

Random band and block matrices with correlated entries

Riccardo Catalano* Michael Fleermann† Werner Kirsch‡

Abstract

We derive limit laws for the empirical spectral distributions of random band and block matrices with correlated entries. In the first part of the paper, we study band matrices with *approximately uncorrelated* entries. We strengthen previously obtained results while requiring weaker assumptions, which is made possible by a refined application of the method of moments. In the second part of the paper, we introduce a new two-layered correlation structure we call *SSB-HKW correlated*, which enables the study of structured random matrices with correlated entries. Our results include semicircle laws in probability and almost surely, but we also obtain other limiting spectral distributions depending on the conditions. Simple necessary and sufficient conditions for the limit law to be the semicircle are provided.

Keywords: random band matrices; block matrices; correlated entries; semicircle law.

MSC2020 subject classifications: 60B20.

Submitted to EJP on July 26, 2022, final version accepted on January 4, 2024.

1 Introduction

This paper is concerned with the study of random band and block matrices with correlated entries. We extend the works of [16], [13], and [25] into various directions, and we include and extend the so far unpublished results of the dissertation [6] of the first author. The authors of [25] have shown the semicircle law in expectation for an ensemble of random matrices that exhibit dependencies within certain equivalence classes of matrix entries, while independence is assumed between different equivalence classes. The authors of [16] have developed an ensemble they called *approximately uncorrelated*, allowing for equally strong dependencies between all matrix entries irrelevant of their distance, and the semicircle law in probability was observed. In [13], the setup of random band matrices with approximately uncorrelated entries has been analyzed deeper, deriving conditions for the semicircle law to hold in probability

*FernUniversität in Hagen, Germany. E-mail: riccardo.catalano@fernuni-hagen.de

†Saxion University of Applied Sciences, The Netherlands. E-mail: m.a.fleermann@saxion.nl

‡FernUniversität in Hagen, Germany. E-mail: werner.kirsch@fernuni-hagen.de

and almost surely, extending the work of [16]. For some of the results, the notion of approximately uncorrelated ensembles was tightened to ensembles called α -almost uncorrelated.

By now, the literature on random matrices with correlated entries has grown tremendously, while in the case of structured random matrices such as band and block matrices, results remain sparse for the correlated case. Therefore, one of our aims is to contribute to this domain. We give a short summary of relevant previous studies: After the seminal paper [25], random matrices with more specific correlation structures have been studied in [15], [14], and [22], all of which consider the case of symmetric random matrices with independent diagonals, but with different assumption on the correlations within these diagonals. Further, the authors of [3] study the case of symmetric random matrices with entries which are functions of shifted i.i.d. random fields. The papers [16], [21], and [13] focus on semicircle laws for approximately uncorrelated entries and in particular Curie-Weiss models, with a local version of their results contained in [12]. Further results in the local regime were achieved in [7] for finite range correlations, in [1] for correlated Gaussian entries and in [8] for more general correlations. For structured random matrices with correlations, see [18] (and references therein) for arbitrary correlations within log-size families of matrix entries, but with independence between these families, [20] for random band matrices with exchangeable entries and [13] for random band matrices with approximately uncorrelated entries.

This paper is organized as follows: Section 2 contains the first main theorem of this paper, Theorem 2.1. There, we show that the additional assumptions made in [13] can be dropped without replacement (thus only imposing the conditions as in [16]), while obtaining stronger results. These improvements are achieved by employing a refined way of applying the method of moments, where we use observations in [10] (which were also presented in [9]). See also [11] for an overview. Further, by a standard perturbation argument, we also derive that our results remain true for non-periodic band matrices. In Section 3 we formulate the second main result of this paper: Theorem 3.3. This theorem was initially motivated by an extension of the results in [13] to band matrices with correlated entries and a bandwidth of proportional growth. However, we study this problem in a much more general setting, including and extending the yet unpublished results [6], also generalizing the findings in [23]. First, we revisit the work [25] and drop the assumption that entries from different equivalence classes should be independent. Instead, matrix entries are assumed to be approximately uncorrelated between different equivalence classes, and arbitrarily correlated within these classes. Thus, our model has two layers of correlations and will be called *SSB-HKW correlated ensemble*, in honor of the authors that invented these correlation structures: Jeffrey Schenker, Hermann Schulz-Baldes w.r.t. [25] and Winfried Hochstättler, Werner Kirsch, Simone Warzel w.r.t. [16]. Second, as in [23], we introduce a weight function which will then allow us to study band and block matrices, and also allow for limits different from the semicircle distribution. Necessary and sufficient conditions for the semicircle law are provided and coincide with the conditions in [23]. Lastly, while the findings in [25] pertained to weak convergence in expectation, we show that their results also hold weakly in probability, and we derive mild sufficient conditions for the convergence results to also hold weakly almost surely. Examples that fit into the framework of our main Theorems 2.1 and 3.3 include Curie-Weiss and correlated Gaussian ensembles.

2 Random band matrices with approximately uncorrelated entries

2.1 Setup and results

Let $(a_n)_n$ be a sequence of random matrices, where a_n is a symmetric $n \times n$ matrix and the random variables $a_n(p, q)$ are real-valued. Using the notation $[\ell] := \{1, \dots, \ell\}$ for $\ell \in \mathbb{N}$, we will call pairs $(p_1, q_1), \dots, (p_\ell, q_\ell) \in [n]^2$ fundamentally different, if for all $1 \leq i \neq j \leq \ell : \{p_i, q_i\} \neq \{p_j, q_j\}$. Note that if $P_1, \dots, P_\ell \in [n]^2$ are fundamentally different, then $a_n(P_1), \dots, a_n(P_\ell)$ are pairwise different matrix elements, even modulo symmetry. We assume the following conditions: There exist sequences of constants $(C(k))_k$ and $(C_n^{(\ell)})_{n,\ell}$ such that for all $\ell \in \mathbb{N}$, $\lim_{n \rightarrow \infty} C_n^{(\ell)} = 0$ and such that for all $N, \ell \in \mathbb{N}$ and all fundamentally different pairs $P_1, \dots, P_\ell \in [N]^2$,

$$\forall \delta_1, \dots, \delta_\ell \in \mathbb{N} : \forall n \geq N : |\mathbb{E}a_n(P_1)^{\delta_1} \dots a_n(P_\ell)^{\delta_\ell}| \leq \frac{C(\delta_1 + \dots + \delta_\ell)}{n^{\frac{1}{2} \#\{i \in [\ell] | \delta_i = 1\}}} \tag{AU1}$$

$$\forall n \geq N : |\mathbb{E}a_n(P_1)^2 \dots a_n(P_\ell)^2 - 1| \leq C_n^{(\ell)} \tag{AU2}$$

We will call a sequence $(a_n)_n$ with the properties above an *approximately uncorrelated triangular scheme* (cf. [16], [13]). Note that (AU1) entails that all random variables $a_n(p, q)$ have uniformly bounded absolute moments of all orders, for example, $\mathbb{E}|a_n(p, q)|^k \leq (\mathbb{E}(a_n(p, q))^{2k})^{1/2} \leq C(2k)^{1/2}$. Further, by (AU1) and (AU2), the entries $a_n(p, q)$ need not have zero expectation nor unit variance, but these properties do hold asymptotically.

For all $n \in \mathbb{N}$, each element $b_n \in \{b \in \mathbb{N} | 1 \leq b < n, b \text{ odd}\} \cup \{n\}$ will be called (n) -bandwidth. If b_n is a bandwidth, we call an index pair $(p, q) \in [n]^2$ b_n -relevant, if $b_n = n$ or $|p - q| \leq (b_n - 1)/2$ or $|p - q| \geq n - (b_n - 1)/2$. In particular, if $b_n = n$, then all pairs $(p, q) \in [n]^2$ are b_n -relevant, but if $b_n < n$, (p, q) is b_n -relevant iff

$$|p - q| \leq \frac{b_n - 1}{2} \quad \text{or} \quad |p - q| \geq n - \frac{b_n - 1}{2}.$$

If b_n is a bandwidth, then an $n \times n$ matrix X will be called periodic band matrix with bandwidth b_n if $X(p, q) = 0$ for all $(p, q) \in [n]^2$ which are not b_n -relevant. For example, an 8×8 periodic band matrix X with bandwidth 5 has the structure

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & 0 & 0 & 0 & x_{1,7} & x_{1,8} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & 0 & 0 & 0 & x_{2,8} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} & x_{3,5} & 0 & 0 & 0 \\ 0 & x_{4,2} & x_{4,3} & x_{4,4} & x_{4,5} & x_{4,6} & 0 & 0 \\ 0 & 0 & x_{5,3} & x_{5,4} & x_{5,5} & x_{5,6} & x_{5,7} & 0 \\ 0 & 0 & 0 & x_{6,4} & x_{6,5} & x_{6,6} & x_{6,7} & x_{6,8} \\ x_{7,1} & 0 & 0 & 0 & x_{7,5} & x_{7,6} & x_{7,7} & x_{7,8} \\ x_{8,1} & x_{8,2} & 0 & 0 & 0 & x_{8,6} & x_{8,7} & x_{8,8} \end{pmatrix}.$$

The bandwidth can be interpreted as the number of allowable non-trivial entries in the “middle row” of the matrix. Given an approximately uncorrelated triangular scheme $(a_n)_n$ as above and bandwidths $(b_n)_n$, we define the periodic band matrices a_n^b by

$$\forall n \in \mathbb{N} : \forall P \in [n]^2 : a_n^b(P) := \begin{cases} a_n(P) & \text{if } P \text{ is } b_n\text{-relevant,} \\ 0 & \text{otherwise.} \end{cases}$$

Finally, if $(a_n)_n$ is an approximately uncorrelated triangular scheme, $(b_n)_n$ is a sequence of bandwidths, we say that a sequence of periodic random band matrices $(X_n)_n$ is based

on the triangular scheme $(a_n)_n$ with bandwidths $(b_n)_n$, if for all $n \in \mathbb{N}$, $X_n = \frac{1}{\sqrt{b_n}} a_n^b$. Note that by choosing the bandwidth $b_n = n$ we obtain $X_n = \frac{1}{\sqrt{n}} a_n$, so the case of full matrices is covered by our setup. The first contribution of this paper is the following theorem, which generalizes Theorem 2.3 in [13], since we obtain a stronger result with fewer assumptions.

Theorem 2.1. *Let $(X_n)_n$ be a sequence of periodic random band matrices which are based on an approximately uncorrelated triangular scheme $(a_n)_n$ with bandwidths $(b_n)_n$, where $b_n \rightarrow \infty$. Then the following statements hold:*

- i) *The semicircle law holds for $(X_n)_n$ in probability.*
- ii) *If there exists a $p \in \mathbb{N}$ such that b_n^{-p} is summable, and if for all $\ell \in \mathbb{N}$, the sequences $C_n^{(\ell)}$ are summable, then the semicircle law holds for $(X_n)_n$ almost surely.*

Remark 2.2. In Theorem 2.1 ii), if all entries in each a_n are $\{\pm 1\}$ -valued, the condition on the summability of the sequences $C_n^{(\ell)}$ can be dropped, since then in (AU2), $C_n^{(\ell)} := 0$ is a valid choice for these constants. Further, we point out that in particular, all bandwidths b_n with $b_n \sim n^\delta$ for some $\delta \in (0, 1)$ are covered by Theorem 2.1 ii).

2.2 Examples

We first note that Wigner matrices fit our framework, since their non-standardized versions satisfy (AU1) and (AU2). For the independent case, statements stronger than those in Theorem 2.1 are known, see [13] and references therein. The main contribution of Theorem 2.1 is thus for the truly dependent case. Therefore, we will now introduce two prominent ensembles of random matrices with dependent (even correlated) entries that fit our framework. These models will also serve as examples for the second part of this paper.

2.2.1 Curie-Weiss ensembles

Definition 2.3. *Let $n \in \mathbb{N}$ be arbitrary and Y_1, \dots, Y_n be random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $\beta > 0$, then we say that Y_1, \dots, Y_n are Curie-Weiss(β, n)-distributed, if for all $y_1, \dots, y_n \in \{-1, 1\}$ we have that*

$$\mathbb{P}(Y_1 = y_1, \dots, Y_n = y_n) = \frac{1}{Z_{\beta,n}} \cdot e^{\frac{\beta}{2n} (\sum y_i)^2},$$

where $Z_{\beta,n}$ is a normalization constant. The parameter β is called inverse temperature.

The Curie-Weiss(β, n) distribution is used to model the behavior of n ferromagnetic particles at inverse temperature β . At large values of β (low temperatures), all spins are likely to have the same alignment, modeling strong magnetization. At high temperatures, however, the spins can act almost independently, resembling weak magnetization. We refer the reader to [19] for a self-contained treatment of the Curie-Weiss distribution.

Definition 2.4. *Let $0 < \beta \leq 1$ and let the random variables $(\tilde{a}_n(i, j))_{1 \leq i, j \leq n}$ be Curie-Weiss(β, n^2)-distributed for each $n \in \mathbb{N}$. Define the triangular scheme $(a_n)_n$ by setting*

$$\forall n \in \mathbb{N} : \forall (i, j) \in [n]^2 : a_n(i, j) = \begin{cases} \tilde{a}_n(i, j) & \text{if } i \leq j \\ \tilde{a}_n(j, i) & \text{if } i > j. \end{cases}$$

Then $(a_n)_n$ will be called Curie-Weiss(β) ensemble.

Corollary 2.5. *Let $(a_n)_n$ be a Curie-Weiss ensemble at inverse temperature $\beta \in (0, 1]$. Let $(b_n)_n$ be a sequence of n -bandwidths and $(X_n)_n$ be the periodic random band matrices which are based on $(a_n)_n$ with bandwidths $(b_n)_n$. Then the following statements hold:*

- i) If $b_n \rightarrow \infty$, then the semicircle law holds for $(X_n)_n$ in probability.
- ii) If for some $p \in \mathbb{N}$, $\frac{1}{b_n^p}$ is summable over n , then the semicircle law holds almost surely for $(X_n)_n$.

Proof. In [16] and [19], it was shown that that $(a_n)_n$ is approximately uncorrelated with constants $C_n^{(\ell)} := 0$ and $C(\ell) := (\ell - 1)!!$ for $\beta \in (0, 1/2]$, $C(\ell) := (\ell - 1)!(\beta/(1 - \beta))^{\ell/2}$ for $\beta \in (1/2, 1)$ and $C(\ell) := 12^{\ell/4}\Gamma((\ell + 1)/4)/\Gamma(1/4)$ for $\beta = 1$, cf. Theorem 5.17 in [19]. Therefore, the statement follows with Theorem 2.1. \square

2.2.2 Correlated Gaussian ensembles

We impose the following assumptions on the triangular scheme $(a_n)_n$, where due to symmetry, we need only specify the upper right triangle:

1. $\forall n \in \mathbb{N} : (a_n(p, q))_{1 \leq p \leq q \leq n} \sim \mathcal{N}(0, \Sigma_n)$, where Σ_n is a positive definite $(n(n+1)/2) \times (n(n+1)/2)$ covariance matrix, indexed by pairs $(p, q), (r, s)$ with $1 \leq p \leq q \leq n$ and $1 \leq r \leq s \leq n$. In this case, we also set $\Sigma_n((q, p), (r, s)) := \Sigma_n((p, q), (r, s))$ and similarly if the second pair (or both pairs) is flipped to account for the symmetry of a_n .
2. We assume the sequence $(\Sigma_n)_n$ satisfies for all $n \in \mathbb{N}$ and all fundamentally different pairs $(p, q), (r, s) \in [n]^2$:

$$|\mathbb{E}a_n(p, q)a_n(r, s)| = |\Sigma_n((p, q), (r, s))| \leq \frac{1}{n}. \tag{2.1}$$

$$\mathbb{E}(a_n(p, q))^2 = |\Sigma_n((p, q), (p, q))| = 1. \tag{2.2}$$

Then we call $(a_n)_n$ an *approximately uncorrelated Gaussian ensemble*. For clarity reasons, the setup we chose here differs from the expositions in [9] and [13]. Note also that indexing the covariance matrices by pairs leads to no significant ambiguity, since positive definiteness and conditions (2.1) and (2.2) are not affected by the index order. But with this new indexing, we are able to give a very brief proof of the following lemma (cf. Lemmas 3.9 and 3.10 in [13]), which we include in Appendix A for the convenience of the reader.

Lemma 2.6. *If $(a_n)_n$ is an approximately uncorrelated Gaussian ensemble, then $(a_n)_n$ is an approximately uncorrelated triangular scheme. More precisely, it satisfies conditions (AU1) with constants $C(\ell) := \mathcal{P}_2(\ell)$ and (AU2) with sequences $C_n^{(\ell)} := \mathcal{P}_2(2\ell)/n^2$, which are summable over n .*

Corollary 2.7. *Let $(a_n)_n$ be an approximately uncorrelated Gaussian ensemble. Let $(b_n)_n$ be a sequence of n -bandwidths and $(X_n)_n$ be the periodic random band matrices which are based on $(a_n)_n$ with bandwidths $(b_n)_n$. Then the following statements hold:*

- i) If $b_n \rightarrow \infty$, then the semicircle law holds for $(X_n)_n$ in probability.
- ii) If for some $p \in \mathbb{N}$, $\frac{1}{b_n^p}$ is summable over n , then the semicircle law holds almost surely for $(X_n)_n$. Observe that this statement applies in particular to full matrices ($b_n = n$), where $p = 2$ may be chosen.

Proof. This is immediate with Lemma 2.6 and Theorem 2.1 \square

2.3 Proof of Theorem 2.1

Denote by $(\sigma_n)_n$ the ESDs of $(X_n)_n$ and by σ the semicircle distribution. In [13] the strategy of proof was to show that (where if Y is a random variable, we denote its variance by $\mathbb{V}Y$)

$$\forall k \in \mathbb{N} : \mathbb{E} \langle \sigma_n, x^k \rangle \xrightarrow{n \rightarrow \infty} \langle \sigma, x^k \rangle, \tag{2.3}$$

$$\forall k \in \mathbb{N} : \mathbb{V} \langle \sigma_n, x^k \rangle \xrightarrow{n \rightarrow \infty} 0. \tag{2.4}$$

Then (2.3) entails weak convergence in expectation and together with (2.4) this ensures weak convergence in probability. Further, if in (2.4) the decay is summably fast, then this entails weak convergence almost surely (see e.g. [9] for details).

However, in [10] it was shown (and also presented in [9]) that $\sigma_n \rightarrow \sigma$ weakly in expectation resp. in probability resp. almost surely if for all $k \in \mathbb{N}$: $\langle \sigma_n, x^k \rangle \rightarrow \langle \sigma, x^k \rangle$ in expectation resp. in probability resp. almost surely. Using this observation, we can refine the above proof strategy as follows: For each $k \in \mathbb{N}$ we identify a number $\ell \in \mathbb{N}$ and a decomposition

$$\langle \sigma_n, x^k \rangle = D_n^{(1)} + D_n^{(2)} + \dots + D_n^{(\ell)} \tag{2.5}$$

into finitely many summands (where ℓ is independent of n), so that these summands are amenable for analysis individually. For example, if we can prove that

$$\forall 1 \leq i \leq \ell - 1 : \mathbb{E} D_n^{(i)} \xrightarrow{n \rightarrow \infty} 0 \quad \text{and} \quad \mathbb{E} D_n^{(\ell)} \xrightarrow{n \rightarrow \infty} \mathcal{C}_{\frac{k}{2}} \mathbb{1}_{2\mathbb{N}}(k) \tag{2.6}$$

$$\forall 1 \leq i \leq \ell : \exists z \in \mathbb{N} : \mathbb{E} |D_n^{(i)} - \mathbb{E} D_n^{(i)}|^z \xrightarrow{n \rightarrow \infty} 0, \tag{2.7}$$

where \mathcal{C}_ℓ is the ℓ -th Catalan number, then (2.6) secures convergence in expectation of $\langle \sigma_n, x^k \rangle$ to $\langle \sigma, x^k \rangle$ and together with (2.7) this ensures convergence in probability of $\langle \sigma_n, x^k \rangle$ to $\langle \sigma, x^k \rangle$. Further, if all decays in (2.7) are summably fast, then we obtain convergence almost surely of $\langle \sigma_n, x^k \rangle$ to $\langle \sigma, x^k \rangle$. This strategy – although seemingly more complicated – has crucial advantages over the strategy pertaining to (2.3) and (2.4). First, note that a direct analysis of $\mathbb{V} \langle \sigma_n, x^k \rangle = \mathbb{V}(D_n^{(1)} + \dots + D_n^{(\ell)})$ requires the analysis of covariances $\text{Cov}(D_n^{(i)}, D_n^{(j)})$ for $i \neq j$. These mixed terms were responsible for serious problems in the analysis of [13], see Step 1/Case 1/Subcase 2 in the proof of their Theorem 2.3, for example, which led to further required conditions (e.g. (AAU3) in [13]). The summand-wise approach merely requires the analysis of the terms $\mathbb{E} D_n^{(i)}, \mathbb{V} D_n^{(i)}, i = 1, \dots, \ell$. The second advantage is that the individual summands $D_n^{(i)}$ – as we will see – can be chosen so that arbitrarily high central moments of $D_n^{(i)}$ are amenable for analysis (this was unwieldy when starting from $\langle \sigma_n, x^k \rangle$ due to the multitude of mixed terms emerging for high central moments). This high-moment analysis in turn allows us to lessen our requirements on the decay rate of the bandwidth $(b_n)_n$. Note that using the former approach as in [13], a higher moment analysis was not just unwieldy, but also not promising, see Remark 4.12 in [13]. But let us begin with the proof: We begin as usual by writing

$$\langle \sigma_n, x^k \rangle = \frac{1}{n} \text{tr} X_n^k = \frac{1}{nb_n^{\frac{k}{2}}} \sum_{\underline{t} \in [n]^k} a_n^b(\underline{t}) \tag{2.8}$$

where $\underline{t} = (t_1, \dots, t_k)$ and $a_n^b(\underline{t}) = a_n^b(t_1, t_2) a_n^b(t_2, t_3) \dots a_n^b(t_k, t_1)$. Note that in (2.8), whenever a pair $(t_\ell, t_{\ell+1})$ is not b_n -relevant, the summand vanishes. Thus, we call a tuple $\underline{t} \in [n]^k$ b_n -relevant, if each pair $(t_\ell, t_{\ell+1}), \ell = 1, \dots, k$, is b_n -relevant, and we set $[n]_b^k := \{\underline{t} \in [n]^k \mid \underline{t} \text{ is } b_n\text{-relevant}\}$. We arrive at

$$\langle \sigma_n, x^k \rangle = \frac{1}{nb_n^{\frac{k}{2}}} \sum_{\underline{t} \in [n]_b^k} a_n^b(\underline{t}). \tag{2.9}$$

Now we identify a tuple \underline{t} with its Eulerian graph $G_{\underline{t}} = (V_{\underline{t}}, E_{\underline{t}}, \phi_{\underline{t}})$ with vertices $V_{\underline{t}} := \{t_1, \dots, t_k\}$, abstract edges $E_{\underline{t}} = \{e_1, \dots, e_k\}$ and incidence function $\phi_{\underline{t}} : E_{\underline{t}} \rightarrow \{U \subseteq V_{\underline{t}} \mid \#U \in \{1, 2\}\}$, where $\phi_{\underline{t}}(e_\ell) = \{t_\ell, t_{\ell+1}\}$ for $\ell = 1, \dots, k$, where $k+1 \equiv 1$. Denote $\pi_\ell(\underline{t}) := \#\{\phi_{\underline{t}}(e) \mid e \text{ is an } \ell\text{-fold edge in } \underline{t}\}$, then we call the vector $\pi(\underline{t}) := (\pi_1(\underline{t}), \dots, \pi_k(\underline{t}))$ the *profile* of \underline{t} . Denote by $\Pi(k) := \{\pi(\underline{t}) \mid \underline{t} \in [n]^k, n \in \mathbb{N}\}$ the set of all possible profiles of k -tuples. Then clearly,

$$\#\Pi(k) \leq (k+1)^k \quad \text{and} \quad \forall \pi \in \Pi(k) : k = \sum_{\ell=1}^k \ell \cdot \pi_\ell. \tag{2.10}$$

For all $\pi \in \Pi(k)$, we define the following sets of tuples:

$$\begin{aligned} \mathcal{T}_n(\pi) &:= \{\underline{t} \in [n]_b^k \mid \pi(\underline{t}) = \pi\} \\ \mathcal{T}_n^d(\pi) &:= \{(\underline{s}, \underline{t}) \in ([n]_b^k)^2 \mid \pi(\underline{t}) = \pi = \pi(\underline{s}), \phi_{\underline{t}}(E_{\underline{t}}) \cap \phi_{\underline{s}}(E_{\underline{s}}) = \emptyset\} \\ \mathcal{T}_n^c(\pi) &:= \{(\underline{s}, \underline{t}) \in ([n]_b^k)^2 \mid \pi(\underline{t}) = \pi = \pi(\underline{s}), \phi_{\underline{t}}(E_{\underline{t}}) \cap \phi_{\underline{s}}(E_{\underline{s}}) \neq \emptyset\} \\ \mathcal{T}_n^{c,\ell}(\pi) &:= \{(\underline{s}, \underline{t}) \in ([n]_b^k)^2 \mid \pi(\underline{t}) = \pi = \pi(\underline{s}), \#(\phi_{\underline{t}}(E_{\underline{t}}) \cap \phi_{\underline{s}}(E_{\underline{s}})) = \ell\} \end{aligned}$$

where the last set is defined for all $\ell \in [k]$. To explain these sets, $\mathcal{T}_n(\pi)$ contains all $\underline{t} \in [n]_b^k$ with profile π , $\mathcal{T}_n^d(\pi)$ contains all edge-disjoint pairs $(\underline{s}, \underline{t})$, where $\underline{s}, \underline{t} \in [n]_b^k$ with profiles π , $\mathcal{T}_n^c(\pi)$ contains all such tuple pairs which share at least one edge and $\mathcal{T}_n^{c,\ell}(\pi)$ all those that share exactly ℓ edges.

In the following, we will say that a $\pi \in \Pi(k)$ admits an odd edge resp. only even edges, if $\pi_\ell \geq 1$ for some $\ell \in [k]$ odd resp. if $\pi_\ell = 0$ for all $\ell \in [k]$ odd. The following lemma holds (see [9] Lemmas 4.31, 4.33, 4.34 and 4.37.)

Lemma 2.8. *Let $n, k \in \mathbb{N}$ and let $\pi \in \Pi(k)$ be arbitrary.*

A) *Let $\underline{t} \in [n]_b^k$ be arbitrary, then*

- i) $\#V_{\underline{t}} \leq 1 + \pi_1(\underline{t}) + \dots + \pi_k(\underline{t})$.
- ii) *If \underline{t} contains at least one odd edge, then $\#V_{\underline{t}} \leq \pi_1(\underline{t}) + \dots + \pi_k(\underline{t})$.*

B) *Let b_n be a bandwidth and $\ell \in [k]$, then $\#\{\underline{t} \in [n]_b^k \mid \#V_{\underline{t}} \leq \ell\} \leq k^k n b_n^{\ell-1}$.*

C) *Let $\pi \in \Pi(k)$ be arbitrary, then*

- i) $\#\mathcal{T}_n(\pi) \leq k^k n b_n^{\pi_1 + \dots + \pi_k}$.
- ii) *If π admits an odd edge, then $\#\mathcal{T}_n(\pi) \leq k^k n b_n^{\pi_1 + \dots + \pi_k - 1}$.*
- iii) $\#\mathcal{T}_n^d(\pi) \leq (\#\mathcal{T}_n(\pi))^2$.
- iv) *If π admits only even edges, then $\#\mathcal{T}_n^c(\pi) \leq k^2 (2k)^{2k} n b_n^{k-1}$.*
- v) *If π admits at least one odd edge, then $\#\mathcal{T}_n^c(\pi) \leq k^2 (2k)^{2k} n b_n^{2(\pi_1 + \dots + \pi_k) - 2}$.*
- vi) *If π admits at least one odd edge, then we have for all $\ell = 1, \dots, k$:*

$$\#\mathcal{T}_n^{c,\ell}(\pi) \leq k^2 (2k)^{2k} n b_n^{2(\pi_1 + \dots + \pi_k) - \ell - 1}$$

We now sort the sum in (2.9) according to profiles $\pi \in \Pi(k)$ and get

$$\langle \sigma_n, x^k \rangle = \sum_{\pi \in \Pi(k)} \frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n(\pi)} a_n^b(\underline{t}). \tag{2.11}$$

We have now achieved a finite decomposition (since $\Pi(k)$ is a finite set) as in (2.5) and will proceed as outlined in (2.6) and (2.7), that is, for each $\pi \in \Pi(k)$, we will analyze the expectation and variance of

$$\frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n(\pi)} a_n^b(\underline{t}). \tag{2.12}$$

Case 1: π admits only even edges.

Subcase 1: $\pi = (0, k/2, 0, \dots, 0)$.

In this subcase, each $\underline{t} \in \mathcal{T}_n(\pi)$ consists of $k/2$ double edges. By Lemma 2.8, a \underline{t} in this set has at most $k/2 + 1$ vertices. We construct the more detailed subsets

$$\mathcal{T}_n^{\leq k/2}(\pi) := \{\underline{t} \in \mathcal{T}_n(\pi) \mid \#V_{\underline{t}} \leq k/2\} \quad \text{and} \quad \mathcal{T}_n^{k/2+1}(\pi) := \{\underline{t} \in \mathcal{T}_n(\pi) \mid \#V_{\underline{t}} = k/2 + 1\}.$$

Then $\mathcal{T}_n^{\leq k/2}(\pi) \leq k^k n b_n^{\frac{k}{2}-1}$ by Lemma 2.8. Thus,

$$\frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n^{\leq k/2}(\pi)} a_n^b(\underline{t}) \xrightarrow[n \rightarrow \infty]{} 0$$

in expectation and probability, and also almost surely if b_n^{-p} is summable for some $p \in \mathbb{N}$ by Lemma B.1. Further, by [9] p.81 f.,

$$\frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n^{k/2+1}(\pi)} a_n^b(\underline{t}) \xrightarrow[n \rightarrow \infty]{} \mathcal{C}_{\frac{k}{2}}.$$

in expectation, and the variance of each of this sum is clearly upper bounded by

$$\frac{1}{n^2 b_n^k} \sum_{(\underline{s}, \underline{t}) \in \mathcal{T}_n^d(\pi)} |\mathbb{E}a_n^b(\underline{s})a_n^b(\underline{t}) - \mathbb{E}a_n^b(\underline{s})\mathbb{E}a_n^b(\underline{t})| \tag{2.13}$$

$$+ \frac{1}{n^2 b_n^k} \sum_{(\underline{s}, \underline{t}) \in \mathcal{T}_n^c(\pi)} |\mathbb{E}a_n^b(\underline{s})a_n^b(\underline{t}) - \mathbb{E}a_n^b(\underline{s})\mathbb{E}a_n^b(\underline{t})|. \tag{2.14}$$

Considering $\#\mathcal{T}_n^d(\pi) \leq k^{2k} n^2 b_n^k$ by Lemma 2.8, the sum in (2.13) converges to zero by (AU1) and (AU2), and the decay is summably fast if $C_n^{(k/2)}$ decays to zero summably fast. Next, considering $\#\mathcal{T}_n^c(\pi) \leq k^2 (2k)^{2k} n b_n^{k-1}$ by Lemma 2.8, the sum (2.14) converges to zero, and this convergence is summably fast if $(n b_n)^{-1}$ is summable, which is the case if b_n^{-p} is summable for some $p \in \mathbb{N}$ (Young's inequality).

Subcase 2: $\pi_\ell \geq 1$ for some $\ell \geq 4$.

Then since $\pi_1 + \dots + \pi_k \leq 1 + (k - 4)/2 = k/2 - 1$, we obtain the bound $\#\mathcal{T}_n(\pi) \leq k^k n b_n^{k/2-1}$ by Lemma 2.8. Thus,

$$\frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n(\pi)} a_n^b(\underline{t}) \xrightarrow[n \rightarrow \infty]{} 0$$

in expectation and in probability, and also almost surely if b_n^{-p} is summable for some $p \in \mathbb{N}$ (Lemma B.1).

Case 2: $\pi_\ell \geq 1$ for some $\ell \in \mathbb{N}$ odd.

Then by Lemma 2.8, $\#\mathcal{T}_n(\pi) \leq k^k n b_n^{\pi_1 + \dots + \pi_k - 1}$. Further, by condition (AU1),

$$\forall \underline{t} \in \mathcal{T}_n(\pi) : |\mathbb{E}a_n^b(\underline{t})| \leq \frac{C(k)}{n^{\frac{1}{2}\pi_1}},$$

so

$$\frac{1}{n b_n^{\frac{k}{2}}} \sum_{\underline{t} \in \mathcal{T}_n(\pi)} |\mathbb{E}a_n^b(\underline{t})| \leq \frac{k^k C(k)}{n b_n^{\frac{k}{2}}} \cdot \frac{n b_n^{\pi_1 + \dots + \pi_k - 1}}{b_n^{\frac{1}{2}\pi_1}} \leq \frac{k^k C(k)}{b_n}, \tag{2.15}$$

using $\pi_1/2 + \pi_2 + \dots + \pi_k \leq k/2$ by (2.10). Since $b_n \rightarrow \infty$, convergence in expectation to zero follows. Next, instead of analyzing the variance of the sum in question, we analyze

an arbitrary high even central moment. To this end, let $z \in 2\mathbb{N}$ be arbitrary, then

$$\begin{aligned}
 & \mathbb{E} \left(\frac{1}{nb_n^{\frac{z}{2}}} \sum_{\underline{t} \in \mathcal{T}_n(\pi)} (a_n^b(\underline{t}) - \mathbb{E}a_n^b(\underline{t})) \right)^z \\
 &= \left| \frac{1}{n^z b_n^{\frac{zk}{2}}} \sum_{\underline{t}^{(1)}, \dots, \underline{t}^{(z)} \in \mathcal{T}_n(\pi)} \mathbb{E} \prod_{s=1}^z [a_n^b(\underline{t}^{(s)}) - \mathbb{E}a_n^b(\underline{t}^{(s)})] \right| \\
 &\leq \sum_{S \subseteq [z]} \frac{1}{n^z b_n^{\frac{zk}{2}}} \sum_{\underline{t}^{(1)}, \dots, \underline{t}^{(z)} \in \mathcal{T}_n(\pi)} \left| \mathbb{E} \prod_{s \in S} a_n^b(\underline{t}^{(s)}) \right| \cdot \prod_{s \notin S} |\mathbb{E}a_n^b(\underline{t}^{(s)})| \\
 &= \underbrace{\sum_{S \subseteq [z]} \frac{1}{n^{\#S} b_n^{\frac{k\#S}{2}}} \sum_{\substack{\underline{t}^{(i)} \in \mathcal{T}_n(\pi) \\ i \in S}} \left| \mathbb{E} \prod_{i \in S} a_n^b(\underline{t}^{(i)}) \right|}_{=: A_n(S)} \cdot \underbrace{\frac{1}{n^{\#S^c} b_n^{\frac{k\#S^c}{2}}} \sum_{\substack{\underline{t}^{(j)} \in \mathcal{T}_n(\pi) \\ j \in S^c}} \prod_{j \in S^c} |\mathbb{E}a_n^b(\underline{t}^{(j)})|}_{=: B_n(S)},
 \end{aligned} \tag{2.16}$$

where for any $S \subseteq [z]$, $S^c := [z] \setminus S$. We will analyze the terms $A_n(S)$ and $B_n(S)$ separately. Notationally, we write $s := \#S$ and $s^c := \#S^c$ and note that $s + s^c = z$. For $B_n(S)$ we find an upper bound using (2.15):

$$B_n(S) \leq \frac{(k^k C(k))^{s^c}}{b_n^{s^c}}.$$

For $A_n(S)$ we need to account for common edges among the $\underline{t}^{(i)}$, $i \in S$. These have the effect that on the one hand, overlaps lead to fewer possible tuples, but on the other hand, single edges might overlap, negating a possible decay which existed due to (AU1). If $S = \emptyset$, then $A_n(S) = 1$ as an empty product. If $S = \{i\}$ for some $i \in [z]$, then by (2.15),

$$A_n(S) \leq \frac{k^k C(k)}{b_n}, \quad \text{so} \quad A_n(S)B_n(S) \leq \frac{(k^k C(k))^z}{b_n^z} \tag{2.17}$$

in both cases $\#S = 0$ and $\#S = 1$. We now assume that $s = \#S \geq 2$. Write $S = \{i_1, \dots, i_s\}$, where $i_j \in [z]$ and $i_1 < \dots < i_s$. Then for all $\ell_2, \dots, \ell_s \in \{0, 1, \dots, \sum_i \pi_i\}$ we denote by $\mathcal{T}_n^{(\ell)}(\pi)$ the set of tuples $(\underline{t}^{(i_1)}, \dots, \underline{t}^{(i_s)})$, where $\underline{t}^{(i_j)} \in \mathcal{T}_n(\pi)$ and

$$\forall j \in \{2, \dots, s\} : \#\phi_{\underline{t}^{(i_j)}}(E_{\underline{t}^{(i_j)}}) \cap \bigcup_{1 \leq r < j} \phi_{\underline{t}^{(i_r)}}(E_{\underline{t}^{(i_r)}}) = \ell_j,$$

in words, each $\underline{t}^{(i_j)}$ has exactly ℓ_j of its $\sum_i \pi_i$ different edges in common with previous tuples $\underline{t}^{(i_r)}$, $r < j$. Then

$$A_n(S) = \sum_{\ell \in \{0, \dots, \sum_i \pi_i\}^{s-1}} \frac{1}{n^s b_n^{\frac{ks}{2}}} \underbrace{\sum_{(\underline{t}^{(i_j)})_{j \in [s]} \in \mathcal{T}_n^{(\ell)}(\pi)} \left| \mathbb{E} \prod_{j=1}^s a_n^b(\underline{t}^{(i_j)}) \right|}_{=: A_n^{(\ell)}(S)}.$$

If $\ell = (0, \dots, 0)$, this entails that all tuples are edge-disjoint, so we use the trivial upper bound

$$\#\mathcal{T}_n^{(\ell)}(\pi) \leq (\#\mathcal{T}_n(\pi))^s \leq (k^k n b_n^{\pi_1 + \dots + \pi_k - 1})^s,$$

and since no single edges may vanish due to overlaps, we obtain by (AU1)

$$\left| \mathbb{E} \prod_{j=1}^s a_n^b(\underline{t}^{(i_j)}) \right| \leq \frac{C(ks)}{n^{\frac{1}{2} \pi_1 s}}, \tag{2.18}$$

so

$$A_n^{(\underline{\ell})}(S) \leq \frac{1}{n^s b_n^{\frac{k}{2}s}} (k^k n b_n^{\pi_1 + \dots + \pi_k - 1})^s \cdot \frac{C(ks)}{n^{\frac{1}{2}\pi_1 s}} \leq \frac{k^{ks} C(ks)}{b_n^s},$$

and subsequently

$$A_n^{(\underline{\ell})}(S) \cdot B_n(S) \leq \frac{k^{ks} C(ks)}{b_n^s} \cdot \frac{(k^k C(k))^{s^c}}{b_n^{s^c}} = \frac{k^{zk} C(k)^{s^c} C(ks)}{b_n^z}.$$

If $\underline{\ell} \neq (0, \dots, 0)$, then some of the tuples in $(\underline{t}^{(i_j)})_{j \in [s]}$ have common edges. As mentioned, this has two effects: On the one hand, the corresponding product on the l.h.s. of (2.18) cannot be guaranteed to decay at the speed given on the r.h.s. of (2.18), since single edges in different tuples might overlap, negating the decay effect guaranteed by (AU1). To be more precise, for each overlap at most two single edges may be eradicated, leading to at least $s \cdot \pi_1 - 2\ell_2 - 2\ell_3 \dots - 2\ell_s$ remaining single random variables in the product on the l.h.s. of (2.18). So if $\underline{\ell} \neq (0, \dots, 0)$, then (2.18) becomes

$$\left| \mathbb{E} \prod_{j=1}^s a_n^b(\underline{t}^{(i_j)}) \right| \leq \frac{C(ks)}{n^{\frac{1}{2} \max(s\pi_1 - 2\ell_2 - \dots - 2\ell_s, 0)}} \tag{2.19}$$

The second effect is that overlaps of edges entail fewer possible vertices, which decreases the count $\#\mathcal{T}_n^{(\underline{\ell})}(\pi)$:

Lemma 2.9. *Let $k, n \in \mathbb{N}$ and $\pi \in \Pi(k)$ be arbitrary so that $\pi_i \neq 0$ for some $i \in [k]$ odd. Further, let $\underline{\ell} \in \{0, \dots, \sum_i \pi_i\}^{s-1}$ be arbitrary for some $s \geq 2$, where $\underline{\ell} \neq (0, \dots, 0)$. Then*

$$\#\mathcal{T}_n^{(\underline{\ell})}(\pi) \leq k^k n b_n^{\sum_i \pi_i - 1} \cdot \prod_{\substack{i \in \{2, \dots, s\} \\ \ell_i = 0}} k^k n b_n^{\sum_i \pi_i - 1} \cdot \prod_{\substack{i \in \{2, \dots, s\} \\ \ell_i \geq 1}} (2k)^k k(ks)^{\ell_i} b_n^{\pi_1 + \dots + \pi_k - \ell_i}.$$

Proof. The strategy of the proof is to derive an upper bound on the number of possibilities to construct an element $(\underline{t}^{(1)}, \dots, \underline{t}^{(s)}) \in \mathcal{T}_n^{(\underline{\ell})}(\pi)$.

We proceed step by step: For $\underline{t}^{(1)}$ we have at most $k^k n b_n^{\pi_1 + \dots + \pi_k - 1}$ possibilities by Lemma 2.8. Of course, the bound also applies for all $\underline{t}^{(i)}$ with $i \geq 2$, but we will only use it again for those $\underline{t}^{(i)}$, $i \geq 2$, for which $\ell_i = 0$.

Now assume that $i \geq 2$ with $\ell_i \geq 1$. Note that so far, at most $(i - 1) \cdot k \leq sk$ vertices have been picked for previous tuples. Each tuple has at most $\pi_1 + \dots + \pi_k$ vertices by Lemma 2.8. We now obtain an upper bound on the new vertices that $\underline{t}^{(i)}$ may contain. To this end, we start a cyclic tour along the tuple $\underline{t}^{(i)}$, starting at an end point $t_{j+1}^{(i)}$ of an edge e_j which is a common edge with a previously determined tuple. Then $t_{j+1}^{(i)}$ is not a newly observed vertex, and as we proceed cyclicly along the tuple, we can observe at most $\pi_1 + \dots + \pi_k - \ell_i$ new vertices.

We now bound the possibilities to construct a $\underline{t} \in [n]_b^k$ with at most $\pi_1 + \dots + \pi_k$ vertices, from which at most $\pi_1 + \dots + \pi_k - \ell_i$ vertices are new and all other vertices are old (i.e. appeared in some previous tuple), but at least one vertex must be old (since we have at least one overlap). Since there is at least one old vertex, we start with such a vertex for t_1 and in the end allow a cyclic permutation (e.g. $(1, 2, 3, 4) \rightarrow (3, 4, 1, 2)$) of the tuple to count all possibilities. For the construction of \underline{t} we first fix a map $f : \{1, \dots, k\} \rightarrow \{1, \dots, \pi_1 + \dots + \pi_k\}$ indicating which places in $\underline{t}^{(i)}$ should receive equal or different vertices, i.e. $t_i = t_j \Leftrightarrow f(i) = f(j)$. We assume that the coloring f has standard form, that is, $f(1) = 1$ and if $f(\ell) \notin \{f(1), \dots, f(\ell - 1)\}$, then $f(\ell) = \max\{f(1), \dots, f(\ell - 1)\} + 1$. This choice of f admits at most k^k possibilities. Note that $\max f([k])$ is the number of different vertices in \underline{t} . To determine which of these should be new and which should be old, we fix another map $g : \{1, \dots, \max f([k])\} \rightarrow \{0, 1\}$

with $g(1) = 0$ and $\#\{i \in \{1, \dots, \max f([k])\} \mid g(1) = 1\} \leq \pi_1 + \dots + \pi_k - \ell_i$, which admits at most 2^k possibilities. We then proceed as follows: For t_1 we choose an old vertex arbitrarily, yielding at most $(i - 1)k \leq sk$ choices. Then if t_1, \dots, t_ℓ have already been constructed, where $\ell \in \{1, \dots, k - 1\}$, we construct $t_{\ell+1}$ as follows: If $f(\ell + 1) = f(\ell')$ for some $\ell' \in \{1, \dots, \ell\}$, then $t_{\ell+1}$ must equal $t_{\ell'}$, leaving no choice for $t_{\ell+1}$. If $f(\ell + 1) \notin \{f(1), \dots, f(\ell)\}$, choose $t_{\ell+1}$ different from all previous vertices in \underline{t} . If $g(f(\ell + 1)) = 0$, pick an old vertex from some previous tuple, yielding at most sk possibilities. If $g(f(\ell + 1)) = 1$, pick a new vertex, yielding at most b_n possibilities. This procedure to construct \underline{t} with help of f and g admits at most $(ks)^{\ell_i} b_n^{\pi_1 + \dots + \pi_k - \ell_i}$ possibilities. Cyclic permutation of this tuple admits at most k choices, picking f and g at most $(2k)^k$ choices, so all together we had at most $(2k)^k k (ks)^{\ell_i} b_n^{\pi_1 + \dots + \pi_k - \ell_i}$ choices for \underline{t} . Since $\underline{t}^{(i)}$ is of this form, this concludes the proof. \square

Combining Lemma 2.9 with (2.19) and the prefactor in the definition of $A_n^{(\ell)}(S)$ yields, setting $\ell(0) := \#\{i \in \{2, \dots, s\} \mid \ell_i = 0\}$ and $\ell(\geq 1) := \#\{i \in \{2, \dots, s\} \mid \ell_i \geq 1\}$,

$$\begin{aligned} A_n^{(\ell)}(S) &\leq \frac{1}{n^s b_n^{\frac{k}{2}s}} k^k n b_n^{\sum_i \pi_i - 1} \cdot \prod_{\substack{i \in \{2, \dots, s\} \\ \ell_i = 0}} k^k n b_n^{\sum_i \pi_i - 1} \\ &\cdot \prod_{\substack{i \in \{2, \dots, s\} \\ \ell_i \geq 1}} (2k)^k k (ks)^{\ell_i} b_n^{\sum_i \pi_i - \ell_i} \cdot \frac{C(ks)}{n^{\frac{1}{2} \max(s\pi_1 - 2\ell_2 - \dots - 2\ell_s, 0)}} \\ &\leq \frac{K(k, s)}{n^{s - \ell(0) - 1} b_n^{\frac{k}{2}s}} b_n^{(\sum_i \pi_i - 1)(\ell(0) + 1)} \cdot b_n^{(\sum_i \pi_i)(\ell(\geq 1))} \frac{1}{b_n^{\max(s\pi_1/2, \ell_2 + \dots + \ell_s)}}, \end{aligned}$$

where

$$K(k, s) := k^{s^{k+1}} (2k)^{ks} s^{ks} C(ks) \geq k^{k+k\ell(0)+1+k\ell(\geq 1)} (2k)^{k\ell(\geq 1)} s^{k\ell(\geq 1)} C(ks),$$

for which we used $\ell(0) + \ell(\geq 1) + 1 = s$. This inequality together with $\pi_1 + \dots + \pi_k \leq \pi_1 + (k - \pi_1)/2 = \pi_1/2 + k/2$ yields

$$A_n^{(\ell)}(S) \leq \frac{K(k, s)}{n^{s - \ell(0) - 1}} \cdot \frac{b_n^{\frac{s\pi_1}{2} - \ell(0) - 1}}{b_n^{\max(s\pi_1/2, \ell_2 + \dots + \ell_s)}} \leq \frac{K(k, s)}{b_n^s},$$

and so

$$A_n^{(\ell)}(S) B_n(S) \leq \frac{K(k, s)}{b_n^s} \cdot \frac{(k^k C(k))^{s^c}}{b_n^{s^c}} = \frac{K(k, s) (k^k C(k))^{s^c}}{b_n^z}. \tag{2.20}$$

We have recognized the z -th central moment in (2.16) as a finite sum of terms $A_n(S) B_n(s)$ and $A_n^{(\ell)}(S) B_n(S)$, where the number of these terms is independent of n , and so that each summand decays at a speed of b_n^{-z} . This completes the proof of parts *i*) and *ii*) of Theorem 2.1 when choosing $z \geq p$ with $z \in 2\mathbb{N}$.

2.4 Non-periodic band matrices

In this section, we will see that Theorem 2.1 remains true for *non-periodic* random matrices with approximately uncorrelated entries. To start, the concept of a bandwidth should be replaced by the concept called *halfwidth*, which we adopted from [21]. Roughly, the halfwidth $h = (h_n)_n$ is half of the bandwidth $b = (b_n)_n$, hence the name. Two 6×6

non-periodic band matrices with halfwidths 2 resp. 4 have the structure

$$\begin{pmatrix} x_{1,1} & x_{1,2} & 0 & 0 & 0 & 0 \\ x_{2,1} & x_{2,2} & x_{2,3} & 0 & 0 & 0 \\ 0 & x_{3,2} & x_{3,3} & x_{3,4} & 0 & 0 \\ 0 & 0 & x_{4,3} & x_{4,4} & x_{4,5} & 0 \\ 0 & 0 & 0 & x_{5,4} & x_{5,5} & x_{5,6} \\ 0 & 0 & 0 & 0 & x_{6,5} & x_{6,6} \end{pmatrix}, \text{ resp. } \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & 0 & 0 \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & x_{2,5} & 0 \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} & x_{3,5} & x_{3,6} \\ x_{4,1} & x_{4,2} & x_{4,3} & x_{4,4} & x_{4,5} & x_{4,6} \\ 0 & x_{5,2} & x_{5,3} & x_{5,4} & x_{5,5} & x_{5,6} \\ 0 & 0 & x_{6,3} & x_{6,4} & x_{6,5} & x_{6,6} \end{pmatrix}.$$

The halfwidth should be interpreted as the number of allowable non-trivial entries in the first row of the matrix. As before, the bandwidth describes the number of such entries in the “middle row”.

Definition 2.10. Let $n \in \mathbb{N}$ be arbitrary, then an $h_n \in \mathbb{N}$ is called (n) -halfwidth, if $h_n \in \{1 \dots n\}$. Given a sequence of halfwidths $h = (h_n)_n$, we set

$$\forall n \in \mathbb{N} : b_n := \min(2h_n - 1, n)$$

and call b_n the bandwidth associated with the halfwidth h_n .

It is clear that for any $n \in \mathbb{N}$, the bandwidth b_n that is associated with a halfwidth h_n is either n itself or an odd number in the set $\{1, \dots, n\}$, thus coincides with the concept of a bandwidth in previous sections.

The difference between periodic and non-periodic matrices is that in the latter case, the triangular areas in the upper right and lower left corner of the matrices are missing, leading to the possibility that the *inner band* is so wide that it reaches the top right and lower left corners of the matrix.

Definition 2.11. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $(a_n)_{n \in \mathbb{N}}$ a triangular scheme, $(h_n)_n$ be a sequence of n -halfwidths with associated bandwidths $(b_n)_n$.

1. We define the non-periodic random matrices which are based on the triangular scheme $(a_n)_{n \in \mathbb{N}}$ with halfwidths $(h_n)_n$ as

$$\forall n \in \mathbb{N} : \forall (i, j) \in [n]^2 : X_n^{NP}(i, j) := \begin{cases} \frac{1}{\sqrt{b_n}} a_n(i, j) & \text{if } |i - j| \leq h_n - 1, \\ 0 & \text{otherwise.} \end{cases}$$

2. We define the periodic random matrices which are based on the triangular scheme $(a_n)_{n \in \mathbb{N}}$ with associated bandwidth b as

$$\forall n \in \mathbb{N} : \forall (i, j) \in [n]^2 : X_n^P(i, j) := \begin{cases} \frac{1}{\sqrt{b_n}} a_n(i, j) & \text{if } |i - j| \leq h_n - 1, \\ \frac{1}{\sqrt{b_n}} a_n(i, j) & \text{if } |i - j| \geq \max(n - h_n + 1, h_n), \\ 0 & \text{otherwise.} \end{cases}$$

Note that the definition of periodic random matrices has not changed in comparison to previous sections as it is not hard to check that if h_n is an n -halfwidth with associated bandwidth b_n , then an index pair $(p, q) \in [n]^2$ is b_n -relevant iff $|p - q| \leq h_n - 1$ or $|p - q| \geq \max(n - h_n + 1, h_n)$.

In [5] it was shown that for the i.i.d. case, the semicircle law holds in probability for $(X_n^{NP})_n$ if

$$\lim_{n \rightarrow \infty} h_n = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{h_n}{n} \in \{0, 1\}, \tag{2.21}$$

whereas the semicircle law does not hold if $\lim_n h_n/n = p$ for some $p \in (0, 1)$. The analysis of this subsection derives the case (2.21) for the approximately uncorrelated

setup. The case that $\lim_n h_n/n = p$ for some $p \in (0, 1)$ is a corollary of our treatment in the second part of this paper.

Comparing non-periodic and periodic band matrices given some halfwidth h_n and associated bandwidth b_n , we realize that both matrices contain a non-trivial area with indices $|i - j| \leq h_n - 1$, which is the band in the middle of the matrix, and additionally, periodic matrices contain non-trivial triangular areas with indices $|i - j| \geq \max(n - h_n + 1, h_n)$. Therefore, the matrix $X_n^P - X_n^{NP}$ has rank $2 \cdot \min(h_n, n - h_n + 1)$. We now use a well-known rank inequality (e.g. [13]):

Lemma 2.12. *Let Y and \mathcal{E} be real symmetric $n \times n$ matrices, where \mathcal{E} has rank k . Then it holds for all $z \in \mathbb{C}_+ := \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$:*

$$\left| \frac{1}{n} \text{tr} [(Y - z)^{-1}] - \frac{1}{n} \text{tr} [(Y + \mathcal{E} - z)^{-1}] \right| \leq \frac{2k}{n \text{Im}(z)}.$$

With this lemma, the following theorem follows directly from Theorem 2.1:

Theorem 2.13. *Let $(a_n)_n$ be an approximately uncorrelated triangular scheme, $(h_n)_n$ a sequence of n -halfwidths and $(X_n^{NP})_n$ the non-periodic random matrices which are based on $(a_n)_n$ with halfwidth h . We assume that*

$$h_n \rightarrow \infty \quad \text{but} \quad \lim_{n \rightarrow \infty} \frac{h_n}{n} \in \{0, 1\}.$$

Then we obtain the following results:

- i) *The semicircle law holds for $(X_n^{NP})_n$ in probability.*
- ii) *If $(1/h_n)_n^p$ is summable for some $p \in \mathbb{N}$, and the sequences $(C_n^{(\ell)})_n$ from condition (AU2) are summable for all $\ell \in \mathbb{N}$, then the semicircle law holds for X_n^{NP} almost surely.*

Proof. Note that $h_n \rightarrow \infty \Leftrightarrow b_n \rightarrow \infty$ and for all $p \in \mathbb{N}$, $(h_n^{-p})_n$ is summable iff $(b_n^{-p})_n$ is summable. Therefore, under the conditions of Theorem 2.13 i) resp. ii), the conclusions of Theorem 2.1 i) resp. ii) hold. It follows with the discussion before Lemma 2.12 that $X_n^P - X_n^{NP}$ has rank $2 \cdot \min(h_n, n - h_n + 1)$. Therefore, if s_n^P resp. s_n^{NP} denote the Stieltjes transforms of the ESDs of X_n^P resp. X_n^{NP} , then we find with Lemma 2.12 that for $z \in \mathbb{C}_+$ arbitrary,

$$|s_n^P(z) - s_n^{NP}(z)| \leq \frac{2}{\text{Im}(z)} \min\left(\frac{h_n}{n}, 1 - \frac{h_n}{n} + \frac{1}{n}\right) \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{surely.}$$

This concludes the proof, since under the conditions of Theorem 2.13 i) resp. ii), $s_n^P(z)$ converges to $s^\sigma(z)$ in probability resp. almost surely, where s^σ denotes the Stieltjes transform of the semicircle distribution. □

Corollary 2.14. *Let $(a_n)_n$ be a Curie-Weiss ensemble with inverse temperature $\beta \in (0, 1]$ (cf. Section 2.2.1) or an approximately uncorrelated Gaussian ensemble (cf. Section 2.2.2). Let $h = (h_n)_n$ be a sequence of n -halfwidths with $h_n \rightarrow \infty$ and $\lim_n h_n/n \in \{0, 1\}$. Let $(X_n^{NP})_n$ be the non-periodic random band matrices which are based on $(a_n)_n$ with halfwidth h . Then the following statements hold:*

- i) *The semicircle law holds for $(X_n^{NP})_n$ in probability.*
- ii) *If $\frac{1}{h_n^p}$ is summable over n for some $p \in \mathbb{N}$, then the semicircle law holds almost surely for $(X_n^{NP})_n$.*

Proof. This is a direct consequence Theorem 2.13 i) and ii), the proof of Corollary 2.5 and Lemma 2.6. □

3 Weighted ensembles with two layers of correlation

3.1 Setup and results

We consider the following setup: We assume that $(a_n)_n$ is a sequence of real-symmetric $n \times n$ random matrices. We do not assume the families of random variables $(a_n(i, j))_{1 \leq i \leq j \leq n}$ to be independent, nor do we require them to be standardized. Rather, we allow arbitrarily high correlation (even equality) of random variables which belong to certain subfamilies, and that random variables from different subfamilies are approximately uncorrelated. To make this precise, we assume that for all $n \in \mathbb{N}$, \sim_n is an equivalence relation on $[n]^2$ which satisfies the following conditions (when n can be derived from the context, we write \sim instead \sim_n): There exists a $B \in \mathbb{N}$ independent of n such that

$$\begin{aligned} (E1) \quad & \max_{p \in [n]} \# \{(q, r, s) \in [n]^3 \mid (p, q) \sim (r, s)\} = o(n^2), \\ (E2) \quad & \max_{p, q, r \in [n]} \# \{s \in [n] \mid (p, q) \sim (r, s)\} \leq B, \\ (E3) \quad & \# \{(p, q, r) \in [n]^3 \mid (p, q) \sim (q, r) \wedge r \neq p\} = o(n^2). \end{aligned}$$

These are exactly the same conditions as (C1), (C2) and (C3) in [25]. For some of our results, we also require the following conditions: There exists a fixed $\delta > 0$ independent of n such that

$$\begin{aligned} (E1') \quad & \max_p \# \{(q, r, s) \in [n]^3 \mid (p, q) \sim (r, s)\} = O(n^{2-\delta}), \\ (E3') \quad & \# \{(p, q, r) \in [n]^3 \mid (p, q) \sim (q, r) \wedge r \neq p\} = O(n^{2-\delta}). \end{aligned}$$

We observe that $(E1')$ and $(E3')$ are slightly stronger than their counterparts $(E1)$ and $(E3)$. The stronger conditions will be used to derive almost sure convergence results.

In the setup of [25], the entries of a_n were assumed to be standardized, have uniformly bounded absolute moments of all orders, and that \sim was required to satisfy $(E1)$, $(E2)$, $(E3)$. Further, it was assumed that the families

$$(a_n(P))_{P \in M}, \quad M \in [n]^2 / \sim \tag{3.1}$$

be independent while no independence requirement was made for members of the same equivalence class (for example, they could be all the same random variable). Due to independence between different equivalence classes it was also necessary to assume that $(p, q) \sim (q, p)$ for all $(p, q) \in [n]^2$, since the matrices a_n are symmetric.

In our setup, we also assume that \sim satisfies $(E1)$, $(E2)$ and $(E3)$ (and for some results $(E1')$, $(E2)$ and $(E3')$) and that $(p, q) \sim (q, p)$ for all $(p, q) \in [n]^2$. However, we drop the standardization requirement and the requirement of independence between equivalence classes. We instead require entries from different equivalence classes to be approximately uncorrelated in the sense of Section 2.1: Let $\ell, s \geq 0$ be arbitrary, $P_1, \dots, P_s, Q_1, \dots, Q_\ell \in [n]^2$ be distinct index pairs, where P_1, \dots, P_s stem from distinct \sim_n -equivalence classes, and let $\delta_1, \dots, \delta_\ell \in \mathbb{N}$, then

$$|\mathbb{E} a_n(P_1) \cdots a_n(P_s) a_n(Q_1)^{\delta_1} \cdots a_n(Q_\ell)^{\delta_\ell}| \leq \frac{C(s + \delta_1 + \dots + \delta_\ell)}{n^{\frac{s}{2}}}, \tag{A1}$$

$$|\mathbb{E} a_n(P_1)^2 \cdots a_n(P_s)^2 - 1| \leq C_n^{(s)}, \tag{A2}$$

where $C(s)$ resp. $(C_n^{(s)})_n$ are constants resp. sequences for all $s \in \mathbb{N}$, where for all s , $C_n^{(s)} \rightarrow 0$ as $n \rightarrow \infty$. Note that (A1) implies that all entries in a_n have uniformly bounded (absolute) moments of all orders and (A1) and (A2) together imply that all

entries in a_n are asymptotically standardized. Clearly, the case of standardized entries with independent families (3.1) – the setup of [25] – is included in above setup as a special case.

Definition 3.1. A triangular scheme $(a_n)_n$ satisfying (A1) and (A2), where equivalence within $[n]^2$ is governed by a relation \sim_n satisfying (E1), (E2), and (E3) (or (E1'), (E2) and (E3')) when this is explicitly stated), will be called SSB-HKW correlated. The acronym honors the authors of [25, 16].

Next, we assume that $w : [0, 1] \rightarrow \mathbb{R}$ is a Riemann integrable weight function. In particular, w is bounded, $|w| \leq W$ for some $W \in \mathbb{R}_+$. We consider weighted matrices of the form

$$X_n := \frac{1}{\sqrt{n}} \left[w \left(\frac{|i-j|}{n} \right) a_n(i, j) \right]_{1 \leq i, j \leq n}. \tag{3.2}$$

We observe that within the diagonals of X_n , the same weight is applied. This could be generalized, as has been done in [27] via graphon theory. However, our study is motivated primarily by band matrices which makes our modeling natural. Throughout the remainder of this text, we will make use of partitions of the index set $[k] = \{1, 2, \dots, k\}$. Therefore, we introduce the following definition.

Definition 3.2. Let $k \in \mathbb{N}$ be fixed. Then we denote by $\mathcal{P}(k)$ the set of partitions of the set $[k]$. A partition $\pi \in \mathcal{P}(k)$ is called pair partition if each block $B \in \pi$ satisfies $\#B = 2$. We denote by $\mathcal{P}_2(k) \subseteq \mathcal{P}(k)$ the set of pair partitions of the set $[k]$. We call a pair partition $\pi \in \mathcal{P}_2(k)$ crossing if there are $1 \leq a < b < c < d \leq k$ such that $\{a, c\}, \{b, d\} \in \pi$. Otherwise, π is called non-crossing. We denote by $\mathcal{NC}_2(k) \subseteq \mathcal{P}_2(k)$ the subset of all non-crossing pair partitions.

Under the setup we just described, we now formulate our main theorem, which summarizes all the results of this second part of the paper:

Theorem 3.3. Let σ_n be the ESD of X_n as in (3.2), where $(a_n)_n$ is SSB-HKW correlated. Then σ_n converges weakly in probability to a symmetric and compactly supported probability measure μ on $(\mathbb{R}, \mathcal{B})$, which is uniquely determined by its moments

$$\forall k \in \mathbb{N} : \langle \mu, x^k \rangle = \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi),$$

where the constants $J_w(\pi)$ depend on the weight function w and the partition π , and can be calculated recursively as described in Lemma 3.11 below. Further, set

$$\forall x \in [0, 1] : \varphi(x) := \int_0^1 w^2(|x-y|) dy$$

and $\varphi_0 := \int_0^1 \varphi(x) dx$. Then the limiting variance is given by φ_0 , i.e. $\langle \mu, x^2 \rangle = \varphi_0$, and can be calculated by

$$\varphi_0 = 2 \int_0^1 (1-x)w^2(x) dx.$$

In particular, $\varphi_0 = 0$ if and only if $w = 0$ on $[0, 1]$ \mathbb{N} -almost surely. In the case that $\varphi_0 > 0$, then the following statements are equivalent:

- a) The semicircle law holds for $\frac{1}{\sqrt{\varphi_0}} X_n$ in probability.
- b) φ is constant, in particular, $\varphi \equiv \varphi_0$.
- c) $\varphi \equiv \int_0^1 w^2(x) dx$,
- d) w^2 is \mathbb{N} -a.s. symmetric around $1/2$, i.e. $w^2(x) = w^2(1-x)$ for \mathbb{N} -a.a. $x \in [0, 1]$.

Further, if (E1) and (E3) are replaced by (E1') and (E3'), all statements about weak convergence in probability can be replaced by weak convergence almost surely.

3.2 Examples

In this subsection we study various types of examples and counterexamples which illustrate the reach of Theorem 3.3.

3.2.1 Repetitions of approximately uncorrelated entries

Assume that \sim_n and w are as above and that a'_n is an approximately uncorrelated triangular scheme, for example a Curie-Weiss(β)-ensemble as in Section 2.2.1 or an approximately uncorrelated Gaussian ensemble as in Section 2.2.2. Let $\ell \in \mathbb{N}$ be the number of equivalence classes induced by \sim_n , and let $P_1, \dots, P_\ell \in [n]^2$ be representatives. Then for all $P \in [n]^2$ we set $a_n(P) := a'_n(P_i)$, where $i \in [\ell]$ is the unique index with $P \sim_n P_i$. Then $(a_n)_n$ is an SSB-HKW correlated ensemble as in Definition 3.1.

3.2.2 The k -band model

The k -band model generalizes periodic and non-periodic band matrices to multiple bands. We assume that a_n is an SSB-HKW correlated ensemble and that $w = \mathbb{1}_{I_1} + \dots + \mathbb{1}_{I_k}$, where $(I_j)_{j \in [k]}$ are non-degenerate intervals in $[0, 1]$ that satisfy $I_1 < I_2 < \dots < I_k$ (where the inequalities are meant element-wise) and we assume that there is an $\epsilon > 0$ such that for all $i \neq j \in [k]$, $\text{dist}(I_i, I_j) \geq \epsilon$, where $\text{dist}(I_i, I_j) := \inf\{|x - y| \mid x \in I_i, y \in I_j\}$. In this setup, Theorem 3.3 ensures convergence of the ESDs of the random matrices X_n based on a_n as in (3.2). By the symmetry characterization in Theorem 3.3 it is immediately clear – without further calculations – that the semicircle law holds iff the intervals are a.s. symmetric around $1/2$, which means that for \mathbb{K} -almost all for all $x \in [0, 1]$: $x \in I_i$ for some $i \in [k]$ iff $(1 - x) \in I_j$ for some $j \in [k]$. As an example, if

$$w = \mathbb{1}_{[\frac{1}{10}, \frac{2}{10}]} + \mathbb{1}_{[\frac{3}{10}, \frac{4}{10}]} + \mathbb{1}_{[\frac{6}{10}, \frac{7}{10}]} + \mathbb{1}_{[\frac{8}{10}, \frac{9}{10}]}, \tag{3.3}$$

then w is symmetric around $1/2$, so the semicircle law holds for the associated X_n . But if exactly one of the indicators in (3.3) is dropped, the semicircle law will fail to hold for X_n .

An important application of this observation is for periodic and non-periodic band matrices as in Definition 2.11 with a halfwidth h_n which grows proportionally with n . So we assume $h_n/n \rightarrow \rho$ with $\rho \in (0, 1)$. If X_n^{NP} is based on an approximately uncorrelated triangular scheme a_n with halfwidth h_n , then X_n^{NP} and X_n share the same limiting spectral distribution (LSD) if X_n is as in (3.2), where a_n is the same approximately uncorrelated triangular scheme, \sim_n is the trivial equivalence relation $(a, b) \sim_n (p, q)$ iff $(a, b) = (p, q)$ or $(a, b) = (q, p)$, and where $w = \mathbb{1}_{[0, \rho]}$. This is easily seen, since by Definition 2.11,

$$X_n^{NP}(a, b) \neq 0 \Leftrightarrow |a - b| \leq h_n - 1 \Leftrightarrow \frac{|a - b|}{n} \leq \frac{h_n - 1}{n} \rightarrow \rho,$$

and $X_n(a, b) \neq 0$ iff $|a - b|/n \leq \rho$ by (3.2). Therefore, the limiting spectral distributions of X_n^{NP} and X_n are the same, as can be seen by using standard perturbation arguments (e.g. Hoffman-Wielandt inequality). With the same argument we obtain that the LSDs of X_n^P and X_n' are the same, where X_n' is constructed the same as X_n , but we change the weight w to the weight $w' = \mathbb{1}_{[0, \rho] \cup [1 - \rho, 1]}$. Since w' is symmetric around $1/2$, but w is not, the semicircle law holds for X_n^P , but not for X_n^{NP} . Of course, for independent entries, these statements about X_n^P and X_n^{NP} are well-known [4].

3.2.3 Block matrices

We consider multiple types of block matrices – all with finitely many different blocks – which can be treated with Theorem 3.3, and we will also construct a type of block matrix which cannot be treated with Theorem 3.3, exemplifying the boundaries of this theorem.

Toeplitz block matrices with k diagonals. A Toeplitz block matrix has the same blocks along diagonals, thus is of the form

$$T_n^{(k)} = \frac{1}{\sqrt{kn}} \begin{pmatrix} A_n^{(1)} & A_n^{(2)} & A_n^{(3)} & \cdots & A_n^{(k)} \\ A_n^{(2)T} & A_n^{(1)} & A_n^{(2)} & \cdots & A_n^{(k-1)} \\ A_n^{(3)T} & A_n^{(2)T} & A_n^{(1)} & \cdots & A_n^{(k-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_n^{(k)T} & A_n^{(k-1)T} & A_n^{(k-2)T} & \cdots & A_n^{(1)} \end{pmatrix},$$

where the matrix $A_n^{(1)}$ is symmetric and T indicates the transpose of a given matrix. We assume that the family

$$\{A_n^{(1)}(i, j) \mid 1 \leq i \leq j \leq n\} \cup \{A^{(\ell)}(P) \mid P \in [n]^2, \ell \in \{2, \dots, k\}\}$$

is approximately uncorrelated. Then the semicircle holds almost surely for $T_n^{(k)}$ by Theorem 3.3. To see this, we need to verify that \sim_{kn} induced by the structure of $T_n^{(k)}$ satisfies the conditions $(E1')$, $(E2)$ and $(E3')$. For $(E1')$, fix a $p \in [kn]$, then for any $q \in [kn]$ there are at most $4k$ pairs (r, s) in $[n]^2$ equivalent to (p, q) . Since $4k^2n = O((kn)^{1.9})$, $(E1')$ is satisfied. For $(E2)$, fix $p, q, r \in [kn]$. Then there are at most two $s \in [kn]$ so that (p, q) is equivalent to (r, s) . Hence, $(E2)$ is satisfied with $B = 2$. For $(E3')$, we note that if (p, q) is fixed, then there is no $r \neq q$ so that (p, q) is equivalent to (q, r) , so that $(E3')$ is satisfied.

Hankel block matrices with k skew diagonals. For k odd, a Hankel block matrix has the same blocks along its skew diagonals, thus has the form

$$H_n^{(k)} = \frac{1}{\sqrt{\frac{k+1}{2}n}} \begin{pmatrix} A_n^{(1)} & A_n^{(2)} & A_n^{(3)} & \cdots & A_n^{((k+1)/2)} \\ A_n^{(2)T} & A_n^{(3)} & A_n^{(4)} & \cdots & A_n^{((k+3)/2)} \\ A_n^{(3)T} & A_n^{(4)T} & A_n^{(5)} & \cdots & A_n^{((k+5)/2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_n^{((k+1)/2)T} & A_n^{((k+3)/2)T} & A_n^{((k+5)/2)T} & \cdots & A_n^{(k)} \end{pmatrix},$$

where the matrices $A_n^{(1)}, A_n^{(3)}, \dots, A_n^{(k)}$ are symmetric. We assume that the family

$$\{A_n^{(\ell)}(i, j) \mid 1 \leq i \leq j \leq n, \ell \in \{1, 3, \dots, k\}\} \cup \{A^{(\ell)}(P) \mid P \in [n]^2, \ell \in \{2, 4, \dots, k-1\}\}$$

is approximately uncorrelated. Then the semicircle law holds almost surely for $H_n^{(k)}$ by Theorem 3.3. To see this, we again need to verify that \sim_{kn} induced by the structure of $T_n^{(k)}$ satisfies the conditions $(E1')$, $(E2)$ and $(E3')$. For $(E1')$, fix a $p \in [kn]$, then for any $q \in [kn]$ there are at most $4k$ pairs (r, s) in $[n]^2$ equivalent to (p, q) . Since $4k^2n = O(((k+1)/2)n)^{1.9}$, $(E1')$ is satisfied. For $(E2)$, fix $p, q, r \in [kn]$. Then there are at most two $s \in [kn]$ so that (p, q) is equivalent to (r, s) . Hence, $(E2)$ is satisfied with $B = 2$. For $(E3')$, we note that if (p, q) is fixed, then there is no $r \neq q$ so that (p, q) is equivalent to (q, r) , so that $(E3')$ is satisfied.

Homogeneous block model. We assume that the block matrix is made up of $k \times k$ blocks, where the diagonal blocks are the same, and also the non-diagonal blocks are

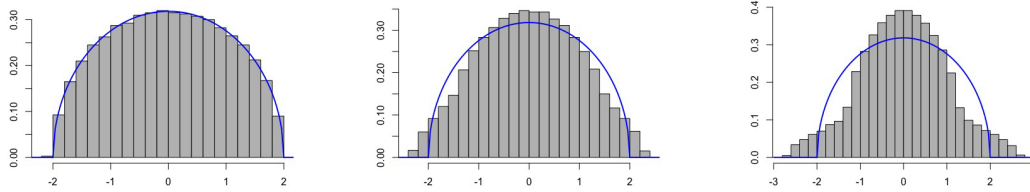


Figure 1: Histograms (in grey) of simulated spectra of $M_n^{(k)}$ for $n = 1000$ and $k = 2, 3, 4$ and i.i.d. Rademacher family (3.4), density (in blue) of the semicircle distribution.

the same (up to transposition).

$$M_n^{(k)} := \frac{1}{\sqrt{kn}} \begin{pmatrix} A_n & B_n & B_n & \cdots & B_n \\ B_n^T & A_n & B_n & \cdots & B_n \\ B_n^T & B_n^T & A_n & \cdots & B_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_n^T & B_n^T & B_n^T & \cdots & A_n \end{pmatrix},$$

where A_n is symmetric. We assume that the family

$$\{A_n(i, j) \mid 1 \leq i \leq j \leq n\} \cup \{B_n(P) \mid P \in [n]^2\} \tag{3.4}$$

is approximately uncorrelated. We check if \sim_{kn} induced by the structure of $M_n^{(k)}$ satisfies the conditions (E1), (E2) and (E3). For (E1), fix a $p \in [kn]$, then for any $q \in [kn]$ there are at most $4k^2$ pairs (r, s) in $[n]^2$ equivalent to (p, q) . Since $4k^2n = O((kn)^{1.9})$, even (E1') is satisfied. For (E2), fix $p, q, r \in [kn]$. Then are at most $2(k - 1)$ elements $s \in [kn]$ so that (p, q) is equivalent to (r, s) . Hence, (E2) is satisfied with $B = 2(k - 1)$. For (E3'), we note that if (p, q) is fixed such that it falls within a block B_n or B_n^T (for which we have $(nk)^2 - kn^2$ choices), then for $k \geq 3$ there are exactly $k - 2$ elements $r \neq q$ so that (p, q) is equivalent to (q, r) , so that (E3) is not satisfied for $k \geq 3$, since then $((nk)^2 - kn^2)(k - 2) \neq o((kn)^2)$. So Theorem 3.3 is not applicable for $k \geq 3$. (E3') is, however, satisfied for $k \in \{1, 2\}$, so that the almost sure semicircle holds in that case. A simulation of $M_n^{(k)}$ for $k = 2, 3, 4$, $n = 1000$ and an i.i.d. Rademacher family (3.4) indicates that indeed, the semicircle law is likely to fail for $k \geq 3$, see Figure 1.

3.3 Proof of Theorem 3.3

In Section 3.3.1 we derive the convergence of the moments of the ESDs studied in Theorem 3.3. In Section 3.3.2 we analyze the limiting moments, argue that these uniquely determine the limiting distribution which has to have compact support and be symmetric, and derive the weak convergence statements of Theorem 3.3. Section 3.3.3 is devoted to the last statement of Theorem 3.3, the characterization of the weak limit being the semicircle distribution.

3.3.1 Convergence of moments

In this subsection, we will prove the weak convergence statements in Theorem 3.3. To this end, we need to develop some notation and combinatorics. For every tuple $\underline{t} \in [n]^k$, denote by $\pi_{\underline{t}}$ the partition on $[k]$ with

$$i \sim_{\pi_{\underline{t}}} j \iff e_i \sim_n e_j \quad (\Leftrightarrow (t_i, t_{i+1}) \sim_n (t_j, t_{j+1})),$$

where $k + 1$ is identified with 1 and we recall the interpretation of \underline{t} as a graph as in the description right below (2.9), so that e_1, \dots, e_k are the edges of \underline{t} . Also, note that $\pi_{\underline{t}}$ depends on n , but we will suppress this dependence in the notation. Now if π is any partition on $[k]$, then define

$$T_{n,k}(\pi) := \{\underline{t} \in [n]^k \mid \pi_{\underline{t}} = \pi\}.$$

Then

$$[n]^k = \dot{\cup}_{\pi \in \mathcal{P}(k)} T_{n,k}(\pi),$$

where as per Definition 3.2, $\mathcal{P}(k)$ denotes the set of partitions of the set $[k]$. Further, define for all $\underline{t} \in [n]^k$:

$$a_n(\underline{t}) := a_n(t_1, t_2) a_n(t_2, t_3) \cdots a_n(t_k, t_1) = \prod_{j=1}^k a_n(t_j, t_{j+1}),$$

$$w(\underline{t}) := w\left(\frac{|t_1 - t_2|}{n}\right) w\left(\frac{|t_2 - t_3|}{n}\right) \cdots w\left(\frac{|t_k - t_1|}{n}\right) = \prod_{j=1}^k w\left(\frac{|t_j - t_{j+1}|}{n}\right).$$

Then

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in [n]^k} w(\underline{t}) a_n(\underline{t}) = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{P}(k)} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}). \quad (3.5)$$

Lemma 3.4. *Let $\pi \in \mathcal{P}(k)$ be a partition with $\#\pi = r$ and $n \geq B$, then we have*

$$\#T_{n,k}(\pi) \leq n^{r+1} B^{k-r-1}.$$

Further, if π contains a singleton, we have

$$\#T_{n,k}(\pi) \leq n^r B^{k-r}.$$

Proof. The first statement was proved in [25, p.4]. The second statement is clear for $k = 1$. If $k \geq 2$, write $\pi = \{B_1, \dots, B_r\}$ in increasing order, that is, $1 \in B_1$, $\min\{1, \dots, k\} \setminus B_1 \in B_2$, and so forth. Now for some $s \in [r]$, $\#B_s = 1$, so $B_s = \{i^*\}$ for some $i^* \in [k]$. We begin the construction of \underline{t} at the edge $e_{i^*+1} = (t_{i^*+1}, t_{i^*+2})$, for which we have n^2 choices. For each $\ell \in \{i^* + 2, \dots, i^* - 1\}$ we have to choose the destination of edge e_ℓ , which is $t_{\ell+1}$ (note that the destination of edge e_{i^*} has already been picked). Each time ℓ enters a block of π which has not been visited before, we have at most n choices for $t_{\ell+1}$, and this happens exactly $r - 2$ times. If ℓ does not enter a new block, then e_ℓ is π -equivalent to some $e_{\ell'}$ with $\ell' \in \{i^* + 1, \dots, \ell - 1\}$, leaving at most B choices for $t_{\ell+1}$ due to condition (E2), which happens exactly $k - 2 - (r - 2)$ times, yielding at most $n^2 \cdot n^{r-2} B^{k-r}$ choices. \square

Next, we analyze the following situation: Assume \underline{t} and $\underline{t}' \in T_{n,k}(\pi)$ are two tuples. Then we know that *separately*, the edges e_1, \dots, e_k and e'_1, \dots, e'_k are compatible with π . But what can be said about the partition $\hat{\pi} \in \mathcal{P}(2k)$ that describes the \sim_n relations of the entirety $e_1, \dots, e_k, e'_1, \dots, e'_k$ (where $e_k = (t_k, t_1)$ and $e'_k = (t'_k, t'_1)$)? Surely, restricting $\hat{\pi}$ to the lower and upper half of $[2k]$ by setting

$$\hat{\pi}_{\leq k} := \{B \cap \{1, \dots, k\}, B \in \hat{\pi}\} \quad \text{and} \quad \hat{\pi}_{>k} := \{B \cap \{k+1, \dots, 2k\}, B \in \hat{\pi}\},$$

then $\hat{\pi}_{\leq k} = \pi = \hat{\pi}_{>k}$, where $\{k+1, \dots, 2k\}$ is identified with $[k]$. In this situation, that is, we are given a $\pi \in \mathcal{P}(k)$ and a $\hat{\pi} \in \mathcal{P}(2k)$ with $\hat{\pi}_{\leq k} = \pi = \hat{\pi}_{>k}$, we say that $\hat{\pi} \in \mathcal{P}(2k)$ is a *square partition* of π . Note that in general, there are many possible square partitions to a given partition. We say that a square partition $\hat{\pi}$ is *traversing*, if there is a block $B \in \hat{\pi}$

with $B \cap \{1, \dots, k\} \neq \emptyset$ and $B \cap \{k + 1, \dots, 2k\} \neq \emptyset$. Otherwise, we call $\hat{\pi}$ *non-traversing*. We are now interested in bounds of the cardinality of the following sets:

$$\hat{T}_{n,2k}(\hat{\pi}) := \{(\underline{t}, \underline{t}') \in (T_{n,k}(\pi))^2, (e_1, \dots, e_k, e'_1, \dots, e'_k) \sim_n \hat{\pi}\}.$$

Our bounds will depend on the number of unifications of blocks in $\hat{\pi}_{\leq k}$ with blocks in $\hat{\pi}_{>k}$ carried out in $\hat{\pi}$. Note that there may be at most $\#\pi$ unifications.

Lemma 3.5. *Let $\hat{\pi} \in \mathcal{P}(2k)$ be a square partition of $\pi \in \mathcal{P}(k)$ with $\#\pi = r$.*

1. *Let $\hat{\pi}$ be non-traversing. Then $\#\hat{\pi} = 2r$, and for all $n \geq B$*

$$\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r+2} B^{2k-2r-2}.$$

Further, if $\hat{\pi}$ contains a singleton, we have

$$\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r} B^{2k-2r}.$$

2. *If $k \geq 2$, $\hat{\pi}$ is traversing, and $1 \leq u \leq r$ is the number of unifications, then $\#\hat{\pi} = 2r - u$ and for $n \geq B$ we have*

$$\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r-u} o(n^2) B^{2k-2r+u-3}.$$

Further, if $s_1(\pi) \geq 2$, that is, π contains at least two singletons, it holds

$$\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r-u-1} o(n^2) B^{2k-2r+u-2}.$$

In addition, the terms $o(n^2)$ may be replaced by $O(n^{2-\delta})$ if $(E1')$ is assumed instead of $(E1)$.

Proof. The statements in (1) follow directly from Lemma 3.4 and the inequality

$$\#\hat{T}_{n,2k}(\hat{\pi}) \leq \#T_{n,k}(\pi) \cdot \#T_{n,k}(\pi),$$

which is a trivial upper bound, since for any pair of tuples $(\underline{t}, \underline{t}') \in \hat{T}_{n,2k}(\hat{\pi})$ we have $\underline{t}, \underline{t}' \in T_{n,k}(\pi)$. For the statements in (2) we analyze how many possibilities we have to construct such a tuple pair. For the first statement we begin with the smallest $\ell \in [k]$ which is $\hat{\pi}$ -equivalent to some $\ell' \in \{k + 1, \dots, 2k\}$. To fix $t_\ell, t_{\ell+1}, t'_{\ell'}, t'_{\ell'+1}$, we have $n \cdot o(n^2)$ possibilities by condition (E3). Continuing through \underline{t} , we have at most $n^{r-1} B^{k-r-1}$ possibilities. When continuing through the edges of \underline{t}' , there are $r - 1$ new blocks to be entered, $u - 1$ of which will be connected with some block in $\hat{\pi}_{\leq k}$, which will then admit B^{u-1} choices by condition (E2). The remaining $r - u$ independent new blocks will admit at most n^{r-u} choices. The remaining edges will admit at most B^{k-r-1} choices by condition (E2). This yields at most

$$n \cdot o(n^2) \cdot n^{r-1} B^{k-r-1} \cdot B^{u-1} \cdot n^{r-u} \cdot B^{k-r-1} = n^{2r-u} \cdot o(n^2) \cdot B^{2k-2r+u-3}$$

choices to construct a pair of tuples $(\underline{t}, \underline{t}') \in \hat{T}_{n,2k}(\hat{\pi})$.

If π contains at least two singletons, we proceed similarly. Starting at the first $\ell \in [k]$ which is $\hat{\pi}$ -equivalent to some $\ell' \in \{k + 1, \dots, 2k\}$, we again fix $t_\ell, t_{\ell+1}, t'_{\ell'}, t'_{\ell'+1}$ for which we have at most $n \cdot o(n^2)$ possibilities by condition (E3). To complete the tuple \underline{t} we pick an $m \in [k] \setminus \{\ell\}$ such that e_m is a single edge and continue exactly as in the case above except that we do not cyclicly walk through the remaining edges of \underline{t} , but from e_ℓ forward until reaching the node t_m , and then from e_ℓ backwards until reaching the node t_{m+1} . We have then completely determined the tuple while only entering $r - 2$ new blocks, leading to at most $n^{r-2} B^{k-r}$ possibilities. To complete \underline{t}' , we proceed exactly as in the proof of the first statement in (2), yielding a total of at most

$$n \cdot o(n^2) \cdot n^{r-2} B^{k-r} \cdot B^{u-1} \cdot n^{r-u} \cdot B^{k-r-1} = n^{2r-u-1} \cdot o(n^2) \cdot B^{2k-2r+u-2}$$

possibilities to construct the pair $(\underline{t}, \underline{t}') \in \hat{T}_{n,2k}(\hat{\pi})$. □

We proceed to analyze the sum in (3.5). Note that for each summand,

$$|\mathbb{E}w(\underline{t})a_n(\underline{t})| \leq W^k C(k), \tag{3.6}$$

which we call *universal bound*. It will be used frequently throughout the text. For a partition $\pi \in \mathcal{P}(k)$ and $\ell \in [k]$, denote by $s_\ell(\pi)$ the number of blocks of size ℓ in π , so

$$s_\ell(\pi) := \#\{B \in \pi \mid |B| = \ell\},$$

so that

$$k = \sum_{\ell=1}^k s_\ell(\pi) \cdot \ell \quad \text{and} \quad \#\pi = \sum_{\ell=1}^k s_\ell(\pi). \tag{3.7}$$

The equalities in (3.7) yield some useful inequalities. For example, for any $\pi \in \mathcal{P}(k)$,

$$\#\pi \leq s_1(\pi) + \frac{k - s_1(\pi)}{2}. \tag{3.8}$$

We will next analyze the sum in (3.5). We would like to see this k -th empirical moment converge to limit moments in expectation, in probability or almost surely.

We now proceed to analyze the convergence in (3.5).

Lemma 3.6. *Let $\pi \in \mathcal{P}(k)$ with $\#\pi =: r \neq k/2$, then*

$$\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t})a_n(\underline{t}) \xrightarrow{n \rightarrow \infty} 0 \tag{3.9}$$

in expectation and in probability, and the statement holds also almost surely if $(E1')$ is assumed instead of $(E1)$.

Proof. For $k = 1$ the term in question is

$$\frac{1}{n^{\frac{3}{2}}} \sum_{t \in [n]} w(0)a_n(t, t)$$

which converges to zero in expectation and almost surely by Lemma B.1 b). So for the remainder of the proof we assume that $k \geq 2$.

If $r < k/2$ we obtain with Lemma 3.4 that $\#T_{n,k}(\pi) \leq n^{r+1} B^{k-r-1} \leq n^{k/2+1/2} B^{k/2-1/2}$, which leads to convergence in expectation and almost sure convergence to zero in (3.9) by Lemma B.1 b).

If $r > k/2$, then (3.8) yields $s_1(\pi) \geq 2r - k \geq 1$, so by (A1),

$$|\mathbb{E}w(\underline{t})a_n(\underline{t})| \leq \frac{W^k C(k)}{n^{\frac{1}{2}(2r-k)}}$$

and therefore by Lemma 3.4,

$$\left| \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in T_{n,k}(\pi)} \mathbb{E}w(\underline{t})a_n(\underline{t}) \right| \leq \frac{1}{n^{\frac{k}{2}+1}} n^r B^{k-r} \frac{W^k C(k)}{n^{r-\frac{k}{2}}} \leq \frac{W^k C(k) B^{k-r}}{n},$$

so the sum in question converges to zero in expectation. We will now show that its variance decays to zero. Let $\hat{\pi}$ be a square partition of π , then it suffices to show

$$\frac{1}{n^{k+2}} \sum_{(\underline{t}, \underline{t}') \in \hat{T}_{n,2k}(\hat{\pi})} |\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - \mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \xrightarrow{n \rightarrow \infty} 0. \tag{3.10}$$

If $\hat{\pi}$ is non-traversing, then $s_1(\hat{\pi}) = 2s_1(\pi)$, so

$$|\mathbb{E}a_n(\underline{t})a_n(\underline{t}')| \leq \frac{C(2k)}{n^{2r-k}} \quad \text{and} \quad |\mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \leq \frac{C(k)^2}{n^{2r-k}}.$$

Further, $\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r}B^{2k-2r}$ by Lemma 3.5, so (3.10) holds summably fast.

If $\hat{\pi}$ is traversing, then we distinguish two cases:

Case 1: $s_1(\pi) \geq 2$.

There is a block in $\hat{\pi}$ which is the union of a block in $\hat{\pi}_{\leq k}$ and a block in $\hat{\pi}_{>k}$. Each union decreases $\#\hat{\pi}$ by 1, and for each union at most 2 singletons are paired. Let u with $1 \leq u \leq r$ be the number of unions, then $\#\hat{\pi} = 2r - u$ and $s_1(\hat{\pi}) \geq \max(0, 4r - 2k - 2u)$. Hence,

$$|\mathbb{E}a_n(\underline{t})a_n(\underline{t}')| \leq \frac{C(2k)}{n^{\max(0, 2r-k-u)}} \quad \text{and} \quad |\mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \leq \frac{C(k)^2}{n^{\max(0, 2r-k-u)}}.$$

By Lemma 3.5, $\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{2r-u-1}o(n^2)B^{2k-2r+u-2}$. Then since

$$\frac{n^{2r-u-1} \cdot o(n^2)}{n^{k+2} \cdot n^{\max(0, 2r-k-u)}} = \frac{1}{n} \cdot \frac{o(n^2)}{n^2} \cdot \frac{1}{n^{\max(0, 2r-k-u)-(2r-k-u)}},$$

we obtain (3.10), and this convergence is summably fast if $(E3')$ is assumed instead of $(E3)$, since the term $o(n^2)$ can then be replaced by $O(n^{2-\delta})$.

Case 2: $s_1(\pi) = 1$.

Since $k \geq 2$ and π has exactly one singleton block, π contains at least one other block of higher cardinality than 1. Setting $r := \#\pi$, the upper bound $r \leq 1 + (k-1)/2 = (k+1)/2$ holds. Therefore, Lemma 3.4 yields for $n \geq B$ that $\#T_{n,k} \leq n^r B^{k-r} \leq n^{(k+1)/2} B^{(k-1)/2}$. With these observations, we can show that the fourth central moment of the sum in (3.9) decays summably fast, where we can drop the factor $w(\underline{t})$ from the analysis, since it is bounded by W^k :

$$\begin{aligned} & \mathbb{E} \left(\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in T_{n,k}(\pi)} (a_n(\underline{t}) - \mathbb{E}a_n(\underline{t})) \right)^4 \\ &= \left| \frac{1}{n^{2k+4}} \sum_{\underline{t}^{(1)}, \dots, \underline{t}^{(4)} \in T_{n,k}(\pi)} \mathbb{E} \prod_{s=1}^4 [a_n(\underline{t}^{(s)}) - \mathbb{E}a_n(\underline{t}^{(s)})] \right| \\ &\leq \frac{1}{n^{2k+4}} \cdot n^{2(k+1)} B^{2(k-1)} \cdot \sum_{S \subseteq \{1, \dots, 4\}} C(k\#S) \cdot C(k)^{\#S^c}, \end{aligned}$$

where $S^c := \{1, \dots, 4\} \setminus S$ and we use that for any $\underline{t}^{(1)}, \dots, \underline{t}^{(4)} \in T_{n,k}(\pi)$ we find

$$\left| \mathbb{E} \prod_{s=1}^4 [a_n(\underline{t}^{(s)}) - \mathbb{E}a_n(\underline{t}^{(s)})] \right| = \sum_{S \subseteq \{1, \dots, 4\}} \left| \mathbb{E} \prod_{s \in S} a_n(\underline{t}^{(s)}) \right| \cdot \prod_{s \notin S} |\mathbb{E}a_n(\underline{t}^{(s)})|,$$

and then (AU1), also setting $C(0) := 1$. □

Lemma 3.6 shows that an asymptotic contribution of the empirical k -th moment may only stem from those $\pi \in \mathcal{P}(k)$ with $\#\pi = k/2$. In particular, odd empirical moments vanish as $n \rightarrow \infty$. Next, we show that only pair partitions need to be considered:

Lemma 3.7. *If $\pi \in \mathcal{P}(k)$ with $\#\pi = k/2$ but $\#B \neq 2$ for some $B \in \pi$, then*

$$\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t})a_n(\underline{t}) \xrightarrow{n \rightarrow \infty} 0$$

in expectation and almost surely.

Proof. If $\#B \neq 2$ for some $B \in \pi$, then necessarily $s_1(\pi) > 1$, since $\#\pi = k/2$. Thus by Lemma 3.4, $\#T_{n,k}(\pi) \leq n^{k/2} B^{k/2}$. The statement follows with Lemma B.1. \square

We remind the reader that as per Definition 3.2, we denote by $\mathcal{P}_2(k)$ resp. $\mathcal{NC}_2(k)$ the set of all pair partitions resp. non-crossing pair partitions of the set $[k]$. Returning to (3.5), we have shown so far that for each $\pi \in \mathcal{P}(k)$ with $\pi \notin \mathcal{P}_2(k)$,

$$\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) \xrightarrow{n \rightarrow \infty} 0 \quad \text{in exp./prob./a.s.,}$$

where we used properties (E1) and (E3) for convergence in expectation and in probability and the stronger conditions (E1') and (E3') for convergence almost surely. We obtain

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{P}_2(k)} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) + R_n^{(1)}, \quad (3.11)$$

where $R_n^{(1)}$ is a remainder term that tends to zero in expectation and in probability under (E1) and (E3), and also almost surely under (E1') and (E3'). It remains to investigate the first summand on the r.h.s. of (3.11). To this end, if $\pi \in \mathcal{NC}_2(k)$, we define

$$\begin{aligned} XT_{n,k}(\pi) &:= \{\underline{t} \in T_{n,k}(\pi) \mid \exists r \neq s \in [k] : r \sim_\pi s \wedge (t_r, t_{r+1}) \neq (t_{s+1}, t_s)\}, \\ FT_{n,k}(\pi) &:= \{\underline{t} \in T_{n,k}(\pi) \mid \{r, s\} \in \pi \Rightarrow (t_r, t_{r+1}) = (t_{s+1}, t_s), \#V_{\underline{t}} \leq k/2\}, \\ BT_{n,k}(\pi) &:= \{\underline{t} \in T_{n,k}(\pi) \mid \{r, s\} \in \pi \Rightarrow (t_r, t_{r+1}) = (t_{s+1}, t_s), \#V_{\underline{t}} = k/2 + 1\}. \end{aligned}$$

Lemma 3.8. *Recall the constant B from condition (E2). Then we find:*

a) *If $\pi \in \mathcal{P}_2(k)$ is crossing, then*

$$\#T_{n,k}(\pi) \leq n^{\frac{k}{2}-1} \cdot K(k, B) \cdot o(n^2),$$

where $o(n^2)$ is the term from condition (E3), which may be replaced by the term $O(n^{2-\delta})$ from condition (E3') if the latter condition is assumed. Further, $K(k, B)$ is a constant which depends only on k and B .

b) *If $\pi \in \mathcal{NC}_2(k)$, then*

$$\#XT_{n,k}(\pi) \leq n^{\frac{k}{2}-1} \cdot K(k, B) \cdot o(n^2),$$

where $o(n^2)$ is the term from condition (E3), which may be replaced by the term $O(n^{2-\delta})$ from condition (E3') if the latter condition is assumed. Further, $K(k, B)$ is a constant which depends only on k and B .

c) *If $\pi \in \mathcal{NC}_2(k)$, then*

$$\#FT_{n,k}(\pi) \leq K(k) \cdot n^{\frac{k}{2}},$$

where $K(k)$ is a constant depending only on k .

Proof. a) Is more detailed version of Lemma 2 in [25], page 5. Their Lemma 2 is proved with a reduction argument (their Lemma 1 on page 4). The constant $K(k, B)$ will depend on the number of reductions, which depends on the specific non-crossing π , but is upper bounded by $k/2 - 1$.

b) This statement is a more detailed version of Lemma 3 in [25], where the reasoning for the constant $K(k, B)$ is similar to the case of a).

c) We bound $\#FT_{n,k}(\pi)$ by the number of possibilities to construct a tuple with at most $k/2$ vertices, which is in turn bounded by $k^k n^{k/2}$ by Lemma 2.8 B). \square

Applying Lemma 3.8 a) and Lemma B.1, we obtain from (3.11):

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{NC}_2(k)} \sum_{\underline{t} \in T_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) + R_n^{(1)} + R_n^{(2)}, \quad (3.12)$$

where $R_n^{(2)} \rightarrow 0$ in expectation and in probability under (E3) and also almost surely under (E3'). Applying Lemma 3.8 b) and Lemma B.1, we obtain from (3.12):

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{NC}_2(k)} \sum_{\substack{\underline{t} \in FT_{n,k}(\pi) \\ \cup BT_{n,k}(\pi)}} w(\underline{t}) a_n(\underline{t}) + R_n^{(1)} + R_n^{(2)} + R_n^{(3)}, \quad (3.13)$$

where $R_n^{(3)} \rightarrow 0$ in expectation and in probability under (E3) and also almost surely under (E3'). Finally, applying Lemma 3.8 c) and Lemma B.1, we obtain from (3.13):

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{NC}_2(k)} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) + R_n^{(1)} + R_n^{(2)} + R_n^{(3)} + R_n^{(4)}, \quad (3.14)$$

where $R_n^{(4)} \rightarrow 0$ in expectation and almost surely. In total, setting $R_n := R_n^{(1)} + R_n^{(2)} + R_n^{(3)} + R_n^{(4)}$, we obtain from above observations that

$$\langle \sigma_n, x^k \rangle = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\pi \in \mathcal{NC}_2(k)} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) + R_n, \quad (3.15)$$

where R_n is a random variable that converges to zero in expectation and in probability if (E1) and (E3) are assumed and also almost surely if (E1') and (E3') are assumed.

It remains to investigate the first summand on the r.h.s. of (3.15). To this end, define for each $\pi \in \mathcal{NC}_2(k)$:

$$J_w(\pi) := \lim_{n \rightarrow \infty} \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}). \quad (3.16)$$

We will show below in Lemma 3.11 that this limit actually exists and how it can be calculated recursively. For now, we take existence for granted. In the next lemma we study the set $BT_{n,k}(\pi)$. We will call elements $\underline{t} \in BT_{n,k}(\pi)$ π -backtracking. Note that any such \underline{t} has $k/2$ edges, where each edge is traversed exactly twice. Further, it has $k/2 + 1$ vertices, so that the graph of \underline{t} spans a double edged tree. For each $\pi \in \mathcal{NC}_2(k)$, we denote by $\underline{t}^\pi \in BT_{n,k}(\pi)$ the canonical π -backtracking path constructed as follows: Set $t_1^\pi = 1$ and $t_2^\pi = 2$, thus determining the edge e_1 of \underline{t}^π . Then if $t_1^\pi, \dots, t_\ell^\pi$, $2 \leq \ell < k$ have been constructed, proceed for $t_{\ell+1}^\pi$ as follows: If $e_\ell \sim_\pi e_{\ell'}$ for some $1 \leq \ell' < \ell$, then set $t_{\ell+1}^\pi := t_{\ell'}^\pi$. Otherwise, set $t_{\ell+1}^\pi := \max(t_1^\pi, \dots, t_\ell^\pi) + 1$. Reminding the reader that VY denotes the variance of a random variable Y , we now formulate the following lemma.

Lemma 3.9. *Let $\pi \in \mathcal{NC}_2(k)$ be arbitrary.*

a) $BT_{n,k}(\pi) = \{(g(t_1^\pi), \dots, g(t_k^\pi)) \mid g : [k/2 + 1] \rightarrow [n] \text{ is injective}\}.$

In particular: $\forall n \geq k : \#BT_{n,k}(\pi) = n(n-1) \cdots (n-k/2).$

b) $\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) \mathbb{E} a_n(\underline{t}) \xrightarrow{n \rightarrow \infty} J_w(\pi).$

c) $\text{V} \left(\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) a_n(\underline{t}) \right) \xrightarrow{n \rightarrow \infty} 0.$

and in c) the convergence is summably fast if condition (E3') is assumed instead of (E3).

Proof. Statement a) is clear.

For b), we calculate using condition (A2):

$$\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) \mathbb{E}a_n(\underline{t}) = \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) (\mathbb{E}a_n(\underline{t}) - 1) + \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t})$$

Since the second summand on the r.h.s. converges to $J_w(\pi)$, it suffices to show that the first summand on the r.h.s. converges to zero, which follows from condition (A2) and part a):

$$\frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} |w(\underline{t})(\mathbb{E}a_n(\underline{t}) - 1)| \leq \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} W^k C_n^{(k/2)} \xrightarrow{n \rightarrow \infty} 0.$$

For c), let $\hat{\pi}$ be a square partition of π , then if $\hat{\pi}$ is non-traversing and $\underline{t}, \underline{t}' \in BT_{n,k}(\pi)$ with $(\underline{t}, \underline{t}') \sim \hat{\pi}$, then

$$\begin{aligned} |\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - \mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| &\leq |\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - 1| + |\mathbb{E}a_n(\underline{t}) - 1||\mathbb{E}a_n(\underline{t}')| + |\mathbb{E}a_n(\underline{t}') - 1| \\ &\leq C_n^{(k)} + C_n^{(k/2)}C(k) + C_n^{(k/2)}, \end{aligned}$$

so

$$\begin{aligned} &\frac{1}{n^{k+2}} \sum_{\substack{\underline{t}, \underline{t}' \in BT_{n,k}(\pi) \\ (\underline{t}, \underline{t}') \sim \hat{\pi}}} |w(\underline{t})w(\underline{t}')||\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - \mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \\ &\leq \frac{(\#BT_{n,k}(\pi))^2}{n^{k+2}} W^{2k} (C_n^{(k)} + C_n^{(k/2)}C(k) + C_n^{(k/2)}) \end{aligned}$$

which converges to zero, and this convergence is summably fast if the sequences $(C_n^{(\ell)})_n$ converge to zero summably fast for all ℓ .

Now if $\hat{\pi}$ is traversing, we can merely achieve the bound

$$|\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - \mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \leq C(2k) + C(k)^2.$$

On the other hand, by Lemma 3.5, $\#\hat{T}_{n,2k}(\hat{\pi}) \leq n^{k-1}o(n^2)B^{k-2}$ where $o(n^2)$ can be replaced by $O(n^{2-\delta})$ in case we assume condition (E3') to hold instead of (E3). Therefore,

$$\begin{aligned} &\frac{1}{n^{k+2}} \sum_{\substack{\underline{t}, \underline{t}' \in BT_{n,k}(\pi) \\ (\underline{t}, \underline{t}') \sim \hat{\pi}}} |w(\underline{t})w(\underline{t}')||\mathbb{E}a_n(\underline{t})a_n(\underline{t}') - \mathbb{E}a_n(\underline{t})\mathbb{E}a_n(\underline{t}')| \\ &\leq \frac{\#\hat{T}_{n,2k}(\hat{\pi})}{n^{k+2}} W^{2k} (C(2k) + C(k)^2) \leq \frac{n^{k-1}o(n^2)}{n^{k+2}} B^{k-2} W^{2k} (C(2k) + C(k)^2) \end{aligned}$$

which converges to zero, and this convergence is summably fast if condition (E3) is replaced by (E3'), since then $o(n^2)$ can be replaced by $O(n^{2-\delta})$. \square

Theorem 3.10. *Let $k \in \mathbb{N}$ be arbitrary, then it holds*

$$\langle \sigma_n, x^k \rangle \xrightarrow{n \rightarrow \infty} \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi)$$

in expectation and in probability, and also almost surely if the conditions (E1) and (E3) are replaced by their stronger counterparts (E1') and (E3'), and the sequences $(C_n^{(\ell)})_n$ from condition (A2) are assumed to converge to zero summably fast.

Proof. Starting from (3.15) and using the definition in (3.16), the statement follows with Lemma 3.9 and Lemma B.2. \square

3.3.2 Analysis of the limiting moments

First, we establish that the limits $J_w(\pi)$ for $\pi \in \mathcal{NC}_2(k)$ which were defined in (3.16) actually exist. It turns out this is a limit of a Riemann sum (see also [4]), thus a Riemann integral. We calculate

$$\begin{aligned} \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} w(\underline{t}) &= \frac{1}{n^{\frac{k}{2}+1}} \sum_{\underline{t} \in BT_{n,k}(\pi)} \prod_{\{t_i, t_j\} \in E_{\underline{t}}} w\left(\frac{|t_i - t_j|}{n}\right) \\ &= \frac{1}{n^{\frac{k}{2}+1}} \sum_{\substack{v_1, \dots, v_{k/2+1}=1 \\ \text{all distinct}}}^n \prod_{\{r,s\} \in E_{\underline{t}\pi}} w\left(\frac{|v_r - v_s|}{n}\right) \\ &= \frac{1}{n^{\frac{k}{2}+1}} \sum_{v_1, \dots, v_{k/2+1}=1}^n \prod_{\{r,s\} \in E_{\underline{t}\pi}} w\left(\frac{|v_r - v_s|}{n}\right) + o(1) \\ &= \int_{[0,1]^{\frac{k}{2}+1}} \prod_{\{r,s\} \in E_{\underline{t}\pi}} w(|x_r - x_s|) \prod_{u \in V_{\underline{t}\pi}} dx_u + o(1) \end{aligned} \tag{3.17}$$

In the first step we used the definition of $w(\underline{t})$, and by slight abuse of notation we identify the abstract edge set $E_{\underline{t}} = \{e_1, \dots, e_k\}$ with the family $(\phi_{\underline{t}}(e_i))_{i \in [k]}$. For the second step we used Lemma 3.9 a), in the third step we used that the difference of the two terms in question is bounded by

$$\frac{W^{\frac{k}{2}}}{n^{\frac{k}{2}+1}} \cdot \#\left\{ \underline{v} \in [n]^{\frac{k}{2}+1} \mid \#\underline{V}_{\underline{v}} \leq k \right\} = \frac{W^{\frac{k}{2}}}{n^{\frac{k}{2}+1}} \cdot \left(n^{\frac{k}{2}+1} - n(n-1) \cdots (n-k/2) \right) \xrightarrow{n \rightarrow \infty} 0,$$

and in the fourth step we recognize the term as a Riemann sum and provide the limit. The next lemma summarizes our findings and shows how the Riemann integral may be calculated recursively.

Lemma 3.11. *Let $\pi \in \mathcal{NC}_2(k)$ be arbitrary. Then the limit in the definition of $J_w(\pi)$ in (3.16) exists and it holds*

$$J_w(\pi) = \int_{[0,1]^{\frac{k}{2}+1}} \prod_{\{r,s\} \in \tilde{E}_{\underline{t}\pi}} w^2(|x_r - x_s|) \prod_{u \in V_{\underline{t}\pi}} dx_u,$$

where $\tilde{E}_{\underline{t}\pi} := \{\phi_{\underline{t}\pi}(e_i) \mid i \in [k]\}$ as a set. In particular,

$$J_w(\{\{1, 2\}\}) = \int_0^1 \int_0^1 w^2(|x - y|) dx dy.$$

Further, for any block of the form $\{m, m + 1\} \in \pi$, it holds with $\tilde{\pi} := \pi \setminus \{m, m + 1\}$ (which denotes the partition on $[k - 2]$ after eliminating the block $\{m, m + 1\}$ from π and relabeling the elements $\{m + 2, \dots, k\}$ according to $a \mapsto a - 2$),

$$J_w(\pi) = \int_0^1 \varphi(x_{t_m^\pi}) J_w(\tilde{\pi} | x_{t_m^\pi}) dx_{t_m^\pi},$$

where

$$\begin{aligned} \varphi(x) &= \int_0^1 w^2(|x - y|) dy \\ J_w(\pi | x_{t_m^\pi}) &= \int_{[0,1]^{\frac{k}{2}}} \prod_{\{r,s\} \in \tilde{E}_{\underline{t}\pi}} w^2(|x_r - x_s|) \prod_{u \in V_{\underline{t}\pi} \setminus \{t_m^\pi\}} dx_u \end{aligned}$$

In particular, if $\varphi \equiv c$ is constant for some $c \in \mathbb{R}$, then $J_w(\pi) = c^{k/2}$, and the following upper bound is always valid:

$$\forall \pi \in \mathcal{NC}_2(k) : |J_w(\pi)| \leq W^k.$$

Proof. The integral equation for $J_w(\pi)$ has been derived in the calculation of (3.17). For the recursion, note that $\{m, m + 1\} \in \pi$ implies that $t_m^\pi = t_{m+2}^\pi$ and t_{m+1}^π is unique in the tuple t^π . Further, $t_m^\pi = t_{\tilde{m}}^\pi$, where $\tilde{\pi} := \pi \setminus \{m, m + 1\}$. Therefore,

$$\begin{aligned} J_w(\pi) &= \int_{[0,1]^{\frac{k}{2}+1}} \prod_{\{r,s\} \in E_{t^\pi}} w(|x_r - x_s|) \prod_{u \in V_{t^\pi}} dx_u \\ &= \int_{[0,1]^{\frac{k}{2}}} \int_{[0,1]} w(|x_{t_m^\pi} - x_{t_{m+1}^\pi}|) w(|x_{t_{m+1}^\pi} - x_{t_{m+2}^\pi}|) dx_{t_{m+1}^\pi} \dots \\ &\quad \dots \prod_{\ell \neq m, m+1} w(|x_{t_\ell^\pi} - x_{t_{\ell+1}^\pi}|) \prod_{u \in V_{t^\pi} \setminus \{t_{m+1}^\pi\}} dx_u \\ &= \int_{[0,1]} \varphi(x_{t_m^\pi}) \int_{[0,1]^{\frac{k}{2}-1}} \prod_{\ell \neq m, m+1} w(|x_{t_\ell^\pi} - x_{t_{\ell+1}^\pi}|) \left(\prod_{\substack{u \in V_{t^\pi} \\ u \neq t_m^\pi, t_{m+1}^\pi}} dx_u \right) dx_{t_m^\pi} \\ &= \int_0^1 \varphi(x_{t_m^\pi}) J_w(\tilde{\pi} | x_{t_m^\pi}) dx_{t_m^\pi}. \quad \square \end{aligned}$$

From the findings of this section and Theorem 3.10, we conclude

Corollary 3.12. *In the setting of Theorem 3.3, the ESDs $(\sigma_n)_n$ converge weakly in probability to a deterministic probability measure μ on \mathbb{R} with moments*

$$\forall k \in \mathbb{N} : \langle \mu, x^k \rangle = \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi).$$

Further, $\sigma_n \rightarrow \mu$ weakly almost surely if conditions (E1) and (E3) are strengthened to (E1') and (E3') respectively. The weak limit μ has compact support and vanishing odd moments. In particular, it is symmetric and uniquely determined by its moments.

Proof. The limiting moments in Theorem 3.10 are bounded by $W^k \# \mathcal{NC}_2(k) \mathbb{1}_{2\mathbb{N}}(k) \leq (4W)^k \mathbb{1}_{2\mathbb{N}}(k)$, where we used Lemma 3.11 and a well-known bound on the Catalan-numbers (e.g. [2]). Therefore, the moments satisfy the Carleman condition, thus admit at most one probability measure. By the method of moments for random probability measures (Theorem 3.5 in [9]), the weak convergence statements in Corollary 3.12 follow from the stochastic moment convergence in Theorem 3.10. Also, by the bound on the limiting moments which we identified in the beginning of the proof, and by Lemma 3.13 in [24], μ has compact support. Therefore, the vanishing odd moments allow to conclude that μ is symmetric (e.g. [26, p.134]). \square

3.3.3 Characterization of the semicircle law

We have seen that the ESDs of the ensemble $(X_n)_n$ as in Theorem 3.3 converge weakly to the unique symmetric probability distribution μ on $(\mathbb{R}, \mathcal{B})$ with moments

$$\forall k \in \mathbb{N} : \langle \mu, x^k \rangle = \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi).$$

and limiting variance

$$\varphi_0 := \int_0^1 \varphi(x) dx,$$

where

$$\forall x \in [0, 1] : \varphi(x) = \int_0^1 w^2(|x - y|)dy.$$

It follows that the ESDs of $(\varphi_0^{-1/2}X_n)_n$ have limiting variance 1, and the question now is when the semicircle law holds:

Lemma 3.13. *Let $(X_n)_n$ be an ensemble as in Theorem 3.3.*

1. *The asymptotic variance φ_0 can be calculated by*

$$\varphi_0 = 2 \int_0^1 (1 - x)w^2(x)dx.$$

In particular, the following statements are equivalent:

- a) $\varphi_0 = 0$.
- b) $w = 0$ on $[0, 1]$ \mathbb{A} -almost surely.

2. *Assume that $\varphi_0 > 0$. Then the following statements are equivalent:*

- a) *The semicircle law holds for $\frac{1}{\sqrt{\varphi_0}}X_n$.*
- b) *φ is constant, in particular, $\varphi \equiv \varphi_0$.*
- c) *$\varphi \equiv \int_0^1 w^2(x)dx$,*
- d) *w^2 is \mathbb{A} -a.s. symmetric around $1/2$, i.e. $w^2(x) = w^2(1 - x)$ for \mathbb{A} -a.a. $x \in [0, 1]$.*

Proof. We prove (2) first: Denote by $\tilde{\mu}$ the limiting spectral distribution of the ESDs of $\frac{1}{\sqrt{\varphi_0}}X_n$. Then

$$\forall k \in \mathbb{N} : \langle \tilde{\mu}, x^k \rangle = \frac{1}{\varphi_0^{k/2}} \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi). \tag{3.18}$$

b) \Rightarrow a) This follows immediately with (3.18) and the last statement in Lemma 3.11.

a) \Rightarrow b): Assume that φ is not \mathbb{A} -a.s. constant on $[0, 1]$. Denote by $\tilde{\mu}$ the LSD of $\varphi_0^{-1/2}X_n$. Then

$$\langle \tilde{\mu}, x^4 \rangle = \frac{1}{\varphi_0^2} \langle \mu, x^2 \rangle = \frac{1}{\varphi_0^2} \sum_{\pi \in \mathcal{NC}_2(k)} J_w(\pi).$$

Since $\mathcal{NC}_2(k) = \{\pi_1, \pi_2\}$, where $\pi_1 = \{\{1, 2\}, \{3, 4\}\}$ and $\pi_2 = \{\{1, 4\}, \{2, 3\}\}$ and $t^{\pi_1} = (1, 2, 1, 3)$, $t^{\pi_2} = (1, 2, 3, 2)$, we obtain

$$J_w(\pi_1) = \int_{[0,1]^3} w^2(|x - y|)w^2(|x - z|)dx dy dz = \int_0^1 \varphi^2(x)dx.$$

and

$$J_w(\pi_2) = \int_{[0,1]^3} w^2(|x - y|)w^2(|y - z|)dx dy dz = \int_0^1 \varphi^2(y)dy.$$

As a result,

$$\langle \tilde{\mu}, x^4 \rangle = \frac{2}{\varphi_0^2} \int_0^1 \varphi^2(x)dx > \frac{2}{\varphi_0^2} \left| \int_0^1 \varphi(x)dx \right|^2 = 2 = \#\mathcal{NC}_2(4) = \langle \sigma, x^4 \rangle$$

where the second step follows from Jensen's strict inequality, since φ is not constant \mathbb{A} -a.s.

b) ⇔ d) ⇒ c): We calculate

$$\begin{aligned} \varphi(x) &= \int_0^x w^2(x-y)dy + \int_x^1 w^2(y-x)dy \\ &= \int_0^x w^2(z)dz + \int_0^{1-x} w^2(z)dz \end{aligned} \tag{3.19}$$

$$\begin{aligned} &= \int_0^1 w^2(z)dz - \int_x^1 w^2(z)dz + \int_x^1 w^2(1-z)dz \\ &= \int_0^1 w^2(z)dz + \int_x^1 [w^2(1-z) - w^2(z)] dz \end{aligned} \tag{3.20}$$

Now let $v(z) := w^2(1-z) - w^2(z)$ for all $v \in [0, 1]$. Since v is Riemann integrable on $[0, 1]$, it is continuous on a set $C \subseteq [0, 1]$ with $\lambda(C) = 1$. But then with (3.20), φ is differentiable at every point $z \in C$ with derivative $w^2(z) - w^2(1-z)$.

Now if b) holds, that is, φ is constant, then

$$\forall z \in C : 0 = \varphi'(z) = w^2(z) - w^2(1-z),$$

which shows statement d). On the other hand, if d) holds, then the second integral in (3.20) vanishes, so φ is constant with $\varphi(x) = \int_0^1 w^2(z)dz$ for all $x \in [0, 1]$, which shows c).

c) ⇒ a) This is immediate. For (1), we start with equation (3.20) and obtain

$$\begin{aligned} \varphi_0 &= \int_0^1 \varphi(x)dx \\ &= \int_0^1 \int_0^1 w^2(z)dzdx + \int_0^1 \int_x^1 [w^2(1-z) - w^2(z)] dzdx \\ &= \int_0^1 w^2(z)dz + \int_0^1 \int_0^z [w^2(1-z) - w^2(z)] dx dz \\ &= \int_0^1 (1-z)w^2(z)dz + \int_0^1 zw^2(1-z)dz = 2 \int_0^1 (1-z)w^2(z)dz \end{aligned}$$

and the last integral is zero iff $w^2(z) = 0$ λ -a.s. iff $w(z) = 0$ λ -a.s. □

A Proof of Lemma 2.6

Proof. Let $\ell, n \in \mathbb{N}$ and $P_1, \dots, P_\ell \in [n]^2$ be fundamentally different, and let $\delta_1, \dots, \delta_\ell \in \mathbb{N}$ be arbitrary. Then we calculate (with explanations below)

$$\begin{aligned} \mathbb{E}a_n(P_1)^{\delta_1} \dots a_n(P_\ell)^{\delta_\ell} &= \mathbb{E} \prod_{j=1}^{\delta_1+\dots+\delta_\ell} a_n(P_{i(j)}) \\ &= \sum_{\pi \in \mathcal{P}_2(\delta_1+\dots+\delta_\ell)} \prod_{\{r,s\} \in \pi} \mathbb{E}a_n(P_{i(r)})a_n(P_{i(s)}) \stackrel{\leq}{|\dots|} \frac{\#\mathcal{P}_2(\delta_1+\dots+\delta_\ell)}{n^{\frac{1}{2}\#\{i \mid \delta_i=1\}}} \end{aligned}$$

For the first step, we set $i(1) = \dots = i(\delta_1) = 1, i(\delta_1 + 1) = \dots = i(\delta_1 + \delta_2) = 2$, and so on. In the second step, we apply Isserlis formular, see [17] or [24]. The third step holds after taking the absolute value on the l.h.s., and then the inequality follows since for all $\pi \in \mathcal{P}_2(\delta_1 + \dots + \delta_\ell)$,

$$\prod_{\{r,s\} \in \pi} |\mathbb{E}a_n(P_{i(r)})a_n(P_{i(s)})| = \prod_{\{r,s\} \in \pi} |\Sigma_n(P_{i(r)}, P_{i(s)})| \leq \frac{1}{n^{\frac{1}{2}\#\{i \mid \delta_i=1\}}},$$

since in the worst case, all single random variables are paired by the partition π . This proves (AU1) with constants $C(\ell) := \#\mathcal{P}_2(\ell)$, and for (AU2) we calculate

$$\begin{aligned} \mathbb{E}a_n(P_1)^2 \cdots a_n(P_\ell)^2 &= \mathbb{E} \prod_{j=1}^{2\ell} a_n(P_{i(j)}) = \sum_{\pi \in \mathcal{P}_2(2\ell)} \prod_{\{r,s\} \in \pi} \mathbb{E}a_n(P_{i(r)})a_n(P_{i(s)}) \\ &= 1 + \sum_{\pi \in \mathcal{P}_2(2\ell) \setminus \{\pi^*\}} \prod_{\{r,s\} \in \pi} \mathbb{E}a_n(P_{i(r)})a_n(P_{i(s)}), \end{aligned}$$

where in the first step we set $i(1) = i(2) = 1, i(3) = i(4) = 2$, etc., in the second step we apply Isserlis' formula, and for the third step we write $\pi^* = \{\{1, 2\}, \dots, \{2\ell - 1, 2\ell\}\}$. Since each $\pi \in \mathcal{P}_2(2\ell) \setminus \{\pi^*\}$ has at least two blocks which do not pair the same index, we find

$$\prod_{\{r,s\} \in \pi} |\mathbb{E}a_n(P_{i(r)})a_n(P_{i(s)})| \leq \prod_{\{r,s\} \in \pi} |\Sigma_n(P_{i(r)}, P_{i(s)})| \leq \frac{1}{n^2}.$$

Consequently, (AU2) holds with sequences $C_n^{(\ell)} := \mathcal{P}_2(2\ell)/n^2$, which are all summable over n . □

B Auxiliary lemmata

Lemma B.1. *Let $(I_n)_n$ be a sequence of finite index sets and let for all $n \in \mathbb{N}$, $(Y_n(i))_{i \in I_n}$ be a family of random variables with uniformly bounded absolute moments of all orders.*

a) *If $(m_n)_n$ is a sequence of positive real numbers with $\#I_n = o(m_n)$, then*

$$\frac{1}{m_n} \sum_{i \in I_n} Y_n(i) \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{in expectation and in probability.}$$

b) *If $(m_n)_n$ is a sequence of real numbers so that for some $p \in \mathbb{N}$, $(\#I_n/m_n)^p$ is summable (e.g. $m_n = n^\delta \#I_n$ for some $\delta > 0$), then*

$$\frac{1}{m_n} \sum_{i \in I_n} Y_n(i) \xrightarrow[n \rightarrow \infty]{} 0 \quad \text{in expectation and almost surely.}$$

Proof. Let $(D_k)_k$ be positive constants for all $k \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, $i \in I_n$, and $k \in \mathbb{N}$: $\mathbb{E}|Y_n(i)|^k \leq D_k$. Choose $\epsilon > 0$ and $k \in \mathbb{N}$ arbitrarily. Then

$$\begin{aligned} \mathbb{P} \left(\left| \frac{1}{m_n} \sum_{i \in I_n} Y_n(i) \right| > \epsilon \right) &\leq \left(\frac{\#I_n}{m_n} \right)^k \frac{1}{\epsilon^k} \cdot \frac{1}{(\#I_n)^k} \sum_{i \in I_n^k} \mathbb{E}|Y_n(i_1) \cdots Y_n(i_k)| \\ &\leq \left(\frac{\#I_n}{m_n} \right)^k \cdot \frac{D_k}{\epsilon^k}. \end{aligned}$$

For part a) choose $k = 1$, and for part b) choose $k = p$. Convergence in expectation to 0 is trivial. □

Lemma B.2. *Let $z \in \mathbb{N}$ and $(Y_n)_n$ be random variables with $\mathbb{E}|Y_n|^z < \infty$ for all $n \in \mathbb{N}$. If $\mathbb{E}Y_n \rightarrow y$ and $\mathbb{E}|Y_n - \mathbb{E}Y_n|^z \rightarrow 0$, then $Y_n \rightarrow y$ in probability. If in addition, $\mathbb{E}|Y_n - \mathbb{E}Y_n|^z$ is summable, then $Y_n \rightarrow y$ almost surely.*

Proof. Using Markov's inequality, we calculate for $\epsilon > 0$ arbitrary:

$$\begin{aligned} \mathbb{P}(|Y_n - y| > \epsilon) &\leq \mathbb{P} \left(|Y_n - \mathbb{E}Y_n| > \frac{\epsilon}{2} \right) + \mathbb{P} \left(|\mathbb{E}Y_n - y| > \frac{\epsilon}{2} \right) \\ &\leq \frac{2^z}{\epsilon^z} \mathbb{E}|Y_n - \mathbb{E}Y_n|^z + \mathbb{P} \left(|\mathbb{E}Y_n - y| > \frac{\epsilon}{2} \right). \end{aligned}$$

The statement follows (also using Borel-Cantelli), since the very last summand vanishes for all n large enough. □

References

- [1] Oskari H. Ajanki, László Erdős, and Torben Krüger. “Local Spectral Statistics of Gaussian Matrices with Correlated Entries”. In: *Journal of Statistical Physics* 163.2 (Apr. 2016), pp. 280–302. MR3478311
- [2] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, 2010. MR2760897
- [3] Marwa Banna, Florence Merlevède, and Magda Peligrad. “On the limiting spectral distribution for a large class of symmetric random matrices with correlated entries”. In: *Stochastic Processes and their Applications* 125.7 (July 2015), pp. 2700–2726. MR3332852
- [4] Leonid Bogachev, Stanislav Molchanov, and Leonid Pastur. “On the level density of random band matrices”. In: *Mathematical Notes* (1991). MR1150631
- [5] Vladimir Bogachev. *Measure Theory*. Vol. 2. Springer, 2006. MR2267655
- [6] Riccardo Catalano. “On Weighted Random Band-Matrices with Dependencies”. PhD thesis. FernUniversität in Hagen, Germany, 2016. URL: <https://arxiv.org/pdf/1605.03349.pdf>.
- [7] Ziliang Che. “Universality of random matrices with correlated entries”. In: *Electronic Journal of Probability* 22.30 (2017), pp. 1–38. MR3629874
- [8] László Erdős, Torben Krüger, and Dominik Schröder. “Random Matrices with Slow Correlation Decay”. In: *Forum of Mathematics, Sigma* 7 (2019), e8. MR3941370
- [9] Michael Fleermann. “Global and Local Semicircle Laws for Random Matrices with Correlated Entries”. PhD thesis. FernUniversität in Hagen, Germany, 2019. DOI: 10.18445/20190612-122137-0.
- [10] Michael Fleermann. “The empirical spectral distribution of symmetric random matrices with correlated entries. An asymptotic analysis employing the method of moments.” MA thesis. University of Münster, Germany, 2015.
- [11] Michael Fleermann and Werner Kirsch. “Proof methods in random matrix theory”. In: *Probability Surveys* 20 (2023), pp. 291–381. DOI: 10.1214/23-PS16. URL: <https://doi.org/10.1214/23-PS16>. MR4563528
- [12] Michael Fleermann, Werner Kirsch, and Thomas Kriecherbauer. “Local Semicircle Law for Curie-Weiss Type Ensembles”. Apr. 2021. URL: <https://arxiv.org/abs/1907.08782>. MR4406239
- [13] Michael Fleermann, Werner Kirsch, and Thomas Kriecherbauer. “The Almost Sure Semicircle Law for Random Band Matrices with Dependent Entries”. In: *Stochastic Processes and their Applications* 131 (2021), pp. 172–200. MR4153882
- [14] Olga Friesen and Matthias Löwe. “A phase transition for the limiting spectral density of random matrices”. In: *Electronic Journal of Probability* 18.17 (2013), pp. 1–17. MR3035745
- [15] Olga Friesen and Matthias Löwe. “The semicircle law for matrices with independent diagonals”. In: *Journal of Theoretical Probability* 26.4 (2013). MR3119985
- [16] Winfried Hochstättler, Werner Kirsch, and Simone Warzel. “Semicircle Law for a Matrix Ensemble with Dependent Entries”. In: *Journal of Theoretical Probability* (2016). URL: <https://dx.doi.org/10.1007/s10959-015-0602-3>. MR3540489
- [17] Leon Isserlis. “On a formula for the product-moment coefficient of each order of a normal frequency distribution in every number of variables”. In: *Biometrika* 12 (1918), pp. 134–139.
- [18] Todd Kemp and David Zimmermann. “Random matrices with log-range correlations, and log-Sobolev inequalities”. In: *Annales Mathématiques Blaise Pascal* 27.2 (2020), pp. 207–232. MR4231819
- [19] Werner Kirsch. “A Survey on the Method of Moments”. Oct. 2015. URL: <https://www.fernuni-hagen.de/stochastik/docs/pub/momente.pdf>.
- [20] Werner Kirsch and Thomas Kriecherbauer. “Random matrices with exchangeable entries”. In: *Reviews in Mathematical Physics* 32.7 (2020). MR4127501
- [21] Werner Kirsch and Thomas Kriecherbauer. “Semicircle Law for Generalized Curie-Weiss Matrix Ensembles at Subcritical Temperature”. In: *Journal of Theoretical Probability* 31.4 (2018), pp. 2446–2458. MR3866621

- [22] Matthias Löwe. “The semicircle law for matrices ergodic entries”. In: *Statistics & Probability Letters* 141 (2018). MR3825490
- [23] S.A. Molchanov, L.A. Pastur, and A.M. Khorunzhy. “Limiting eigenvalue distribution for random band matrices”. In: *The Annals of Probability* 90 (1992), pp. 108–118.
- [24] Alexandru Nica and Roland Speicher. *Lectures on the Combinatorics of Free Probability*. Cambridge University Press, 2006. MR2266879
- [25] Jeffrey H. Schenker and Hermann Schulz-Baldes. “Semicircle law and freeness for random matrices with symmetries or correlations”. In: *Mathematical Research Letters* 12 (2005), pp. 531–542. MR2155229
- [26] Terence Tao. *Topics in Random Matrix Theory*. American Mathematical Society, 2012. MR2906465
- [27] Yizhe Zhu. “A graphon approach to limiting spectral distributions of Wigner-type matrices”. In: *Random Structures & Algorithms* 56 (2020), pp. 251–279. MR4052853

Acknowledgments. We are grateful for the feedback from the referee which helped to improve this paper. Further, this work was carried out while the second author was still employed at the FernUniversität in Hagen, Germany.