Local tail statistics of heavy-tailed random matrix ensembles with unitary invariance

M Kieburg and A Monteleone

School of Mathematics and Statistics, University of Melbourne, 813 Swanston Street, Parkville, Melbourne VIC 3010, Australia

E-mail: m.kieburg@unimelb.edu.au and monteleonea@student.unimelb.edu.au

Received 7 March 2021, revised 14 June 2021
Accepted for publication 22 June 2021
Published 12 July 2021

Abstract

We study heavy-tailed Hermitian random matrices that are unitarily invariant. The invariance implies that the eigenvalue and eigenvector statistics are decoupled. The motivating question has been whether a freely stable random matrix has stable eigenvalue statistics for the largest eigenvalues in the tail. We investigate this question through the use of both numerical and analytical means, the latter of which makes use of the supersymmetry method. A surprising behaviour is uncovered in that a freely stable random matrix does not necessarily yield stable statistics and if it does then it might exhibit Poisson or Poisson-like statistics. The Poisson statistics have been already observed for heavy-tailed Wigner matrices. We conclude with two conjectures on this peculiar behaviour.

Keywords: random matrices, local spectral statistics, heavy-tailed distributions, stable distributions, supersymmetry method

(Some figures may appear in colour only in the online journal)

1. Introduction

Gaussian random matrices are a well-studied topic due to the enormous number of applications in engineering, mathematics and physics. This large variety in scope has led to a multitude of modifications of random matrix models veering away from the traditional Gaussian ensembles into models with greater generality, e.g. see the textbooks [1–3]. This has resulted in various studies of spectral statistics such as bulk statistics, hard- and soft-edge statistics, as well as the statistics of multicritical points where for instance spectral supports merge (also known as cuts) or an outlier (a separate eigenvalue not belonging to the bulk of the spectrum) is absorbed into the bulk of the spectrum.

∗ Author to whom any correspondence should be addressed.
What is less well-studied is the tail statistics of eigenvalues of a random matrix that exhibits a heavy tail. Nevertheless, heavy-tailed random matrices have many important applications especially when systems are non-stationary or open. For such systems one can expect heavy-tails being more natural to occur rather than ensembles for which all moments exist. Examples admitting heavy tails being time series analysis, disordered systems, quantum field theory and more recently deep neural networks e.g. see [4–16]. From a theoretical point of view the ensembles stable under matrix addition are of particular importance since they are the fixed points of their respective domains of attraction via the multivariate central limit theorem [17]. The classification of these domains as well as the stable distributions is still poorly understood from the perspective of spectral statistics. We aim to unveil one part of this incomplete picture, namely the statistics of the largest eigenvalues for unitarily invariant ensembles.

There are several works [4, 16, 18–26, 28–31], in physics and mathematics that have studied heavy-tailed Wigner matrices in detail (meaning matrices whose entries are independently and identically distributed along a heavy-tailed univariate probability measure). For these matrices it has been shown [19, 20, 23, 27] that the largest eigenvalues in the heavy tail converge to Poisson statistics. Moreover the eigenvectors become localised [9, 18, 27] while those in the bulk become delocalised [29]. One may argue that the Poisson statistics of the eigenvalues is due to the localisation of the eigenvectors. This perspective is supported by the fact that the distribution of these largest eigenvalues are shared by the distribution of the largest matrix entries [23]. In the present work, we will argue that this is not necessarily the case and that the Poisson statistics or at least a very diminished level repulsion can also be found for unitarily invariant random matrix ensembles. Such ensembles have been discussed in [8, 11, 21, 32–37]. The unitary invariance takes the eigenvectors out of the picture as they are still Haar distributed and thus delocalised. In [38], it was also shown that exactly such matrices maximise the Shannon entropy when only a level density is given as an input.

In our present work we consider two specific random matrix ensembles. The first one is about the singular values of a product of complex inverse Ginibre matrices, see [39–42] for an analytical computation of the finite $N$ (matrix dimension) statistics as well as the hard edge statistics. This ensemble is not stable for finite matrix dimension, it is known [43, 44] that the limiting macroscopic level density (see [2] for the definition of the macroscopic statistics) is stable under free convolution (sum of identical and independent copies of the random matrix in the large $N$ limit). Numerically we have confirmed that this asymptotic stability also holds for the local statistics at the soft-edge and some part of the bulk. Indeed in reference [42], those local spectral statistics have been proven for this kind of random matrix. However the tail statistics do not share this behaviour. In the tail the eigenvalues degenerate statistically, meaning clusters of eigenvalues show a diminished level repulsion which is vanishing completely for $N \to \infty$. This number of eigenvalues inside such a cluster is equal to the number of copies of the random matrix that has been added. Through using the supersymmetry method we have analytically confirmed our numerical observations. Namely inside the tail the sum of random matrices agrees with the direct sum of exactly the same matrices. This agreement is however not true in the bulk or at the soft-edge. Therefore there is a transition. The scaling of the eigenvalues that belong to this critical regime of the transition has been identified and exhibits a dependence on the stability exponent.

Furthermore, we will address the question of the central limit theorem for the tail statistics of the sum of these random matrices, which happens to converge to a Poisson point process for the largest eigenvalues as it has already been known for heavy-tailed Wigner matrices [19, 20, 23, 27]. To confirm whether this picture is true more generally, we have consider a second random matrix ensemble which is a Gaussian unitary ensemble (GUE) whose variance is averaged over a stable one-sided distribution. Similar averages have also been discussed in [6–8, 33, 34]. This
construction yields a random matrix ensemble that is already stable at fixed matrix dimension \( N \) so that its macroscopic as well as its microscopic statistics must be stable, see \([2]\) for the definition of the microscopic statistics. There is however the downside that two copies of the matrix are not free in the sense of free probability \([45]\). Whether their largest eigenvalues will go exactly to the Poisson statistics will depend on whether or not the largest eigenvalues live on a scale that is bigger than the one of the bulk. This will be shown with a supersymmetry calculation. The Monte Carlo simulations we have generated suggest that the average position of the largest eigenvalues might saturate at a finite value.

The present work is built up as follows. In section 2, we describe the numerical experiment we have carried out for a sum of inverse complex Ginibre matrices in order to get a feeling what is happening. We do not only confirm that the macroscopic level density as well as the soft edge statistics are stable but show how the tail statistics change by adding several independent copies of these random matrices. To get an analytical confirmation, we compute the average of the ratio of characteristic polynomials in section 3. Therein, we consider a more general situation of the sum of products of inverse Ginibre matrices as those are also known to be stable under free convolution \([43, 44]\). Those averages encode the whole eigenvalue statistics and are computed via the supersymmetry method \([46–48]\). In section 4, we study the limit of the sum of an infinite number of matrices analytically as well as numerically. In particular we investigate the question regarding how universal are the Poisson statistics for the largest eigenvalues of heavy-tailed ensembles and what are the scales to finding them. We conclude in section 5 by formulating two conjectures for heavy-tailed ensembles.

### 2. Numerical observations

We begin this section with a short numerical experiment (in fact the same one that led to the discovery of this surprising result) as it will give us some insight into what is going on when we add random matrices with heavy tailed macroscopic level densities.

We consider a sum of \( L \) identically and independently distributed inverse complex Ginibre matrices \( X_j \in \mathbb{C}^{N \times N} \) with the probability density

\[
P(X_j) = \frac{1}{\pi N^2} \frac{\exp[-\text{tr}(X_j^\dagger X_j)^{-1}]}{\det(X_j^\dagger X_j)^{2N}}.
\]

Certainly, the inverse \( X_j^{-1} \) is a complex Ginibre matrix \([49]\) and the product \( (X_j^\dagger X_j)^{-1} \) describes a complex (Wishart) Laguerre matrix \([1, 50]\). Thus, the full spectral statistics of a single matrix \((L = 1)\) is completely known, see \([1, 2]\).

For instance, the macroscopic level density of \((X_j^\dagger X_j)^{-1}\) is the Marcenko–Pastur law \([51]\),

\[
\rho_{X_j^\dagger X_j^{-1}}(\lambda) = \lim_{N \to \infty} \frac{1}{N} \text{tr} \left( \lambda I_N - \frac{1}{N}(X_j^\dagger X_j)^{-1} \right) = \frac{\sqrt{4-\lambda}}{2\pi\sqrt{\lambda}} \Theta[\lambda(4-\lambda)]
\]

with the Heaviside step function \( \Theta \). This implies that the macroscopic level density of \( X_j^\dagger X_j \) is equal to

\[
\rho_{X_j^\dagger X_j}(\lambda) = \lim_{N \to \infty} \frac{1}{N} \text{tr} \left( \lambda I_N - NX_j^\dagger X_j \right) = \frac{\sqrt{4-\lambda^{-1}}}{2\pi\sqrt{\lambda^{3/2}}} \Theta[\lambda(4-\lambda^{-1})].
\]

Comparing this result with the definition of the stability exponent \( \alpha \), see \([43, \text{appendix A}]\) for the definition of the stability exponent \( \alpha \), which shows in the asymptotic approximation \( \rho(\lambda) \propto |\lambda|^{-1-\alpha} \) for large \( |\lambda| \gg 1 \), the current ensemble corresponds to \( \alpha = 1/2 \).
Figure 1. The macroscopic level density of the random matrix sum (4) has been simulated by Monte Carlo simulations (coloured symbols) for $L$ numbers of matrices added together. This is compared to the analytical result (3) (black solid curve) which should agree for any $L$ via free probability. We generated for each setting $10^6$ configurations of dimension $N = 200$. The bin size is equal to 0.1. Thus, the statistical and systematic error is below one percent.

From free probability [43, 44], we know that this distribution is stable under free convolution, meaning when we add two or more copies of the matrix $X_j^\dagger X_j$, i.e.

$$Y_L = \frac{1}{L^{1/\alpha}} \sum_{j=1}^L X_j^\dagger X_j, \quad (4)$$

the resulting matrix $Y_L$ shares the same macroscopic level density as each single $X_j^\dagger X_j$. This can be readily checked via the $\mathcal{R}$-transform which is implicitly defined with the help of the Green function [45]

$$G(z) = \int_{-\infty}^{\infty} \rho(\lambda) d\lambda \frac{1}{z - \lambda}, \quad (5)$$

namely

$$\mathcal{R}[G(z)] = z - \frac{1}{G(z)} \quad (6)$$

The $\mathcal{R}$-transform of $X_j^\dagger X_j$ is [44]

$$\mathcal{R}_{X_j^\dagger X_j}(y) = -e^{i\pi/2} y^{-1/2}. \quad (7)$$

Due to the rule for a sum of two asymptotically free random matrices $A$ and $B$, 

$$\mathcal{R}_{A+B}(y) = \mathcal{R}_A(y) + \mathcal{R}_B(y), \quad (8)$$
Figure 2. Microscopic level density at the soft edge for the Monte Carlo simulations (coloured symbols) of the random matrix sums (4) and the analytical prediction (11) (solid black curve). We have employed the same configurations generated for figure 1. The bin size is this time 0.2.

and the scaling rule of a random matrix $A$ with a scalar $\mu$,

$$R_{\mu A}(y) = \mu R_A(\mu y),$$  

we have

$$R_{YL}(y) = -e^{y/2}y^{-1/2},$$

too. We have numerically illustrated this for $L = 1, 2, 3, 4$ in figure 1.

Other statistics that can be checked to be stable when performing the sum (4) are those in the bulk and at the soft-edge. In reference [42], those have been proven to be those shared with the GUE. For instance, the soft edge lies at $\lambda_{\text{min}} = 1/4$ in the macroscopic scaling for any $L \in \mathbb{N}$. Thus, we should find the microscopic level density $[1, 2]$

$$\rho_{\text{Airy}}(\lambda) = \lim_{N \to \infty} \left( \text{tr} \left( \lambda 1_N - \frac{N^{2/3}}{2^{1/3}} \left[ 1_N - \frac{4N}{L^2} \sum_{j=1}^{L} X_j^\dagger X_j \right] \right) \right)$$

$$= \left[ \text{Ai}'(\lambda) \right]^2 - \lambda \left[ \text{Ai}(\lambda) \right]^2$$  

(11)

with the Airy function $\text{Ai}(x)$ and its derivative $\text{Ai}'(x)$. This stability has been corroborated in figure 2 with the help of Monte Carlo simulations.

The question is what are the eigenvalue statistics in the tail? For $L = 1$ it is rather trivial because $(X_1^\dagger X_1)^{-1}$ is an ordinary Laguerre ensemble that exhibits a hard edge with the microscopic level density $[1, 2]$ of the Bessel kernel,

$$\rho_{\text{Bessel}}(\lambda) = \lim_{N \to \infty} \left( \text{tr} \left( \lambda 1_N - \frac{2\sqrt{N}}{\pi} X_1^\dagger X_1 \right) \right)^{1/2} = \pi^2 \lambda [J_0'(\pi \lambda) + J_1^2(\pi \lambda)].$$  

(12)
Figure 3. The microscopic level density of the Monte Carlo simulations of figures 1 and 2 (coloured histograms) compared to the inverse Bessel statistic result (13) (black solid curve). The bin size is this time 0.2. We underline that we have employed the unfolded scale meaning that we have inverted the spectrum so that the largest eigenvalues are those closest to the origin.

We have used the Bessel function of the first kind $J_\nu(x)$. We also note that we have unfolded in such a way that the asymptotic of the level density for large $\lambda$ is $\rho_{\text{Bessel}}(\lambda) \to 1$, meaning the mean level spacing is approximately one.

From (12) we can read off the microscopic level density in the tail which we ‘baptise’ the inverse Bessel statistics,

$$\rho_{\text{invB}}(\lambda) = \frac{1}{\lambda^2} \rho_{\text{Bessel}}(\lambda^{-1}) = \frac{\pi^2}{2\lambda^3} [J_0^2(\pi \lambda^{-1}) + J_1^2(\pi \lambda^{-1})].$$

(13)

We note that the tail of the largest eigenvalues is decaying slightly stronger than the macroscopic level density, namely $\rho_{\text{invB}}(\lambda) \propto \lambda^{-3}$ in contrast to $\rho_{X^\dagger X}(\lambda) \propto \lambda^{-3/2}$. One reason is that the unfolding involves taking the square root of the eigenvalues. Yet, this does not completely explain everything. The largest eigenvalues are far in the tail and the different tail behaviours lie on a different scale than the heavy-tail of the macroscopic level density. This is known already for the hard edge where the macroscopic level density can have a totally different behaviour (for instance it goes to zero instead of diverging as a square root singularity) than the microscopic ones due to the very different scales.

We would like investigate whether the tail statistics are also stable and particularly if equation (13) still holds for the sum (4) for any $L \in \mathbb{N}$. Unfortunately and quite surprisingly this is not the case. Numerically we have observed that the microscopic level density in the tail follows the law

$$\rho_{\text{invB}}^{(T)}(\lambda) = L^2 \rho_{\text{invB}}(L\lambda) = \frac{1}{\lambda^2} \rho_{\text{Bessel}}(L^2 \lambda^{-1}) = \frac{\pi^2}{2L^3 \lambda^3} [J_0^2((\pi L \lambda)^{-1}) + J_1^2((\pi L \lambda)^{-1})].$$

(14)

see figure 3 and left plots of figure 4. Note that this is not a simple rescaling for the level density as then we would have multiplied the density by $L$ and not $L^2$ resulting from the Jacobian.
Therefore there has to be a different and perhaps more fundamental change in the spectral statistics. What is actually going on in the tail? To analyse this first numerically we have measured the distributions of the four largest eigenvalues and the three level spacing distributions between
these four eigenvalues of $Y_L$ for $L = 1, 2, 3, 4$, see figure 4. Each single maximum of the microscopic Bessel level density (12) and thus the inverse Bessel result (14) is now described by $L$ eigenvalues and not just a single one. The level spacing distribution indeed corroborates that the eigenvalues corresponding to one maximum have a diminished level repulsion, see to the right plots of figure 4.

Our interpretation is that in the large $N$-limit of $Y_L$ the eigenvalue tail statistics asymptotes to the statistics of the direct sum

$$
\hat{Y}_L = \bigoplus_{j=1}^L X_j^\dagger X_j = \begin{bmatrix} X_1^\dagger X_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & X_{L}^\dagger X_L \end{bmatrix}.
$$

We have also simulated these random matrices and measured the same level spacing distributions of the largest eigenvalues for comparison. In the right plots of figure 4 we see indeed similarities such that our point is substantiated. The deviations can be understood as residual level repulsions which are suppressed onto very small scales. Therefore we expect a convergence to the statistics of (15) though it will not be uniform about very small spacings. For reference we have added the level spacing distribution for the Poisson statistics (statistically independent eigenvalues).

$$
p_{\text{Poisson}}(s) = e^{-s}
$$

and the Wigner surmise for the GUE

$$
p_{\text{Poisson}}(s) = \frac{32}{\pi^2} s^2 \exp \left[-\frac{4}{\pi} s^2 \right],
$$

which is pretty close to the universal bulk distribution [52] as well as to the level spacing distributions at the hard edge [53].

In section 3 we have derived an analytical argument based on the supersymmetry method that verifies our understanding. The question that remains is if this observation is true in general and whether it is perhaps governed by a more universal behaviour. This will be discussed later, see section 5.

### 3. Analytical corroborations

In section 3.1 we define and discuss the random matrix ensembles namely the sum of products of inverse Wishart–Laguerre matrices that we study with the supersymmetry method in section 3.2. After we have derived the corresponding supermatrix integral we carry out the large $N$-limit in section 3.3. Finally in section 3.4 we look for the critical scale where the spectral statistics of this ensemble changes from stable (meaning the sum exhibits the same statistics as each matrix in the sum) to unstable.

#### 3.1. Matrix model and the scaling of the tail statistics

After we have seen a glimpse of the peculiarities of the spectral statistics in heavy tails we will show in this section that first the proper interpretation is indeed the one we have concluded with in the last section and secondly that it works for more general ensembles. For this reason we consider a sum of product matrices of the form

$$
X^{(M)}_l = X_{l,M} X_{l,M-1} \cdots X_{l,1}, \quad \text{for } l = 1, \ldots, L,
$$
where each $X_{i,m} \in (\mathbb{C}^{N \times N}$ is independently drawn from (1). Then the sum is given by

$$Y_{1M}^{(M)} = \frac{1}{L} \sum_{i=1}^{L} (X_{1}^{(M)})_{i} (X_{1}^{(M)}).$$

(19)

Comparison with (4) shows that the stability exponent will be $\alpha = 1/(M + 1)$. Certainly it has been shown that the macroscopic level density of $Y_{1M}^{(M)}$,

$$\rho_{Y_{1M}^{(M)}}(\lambda) = \lim_{N \to \infty} \left( \frac{1}{N} \text{tr} \left( \lambda I_{N} - N^{M} Y_{1M}^{(M)} \right) \right),$$

(20)

is stable under free convolution, too, see [43, 44]. The reason is the $S$-transform, another transform in free probability indirectly defined via the $\mathcal{R}$-transform [45]

$$\mathcal{R}(y) = \frac{1}{S(y \mathcal{R}(y))} \Leftrightarrow S(\chi) = \frac{1}{\mathcal{R}[\chi \mathcal{S}(\chi)]}.$$ 

(21)

It has the property [45]

$$S_{1M}^{(M)}(\chi) = S_{1M}^{(M-1)} X_{1M}(\chi) = S_{1M}^{(M-1)}(\chi) S_{1M}^{(M)} X_{1M}(\chi)$$

(22)

as $X_{1M}$ and $Y_{1M}^{(M)}$ are asymptotically free. In our case we have

$$S_{X_{1M}}(\chi) = -\chi \Rightarrow S_{Y_{1M}^{(M)}}(\chi) = (-\chi)^{M} \Rightarrow \mathcal{R}_{Y_{1M}^{(M)}}(y) = -e^{\frac{y}{M+1} - \frac{y}{M+1}}.$$ 

(23)

This result reflects the stability when considering the classification in [43, appendix A].

The macroscopic level density of the inverse of $Y_{1M}^{(M)}$ is given in terms of the Meijer G-function [55]

$$\rho_{Y_{1M}^{-1}(M)}(\lambda) = \frac{1}{\sqrt{2\pi}} \frac{M^{M-3/2}}{(M+1)^{M+1/2}} G_{M,0}^{M,0} \left( \frac{(1 + M - j)/M}{(j - M)/(M + 1)} \right)_{j=1,...,M} \frac{M^{M}}{(M + 1)^{M+1}} \lambda.$$ 

(24)

The Meijer G-function [54] is essentially an inverse Mellin transform of ratios of gamma functions, i.e.

$$\frac{G_{m,n}^{p,q} \left( a_{1}, \ldots, a_{p}; b_{1}, \ldots, b_{q} \mid c_{1}, \ldots, c_{n}; d_{1}, \ldots, d_{m} \right) z^{\lambda}}{2\pi i}$$

(25)

where the contour $\mathcal{C}$ starts at $-\infty$ and finishes at $+\infty$ while having the poles of $\Gamma[c_{j} + s]$ on the left side of the path and $\Gamma[1 - a_{j} - s]$ are on the right side. The distribution (24) is called the Fuss–Catalan distribution since it has the Fuss–Catalan numbers [55],

$$\text{FC}_{M}(n) = \frac{\Gamma[(M + 1)n - M]}{\Gamma[Mn - M + 2\Gamma[n]]} \text{ with } n \in \mathbb{N}_{0},$$

(26)

as its moments and it has a support on $\lambda \in [0, (M + 1)^{M+1}/M^{M}]$. Additionally, its behaviour at the origin diverges like

$$\rho_{Y_{1M}^{-1}(M)}(\lambda) \approx \frac{\lambda^{-M/(M+1)}}{\Gamma(M + 2)/(M + 1)\Gamma[(M + 1)/(M + 1)]} \text{ for } \lambda \ll 1,$$

(27)
this implies that the tail behaviour of the matrix $Y_1^{(M)}$ will be

$$
\rho_{1}^{(M)}(\lambda) = \lambda^{-2} \rho_{1}^{(M-1)}(\lambda^{-1})
\approx \frac{\lambda^{M/(1+M)-2}}{\Gamma[(M+2)/(M+1)]^2 \Gamma[M/(M+1)]}
$$

for $\lambda \gg 1$ \hspace{1cm} (28)

, where we can again read off the stability exponent $\alpha = 1/(M+1)$ which is consistent with the other discussion.

As a result from the above discussion, the scaling of the largest eigenvalues of $Y_1^{(M)}$ as well as $Y_{L}^{(M)}$ will be $N$ regardless of $M$. This scaling can be obtained by combining $N\rho_{1}^{(M)}(\lambda)d\lambda \propto d(N\lambda^{-1/(M+1)})$ for $\lambda \gg 1$ and equation (20). We need this scale to properly unfold the spectrum as well as to find the largest eigenvalues. For instance, the unfolded microscopic tail level density of $Y_1^{(M)}$ is given by the so-called Meijer G-kernel result [56, theorem 5.3] ($\nu_j = 0$ for all $j$) of the hard edge microscopic level density of $Y_1^{(M)}$ which is

$$
\rho_{\text{Meijer } G}^{(M)}(\lambda) = \lim_{N \to \infty} \left( \frac{\text{tr} \delta(\lambda I_N - \frac{N^{1/(M+1)}}{c_M} Y_1^{(M)} - 1/(M+1))}{\Gamma[(M+2)/(M+1)]^2 \Gamma[M/(M+1)]} \right)
\approx \frac{\lambda^{M/(1+M)-2}}{\Gamma[(M+2)/(M+1)]^2 \Gamma[M/(M+1)]}
$$

for $\lambda \gg 1$. For $M = 2$, this formula reduces to (12). Due to the proper unfolding the spectrum becomes the half sided picket fence spectrum for $M \to \infty$ [57–60]

$$
\lim_{M \to \infty} \rho_{\text{Meijer } G}^{(M)}(\lambda) = \sum_{j=1}^{\infty} \delta(\lambda - j + 0.5).
$$

The shift by 0.5 reflects the level repulsion from the origin and has thus a strong resemblance to the spectrum of the quantum harmonic oscillator.

The microscopic tail level density is given then by

$$
\rho_{\text{inv-MG}}^{(M)}(\lambda) = \frac{1}{\lambda^2} \rho_{\text{Meijer } G}^{(M)}(\lambda^{-1}).
$$

This is the one that can be expected when studying the matrix $Y_1^{(M)}$. For the sum of $L$ copies of matrices, meaning $Y_{L}^{(M)}$, we will find that the averaged spectrum behaves as if we would have directly summed these random matrices, cf. subsection 3.3.
3.2. Supersymmetry method

Instead of computing the level density or more generally the \( k \)-point correlation functions of \( Y_L^{(M)} \) we consider the partition function

\[
Z_{Y_L^{(M)}}^{(k)}(\kappa) = \left\langle \prod_{j=1}^{k} \frac{\det(Y_L^{(M)} - \kappa_{F,j})}{\det(Y_L^{(M)} - \kappa_{B,j})} \right\rangle
\]

with \( \kappa_{F,j}, \kappa_{B,j} \in \mathbb{C} \) and \( \kappa_{B,j} \notin [0, +\infty) \) for all \( j = 1, \ldots, k \). It is a well-known fact these kinds of partition functions can generate several different quantities such as the \( k \)-point correlation functions and specifically the level density, see references [1, 46–48]. The variables \( \kappa_{F,j} \) and \( \kappa_{B,j} \) usually contain the spectral variables that correspond to the eigenvalue statistics, regularisations such as an imaginary shift away from the positive real axis, as well as some source variables that can be expanded and can create Green’s functions. When arranging the \( \kappa_{F,j} \) and \( \kappa_{B,j} \) in the form of a diagonal supermatrix \( \kappa = \text{diag}(\kappa_{B,1}, \ldots, \kappa_{B,k}; \kappa_{F,1}, \ldots, \kappa_{F,k}) \), we can write the average in terms of a superdeterminant

\[
Z_{Y_L^{(M)}}^{(k)}(\kappa) = \left\langle \text{Sdet}(Y_L^{(M)} \otimes I_{k} - I_{N} \otimes \kappa)^{-1} \right\rangle.
\]

For an introduction to superalgebra and superanalysis we refer to [61]. As the conventions slightly vary in the literature, we briefly summarise ours in the present work. A \( (p_1 | q_1) \times (p_2 | q_2) \) supermatrix \( A \) can be arranged into four blocks

\[
A = \begin{bmatrix} A_{BB} & A_{BF} \\ A_{FB} & A_{FF} \end{bmatrix}
\]

with the Boson–Boson \( (A_{BB} : p_1 \times p_2) \) and the Fermion–Fermion \( (A_{FF} : q_1 \times q_2) \) blocks containing only commuting variables and the Boson–Fermion \( (A_{BF} : p_1 \times q_2) \) and the Fermion–Boson \( (A_{FB} : q_1 \times p_2) \) comprising of only anti-commuting variables. Then for \( p_1 = p_2 = p \) and \( q_1 = q_2 = q \), the supertrace and the superdeterminant are given by

\[
\text{Str} A = \text{tr} A_{BB} - \text{tr} A_{FF} \quad \text{and} \quad \text{Sdet} A = \frac{\text{det}(A_{BB} - A_{BF}A_{FF}^{-1}A_{FB})}{\text{det} A_{FF}},
\]

respectively. Each commuting variable consists of a numerical part, which can be real or complex, and a nilpotent one which is an even polynomial of the underlying Grassmann variables \( \{\eta_j\} \) (algebraic basis of the anti-commuting variables), while an anti-commuting variable only contains a nilpotent part which is an odd polynomial of the Grassmann variables. The integral over a Grassmann variable is defined by the two axiomatic identities

\[
\int \eta_j \, d\eta = 0 \quad \text{and} \quad \int \eta_j \, d\eta = 1.
\]

These two equalities are enough to fix the integration over Grassmann variables as any function of Grassmann variables is understood as a finite Taylor series due to the nilpotence of the Grassmann variables, i.e. \( \eta_j^2 = 0 \).

Coming back to (34), we can readily generalise the partition function to an arbitrary \( (k|k) \times (k|k) \) supermatrix \( \kappa \) as long as the numerical part of the eigenvalues of the Boson–Boson block \( \kappa_{BB} \) does not lie on the positive real line or the origin.

The idea of the supersymmetry method in random matrix theory is to map the average over an ordinary random matrix to an average over a supermatrix whose dimension is independent
of $N$. In our case we essentially have products of matrices of the form $WW^\dagger$. In references [62, 63], one of the present authors has introduced a short cut called the supersymmetric projection formula that essentially states for a random matrix $W \in \mathbb{C}^{N \times N}$ that is distributed along a unitarily invariant density $P(WW^\dagger) = P(VV^\dagger)$ (for all $V \in U(N)$ and $W \in \mathbb{C}^{N \times N}$), we can find a superfunction $Q(U)$ for a $(k|k) \times (k|k)$ supermatrix $U$ such that

$$\langle S\text{det} (WW^\dagger \otimes 1_{k|k} + \tilde{\kappa})^{-1} \rangle = \langle S\text{det} (1_N \otimes U + \tilde{\kappa})^{-1} \rangle .$$

(38)

Here, $\tilde{\kappa}$ can be a much larger supermatrix of dimensions $(kN|kN) \times (kN|kN)$. We underline that on the left-hand side we average over the ordinary random matrix $W$ while on the right-hand side we average over the supermatrix $U$.

Let $\text{Herm}(k)$ be the set of $k \times k$ Hermitian matrices. The supermatrix $U$ is in the current situation relatively simple, namely its Boson–Boson block is a positive definite Hermitian matrix $U_{BB} = U_{BB}^\dagger \in \text{Herm}(k)$ with no Grassmann variables and its Fermion–Fermion block is a unitary matrix $U_{FF} \in U(k)$ also containing no Grassmann variables. The Boson–Fermion and Fermion–Boson blocks only comprise independent Grassmann variables with no further symmetries. Therefore, the supermatrix space described by $U$ is the supersymmetric coset $\text{Herm}_{\nu}(k) = \text{Gl}(k|k)/U(k|k)$, see [64]. The superfunction $Q(U)$ is given via the supersymmetric projection formula [62, 63] times the measure $S\text{det} U^N d[U]$ on $\text{Herm}_{\nu}(k|k)$,

$$Q(U) = \int_{\mathbb{C}^{N \times N}} d[W] \int_{\mathbb{C}^{N \times k|k}} d[\tilde{W}] \tilde{P} \left( \begin{array}{cc} WW^\dagger + \tilde{W}\tilde{W}^\dagger & \tilde{W} \\ U\tilde{W} & U \end{array} \right) ,$$

(39)

where $d[U]$ is the product of the differentials of all supermatrix elements. The supermatrix $\tilde{W}$ is a rectangular matrix where its first $k$ columns are ordinary $N$ dimensional complex vectors and the last $k$ columns are $N$ dimensional vectors with independent complex Grassmann variables as their entries, this set has been denoted by $\mathbb{C}^{N \times (k|k)}$. The superfunction $\tilde{P}$ is a supersymmetric extension of $P$ that satisfies

$$P(VV^\dagger) = \tilde{P}(V^\dagger V), \quad \text{for any } V \in \mathbb{C}^{N \times (N+k|k)} .$$

(40)

Such a supersymmetric extension is commonly not unique but there are usually natural choices as we will see below for the present situation of the inverse Ginibre matrices.

We approach the average (34) inductively by writing $Y_L^{(M)}$ as a sum and a product of inverse Ginibre matrices,

$$Z_{Y,L,M}^{(k|N)}(\kappa) = \left\langle S\text{det} \left( X_L^{(M)}X_L^{(M)} \otimes 1_{k|k} + L^{M+1}Y_{L-1}^{(M)} \otimes 1_{k|k} - L^{M+1}1_N \otimes \kappa \right)^{-1} \right\rangle$$

$$= \left\langle S\text{det} \left( X_L^I X_L^I \otimes 1_{k|k} + \tilde{\kappa}_{LM} \right)^{-1} \right\rangle$$

(41)

with

$$\tilde{\kappa}_{LM} = L^{M+1} \left( X_L^{(M-1)} \right)^{-1} Y_{L-1}^{(M)} X_L^{(M-1)} \otimes 1_{k|k} - \left( X_L^{(M-1)} \right)^{-1} \otimes \kappa .$$

(42)

For the rearrangement of the matrices inside the superdeterminant we have employed the identities

$$\text{Sdet} (AB) = \text{Sdet} (A) \text{Sdet} (B) \quad \text{and} \quad \text{Sdet} (H \otimes 1_{k|k}) = 1$$

(43)

for any two square supermatrices $A$ and $B$ and any ordinary square matrix $H$. 

12
Next, we apply (38) to (41) and obtain
\[
Z_{\gamma_L}^{(k,N)}(\kappa) = \left\langle \text{Sdet} \left( I_N \otimes U_{LM} + \tilde{\kappa} \right)^{-1} \right\rangle = \left\langle \text{Sdet} \left( X_L^{(M-1)} \otimes U_{LM} + L^{M+1} \left[ Y_{L-1}^{(M)} \otimes I_{\kappa k} - I_N \otimes \kappa \right] \right)^{-1} \right\rangle.
\]
We repeat this procedure until each integration over the random matrix $X_{L,j}$ is transferred into an integration over a supermatrix $U_{LM}$. The only thing that changes is the supermatrix
\[
\tilde{\kappa}_{L,j} = L^{M+1} \left[ (X_L^{(j-1)})^{-1} Y_{L-1}^{(M)} (X_L^{(j-1)})^{-1} \otimes (U_{L,j+1} \ldots U_{LM})^{-1} \right.
\]
\[
\left. - (X_L^{(j-1)})^{-1} \otimes \kappa (U_{L,j+1} \ldots U_{LM})^{-1} \right].
\]
This eventually leads to
\[
Z_{\gamma_L}^{(k,N)}(\kappa) = \left\langle \text{Sdet} \left( L^{-M-1} I_N \otimes U_{LM}^{(M)} + Y_{L-1}^{(M)} \otimes I_{\kappa k} - I_N \otimes \kappa \right)^{-1} \right\rangle
\]
with
\[
U_{LM}^{(M)} = U_{LM}^{(M-1)} U_{LM}. \tag{47}
\]
The average over $Y_{L}^{(M)}$ looks now like the one over $Y_{L}^{(M)}$. Therefore, we can proceed as before and find eventually
\[
Z_{\gamma_L}^{(k,N)}(\kappa) = \left\langle \text{Sdet} \left( V_L^{(M)} - \kappa \right)^{-N} \right\rangle
\]
with
\[
V_L^{(M)} = \frac{1}{L^{M+1}} \sum_{j=1}^{L} U^{(M)}_j \quad \text{and} \quad U_{LM}^{(M)} = U_{LM}^{(M-1)} U_{LM} \tag{49}
\]
in analogy to (18) and (19).

The remaining ingredient to be calculated is the superfunction $Q$ for the inverse Ginibre ensemble. For this purpose, it helps to know the corresponding superfunction for the Ginibre ensemble that has been computed with the projection formula in [63], which follows from the Gaussian structure
\[
\exp[-\text{tr} \; V V^\dagger] = \exp[- \text{Str} \; V^\dagger V], \quad \text{for any} \; V \in \mathbb{C}^{N \times (N+k)}.
\]

thence,
\[
Q_{\text{Gin}}(U_{\text{Gin}}) = \int_{\mathbb{C}^{N \times N}} d[W] \int_{\mathbb{C}^{N \times (N+k)}} d[\tilde{W}] \exp \left( -\text{Str} \begin{bmatrix} WW^\dagger + \tilde{W} \tilde{W}^\dagger & \tilde{W} \\ U_{\text{Gin}} \tilde{W}^\dagger & U_{\text{Gin}} \end{bmatrix} \right)
\]
\[
\propto \exp[- \text{Str} \; U_{\text{Gin}}]. \tag{51}
\]
We can make use of it by noticing that the averages (38) between Ginibre and inverse Ginibre
are related,
\[
\left\langle \text{Sdet} \left( WW^\dagger \otimes 1_{k|k} + \hat{\kappa} \right)^{-1} \right\rangle = \left\langle \text{Sdet} \left( \left( WW^\dagger \right)^{-1} \otimes 1_{k|k} + \hat{\kappa}^{-1} \right)^{-1} \text{Sdet} \hat{\kappa}^{-1} \right\rangle \\
= \left\langle \text{Sdet} \left( I_N \otimes U_{\text{Gin}} + \hat{\kappa}^{-1} \right)^{-1} \right\rangle \\
= \left\langle \left( \text{Sdet} U_{\text{Gin}} \right)^{-N} \text{Sdet} \left( I_N \otimes U_{\text{Gin}} + \hat{\kappa}^{-1} \right)^{-1} \right\rangle.
\] (52)

Simple comparison with the general duality (38) yields the identification that each supermatrices \( U_{lm} \) in (48) is a copy of \( U_{\text{Gin}}^{-1} \) which we coin \( V_{lm} \).

Summarising, the partition function takes the form
\[
Z^{(N)}_{\text{LM}}(\kappa) = \frac{\int_{\text{Herm}_{k|l}} \text{Sdet}(V_L^{(M)} - \kappa)^{-N} \prod_{m=1}^M \prod_{i=1}^L e^{-\text{Str} V_{lm}} d[V_{lm}]}{\left( \int_{\text{Herm}_{k|l}} \text{Sdet} V^{(M)} e^{-\text{Str} V} d[V] \right)^{LM}}
\] (53)
with
\[
V_L^{(M)} = \frac{1}{LM+1} \sum_{j=1}^L V_{ji}^{-1} \ldots V_{jM}^{-1}.
\]

The normalisation \( Z^{(N)}_{\text{LM}}(z) = 1 \) with an arbitrary complex \( z \in [0, +\infty] \) follows from the Wegner integration theorems [65–71] for supergroup invariant integrands. These theorems are essentially multidimensional Cauchy-like identities which tell us that the integral is essentially the integrand at \( V = I_{k|k} \) times an integrand independent constant. The proportionality constant cancels in the invariants like the supertrace or the superdeterminant as can be readily checked with (36). Therefore, the denominator in (53) is only an \( N \) independent constant, i.e. it is the power of the integral
\[
\int_{\text{Herm}_{k|l}} \left( \text{Sdet} V \right)^N e^{-\text{Str} V} d[V]
\]
\[
= \int_{\text{Herm}_{k|l}} \left( \frac{\det(V_{BB} - V_{FF}V_{EF}^{-1}V_{BF})}{\det V_{FF}} \right)^N e^{-tr V_{BB} + tr V_{FF}} d[V]
\]
\[
= \int_{\text{Herm}_{k|l}} (\det V_{BB})^{N-k} e^{-tr V_{BB}} d[V_{BB}]
\]
\[
\times \int_{U_{l|k}} (\det V_{FF})^{-N+k} e^{tr V_{FF}} d[V_{FF}]
\]
\[
\times \int_{C(k|0) \otimes (0|k)} \det(I_k - V_{BF}V_{FB})^{N} d[V_{BF}, V_{FB}].
\] (55)

The latter equality can be found by substituting \( V_{BF} \rightarrow V_{BB}V_{BF}V_{FF} \). The first two integrals are given by the Selberg integrals [1, 2]
\[
\int_{\text{Herm}_{k|l}} (\det V_{BB})^{N-k} e^{-tr V_{BB}} d[V_{BB}]
= \frac{1}{k!} \left( \prod_{j=0}^{k-1} \frac{\pi^j}{j!} \right) \int_{\mathbb{R}_+^k} \det(x)^{N-k} e^{-tr x \Delta^2(x)} d[x]
= \prod_{j=0}^{k-1} \frac{\pi^j (N-j-1)!}{j!}
\] (56)
and

\[
\int_{U(k)} (\det V_{BF})^{-N-k} e^{i\Delta^2_k} d[V_{BF}] = \frac{1}{k!} \left( \prod_{j=0}^{k-1} \frac{\pi^j}{j!} \right) \int_{[0,2\pi]} \det (e^{i\varphi})^{-N-k} e^{i\Delta^2_k} d[e^{i\varphi}] \\
= (2\pi i)^k \prod_{j=0}^{k-1} \frac{\pi^j}{(N+j)!} 
\]  

(57)

where we have first diagonalised the matrices and then integrated over their eigenvalues. To compute the remaining integral over the Grassmann variables, we employ that the Gaussian of Grassmann variables is equal to

\[
\int C^k(0|0) \times (0|k) \exp(-\text{tr} V_{BF} V_{FB}) d[V_{BF}, V_{FB}] = (2\pi i)^k \prod_{a,b=1}^N \frac{(N+\lambda^2_k)}{(N-\lambda^2_k)}! 
\]  

(58)

Additionally we exploit the superbosonisation formula [72–74] which tells us how to replace the product \( V_{BF} V_{FB} \) by a Haar distributed unitary matrix \( U \in U(k) \) at the cost of an additional factor of \( \det U^{-k} \). In particular we compute

\[
\int_{C^k(0|0)} \det (I_k - V_{BF} V_{FB})^N d[V_{BF}, V_{FB}] \\
= \int_{C^k(0|0)} \det (I_k - U) \det U^{-k} d\mu(U) \\
= \prod_{j=0}^{k-1} \frac{(N+j)!}{(N-j-1)!} 
\]  

(59)

The introduction of the denominator guarantees that we do not get any additional constants from the superbosonisation formula and the diagonalisation of the unitary matrix \( U \) as both steps are independent of the integrand. Therefore, the result (53) simplifies to

\[
Z_{\gamma, M}^{(k,N)}(\kappa) = \int_{\text{Herm}^M(k|k)} \text{sdet} (V_{LM} - \kappa)^{-N} \prod_{m=1}^M \prod_{r=1}^L e^{-\text{str} V_{lm}} \frac{d[V_{LM}]}{(2\pi i)^M k^2}. 
\]  

(60)

### 3.3. Microscopic tail statistics

The tail asymptotic results from combining expression (60) for the partition function (33) with the scaling shown in (29) for the largest eigenvalues inside the tail. This implies the scaling

\[
\kappa = N(c M \lambda)^{-M-1} 
\]  

(61)
with \( c_M \) as in (30) and the supermatrix \( \hat{\lambda} \) being fixed. Plugging this into (60), we can readily carry out the large \( N \) limit and find

\[
\lim_{N \to \infty} Z^{(N)}_{Y^{(M)}} \left( N \left( c_M \hat{\lambda} \right)^{-M-1} \right) = \lim_{N \to \infty} \int_{\Herm^{M}} \det \left( \frac{I_{k|k} - (c_M \hat{\lambda})^{M+1}}{N} V_l^{(M)} \right)^{-N} \\
\times \prod_{m=1}^{M} \prod_{l=1}^{L} e^{-\text{Str} V_{lm}} \frac{d[V_{lm}]}{(2i)^L \pi^L} \\
= \int_{\Herm^{M}} \exp \left[ \text{Str} \left( c_M \hat{\lambda} \right)^{M+1} V_l^{(M)} \right] \prod_{m=1}^{M} \prod_{l=1}^{L} e^{-\text{Str} V_{lm}} \frac{d[V_{lm}]}{(2i)^L \pi^L} \\
= \left[ \int_{\Herm^{M}} \exp \left[ \text{Str} \left( \frac{c_M \hat{\lambda}}{L} \right)^{M+1} V_l^{(M)} \right] \prod_{m=1}^{M} \prod_{l=1}^{L} e^{-\text{Str} V_{lm}} \frac{d[V_{lm}]}{(2i)^L \pi^L} \right]^L. \tag{62}
\]

Certainly, this limit works only when \( \hat{\lambda}_{BB}^{M+1} + [\hat{\lambda}_{BB}^{M+1}] < 0 \) is negative definite. This is usually not the case since the imaginary part of \( \hat{\lambda} \) should eventually be set to 0. Therefore we deform the contours by a slight rotation \( V_{1,j} \to e^{-i\pi/(M+1)} V_{1,j} \) so that for a positive definite Hermitian \( \hat{\lambda} \) we obtain the convergent result

\[
\lim_{N \to \infty} Z^{(N)}_{Y^{(M)}} \left( N \left( c_M \hat{\lambda} \right)^{-M-1} \right) = \left[ \int_{\Herm^{M}} \prod_{m=1}^{M} \frac{d[V_{lm}]}{(2i)^L \pi^L} \exp \left[ -e^{-i\pi/(M+1)} \text{Str} \left( \frac{c_M \hat{\lambda}}{L} \right)^{M+1} V_l^{(M)} \right] \prod_{m=1}^{M} \prod_{l=1}^{L} e^{-\text{Str} V_{lm}} \frac{d[V_{lm}]}{(2i)^L \pi^L} \right]^L. \tag{63}
\]

In the last line, we recognise our claim in section 2 that the tail eigenvalue statistics agree with those of a direct sum, which is namely

\[
\left( \prod_{j=1}^{L} \frac{\det(\bigoplus_{i=1}^{L}(X_{j}^{(M)})^{X_{i}^{(M)}} - \kappa_{E,j} \otimes I_{L})}{\det(\bigoplus_{i=1}^{L}(X_{j}^{(M)})^{X_{i}^{(M)}} - \kappa_{B,j} \otimes I_{L})} \right)^L = \left( \prod_{j=1}^{L} \frac{\det((X_{j}^{(M)})^{X_{i}^{(M)}} - \kappa_{E,j} \otimes I_{L})}{\det((X_{j}^{(M)})^{X_{i}^{(M)}} - \kappa_{B,j} \otimes I_{L})} \right)^L = \left[ \frac{Z^{(N)}_{Y^{(M)}} (\kappa)}{Z^{(N)}_{Y^{(M)}} (\kappa)} \right]^L. \tag{64}
\]
For \( M = L = 1 \), this result agrees with the Bessel kernel result [63] which has been an important result in the study of quantum chromodynamics [75].

The result (63) is the analytical corroboration which we previously mentioned despite that we considered here a particular kind of ensemble where we could carry out the computation. However, we are rather sure that this is the generic behaviour of heavy-tailed statistics. The eigenvalues seem to be too diluted to show their level repulsion which is reflected in the Vandermonde determinant of their joint probability densities, so that effectively they behave as if this level repulsion never existed.

### 3.4. Critical regime of the spectral statistics

Now we investigate how far into the bulk the limiting statistics to a direct sum of random matrices carries over and what the critical regime is where we eventually enter stable statistics such as the sine kernel in the bulk. For this aim we consider a general scaling

\[ \kappa = N^\gamma \kappa_0 \mathbf{1}_{[k]} + N^\delta \tilde{\kappa} \]  

(65)

with \( \delta < \gamma < 1 \). The point \( \kappa_0 > 0 \) is where we zoom into the spectrum and hence the condition \( \delta < \gamma \) where \( \delta \) has to be determined so that the spectral fluctuations on the scale of the local mean level spacing are resolved. For \( \gamma = 1 \) this scale has been \( \delta = \gamma \) as we have seen and for \( \gamma > 1 \) we look into the tail of the largest eigenvalue only implying that we do not see anything of much interest.

Choosing the slightly rotated version of the supermatrices \( V_{lm} \), cf, equation (63), we rescale the supermatrices by \( N^{(1-\gamma)/(M+1)} \)

\[
Z^{(k,N)}_{LM}(N^\gamma \kappa_0 \mathbf{1}_{[k]} + N^\delta \tilde{\kappa}) \\
= \int_{\text{Herm}^{M}\left(\mathbb{C}\right)} \text{Sdet} \left( \mathbf{1}_{[k]} + e^{-i\pi/(M+1)} N^{\delta-\gamma} V_{LM} \right)^{-N} \\
\times \prod_{m=1}^{M} \prod_{l=1}^{L} \exp \left[ -e^{-i\pi/(M+1)} N^{(1-\gamma)/(M+1)} \text{Str} \ V_{lm} \right] \frac{d[V_{lm}]}{(2i)^2 \pi^M}.
\]  

(66)

When Taylor expanding the logarithm of the superdeterminant we notice that only the linear term survives in the large \( N \) limit since \( 1 - j(M + \gamma)/(M + 1) < 0 \) for all \( j \geq 2 \) and any \( M \geq 1 \) and \( \gamma > 1 - M \), so that

\[
Z^{(k,N)}_{LM}(N^\gamma \kappa_0 \mathbf{1}_{[k]} + N^\delta \tilde{\kappa}) \\
\approx \left[ Z^{(k,N)}_{LM} \left( L^{M+1}(N^\gamma \kappa_0 \mathbf{1}_{[k]} + N^\delta \tilde{\kappa}) \right) \right]^{L}. 
\]  

(67)
In this expression, we can read off the local scale given by the exponent \( \delta = (M + 2) \gamma - 1) / (M + 1) \) as then the Taylor expansion in \( \tilde{\kappa} \) terminates with the linear term in the asymptotic limit \( N \rightarrow \infty \).

Equation (67) already shows that even some part of the bulk statistics close to the tail still follows the statistics of a direct sum. A more detailed saddle point analysis would show that we get the direct sum of \( L \) independent sine-kernel statistics. The condition of the scaling exponent \( \gamma \) which has to be satisfied for these kind of statistics is \( \gamma > 1 - M \). Hence, for \( \gamma \leq 1 - M \) we have to go to higher order expansions of the logarithm of the superdeterminant. Those terms couple the supermatrices.

For \( \gamma < 1 - M \), those higher order terms are also large so that we need to carry out an additional saddle point expansion where all \( V_{lm} \) become equal. Thence we would find the statistics of a single sine-kernel, see [64, 75] for the supersymmetric integral expression of these statistics. One needs to be careful, as Rothstein vector fields [76] will occur in the saddle point expansion as they account for all Efetov–Wegner boundary terms [66, 67] that correspond with the diagonalisation of a supermatrix. All those terms have been explicitly computed for diagonalising Hermitian supermatrices in [71].

For \( \gamma = 1 - M \) and, hence, \( \delta = (1 - M - M^2) / (M + 1) \), the quadratic term of the Taylor expansion is of order 1 so that all supermatrices are coupled,

\[
Z_{\gamma \nu M}^{(2N)}(N^{1-M}k_01_{(k)} + N^{(1-M-M^2)/(M+1)}k) \\
N^{\geq 1} \approx \int \operatorname{Herm}_{M(E)}^{(k)} \exp \left[ - e^{-i\pi/(M+1)}N^{M/(M+1)} \right] \\
\times \operatorname{Str} \left( \frac{1}{\kappa_0} V_L^{(M)} + \sum_{m=1}^{M} \sum_{l=1}^{L} V_{lm} \right) \\
\times \exp \left[ e^{-i\pi/(M+1)} \frac{\kappa_0}{2} \operatorname{Str} V_L^{(M)} + e^{-2i\pi/(M+1)} \frac{\kappa_0}{2} \operatorname{Str} (V_L^{(M)})^2 \right] \\
\times \prod_{m=1}^{M} \prod_{l=1}^{L} \frac{d[V_{lm}]}{(2i)^{\pi \kappa_0}}.
\]

This expression can be decoupled by a Hubbard–Stratonovich transformation [77, 78]

\[
\exp \left[ \frac{e^{-2i\pi/(M+1)}}{2} \operatorname{Str} (V_L^{(M)})^2 \right] \\
= \frac{\int_{\operatorname{Herm}(k)} \exp \left[ - \frac{1}{2} \operatorname{Str} \Xi^2 + \frac{e^{-i\pi/(M+1)}}{\kappa_0} \operatorname{Str} \Xi V_L^{(M)} \right] d[\Xi]}{\int_{\operatorname{Herm}(k)} \exp \left[ - \frac{1}{2} \operatorname{Str} \Xi^2 \right] d[\Xi]},
\]

with \( \Xi \) a supermatrix whose Boson–Boson block is an arbitrary Hermitian matrix \( \Xi_{BB} = \Xi_{BB} \in \operatorname{Herm}(k) \) and the Fermion–Fermion block is an arbitrary anti-Hermitian matrix \( \Xi_{FF} = -\Xi_{FF} \in i \operatorname{Herm}(k) \). The Boson–Fermion and Fermion–Boson blocks again contain independent Grassmann variables and the measure \( d[\Xi] \) is the product of all differentials of the supermatrix entries. The partition function can then be approximated anew by the partition function of a direct sum of identical random matrices which however are now coupled by the
supermatrix \( \Xi \), i.e.
\[
Z^{(k,N)}_{Y^{(M)}} \left( N^{1-M} \kappa_0 \mathbf{1}_{k\ell} + N^{(1-M-M^2)/(M+1)} \tilde{\kappa} \right)
\]
\[
N \gg 1 \approx \int \text{Herm}(k|k) \left[ Z^{(k,N)}_{Y^{(M)}} (K(\Xi)) \right]^L \exp \left[ -\frac{1}{2} \text{Str} \, \Xi^2 \right] \, d[\Xi]
\]
(70)

with
\[
K(\Xi) = L^{M+1} \left[ N^{1-M} \kappa_0 \mathbf{1}_{k\ell} + N^{(1-M-M^2)/(M+1)} (\tilde{\kappa} + \kappa_0 \Xi) \right].
\]
(71)

This result resembles those in [79] where the transition between independent diagonal blocks of random matrices between a full matrix have been considered. This underlines that our understanding of a transition between a direct sum to a full matrix without block structures in the tail is ostensibly correct.

One last comment on the critical scale of the eigenvalues \( \lambda \propto N^{1-M} \) of the random matrix \( Y^{(M)} \). It is only slightly larger than the scale \( N^{-M} \) of the macroscopic level density (20) which is freely stable. Although we are already then deep in the tail we are far away from the scale of the largest eigenvalue which scales like \( N \) for the chosen reference scale of the product of inverse Ginibre ensembles, cf., equation (1). We believe that the relative scales should hold for other ensembles too and that in particular the ratio of the scale between the largest eigenvalues and the critical scale should follow the law \( \text{scale}_{\text{largest eigenvalue}}/\text{scale}_{\text{critical}} = N^{1/(2\alpha)} \) where \( \alpha \) is the stability exponent. The Taylor expansion should follow the same mechanism however the probability density \( P(H) \) and thus the corresponding superfunction \( Q(U) \) may vary.

4. Stable ensembles and Poisson statistics in the tail

The last thing we would like to address in the present article is the limit \( L \to \infty \). The multivariate central limit theorem [17] tells us that if the limit of a sum of random matrices exists, the sum should converge to one of the stable random matrix ensembles and this has to hold already at finite matrix dimension \( N \). Specifically, a (strictly) stable Hermitian random matrix \( H \in \text{Herm}(N) \) associated to the stability exponent \( \alpha \) is given when drawing two copies \( H_1 \) and \( H_2 \) of \( H \), then the sum \( (H_1 + H_2)/2^{1/\alpha} \) is also a copy of \( H \), implying that it exhibits the very same statistics, including but not limited to eigenvalues, eigenvectors, and matrix entries (also for finite \( N \)). Hence this behaviour should carry over to the large \( N \) limit. The only question is whether the two limits \( L \to \infty \) and \( N \to \infty \) commute. Our numerical and analytical simulations suggest that this might be true for some cases of the microscopic and macroscopic spectral scales depending on the averaged position of the largest eigenvalues in the tail.

Additionally, mesoscopic spectral scales might arise which are reminiscent to the order of the two limits. What underlines the latter point is the fact that the Bessel result for the level density (12) of the case \( M = 1 \) can be found for all \( L \), see left plots in figure 4. One needs only to cluster the eigenvalues in \( L \) consecutive couples.

The limit \( L \to \infty \) for the model considered in section 3 and in particular for the result (64), is carried out in subsection 4.1, while we consider a random matrix ensemble that is already stable at finite \( N \) in subsection 4.2.

4.1. Limiting statistics for the model of section 3

What does the discussion above on the limiting stable distributions imply for the tail statistics? For that reason, we start from the knowledge that the microscopic tail statistics follows
approximately the one of a direct sum of random matrices, i.e. we start from (63). The spectral
variables are scaled as follows
\[ \kappa = N \left( \lambda_0 1_{k,k} + \frac{\tilde{\lambda}}{\lambda_0 (M + 1)L} \right)^{-M-1}, \tag{72} \]
where \( \lambda_0 > 0 \) is the base point where we zoom into the spectrum and \( \tilde{\lambda} \) measures the spectral
fluctuations. Plugging this into (63) and expanding for large \( L \), we obtain
\[
\lim_{L \to \infty} \lim_{N \to \infty} Z_{\text{L}(k,N)}^{(L)} (\kappa) = \lim_{L \to \infty} \left( \prod_{m=1}^{M} \frac{dV_{1,m}}{(2i)^{1/2}} \left( 1 - \frac{1}{L} e^{-i\pi \frac{\pi}{2} \text{Str} \lambda V_{1,1}^{1-1} \cdots V_{1,M}^{1-1}} \right) \right)^L \\
\times \exp \left[ -e^{-i\pi/(M+1)} \text{Str} \left( \lambda_0^{M+1} V_{1,1}^{1-1} \cdots V_{1,M}^{1-1} + \sum_{m=1}^{M} V_{1,m} \right) \right] \\
= \exp \left[ - \left( \prod_{m=1}^{M} \frac{dV_{1,m}}{(2i)^{1/2}} e^{-i\pi \frac{\pi}{2} \text{Str} \lambda V_{1,1}^{1-1} \cdots V_{1,M}^{1-1}} \right) \right] \\
\times \exp \left[ -e^{-i\pi/(M+1)} \text{Str} \left( \lambda_0^{M+1} V_{1,1}^{1-1} \cdots V_{1,M}^{1-1} + \sum_{m=1}^{M} V_{1,m} \right) \right]. \tag{73} \]

For the second equality we have exploited the fact that the integrand is normalised for
any \( \kappa \propto 1_{k,k} \) because then all determinants in (33) cancel. The average in the exponent
can be simplified due to the supergroup invariance of the integrand that is only broken by
\( \text{Str} \lambda V_{1,1}^{1-1} \cdots V_{1,M}^{1-1} \) so that eventually we have
\[
\lim_{L \to \infty} \lim_{N \to \infty} Z_{\text{L}(k,N)}^{(L)} (\kappa) = \exp[C \text{Str} \lambda] \tag{74} \]
with a constant \( C \) that depends only on \( k, M \) and \( \lambda_0 \).

Let us compare (74) with the partition function of a Poisson distributed spectrum, meaning
the random matrix \( H = \text{diag}(E_1, \ldots, E_N) \) is diagonal and each eigenvalue \( E_j \) is independently
and identically distributed by \( F(E) \). We consider the average
\[
Z_{\text{Poisson}}^{(k,N)} (\kappa) = \left\langle \text{Sdet} (H \otimes 1_{k,k} - 1_{N} \otimes \kappa)^{-1} \right\rangle = \left[ \left\langle \text{Sdet} (E_1 1_{k,k} - \kappa)^{-1} \right\rangle \right]^N, \tag{75} \]
which is the counterpart of (34). The average on the right-hand side is over a single eigenvalue only.

The local spectral fluctuations of a Poisson ensemble happen on the scale \( 1/N \) when the
distribution \( F(E) \) is \( N \) independent. Therefore, we choose the scaling \( \kappa = \lambda_0 1_{k,k} + \tilde{\lambda}/N \) with
\( \lambda_0 \) the base point with a tiny imaginary increment for the regularization then the limit \( N \to \infty \)
of (75) leads to
\[
\lim_{N \to \infty} Z_{\text{Poisson}}^{(k,N)} (\kappa) = \exp \left[ - \int \frac{F(E) dE}{E - \lambda_0} \text{Str} \lambda \right]. \tag{76} \]
Comparison with the result (74) underlines our point that in the large $L$ limit we indeed find the Poisson statistics.

Since the critical scaling $\kappa \propto N^{1-M}$ where the transition to the sine-kernel statistics happens, is independent of $L$ we expect that it is the same critical scale where the Poisson statistics should turn over into the sine-kernel as well. This certainly deserves more investigation, yet we skip it here as it exceeds the scope of the present work.

4.2. Tail statistics for a stable random matrix model

To understand better whether the Poisson statistic holds true generally in the heavy tail as we know it does for heavy-tailed Wigner ensembles [19, 20, 23, 27], we have also numerically checked whether an already stable heavy-tailed unitarily invariant random matrix exhibits the Poisson statistics in its tail. Therefore we would like to point out that such an ensemble can be readily constructed with the help of the GUE. Its probability density is

$$P_{\text{GUE}}(H; \sigma) = \frac{\exp[-\operatorname{tr} H^2/(2\sigma^2)]}{2^{N/2}(\pi\sigma^2)^{N/2}}, \quad H \in \text{Herm}(N),$$

with an arbitrary standard deviation $\sigma > 0$. This standard deviation is now drawn from a stable totally asymmetric univariate density with a stability exponent $\tilde{\alpha} \in ]0, 1]$ and asymmetry parameter $\tilde{\beta} = 1$. The distribution of such a density is given by the Fourier transform [80]

$$\hat{p}_{\tilde{\alpha}}(x) = \int_{-\infty}^{\infty} \exp\left[-ix\omega - \frac{(-i\omega)^{\tilde{\alpha}}}{\cos(\pi\tilde{\alpha}/2)}\right] d\omega/2\pi.$$  

(78)

The condition $\tilde{\alpha} < 1$ is important because only then the support is restricted on the positive real line $\mathbb{R}_+$. When determining the variance $\sigma^2$ from $p_{\tilde{\alpha}}(x)$, we obtain a symmetric unitarily invariant random matrix which is stable with stability exponent $\alpha = 2\tilde{\alpha}$. In particular we consider the random matrix distribution

$$P_{\alpha}(H) = \int_{0}^{\infty} \frac{\exp[-\operatorname{tr} H^2/(2\alpha x)]}{2^{N/2}(\pi x)^{N/2}} \hat{p}_{\alpha/2}(x) dx.$$  

(79)

A simple computation readily proves its stability

$$2^{N^2/\alpha} P_{\alpha} * P_{\alpha}(2^{1/\alpha} H) = 2^{N^2/\alpha} \int_{\text{Herm}(N)} P_{\alpha}(H') P_{\alpha}(2^{1/\alpha} H - H') dH'$$

$$= \int_{0}^{\infty} dx_1 \int_{0}^{\infty} dx_2 \frac{\exp\left[-\frac{u H^2}{2^{2/\alpha}(x_1 + x_2)}\right]}{2^{N^2/2}(2^{-2/\alpha}\pi(x_1 + x_2))^{N/2}} \times \hat{p}_{\alpha}(x_1) \hat{p}_{\alpha}(x_2)$$

$$= P_{\alpha}(H).$$  

(80)

The factor $2^{N^2/\alpha}$ is the Jacobian due to the rescaling $H \rightarrow 2^{1/\alpha} H$. Additionally the second equality results from the convolution rules of two Gaussians and the third equality takes into account the stability of $\hat{p}_{\alpha/2}(x)$.

The construction (79), where one averages over the variance has also been studied for similar ensembles in [6–8, 33, 34] though the authors of these works did not aim for stable distributions.

We have employed the construction above for the Monte Carlo simulations which readily can be numerically generated by noticing that $\sqrt{\sigma} H$ with $\sigma$ and $H$ independently drawn
Figure 5. The level spacing distributions of the unfolded eigenvalues (83) for the stable random matrix ensemble (79) for the four stability exponents $\alpha = 0.5, 1.0, 1.5, 1.8$ and the matrix dimensions $N = 50, 500$. The Monte Carlo simulations (coloured symbols) are the same of figure 7. The nomenclature $p_k(s)$ refers to the spacing of the $k$th and $(k+1)$st smallest eigenvalue. Due to symmetry of the ensemble, the two distributions $p_k(s)$ and $p_{N-k}(s)$ must agree as they indeed approximately do (same coloured circles and triangles). As a comparison we have drawn the Poisson distribution (16) (solid curve) and the Wigner surmise (17) (dashed curve).

From $\tilde{p}_{\alpha/2}(x)$ and $P_{\text{GUE}}(H; 1)$, respectively, leads to the same random matrix. For four different stability exponents $\alpha = 0.5, 1, 1.5, 1.8$, we have numerically simulated the level spacing distribution as well as their level density for the four largest and four smallest eigenvalues. Those are drawn in figure 5. Since the ensemble is symmetric about the origin the ensemble exhibits on both sides heavy-tails.
Figure 6. The macroscopic level densities $\rho_1(\lambda)$ (blue solid curve, see equation (81)) and $\rho_{CL}(\lambda)$ (red dashed curve, see equation (82)). Despite the corresponding ensembles are stable only those yielding $\rho_{CL}(\lambda)$ can be stable under free convolution.

The macroscopic level density is an averaged Wigner semicircle

$$
\rho_\alpha(\lambda) = \lim_{N \to \infty} \frac{1}{N} \left\langle \text{tr} \delta(\lambda \mathbf{1}_N - N^{-1/2} H) \right\rangle = \int_{\lambda^{-1/2}}^{\infty} \frac{\sqrt{4\lambda - \lambda^2}}{2\pi\lambda} \rho_{\alpha/2}(x) \, dx.
$$

(81)

Evidently, the asymptotic behavior is $\rho_\alpha(\lambda) \propto |\lambda|^{-1-\alpha}$ for $|\lambda| \gg 1$.

Despite the considered ensemble being stable, it is not ‘freely stable’ meaning two copies of the matrix are not free random variables. One can convince oneself fairly easily by considering the case $\alpha = 1$. The symmetric distribution which is stable under free convolution is uniquely given (up to a scaling parameter $c > 0$) by the Lorentz function [43] (also known as the Cauchy or Breit–Wigner distribution)

$$
\rho_{CL}(\lambda) = \frac{1}{\pi} \frac{c}{1 + c^2 \lambda^2}.
$$

(82)

One can then show that it is always the case $\rho_1(\lambda) \neq \rho_{CL}(\lambda)$ regardless of what scaling $c > 0$ is chosen, cf, figure 6 where $c$ is fixed by $\rho_1(0) = c/\pi$.

When computing the (unfolded) largest eigenvalues, we have been very surprised by the fact that the largest eigenvalues are still deep in the bulk of the distribution $\rho_{\alpha}(\lambda)$ in contrast to the model in section 3. Thence, the approximation of this distribution by its leading asymptotic behavior $\rho_{\alpha}(\lambda) \propto \lambda^{-1-\alpha}$ has not been suitable for the unfolding. Instead, we have unfolded the entire spectrum with the new variables

$$
\mu(\lambda) = \int_0^\lambda \rho_\alpha(\lambda') \, d\lambda'.
$$

(83)

This substitution maps the spectrum from the real line $\mathbb{R}$ to the open interval $]-1/2, 1/2[$ with a uniform level density. The mean of the largest eigenvalue after this mapping is well-defined.
Figure 7. The non-unfolded (left plots) and unfolded (right plots) macroscopic level densities for the stable ensemble (79) for the four stability exponents $\alpha = 0.5, 1.0, 1.5, 1.8$ and the matrix dimensions $N = 50, 500$. The analytical curves (solid black curves) are generated via the integral (81). The unfolded eigenvalues $\mu$ are given by (83). For each ensemble (coloured symbols) we have created $10^5$ configurations. The scattering in the unfolded densities close to the boundaries $\mu = \pm 1/2$ can be explained due to very rare events as those points correspond to the heavy tails. The slight dip at $\mu = 0$ for $\alpha = 0.5$ can be understood by the very narrow peak which is not sufficiently resolved by the chosen bin. The vertical lines show the mean positions of the largest and smallest eigenvalue in the unfolded variables $\mu$. In the left plots we have mapped these positions back to $\lambda$.

and after mapping this back to the original spectrum, they have been indicated as vertical lines in figure 7.

The reason why the largest eigenvalues do not lie at the utmost end of the heavy-tail (for the unfolded variables at $\mu = \pm 1/2$) is that the tails are a superposition of almost all eigenvalues
instead of (essentially) only the largest one. Therefore the situation is not clear as to whether we should find the Poisson statistics when taking the limit $N \to \infty$. To further understand the problem in more detail we compute the partition function

$$Z_{\alpha}^{(k, \mathcal{N})}(\kappa) = \langle \text{Sdet} (H \otimes \mathbf{1}_{k} \pm \mathbf{1}_{\mathcal{N}} \otimes \kappa)^{-1} \rangle = \int_{0}^{\infty} \langle \text{Sdet} (H \otimes \mathbf{1}_{k} \pm \mathbf{1}_{\mathcal{N}} \otimes \kappa)^{-1} \rangle_s \rho_{\alpha/2}(x) dx$$

with respect to the considered stable ensemble. The notation $\langle \cdot \rangle_s$ denotes the average over the GUE with variance $x$.

The average over the GUE can be cast into a supermatrix integral like in [74] with the help of the superbosonisation formula. In doing so we assume that the Boson–Boson block $\kappa_{BB}$ is diagonalised which is always possible when its Jordan normal form is diagonal. Then we define the diagonal matrix $\hat{S} = (\text{sign}(\text{Im}[\kappa_{BB}]), \mathbf{1}_k)$ which comprises all signs of the imaginary parts of $\kappa_{BB}$. In this way, the Hermitian numerical part of $(iH \otimes \hat{S} - iN \otimes \hat{S}\kappa)^{BB}$ is positive definite. This allows us to write the superdeterminant in terms of an average over a Gaussian integral of a rectangular supermatrix $V$ of size $(N/0) \times (k/k)$ as the convergence is guaranteed now,

$$\text{Sdet} (H \otimes \mathbf{1}_{k} \pm \mathbf{1}_{\mathcal{N}} \otimes \kappa)^{-1} = \frac{\int \exp[i \text{Str} V^\dagger V \hat{S} \kappa - \text{itr} H V \hat{S} V^\dagger] d[V]}{\int \exp[- \text{Str} V^\dagger V] d[V]}.$$

After averaging over $H$, we arrive at

$$\langle \text{Sdet} (H \otimes \mathbf{1}_{k} \pm \mathbf{1}_{\mathcal{N}} \otimes \kappa)^{-1} \rangle_s = \frac{\int \exp[i \text{Str} V^\dagger V \hat{S} \kappa - x \text{Str}(\hat{S} V^\dagger V)^2/2] d[V]}{\int \exp[- \text{Str} V^\dagger V] d[V]}.$$ 

In the last step, we employ the superbosonisation formula [72–74] and replace $V^\dagger V$ by $\gamma U$ with $\gamma > 0$ a scaling that needs to be adjusted and the supermatrix $U \in \text{Herm}_{\alpha}(k/k)$ which is realized from the very same set as in the supersymmetric projection formula (39). Thus we eventually arrive at

$$Z_{\alpha}^{(k, \mathcal{N})}(\kappa) = \int_{0}^{\infty} dx \int_{\text{Herm}_{\alpha}(k/k)(2^{k})^k} d[U] \rho_{\alpha/2}(x) \text{Sdet} U^N \exp \left[ -\frac{x \gamma^2}{2} \text{Str}(\hat{S} U)^2 + i \gamma \text{Str} U \hat{S} \kappa \right].$$

The scale of the macroscopic level density is obtained for $\kappa \propto \sqrt{N} \kappa_0 \in \mathbb{R}$. Then we choose $\gamma = \sqrt{N}$ and perform the saddle point analysis for $N \gg 1$. The two corresponding saddle point solutions for the eigenvalues of $U \hat{S}$ are $z_{\pm} = (i\kappa_0 \pm \sqrt{4x - \kappa_0^2})/(2x)$. Here we have to split the discussion into cases depending on whether $4x$ is larger or smaller than $\kappa_0^2$.

If $4x > \kappa_0^2$, then the real parts of the eigenvalues in the Boson–Boson block of $U \hat{S}$ must have the signature $\hat{S}_{BB}$ as otherwise the maximum of the integrand along the contour is never acquired at the saddle point. Moreover the eigenvalues of the Fermion–Fermion block need to be the same eigenvalues since only then the Berezinian of the diagonalisation of $U \hat{S}$ (Jacobian in superspace, see [61]) is of order one. Other solutions will yield higher orders in $1/N$. To summarise the saddle point manifold is given by

$$U \hat{S} = \frac{i\kappa_0}{2x} \mathbf{1}_{k} \pm \frac{\sqrt{4x - \kappa_0^2}}{2x} U \text{diag}(\hat{S}_{BB}, \hat{S}_{BB}) \hat{U}^{-1}.$$
where $\widehat{U} \in \text{U}(k_+|k_-)\times \text{U}(k_-|k_+)$ is a Haar distributed unitary supermatrix with $k_+$ and $k_-$ the number of plus and minus signs in $\widehat{S}\text{BB}$. We would like to point out that the change of coordinates for (88) involves a Rothstein vector field [76] which we denote by $\mathcal{V}_U$. It is however $N$-independent because it only corresponds to the substitution and not the integrand as we know it for Jacobians.

If $4x \leq \kappa_0^2$, the solutions become entirely imaginary though only one of them is a maximum of the integrand for the integration variables. The second derivative of the exponential term is at the two saddle points

$$\partial_z^2 \left( -\frac{x^2}{2} + i\kappa_0 z + \ln(z) \right) \bigg|_{z=\pm \kappa} = -x + \frac{4x^2}{(\kappa_0 \pm \sqrt{\kappa_0^2 - 4x})^2}. \tag{89}$$

Combining this with the fact that the bosonic eigenvalues run through these points along the imaginary line and the fermionic ones parallel to the real line, both amount to an additional minus sign in the second term of the Taylor expansion about the saddle point in (87). Therefore only

$$U\widehat{S} = i\frac{\kappa_0 - \sqrt{\kappa_0^2 - 4x}}{2x} 1_{kk} \tag{90}$$

can be a maximum along the contours. Here, we would like to underline that no Rothstein vector field is needed as we do not need to diagonalise the supermatrix to reach this saddle point in contrast to the case $4x > \kappa_0^2$.

The spectral fluctuations can be obtained by setting $\kappa = \sqrt{N}\kappa_0 1_{kk} + \tilde{\kappa}/\sqrt{N}$. We expand

$$U\widehat{S} = i\frac{\kappa_0 - \sqrt{\kappa_0^2 - 4x}}{2x} 1_{kk} + \sqrt{\frac{4x - \kappa_0^2}{2x}} \tilde{U} \text{diag}(\tilde{S}\text{BB}, \tilde{S}\text{BB}) \tilde{U}^{-1} + \frac{1}{\sqrt{N}} \tilde{U} \delta Q \tilde{U}^{-1} \tag{91}$$

for $4x > \kappa_0^2$ and

$$U\widehat{S} = i\frac{\kappa_0 - \sqrt{\kappa_0^2 - 4x}}{2x} 1_{kk} + \frac{1}{\sqrt{N}} \delta Q \tag{92}$$

for $4x < \kappa_0^2$ up to second order in the massive modes $\delta Q$. The supermatrix $\delta Q$ describes essentially the superspaces $\text{Herm}(k_+|k_-) \times \text{Herm}(k_-|k_+)$ and $\text{Herm}(k|k)$, respectively for the two situations. Their integrations yield 1 due to the normalisation and we eventually obtain

$$\lim_{N \to \infty} Z^{(1,N)}(\kappa) = \int_0^{\kappa_0^2/4} dx \tilde{p}_{\kappa/2}(x) \exp \left[ -\frac{\kappa_0 - \sqrt{\kappa_0^2 - 4x}}{2x} \text{Str}(\tilde{\kappa}) \right]$$

$$+ \int_{\kappa_0^2/4}^{\infty} dx \tilde{p}_{\kappa/2}(x) \exp \left[ -\frac{\kappa_0}{2x} \text{Str}(\tilde{\kappa}) \right] \int_{\nu(\kappa) + k_+ \pm k_-}^{\kappa_0 - \kappa \pm \sqrt{\nu(\kappa)}} \frac{d\nu}{\sqrt{\nu^2 - k_+^2 - k_-^2}}$$

$$\times \exp \left[ \frac{4x - \kappa_0^2}{2x} \text{Str}(\tilde{U} \text{diag}(\tilde{S}\text{BB}, \tilde{S}\text{BB}) \tilde{U}^{-1} \tilde{\kappa}) \right] \frac{d\mu(\tilde{U})}{\nu^2 - k_+^2 - k_-^2}. \tag{93}$$

The exponential term $\exp[\mathcal{V}_U]$ is the application of the Rothstein vector field $\mathcal{V}_U$ which takes care of all Efetov–Wegner boundary terms [66, 67] that result from the corresponding change of coordinates.
The result (93) is a superposition of the Poisson partition function (76) convolved with the stable distribution \( p_{\nu/2}(x) \) and the sine kernel partition function in its supersymmetric form \([64, 75]\) convolved again with \( p_{\nu/2}(x) \). For the sine kernel, usually the vector field is dropped as it only generates lower point correlations than the \( k \)-point correlation function that can be derived by taking derivatives in \( \kappa \) and then setting \( \kappa_{BB} = \kappa_{FF}, \) see \([46–48]\).

For the computation above, we have assumed that \( \kappa_0 \) is of order \( O(1) \) in \( N \). We could also choose that \( \kappa_0 \) is of a larger order since the spectrum has a heavy tail so that eigenvalues can indeed lie very deep in the tail. For instance this is the case for the product of inverse Ginibre matrices where the ratio of the scale of the largest eigenvalue and the bulk is of order \( \text{scale largest eigenvalue}/\text{scale bulk} = N \).

Assuming \( \kappa_0 \gg 1 \) in (93) with \( \tilde{\kappa} = O(\kappa_0) \), we can Taylor expand the square root in the first term which makes the exponent \( x \) independent so that we get the Poisson partition function (76),

\[
\int_0^{\kappa_0^{3/4}} dx \, \hat{p}_{\nu/2}(x) \exp \left[ -\frac{\kappa_0 - \sqrt{\kappa_0^2 + 4x}}{2x} \text{Str } \tilde{\kappa} \right] \approx \exp \left[ -\frac{1}{\kappa_0} \text{Str } \tilde{\kappa} \right].
\]  

For the second term we rescale \( x \to \kappa_0^2 x/2 \) then the Jacobian together with the approximation \( \hat{p}_{\nu/2}(\kappa_0^2 x/2) \propto (\kappa_0^2 x)^{-1/2} \) is of the size \( \kappa_0^{-1} \) while the integrand is of order one. Thus it is a lower order term and vanishes when \( \kappa \) is of an order larger than \( \sqrt{N} \).

In summary, for \( \kappa \gg \sqrt{N} \), the partition function becomes the one for the Poisson statistic,

\[
Z_{\alpha}(N) \approx \exp \left[ -\frac{1}{\kappa_0} \text{Str } \tilde{\kappa} \right] + O(\kappa_0^{-\alpha}).
\]  

The problem is that the convergence is slower for smaller \( \alpha \).

The question is whether the largest eigenvalues are now of larger order than the scale of the macroscopic level density for the ensemble (79). Using the averaged position of the largest and smallest eigenvalue in the unfolded variables (83), we see that their position (horizontal lines in figure 7) moves to the extreme values at \( \pm 0.5 \). As their change is however very tiny it could be very likely that those positions saturate at a certain value which implies the mixed statistics for the level spacing distribution between the four largest and smallest consecutive eigenvalues seen in figure 5 will persist when taking the limit \( N \to \infty \).

We have also simulated the ensemble for \( \alpha = 0.5 \) and \( N = 5000 \) and it seems that this kind of saturation is taking place. Nonetheless, we can confirm that the similarity to the Poisson statistics is diminished for a smaller \( \alpha \) which can be indeed understood by the resulting error term in equation (95).

A more detailed analysis is needed to decide which scenario either (93) or (95), is actually realised.

5. Conclusions and two conjectures

We investigated heavy-tailed unitarily invariant random matrices and their limiting spectral statistics in the tail. In particular we addressed the question whether the statistics are stable when all the remaining statistics are stable. To achieve this we considered two classes of random matrices. One is a product of inverse Ginibre matrices that are known \([43, 44]\) to yield a freely stable macroscopic level density. Surprisingly in the tail the spectrum is not stable. When \( L \) is the number of matrices added in order to check the stability, the eigenvalues in the tail cluster into groups of \( L \) eigenvalues. Eigenvalues inside this cluster become statistically independent in the limit of large matrix dimension \( N \) while eigenvalues in different clusters
are still correlated. Our interpretation is that the sum of heavy-tailed random matrices behave in the tail like a direct sum of the same types of random matrices. Our analytical computations with the supersymmetry method confirm this. When looking at the particular details of the computation one notices that this easily carries over to a sum of heavy-tailed random matrices that do not necessarily have to be equally distributed nor do not need to have the same stability exponent. We believe that this is even true for real and quaternionic matrices. We now arrive at our first conjecture.

**Conjecture 1 (Tail statistics of a sum of heavy-tailed random matrix ensembles).** Let $X_1, \ldots, X_L$ be independently (not necessarily identically) distributed random matrices with heavy-tails, and unitary invariance (eigenvalues and eigenvectors are uncorrelated) such that the position of the largest eigenvalues is on a scale larger than that of the eigenvalues in the bulk. Then the statistics of the largest eigenvalues in the tail of the sum $\sum_{j=1}^L X_j$ and of the direct sum $\bigoplus_{j=1}^L X_j$ will be the same up to a scaling.

To corroborate this conjecture we have additionally performed Monte Carlo simulations for unitary Jacobi matrices, which are $N \times N$ sub-blocks of $2N \times 2N$ Haar distributed unitary matrices, and for complex $N \times N$ Bernoulli matrices, where all entries are independently and identically distributed random variables uniformly drawn from the set $\{0, \pm 1, \pm i, \pm 1 + i, \pm 1 - i\}$. Similarly to the Ginibre ensemble, we have inverted these matrices and added independently sampled copies of these random matrices. In figure 8, we have plotted the level spacing distribution between the four largest eigenvalues. The results look very similar to those found in figure 4 for the Ginibre case. With the Jacobi ensemble we have ruled out that the effect is due to the independence of the matrix elements of the Ginibre matrices and the Bernoulli ensemble breaks the $U(N)$ group invariance highlighting that the effect is also independent of this invariance.

As we have seen there are also stable statistics in the product of inverse Ginibre matrices. The scale ‘scale critical’ of the transition from stable to unstable spectral statistics has been quantified by the following ratio of scales $\text{scale}_{\text{largest eigenvalue}} / \text{scale}_{\text{critical}} = N^{1/(2\alpha)}$, where $\alpha$ is the stability exponent. We think that this critical scaling might be universal. Certainly a further investigation is needed for more general classes of random matrix ensembles like the multiplicative Pólya ensembles [81, 82], that also comprise several heavy-tailed ensembles.

It seems to be paramount in conjecture 1 that the largest eigenvalues are on a larger scale than the bulk as we have observed with the second class of ensembles. This class consists of averaged GUE’s, where one integrates over the variance with a stable univariate distribution. This construction is very similar to the one in [6–8, 33, 34], where other averaging distributions have been studied. This led to a heavy-tailed random matrix that is already stable for fixed matrix dimension $N$. With the help of this class, we wanted to examine whether the limiting statistics for $L \to \infty$ follows the Poisson statistics as it is the natural choice for a direct sum of random matrices. Our numerical simulations suggest that this is not true. Through our analytical computations we have found out that it has to follow the Poisson statistics if the largest eigenvalue scales much larger than the bulk otherwise one should find a mixture of Poisson statistics and a kind of average of the sine-kernel result. The latter seems to be the case of this average of the GUE with a heavy-tailed standard deviation. Thus, we state our second conjecture.

**Conjecture 2 (Tail statistics of stable random matrix ensembles).** Let $X$ be a heavy-tailed stable random matrix with unitary invariance (eigenvalues and eigenvectors are uncorrelated). If the largest eigenvalues are considerably larger than the bulk of eigenvalues then the local spectral statistics in the tail follows Poisson statistics.
Figure 8. Level spacing distributions between the largest consecutive unfolded eigenvalues (histograms, $p_{j,j+1}$ is the level spacing distribution between the $j$th and $(j+1)$st largest eigenvalue) of a sum of $L = 1, 2, 3, 4$ inverse complex Bernoulli matrices (left plots) and inverse unitary Jacobi matrices (right plots). The Monte Carlo simulations have been performed for matrices with matrix dimension $N = 200$ and an ensemble size of $10^5$. For reference we have also drawn the spacing distributions of the Poisson statistics (16) and the Wigner surmise (17) of the GUE.

These two conjectures should certainly also carry over in some way to the other symmetry classes for the Hermitian, see [64], as well as non-Hermitian random matrix ensembles. Surely for complex spectra other mechanisms will enter the game. Nevertheless in two dimensions the added spatial capacity will further facilitate an increase in the decorrelation of eigenvalues. The supersymmetry method will be certainly helpful also in these cases.
Acknowledgments

MK acknowledges fruitful discussions with Jiyuan Zhang and Holger Kösters.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

ORCID iDs

M Kieburg https://orcid.org/0000-0001-8864-6022

References

[1] Mehta M L 2004 Random Matrices 3rd edn (New York: Academic)
[2] Forrester P J 2010 Log-Gases and Random Matrices (Princeton, NJ: Princeton University Press)
[3] Akemann G, Baik J and Di Francesco P (ed) 2011 The Oxford Handbook of Random Matrix Theory (Oxford: Oxford University Press)
[4] Burda Z, Jurkiewicz J, Nowak M A, Papp G and Zahed I 2001 Lévy matrices and financial covariances Acta Phys. Pol. B 34 4747
[5] Meerschaert M M and Scheffler H-P 2003 Portfolio modeling with heavy tailed random vectors Handbook of Heavy Tailed Distributions in Finance ed S T Rachev ch 15 (Amsterdam: Elsevier)
[6] Burda Z, Görlich A T and Wacław B 2006 Spectral properties of empirical covariance matrices for data with power-law tails Phys. Rev. E 74 041129
[7] Bohigas O, de Carvalho J X and Pato M P 2008 Disordered ensembles of random matrices Phys. Rev. E 77 011122
[8] Akemann G, Fischmann J and Vivo P 2010 Universal correlations and power-law tails in financial covariance matrices Physica A 389 2566–79
[9] Birolí G and Tarzia M 2012 The Lévy–Rosenzweig–Porter random matrix ensemble (arXiv:2012.12841)
[10] Münix MC, Schäfer R and Guhr T 2014 A random matrix approach to credit risk PloS One 9 e98030
[11] Kanazawa T 2016 Heavy-tailed chiral random matrix theory J. High Energy Phys. JHEP05(2016)166
[12] Oymak S and Tropp J A 2017 Universality laws for randomized dimension reduction, with applications Int. Inference 7 337–446
[13] Minsker S 2018 Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries Ann. Stat. 46 2871–903
[14] Martin C H and Mahoney M W 2018 Implicit self-regularization in deep neural networks: evidence from random matrix theory and implications for learning (arXiv:1810.01075)
[15] Martin C H and Mahoney M W 2019 Traditional and heavy-tailed self regularization in neural network models Proc. 36th Int. Conf. on Machine Learning, Long Beach, California, PMLR vol 97
[16] Heiny J 2019 Random matrix theory for heavy-tailed time series J. Math. Sci. 237 652–66
[17] Rvačeva E L 1954 On domains of attraction of multidimensional distributions L’Vov. Gos. Univ. Žap. 29 5
[18] Cizeau P and Bouchaud J P 1994 Theory of Lévy matrices Phys. Rev. E 50 1810
[19] Soshnikov A 2004 Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails Electron. Commun. Probab. 9 82–91
[20] Birolí G, Bouchaud J-P and Potters M 2007 On the top eigenvalue of heavy-tailed random matrices Europhys. Lett. 78 10001
[21] Burda Z, Jurkiewicz J, Nowak M A, Papp G and Zahed I 2007 Random Lévy matrices revisited Phys. Rev. E 75 051126
[22] Arous G B and Guionnet A 2008 The spectrum of heavy tailed random matrices Commun. Math. Phys. 278 715–51
[23] Auffinger A, Ben Arous G and Péché S 2009 Poisson convergence for the largest eigenvalues of heavy tailed random matrices Ann. Inst. Henri Poincare 45 589–610
[24] Vershynin R 2012 Introduction to the non-asymptotic analysis of random matrices Compressed Sensing, Theory and Applications ed Y Eldar and G Kutyniok ch 5 (Cambridge: Cambridge University Press)
[25] Benaych-Georges F, Guionnet A and Male C 2014 Central limit theorems for linear statistics of heavy tailed random matrices Stoch. Process. Appl. 126 3331–52
[26] Tarquini E, Biroli G and Tarzia M 2016 Level statistics and localization transitions of Lévy matrices

[28] Heiny J and Mikosch T 2017 Eigenvalues and eigenvectors of heavy-tailed sample covariance matrices with general growth rates: the iid case Stoch. Process. Appl. 127 2179–207
[29] Bordenave C and Guionnet A 2017 Delocalization at small energy for heavy-tailed random matrices Commun. Math. Phys. 354 115–59
[30] Male C 2017 The limiting distributions of large heavy Wigner and arbitrary random matrices J. Funct. Anal. 272 1–46
[31] Guédon O, Litvak A, Pajor A and Tomczak-Jaegermann N 2017 On the interval of fluctuation of the singular values of random matrices J. Eur. Math. Soc. 19 1469–505
[32] Burda Z, Janik R A, Jurkiewicz J, Nowak M A, Papp G and Zahed I 2002 Free random Lévy matrices Phys. Rev. E 65 021106
[33] Akemann G and Vivo P 2008 Power law deformation of Wishart–Laguerre ensembles of random matrices J. Stat. Mech. P09002
[34] Abul-Magd A Y, Akemann G and Vivo P 2009 Superstatistical generalizations of Wishart–Laguerre ensembles of random matrices J. Phys. A: Math. Theor. 42 175207
[35] Choi J and Muttilah K A 2009 Rotationally invariant family of Lévy-like random matrix ensembles J. Phys. A: Math. Theor. 42 152001
[36] Guhr T and Schell A 2020 Matrix moments in a real, doubly correlated algebraic generalization of the Wishart model (arXiv:2011.07573)
[37] Gupta A K and Nagar D K 1999 Matrix Variate Distributions (Monographs and Surveys in Applied and Pure Mathematics vol 104) (Boca Raton, FL: CRC Press)
[38] Arizmendi E O and Pérez-Abreu V 2009 The S-transform of symmetric probability measures with unbounded supports Proc. Am. Math. Soc. 137 3057
[39] Speicher R 2011 Free Probability Theory Oxford: Oxford University Press chapter 22 of reference[3]
[40] Efetov KB 1997 Supersymmetry in Disorder and Chaos 1st edn (Cambridge: Cambridge University Press)
[41] Zirnbauer M R 2006 The Supersymmetry Method of Random Matrix Theory (Encyclopedia of Mathematical Physics vol 5) ed J-P Françoise, G L Naber and S T Tsou (Oxford: Elsevier) p 151
[42] Ginibre J 1965 Statistical ensembles of complex, quaternion, and real matrices J. Math. Phys. 6 440–9
[43] Wishart J 1928 The generalised product moment distribution in samples from a normal multivariate population Biometrika 20A 32–52
Marčenko V A and Pastur L A 1967 Distribution of eigenvalues for some sets of random matrices
Math. USSR Sbornik 1 457–83
Marčenko V A and Pastur L A 1967 Mat. Sb. 72 507–36 (Engl. transl.)

Dietz B and Haake F 1990 Taylor and Padé analysis of the level spacing distributions of random-matrix ensembles Z. Phys. B 80 153–8

Akemann G, Gorski V and Kieburg M 2021 in preparation

Abramowitz M and Stegun I A 1965 Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (New York: Dover)
Penson K A and Žyczkowski K 2011 Product of Ginibre matrices: Fuss–Catalan and Raney distributions Phys. Rev. E 83 061118
Kuijlaars A B J and Zhang L 2014 Singular values of products of Ginibre random matrices, multiple orthogonal polynomials and hard edge scaling limits Commun. Math. Phys. 332 759–81

Akemann G, Burda Z and Kieburg M 2014 Universal distribution of Lyapunov exponents for products of Ginibre matrices J. Phys. A: Math. Theor. 47 395202

Akemann G, Burda Z and Kieburg M 2019 From integrable to chaotic systems: universal local statistics of Lyapunov exponents Europhys. Lett. 126 40001
Liu D Z, Wang D and Wang Y 2018 Lyapunov exponent, universality and phase transition for products of random matrices (arXiv:1810.00433)

Akemann G, Burda Z and Kieburg M 2020 Universality of local spectral statistics of products of random matrices Phys. Rev. E 102 052134

Berezin F A 1987 Introduction to Superanalysis 1st edn (Dordrecht: Reidel)
Kaymak V, Kieburg M and Guhr T 2014 The supersymmetry method for chiral random matrix theory with arbitrary rotation-invariant weights J. Phys. A: Math. Theor. 47 295201

Kieburg M 2015 Supersymmetry for products of random matrices Acta Phys. Pol. B 46 1709–28
Zirnbauer M R 1996 Riemannian symmetric superspaces and their origin in random-matrix theory J. Math. Phys. 37 4086

Parisi G and Sourlas N 1979 Phys. Rev. Lett. 43 744

Wegner F 1983 unpublished
Efetov K B 1983 Supersymmetry and theory of disordered metals Adv. Phys. 32 53
Constantinescu F 1988 The supersymmetric transfer matrix for linear chains with nondiagonal disorder J. Stat. Phys. 50 1167–77
Constantinescu F and de Groote H F 1989 The integral theorem for supersymmetric invariants J. Math. Phys. 30 981–92
Kieburg M, Kohler H and Guhr T 2009 Integration of Grassmann variables over invariant functions on flat superspaces J. Math. Phys. 50 013528
Kieburg M 2011 On the Efetov–Wegner terms by diagonalizing a Hermitian supermatrix J. Phys. A: Math. Theor. 44 285210

Sommers H-J 2007 Acta Phys. Pol. B 38 4105
Littlennann P, Sommers H-J and Zirnbauer M R 2008 Commun. Math. Phys. 283 343
Kieburg M, Sommers H-J and Guhr T 2009 J. Phys. A: Math. Theor. 42 275206
Verbaarschot J J M 2004 The supersymmetric method in random matrix theory and applications to QCD AIP Conf. Proc. 744 277–362
Rothstein M J 1987 Integration on noncompact supermanifolds Trans. Am. Math. Soc. 299 387
Hubbard J 1959 Calculation of partition functions Phys. Rev. Lett. 3 77–8
Stratonovich B L 1957 On a method of calculating quantum distribution functions Sov. Phys. - Dokl. 2 416

Kieburg M, Verbaarschot J J M and Zafeiropoulos S 2013 Spectral properties of the Wilson Dirac operator and random matrix theory Phys. Rev. D 88 094502
Lévy P 1925 Calcul des Probabilités (Paris: Gauthier-Villars)
Kieburg M and Kösters K 2019 Products of random matrices from polynomial ensembles Ann. Inst. Henri Poincare 55 98–126
Fürstner Y-P, Kieburg M and Kösters K 2020 Polynomial ensembles and Pólya frequency functions J. Theor. Probab.