T}_{k+1}} \cdots \tilde{L}_{\alpha_1, \beta_1, \gamma_1, \delta_1}^{\tilde{T}_1} Q(X^{N, \tilde{T}_{k+1}}, Z^N) \right\| \\ &\leq \prod_{r=1}^{k+1} \frac{e^{-t_{2r}}}{2} \times \frac{|y|}{1 + |y|} \\ &\quad \times \left((1 + |y|) C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)} (4(k+1))! \times D_N^{4(k+1) \deg P} \end{aligned}$$

Consequently after integrating over $\alpha_n, \beta_n, \gamma_n, \delta_n$, we get that

$$\begin{aligned} &\left| \int_{\mathbb{R}} \int_{A_{k+1}} \mathbb{E} \left[\tau_N \left(\left(L^{\tilde{T}_{k+1}} \cdots L^{\tilde{T}_1} \right) (e^{iyP}) (X^{N, \tilde{T}_{k+1}}, Z^N) \right) \right] dt_1 \cdots dt_{2(k+1)} d\mu(y) \right| \\ &\leq \int_{t_{k+1} \geq t_{2k} \geq \dots t_2 \geq 0} \prod_{r=1}^{k+1} \frac{t_{2r} e^{-t_{2r}}}{2} dt_2 dt_4 \cdots dt_{2k+2} \times \int_{\mathbb{R}} |y|(1 + |y|)^{4k+3} d|\mu|(y) \\ &\quad \times \left(C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)} (4(k+1))! \times \mathbb{E} \left[D_N^{4(k+1) \deg P} \right]. \end{aligned}$$

Besides

$$\begin{aligned} \int_{t_{2k+2} \geq t_{2k} \geq \dots t_2 \geq 0} \prod_{r=1}^{k+1} \frac{t_{2r} e^{-t_{2r}}}{2} dt_1 \cdots dt_{k+1} &= 2^{-k-1} \int_{0 \leq t_1 \leq \dots \leq t_{k+1}} \prod_{r=1}^{k+1} t_r e^{-t_r} dt_1 \cdots dt_{k+1} \\ &\leq 2^{-k-1} \int_{0 \leq t_1 \leq \dots \leq t_{k+1}} \prod_{r=1}^{k+1} e^{-t_r/2} dt_1 \cdots dt_{k+1} \\ &= \int_{0 \leq t_1 \leq \dots \leq t_{k+1}} \prod_{r=1}^{k+1} e^{-t_r} dt_1 \cdots dt_{k+1} \\ &= \frac{1}{(k+1)!}, \end{aligned}$$

and

$$\int_{\mathbb{R}} |y|(1 + |y|)^{4k+3} d|\mu|(y) \leq 2^{4k+3} \int_{\mathbb{R}} (|y| + y^{4(k+1)}) d|\mu|(y).$$

Thanks to Proposition 4.2.18 we can find constants K and c such that with

$$K_N = \max\{K, \|Z_1^N\|, \dots, \|Z_q^N\|\},$$

then for any $k \leq c(\deg P)^{-1}N$,

$$\mathbb{E} \left[D_N^{4(k+1) \deg P} \right] \leq K_N^{4(k+1) \deg P}.$$

Thus thanks to Stirling formula, there exists a constant C such that

$$\begin{aligned} & \left| \int_{\mathbb{R}} \int_{A_{k+1}} \mathbb{E} \left[\tau_N \left(\left(L^{\tilde{T}_{k+1}} \dots L^{\tilde{T}_1} \right) (e^{iyP})(X^{N, \tilde{T}_{k+1}}, Z^N) \right) \right] dt_1 \dots dt_{2(k+1)} d\mu(y) \right| \\ & \leq \int_{\mathbb{R}} (|y| + y^{4(k+1)}) d|\mu|(y) \times \left(C \times K_N^{\deg P} C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)} \times k^{3k}. \end{aligned}$$

Hence we get equation (4.21). We get equation (4.22) very similarly. Finally to prove the last affirmation, we only need to consider a function which takes value 0 on a neighborhood of the spectrum of $P(x, Z^N)$. Let X^{lN} be independent GUE matrices of size lN , then we get that for any k such that f is smooth enough, thanks to equation (4.21),

$$\mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] = \sum_{0 \leq i \leq k} \frac{1}{(lN)^{2i}} \alpha_i^P(f, Z^N \otimes I_l) + \mathcal{O}(l^{-2(k+1)}).$$

But in the sense of Definition 4.2.1, for any i , $(x^{T_i}, Z^N \otimes I_l)$ and (x^{T_i}, Z^N) have the same distribution, hence

$$\mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] = \sum_{0 \leq i \leq k} \frac{1}{(lN)^{2i}} \alpha_i^P(f, Z^N) + \mathcal{O}(l^{-2(k+1)}).$$

Consequently, if there exists i such that $\alpha_i^P(f, Z^N) \neq 0$, then we can find constants c and k (dependent on N) such that

$$\mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] \sim_{l \rightarrow \infty} c \times l^{-2k}. \quad (4.28)$$

We are going to show that the left hand side decays exponentially fast in l , hence proving a contradiction. Now if we set E the support of f , then

$$\left| \mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] \right| \leq \|f\|_{\infty} \mathbb{P} \left(\sigma \left(P(X^{lN}, Z^N \otimes I_l) \right) \cap E \neq \emptyset \right).$$

However thanks to Proposition 4.2.18, there exist constants A and B such that for any l ,

$$\mathbb{P} \left(\|P(X^{lN}, Z^N \otimes I_l)\| \geq A \right) \leq e^{-Bl}.$$

Thus,

$$\left| \mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] \right| \leq \|f\|_{\infty} \left(\mathbb{P} \left(\sigma \left(P(X^{lN}, Z^N \otimes I_l) \right) \cap E \cap [-A, A] \neq \emptyset \right) + e^{-Bl} \right).$$

Let g be a \mathcal{C}^∞ -function, with compact support disjoint from the spectrum of $P(x, Z^N)$ such that $g|_{E \cap [-A, A]} = 1$. Then,

$$\left| \mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] \right| \leq \|f\|_\infty \mathbb{P} \left(\|g(P(X^{lN}, Z^N \otimes I_l))\| \geq 1 \right) + e^{-Bl}.$$

Since g is \mathcal{C}^∞ and has compact support, thanks to the Fourier transform it satisfies (4.20) for any k , thus for any self-adjoint matrices U and V ,

$$\begin{aligned} \|g(U) - g(V)\| &= \left\| \int y \int_0^1 e^{iyU\alpha}(U - V)e^{iyV(1-\alpha)} d\alpha d\mu(y) \right\| \\ &\leq \|U - V\| \int |y| d|\mu|(y). \end{aligned}$$

Hence there is a constant C_B such that for any self-adjoint matrices $X_i, Y_i \in \mathbb{M}_{lN}(\mathbb{C})$ whose operator norm is bounded by B ,

$$\|g(P(X, Z^N)) - g(P(Y, Z^N))\| \leq C_B \sum_i \|X_i - Y_i\|.$$

Consequently, with a proof very similar to the one of Proposition 4.6 of [7], we get that there exist constant D and S such that for any $\delta > 0$,

$$\mathbb{P} \left(\left| \|g(P(X^{lN}, Z^N \otimes I_l))\| - \mathbb{E} \left[\|g(P(X^{lN}, Z^N \otimes I_l))\| \right] \right| \geq \delta + De^{-N} \right) \leq pe^{-2N} + e^{-S\delta^2l}.$$

But then thanks to Theorem 1.6 of [28] and Weierstrass theorem, we know that almost surely $\|g(P(X^{lN}, Z^N \otimes I_l))\|$ converges towards 0. Hence thanks to Proposition 4.2.18 and dominated convergence theorem, we get that $\mathbb{E} \left[\|g(P(X^{lN}, Z^N \otimes I_l))\| \right]$ also converges towards 0. Hence for l large enough, there exist a constant S such that

$$\mathbb{P} \left(\|g(P(X^{lN}, Z^N \otimes I_l))\| \geq 1 \right) \leq e^{-Sl}.$$

Consequently we get that there exist constants A and B such that

$$\left| \mathbb{E} \left[\tau_N \left(f(P(X^{lN}, Z^N \otimes I_l)) \right) \right] \right| \leq Ae^{-Bl},$$

which is in contradiction with equation (4.28). Hence the conclusion. \square

We can now prove Theorem 4.1.1, the only difficulty of the proof is to use the hypothesis of smoothness to replace our function f by a function which satisfies (4.20) without losing too much on the constants.

Proof of Theorem 4.1.1. To begin with, let

$$h : x \rightarrow \begin{cases} e^{-x^{-4} - (1-x)^{-4}} & \text{if } x \in (0, 1), \\ 0 & \text{else.} \end{cases}$$

Let H be the primitive of h which takes value 0 on \mathbb{R}^- and renormalized such that it takes value 1 for $x \geq 1$. Then given a constant m one can define the function $g : x \rightarrow H(m+1-x)H(m+1+x)$ which takes value 1 on $[-m, m]$ and 0 outside of $(-m-1, m+1)$. Let B be the union over i

of the events $\{\|X_i^N\| \geq D + \alpha^{-1}\}$ where D and α where defined in Proposition 4.2.18. Thus $\mathbb{P}(B) \leq pe^{-N}$. By adjusting the constant K defined in Theorem 4.3.4 we can always assume that it is larger than $D + \alpha^{-1}$, thus if for any i , $\|X_i^N\| \leq D + \alpha^{-1}$, $\|P(X^N, Z^N)\| \leq \mathbf{m}C_{\max}K_N^n$. We fix $m = \mathbf{m}C_{\max}K_N^n$, thus if $P(X^N, Z^N)$ has an eigenvalue outside of $[-m, m]$, necessarily $X^N \in B$. Thus

$$\mathbb{E} \left[\tau_N \left(f(1-g)(P(X^N, Z^N)) \right) \right] \leq \|f\|_{\infty} \mathbb{P}(B) \leq \|f\|_{\infty} p \times e^{-N}. \quad (4.29)$$

Since fg has compact support and is a function of class $\mathcal{C}^{4(k+1)+2}$, we can take its Fourier transform and then invert it so that with the convention $\hat{h}(y) = \frac{1}{2\pi} \int_{\mathbb{R}} h(x)e^{-ixy}dx$, we have

$$\forall x \in \mathbb{R}, \quad (fg)(x) = \int_{\mathbb{R}} e^{ixy} \widehat{fg}(y) dy.$$

Besides, since if h has compact support bounded by $m+1$ then $\|\hat{h}\|_0 \leq \frac{1}{\pi}(m+1)\|h\|_0$, we have

$$\begin{aligned} \int_{\mathbb{R}} (|y| + y^{4(k+1)}) |\widehat{fg}(y)| dy &\leq \int_{\mathbb{R}} \frac{\sum_{i=0}^{4(k+1)+2} |y|^i}{1+y^2} |\widehat{fg}(y)| dy \\ &\leq \int_{\mathbb{R}} \frac{\sum_{i=0}^{4(k+1)+2} |\widehat{(fg)^{(i)}}(y)|}{1+y^2} dy \\ &\leq \frac{1}{\pi} (m+1) \|fg\|_{\mathcal{C}^{4(k+1)+2}} \int_{\mathbb{R}} \frac{1}{1+y^2} dy \\ &\leq (m+1) \|fg\|_{\mathcal{C}^{4(k+1)+2}}, \end{aligned}$$

Hence fg satisfies the hypothesis of Theorem 4.3.4 with $\mu(dy) = \widehat{fg}(y)dy$. Therefore, combining with equation (4.29), by adjusting the constant C , we get that

$$\begin{aligned} &\left| \mathbb{E} \left[\tau_N \left(f(P(X^N, Z^N)) \right) \right] - \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(fg, Z^N) \right| \\ &\leq \frac{1}{N^{2k+2}} \|fg\|_{\mathcal{C}^{4(k+1)+2}} \times \left(C \times K_N^{\deg P} C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)+1} \times k^{3k}. \end{aligned}$$

Since the norm of the family Z^N is uniformly bounded over N , the second line is of order N^{-2k-2} .

Hence if $(\beta_i)_{1 \leq i \leq k}$ is a family of scalar such that $\mathbb{E} \left[\tau_N \left(f(P(X^N, Z^N)) \right) \right] - \sum_{0 \leq i \leq k} N^{-2i} \beta_i$ is also of order N^{-2k-2} , then so is $\sum_{0 \leq i \leq k} (\alpha_i^P(fg, Z^N) - \beta_i) N^{-2i}$. Thus for any i , $\beta_i = \alpha_i^P(fg, Z^N)$. Thus since it does not depends on g , one can set $\alpha_i^P(f, Z^N) = \alpha_i^P(fg, Z^N)$.

Finally, one can write the j -th derivative of $x \rightarrow e^{-x^{-4}}$ on \mathbb{R}^+ as $x \rightarrow Q_j(x^{-1})e^{-x^{-4}}$ for some polynomial Q_j . By studying $Nb(Q_j)$, $C_{\max}(Q_j)$ and $\deg(Q_j)$, as in the proof of Theorem 4.3.4, we get that the infinity norm of the j -th derivative of this function is smaller than $20^j j! (5j/4)^{5j/4}$. Hence by adjusting C and using Stirling formula,

$$\begin{aligned} &\left| \mathbb{E} \left[\tau_N \left(f(P(X^N, Z^N)) \right) \right] - \sum_{0 \leq i \leq k} \frac{1}{N^{2i}} \alpha_i^P(fg, Z^N) \right| \\ &\leq \frac{1}{N^{2k+2}} \|f\|_{\mathcal{C}^{4(k+1)+2}} \times \left(C \times K_N^{\deg P} C_{\max}(P) Nb(P) (\deg P)^2 \right)^{4(k+1)+1} \times k^{12k}. \end{aligned}$$

The other points of the Theorem are a direct consequence of Theorem 4.3.4. \square

4.4 Consequences of Theorem 4.3.4

4.4.1 Proof of corollary 4.1.2

Let g be a non-negative C^∞ -function which takes value 0 on $(-\infty, 1/2]$, 1 on $[1, \infty)$ and in $[0, 1]$ elsewhere. For any $a, b \in \mathbb{R}$, we define $h_{[a,b]}^\varepsilon : x \mapsto g(\varepsilon^{-1}(x - a))g(-\varepsilon^{-1}(x - b))$. Then let \mathcal{I} be the collection of connected components of the complementary set of $\sigma(P(x, A^N))$. Then we define

$$h^\varepsilon = \sum_{I \in \mathcal{I}} h_I^\varepsilon.$$

This function is well-defined since the spectrum of $P(x, A^N)$ is compact, hence its complementary set has a finite number of connected components of measure larger than ε . And since if $b - a \leq \varepsilon$, $h_{[a,b]}^\varepsilon = 0$, the sum over $I \in \mathcal{I}$ is actually a finite sum. Besides, we have that

$$\begin{aligned} \mathbb{P} \left(\sigma(P(X^N, A^N)) \not\subset \sigma(P(x, A^N)) + \varepsilon \right) &\leq \mathbb{P} \left(\|h^\varepsilon(P(X^N, A^N))\| \geq 1 \right) \\ &\leq \mathbb{E} \left[\text{Tr}_N \left(h^\varepsilon(P(X^N, A^N)) \right) \right]. \end{aligned}$$

Besides thanks to Theorem 4.1.1 since the spectrum of $P(x, A^N)$ and the support of h^ε are disjoint, and that the operator norm of the matrices A^N is uniformly bounded over N , for any $k \in \mathbb{N}$, we get that there is a constant C_k such that for any ε and for N large enough,

$$\mathbb{E} \left[\text{Tr}_N \left(h^\varepsilon(P(X^N, A^N)) \right) \right] \leq C_k \frac{\varepsilon^{-4k-2}}{N^{2k-1}}.$$

Thus if we set $\varepsilon = N^{-\alpha}$ with $\alpha < 1/2$, then by fixing k large enough we get that

$$\mathbb{P} \left(\sigma(P(X^N, A^N)) \not\subset \sigma(P(x, A^N)) + N^{-\alpha} \right) = \mathcal{O}(N^{-2}).$$

Hence the conclusion by Borel-Cantelli lemma.

4.4.2 Proof of Corollary 4.1.3

Firstly, we need the following lemma.

Lemma 4.4.1. *Let g be a C^∞ function which takes value 0 on $(-\infty, 1/2]$ and value 1 on $[1, \infty)$, and in $[0, 1]$ otherwise. We set $f_\varepsilon : t \mapsto g(\varepsilon^{-1}(t - \alpha))$ with $\alpha = \|PP^*(x, A^N)\|$, then there exist constants C and c such that for any $k \leq cN$, $\varepsilon > 0$ and N ,*

$$\mathbb{E} \left[\text{Tr}_N \left(f_\varepsilon(PP^*(X^N, A^N)) \right) \right] \leq N \times C^k \left(\frac{\varepsilon^{-2}}{N} \right)^{2k} k^{12k}.$$

Proof. To estimate the above expectation we once again want to use the Fourier transform with a few refinements to have an optimal estimate with respect to ε . We set $f_\varepsilon^\kappa : t \mapsto g(\varepsilon^{-1}(t - \alpha))g(\varepsilon^{-1}(\kappa - t) + 1)$ with $\kappa > \alpha$. Since g has compact support and is sufficiently

smooth we can apply Theorem 4.3.2. Setting $h : t \mapsto g(t - \varepsilon^{-1}\alpha)g(\varepsilon^{-1}\kappa + 1 - t) = f_\varepsilon^\kappa(\varepsilon t)$, we have for $k \in \mathbb{N}^*$,

$$\begin{aligned} \int y^{4k} |\hat{f}_\varepsilon^\kappa(y)| dy &= \frac{1}{2\pi} \int y^{4k} \left| \int g(\varepsilon^{-1}(t - \alpha))g(\varepsilon^{-1}(\kappa - t) + 1)e^{-iyt} dt \right| dy \\ &= \frac{1}{2\pi} \int y^{4k} \left| \int h(t)e^{-iy\varepsilon t} \varepsilon dt \right| dy \\ &= \frac{\varepsilon^{-4k}}{2\pi} \int y^{4k} \left| \int h(t)e^{-iyt} dt \right| dy \\ &\leq \frac{\varepsilon^{-4k}}{2\pi} \int \frac{1}{1+y^2} dy \int (|h^{(4k)}(t)| + |h^{(4k+2)}(t)|) dt \\ &\leq \varepsilon^{-4k} \left(\|h^{(4k)}\|_\infty + \|h^{(4k+2)}\|_\infty \right). \end{aligned}$$

In the last line we used the fact the support of the derivatives of h are included in $[\varepsilon^{-1}\alpha, \varepsilon^{-1}\alpha + 1] \cup [\varepsilon^{-1}\kappa, \varepsilon^{-1}\kappa + 1]$. Thus thanks to Theorem 4.3.4 and by using the function g defined in the proof of Theorem 4.1.1, we get that there exist constants C and c such that for any $k \leq cN$, for any $\kappa > \alpha$,

$$\mathbb{E} \left[\text{Tr}_N \left(f_\varepsilon^\kappa(PP^*(X^N, A^N)) \right) \right] \leq N \times C^k \left(\frac{\varepsilon^{-2}}{N} \right)^{2k} k^{12k}.$$

Hence the conclusion by dominated convergence theorem. □

Consequently, with $x_+ = \max(x, 0)$, for any $r > 0$,

$$\begin{aligned} &\mathbb{E} \left[\left(\|PP^*(X^N, A^N)\| - \|PP^*(x, A^N)\| \right)_+ \right] \\ &\leq r + \int_r^\infty \mathbb{P} \left(\|PP^*(X^N, A^N)\| \geq \|PP^*(x, A^N)\| + \varepsilon \right) d\varepsilon \\ &\leq r + \int_r^\infty \mathbb{P} \left(\text{Tr}_N \left(f_\varepsilon(PP^*(X^N, A^N)) \right) \geq 1 \right) d\varepsilon \\ &\leq r + \int_r^\infty \mathbb{E} \left[\text{Tr}_N \left(f_\varepsilon(PP^*(X^N, A^N)) \right) \right] d\varepsilon \\ &\leq r + r \times N \times C^k \left(\frac{r^{-2}}{N} \right)^{2k} k^{12k}. \end{aligned}$$

Thus by taking $r = N^{-a}$, we get that

$$\mathbb{E} \left[\left(\|PP^*(X^N, A^N)\| - \|PP^*(x, A^N)\| \right)_+ \right] \leq N^{-a} \times \left(1 + N^{1+2k(2a-1)} \times C^k k^{12k} \right).$$

Now we want to pick a and k such that $N^{1+2k(2a-1)} \times C^k k^{12k}$ is bounded by 1 uniformly over N (while keeping in mind that k has to be an integer). It is sufficient to pick a and k such that,

$$\ln C + \frac{\ln N}{k} + 12 \ln k \leq 2 \times (1 - 2a) \ln N.$$

We fix $k = \lceil \ln N \rceil$, then we need to pick a such that

$$\ln C + 1 + 12 \ln \lceil \ln N \rceil \leq 2 \times (1 - 2a) \ln N.$$

Which means that we can pick $a = \frac{1}{2} - 4 \frac{\ln \ln N}{\ln N}$, and for N large enough,

$$\mathbb{E} \left[\left(\left\| PP^*(X^N, A^N) \right\| - \left\| PP^*(x, A^N) \right\| \right)_+ \right] \leq 2N^{-a} = \frac{2 \ln^4 N}{\sqrt{N}}.$$

Thanks to Proposition 4.6 from [7] we have that for $N \geq \ln(p)$, there exist constants K and D such that

$$\mathbb{P} \left(\left| \left\| P^*P(X^N, A^N) \right\| - \mathbb{E} \left[\left\| P^*P(X^N, A^N) \right\| \right] \right| \geq \delta + Ke^{-N} \right) \leq e^{-N} + e^{-D\delta^2 N}.$$

Thus with $x_+ = \max(x, 0)$, we immediately get that

$$\begin{aligned} & \mathbb{P} \left(\left(\left\| P(X^N, A^N) \right\| - \left\| P(x, A^N) \right\| \right) \geq \frac{\delta + Ke^{-N} + \frac{2 \ln^4 N}{\sqrt{N}}}{\left\| P(x, A^N) \right\|} \right) \\ & \leq \mathbb{P} \left(\left(\left\| P(X^N, A^N) \right\| - \left\| P(x, A^N) \right\| \right) \times \left(\left\| P(X^N, A^N) \right\| + \left\| P(x, A^N) \right\| \right) \geq \delta + Ke^{-N} + \frac{2 \ln^4 N}{\sqrt{N}} \right) \\ & \leq \mathbb{P} \left(\left\| P^*P(X^N, A^N) \right\| - \left\| P^*P(x, A^N) \right\| \geq \delta + Ke^{-N} + \frac{2 \ln^4 N}{\sqrt{N}} \right) \\ & \leq \mathbb{P} \left(\left\| P^*P(X^N, A^N) \right\| - \left\| P^*P(x, A^N) \right\| \geq \delta + Ke^{-N} \right. \\ & \quad \left. + \mathbb{E} \left[\left(\left\| P^*P(X^N, A^N) \right\| - \left\| P^*P(x, A^N) \right\| \right)_+ \right] \right) \\ & \leq \mathbb{P} \left(\left\| P^*P(X^N, A^N) \right\| - \left\| P^*P(x, A^N) \right\| \geq \delta + Ke^{-N} + \mathbb{E} \left[\left\| P^*P(X^N, A^N) \right\| - \left\| P^*P(x, A^N) \right\| \right] \right) \\ & \leq \mathbb{P} \left(\left\| P^*P(X^N, A^N) \right\| - \mathbb{E} \left[\left\| P^*P(X^N, A^N) \right\| \right] \geq \delta + Ke^{-N} \right) \\ & \leq e^{-N} + e^{-D\delta^2 N}. \end{aligned}$$

Since the family $(A^N)_N$ is uniformly bounded over N , so is the sequence $(\|P(x, A^N)\|)_N$, hence by replacing δ by $D^{-1/2}\delta$, we get that there is a constant C such that

$$\mathbb{P} \left(\left\| P(X^N, A^N) \right\| - \left\| P(x, A^N) \right\| \geq C \left(\delta + \frac{\ln^4 N}{\sqrt{N}} \right) \right) \leq e^{-N} + e^{-\delta^2 N}.$$

Finally by replacing δ by $\frac{\ln^4 N}{\sqrt{N}}\delta$, we get that

$$\mathbb{P} \left(\frac{\sqrt{N}}{\ln^4 N} \left(\left\| P(X^N, A^N) \right\| - \left\| P(x, A^N) \right\| \right) \geq C(\delta + 1) \right) \leq e^{-N} + e^{-\delta^2 \ln^8 N}.$$

Acknowledgements

The author would like to thanks his PhD supervisors Benoît Collins and Alice Guionnet for proofreading this paper and their continuous help, as well as Mikael de la Salle for helpful discussion. The author was partially supported by a MEXT JASSO fellowship and Labex Milyon (ANR-10-LABX-0070) of Université de Lyon.

Concentration estimates for random subspaces of a tensor product, and application to Quantum Information Theory

Given a random subspace H_n chosen uniformly in a tensor product of Hilbert spaces $V_n \otimes W$, we consider the collection K_n of all singular values of all norm one elements of H_n with respect to the tensor structure. A law of large numbers has been obtained for this random set in the context of W fixed and the dimension of H_n, V_n tending to infinity at the same speed by Belinschi, Collins and Nechita.

In this paper, we provide measure concentration estimates in this context. The probabilistic study of K_n was motivated by important questions in Quantum Information Theory, and allowed to provide the smallest known dimension for the dimension an an ancilla space allowing Minimum Output Entropy (MOE) violation. With our estimates, we are able, as an application, to provide actual bounds for the dimension of spaces where violation of MOE occurs.

This chapter is adapted from [11], which is a joint work with Benoît Collins.

5.1 Introduction

One of the most important questions in Quantum Information Theory (QIT) was to figure out whether one can find two quantum channels Φ_1 and Φ_2 such that

$$H_{min}(\Phi_1 \otimes \Phi_2) < H_{min}(\Phi_1) + H_{min}(\Phi_2), \quad (5.1)$$

where H_{min} is the Minimum Output Entropy (MOE), defined in section 5.4. This problem was solved by [61], with important preliminary work by [62] (see also references therein). This was especially important in QIT, since there was hope such quantum channels did not exist and consequently that the MOE would be additive, i.e. that the inequality of equation (5.1) would be an equality. In this case, it would give a systematic way to compute the classical capacity of a quantum channel. For more explanation we refer to [90].

All proofs available so far are not constructive in the sense that constructions rely on the probabilistic method. After the initial construction of [61], the probabilistic tools involved in the proof have been found to have deep relation with random matrix theory in many respects, including large deviation principle [63], Free probability [64], convex geometry [65] and Operator Algebra [66]. The last two probably give the most conceptual proofs, and in particular convex geometry gives explicit numbers. Free probability gives the best numbers for the output dimension [64] but was unable to give estimates for the input dimension so far. More generally, the optimal violation obtained in [64] relates to a LLN obtained in [67] whose speed of convergence was not explicit, and in turn, did not give any estimate on the smallest dimension of the input space. In order to obtain explicit parameters, measure concentration estimates, ideally large deviation estimates, are required. And from a theoretical point of view, this is the goal of this paper. Our main results (Theorem 5.2.2 and 5.4.3) give precise estimates for the probability of additivity violation and the dimension of the violating channel in a natural random channel model. The proof is based on the far reaching approach of [8] – see as well [7]. As a corollary, we obtain the following important application in Quantum Information Theory:

Theorem 5.1.1 (For the precise statement, see Theorem 5.4.3). *There exist a quantum channel from $\mathbb{M}_{368 \times 10^{51}}(\mathbb{C})$ to $\mathbb{M}_{184}(\mathbb{C})$ such that combined with its conjugate channel, it yields violation of the MOE.*

The paper is organized as follows. After this introduction, section 2 is devoted to introducing necessary notations and state the main theorem. Section 3 contains the proof of the main theorem, and section 4 contains application to Quantum Information Theory.

Acknowledgements: B.C. was supported by JSPS KAKENHI 17K18734, 17H04823 and 20K20882. F.P. was supported by a JASSO fellowship and Labex Milyon (ANR-10-LABX-0070) of Université de Lyon. This work was initiated while the second author was doing his MSc under the supervision of Alice Guionnet and he would like to thank her for insightful comments and suggestions on this work. The authors would also like to thank Ion Nechita for an interesting remark on the minimum dimension with respect to k .

5.2 Notations and main theorem

We denote by H a Hilbert space, which we assume to be finite dimensional. $B(H)$ is the set of bounded linear operators on H , and $D(H) \subset B(H)$ is the collection of trace 1, positive operators – known as *density matrices*. In the case of matrices, we denote it by $\mathcal{D}_k \subset \mathbb{M}_k(\mathbb{C})$.

Let $d, k, n \in \mathbb{N}$, let U be distributed according to the Haar measure on the unitary group of $\mathbb{M}_{kn}(\mathbb{C})$, let P_n be the canonical injection from \mathbb{C}^d to \mathbb{C}^{kn} , that is the matrix with kn lines and d columns with 1 on the diagonal and 0 elsewhere. With Tr_n the trace on $\mathbb{M}_n(\mathbb{C})$, we define the following random linear map,

$$\Phi_n : X \in \mathbb{M}_d(\mathbb{C}) \mapsto id_k \otimes \text{Tr}_n(UP_n X P_n^* U^*) \in \mathbb{M}_k(\mathbb{C}). \quad (5.2)$$

This map is trace preserving, linear and completely positive and as such, it is known as a quantum channel. Let $t \in [0, 1]$. Let $(d_n)_{n \in \mathbb{N}}$ be an integer sequence such that $d \sim tkn$, and define

$$K_{n,k,t} = \Phi_n(\mathcal{D}_{d_n}). \quad (5.3)$$

There is a much more geometric definition of $K_{n,k,t}$ thanks to the following proposition. Actually, while the quantum channel (5.2) is random, we do not use this fact in the proof, and we could very well prove the same result for a quantum channel defined as in equation (5.2) but with a deterministic unitary matrix instead of a Haar unitary matrix U .

Proposition 5.2.1. *We have,*

$$K_{n,k,t} = \{X \in \mathcal{D}_k \mid \forall A \in \mathcal{D}_k, \text{Tr}_k(XA) \leq \|P_n^* U^* A \otimes I_n U P_n\|\}. \quad (5.4)$$

Besides for any $A \in \mathcal{D}_k$, $\{X \in K_{n,k,t} \mid \text{Tr}_k(XA) = \|P_n^* U^* A \otimes I_n U P_n\|\}$ is non-empty.

Proof. Let $Y \in \mathcal{D}_{d_n}$, $A \in \mathcal{D}_k$, then

$$\begin{aligned} \text{Tr}_k(\Phi_n(Y)A) &= \text{Tr}_{kn}(UP_n Y P_n^* U^* \cdot A \otimes I_n) \\ &= \text{Tr}_d(\sqrt{Y} P_n^* U^* \cdot A \otimes I_n \cdot U P_n \sqrt{Y}) \\ &\leq \text{Tr}_d(Y) \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \\ &= \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\|. \end{aligned}$$

Let us write E for the right member of the equation (5.4), we just showed that $K_{n,k,t} = \Phi_n(\mathcal{D}_{d_n}) \subset E$. Besides if P_x is the orthogonal projection on the vector x , we have that

$$\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| = \max_{x \in \mathbb{C}^d} \text{Tr}_d(P_n^* U^* \cdot A \otimes I_n \cdot U P_n P_x) = \max_{x \in \mathbb{C}^d} \text{Tr}_k(A \Phi_n(P_x)).$$

Thus, for every $\varepsilon > 0$ and $A \in \mathcal{D}_k$, we can find an element of $K_{n,k,t}$ in $\{X \in \mathcal{D}_k \mid \text{Tr}_k(XA) \geq \|P_n^* U^* A \otimes I_n U P_n\| - \varepsilon\}$. By compactness of $K_{n,k,t}$, we can even find an element of $K_{n,k,t}$ in $\{X \in \mathcal{D}_k \mid \text{Tr}_k(XA) = \|P_n^* U^* A \otimes I_n U P_n\|\}$.

If we see E as a convex set of $\mathbb{M}_k(\mathbb{C})_{sa}$, let $X \in E$ be an exposed point of E , that is there exists $A \in \mathbb{M}_k(\mathbb{C})_{sa}$ and C such that the intersection of E and $\{Y \in \mathbb{M}_k(\mathbb{C})_{sa} \mid \text{Tr}_k(AY) = C\}$ is reduced to $\{X\}$ and that E is included in $\{Y \in \mathbb{M}_k(\mathbb{C})_{sa} \mid \text{Tr}_k(AY) \leq C\}$. We have the following equality for λ large enough since if $Y \in \mathcal{D}_k$, $\text{Tr}_k(Y) = 1$,

$$\{Y \in \mathcal{D}_k \mid \text{Tr}_k(AY) = C\} = \left\{ Y \in \mathcal{D}_k \mid \text{Tr}_k \left(\frac{A + \lambda I_k}{\text{Tr}_k(A + \lambda I_k)} Y \right) = \frac{C + \lambda}{\text{Tr}_k(A + \lambda I_k)} \right\}.$$

Thus, we can find $B \in \mathcal{D}_k$ and c such that the intersection of E and $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) = c\}$ is reduced to $\{X\}$ and that E is included in $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) \leq c\}$. To summarize:

- The intersection of $K_{n,k,t}$ and $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) = \|P_n^*U^* \cdot B \otimes I_n \cdot UP_n\|\}$ is non-empty.
- $K_{n,k,t} \subset E$, so the intersection of E and $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) = \|P_n^*U^* \cdot B \otimes I_n \cdot UP_n\|\}$ is non-empty.
- The intersection of E and $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) = c\}$ is exactly $\{X\}$.
- E is included in both $\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) \leq c\}$ and

$$\{Y \in \mathcal{D}_k \mid \text{Tr}_k(BY) \leq \|P_n^*U^* \cdot B \otimes I_n \cdot UP_n\|\}.$$

Hence it implies that $c = \|P_n^*U^*B \otimes I_nUP_n\|$ and that $X \in K_{n,k,t}$. Thus we showed that the exposed point of E belongs to $K_{n,k,t}$. By a result of Straszewicz ([91],theorem 18.6) the set of exposed points is dense in the set of extremal points, so the set of extremal points of E is included in $K_{n,k,t}$. Since $K_{n,k,t}$ is convex, E is included in $K_{n,k,t}$. \square

Thanks to Theorem 1.4 of [9], we know that $\|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\|$ converges almost surely towards a limit $\|A\|_{(t)}$, which we now describe in terms of free probability (for the interested reader we refer to [92], but a non expert reader can take $\lim_n \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\|$ as the definition of $\|A\|_{(t)}$ without loss of generality). For $A \in \mathbb{M}_k(\mathbb{C})$, we set:

$$\|A\|_{(t)} := \|p_t A p_t\|, \quad (5.5)$$

where on the right side we took the operator norm of $p_t A p_t$, with p_t a self-adjoint projection of trace t , free from A . Consequently, we define

$$K_{k,t} = \{X \in \mathcal{D}_k \mid \forall A \in \mathcal{D}_k, \text{Tr}_k(XA) \leq \|A\|_{(t)}\}. \quad (5.6)$$

Given their definition, it seems natural to say that $K_{n,k,t}$ converges towards $K_{k,t}$. However it is not quite as straightforward. The convergence for the Hausdorff distance was proved in [67], Theorem 5.2. More precisely the authors proved that given a random subspace of size d_n , $F_{n,k,t}$ the collection of singular values of unit vectors in this subspace converges for the Hausdorff distance towards a deterministic set $F_{k,t}$. It turns out that $K_{n,k,t}$ (respectively $K_{k,t}$) is the convex hull of the self-adjoint matrices whose eigenvalues are in $F_{n,k,t}$ (respectively $F_{k,t}$). However our paper is self-contained and we do not use this theorem. The main result of this paper is a measure concentration estimate and can be stated as follows.

Theorem 5.2.2. *If we assume $d_n \leq tkn$, then for $n \geq 3^4 \times 2^{30} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$,*

$$\mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}},$$

where $K_{k,t} + \varepsilon = \{Y \in \mathcal{D}_k \mid \exists X \in K_{k,t}, \|X - Y\|_2 \leq \varepsilon\}$ with $\|M\|_2 := \sqrt{\text{Tr}_k(M^*M)}$.

While this does not prove the convergence for the Hausdorff distance of $K_{n,k,t}$ towards $K_{k,t}$ since we do not study the probability that $K_{k,t} \not\subset K_{n,k,t} + \varepsilon$. We could adapt our proof to get this result, but without getting estimates with explicit constants which would be detrimental to our aim of finding explicit parameters for violation of the additivity of the MOE.

5.3 Proof of main theorem

We will combine this geometrical description with the following lemma to get an estimate.

Proposition 5.3.1. *If we define $K_{k,t} + \varepsilon = \{Y \in \mathcal{D}_k \mid \exists X \in K_{k,t}, \|X - Y\|_2 \leq \varepsilon\}$ with $\|M\|_2 := \sqrt{\text{Tr}_k(M^*M)}$, then the following implication is true,*

$$\forall A \in \mathcal{D}_k, \|P_n^* U^* A \otimes I_n U P_n\| \leq \|A\|_{(t)} + \frac{\varepsilon}{k} \implies K_{n,k,t} \subset K_{k,t} + \varepsilon.$$

Before proving it, we need a small lemma on the structure of $K_{k,t}$.

Lemma 5.3.2. *Let $A \in \mathcal{D}_k$, then $\{X \in K_{k,t} \mid \text{Tr}_k(XA) = \|A\|_{(t)}\}$ is non-empty.*

Proof. Thanks to Proposition 5.2.1 we know that for any n , $\{X \in K_{n,k,t} \mid \text{Tr}_k(XA) = \|P_n^* U^* A \otimes I_n U P_n\|\}$ is non-empty. Hence there exists X_n such that:

- $\text{Tr}_k(X_n A) = \|P_n^* U^* A \otimes I_n U P_n\|$,
- $\forall B \in \mathcal{D}_k, \text{Tr}_k(X_n B) \leq \|P_n^* U^* B \otimes I_n U P_n\|$.

By compactness of \mathcal{D}_k , we can assume that X_n converges towards a limit X . But then as we said in the previous section, thanks to Theorem 1.4 from [9], $\|P_n^* U^* B \otimes I_n U P_n\|$ converges towards $\|B\|_{(t)}$. Thus X is such that:

- $\text{Tr}_k(XA) = \|A\|_{(t)}$,
- $\forall B \in \mathcal{D}_k, \text{Tr}_k(XB) \leq \|B\|_{(t)}$.

That is, X belongs to $\{X \in K_{k,t} \mid \text{Tr}_k(XA) = \|A\|_{(t)}\}$. □

We can now prove Proposition 5.3.1.

Proof of Proposition 5.3.1. We assume that $K_{n,k,t} \not\subset K_{k,t} + \varepsilon$, then thanks to the compactness of $K_{n,k,t}$ and $K_{k,t}$, we can find $X \in K_{k,t}$ and $Y \in K_{n,k,t}$ such that $\|X - Y\|_2 > \varepsilon$, and $K_{k,t} \cap B(Y, \|X - Y\|_2)$ is empty. We set $U = \frac{Y - X}{\|Y - X\|_2}$, $A = \frac{1}{k}(U + I_k)$, then $A \in \mathcal{D}_k$. We are going to show that $\|P_n^* U^* A \otimes I_n U P_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k}$. To do so we define

$$P_C = \left\{ B \in K_{k,t} \mid \text{Tr}_k(AB) = \frac{C + 1}{k} \right\} = \{B \in K_{k,t} \mid \text{Tr}_k(UB) = C\}.$$

Let us assume that for $C > \text{Tr}_k(UX)$, P_C is not empty, then let $S \in P_C$. We can write $C = \text{Tr}_k(U(X + tU))$ for some $t > 0$, thus $\text{Tr}_k(US) = \text{Tr}_k(U(X + tU))$, that is $\text{Tr}_k((Y - X)(X - S)) = -t\|Y - X\|_2^2$. Hence the following estimate:

$$\begin{aligned} \|Y - (\alpha X + (1 - \alpha)S)\|_2^2 &= \text{Tr}_k \left((Y - X + (1 - \alpha)(X - S))^2 \right) \\ &= \|Y - X\|_2^2 - 2t(1 - \alpha)\|Y - X\|_2 + \mathcal{O}((1 - \alpha)^2). \end{aligned}$$

Consequently since $K_{k,t}$ is convex, for any α , $\alpha X + (1 - \alpha)S \in K_{k,t}$, thus for $1 - \alpha$ small enough we could find an element of $K_{k,t}$ in $B(Y, \|X - Y\|_2)$. Hence the contradiction. Thus for $C > \text{Tr}_k(UX)$, P_C is empty. By Lemma 5.3.2, we get that $\frac{\text{Tr}_k(UX) + 1}{k} \geq \|A\|_{(t)}$. Next we define

$$Q_C = \left\{ B \in K_{n,k,t} \mid \text{Tr}_k(AB) = \frac{C + 1}{k} \right\} = \{B \in K_{n,k,t} \mid \text{Tr}_k(UB) = C\}.$$

Then clearly for $C = \text{Tr}_k(UY)$, Q_C is non-empty since $Y \in Q_{\text{Tr}_k(UY)}$. Hence thanks to the geometric definition (5.4) of $K_{n,k,t}$, we have that $\frac{\text{Tr}_k(UY)+1}{k} \leq \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\|$. Thus we have,

$$\|P_n^*U^*A \otimes I_nUP_n\| \geq \frac{\text{Tr}_k(U(Y-X))}{k} + \|A\|_{(t)} = \frac{\|Y-X\|_2}{k} + \|a\|_{(t)} > \frac{\varepsilon}{k} + \|A\|_{(t)}.$$

□

Actually with a very similar proof, we could even show that almost surely there exist $A \in \mathcal{D}_k$ such that

$$d_H(K_{n,k,t}, K_{k,t}) = k \times \left| \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| - \|A\|_{(t)} \right|,$$

where d_H is the Hausdorff distance associated to the norm $\|\cdot\|_2$ which comes from the scalar product $(U, V) \mapsto \text{Tr}_k(UV)$. However this result will not be useful in this paper since the absolute value would be detrimental for the computation of our estimate. The following lemma is a rather direct consequence of the previous proposition.

Lemma 5.3.3. *Let $u > 0$, let $\mathcal{S}_u = \{uM \mid M \in \mathbb{M}_k(\mathbb{C})_{sa}, \forall i \geq j, \Re(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil], \forall i > j, \Im(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil]\}$, let $P_{\mathcal{D}_k}$ be the convex projection on \mathcal{D}_k . Then with $u = \frac{\sqrt{2\varepsilon}}{3k^2}$,*

$$\mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq \sum_{M \in \mathcal{S}_u} \mathbb{P}\left(\|P_n^*U^*(P_{\mathcal{D}_k}M \otimes I_n)UP_n\| > \|P_{\mathcal{D}_k}M\|_{(t)} + \frac{\varepsilon}{3k}\right).$$

Proof. We immediately get from proposition 5.3.1 that

$$\mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq \mathbb{P}\left(\exists A \in \mathcal{D}_k, \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k}\right).$$

Now, let $A \in \mathcal{D}_k$, by construction of \mathcal{S}_u , there exists $M \in \mathcal{S}_u$ such that the real part and the imaginary part of the coefficients of M are $u/2$ -close from those of A . Thus we have $\|A - M\|_2 \leq \frac{ku}{\sqrt{2}}$. Hence if we fix $u = \frac{\sqrt{2\varepsilon}}{3k^2}$, then we can always find $M \in \mathcal{S}_u$ such that $\|A - M\|_2 \leq \frac{\varepsilon}{3k}$. Besides we have,

$$\left| \|A\|_{(t)} - \|P_{\mathcal{D}_k}M\|_{(t)} \right| \leq \|A - P_{\mathcal{D}_k}M\| \leq \|A - P_{\mathcal{D}_k}M\|_2,$$

$$\left| \|P_n^*U^*A \otimes I_nUP_n\| - \|P_n^*U^*(P_{\mathcal{D}_k}M \otimes I_n)UP_n\| \right| \leq \|A - P_{\mathcal{D}_k}M\| \leq \|A - P_{\mathcal{D}_k}M\|_2.$$

Hence since $P_{\mathcal{D}_k}A = A$ and that $P_{\mathcal{D}_k}$ is 1-lipschitz, we have $\|A - P_{\mathcal{D}_k}M\|_2 \leq \|A - M\|_2 \leq \frac{\varepsilon}{3k}$. Consequently,

$$\left\{ \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k} \right\} \subset \left\{ \|P_n^*U^*(P_{\mathcal{D}_k}M \otimes I_n)UP_n\| > \|P_{\mathcal{D}_k}M\|_{(t)} + \frac{\varepsilon}{3k} \right\}.$$

Hence,

$$\begin{aligned} & \left\{ \exists A \in \mathcal{D}_k, \|P_n^*U^* \cdot A \otimes I_n \cdot UP_n\| > \|A\|_{(t)} + \frac{\varepsilon}{k} \right\} \\ & \subset \bigcup_{M \in \mathcal{S}_u} \left\{ \|P_n^*U^*(P_{\mathcal{D}_k}M \otimes I_n)UP_n\| > \|P_{\mathcal{D}_k}M\|_{(t)} + \frac{\varepsilon}{3k} \right\}. \end{aligned}$$

The conclusion follows.

□

The next lemma shows that there exist a smooth function which verifies some assumptions on the infinite norm of its derivatives.

Lemma 5.3.4. *There exists g a \mathcal{C}^6 function which takes value 0 on $(-\infty, 0]$ and value 1 on $[1, \infty)$, and in $[0, 1]$ otherwise. Besides for any $j \leq 6$, $\|g^{(j)}\|_\infty = 2^{\frac{j(j+1)}{2}}$.*

Proof. Firstly we define,

$$f : t \in [0, 1] \mapsto \begin{cases} 2t & \text{if } t \leq 1/2 \\ 2(1-t) & \text{if } t \geq 1/2 \end{cases},$$

$$H : \mathbb{C}^0([0, 1]) \rightarrow \mathbb{C}^0([0, 1])$$

$$f \mapsto t \mapsto \begin{cases} f(2t) & \text{if } t \leq 1/2 \\ -f(2t-1) & \text{if } t \geq 1/2 \end{cases}.$$

Inspired by Taylor's Theorem, we define

$$h : x \in [0, 1] \mapsto \int_0^x \frac{(x-t)^5}{5!} H^5 f(t) dt.$$

It is easy to see that $h \in \mathcal{C}^6([0, 1])$ with

$$\forall j \leq 5, \quad h^{(j)} : x \in [0, 1] \mapsto \int_0^x \frac{(x-t)^{5-j}}{(5-j)!} H^5 f(t) dt, \quad h^{(6)} = H^5 f.$$

Thus one can easily extend h by 0 on \mathbb{R}^- and h remains \mathcal{C}^6 in 0, as for what happens in 1 it is way less obvious. In order to build g we want to show that

$$\forall 1 \leq j \leq 6, \quad h^{(j)}(1) = 0, \quad h(1) > 0.$$

To do so let $w \in \mathcal{C}^0([0, 1])$, then for any $k \geq 0$,

$$\begin{aligned} \int_0^1 (1-t)^k Hw(t) dt &= \int_0^{1/2} (1-t)^k w(2t) dt - \int_{1/2}^1 (1-t)^k w(2t-1) dt \\ &= \frac{1}{2^{k+1}} \int_0^1 ((2-t)^k - (1-t)^k) w(t) dt \\ &= \frac{1}{2^{k+1}} \int_0^1 \sum_{0 \leq i < k} \binom{k}{i} (1-t)^i w(t) dt. \end{aligned}$$

Thus recursively one can show that $\forall 1 \leq j \leq 6, h^{(j)}(1) = 0$. We also get that

$$h(1) = \int_0^x \frac{(1-t)^5}{5!} H^5 f(t) dt = 2^{-\sum_{2 \leq i \leq 6} i} \int_0^x f(t) dt = 2^{-21}.$$

Hence we fix $g = 2^{21}h$, further studies show that $\|g^{(j)}\|_\infty = 2^{\frac{j(j+1)}{2}}$. □

In the next lemma, we prove a first rough estimate on the deviation of the norm with respect to its limit. It is the only one where we use that $d_n \leq tkn$.

Lemma 5.3.5. *For any $A \in \mathcal{D}_k$, $\varepsilon > 0$,*

$$\mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \leq 3 \times 2^{22} \times \frac{\ln^2(kn)}{kn} \varepsilon^{-4}. \quad (5.7)$$

Proof. For a better understanding of the notations and tools used in this proof, such as free stochastic calculus, we refer to [8]. In particular τ_{kn} is the trace on the free product of $\mathbb{M}_{kn}(\mathbb{C})$ with a C^* -algebra which contains a free unitary Brownian motion, see Definition 2.8 of [8]. As for δ, \mathcal{D} and \boxtimes , see Definition 2.5 of [8] and 2.10 from [10]. If you are not familiar with free probability, it is possible to simply admit equation (5.9) to avoid having to understand the previous notations.

Since $\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| = \|P_n P_n^* U^* \cdot A \otimes I_n \cdot U P_n P_n^*\|$, we will rather work with $P = P_n P_n^*$ since it is a square matrix. To simplify notations, instead of $A \otimes I_n$ we simply write A . Let us now consider a function f such that

$$\forall x \in \mathbb{R}, \quad f(x) = \int_{\mathbb{R}} e^{ixy} d\mu(y), \quad (5.8)$$

for some measure μ . We set $v_s = U \otimes I_l \times V_{t-s}^{lkn} U_s^{lkn}$ and $w_t = U \otimes I_l \times W_{t-s}^{lkn} U_s^{lkn}$ where $(U_t^{mkn})_{t \geq 0}, (V_t^{mkn})_{t \geq 0}, (W_t^{mkn})_{t \geq 0}$ are independent unitary Brownian motions of size mkn started in the identity. We also set $P_{l,l'} = I_{kn} \otimes E_{l,l'}$ where $E_{l,l'}$ is the matrix of size m whom all coefficients are zero but the (l, l') one which is 1. Then thanks to Lemma 4.2, 4.6 and Corollary 3.3 from [8], with u_T a free unitary Brownian motion at time T started in 1, we have the following expression,

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{kn} \text{Tr}_{kn} \left(f(P U^* A U P) \right) \right] - \mathbb{E} \left[\tau_{kn} \left(f(P u_T^* U^* A U u_T P) \right) \right] \\ &= \lim_{m \rightarrow \infty} \frac{1}{2m^2(kn)^3} \sum_{1 \leq l, l' \leq m} \int_0^T \int_0^t \text{Tr}_{mkn} \left(\delta \circ \delta^1 \circ \mathcal{D} \left(e^{iy P v_s^* A v_s P} \right) \#_{P_{l,l'}} \right. \\ & \quad \left. \boxtimes \delta \circ \delta^2 \circ \mathcal{D} \left(e^{iy P w_s^* A w_s P} \right) \#_{P_{l,l'}} \right) ds dt d\mu(y). \end{aligned}$$

Then if we set $R_1^s = P v_s^* A v_s P$ and $R_2^s = P w_s^* A w_s P$, after a lengthy computation,

$$\begin{aligned} & \text{Tr}_{mkn} \left(\delta \circ \delta^1 \circ \mathcal{D} \left(e^{iy P v_s^* A v_s P} \right) \#_{P_{l,l'}} \boxtimes \delta \circ \delta^2 \circ \mathcal{D} \left(e^{iy P w_s^* A w_s P} \right) \#_{P_{l,l'}} \right) \\ &= -iy \text{Tr}_{lkn} \left(\delta(v_s^* A v_s) \#_{P_{l,l'}} \times \delta(P e^{iy R_2^s} P) \#_{P_{l,l'}} \right) \\ & \quad - iy \text{Tr}_{lkn} \left(\delta(P e^{iy R_1^s} P) \#_{P_{l,l'}} \times \delta(w_s^* A w_s) \#_{P_{l,l'}} \right) \\ & \quad + y^2 \int_0^1 \text{Tr}_{lkn} \left(\delta(v_s^* A v_s P e^{iy \alpha R_1^s} P v_s^* A v_s) \#_{P_{l,l'}} \times \delta(P e^{iy(1-\alpha) R_2^s} P) \#_{P_{l,l'}} \right) d\alpha \\ & \quad - y^2 \int_0^1 \text{Tr}_{lkn} \left(\delta(v_s^* A v_s P e^{iy \alpha R_1^s} P) \#_{P_{l,l'}} \times \delta(w_s^* A w_s P e^{iy(1-\alpha) R_2^s} P) \#_{P_{l,l'}} \right) d\alpha \\ & \quad + y^2 \int_0^1 \text{Tr}_{lkn} \left(\delta(P e^{iy \alpha R_1^s} P) \#_{P_{l,l'}} \times \delta(w_s^* A w_s P e^{iy(1-\alpha) R_2^s} P w_s^* A w_s) \#_{P_{l,l'}} \right) d\alpha \\ & \quad - y^2 \int_0^1 \text{Tr}_{lkn} \left(\delta(P e^{iy \alpha R_1^s} P v_s^* A v_s) \#_{P_{l,l'}} \times \delta(P e^{iy(1-\alpha) R_2^s} P w_s^* A w_s) \#_{P_{l,l'}} \right) d\alpha \end{aligned}$$

Since the norm of A, P, v_s and w_s are smaller than 1, and that the rank of $P_{l,\nu}$ is kn , by using the fact that the non-renormalized trace of a matrix of norm smaller than 1 is smaller or equal to its rank, we finally get that for any t ,

$$\begin{aligned} & \frac{1}{kn} \left| \text{Tr}_{mkn} \left(\delta \circ \delta^1 \circ \mathcal{D} \left(e^{iyPv_s^*Av_sP} \right) \# P_{l,\nu} \boxtimes \delta \circ \delta^2 \circ \mathcal{D} \left(e^{iyPw_s^*Aw_sP} \right) \# P_{l,\nu} \right) \right| \\ & \leq 8y^2 + 2y^2 \int_0^1 (4 + 2\alpha|y|) \times 2(1 - \alpha)|y| \, d\alpha + 2y^2 \int_0^1 (2 + 2\alpha|y|) \times (2 + 2(1 - \alpha)|y|) \, d\alpha \\ & = 16y^2 + 16|y|^3 + \frac{8}{3}y^4. \end{aligned}$$

Consequently, we get that

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{kn} \text{Tr}_{kn} \left(f(P U^* A U P) \right) \right] - \mathbb{E} \left[\tau_{kn} \left(f(P u_T^* U^* A U u_T P) \right) \right] \right| \\ & \leq \frac{T^2}{(kn)^2} \int 4y^2 + 4|y|^3 + \frac{2}{3}y^4 \, d|\mu|(y). \end{aligned}$$

Thanks to Proposition 3.3 from [8], we get that

$$\begin{aligned} & \left| \tau_{kn} \left(e^{iyPu_T^*U^*AUu_TP} \right) - \tau \left(e^{iyPu^*AuP} \right) \right| \\ & = \left| \tau_{kn} \left(e^{iyPu_T^*U^*AUu_TP} \right) - \tau_{kn} \left(e^{iyPu^*U^*AUuP} \right) \right| \\ & = \left| y \int_0^1 \tau \left(e^{isyPu_T^*Au_TP} P (u_T^*U^*AUu_T - \tilde{u}_T^*U^*AU\tilde{u}_T) P e^{i(1-s)yP} \tilde{u}_T^*A\tilde{u}_TP \right) ds \right| \\ & \leq 8e^2\pi e^{-T/2}|y|. \end{aligned}$$

Consequently, if the support of f and the spectrum of Pu^*AuP are disjoint, then $\tau \left(f(P u^* A \otimes I_n u P) \right) = 0$, and

$$\begin{aligned} & \left| \mathbb{E} \left[\frac{1}{kn} \text{Tr}_{kn} \left(f(P U^* A \otimes I_n U P) \right) \right] \right| \\ & \leq 8e^2\pi e^{-T/2} \int |y| \, d|\mu|(y) + \left(\frac{T}{kn} \right)^2 \int 4y^2 + 4|y|^3 + \frac{2}{3}y^4 \, d|\mu|(y). \end{aligned} \tag{5.9}$$

Let g be a \mathcal{C}^6 function which takes value 0 on $(-\infty, 0]$ and value 1 on $[1, \infty)$, and in $[0, 1]$ otherwise. We set $f_\varepsilon : t \mapsto g(2\varepsilon^{-1}(t - \alpha) - 1)g(2\varepsilon^{-1}(1 - t) + 1)$ with $\alpha = \|A\|_{(t)}$. Then the support of f is included in $(\|A\|_{(t)}, \infty)$, whereas the spectrum of Pu^*AuP is bounded by $\|Pu^*AuP\| = \|A\|_{d_n(kn)^{-1}} \leq \|A\|_{(t)}$ since $d_n \leq tkn$. Hence f_ε satisfies (5.9). Setting $h : t \mapsto g(t - 2\varepsilon^{-1}\alpha - 1)g(2\varepsilon^{-1} + 1 - t)$, we have with convention $\hat{f}(x) = (2\pi)^{-1} \int_{\mathbb{R}} f(y) e^{-ixy} dy$,

for $0 \leq k \leq 4$ and any $\beta > 0$,

$$\begin{aligned}
\int |y|^k |\hat{f}_\varepsilon(y)| dy &= \frac{1}{2\pi} \int |y|^k \left| \int g(2\varepsilon^{-1}(t-\alpha) - 1) g(2\varepsilon^{-1}(1-t) + 1) e^{-iyt} dt \right| dy \\
&= \frac{1}{2\pi} \int |y|^k \left| \int h(\beta t) e^{-iy\varepsilon\beta t/2} \frac{\varepsilon\beta}{2} dt \right| dy \\
&= \frac{1}{2\pi} 2^k \varepsilon^{-k} \beta^{-k} \int |y|^k \left| \int h(\beta t) e^{-iyt} dt \right| dy \\
&\leq \frac{1}{2\pi} 2^k \varepsilon^{-k} \int \frac{1}{1+y^2} dy \int (|h^{(k)}(\beta t)| + \beta^2 |h^{(k+2)}(\beta t)|) dt \\
&\leq 2^{k-1} \varepsilon^{-k} \left(\beta^{-1} \|g^{(k)}\|_\infty + \beta \|g^{(k+2)}\|_\infty \right).
\end{aligned}$$

In the last line we used the fact that we can always assume that $\alpha + \varepsilon \leq 1$ (otherwise $\mathbb{P}(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \geq \|A\|_{(t)} + \varepsilon) = 0$ and there is no need to do any computation) and thus that the support of $t \mapsto g(t - 2\varepsilon^{-1}\alpha - 1)$ and the derivative of $t \mapsto g(2\varepsilon^{-1} + 1 - t)$ are disjoint. Thus by fixing $\beta = \sqrt{\|g^{(k)}\|_\infty \|g^{(k+2)}\|_\infty^{-1}}$ we get

$$\int |y|^k |\hat{f}_\varepsilon(y)| dy \leq 2^k \varepsilon^{-k} \sqrt{\|g^{(k)}\|_\infty \|g^{(k+2)}\|_\infty}.$$

Consequently, since f_ε satisfies (5.8) with $d\mu(y) = \hat{f}_\varepsilon(y) dy$, by using (5.9) we get

$$\begin{aligned}
&\left| \mathbb{E} \left[\frac{1}{kn} \text{Tr}_{kn} \left(f_\varepsilon(P U^* A \otimes I_n U P) \right) \right] \right| \\
&\leq 16e^2 \pi e^{-T/2} \sqrt{\|g^{(1)}\|_\infty \|g^{(3)}\|_\infty} \varepsilon^{-1} + \left(\frac{T}{kn} \right)^2 16\varepsilon^{-2} \sqrt{\|g^{(2)}\|_\infty \|g^{(4)}\|_\infty} \\
&\quad + \left(\frac{T}{kn} \right)^2 32\varepsilon^{-3} \sqrt{\|g^{(3)}\|_\infty \|g^{(5)}\|_\infty} + \left(\frac{T}{kn} \right)^2 \frac{32}{3} \varepsilon^{-4} \sqrt{\|g^{(4)}\|_\infty \|g^{(6)}\|_\infty}.
\end{aligned}$$

Combined with Lemma 5.3.4 and fixing $T = 4 \ln(kn)$, we get

$$\begin{aligned}
&\left| \mathbb{E} \left[\frac{1}{kn} \text{Tr}_{kn} \left(f_\varepsilon(P U^* A \otimes I_n U P) \right) \right] \right| \\
&\leq 2^{\frac{15}{2}} e^2 \pi \frac{\varepsilon^{-1}}{(kn)^2} + 2^{29/2} \left(\frac{\ln(kn)}{kn} \right)^2 \varepsilon^{-2} + 2^{39/2} \left(\frac{\ln(kn)}{kn} \right)^2 \varepsilon^{-3} + \frac{2^{49/2}}{3} \left(\frac{\ln(kn)}{kn} \right)^2 \varepsilon^{-4}.
\end{aligned}$$

Since for any n , almost surely $\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \leq 1$, we have

$$\begin{aligned}
&\mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \\
&= \mathbb{P} \left(\exists \lambda \in \sigma(P U^* A \otimes I_n U P), f_\varepsilon(\lambda) = 1 \right) \\
&\leq \mathbb{P} \left(\text{Tr}_{kn} \left(f_\varepsilon(P^* U^* A \otimes I_n U P) \right) \geq 1 \right) \\
&\leq \mathbb{E} \left[\text{Tr}_{kn} \left(f_\varepsilon(P U^* A \otimes I_n U P) \right) \right] \\
&\leq 2^{\frac{15}{2}} e^2 \pi \frac{\varepsilon^{-1}}{kn} + \frac{\ln^2(kn)}{kn} \left(2^{29/2} \varepsilon^{-2} + 2^{39/2} \varepsilon^{-3} + \frac{2^{49/2}}{3} \varepsilon^{-4} \right).
\end{aligned}$$

One can always assume that $\ln^2(kn) \geq 1$ since for small value of k and n , (5.7) is easily verified since the right member of the inequality is larger than 1. One can also assume that $\varepsilon < 1$ since almost surely $\|P_n^* U^* A \otimes I_n U P_n\| \leq 1$. We get the conclusion by a numerical computation. \square

We can now refine this inequality by relying on corollary 4.4.28 of [13], we state the part that we will be using in the next proposition.

Proposition 5.3.6. *We set $S\mathbb{U}_N = \{X \in \mathbb{U}_N \mid \det(X) = 1\}$, let f be a continuous, real-valued function on \mathbb{U}_N . We assume that there exists a constant C such that for every $X, Y \in \mathbb{U}_N$,*

$$|f(X) - f(Y)| \leq C \|X - Y\|_2 \quad (5.10)$$

Then if we set ν_G the law of the Haar measure on G , for all $\delta > 0$:

$$\nu_{\mathbb{U}_N} \left(\left| f(\cdot) - \int f(Y \cdot) d\nu_{S\mathbb{U}_N}(Y) \right| \geq \delta \right) \leq 2e^{-\frac{N\delta^2}{4C^2}} \quad (5.11)$$

Lemma 5.3.7. *For any $A \in \mathcal{D}_k$, $\varepsilon > 0$, if $kn \geq 2^{30} \times \ln^2(kn) \times \varepsilon^{-4}$, we have*

$$\mathbb{P} \left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \leq 2e^{-kn \times \frac{\varepsilon^2}{64}}.$$

Proof. We set,

$$f : U \mapsto \|P_n^* U^* A \otimes I_n U P_n\|,$$

$$h : X \in \mathbb{U}_n \mapsto \int f(YX) d\nu_{S\mathbb{U}_{kn}}(Y).$$

If U^1 is a random matrix of law $\nu_{S\mathbb{U}_{kn}}$, and α a scalar of law $\nu_{\mathbb{U}_1}$ independent of U^1 . Then the law of αU^1 is $\nu_{\mathbb{U}_{kn}}$ since its law is invariant by multiplication by a unitary matrix. Consequently for any $X \in \mathbb{U}_{kn}$,

$$h(X) = \mathbb{E}[f(U^1 X)] = \mathbb{E}[f(\alpha U^1 X)] = \mathbb{E}[f(\alpha U^1)] = \int f(Y) d\nu_{\mathbb{U}_{kn}}(Y).$$

The third inequality is true since for any scalar α and $X \in \mathbb{U}_{kn}$, $f(X) = f(\alpha X)$. Besides we also have that for any $U, V \in \mathbb{U}_{kn}$,

$$|f(U) - f(V)| \leq 2 \|U - V\| \leq 2 \|U - V\|_2.$$

Thus by using Proposition 5.3.6, we get

$$\mathbb{P} \left(\left| \|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \mathbb{E} \left[\|P_n^* U^* A \otimes I_n U P_n\| \right] \right| \geq \delta \right) \leq 2e^{-\frac{kn\delta^2}{16}}.$$

Besides if for $x \in \mathbb{R}$, we denote $x_+ = \max(0, x)$, then

$$\begin{aligned}
& \mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \\
& \leq \mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| - \mathbb{E} [\|P_n^* U^* A \otimes I_n U P_n\|] \right. \\
& \quad \left. \geq \varepsilon - \mathbb{E} \left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)} \right)_+ \right] \right) \\
& \leq \mathbb{P} \left(\left| \|P_n^* U^* A \otimes I_n U P_n\| - \mathbb{E} [\|P_n^* U^* A \otimes I_n U P_n\|] \right| \right. \\
& \quad \left. \geq \varepsilon - \mathbb{E} \left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)} \right)_+ \right] \right) \\
& \leq 2e^{-kn \left(\varepsilon - \mathbb{E} \left[\left(\|P_n^* U^* A \otimes I_n U P_n\| - \|A\|_{(t)} \right)_+ \right] \right)^2 / 16}.
\end{aligned}$$

Besides thanks to our first estimate, i.e. Lemma 5.3.5, we get that for any $r > 0$,

$$\begin{aligned}
\mathbb{E} \left[\left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \|A\|_{(t)} \right)_+ \right] & \leq r + \int_r^1 \mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \alpha \right) d\alpha \\
& \leq r + 2^{22} \times \frac{\ln^2(kn)}{kn} \int_r^1 3 \times \alpha^{-4} d\alpha \\
& \leq r + 2^{22} \times \frac{\ln^2(kn)}{kn} r^{-3}
\end{aligned}$$

And after fixing $r = \left(2^{22} \times \frac{\ln^2(kn)}{kn} \right)^{1/4}$, we get that

$$\mathbb{E} \left[\left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| - \|A\|_{(t)} \right)_+ \right] \leq \left(2^{26} \times \frac{\ln^2(kn)}{kn} \right)^{1/4}.$$

Hence if $kn \geq 2^{30} \times \ln^2(kn) \times \varepsilon^{-4}$, we have

$$\mathbb{P} \left(\|P_n^* U^* \cdot A \otimes I_n \cdot U P_n\| \geq \|A\|_{(t)} + \varepsilon \right) \leq 2e^{-kn \times \frac{\varepsilon^2}{64}}.$$

□

We can finally prove Theorem 5.2.2 by using the former lemma in combination with Lemma 5.3.3.

Proof of Theorem 5.2.2. If we set $u = \frac{\sqrt{2}\varepsilon}{3k^2}$, then with $\mathcal{S}_u = \{uM \mid M \in \mathbb{M}_k(\mathbb{C})_{sa}, \forall i \geq j, \Re(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil], \Im(m_{i,j}) \in \{\mathbb{N} + \frac{1}{2}\} \cap [0, \lceil u^{-1} \rceil]\}$, Lemma 5.3.3 tells us that

$$\mathbb{P} (K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq \sum_{M \in \mathcal{S}_u} \mathbb{P} \left(\|P_n^* U^* (P_{\mathcal{D}_k} M \otimes I_n) U P_n\| > \|P_{\mathcal{D}_k} M\|_{(t)} + \frac{\varepsilon}{3k} \right).$$

But thanks to Lemma 5.3.7, we know that for any $A \in \mathcal{D}_k$, if $n \geq 3^4 \times 2^{30} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$, then

$$\mathbb{P} \left(\|P_n^* U^* A \otimes I_n U P_n\| \geq \|A\|_{(t)} + \frac{\varepsilon}{3k} \right) \leq 2e^{-\frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

Thus since the cardinal of \mathcal{S}_u can be bounded by $(u^{-1} + 1)^{k^2}$, we get that for $n \geq 3^4 \times 2^{30} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$,

$$\mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon) \leq 2(u^{-1} + 1)^{k^2} e^{-\frac{n}{k} \times \frac{\varepsilon^2}{576}} \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

□

5.4 Application to Quantum Information Theory

5.4.1 Preliminaries on entropy

For $X \in D(H)$, its *von Neumann entropy* is defined by functional calculus by $H(X) = -\text{Tr}(X \ln X)$, where $0 \ln 0$ is assumed by continuity to be zero. In other words, $H(X) = \sum_{\lambda \in \text{spec}(X)} -\lambda \ln \lambda$ where the sum is counted with multiplicity. A *quantum channel* $\Phi : B(H_1) \rightarrow B(H_2)$ is a completely positive trace preserving linear map. The *Minimum Output Entropy* (MOE) of Φ is

$$H_{\min}(\Phi) = \min_{X \in D(H_1)} H(\Phi(X)). \quad (5.12)$$

During the last decade, a crucial problem in Quantum Information Theory was to determine whether one can find two quantum channels

$$\Phi_i : B(H_{j_i}) \rightarrow B(H_{k_i}), i = \{1, 2\},$$

such that

$$H_{\min}(\Phi_1 \otimes \Phi_2) < H_{\min}(\Phi_1) + H_{\min}(\Phi_2).$$

Let $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^k$ and let

$$x_t^* = \left(\|e_1\|_t, \underbrace{\frac{1 - \|e_1\|_t}{k-1}, \dots, \frac{1 - \|e_1\|_t}{k-1}}_{k-1 \text{ times}} \right). \quad (5.13)$$

If we view x_t^* as a diagonal matrix, then it can be easily checked that $x_t^* \in K_{k,t}$, and the following is the main result of [64]:

Theorem 5.4.1. *For any $p > 1$, the maximum of the l^p norm on $K_{k,t}$ is reached at the point x_t^* .*

By letting $p \rightarrow 1$ it implies that the minimum of the entropy on $K_{k,t}$ is reached at the point x_t^* and this is what we will be using. For the sake of making actual computation, it will be useful to recall the value of $\|e_1\|_t$. For this, we use the following notation:

$$(1^j 0^{k-j}) = (\underbrace{1, 1, \dots, 1}_{j \text{ times}}, \underbrace{0, 0, \dots, 0}_{k-j \text{ times}}) \in \mathbb{R}^k, \quad (5.14)$$

and $1^k = (1^k 0^0)$. It was proved in the early days of free probability theory (see [92]) that for $j = 1, 2, \dots, k$, one has

$$\|(1^j 0^{k-j})\|_{(t)} = \phi(u, t) = \begin{cases} t + u - 2tu + 2\sqrt{tu(1-t)(1-u)} & \text{if } t + u < 1, \\ 1 & \text{if } t + u \geq 1, \end{cases}$$

where $u = j/k$.

5.4.2 Corollary of the main result

The following is a direct consequence of the main theorem in terms of possible entropies of the output set.

Theorem 5.4.2. *With $S_{n,k,t} = \min_{A \in K_{n,k,t}} H(A) = H_{\min}(\Phi_n)$ and $S_{k,t} = \min_{A \in K_{k,t}} H(A)$, if we assume $d_n \leq tkn$, then for $n \geq 3^4 \times 2^{30} \times \ln^2(kn) \times k^3 \varepsilon^{-4}$ where $0 < \varepsilon \leq e^{-1}$,*

$$\mathbb{P}\left(S_{n,k,t} \leq S_{k,t} - 3k\varepsilon |\ln(\varepsilon)|\right) \leq e^{k^2(\ln(3k^2\varepsilon^{-1})) - \frac{n}{k} \times \frac{\varepsilon^2}{576}}.$$

Proof. Let $A, B \in \mathcal{D}_k$ such that $\|A - B\|_2 \leq \varepsilon$ with $\|M\|_2 = \sqrt{\text{Tr}_k(M^*M)}$, with eigenvalues $(\lambda_i)_i$ and $(\mu_i)_i$. Then with $\tilde{x} = \max\{\varepsilon, x\}$,

$$\begin{aligned} \left| |\text{Tr}_k(A \ln(A))| - |\text{Tr}_k(B \ln(B))| \right| &= \left| \sum_i \lambda_i \ln(\lambda_i) - \sum_i \mu_i \ln(\mu_i) \right| \\ &\leq 2k \sup_{x \in [0, \varepsilon]} |x \ln(x)| + \left| \sum_i \tilde{\lambda}_i \ln(\tilde{\lambda}_i) - \sum_i \tilde{\mu}_i \ln(\tilde{\mu}_i) \right| \\ &\leq 2k\varepsilon |\ln(\varepsilon)| + \sum_i |\lambda_i - \mu_i| |\ln(\varepsilon)| \\ &\leq 2k\varepsilon |\ln(\varepsilon)| + k \|A - B\| |\ln(\varepsilon)| \\ &\leq 3k\varepsilon |\ln(\varepsilon)| \end{aligned}$$

Thus if we endow $\mathbb{M}_k(\mathbb{C})$ with the norm $M \mapsto \sqrt{\text{Tr}_k(M^*M)}$, then if $K_{n,k,t} \subset K_{k,t} + \varepsilon$, then $S_{n,k,t} \geq \min_{A \in K_{k,t} + \varepsilon} |\text{Tr}_k(A \ln(A))| \geq S_{k,t} - 3k\varepsilon |\ln(\varepsilon)|$. Hence

$$\mathbb{P}\left(S_{n,k,t} \leq S_{k,t} - 3k\varepsilon |\ln(\varepsilon)|\right) \leq \mathbb{P}(K_{n,k,t} \not\subset K_{k,t} + \varepsilon).$$

Theorem 5.2.2 then allows us to conclude. □

5.4.3 Application to violation of the Minimum Output Entropy of Quantum Channels

In order to obtain violations for the additivity relation of the minimum output entropy, one needs to obtain upper bounds for the quantity $H_{\min}(\Phi \otimes \Psi)$ for some quantum channels Φ and Ψ . The idea of using conjugate channels ($\Psi = \bar{\Phi}$) and bounding the minimum output entropy by the value of the entropy at the Bell state dates back to Werner, Winter and others (we refer to [62] for references). To date, it has been proven to be the most successful method of tackling the additivity problem. The following inequality is elementary and lies at the heart of the method:

$$H_{\min}(\Phi \otimes \bar{\Phi}) \leq H([\Phi \otimes \bar{\Phi}](E_d)), \quad (5.15)$$

where E_d is the maximally entangled state over the input space $(\mathbb{C}^d)^{\otimes 2}$. More precisely, E_d is the projection on the Bell vector

$$\text{Bell}_d = \frac{1}{\sqrt{d}} \sum_{i=1}^d e_i \otimes e_i, \quad (5.16)$$

where $\{e_i\}_{i=1}^d$ is a fixed basis of \mathbb{C}^d .

For random quantum channels $\Phi = \Phi_n$, the random output matrix $[\Phi_n \otimes \bar{\Phi}_n](E_{d_n})$ was thoroughly studied in [93] in the regime $d_n \sim tkn$; we recall here one of the main results of that paper. There, it was proved that almost surely, as n tends to infinity, the random matrix $[\Phi_n \otimes \bar{\Phi}_n](E_{d_n}) \in M_{k^2}(\mathbb{C})$ has eigenvalues

$$\gamma_t^* = \left(t + \frac{1-t}{k^2}, \underbrace{\frac{1-t}{k^2}, \dots, \frac{1-t}{k^2}}_{k^2-1 \text{ times}} \right). \quad (5.17)$$

This result improves on a bound [62] via linear algebra techniques, which states that the largest eigenvalue of the random matrix $[\Phi_n \otimes \bar{\Phi}_n](E_{d_n})$ is at least $d_n/(kn) \sim t$. Although it might be possible to work directly with the bound provided by (5.17) with additional probabilistic consideration, for the sake of concreteness we will work with the bound of [62]. Thus if the largest eigenvalue of $[\Phi_n \otimes \bar{\Phi}_n](E_{d_n})$ is $d_n/(kn)$, since $\text{Tr}_k \otimes \text{Tr}_k([\Phi_n \otimes \bar{\Phi}_n](E_{d_n})) = 1$, the entropy is maximized if we take the remaining $k^2 - 1$ eigenvalues equal to $\frac{1-d_n/(kn)}{k^2-1}$, thus it follows that

$$H_{\min}(\Phi \otimes \bar{\Phi}) \leq H([\Phi \otimes \bar{\Phi}](E_{d_n})) \leq H \left(\frac{d_n}{kn}, \underbrace{\frac{1-\frac{d_n}{kn}}{k^2-1}, \dots, \frac{1-\frac{d_n}{kn}}{k^2-1}}_{k^2-1 \text{ times}} \right)$$

Therefore, it is enough to find n, k, d_n, t such that

$$-\frac{d_n}{kn} \log \left(\frac{d_n}{kn} \right) - \left(1 - \frac{d_n}{kn} \right) \log \left[\left(1 - \frac{d_n}{kn} \right) / (k^2 - 1) \right] < 2H(x_t^*). \quad (5.18)$$

In [64] it was proved with the assistance of a computer that this can be done for any $k \geq 184$, as long as we take t around $1/10$, see figure 1 from [64]. However for k large enough, the difference between the right and left term of (5.18) is maximal for $t = 1/2$. As soon as we take ε such that $3k\varepsilon |\ln(\varepsilon)|$ is less than the difference, we are done. For example, we obtain the following theorem

Theorem 5.4.3. *For the following values $(k, t, n) = (184, 1/10, 2 \times 10^{52}), (185, 1/10, 4 \times 10^{51}), (200, 1/10, 2 \times 10^{47}), (500, 1/10, 8 \times 10^{45}), (500, 1/2, 2 \times 10^{45})$ violation of additivity is achieved with probability at least $1 - \exp(-10^{20})$.*

Proof. We make sure to work with n a multiple of 10 so that we can set $d_n = tkn$, then since $H_{\min}(\Phi_n) = H_{\min}(\bar{\Phi}_n)$,

$$\begin{aligned} & \mathbb{P} \left(H_{\min}(\Phi_n \otimes \bar{\Phi}_n) < H_{\min}(\Phi_n) + H_{\min}(\bar{\Phi}_n) \right) \\ &= \mathbb{P} \left(H_{\min}(\Phi_n \otimes \bar{\Phi}_n) < 2H_{\min}(\Phi_n) \right) \\ &\leq \mathbb{P} \left(-t \log(t) - (1-t) \log \left[(1-t)/(k^2-1) \right] < 2H_{\min}(\Phi_n) \right) \\ &= 1 - \mathbb{P} \left(H_{\min}(\Phi_n) \leq -\frac{t}{2} \log(t) - \frac{1-t}{2} \log \left(\frac{1-t}{k^2-1} \right) \right) \\ &= 1 - \mathbb{P} (S_{n,k,t} \leq S_{k,t} - \delta_{k,t}), \end{aligned}$$

with

$$\delta_{k,t} = \frac{t}{2} \log(t) + \frac{1-t}{2} \log\left(\frac{1-t}{k^2-1}\right) - \|e_1\|_{(t)} \log(\|e_1\|_{(t)}) - (1 - \|e_1\|_{(t)}) \log\left(\frac{1 - \|e_1\|_{(t)}}{k-1}\right).$$

Then we conclude with Theorem 5.4.2 to compute explicit parameters. □

Let us remark that in [64] what was actually proved is that violation of additivity can occur for any $k \geq 183$. However, for the output dimension $k = 183$, a probabilistic argument is needed – namely, the limiting distribution of the output of a Bell state and in turn, the fact that one can give a good estimate in probability for $\liminf_n H_{\min}(\Phi_n \otimes \bar{\Phi}_n)$ (again, we refer to [93, 64] for details). However this estimate is difficult to evaluate explicitly as a function of n , therefore, in this paper, we replace it by a slightly weaker estimate that is always true and that yields violation for any $k \geq 184$. In other words we lose one dimension. To conclude, since our bound is explicit, we solve the problem of supplying actual input dimensions for any valid output dimension, for which the violation of MOE will occur. From a point of view of theoretical probability, this is a step towards a large deviation principle. And although our bound is not optimal, our results presumably give the right speed of deviation. However conjecturing a complete large deviation principle and a rate function seems to be beyond the scope of our techniques.

Convergence for noncommutative rational functions evaluated in random matrices

One of the main application of free probability is to show that for appropriately chosen independent copies of d random matrix models, any noncommutative polynomial in these d variables has a spectral distribution that converges asymptotically and can be described through free convolutions. The goal of this paper is to show that this can be extended to noncommutative rational fractions, hereby answering an open question by Roland Speicher.

This chapter is adapted from [12], which is a joint work with Benoît Collins, Tobias Mai, Akihiro Miyagawa and Sheng Yin.

6.1 Introduction

Free probability was invented by Voiculescu 40 years ago as an attempt to understand the group von Neumann algebra of free product of groups. In general, non-trivial free products of groups yield ICC class groups, and their von Neumann algebra is a factor. Therefore they have only one finite trace, and it was natural to study free product factors from the point of view of non-commutative probability spaces – a pair consisting of an algebra and a trace. This point of view has been spectacularly successful, and arguably one of its biggest – and initially unexpected – achievements was to be able to describe the limiting spectrum of noncommutative polynomials in i.i.d. random matrices. This phenomenon is known as asymptotic freeness. It was initially described in [3, 14]. It has been subsequently enhanced in many directions, and a notable direction of improvement was the study of the norm of random matrix models (strong asymptotic freeness), cf. [9]. This result has allowed to turn an area of pure mathematics into a very useful tool for applied mathematics who rely heavily on random matrix models.

The models whose limiting behavior is well understood thanks to asymptotic freeness involve the arithmetic operations of multiplication, addition, scalar multiplication. Other useful operations have also been successfully studied through free probability, such as taking matrix values, or Hadamard products (asymptotic freeness with amalgamation). We refer to [94] for an exposition of many classical and recent results in this direction. Recently, it has also been possible to involve systematically the composition with smooth functions, even at the level of strong convergence, cf. [7, 8].

However, one arithmetic operation that remains largely unexplored in the context of random matrix models is the inverse. The purpose of this paper is to address this question. A recent result in [95] shows that taking inverse is a stable operation if the convergence in distribution is replaced by the strong convergence for matrices. More precisely, for a sequence of matrices that strongly converges in distribution towards a limiting object – an operator which we also call a (noncommutative) random variable – the sequence of their inverses will eventually be well defined and strongly converges in distribution to the inverse of the limiting random variable, provided that this random variable has a bounded inverse. Furthermore, such a result can be extended from an inverse to noncommutative rational functions in multiple variables (a counterpart of commutative rational functions that has been developed by many pioneers, see, for examples [96, 97]) by a recursive structure of rational functions (and their representing rational expressions) or by a linearization trick for rational functions.

However, in order to go beyond the case of bounded evaluations, a problem that one faces about the inverse is that using it might fail to result in a well-defined model, when the inverse is performed on non-invertible matrices. On the other hand, the limiting object has been at the center of the attention of free probabilists recently, and many breakthroughs have been obtained, see among others [69]. Incidentally, the limiting theory relies on the theory of non-commutative rational functions and the embedding question of the rational functions in the generating operators into the algebra of unbounded operators affiliated with the underlying von Neumann algebra. This embedding question was affirmatively answered long ago in [98], whose goal was to provide an answer to the Atiyah conjecture for some family of groups including the free groups. It was recently noted that this result also provided the answer to the well-definedness question for rational functions in freely independent Haar unitary random variables in the context of free probability. Moreover, in [69] the well-definedness was further proved for a large family of random variables beyond the free Haar unitaries.

Let us also mention that there are many natural random matrix models involving the inverse operation and that this is an important topic nowadays, see e.g. [99, 100]. One goal of our manuscript is to provide an unified approach to the study of the limiting spectral distribution under such generality. Therefore, the natural questions are:

- can we make sense of random matrix models involving inverses?
- do they converge towards their natural limiting candidates, whose properties have been unveiled recently?

These questions have been phrased by Speicher during a meeting at MFO in 2019, [101]. Partial answer has been given under some assumptions such as bounded evaluation and specific random matrix models, cf. [95, 99, 102]. The purpose of this paper is to settle these questions in a very general setup. Our main results can be stated as follows.

Theorem 6.1.1. *Let $X^N = (X_1^N, \dots, X_{d_1}^N)$ be a d_1 -tuple of self-adjoint random matrices and let $U^N = (U_1^N, \dots, U_{d_2}^N)$ be a d_2 -tuple of unitary random matrices. Further, let R be a non-degenerate square matrix-valued noncommutative rational expression in $d = d_1 + d_2$ variables which is self-adjoint of type (d_1, d_2) ; see Definition 6.2.9. Suppose that the following conditions are satisfied:*

1. (X^N, U^N) converges almost surely in $*$ -distribution towards a d -tuple of noncommutative random variables (x, u) in some tracial W^* -probability space (\mathcal{M}, τ) satisfying the regularity condition $\Delta(x, u) = d$; see Sections 6.2.4 and 6.2.5.
2. For N large enough $R(X^N, U^N)$ is well-defined almost surely.

Then $R(x, u)$ is well-defined, and the empirical measure of $R(X^N, U^N)$ converges almost surely in law towards the analytic distribution of $R(x, u)$.

The assumption 2 is satisfied for random matrix models (X^N, U^N) whose law on $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ is absolutely continuous with respect to the product measure of the Lebesgue measure on $M_N(\mathbb{C})_{\text{sa}}$ and the Haar measure on $U_N(\mathbb{C})$.

In particular, the assumptions 1 and 2 are satisfied for random matrix models satisfying the following conditions:

- (X^N, U^N) are almost surely asymptotically free.
- The law of each X_j^N has a density with respect to the Lebesgue measure on $M_N(\mathbb{C})_{\text{sa}}$ and its eigenvalue distribution almost surely converges weakly to some compactly supported probability measure on \mathbb{R} that is non-atomic.
- U^N are i.i.d. Haar distributed.

The main part of Theorem 6.1.1 will be proven in Theorem 6.3.6 and Theorem 6.4.4. In Theorem 6.3.1, we prove that condition 2 is satisfied for “absolutely continuous” random matrix models (X^N, U^N) . In combination with this, Corollary 6.2.16 ensures that the particular random matrix model (X^N, U^N) satisfies the assumptions 1 and 2 of Theorem 6.1.1.

An interpretation of our results is as follows: the free field $\mathbb{C}\langle x_1, \dots, x_d \rangle$ together with a $*$ -structure can be endowed with a non-commutative probability structure through its embedding in the $*$ -algebra of operators affiliated to a II_1 factor. Let us elaborate on this non-commutative

probability structure. To each selfadjoint rational function R of the free field one can associate a real valued probability (i.e. an element of $\mathcal{P}(\mathbb{R})$), which is in general unbounded. One sees that this map $\mathbb{C}\langle x_1, \dots, x_d \rangle_{\text{sa}} \rightarrow \mathcal{P}(\mathbb{R})$ strictly supersedes the data of the tracial map on the $*$ -algebra generated by the generators, because in general, elements of the free field do not have moments. In addition, it allows to define directly a non-commutative probability structure on a $*$ -free field without necessarily resorting to von Neumann algebras and affiliated operators. In this context, our result says that any matrix approximation of the free non-commutative tracial $*$ -algebra generated by the generators of the free field in the sense of non-commutative distribution convergence, can be upgraded into a pointwise convergence of the map $\mathbb{C}\langle x_1, \dots, x_d \rangle_{\text{sa}} \rightarrow \mathcal{P}(\mathbb{R})$. As a consequence, this interpretation may allow us to read information on the algebraic side (the free field) out of information on the probabilistic side (probability measures). For example, the inner rank of a self-adjoint matrix A over $\mathbb{C}\langle x_1, \dots, x_d \rangle$ can be seen from the asymptotic proportion of zero eigenvalues in the spectrum of the evaluation of A at any approximation matrices that model the free field.

For a d -tuple X satisfying $\Delta(X) = d$, we know from [69] that the division closure $D(X)$ of $\mathbb{C}\langle X \rangle$ in the $*$ -algebra $W^*(X)$ of all closed and densely defined operators that are affiliated with $W^*(X)$ provides a model of the free field $\mathbb{C}\langle x \rangle$. It would be interesting to find a criterion similar to [103, 104] which allows us to decide whether an element in $W^*(X)$ belongs to the division closure $D(X)$. If $d \geq 2$ and X are free Haar unitaries, then $W^*(X)$ is isomorphic to the free group factor $L(\mathbb{F}_d)$; in this case, such a criterion was provided by Linnell in [104], building on the paper [103] by Duchamp and Reutenauer in which they proved a conjecture of Connes [105].

This paper is organized as follows: following this introductory section, section 6.2 gathers necessary facts about non-commutative rational functions and expressions; section 6.3 shows that the random matrix model is well-defined for dimension large enough and section 6.4 evaluates the limiting distribution.

Acknowledgements. The problem considered in this paper appeared in the context of the MSc studies of A. Miyagawa under the supervision of B. Collins (cf [106]), and related questions were discussed during the visit of T. Mai in Kyoto in 2019. T. Mai is grateful for the great hospitality of B. Collins and the entire Department of Mathematics at Kyoto University. F. Parraud, T. Mai and S. Yin benefited from the hospitality of MFO during which the conjecture leading to this paper was stated by R. Speicher. The authors would like to thank G. Cébron, A. Connes, M. de la Salle and R. Speicher for useful discussions.

BC was supported by JSPS KAKENHI 17K18734 and 17H04823. FP was partially supported by Labex Milyon (ANR-10-LABX-0070) of Université de Lyon. SY was supported by ANR project MESA.

6.2 Preliminaries

6.2.1 Noncommutative rational functions and expressions

Let us denote by $\mathbb{C}\langle x_1, \dots, x_d \rangle$ the algebra of noncommutative polynomials over \mathbb{C} in the indeterminates x_1, \dots, x_d . It is well-known that for its commutative counterpart, namely the ring of commutative polynomials, one can construct uniquely the field of fractions of this polynomial ring by the quotients of polynomials. However, to construct a skew field of fractions of

$\mathbb{C}\langle x_1, \dots, x_d \rangle$ is highly non-trivial. Moreover, there exist skew fields of fractions of $\mathbb{C}\langle x_1, \dots, x_d \rangle$ which are not isomorphic (see, for example, [97, Exercise 7.2.13]). Nevertheless, there exists a unique field of fractions of $\mathbb{C}\langle x_1, \dots, x_d \rangle$ which has some universal property. We call this skew field of fractions of $\mathbb{C}\langle x_1, \dots, x_d \rangle$ the *free field* and denote it by $\mathbb{C}\langle\!\langle x_1, \dots, x_d \rangle\!\rangle$. An element in the free field is called a *noncommutative rational function*.

Since the precise definition of the universal property of the free field is not relevant to this paper, we refer the interested reader to [97, Chapter 7] for a more detailed description (as well as some ring-theoretic construction) of the free field. We take the existence of the free field for granted and apply some recent results about it.

Like a commutative rational function can be represented by a class of quotients of polynomials, a noncommutative rational function can be represented by a class of noncommutative rational expressions. One can think of a noncommutative rational expression as a representation of a noncommutative rational function. Actually, in [96], Amitsur constructed the free field $\mathbb{C}\langle\!\langle x_1, \dots, x_d \rangle\!\rangle$ from noncommutative rational expressions (see also [107, Section 2]).

More precisely, *noncommutative rational expressions* are syntactically valid combinations of \mathbb{C} and symbols x_1, \dots, x_d with $+$, \cdot , $^{-1}$, and $()$, which are respectively corresponding to addition, multiplication, taking inverse, and ordering these operations. For the sake of completeness, let us mention that polynomial expressions are obtained in precisely the same manner but without involving inverses. We admit that this definition, though easy to grasp, is not entirely rigorous as it relies on the tacit agreement about what is meant by syntactically valid. Thus we refer here to [108] for an alternative definition based on the graph theory, by which arithmetic operations for rational expressions can be interpreted as operations on graphs (see also [102, Section IV.1]). We emphasize that noncommutative rational expressions (in contrast to noncommutative rational functions which we are going to define later) are completely formal objects obeying absolutely no arithmetic rules like commutativity or associativity. For example, $x_1 + (-1) \cdot x_1$ and 0 are two distinct rational expressions though clearly they represent the same function. Similarly, $x_1 \cdot (x_2 \cdot x_1)$ and $(x_1 \cdot x_2) \cdot x_1$ are different noncommutative rational expressions, but since they show the same behavior when evaluated on associative algebras, we will write shorthand $x_1 \cdot x_2 \cdot x_1$ or even $x_1 x_2 x_1$ for better legibility as the inherent ambiguity does not cause any problems.

One can also define *matrix-valued noncommutative rational expressions*; see Definition 2.1 in [109]. Those are possible combinations of symbols $A \otimes 1$ and $A \otimes x_j$ for $j = 1, \dots, d$, for each rectangular matrix A over \mathbb{C} of arbitrary size, with $+$, \cdot , $^{-1}$, and $()$, where the operations are required to be compatible with the matrix sizes. Notice that \otimes has only symbolic meaning here, but will turn into the ordinary tensor product (over \mathbb{C}) under evaluation as will be defined below.

Let us enumerate the rules which allow to recursively compute for every matrix-valued noncommutative rational expression R the domain $\text{dom}_{\mathcal{A}}(R)$ of R for every unital complex algebra \mathcal{A} and evaluations $R(X)$ of R at any point $X \in \text{dom}_{\mathcal{A}}(R)$; note that the evaluation $R(X)$ of a $p \times q$ matrix-valued noncommutative rational expression R and every point $X \in \text{dom}_{\mathcal{A}}(R)$ belongs to $M_{p \times q}(\mathbb{C}) \otimes \mathcal{A} \cong M_{p \times q}(\mathcal{A})$.

- If $R = A \otimes 1$ for some $A \in M_{p \times q}(\mathbb{C})$, then R is a $p \times q$ matrix-valued noncommutative rational expression with $\text{dom}_{\mathcal{A}}(R) := \mathcal{A}^d$ and $R(X) := A \otimes 1_{\mathcal{A}}$ for every $X \in \mathcal{A}^d$.
- If $R = A \otimes x_j$ for some $A \in M_{p \times q}(\mathbb{C})$ and $1 \leq j \leq d$, then R is a $p \times q$ matrix-valued noncommutative rational expression with $\text{dom}_{\mathcal{A}}(R) := \mathcal{A}^d$ and $R(X) := A \otimes X_j$ for every

$$X = (X_1, \dots, X_d) \in \mathcal{A}^d.$$

- If R_1, R_2 are $p \times q$ matrix-valued noncommutative rational expressions, then $R_1 + R_2$ is a $p \times q$ matrix-valued noncommutative rational expression with $\text{dom}_{\mathcal{A}}(R_1 + R_2) := \text{dom}_{\mathcal{A}}(R_1) \cap \text{dom}_{\mathcal{A}}(R_2)$ and $(R_1 + R_2)(X) := R_1(X) +_{\mathcal{A}} R_2(X)$ for every $X \in \text{dom}_{\mathcal{A}}(R_1 + R_2)$, where $+_{\mathcal{A}}$ stands for the addition $M_{p \times q}(\mathcal{A}) \times M_{p \times q}(\mathcal{A}) \rightarrow M_{p \times q}(\mathcal{A})$.
- If R_1, R_2 are $p \times q$ respectively $q \times r$ matrix-valued noncommutative rational expressions, then $R_1 \cdot R_2$ is a $p \times r$ matrix-valued noncommutative rational expression with $\text{dom}_{\mathcal{A}}(R_1 \cdot R_2) := \text{dom}_{\mathcal{A}}(R_1) \cap \text{dom}_{\mathcal{A}}(R_2)$ and $(R_1 \cdot R_2)(X) := R_1(X) \cdot_{\mathcal{A}} R_2(X)$ for every $X \in \text{dom}_{\mathcal{A}}(R_1 \cdot R_2)$, where $\cdot_{\mathcal{A}}$ stands for the matrix multiplication $M_{p \times q}(\mathcal{A}) \times M_{q \times r}(\mathcal{A}) \rightarrow M_{p \times r}(\mathcal{A})$.
- If R is a $p \times p$ matrix-valued noncommutative rational expression, then

$$\text{dom}_{\mathcal{A}}(R^{-1}) := \{X \in \text{dom}_{\mathcal{A}}(R) \mid R(X) \text{ is invertible in } M_p(\mathcal{A})\}$$

$$\text{and } R^{-1}(X) := R(X)^{-1} \text{ for every } X \in \text{dom}_{\mathcal{A}}(R^{-1}).$$

Note that the (scalar-valued) noncommutative rational expressions which we have introduced before belong to the strictly larger class of 1×1 matrix-valued noncommutative rational expressions; see Remark 2.11 in [109].

For reader's convenience, we introduce two types of matrix-valued noncommutative rational expressions which are important in a practical sense.

- A noncommutative rational expression evaluated in formal tensor products of matrices and formal variables like as

$$R = r(A_1 \otimes x_1, A_2 \otimes x_2, \dots, A_d \otimes x_d)$$

where r is a (scalar-valued) noncommutative rational expression and $A_i \in M_p(\mathbb{C})$ for $1 \leq i \leq d$. In other words, in this case we amplify formal variables by matrices and then consider their (scalar-valued) rational expression.

- A matrix which consists of (scalar-valued) noncommutative rational expressions

$$R = (r_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}.$$

This can be seen as a $p \times q$ matrix-valued non-commutative rational expression by identifying with $\sum_{ij} (a_i \otimes 1) r_{ij} (b_j \otimes 1)$ where $a_i \in M_{p \times 1}(\mathbb{C})$ and $b_j \in M_{1 \times q}(\mathbb{C})$ are standard basis of \mathbb{C}^p and \mathbb{C}^q . We will implicitly use this viewpoint later (for example, in the proof of Proposition 6.3.4).

A class of matrix-valued noncommutative polynomial expressions are affine linear pencils. An *affine linear pencil (in d variables with coefficients from $M_k(\mathbb{C})$)* is a $k \times k$ matrix-valued noncommutative polynomial expression of the form

$$A = A_0 \otimes 1 + A_1 \otimes x_1 + \dots + A_d \otimes x_d$$

with coefficient matrices A_0, A_1, \dots, A_d belonging to $M_k(\mathbb{C})$. Notice, once again, that we omit the parentheses for better readability as each syntactically valid placement of parentheses will

produce the same result under evaluation. If \mathcal{A} is any unital complex algebra and $X \in \mathcal{A}^d$, then

$$A(X) = A_0 \otimes 1_{\mathcal{A}} + A_1 \otimes X_1 + \cdots + A_d \otimes X_d \in M_k(\mathbb{C}) \otimes \mathcal{A} \cong M_k(\mathcal{A}).$$

Of particular interest are matrix-evaluations. For each matrix-valued noncommutative rational expression R , we put

$$\text{dom}_{M(\mathbb{C})}(R) := \prod_{N=1}^{\infty} \text{dom}_{M_N(\mathbb{C})}(R),$$

i.e., $\text{dom}_{M(\mathbb{C})}(R)$ is the subset of all square matrices over \mathbb{C} where evaluation of R is well-defined. A matrix-valued noncommutative rational expression R is said to be *non-degenerate* if it satisfies $\text{dom}_{M(\mathbb{C})}(R) \neq \emptyset$. In the sequel, we will make use of the following important fact.

Theorem 6.2.1 (Remark 2.3 in [109]). *Let R be a non-degenerate matrix-valued noncommutative rational expression. Then there exists some $N_0 = N_0(R) \in \mathbb{N}$ such that $\text{dom}_{M_N(\mathbb{C})}(R) \neq \emptyset$ for all $N \geq N_0$.*

Two non-degenerate matrix-valued noncommutative rational expressions R_1, R_2 are called *$M(\mathbb{C})$ -evaluation equivalent* if the condition $R_1(X) = R_2(X)$ is satisfied for all X in the intersection of $\text{dom}_{M(\mathbb{C})}(R_1)$ and $\text{dom}_{M(\mathbb{C})}(R_2)$.

As we mentioned earlier, one can construct the free field out of noncommutative rational expressions. This construction can be done by evaluating (scalar-valued) noncommutative rational expressions on scalar-valued matrices. For a non-degenerate noncommutative rational expression r , we denote by $[r]$ its equivalence class of noncommutative rational expressions with respect to $M(\mathbb{C})$ -evaluation equivalence. We endow the set of all such equivalence classes with the arithmetic operations $+$ and \cdot defined by $[r_1] + [r_2] := [r_1 + r_2]$ and $[r_1] \cdot [r_2] := [r_1 \cdot r_2]$. Notice that the arithmetic operations are indeed well-defined as one has $\text{dom}_{M(\mathbb{C})}(r_1) \cap \text{dom}_{M(\mathbb{C})}(r_2) \neq \emptyset$ for any two non-degenerate scalar-valued noncommutative rational expressions r_1 and r_2 ; see the footnote on page 52 of [107], for instance. It is known (see Proposition 2.2 in [107]) that the set of all equivalence classes of noncommutative rational expressions with respect to $M(\mathbb{C})$ -evaluation equivalence endowed with the arithmetic operations $+$ and \cdot forms the free field $\mathbb{C}\langle x_1, \dots, x_d \rangle$.

6.2.2 Linearization

Let us recall the following terminology that was introduced in [110, Definition 4.10].

Definition 6.2.2 (Formal linear representation). *Let R be a $p \times q$ matrix-valued noncommutative rational expression in the variables x_1, \dots, x_d . A formal linear representation $\rho = (u, A, v)$ of R (of dimension k) consists of an affine linear pencil*

$$A = A_0 \otimes 1 + A_1 \otimes x_1 + \cdots + A_d \otimes x_d$$

in d variables and with coefficients A_0, A_1, \dots, A_d from $M_k(\mathbb{C})$ and matrices $u \in M_{p \times k}(\mathbb{C})$ and $v \in M_{k \times q}(\mathbb{C})$, such that the following condition is satisfied: for every unital complex algebra \mathcal{A} , we have that $\text{dom}_{\mathcal{A}}(R) \subseteq \text{dom}_{\mathcal{A}}(A^{-1})$ and for each $X \in \text{dom}_{\mathcal{A}}(R)$ it holds true that $R(X) = uA(X)^{-1}v$, where $A(X) \in M_k(\mathcal{A})$.

Note that we use here a different sign convention by requiring $R(X) = uA(X)^{-1}v$ instead of $R(X) = -uA(X)^{-1}v$; this, however, does not affect the validity of the particular results that we will take from [110]. Furthermore, as we will exclusively work with formal linear representations for matrix-valued noncommutative rational expressions, we will go without specifying them as matrix-valued formal linear representations like it was done in [110].

It follows from [110, Theorem 4.12] that indeed every matrix-valued noncommutative rational expression R admits a formal linear representation $\rho = (u, A, v)$. For the reader's convenience, we include with Theorem 6.2.3 the precise statement as well as its constructive proof. In doing so, we will see that Algorithm 4.11 in [110], on which the proof of Theorem 4.12 in the same paper relies, provides a formal linear representation $\rho = (u, A, v)$ of the $p \times q$ matrix-valued noncommutative rational expression R with the additional property that the dimension k of ρ is larger than $\max\{p, q\}$ and that both u and v have maximal rank; we will call such ρ *proper*. Note that if R is a scalar-valued rational expression, then a proper formal linear representation ρ simply means that u and v are non-zero vectors. In general, due to the restriction $k \geq \max\{p, q\}$, we have that the rank of u is p and the rank of v is q for any proper formal linear representation $\rho = (u, A, v)$ of R . This notion of proper formal linear representation will be important in the sequel.

Theorem 6.2.3 (Theorem 4.12 in [110]). *Every matrix-valued noncommutative rational expression admits a formal linear representation in the sense of Definition 6.2.2 which is also proper.*

Proof. Here, we give the algorithm which inductively builds a proper linear representations of any matrix-valued noncommutative rational expression. For $R = A \otimes 1$ or $R = A \otimes x_j$ for some $A \in M_{p \times q}(\mathbb{C})$ and $1 \leq j \leq d$ we have

$$R(X) = \begin{pmatrix} I_p & 0_{p \times q} \end{pmatrix} \begin{pmatrix} I_p \otimes 1_A & -R(X) \\ 0_{q \times p} & I_q \otimes 1_A \end{pmatrix}^{-1} \begin{pmatrix} 0_{p \times q} \\ I_q \end{pmatrix}$$

where $I_p \in M_p(\mathbb{C})$ is an identity matrix. Clearly we obtain a proper formal linear representation in this way.

If the $p \times q$ matrix-valued noncommutative expressions R_1 and R_2 admit proper formal linear representations (u_1, A_1, v_1) and (u_2, A_2, v_2) then we have

$$(R_1 + R_2)(X) = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \left(\begin{array}{c|c} A_1(X) & 0_{k_1 \times k_2} \\ \hline 0_{k_2 \times k_1} & A_2(X) \end{array} \right)^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

This gives us a proper formal linear representation since $(u_1 \ u_2)$, resp. $(v_1 \ v_2)^T$ is of rank p , resp. q .

If R_1, R_2 are $p \times q$, resp. $q \times r$ matrix-valued noncommutative rational expressions and admit formal linear representations (u_1, A_1, v_1) , resp. (u_2, A_2, v_2) of dimension k_1 , resp. k_2 then we have

$$(R_1 \cdot R_2)(X) = \begin{pmatrix} u_1 & 0_{p \times k_2} \end{pmatrix} \left(\begin{array}{c|c} A_1(X) & -v_1 u_2 \\ \hline 0_{k_2 \times k_1} & A_2(X) \end{array} \right)^{-1} \begin{pmatrix} 0_{k_1 \times r} \\ v_2 \end{pmatrix}.$$

We obtain a proper formal linear representation since $(u_1 \ 0_{p \times k_2})$, resp. $(0_{k_1 \times r} \ v_2)^T$ is of rank p , resp. q .

If R is a $p \times p$ matrix-valued noncommutative rational expression which admits a formal linear representation (u, A, v) of dimension k , then we have

$$R^{-1}(X) = \left(\begin{array}{c|c} I_p & 0_{p \times k} \\ \hline 0_{p \times p} & u \\ v & A(X) \end{array} \right)^{-1} \begin{pmatrix} -I_p \\ 0_{k \times p} \end{pmatrix},$$

where X belongs to an appropriate domain for each step. It is clear that this gives us a proper formal linear representation.

Finally, since all matrix-valued noncommutative rational expressions can be represented by finitely many of the above steps, any matrix-valued noncommutative rational expression has such a formal linear representation which is proper. \square

In the non-degenerate case, formal linear representations are connected with the concept of representations for noncommutative rational functions which is used, for instance, in [111, 112]; this will be addressed in Remark 6.2.5 and Remark 6.2.8. Before, we need to recall the following terminology.

Definition 6.2.4 (Inner rank and fullness). *Let \mathcal{R} be a ring. For $A \in M_{n \times m}(\mathcal{R})$ we define the inner rank $\rho_{\mathcal{R}}(A)$ by*

$$\rho_{\mathcal{R}}(A) = \min\{r \in \mathbb{N} \mid A = BC, B \in M_{n \times r}(\mathcal{R}), C \in M_{r \times m}(\mathcal{R})\},$$

and $\rho_{\mathcal{R}}(0) = 0$. In addition we call A full if $\rho_{\mathcal{R}}(A) = \min\{n, m\}$.

Remark 6.2.5. 1. *Let A be a matrix over noncommutative polynomials in a tuple $x = (x_1, \dots, x_d)$ of formal variables. According to Theorem 7.5.13 in [97] (see also A.2 in [68]), we have*

$$\rho_{\mathbb{C}\langle x \rangle}(A) = \rho_{\mathbb{C}}(A).$$

For this reason, we just say A is full, for a square matrix A over the noncommutative polynomials, without mentioning which algebra we consider.

2. *Let A be an affine linear pencil in x with coefficients taken from $M_k(\mathbb{C})$. We may view A as an element in $M_k(\mathbb{C}) \otimes \mathbb{C}\langle x \rangle \cong M_k(\mathbb{C}\langle x \rangle)$, i.e., $A = A(x)$ is considered as a matrix over the ring $\mathbb{C}\langle x \rangle$. We notice that if there exists a tuple $X \in M_N(\mathbb{C})^d$ such that $A(X)$ is invertible in $M_k(\mathbb{C}) \otimes M_N(\mathbb{C}) \cong M_{kN}(\mathbb{C})$, or equivalently, if $\text{dom}_{M(\mathbb{C})}(A^{-1}) \neq \emptyset$, then A must be full. In fact, if A is not full, then any factorization $A = BC$ with $B \in M_{k \times r}(\mathbb{C}\langle x \rangle)$ and $C \in M_{r \times k}(\mathbb{C}\langle x \rangle)$ for $r = \rho(A) < k$ yields under evaluation $A(X) = B(X)C(X)$ at any point $X \in M_N(\mathbb{C})^d$, so that $A(X)$ is never invertible.*

On the other hand, if A is full, then A is invertible as a matrix over $\mathbb{C}\langle x \rangle$. Indeed, fullness and invertibility are equivalent for any skew field (see Lemma 5.20 in [68]).

3. *Now, let R be a non-degenerate matrix-valued noncommutative rational expression. From Theorem 6.2.3, we know that there exists a formal linear representation $\rho = (u, A, v)$; in particular, we have that*

$$\begin{aligned} \text{dom}_{M(\mathbb{C})}(R) &\subseteq \text{dom}_{M(\mathbb{C})}(A^{-1}) \\ &= \prod_{N=1}^{\infty} \{X \in M_N(\mathbb{C})^d \mid A(X) \text{ invertible in } M_{kN}(\mathbb{C})\}. \end{aligned}$$

Since R is non-degenerate, we find $X \in \text{dom}_{M(\mathbb{C})}(R)$; from the aforementioned inclusion and 2, we infer that A is a full matrix.

4. Suppose that R is a non-degenerate $p \times p$ matrix-valued noncommutative rational expression such that R^{-1} is non-degenerate as well. Let $\rho = (u, A, v)$ be a formal linear representation of R ; we associate to ρ the affine linear pencil

$$\tilde{A} := \begin{pmatrix} 0_{p \times p} & u \\ v & A \end{pmatrix}.$$

We claim that both A and \tilde{A} are full. For A , we already know from 3 that this is true. To check fullness of \tilde{A} , we use that R^{-1} is non-degenerate, which guarantees the existence of some $X \in \text{dom}_{M(\mathbb{C})}(R^{-1})$. Since in particular $X \in \text{dom}_{M(\mathbb{C})}(R)$, we get as ρ is a formal linear representation of R that $A(X)$ is invertible and $R(X) = uA(X)^{-1}v$. Because $X \in \text{dom}_{M(\mathbb{C})}(R^{-1})$, we know that $R(X)$ is invertible. Hence, by the Schur complement formula, it follows that the matrix $\tilde{A}(X)$ is invertible. Thanks to 2, this implies that the affine linear pencil \tilde{A} is full.

The following lemma explains that non-degenerate matrix-valued noncommutative rational expressions induce in some very natural way matrices over the free field.

Lemma 6.2.6. *Let R be a $p \times q$ matrix-valued noncommutative rational expression in d formal variables. If R is non-degenerate, then $x = (x_1, \dots, x_d) \in \text{dom}_{\mathbb{C}\langle x \rangle}(R)$ and consequently $R(x) \in M_{p \times q}(\mathbb{C}\langle x \rangle)$.*

Proof. Let us denote by \mathfrak{R}_0 the set of all non-degenerate matrix-valued noncommutative rational expressions R which have the property $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(R)$. We want to show that \mathfrak{R}_0 consists of all non-degenerate matrix-valued noncommutative rational expressions. In order to verify this assertion, we proceed as follows. Firstly, we notice that both $R_1 + R_2$ and $R_1 \cdot R_2$ belong to \mathfrak{R}_0 whenever we take $R_1, R_2 \in \mathfrak{R}_0$ for which the respective arithmetic operation is defined. Secondly, we consider some $R \in \mathfrak{R}_0$ which is of size $p \times p$ and has the property that R^{-1} is non-degenerate. By Theorem 6.2.3, there exists a formal linear representation $\rho = (u, A, v)$ of R , say of dimension k , and according to Remark 6.2.5 4 we have that both A and the associated affine linear pencil

$$\tilde{A} := \begin{pmatrix} 0_{p \times p} & u \\ v & A \end{pmatrix}$$

are full, i.e., $A(x) \in M_k(\mathbb{C}\langle x \rangle)$ and $\tilde{A}(x) \in M_{k+p}(\mathbb{C}\langle x \rangle)$ become invertible as matrices over the free field $\mathbb{C}\langle x \rangle$. Since $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(R)$ as $R \in \mathfrak{R}_0$, we get $R(x) = uA(x)^{-1}v$, because ρ is a formal linear representation of R . Putting these observations together, the Schur complement formula yields that $R(x) \in M_p(\mathbb{C}\langle x \rangle)$ must be invertible, i.e., $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(R^{-1})$ and thus $R^{-1} \in \mathfrak{R}_0$, as desired. \square

Remark 6.2.7. *With arguments similar to the proof of Lemma 6.2.6 as based on Remark 6.2.5 4, one finds that if R_1, R_2 are non-degenerate matrix-valued noncommutative rational expressions satisfying $R_1(x) = R_2(x)$, then $R_1 \sim_{M(\mathbb{C})} R_2$. In other words, matrix identities over $\mathbb{C}\langle x \rangle$ are preserved under well-defined matrix evaluations.*

Remark 6.2.8. *In the scalar-valued case, the conclusion of Lemma 6.2.6 can be strengthened slightly. For that purpose, it is helpful to denote the formal variables out of which the noncommutative rational expressions are built by χ_1, \dots, χ_d in order to distinguish them from the variables x_1, \dots, x_d of the free skew field $\mathbb{C}\langle x_1, \dots, x_d \rangle$; note that accordingly $x_j = [\chi_j]$ for*

$j = 1, \dots, d$. Now, if r is any scalar-valued noncommutative rational expression in the formal variables χ_1, \dots, χ_d , then Lemma 6.2.6 tells us that $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(r)$ and $r(x) \in \mathbb{C}\langle x \rangle$. Moreover, we have the equality $r(x) = [r]$. This can be shown with a recursive argument similar to the proof of Lemma 6.2.6; notice that if a non-degenerate rational expression r satisfies $r(x) = [r]$ and has the additional property that r^{-1} is non-degenerate, then $r(x) = [r]$ is invertible in $\mathbb{C}\langle x \rangle$, which implies $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(r^{-1})$ with $r^{-1}(x) = [r]^{-1} = [r^{-1}]$.

This has the consequence that every formal linear representation $\rho = (u, A, v)$ of r satisfies $[r] = r(x) = uA(x)^{-1}v$. In the language of [111, 112], this means that the formal linear representation ρ of r induces a (pure and linear) representation of the corresponding noncommutative rational function $[r]$.

6.2.3 Self-adjointness for matrix-valued noncommutative rational expressions

When evaluations of matrix-valued noncommutative rational expressions R at points $X = (X_1, \dots, X_d) \in \text{dom}_{\mathcal{A}}(R)$ for $*$ -algebras \mathcal{A} are considered, it is natural to ask for conditions which guarantee that the result $R(X)$ is self-adjoint, i.e., $R(X)^* = R(X)$. Those conditions shall concern the matrix-valued noncommutative rational expression R itself, but depending on the particular type of its the arguments X_1, \dots, X_d . The case when X_1, \dots, X_d are all self-adjoint was discussed in [110, Section 2.5.7]. The following definition generalizes the latter to matrix-valued noncommutative rational expressions in self-adjoint and unitary variables.

Recall that an element X in a complex $*$ -algebra \mathcal{A} with unit $1_{\mathcal{A}}$ is called *self-adjoint* if $X^* = X$, and $U \in \mathcal{A}$ is said to be *unitary* if $U^*U = 1_{\mathcal{A}} = UU^*$.

Definition 6.2.9 (Self-adjoint matrix-valued noncommutative rational expressions). *Let R be a square matrix-valued noncommutative rational expression in $d_1 + d_2$ formal variables which we denote by $x_1, \dots, x_{d_1}, u_1, \dots, u_{d_2}$. We say that R is self-adjoint of type (d_1, d_2) , if for every unital complex $*$ -algebra \mathcal{A} and all tuples $X = (X_1, \dots, X_{d_1})$ and $U = (U_1, \dots, U_{d_2})$ of self-adjoint respectively unitary elements in \mathcal{A} , the following implication holds:*

$$(X, U) \in \text{dom}_{\mathcal{A}}(R) \quad \implies \quad R(X, U)^* = R(X, U)$$

One comment on this definition is in order. The reader might wonder why the matrix-valued noncommutative rational expressions do not explicitly involve further variables $u_1^*, \dots, u_{d_2}^*$ serving as placeholder for the adjoints of u_1, \dots, u_{d_2} . In fact, for (scalar-valued) noncommutative rational expressions such an approach was presented, for instance, in the appendix of [99] (a version for noncommutative polynomials appears also in [113]); more precisely, noncommutative rational expressions in collections of self-adjoint variables x , non-self-adjoint variables y , and their adjoints y^* were considered. For our purpose, however, this has the slight disadvantage that non-degenerate noncommutative rational expressions of this kind (take $r(y, y^*) = (yy^* - 1)^{-1}$, for example) may have no unitary elements in their domain. On the other hand, there are noncommutative rational expressions (or even noncommutative polynomial expressions such as $yyy^* + y^*yy^*$) which are not self-adjoint on their entire domain but self-adjoint on unitaries.

The following example illustrates Definition 6.2.9 and highlights the effect of having two types of variables.

Example 6.2.10. $x_1 + x_2^{-1}$, $i(u_1 - u_1^{-1})$ and $u_1^{-1}x_1^{-1}u_1$ are self-adjoint non-commutative rational expressions since we have for self-adjoint elements X_1, X_2 and a unitary U_1 in their domain,

$$\begin{aligned} (X_1 + X_2^{-1})^* &= X_1^* + (X_2^*)^{-1} = X_1 + X_2^{-1} \\ i(U_1 - U_1^{-1})^* &= -i(U_1^* - (U_1^*)^{-1}) = i(U_1 - U_1^{-1}) \\ (U_1^{-1}X_1^{-1}U_1)^* &= U_1^*(X_1^*)^{-1}(U_1^*)^{-1} = U_1^{-1}X_1^{-1}U_1. \end{aligned}$$

However, $u_1 + u_2^{-1}$, $i(x_1 - x_1^{-1})$ and $x_1^{-1}u_1^{-1}x_1$ are not self-adjoint in our definition. So we need to be careful to the roles of formal variables when we consider self-adjoint rational expressions. For the matrix-valued case, the 2×2 respectively 1×1 matrix-valued noncommutative rational expressions

$$\begin{pmatrix} x_1^{-1} & u_1 \\ u_1^{-1} & x_2^{-1} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} u_1 & x_1 + iu_2 \\ iu_1^{-1} & x_2 \end{pmatrix}^{-1} \begin{pmatrix} u_1^{-1} \\ x_1 - iu_2^{-1} \end{pmatrix}$$

are self-adjoint of type (2, 1) and (2, 2), respectively.

Like in [110, Definition 4.13] for the case of self-adjoint arguments, we can introduce self-adjoint formal linear representations; see also [99, Definition A.5] for the scalar-valued case.

Note that in order to make the machinery of self-adjoint linearizations ready for further applications, we will switch from now on to a more general situation.

Definition 6.2.11 (Self-adjoint formal linear representation). *Let R be a $p \times p$ matrix-valued noncommutative rational expression in d formal variables x_1, \dots, x_d . A tuple $\rho = (Q, w)$ consisting of an affine linear pencil*

$$Q = A_0 \otimes 1 + \sum_{j=1}^d (B_j \otimes x_j + B_j^* \otimes x_j^*)$$

in the formal variables x_1, \dots, x_d and x_1^, \dots, x_d^* , with coefficients being (not necessarily self-adjoint) matrices B_1, \dots, B_d in $M_k(\mathbb{C})$ for some $k \in \mathbb{N}$, some self-adjoint matrix $A_0 \in M_k(\mathbb{C})$ and some matrix $w \in M_{k \times p}(\mathbb{C})$ is called a self-adjoint formal linear representation of R (of dimension k) if the following condition is satisfied: for every unital complex $*$ -algebra \mathcal{A} and all tuples $X = (X_1, \dots, X_d)$ of (not necessarily self-adjoint) elements in \mathcal{A} , one has*

$$X \in \text{dom}_{\mathcal{A}}(R) \quad \implies \quad (X, X^*) \in \text{dom}_{\mathcal{A}}(Q^{-1})$$

and for every $X \in \text{dom}_{\mathcal{A}}(R)$ for which $R(X)$ is self-adjoint, it holds true that

$$R(X) = w^*Q(X, X^*)^{-1}w.$$

We point out that in contrast to the related concept introduced in [110, Definition 4.13] the existence of a self-adjoint formal linear representation in the sense of the previous Definition 6.2.11 does not enforce R to be self-adjoint at any distinguished points in its domain. In fact, we have the following theorem which says that every square matrix-valued noncommutative rational expression admits a self-adjoint formal linear representation; this is analogous to [110, Theorem 4.14].

Like for formal linear representations, we will say that a self-adjoint formal linear representation $\rho = (Q, w)$ of a self-adjoint $p \times p$ matrix-valued noncommutative rational expression R is *proper* if the dimension k of ρ is larger than p and if w has full rank (i.e., the rank of w is p).

Theorem 6.2.12. *Every square matrix-valued noncommutative rational expression in d formal variables admits a self-adjoint formal linear representation in the sense of Definition 6.2.11 which is proper.*

Proof. Let $\rho = (v, Q, w)$ be a formal linear representation of R in the variables x_1, \dots, x_d with the affine linear pencil Q being of the form

$$Q = A_0 \otimes 1 + \sum_{j=1}^d B_j \otimes x_j.$$

We consider $\tilde{\rho} = (\tilde{Q}, \tilde{w})$ with the affine linear pencil

$$\tilde{Q} = \tilde{A}_0 \otimes 1 + \sum_{j=1}^d (\tilde{B}_j \otimes x_j + \tilde{B}_j^* \otimes x_j^*)$$

in the variables $x_1, \dots, x_d, x_1^*, \dots, x_d^*$ given by

$$\tilde{A}_0 := \begin{pmatrix} 0 & A_0^* \\ A_0 & 0 \end{pmatrix}, \quad \tilde{B}_j := \begin{pmatrix} 0 & 0 \\ B_j & 0 \end{pmatrix}, \quad \text{and} \quad \tilde{w} := \begin{pmatrix} \frac{1}{2}v^* \\ w \end{pmatrix}.$$

One verifies that $\tilde{\rho} = (\tilde{Q}, \tilde{w})$ is a self-adjoint formal linear representation of R which is moreover proper whenever ρ is proper. \square

Notice that if R is a $p \times p$ matrix-valued noncommutative rational expression in $d_1 + d_2$ formal variables $x_1, \dots, x_{d_1}, u_1, \dots, u_{d_2}$ which is self-adjoint of type (d_1, d_2) , then each self-adjoint formal linear representation of R can be brought into the simplified form $\rho = (Q, w)$ with an affine linear pencil

$$Q = A_0 \otimes 1 + \sum_{j=1}^{d_1} A_j \otimes x_j + \sum_{j=1}^{d_2} (B_j \otimes u_j + B_j^* \otimes u_j^*)$$

in the formal variables $x_1, \dots, x_{d_1}, u_1, \dots, u_{d_2}, u_1^*, \dots, u_{d_2}^*$ with coefficients being self-adjoint matrices A_0, A_1, \dots, A_{d_1} and (not necessarily self-adjoint) matrices B_1, \dots, B_{d_2} in $M_k(\mathbb{C})$ for some $k \in \mathbb{N}$ and some matrix $w \in M_{k \times p}(\mathbb{C})$; indeed Theorem 6.2.12 yields a self-adjoint formal linear representation of R with an affine linear pencil in the formal variables $x_1, \dots, x_{d_1}, x_1^*, \dots, x_{d_1}^*$ and $u_1, \dots, u_{d_2}, u_1^*, \dots, u_{d_2}^*$, from which we obtain Q of the asserted form by replacing $x_1^*, \dots, x_{d_1}^*$ by x_1, \dots, x_{d_1} and merging their coefficients. In particular, we have

$$(X, U) \in \text{dom}_{\mathcal{A}}(R) \quad \implies \quad (X, U, U^*) \in \text{dom}_{\mathcal{A}}(Q^{-1})$$

and for every $(X, U) \in \text{dom}_{\mathcal{A}}(R)$ it holds true that

$$R(X, U) = w^* Q(X, U, U^*)^{-1} w.$$

Example 6.2.13. *We return to the self-adjoint noncommutative rational expressions presented in Example 6.2.10. Let us construct a self-adjoint formal linearization of $x_1 + x_2^{-1}$. Using the algorithm from [110] which we recalled in the proof of Theorem 6.2.3, we obtain first a formal linear representation*

$$X_1 + X_2^{-1} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -X_1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & X_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Out of the latter, we construct with the help of Theorem 6.2.12 the self-adjoint formal linear representation

$$X_1 + X_2^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -X_1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & X_2 \\ 1 & -X_1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & X_2 & 0 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

The second example is $u_1 + u_1^{-1}$. Since we have for unitary U_1 in any $*$ -algebra

$$U_1 + U_1^{-1} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -U_1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & U_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

we have a formal self-adjoint linearization

$$U_1 + U_1^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -U_1^* & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & U_1^* \\ 1 & -U_1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & U_1 & 0 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

6.2.4 Unbounded random variables

In this subsection, we set (\mathcal{M}, τ) to be a tracial W^* -probability space (i.e., a von Neumann algebra \mathcal{M} that is endowed with a faithful normal tracial state $\tau : \mathcal{M} \rightarrow \mathbb{C}$). The condition that τ is a trace is necessary since we are going to consider closed and densely defined operators affiliated with the von Neumann algebra \mathcal{M} . We will simply call these operators unbounded operators. In general, unbounded operators might not well-behave under either addition or composition. However, in the case of tracial W^* -probability space, they form a $*$ -algebra, denoted by $\widetilde{\mathcal{M}}$, which provides us a framework in which one has well-defined evaluations of rational expressions.

In a language of probability, this framework allows us to consider random variables that may not have compact support or even finite moments. For a normal random variable X in a W^* -probability space (\mathcal{M}, τ) , we know that X has finite moments of all orders and its analytic distribution μ_X determined by the moments (i.e., the probability measure associated to X by a representation theorem of Riesz) has a compact support. For an (unbounded) operator X in $\widetilde{\mathcal{M}}$, it may not have finite moments. But we could still associate a probability measure to X via the spectral theorem. We refer the interested reader to [114, 115] for more details on unbounded operators (which are also known as measurable operators as the noncommutative analogue of measurable functions, cf. [116]).

Let $\mathcal{P}(\mathcal{M})$ denote the set of self-adjoint projections in \mathcal{M} and let $\widetilde{\mathcal{M}}_{sa}$ be the set of self-adjoint elements in $\widetilde{\mathcal{M}}$. Given an element $X \in \widetilde{\mathcal{M}}_{sa}$, for a Borel set B on \mathbb{R} , we denote by $\mathbf{1}_B(X) \in \mathcal{P}(\mathcal{M})$ the spectral projection of X on B given by the spectral theorem (see, for example, [117]). Then we can associate a probability measure μ_X to X as follows.

Definition 6.2.14. For $X \in \widetilde{\mathcal{M}}_{\text{sa}}$, we define its analytic distribution μ_X by

$$\mu_X(B) := \tau(\mathbf{1}_B(X)), \quad \text{for all Borel sets } B \subseteq \mathbb{R}.$$

Furthermore, we define the cumulative distribution function of X as the function $\mathcal{F}_X : \mathbb{R} \rightarrow [0, 1]$ given by,

$$\mathcal{F}_X(t) := \int_{-\infty}^t 1 d\mu_X(s) = \tau(\mathbf{1}_{(-\infty, t]}(X)).$$

In particular, if we take $\mathcal{M} = L^\infty(\Omega, \mathbb{P})$ and $\tau = \mathbb{E}$ for some probability measure space (Ω, \mathbb{P}) , then $\widetilde{\mathcal{M}}$ is the $*$ -algebra consisting of all measurable functions, i.e., classical random variables. Moreover, the analytic distribution and cumulative distribution defined above coincide with their classical counterparts.

Recall that for a probability measure μ on \mathbb{R} . A number $\lambda \in \mathbb{R}$ is called an *atom* of μ if $\mu(\{\lambda\}) \neq 0$. Thus for a random variable X in $\widetilde{\mathcal{M}}_{\text{sa}}$, we say that $\lambda \in \mathbb{R}$ is an atom for X if λ is an atom for μ_X . Moreover, we see that X has an atom $\lambda \in \mathbb{R}$ if and only if $p_{\ker(\lambda - X)} \neq 0$, where $p_{\ker(\lambda - X)} \in \mathcal{P}(\mathcal{M})$ is the orthogonal projection onto the kernel of $\lambda - X$ (in the Hilbert space $L^2(\mathcal{M}, \tau)$). For an atom λ of X , we have

$$\mu_X(\{\lambda\}) = \tau(p_{\ker(\lambda - X)}).$$

A closely related notion is a rank defined via the image. That is, we define

$$\text{rk}(X) := \tau(p_{\overline{\text{im}X}}),$$

where $p_{\overline{\text{im}X}}$ is the orthogonal projection onto the closure of the image of X . The following alternative description of this rank will be needed later:

$$\text{rk}(X) = \inf\{\tau(r) \mid r \in \mathcal{P}(\mathcal{M}), rX = X\}. \quad (6.1)$$

Clearly, since $p_{\overline{\text{im}(X)}}X = X$, we have $\inf\{\tau(r) \mid r \in \mathcal{P}(\mathcal{M}), rX = X\} \leq \tau(p_{\overline{\text{im}(X)}}) = \text{rk}(X)$. To see it is an equality, note that for any $r \in \mathcal{P}(\mathcal{M})$ satisfying $rX = X$, $\text{im}(X) \subseteq \text{im}(r)$, which implies that $p_{\overline{\text{im}(X)}} \leq r$.

6.2.5 The quantity Δ

The regularity condition which we impose in Theorem 6.4.4 on the limit of the considered random matrix model involves the quantity Δ which was introduced by Connes and Shlyakhtenko in [118]. We briefly recall the definition. Let (\mathcal{M}, τ) be a tracial W^* -probability space and consider a tuple $x = (x_1, \dots, x_d)$ of (not necessarily self-adjoint) noncommutative random variables in \mathcal{M} . We denote by $\mathcal{F}(L^2(\mathcal{M}, \tau))$ the ideal of all finite rank operators on $L^2(\mathcal{M}, \tau)$ and by J Tomita's conjugation operator, i.e., the conjugate-linear map $J : L^2(\mathcal{M}, \tau) \rightarrow L^2(\mathcal{M}, \tau)$ that extends isometrically the conjugation $x \mapsto x^*$ on \mathcal{M} . We then put

$$\Delta(x) := d - \dim_{\mathcal{M} \overline{\otimes} \mathcal{M}^{\text{op}}} \overline{\left\{ (T_1, \dots, T_d) \in \mathcal{F}(L^2(\mathcal{M}, \tau))^d \mid \sum_{j=1}^d [T_j, Jx_j^*J] = 0 \right\}}^{\text{HS}},$$

where the closure is taken with respect to the Hilbert-Schmidt norm. Note that in contrast to [118], we do not require the set $\{x_1, \dots, x_d\}$ to be closed under the involution $*$; see also [69]. Despite this slight deviation from the setting of [118], the following result remains true.

Theorem 6.2.15 (Theorem 3.3 (e) in [118]). *Let $1 \leq k < d$ and suppose that the sets $\{x_1, \dots, x_k\}$ and $\{x_{k+1}, \dots, x_d\}$ are freely independent, then*

$$\Delta(x_1, \dots, x_d) = \Delta(x_1, \dots, x_k) + \Delta(x_{k+1}, \dots, x_d).$$

Further, we recall from [69, Corollary 6.4] that $\Delta(u) = d$ for every d -tuple u of freely independent Haar unitary elements in (\mathcal{M}, τ) .

In the particular case of a d -tuple x consisting of self-adjoint operators in \mathcal{M} , Corollary 4.6 in [118] says that $d \geq \Delta(x) \geq \delta(x)$, where $\delta(x)$ denotes the so-called *microstates free entropy dimension* which was introduced by Voiculescu in [119, Definition 6.1]. Now, if the x_1, \dots, x_d are freely independent, then Proposition 6.4 in [119] tells us that

$$\delta(x) = d - \sum_{j=1}^d \sum_{t \in \mathbb{R}} \mu_{x_j}(\{t\})^2,$$

where μ_{x_j} is the analytic distribution of the operator x_j in the sense of Definition 6.2.14. We infer that $\Delta(x_1, \dots, x_d) = d$ if x_1, \dots, x_d are self-adjoint, freely independent and their individual analytic distributions $\mu_{x_1}, \dots, \mu_{x_d}$ are all non-atomic. For reference, we summarize these observations by the following corollary.

Corollary 6.2.16. *Let $x = (x_1, \dots, x_{d_1})$ be a d_1 -tuple of self-adjoint and freely independent elements in (\mathcal{M}, τ) with $\mu_{x_1}, \dots, \mu_{x_{d_1}}$ being non-atomic. Further, let $u = (u_1, \dots, u_{d_2})$ be a d_2 -tuple of freely independent Haar unitary elements in (\mathcal{M}, τ) . Suppose that x and u are freely independent. Then $\Delta(x, u) = d_1 + d_2$.*

6.3 Evaluations of non-degenerate matrix-valued non-commutative rational expressions

By definition, every non-degenerate matrix-valued noncommutative rational expression has a non-empty domain when evaluations in matrices of sufficiently large size are considered. In this section, we show that actually much more is true. Namely, we establish that the assumptions of Theorem 6.4.4 are satisfied in very general situations.

6.3.1 Evaluations in random matrices

The following result asserts loosely spoken that one can almost surely evaluate every non-degenerated matrix-valued noncommutative rational expressions in “absolutely continuous” random matrix models, provided that their size is large enough. The precise statement reads as follows.

Theorem 6.3.1. *Let R be a matrix-valued noncommutative rational expression in $d = d_1 + d_2$ formal variables which is non-degenerate. Suppose that μ_{d_1, d_2}^N is a probability measure on $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ which is absolutely continuous with respect to the product measure of the Lebesgue measure on $M_N(\mathbb{C})_{\text{sa}}$ and the Haar measure on $U_N(\mathbb{C})$. If (X^N, U^N) is a tuple of random matrices in $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ with law μ_{d_1, d_2}^N , then there exists some $N_0 \in \mathbb{N}$ such that almost surely $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R)$ for all $N \geq N_0$.*

Remark 6.3.2. For the validity of Theorem 6.3.1, it is essential to work over the field of complex numbers. In order to see this, consider the scalar-valued noncommutative rational expression $r = (x_1x_2 - x_2x_1)^{-1}$. Note that there are real matrices X_1, X_2 at which one can evaluate r , but in $M_N(\mathbb{R})$ for N odd there cannot exist symmetric real matrices X_1, X_2 at which evaluation $r(X_1, X_2)$ would be defined, since necessarily $\det(X_1X_2 - X_2X_1) = 0$ because

$$\det(X_1X_2 - X_2X_1) = \det\left((X_1X_2 - X_2X_1)^T\right) = -\det(X_1X_2 - X_2X_1).$$

It is consistent with this observation that the proof of Proposition 6.3.3, on which Theorem 6.3.1 relies, will make use of complex analysis techniques.

One can also see an algebraic description of existence of a symmetric matrix in a domain of noncommutative rational expression in [120, Remark 6.7].

The proof of Theorem 6.3.1 relies on a study of evaluations of affine linear pencils. The first step is the following proposition, which requires some notation. Consider an affine linear pencil

$$Q = A_0 \otimes 1 + \sum_{j=1}^{d_1} A_j \otimes x_j + \sum_{j=1}^{d_2} B_j \otimes u_j \quad (6.2)$$

in the variables $x = (x_1, \dots, x_{d_1})$ and $u = (u_1, \dots, u_{d_2})$, say with coefficients A_0, A_1, \dots, A_{d_1} and B_1, \dots, B_{d_2} taken from $M_k(\mathbb{C})$. We regard Q as an element in

$$M_k(\mathbb{C}) \otimes \mathbb{C}\langle x, u \rangle \cong M_k(\mathbb{C}\langle x, u \rangle).$$

Given an d -tuple $Z = (Z', Z'')$ of matrices in $M_N(\mathbb{C})$, we consider the evaluation of Q at Z which is given by

$$Q(Z) := A_0 \otimes 1 + \sum_{j=1}^{d_1} A_j \otimes Z'_j + \sum_{j=1}^{d_2} B_j \otimes Z''_j,$$

where $Q(Z)$ lies in $M_k(\mathbb{C}) \otimes M_N(\mathbb{C}) \cong M_{kN}(\mathbb{C})$. Building on such evaluations, we associate to Q functions

$$\phi_Q^{(N)} : M_N(\mathbb{C})^d \longrightarrow \mathbb{C}, \quad Z \longmapsto \det(Q(Z))$$

for every $N \in \mathbb{N}$. Notice that $\phi_Q^{(N)}$ is a holomorphic commutative polynomial in the dN^2 complex matrix entries appearing in the tuple Z . This allows us to use the complex analysis machinery in order to relate $\phi_Q^{(N)}$ and its restriction to the real space $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$.

Proposition 6.3.3. Let Q be an affine linear pencil of the form (6.2) in $M_k(\mathbb{C}) \otimes \mathbb{C}\langle x, u \rangle$ and let $N \in \mathbb{N}$. If $\phi_Q^{(N)}|_{M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}} \equiv 0$, then $\phi_Q^{(N)} \equiv 0$.

Proof. Fix any $Z = (Z', Z'') \in M_N(\mathbb{C})^{d_1} \times M_N(\mathbb{C})^{d_2}$ and suppose that the d_2 -tuple Z'' consists of invertible matrices. We write $Z' = X + iY$ with the tuples $X = (X_1, \dots, X_{d_1}), Y = (Y_1, \dots, Y_{d_1}) \in M_N(\mathbb{C})_{\text{sa}}^{d_1}$ that are given by $X_j := \Re(Z'_j)$ and $Y_j := \Im(Z'_j)$ for $j = 1, \dots, d_1$. Further, for $j = 1, \dots, d_2$, we consider the polar decomposition $Z''_j = P_j U_j$ of Z''_j with a positive definite matrix $P_j \in M_N(\mathbb{C})$ and $U_j \in U_N(\mathbb{C})$. As the matrices P_1, \dots, P_{d_2} are positive definite, we can define a holomorphic function $f : \mathbb{C} \rightarrow \mathbb{C}$ by

$$f(z) := \phi_Q^{(N)}\left(X_1 + zY_1, \dots, X_{d_1} + zY_{d_1}, \exp(-iz \log(P_1))U_1, \dots, \exp(-iz \log(P_{d_2}))U_{d_2}\right)$$

for $z \in \mathbb{C}$. Due to the assumption that $\phi_Q^{(N)}|_{M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}} \equiv 0$, we have that $f|_{\mathbb{R}} \equiv 0$. Thus, by the identity principle, it follows that f vanishes identically on \mathbb{C} . In particular, $\phi_Q^{(N)}(Z) = f(i) = 0$. This shows that $\phi_Q^{(N)}$ vanishes on all d -tuples $Z = (Z', Z'') \in M_N(\mathbb{C})^{d_1} \times M_N(\mathbb{C})^{d_2}$ satisfying the condition that Z'' consists of invertible matrices. Since those are dense in $M_N(\mathbb{C})^d$, the assertion follows. \square

With the help of Proposition 6.3.3, we see that fullness of affine linear pencils Q can be detected by evaluations of Q at points in $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$.

Proposition 6.3.4. *Let Q be an affine linear pencil of the form (6.2) in $M_k(\mathbb{C}) \otimes \mathbb{C}\langle x, u \rangle$ which is full. Then there exists $N_0 \in \mathbb{N}$ with the following property: for each $N \geq N_0$, we have that $\phi_Q^{(N)}|_{M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}} \not\equiv 0$, i.e., one can find some d -tuple $(X^N, U^N) \in M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ for which $Q(X^N, U^N)$ becomes invertible in $M_{kN}(\mathbb{C})$.*

Proof. First, we note that there exists some $N_0 \in \mathbb{N}$ such that $\phi_Q^{(N)} \not\equiv 0$ for all $N \geq N_0$. This fact is well-known (see Proposition 2.4 in [121], for instance), but we include the argument for the sake of completeness. Since Q is full, $Q(x)$ is invertible as a matrix over the free skew field $\mathbb{C}\langle x, u \rangle$; see Remark 6.2.5 2. Its inverse $Q(x)^{-1} \in M_k(\mathbb{C}\langle x, u \rangle)$ is represented by some non-degenerate $k \times k$ matrix-valued noncommutative rational expression R , i.e., we have $Q^{-1}(x) = R(x)$; this follows by applying Remark 6.2.8 entrywise. From Theorem 6.2.1, we know that there exists some $N_0 \in \mathbb{N}$ such that $\text{dom}_{M_N(\mathbb{C})}(R) \neq \emptyset$ for all $N \geq N_0$. Thanks to Remark 6.2.7, the identity $Q(x, u)R(x, u) = I_k$ over $\mathbb{C}\langle x, u \rangle$ continues to hold on $\text{dom}_{M_N(\mathbb{C})}(R)$, and by applying determinants, we infer that $\phi_Q^{(N)} \not\equiv 0$ for all $N \geq N_0$, as desired.

Having this, Proposition 6.3.3 guarantees that $\phi_Q^{(N)}$ does not vanish identically on all of $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$, as we wished to show. \square

In the next step, we involve the concrete random matrix model that we want to consider.

Proposition 6.3.5. *Let Q be an affine linear pencil of the form (6.2) in $M_k(\mathbb{C}) \otimes \mathbb{C}\langle x, u \rangle$ which is full. For $N \in \mathbb{N}$, let (X^N, U^N) be a random matrix in $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ with an absolutely continuous law μ_{d_1, d_2}^N like in Theorem 6.3.1. Then there exists $N_0 \in \mathbb{N}$ such that almost surely $Q(X^N, U^N)$ is invertible in $M_k(\mathbb{C}) \otimes M_N(\mathbb{C}) \cong M_{kN}(\mathbb{C})$ for all $N \geq N_0$.*

Proof. Thanks to Proposition 6.3.4, since Q is supposed to be full, there is an $N_0 \in \mathbb{N}$ such that none of the functions $\phi_Q^{(N)}|_{M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}}$ for $N \geq N_0$ can vanish identically. Notice that $M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}$ is a real manifold of dimension dN^2 . In suitable local charts, we see that $\phi_Q^{(N)}|_{M_N(\mathbb{C})_{\text{sa}}^{d_1} \times U_N(\mathbb{C})^{d_2}}$ induces a real analytic functions on an open subset of \mathbb{R}^{dN^2} and can therefore vanish only on a set of Lebesgue measure 0. Due to the choice of μ_{d_1, d_2}^N , we conclude that, for each $N \geq N_0$, the random matrix $Q(X^N, U^N)$ is almost surely invertible in $M_{kN}(\mathbb{C})$. \square

Proof of Theorem 6.3.1. We define the set \mathfrak{R}_0 of all non-degenerate matrix-valued noncommutative rational expressions R for which the conclusion of Theorem 6.3.1 is true, i.e., there exists $N_0 \in \mathbb{N}$ such that almost surely $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R)$ for all $N \geq N_0$. We have to prove that \mathfrak{R}_0 consists in fact of all non-degenerate matrix-valued noncommutative rational expressions.

Notice that obviously all matrix-valued noncommutative polynomial expressions belong to \mathfrak{R}_0 . Further, it is easily seen that both $R_1 + R_2$ and $R_1 \cdot R_2$ are in \mathfrak{R}_0 whenever we take $R_1, R_2 \in$

\mathfrak{R}_0 for which the respective arithmetic operation makes sense. Therefore, it only remains to prove that if $R \in \mathfrak{R}_0$ is square and enjoys the property that R^{-1} is non-degenerate, then necessarily $R^{-1} \in \mathfrak{R}_0$. In order to verify this, we take any square matrix-valued noncommutative rational expression R belonging to \mathfrak{R}_0 for which R^{-1} is non-degenerate. Further, let $\rho = (v, Q, w)$ be a formal linear representation of R in the sense of Definition 6.2.2, say of dimension k ; see Theorem 6.2.3.

By assumption, we have that R^{-1} is a non-degenerate matrix-valued noncommutative rational expression. Thus, Remark 6.2.5 4 gives us that the affine linear pencil in d variables with coefficients from $M_{k+p}(\mathbb{C})$ which is given by

$$\tilde{Q} := \begin{pmatrix} 0_{p \times p} & v \\ w & Q \end{pmatrix}$$

is full. Therefore, Proposition 6.3.5 tells us that an $N_0 \in \mathbb{N}$ exists such that almost surely $\tilde{Q}(X^N, U^N)$ is invertible in $M_{(k+p)N}(\mathbb{C})$ for all $N \geq N_0$. Since $R \in \mathfrak{R}_0$, we may suppose that (after enlarging N_0 if necessary) that at the same time almost surely $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R)$ for all $N \geq N_0$. Because ρ is a formal linear representation, the latter implies that almost surely $Q(X^N, U^N)$ is invertible and $R(X^N, U^N) = vQ(X^N, U^N)^{-1}w$ for all $N \geq N_0$. Putting these observations together, we see, again with the help of the Schur complement formula, that almost surely $R(X^N, U^N)$ is invertible for all $N \geq N_0$. In other words, we have almost surely that $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R^{-1})$ for all $N \geq N_0$. The latter means that $R^{-1} \in \mathfrak{R}_0$, as desired. \square

6.3.2 Evaluation in operators with maximal Δ

It follows from [69, Theorem 1.1] that for any d -tuple $X = (X_1, \dots, X_d)$ of (not necessarily self-adjoint) operators in some W^* -probability space (\mathcal{M}, τ) which satisfy the ‘‘regularity condition’’ $\Delta(X) = d$, where Δ stands for a quantity that was introduced in [118] and which we discussed in Section 6.2.5, then the canonical evaluation homomorphism

$$\text{ev}_X : \mathbb{C}\langle x_1, \dots, x_d \rangle \rightarrow \mathcal{M}$$

which is determined by $1 \mapsto 1$ and $x_j \mapsto X_j$ for $j = 1, \dots, d$ extends to an injective homomorphism

$$\text{Ev}_X : \mathbb{C}\langle\!\langle x_1, \dots, x_d \rangle\!\rangle \rightarrow \widetilde{\mathcal{M}}$$

into the $*$ -algebra $\widetilde{\mathcal{M}}$ of all closed and densely defined operators affiliated with \mathcal{M} ; see Section 6.2.4.

While the result of [69] addresses evaluations of noncommutative rational functions, it leaves open the question whether also all non-degenerate rational expressions can be evaluated; indeed, this is not immediate as the domain of a rational function is larger than the domain of any of its representing non-degenerate noncommutative rational expressions. This question is answered to the affirmative by the next theorem, which gives the conclusion even in the matrix-valued case. For that purpose, we will consider the canonical amplifications

$$\text{Ev}_X^\bullet : M_\bullet(\mathbb{C}\langle\!\langle x_1, \dots, x_d \rangle\!\rangle) \rightarrow M_\bullet(\widetilde{\mathcal{M}}).$$

Theorem 6.3.6. *Let $X = (X_1, \dots, X_d)$ be a d -tuple of (not necessarily self-adjoint) operators in some tracial W^* -probability space (\mathcal{M}, τ) satisfying $\Delta(X) = d$. Then, for every non-degenerate*

matrix-valued noncommutative rational expression R , we have that $X \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$ and $R(X) = \text{Ev}_X^\bullet(R(x))$, where $R(x)$ is the matrix over $\mathbb{C}\langle x_1, \dots, x_d \rangle$ associated to R via Lemma 6.2.6.

Proof. The proof is similar to the proof of Theorem 6.3.1. Here, we consider the set \mathfrak{R}_0 of all non-degenerate matrix-valued noncommutative rational expressions r for which the conclusion of Theorem 6.3.6 is true, i.e., we have $X \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$ and $R(X) = \text{Ev}_X^\bullet(R(x))$. We want to show that \mathfrak{R}_0 consists of all non-degenerate matrix-valued noncommutative rational expressions. This can be done in almost the same way as in Theorem 6.3.1, except some slight modification in the last step. Suppose that $R \in \mathfrak{R}_0$ is of size $p \times p$ and has the property that R^{-1} is non-degenerate. Consider a formal linear representation $\rho = (u, A, v)$ of r , say of dimension k . Like in the proof of Theorem 6.3.1, we deduce from Remark 6.2.5 4 that the associated affine linear pencil

$$\tilde{A} := \begin{pmatrix} 0 & u \\ v & A \end{pmatrix}$$

is full. Now, by applying [69, Theorem 5.6] instead of Proposition 6.3.5, we get that $\tilde{A}(X)$ is invertible. Having this, we can proceed again like in the proof of Theorem 6.3.1 and we arrive at $X \in \text{dom}_{\widetilde{\mathcal{M}}}(R^{-1})$. Moreover, since $R(X) = \text{Ev}_X^\bullet(R(x))$ by the assumption $R \in \mathfrak{R}_0$, we further get that $R^{-1}(X) = R(X)^{-1} = \text{Ev}_X^\bullet(R(x))^{-1} = \text{Ev}_X^\bullet(R(x)^{-1}) = \text{Ev}_X^\bullet(R^{-1}(x))$; notice that $R(x)$ is invertible because Lemma 6.2.6 guarantees that $x \in \text{dom}_{\mathbb{C}\langle x \rangle}(R^{-1})$ as R^{-1} was assumed to be non-degenerate. In summary, we see that $R^{-1} \in \mathfrak{R}_0$. \square

6.4 Convergence in law of the spectral measure

6.4.1 Estimate on the cumulative distribution function of the spectral measure of self-adjoint operators

In this subsection we simply list and prove a few properties that we need in the next subsection to prove Theorem 6.4.4 about the convergence of the empirical measure of a self-adjoint non-degenerate matrix-valued noncommutative rational expression evaluated in matrices towards the analytic distribution of the limiting operator.

Lemma 6.4.1 (Lemma 3.2 in [122]). *For $X \in \widetilde{\mathcal{M}}_{\text{sa}}$ and $t \in \mathbb{R}$ we have*

$$\mathcal{F}_X(t) = \max\{\tau(p) \mid p \in \mathcal{P}(\mathcal{M}), p(t - X)p \geq 0\}.$$

The crux of the proof of Theorem 6.4.4 lies in the following two lemmas.

Lemma 6.4.2. *Let $X, Y \in \widetilde{\mathcal{M}}_{\text{sa}}$, then*

$$\sup_{t \in \mathbb{R}} |\mathcal{F}_X(t) - \mathcal{F}_{X+Y}(t)| \leq \text{rk}(Y).$$

Proof. We fix $t \in \mathbb{R}$. Let $r \in \mathcal{P}(\mathcal{M})$ be such that $rY = Y$ and $q \in \mathcal{P}(\mathcal{M})$ such that $q(t - X)q \geq 0$. Then if we set $p = q \wedge (1 - r)$, we have $pY = 0$ and $pq = p$, thus

$$p(t - X - Y)p = p(t - X)p = pq(t - X)qp \geq 0.$$

Consequently

$$\mathcal{F}_{X+Y}(t) \geq \tau(p) \geq \tau(q) - \tau(r).$$

By taking the supremum over q and the infimum over r we get that

$$\mathcal{F}_{X+Y}(t) \geq \mathcal{F}_X(t) - \text{rk}(Y).$$

Now let's assume that q is such that $q(t - X - Y)q \geq 0$, then similarly with $p = q \wedge (1 - r)$,

$$p(t - X)p = p(t - X - Y)p = pq(t - X - Y)pq \geq 0.$$

Hence

$$\mathcal{F}_X(t) \geq \tau(p) \geq \tau(q) - \tau(r).$$

And once again by taking the supremum over q and the infimum over r we get that

$$\mathcal{F}_X(t) \geq \mathcal{F}_{X+Y}(t) - \text{rk}(Y).$$

Hence the conclusion. □

The authors are indebted to Mikael de la Salle for indicating them the following lemma.

Lemma 6.4.3. *Let $p \in \mathcal{P}(\mathcal{M})$, $X \in \widetilde{\mathcal{M}}_{sa}$, then $\text{rk}(pXp) \leq \text{rk}(X)$.*

Proof. Let $q \in \mathcal{P}(\mathcal{M})$ be such that $qX = X$, $r = p \wedge (1 - q)$, then $r + 1 - p$ is such that

$$(r + 1 - p)pXp = rpXp = rXp = rqXp = 0.$$

Consequently $(p - r)pXp = pXp$. And since $p \geq r$, $p - r$ is a self-adjoint projection, hence

$$\text{rk}(pXp) \leq \tau(p - r) \leq \tau(q).$$

Hence the conclusion by taking the infimum over q . □

6.4.2 Main result

This subsection focuses on proving the convergence in law of the empirical measure of matrix-valued noncommutative rational expressions evaluated in matrices satisfying some assumptions. Theorem 6.4.4 is for deterministic matrices, but it can easily be extended to random matrices by applying this result almost surely.

Theorem 6.4.4. *Let $X^N = (X_1^N, \dots, X_{d_1}^N)$ be a d_1 -tuple of deterministic self-adjoint matrices and let $U^N = (U_1^N, \dots, U_{d_2}^N)$ be a d_2 -tuple of deterministic unitary matrices. Further, let R be a non-degenerate square matrix-valued noncommutative rational expression in $d = d_1 + d_2$ variables which is self-adjoint of type (d_1, d_2) in the sense of Definition 6.2.9. Suppose that the following conditions are satisfied:*

1. (X^N, U^N) converges in $*$ -distribution towards a d -tuple of noncommutative random variables (x, u) in some tracial W^* -probability space (\mathcal{M}, τ) satisfying $\Delta(x, u) = d$.
2. For N large enough $R(X^N, U^N)$ is well-defined, i.e., there exists $N_0 \in \mathbb{N}$ such that $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R)$ for all $N \geq N_0$.

Then $(x, u) \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$, so that $R(x, u)$ is well-defined, and the empirical measure of $R(X^N, U^N)$ converges in law towards the analytic distribution of $R(x, u)$.

The fact that $(x, u) \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$ holds was established already in Theorem 6.3.6. Accordingly, the main statement of Theorem 6.4.4 is the convergence of the empirical measure of $R(X^N, U^N)$ towards the spectral measure of $R(x, u)$. This convergence result actually holds in a more general setting than the above theorem. We summarize it as the following proposition.

Proposition 6.4.5. *For each $N \in \mathbb{N}$, let $X^N = (X_1^N, \dots, X_d^N)$ be a d -tuple of noncommutative random variables in some tracial W^* -probability space $(\mathcal{M}^{(N)}, \tau^{(N)})$. Further, let X^N converge in $*$ -distribution towards a d -tuple $X = (X_1, \dots, X_d)$ of noncommutative random variables in some tracial W^* -probability space (\mathcal{M}, τ) . Let R be a square matrix-valued noncommutative rational expression in d variables such that, for all $N \in \mathbb{N}$ which are sufficiently large,*

1. $X^N \in \text{dom}_{\widetilde{\mathcal{M}^{(N)}}}(R)$ and $X \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$,
2. $R(X^N)$ and $R(X)$ are self-adjoint.

Then the analytic distribution of $R(X^N)$ converges in law towards the analytic distribution of $R(X)$.

Once Proposition 6.4.5 is shown, the statement on the convergence in Theorem 6.4.4 follows immediately. Indeed, the condition formulated in Item 1 of Proposition 6.4.5 is satisfied as we have $(X^N, U^N) \in \text{dom}_{M_N(\mathbb{C})}(R)$ for all $N \geq N_0$ by Item 2 of Theorem 6.4.4 and $(x, u) \in \text{dom}_{\widetilde{\mathcal{M}}}(R)$ by Theorem 6.3.6; further, we have that $R(X^N, U^N)$ for all $N \geq N_0$ and $R(x, u)$ are self-adjoint thanks to Definition 6.2.9 as R is supposed to be self-adjoint of type (d_1, d_2) , so that the condition in Item 2 of Proposition 6.4.5 is fulfilled as well.

Let us provide an outline of the proof of Proposition 6.4.5. Let $\rho = (Q, w)$ be a self-adjoint formal linear representation of R in the sense of Definition 6.2.11 which is moreover proper as given by Theorem 6.2.12. Thanks to Lemma 6.4.2, we can ignore the singularity in 0 of $Q(X, X^*)^{-1}$. More precisely, as long as the spectral measure of $Q(X, X^*)$ has no atom at 0, we can use Lemma 6.4.2 to prove that the cumulative distribution function of $w^*Q(X, X^*)^{-1}w$ is close to the one of $w^*f_\varepsilon(Q(X^N, X^{N*}))w$ where f_ε is a continuous function which is equal to $t \mapsto t^{-1}$ outside of a neighborhood of 0 of size ε . Then we can use the convergence in $*$ -distribution of X^N to show that the cumulative distribution function of $w^*f_\varepsilon(Q(X^N, X^{N*}))w$ converges towards the correct limit when we let N go to infinity and ε go to 0.

It is important to note that in this subsection, by convergence in $*$ -distribution of X^N of noncommutative random variables $X^N = (X_1^N, \dots, X_d^N)$, we mean that the trace of any noncommutative $*$ -polynomial P evaluated in X^N converges towards the trace of $P(X, X^*)$ where X is a d -tuple of noncommutative random variables in some tracial W^* -probability space. In particular, this does *not* exclude the case where the operator norm of X_i^N is not bounded over N . This forces us to do a few more computations since the convergence in law of the analytic measure of $P(X^N, X^{N*})$ towards the analytic measure of the limiting operator, while still true, is not immediate anymore.

Proof of Proposition 6.4.5. Let $\rho = (Q, w)$ be a proper self-adjoint formal linear representation (of dimension k) of R . If $p \in \mathbb{N}$ is the size of R , then since $k \geq p$ and w has full rank, there exist a matrix $T \in \text{GL}_k(\mathbb{C})$ such that $w = Tw_0$ where $w_0 \in M_{k \times p}(\mathbb{C})$ is the rectangular matrix whose diagonal coefficients are all 1, and non diagonal coefficients are all 0. By replacing Q by T^*QT , one can assume without loss of generality that $w = w_0$.

Notice that by assumption $Q_N := Q(X^N, X^{N*})$ is invertible in $\widetilde{\mathcal{M}}^{(N)}$ and $R(X^N) = w^*Q_N^{-1}w$. Further, we have also that $Q_\infty := Q(X, X^*)$ is invertible in $\widetilde{\mathcal{M}}$ and $R(X) = w^*Q_\infty^{-1}w$. To prove

the convergence in law, we need to prove that $\mathcal{F}_{w^*Q_N w}(t)$ converges towards $\mathcal{F}_{w^*Q_\infty w}(t)$ for $t \in \mathbb{R}$ such that the function $s \mapsto \mathcal{F}_{w^*Q_\infty w}(s)$ is continuous in t . To do so, let $g : t \mapsto t^{-1}$ and $f_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function such that on the complementary set of $[-\varepsilon, \varepsilon]$, $f_\varepsilon = g$. We have for any $t \in \mathbb{R}$,

$$\begin{aligned} \left| \mathcal{F}_{w^*Q_N^{-1}w}(t) - \mathcal{F}_{w^*Q_\infty^{-1}w}(t) \right| &\leq \left| \mathcal{F}_{w^*f_\varepsilon(Q_N)w}(t) - \mathcal{F}_{w^*f_\varepsilon(Q_\infty)w}(t) \right| \\ &\quad + \left| \mathcal{F}_{w^*Q_N^{-1}w}(t) - \mathcal{F}_{w^*f_\varepsilon(Q_N)w}(t) \right| \\ &\quad + \left| \mathcal{F}_{w^*Q_\infty^{-1}w}(t) - \mathcal{F}_{w^*f_\varepsilon(Q_\infty)w}(t) \right|. \end{aligned}$$

Thus thanks to Lemma 6.4.2,

$$\begin{aligned} \left| \mathcal{F}_{w^*Q_N^{-1}w}(t) - \mathcal{F}_{w^*Q_\infty^{-1}w}(t) \right| &\leq \left| \mathcal{F}_{w^*f_\varepsilon(Q_N)w}(t) - \mathcal{F}_{w^*f_\varepsilon(Q_\infty)w}(t) \right| \\ &\quad + \text{rk}(w^*(f_\varepsilon - g)(Q_N)w) \\ &\quad + \text{rk}(w^*(f_\varepsilon - g)(Q_\infty)w). \end{aligned}$$

Since $w = w_0$, we have that for any $X \in M_p(\widetilde{\mathcal{M}})$,

$$\begin{aligned} \text{rk}(wXw^*) &= \text{rk} \begin{pmatrix} X & 0_{p \times (k-p)} \\ 0_{(k-p) \times p} & 0_{k-p} \end{pmatrix} \\ &= \frac{p}{k} \text{rk}(X). \end{aligned}$$

This implies that

$$\text{rk}(w^*(f_\varepsilon - g)(Q_\infty)w) = \frac{k}{p} \times \text{rk}(ww^*(f_\varepsilon - g)(Q_\infty)ww^*) \leq \frac{k}{p} \times \text{rk}((f_\varepsilon - g)(Q_\infty)),$$

where in the last inequality we used Lemma 6.4.3. Besides $\mathbf{1}_{[-\varepsilon, \varepsilon]}(Q_\infty)$ is a self-adjoint projection such that $\mathbf{1}_{[-\varepsilon, \varepsilon]}(Q_\infty)(f_\varepsilon - g)(Q_\infty) = (f_\varepsilon - g)(Q_\infty)$. Consequently with Tr_k the non-normalized trace on $M_k(\mathbb{C})$ and τ the trace on \mathcal{M} ,

$$\text{rk}(w^*(f_\varepsilon - g)(Q_\infty)w) \leq \frac{1}{p} \text{Tr}_k \otimes \tau(\mathbf{1}_{[-\varepsilon, \varepsilon]}(Q_\infty)).$$

Let h_ε be a continuous function which takes value 1 on $[-\varepsilon, \varepsilon]$, 0 outside of $[-2\varepsilon, 2\varepsilon]$ and in $[0, 1]$ elsewhere, then

$$\text{rk}(w^*(f_\varepsilon - g)(Q_\infty)w) \leq \frac{1}{p} \text{Tr}_k \otimes \tau(h_\varepsilon(Q_\infty)). \quad (6.3)$$

Hence with similar computations we get

$$\begin{aligned} \left| \mathcal{F}_{w^*Q_N^{-1}w}(t) - \mathcal{F}_{w^*Q_\infty^{-1}w}(t) \right| &\leq \left| \mathcal{F}_{w^*f_\varepsilon(Q_N)w}(t) - \mathcal{F}_{w^*f_\varepsilon(Q_\infty)w}(t) \right| \\ &\quad + \frac{1}{p} \text{Tr}_k \otimes \tau(h_\varepsilon(Q_\infty)) \\ &\quad + \frac{1}{p} \text{Tr}_k \otimes \tau^{(N)}(h_\varepsilon(Q_N)). \end{aligned}$$

In order to use the Portmanteau theorem, we want to prove that the analytic distribution of $w^* f_\varepsilon(Q_N)w$ converges towards the analytic distribution of $w^* f_\varepsilon(Q_\infty)w$. However since this self-adjoint operator is uniformly bounded over N , we simply need to prove the convergence of the moments. That is, that

$$\lim_{N \rightarrow \infty} \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^l \right) = \frac{1}{p} \operatorname{Tr}_p \otimes \tau \left((w^* f_\varepsilon(Q_\infty)w)^l \right)$$

for any l . The strategy consists in approximating f_ε by a polynomial then use the convergence in $*$ -distribution of X^N . However, the fact that we did not assume the operator norm of the matrices X_i^N to be bounded over N , forces us to make additional estimates.

Let $C = \|Q_\infty\| + 1$, and h be a non-negative continuous function which takes value 0 on $[-C, C]$, 1 outside of $[-C-1, C+1]$ and in $[0, 1]$ elsewhere. Let P_m be a polynomial such that $\|f_\varepsilon - P_m\|_{C^0([-C-1, C+1])} \leq 1/m$. We set

$$B^N := (f_\varepsilon(Q_N) - P_m(Q_N))(1 - h(Q_N)) \quad \text{and} \quad C^N := (f_\varepsilon(Q_N) - P_m(Q_N))h(Q_N),$$

then

$$\begin{aligned} & \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^l \right) - \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* P_m(Q_N)w)^l \right) \\ &= \sum_{i=1}^l \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^{i-1} w^* (B^N + C^N) w (w^* P_m(Q_N)w)^{l-i} \right). \end{aligned}$$

Thanks to the Cauchy-Schwarz inequality, we have for any i ,

$$\begin{aligned} & \left| \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^{i-1} w^* (B^N + C^N) w (w^* P_m(Q_N)w)^{l-i} \right) \right| \\ & \leq \left(\sqrt{\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} (w^* B^N w w^* B^N w)} + \sqrt{\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} (w^* C^N w w^* C^N w)} \right) \\ & \quad \times \sqrt{\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* P_m(Q_N)w)^{2(l-i)} (w^* f_\varepsilon(Q_N)w)^{2(i-1)} \right)} \\ & \leq \left(\sqrt{\operatorname{Tr}_k \otimes \tau^{(N)} ((B^N)^2)} + \sqrt{\operatorname{Tr}_k \otimes \tau^{(N)} ((C^N)^2)} \right) \\ & \quad \times \left(\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* P_m(Q_N)w)^{4(l-i)} \right) \right)^{1/4} \\ & \quad \times \left(\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^{4(i-1)} \right) \right)^{1/4}. \end{aligned}$$

Since f_ε is bounded by a constant K , we have that

$$\frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N)w)^{4(i-1)} \right) \leq K^{4(i-1)}.$$

Thanks to the convergence in $*$ -distribution of X^N , and since the expression $w^* P_m(Q_N)w$ is a matrix of polynomials in X^N , we have

$$\lim_{N \rightarrow \infty} \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* P_m(Q_N)w)^{4(l-i)} \right) = \frac{1}{p} \operatorname{Tr}_p \otimes \tau \left((w^* P_m(Q_\infty)w)^{4(l-i)} \right),$$

which means that

$$\lim_{N \rightarrow \infty} \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* P_m(Q_N) w)^{4(l-i)} \right) \leq (K + 1/m)^{4(l-i)}.$$

We also have

$$\operatorname{Tr}_k \otimes \tau^{(N)} \left((B^N)^2 \right) \leq \frac{k}{m}.$$

Finally since f_ε is bounded, there exists an integer g such that for any $t \in \mathbb{R}$, $|f_\varepsilon(t) - P_m(t)| \leq (1 + t^2)^g$, thus for any $r \geq 0$,

$$\begin{aligned} \operatorname{Tr}_k \otimes \tau^{(N)} \left((C^N)^2 \right) &\leq \frac{\operatorname{Tr}_k \otimes \tau^{(N)} \left((C^N)^2 Q_N^{2r} \right)}{N C^{2r}} \\ &\leq \frac{\operatorname{Tr}_k \otimes \tau^{(N)} \left((1 + Q_N^2)^{2g} Q_N^{2r} \right)}{N C^{2r}}. \end{aligned}$$

And so for any $r \geq 0$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \operatorname{Tr}_k \otimes \tau^{(N)} \left((C^N)^2 \right) &\leq \frac{\operatorname{Tr}_k \otimes \tau \left((1 + Q_\infty^2)^{2g} Q_\infty^{2r} \right)}{C^{2r}} \\ &\leq k \left\| (1 + Q_\infty^2)^{2g} \right\| \frac{(C - 1)^{2r}}{C^{2r}}. \end{aligned}$$

So by letting r go to infinity, we get

$$\lim_{N \rightarrow \infty} \operatorname{Tr}_k \otimes \tau^{(N)} \left((C^N)^2 \right) = 0.$$

By combining those results, we obtain

$$\limsup_{N \rightarrow \infty} \left| \frac{1}{p} \operatorname{Tr}_p \otimes \tau^{(N)} \left((w^* f_\varepsilon(Q_N) w)^l \right) - \frac{1}{p} \operatorname{Tr}_p \otimes \tau \left((w^* f_\varepsilon(Q_\infty) w)^l \right) \right| = \mathcal{O}(1/m).$$

Thus, by letting m go to infinity we get the convergence of the moments. This implies that the analytic distribution of $w^* f_\varepsilon(Q_N) w$ converges towards the analytic distribution of $w^* f_\varepsilon(Q_\infty) w$. Thanks to Portmanteau's theorem and Lemma 6.4.1, we have

$$\begin{aligned} \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(t) &\geq \limsup_{N \rightarrow \infty} \mathcal{F}_{w^* f_\varepsilon(Q_N) w}(t) \\ &\geq \liminf_{N \rightarrow \infty} \mathcal{F}_{w^* f_\varepsilon(Q_N) w}(t) \geq \lim_{s \rightarrow t, s < t} \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(s). \end{aligned}$$

Consequently,

$$\begin{aligned} &\limsup_{N \rightarrow \infty} \left| \mathcal{F}_{w^* Q_N^{-1} w}(t) - \mathcal{F}_{w^* Q_\infty^{-1} w}(t) \right| \\ &\leq \lim_{s \rightarrow t, s < t} \left| \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(t) - \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(s) \right| + \frac{2}{p} \operatorname{Tr}_k \otimes \tau \left(h_\varepsilon(Q_\infty) \right), \end{aligned}$$

where we used the convergence in $*$ -distribution of X^N once again in the last line, coupled with an argument similar to the one which let us prove the convergence of the moments of $w^* f_\varepsilon(Q_N) w$. But by using Lemma 6.4.2 one more time we have

$$\begin{aligned} &\left| \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(t) - \mathcal{F}_{w^* f_\varepsilon(Q_\infty) w}(s) \right| \\ &\leq \left| \mathcal{F}_{w^* Q_\infty^{-1} w}(t) - \mathcal{F}_{w^* Q_\infty^{-1} w}(s) \right| + 2 \operatorname{rk}(w^*(f_\varepsilon - g)(Q_\infty)w). \end{aligned}$$

Hence by using equation (6.3), we have that

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \left| \mathcal{F}_{w^* Q_N^{-1} w}(t) - \mathcal{F}_{w^* Q_\infty^{-1} w}(t) \right| \\ & \leq \lim_{s \rightarrow t, s < t} \left| \mathcal{F}_{w^* Q_\infty w}(t) - \mathcal{F}_{w^* Q_\infty w}(s) \right| + \frac{4}{p} \text{Tr}_k \otimes \tau(h_\varepsilon(Q_\infty)). \end{aligned}$$

Since we made the assumption that t was a point of continuity of the function $s \mapsto \mathcal{F}_{w^* Q_\infty w}(s)$, we have that $\lim_{s \rightarrow t, s < t} \left| \mathcal{F}_{w^* Q_\infty w}(t) - \mathcal{F}_{w^* Q_\infty w}(s) \right| = 0$. Besides, by the dominated convergence theorem, $\lim_{\varepsilon \rightarrow 0} \text{Tr}_k \otimes \tau(h_\varepsilon(Q_\infty)) = \text{Tr}_k \otimes \tau(\mathbf{1}_{\{0\}}(Q_\infty))$, which is equal to 0 since otherwise the distribution of Q_∞ would have an atom in 0, in contradiction to the invertibility of Q_∞ ; indeed, analogous to the proof of [69, Corollary 5.13], we notice that $Q_\infty \mathbf{1}_{\{0\}}(Q_\infty) = 0$ and conclude from the latter that since Q_∞ is invertible over $\widetilde{\mathcal{M}}$ we necessarily have that $\mathbf{1}_{\{0\}}(Q_\infty) = 0$ and hence $\mu_{Q_\infty}(\{0\}) = \frac{1}{k} \text{Tr}_k \otimes \tau(\mathbf{1}_{\{0\}}(Q_\infty)) = 0$. \square

Bibliography

- [1] J. Wishart, “The generalized product moment distribution in samples from a normal multivariate population,” *Biometrika*, vol. 20A, no. 1/2, pp. 32–52, 1928.
- [2] E. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions,” *The Annals of Mathematics*, vol. 62, no. 3, pp. 548–564, 1955.
- [3] D. Voiculescu, “Limit laws for random matrices and free products,” *Inventiones mathematicae*, vol. 104, no. 1, pp. 201–220, 1991.
- [4] G. Akemann, J. Baik, and P. D. Francesco, *The Oxford Handbook of Random Matrix Theory*. Oxford Handbooks in Mathematics Series, Oxford University Press, 2015.
- [5] B. Collins and I. Nechita, “Random matrix techniques in quantum information theory,” *Journal of Mathematical Physics*, vol. 57, no. 1, 2016.
- [6] U. Haagerup and S. Thorbjørnsen, “A new application of random matrices: $\text{ext}(c_{\text{red}}^*(\mathcal{U}_2))$ is not a group,” *Annals of Mathematics*, vol. 162, no. 2, pp. 711–775, 2005.
- [7] B. Collins, A. Guionnet, and F. Parraud, “On the operator norm of non-commutative polynomials in deterministic matrices and iid gue matrices,” *arXiv:1912.04588*, 2019.
- [8] F. Parraud, “On the operator norm of non-commutative polynomials in deterministic matrices and iid Haar unitary matrices,” *arXiv:2005.13834*, 2020.
- [9] B. Collins and C. Male, “The strong asymptotic freeness of Haar and deterministic matrices,” *Annales Scientifiques de l’Ecole Normale Supérieure*, vol. 47, no. 1, pp. 147–163, 2014.
- [10] F. Parraud, “Asymptotic expansion of smooth functions in polynomials in deterministic matrices and iid gue matrices,” *arXiv:2011.04146*, 2020.
- [11] B. Collins and F. Parraud, “Concentration estimates for random subspaces of a tensor product, and application to quantum information theory,” *arXiv:2012.00159*, 2020.
- [12] B. Collins, T. Mai, A. Miyagawa, F. Parraud, and S. Yin, “Convergence for non-commutative rational functions evaluated in random matrices,” *arXiv:2103.05962*, 2021.
- [13] G. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, vol. 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2010.

- [14] D. Voiculescu, “A strengthened asymptotic freeness result for random matrices with applications to free entropy,” *International Mathematics Research Notices*, vol. 1998, no. 1, pp. 41–63, 1998.
- [15] Z. Füredi and J. Komlós, “The eigenvalues of random symmetric matrices,” *Combinatorica*, vol. 1, no. 3, pp. 233–241, 1981.
- [16] Z.-D. Bai and Y.-Q. Yin, “Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix,” *The Annals of Probability*, vol. 16, no. 4, pp. 1729–1741, 1988.
- [17] C. A. Tracy and H. Widom, “Level-spacing distributions and the Airy kernel,” *Communications in Mathematical Physics*, vol. 159, no. 1, pp. 151–174, 1994.
- [18] A. Soshnikov, “Universality at the edge of the spectrum in Wigner random matrices,” *Communications in mathematical physics*, vol. 207, no. 3, pp. 697–733, 1999.
- [19] L. Erdős and H.-T. Yau, *A dynamical approach to random matrix theory*, vol. 28 of *Courant Lecture Notes*. American Mathematical Soc., 2017.
- [20] T. Tao and V. Vu, “Random matrices: Universality of local eigenvalue statistics up to the edge,” *Communications in Mathematical Physics*, vol. 298, no. 2, pp. 549–572, 2010.
- [21] J. O. Lee and J. Yin, “A necessary and sufficient condition for edge universality of Wigner matrices,” *Duke Mathematical Journal*, vol. 163, no. 1, pp. 117–173, 2014.
- [22] L. Erdős, B. Schlein, and H.-T. Yau, “Wegner estimate and level repulsion for Wigner random matrices,” *International Mathematics Research Notices*, vol. 2010, no. 3, pp. 436–479, 2010.
- [23] A. Figalli and A. Guionnet, “Universality in several-matrix models via approximate transport maps,” *Acta mathematica*, vol. 217, no. 1, pp. 81–176, 2016.
- [24] L. Erdős, T. Krüger, and Y. Nemish, “Local laws for polynomials of Wigner matrices,” *Journal of Functional Analysis*, vol. 278, no. 12, p. 108507, 2020.
- [25] H. Schultz, “Non-commutative polynomials of independent Gaussian random matrices. the real and symplectic cases.,” *Probability theory and related fields*, vol. 131, no. 2, pp. 261–309, 2005.
- [26] M. Capitaine and D. Martin, “Strong asymptotic freeness for Wigner and Wishart matrices,” *Indiana university mathematics journal*, vol. 56, no. 2, pp. 767–804, 2007.
- [27] G. W. Anderson, “Convergence of the largest singular value of a polynomial in independent Wigner matrices,” *The Annals of Probability*, vol. 41, no. 3B, pp. 2103–2181, 2013.
- [28] C. Male, “The norm of polynomials in large random and deterministic matrices,” *Probability Theory and Related Fields*, vol. 154, no. 3-4, pp. 477–532, 2012.

-
- [29] S. T. Belinschi and M. Capitaine, “Spectral properties of polynomials in independent Wigner and deterministic matrices,” *Journal of Functional Analysis*, vol. 273, no. 12, pp. 3901–3963, 2017.
- [30] C. Stein, “A bound for the error in the normal approximation to the distribution of a sum of dependent random variables,” *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 2, pp. 583–602, 1972.
- [31] G. Pisier, “Random matrices and subexponential operator spaces,” *Israel Journal of Mathematics*, vol. 203, no. 1, pp. 223–273, 2012.
- [32] B. Hayes, “A random matrix approach to the Peterson-Thom conjecture,” *arXiv:2008.12287*, 2020.
- [33] J. Harer and D. Zagier, “The euler characteristic of the moduli space of curves,” *Inventiones mathematicae*, vol. 85, no. 3, pp. 457–485, 1986.
- [34] A. Zvonkin, “Matrix integrals and map enumeration: an accessible introduction,” *Mathematical and Computer Modelling*, vol. 26, no. 8-10, pp. 281–304, 1997.
- [35] G. ’t Hooft, “Magnetic monopoles in unified gauge theories,” *Nuclear Physics: B*, vol. 79, no. 2, pp. 276–284, 1974.
- [36] E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber, “Planar diagrams,” *Communications in Mathematical Physics*, vol. 59, no. 1, pp. 35–51, 1978.
- [37] N. M. Ercolani and K.-R. McLaughlin, “Asymptotics of the partition function for random matrices via Riemann-Hilbert techniques and applications to graphical enumeration,” *International Mathematics Research Notices*, vol. 2003, no. 14, pp. 755–820, 2003.
- [38] S. Albeverio, L. Pastur, and M. Shcherbina, “On the $1/n$ expansion for some unitary invariant ensembles of random matrices,” *Communications in Mathematical Physics*, vol. 224, no. 1, pp. 271–305, 2001.
- [39] A. Guionnet and E. Maurel-Segala, “Combinatorial aspects of matrix models,” *Alea*, vol. 1, pp. 241–279, 2006.
- [40] A. Guionnet and E. Maurel-Segala, “Second order asymptotics for matrix models,” *The Annals of Probability*, vol. 35, no. 6, pp. 2160–2212, 2007.
- [41] E. Maurel-Segala, “High order asymptotics of matrix models and enumeration of maps,” *arXiv:math/0608192v1*, 2006.
- [42] B. Collins, A. Guionnet, and E. Maurel-Segala, “Asymptotics of unitary and orthogonal matrix integrals,” *Advances in Mathematics*, vol. 222, no. 1, pp. 172–215, 2009.
- [43] A. Guionnet and J. Novak, “Asymptotics of unitary multimatrix models: The Schwinger–Dyson lattice and topological recursion,” *Journal of Functional Analysis*, vol. 268, no. 10, pp. 2851–2905, 2015.

- [44] M. Shcherbina, “Asymptotic expansions for β -matrix models and their applications to the universality conjecture, random matrix theory, interacting particle systems, and integrable systems,” *Mathematical Sciences Research Institute Publications*, vol. 65, pp. 463–482, 2014.
- [45] L. O. Chekhov, B. Eynard, and O. Marchal, “Topological expansion of the β -ensemble model and quantum algebraic geometry in the sectorwise approach,” *Theoretical and Mathematical Physics*, vol. 166, no. 2, pp. 141–185, 2011.
- [46] G. Borot and A. Guionnet, “Asymptotic expansion of β matrix models in the one-cut regime,” *Communications in Mathematical Physics*, vol. 317, no. 2, pp. 447–483, 2013.
- [47] G. Borot and A. Guionnet, “Asymptotic expansion of β matrix models in the multi-cut regime,” *arXiv:1303.1045*, 2013.
- [48] G. Borot, A. Guionnet, and K. K. Kozłowski, “Large- N asymptotic expansion for mean field models with Coulomb gas interaction,” *International Mathematics Research Notices*, vol. 2015, no. 20, pp. 10451–10524, 2015.
- [49] G. Borot, A. Guionnet, and K. K. Kozłowski, *Asymptotic expansion of a partition function related to the sinh-model*. Mathematical Physics Studies, Springer, 2016.
- [50] F. David, “Loop equations and non perturbative effects in two-dimensional quantum gravity,” *Modern Physics Letters A*, vol. 5, no. 13, pp. 1019–1029, 1990.
- [51] V. A. Kazakov, “The appearance of matter fields from quantum fluctuations of 2d-gravity,” *Modern Physics Letters A*, vol. 4, no. 22, pp. 2125–2139, 1989.
- [52] B. Eynard and N. Orantin, “Invariants of algebraic curves and topological expansion,” *Communications in Number Theory and Physics*, vol. 1, no. 2, pp. 347–452, 2007.
- [53] B. Eynard, “Topological expansion for the 1-hermitian matrix model correlation functions,” *Journal of High Energy Physics*, vol. 2004, no. 11, p. 031, 2005.
- [54] L. Chekhov and B. Eynard, “Matrix eigenvalue model: Feynman graph technique for all genera,” *Journal of High Energy Physics*, vol. 2006, no. 12, p. 026, 2006.
- [55] G. Borot, B. Eynard, and N. Orantin, “Abstract loop equations, topological recursion and new applications,” *Communications in Number Theory and Physics*, vol. 9, no. 1, pp. 51–187, 2015.
- [56] L. Chekhov and B. Eynard, “Hermitian matrix model free energy: Feynman graph technique for all genera,” *Journal of High Energy Physics*, vol. 2006, no. 3, p. 014, 2006.
- [57] B. Eynard and N. Orantin, “Topological expansion of the 2-matrix model correlation functions: diagrammatic rules for a residue formula,” *Journal of High Energy Physics*, vol. 2005, no. 12, p. 034, 2005.
- [58] J. Ambjørn, L. Chekhov, C. Kristjansen, and Y. Makeenko, “Matrix model calculations beyond the spherical limit,” *Nuclear Physics, Section B*, vol. 404, no. 1-2, pp. 127–172, 1993.

-
- [59] U. Haagerup and S. Thorbjørnsen, “Asymptotic expansions for the Gaussian unitary ensemble,” *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, vol. 15, no. 01, p. 1250003, 2012.
- [60] A. Nica and R. Speicher, *Lectures on the combinatorics of free probability*, vol. 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 2006.
- [61] M. B. Hastings, “Superadditivity of communication capacity using entangled inputs,” *Nature Physics*, vol. 5, no. 4, pp. 255–257, 2009.
- [62] P. Hayden and A. Winter, “Counterexamples to the maximal p -norm multiplicativity conjecture for all $p > 1$,” *Communications in mathematical physics*, vol. 284, no. 1, pp. 263–280, 2008.
- [63] F. G. Brandao and M. Horodecki, “On Hastings’ counterexamples to the minimum output entropy additivity conjecture,” *Open Systems & Information Dynamics*, vol. 17, no. 01, pp. 31–52, 2010.
- [64] S. T. Belinschi, B. Collins, and I. Nechita, “Almost one bit violation for the additivity of the minimum output entropy,” *Communications in Mathematical Physics*, vol. 341, no. 3, pp. 885–909, 2016.
- [65] G. Aubrun, S. Szarek, and E. Werner, “Hastings’s additivity counterexample via Dvoretzky’s theorem,” *Communications in mathematical physics*, vol. 305, no. 1, pp. 85–97, 2011.
- [66] B. Collins, “Haagerup’s inequality and additivity violation of the minimum output entropy,” *Houston Journal of Mathematics*, vol. 44, no. 1, pp. 253–261, 2016.
- [67] S. Belinschi, B. Collins, and I. Nechita, “Eigenvectors and eigenvalues in a random subspace of a tensor product,” *Inventiones mathematicae*, vol. 190, no. 3, pp. 647–697, 2012.
- [68] T. Mai, R. Speicher, and S. Yin, “The free field: zero divisors, Atiyah property and realizations via unbounded operators,” *arXiv:1805.04150*, 2018.
- [69] T. Mai, R. Speicher, and S. Yin, “The free field: realization via unbounded operators and Atiyah property,” *arXiv:1905.08187*, 2019.
- [70] T. Cabanal-Duvillard, “Fluctuations de la loi empirique de grandes matrices aléatoires,” in *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, vol. 37, pp. 373–402, 2001.
- [71] A. Guionnet, “Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices,” in *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, vol. 38, pp. 341–384, 2002.
- [72] J. A. Mingo and R. Speicher, “Second order freeness and fluctuations of random matrices: I. Gaussian and Wishart matrices and cyclic Fock spaces,” *Journal of Functional Analysis*, vol. 235, no. 1, pp. 226–270, 2006.

- [73] P. Biane and R. Speicher, “Free diffusions, free entropy and free Fisher information,” in *Annales de l’Institut Henri Poincaré (B) Probability and Statistics*, vol. 37, pp. 581–606, 2001.
- [74] D. Shlyakhtenko and P. Skoufranis, “Freely independent random variables with non-atomic distributions,” *Transactions of the American Mathematical Society*, vol. 367, no. 9, pp. 6267–6291, 2015.
- [75] N. P. Brown and N. Ozawa, *C*-Algebras and Finite-Dimensional Approximations*, vol. 88 of *Graduate studies in mathematics*.
- [76] G. J. Murphy, *C*-algebras and operator theory*. Elsevier Science, 1990.
- [77] A. Guionnet, “Large random matrices: Lectures on macroscopic asymptotics,” *Ecole d’Été de Probabilités de Saint-Flour XXXVI–2006*, 2009.
- [78] G. Pisier, *Introduction to operator space theory*. No. 294, Cambridge University Press, 2003.
- [79] J. A. Mingo, P. Śniady, and R. Speicher, “Second order freeness and fluctuations of random matrices: II. unitary random matrices,” *Advances in Mathematics*, vol. 209, no. 1, pp. 212–240, 2007.
- [80] A. Guionnet, *Asymptotics of Random Matrices and Related Models: The Uses of Dyson-Schwinger Equations*, vol. 130. American Mathematical Society, 2019.
- [81] P. Biane and R. Speicher, “Stochastic calculus with respect to free Brownian motion and analysis on Wigner space,” *Probability theory and related fields*, vol. 112, no. 3, pp. 373–409, 1998.
- [82] Y. Dabrowski, “A free stochastic partial differential equation,” in *Annales de l’IHP Probabilités et statistiques*, vol. 50, pp. 1404–1455, 2014.
- [83] P. Biane, “Segal–Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems,” *Journal of Functional Analysis*, vol. 144, no. 1, pp. 232–286, 1997.
- [84] B. Collins, A. Dahlqvist, and T. Kemp, “Strong convergence of unitary Brownian motion,” *Probability Theory and Related Fields*, vol. 170, no. 1-2, pp. 49–93, 2015.
- [85] T. Lévy and M. Maïda, “Central limit theorem for the heat kernel measure on the unitary group,” *Journal of Functional Analysis*, vol. 259, no. 12, pp. 3163–3204, 2010.
- [86] G. Cébron and T. Kemp, “Fluctuations of Brownian motions on $\mathbb{G}\mathbb{L}_N$,” *arXiv:1409.5624*, 2014.
- [87] E. A. Nikitopoulos, “Noncommutative C^k functions and Fréchet derivatives of operator functions,” *arXiv:2011.03126*, 2020.
- [88] P. Biane, “Free Brownian motion, free stochastic calculus and random matrices, in free probability theory,” *Fields Institute Communications*, vol. 12, pp. 1–19, 1997.

-
- [89] F. Bekerman, A. Figalli, and A. Guionnet, “Transport maps for β -matrix models and universality,” *Communications in mathematical physics*, vol. 338, no. 2, pp. 589–619, 2015.
- [90] B. Collins and S.-G. Youn, “Additivity violation of the regularized minimum output entropy,” *arXiv:1907.07856*, 2019.
- [91] R. T. Rockafellar, *Convex analysis*, vol. 36. Princeton university press, 1970.
- [92] D. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*. No. 1, American Mathematical Society, 1992.
- [93] B. Collins and I. Nechita, “Random quantum channels I: Graphical calculus and the bell state phenomenon,” *Communications in Mathematical Physics*, vol. 297, no. 2, pp. 345–370, 2010.
- [94] J. A. Mingo and R. Speicher, *Free probability and random matrices*, vol. 35 of *Fields Institute Monographs*. Springer, 2017.
- [95] S. Yin, “Non-commutative rational functions in strong convergent random variables,” *Advances in Operator Theory*, vol. 3, no. 1, pp. 178–192, 2018.
- [96] S. A. Amitsur, “Rational identities and applications to algebra and geometry,” *Journal of Algebra*, vol. 3, no. 3, pp. 304–359, 1966.
- [97] P. M. Cohn, *Free ideal rings and localization in general rings*, vol. 3. Cambridge university press, 2006.
- [98] P. A. Linnell, “Division rings and group von Neumann algebras,” in *Forum Math*, vol. 5, pp. 561–576, 1993.
- [99] L. Erdős, T. Krüger, and Y. Nemish, “Scattering in quantum dots via noncommutative rational functions,” *arXiv:1911.05112*, 2019.
- [100] Z. Ma and F. Yang, “Sample canonical correlation coefficients of high-dimensional random vectors with finite rank correlations,” *arXiv:2102.03297*, 2021.
- [101] R. Speicher, “Regularity of non-commutative distributions and random matrices,” in *Publications of MFO, OWR-2019-56*, pp. 3513–3514, 2019.
- [102] S. Yin, “On the rational functions in non-commutative random variables,” 2020.
- [103] G. Duchamp and C. Reutenauer, “Un critere de rationalité provenant de la géométrie non commutative,” *Inventiones Mathematicae*, vol. 128, no. 3, pp. 613–622, 1997.
- [104] P. A. Linnell, “A rationality criterion for unbounded operators,” *Journal of Functional Analysis*, vol. 171, no. 1, pp. 115–123, 2000.
- [105] A. Connes, *Noncommutative geometry. Transl. from the French by Sterling Berberian*. San Diego, CA: Academic Press, 1994.

- [106] A. Miyagawa, “The estimation of non-commutative derivatives and the asymptotics for the free field in free probability theory,” Master’s thesis, Kyoto University, 2021.
- [107] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov, “Noncommutative rational functions, their difference-differential calculus and realizations,” *Multidimensional Systems and Signal Processing*, vol. 23, no. 1-2, pp. 49–77, 2012.
- [108] P. Hrubes and A. Wigderson, “Non-commutative arithmetic circuits with division,” in *Proceedings of the 5th conference on Innovations in theoretical computer science*, pp. 49–66, 2014.
- [109] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov, “Singularities of rational functions and minimal factorizations: the noncommutative and the commutative setting,” *Linear Algebra and its Applications*, vol. 430, no. 4, pp. 869–889, 2009.
- [110] J. W. Helton, T. Mai, and R. Speicher, “Applications of realizations (aka linearizations) to free probability,” *Journal of Functional Analysis*, vol. 274, no. 1, pp. 1–79, 2018.
- [111] P. M. Cohn and C. Reutenauer, “A normal form in free fields,” *Canadian Journal of Mathematics*, vol. 46, no. 3, pp. 517–531, 1994.
- [112] P. M. Cohn and C. Reutenauer, “On the construction of the free field,” *International journal of Algebra and Computation*, vol. 9, no. 3-4, pp. 307–323, 1999.
- [113] C. Vargas, “A general solution to (free) deterministic equivalents,” *Contributions of Mexican Mathematicians Abroad in Pure and Applied Mathematics*, vol. 709, p. 131, 2018.
- [114] F. J. Murray and J. von Neumann, “On rings of operators,” *Annals of Mathematics*, vol. 37, no. 1, pp. 116–229, 1936.
- [115] B. Blackadar, *Operator algebras: theory of C^* -algebras and von Neumann algebras*, vol. 122. Berlin:Springer, 2006.
- [116] M. Terp, “ l_p spaces associated with von Neumann algebras,” *Math. Institute, Copenhagen Univ.*, vol. 3, no. 4, 1981.
- [117] J. B. Conway, *A course in functional analysis*. New York etc.: Springer–Verlag, 1990.
- [118] A. Connes and D. Shlyakhtenko, “ l^2 -homology for von Neumann algebras,” *Journal für die reine und angewandte Mathematik*, vol. 2005, no. 586, pp. 125–168, 2005.
- [119] D. Voiculescu, “The analogues of entropy and of Fisher’s information measure in free probability theory, II,” *Inventiones mathematicae*, vol. 118, no. 1, pp. 411–440, 1994.
- [120] J. Volcic, “Matrix coefficient realization theory of noncommutative rational functions,” *Journal of Algebra*, vol. 499, pp. 397–437, 2018.
- [121] J. Volcic, “Hilbert’s 17th problem in free skew fields,” *arXiv:2101.02314*, 2021.
- [122] H. Bercovici and D. Voiculescu, “Free convolution of measures with unbounded support,” *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 733–773, 1993.