GRAPHONS, PERMUTONS AND THE THOMA SIMPLEX: THREE MOD-GAUSSIAN MODULI SPACES

VALENTIN FÉRAY, PIERRE-LOÏC MÉLIOT, AND ASHKAN NIKEGHBALI

ABSTRACT. In this paper, we show how to use the framework of mod-Gaussian convergence in order to study the fluctuations of certain models of random graphs, of random permutations and of random integer partitions. We prove that, in these three frameworks, a generic homogeneous observable of a generic random model is mod-Gaussian under an appropriate renormalisation. This implies a central limit theorem with an extended zone of normality, a moderate deviation principle, an estimate of the speed of convergence, a local limit theorem and a concentration inequality. The universal asymptotic behavior of the observables of these models gives rise to a notion of mod-Gaussian moduli space.

CONTENTS

1. INTRODUCTION

The purpose of this paper is to show that many random models stemming from combinatorics exhibit asymptotic fluctuations which one can study in the framework of *mod-Gaussian*

convergence. We start our introduction by a discussion of this notion and of the probabilistic estimates which follow from it.

1.1. **Mod-Gaussian convergence, cumulants and dependency graphs.** Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real random variables. The notion of mod-Gaussian convergence was introduced in [JKN11], and later developed in [DKN15; MN15; FMN16; FMN19; BMN19]. Mod-Gaussian convergence yields precise quantitative information about the convergence in distribution of an appropriate renormalisation of X_n towards the standard Gaussian law: we get normality zone estimates, precise moderate deviations, bounds on the speed of convergence, or local limit theorems.

Though the original definition only involved the Fourier transforms $\mathbb{E}[e^{i\xi X_n}]$, in this paper, we shall work with complex exponential generating functions (Laplace transforms):

Definition 1 (Mod-Gaussian convergence). *The sequence of random variables $(X_n)_{n\in\mathbb{N}}$ converges in the mod-Gaussian sense with parameters $t_n \to +\infty$ and limit $\psi(z)$ if, locally uniformly on the complex plane,*

$$\psi_n(z) := \mathbb{E}[e^{zX_n}]\, e^{-\frac{t_n z^2}{2}} \to_{n\to\infty} \psi(z),$$

where $\psi(z)$ is a continuous function with $\psi(0) = 1$.

In [FMN16], the local uniform convergence $\psi_n \to \psi$ was only asked to occur on a band $\mathcal{S}_{(c,d)} = \{z \in \mathbb{C} \,|\, c < \mathrm{Re}(z) < d\}$), but in all instances of mod-Gaussian convergence in this paper, the convergence holds on the whole complex plane, so we restrict to this case here. In particular, we only consider random variables X_n with an *entire* exponential generating function.

One of the ways to prove the mod-Gaussian convergence of $(X_n)_{n\in\mathbb{N}}$ consists in introducing an adequate renormalisation S_n of X_n and estimating the coefficients of the Taylor expansion of $\log(\mathbb{E}[e^{zS_n}])$, which are called the *cumulants* of S_n. Formally, if

$$\log \mathbb{E}[e^{zS_n}] = \sum_{r=1}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!}\, z^r$$

for $|z|$ sufficiently small, then the coefficient $\kappa^{(r)}(S_n)$ is called *cumulant of order r of S_n*. In [FMN16; FMN19], we explain how estimates on cumulants can be used to prove mod-Gaussian convergence:

Definition 2 (Method of cumulants). *Let $(S_n)_{n\in\mathbb{N}}$ be a sequence of real random variables. Let A be a positive constant, and $(D_n)_{n\in\mathbb{N}}$ and $(N_n)_{n\in\mathbb{N}}$ be two positive sequences such that $\lim_{n\to\infty} \frac{D_n}{N_n} = 0$. We say that the sequence $(S_n)_{n\in\mathbb{N}}$ satisfies the hypotheses of the method of cumulants with parameters (D_n, N_n, A) and limits (σ^2, L) if:*

- *For any $r \geq 2$, we have the uniform bounds on cumulants (see [FMN19, Definition 28]):*

$$|\kappa^{(r)}(S_n)| \leq N_n\, (2D_n)^{r-1}\, r^{r-2}\, A^r. \tag{MC1}$$

- *There exists two real numbers $\sigma^2 \geq 0$ and L such that:*

$$\frac{\kappa^{(2)}(S_n)}{N_n\, D_n} = (\sigma_n)^2 = \sigma^2 \left(1 + o\left(\left(\frac{D_n}{N_n}\right)^{1/3}\right)\right); \tag{MC2}$$

$$\frac{\kappa^{(3)}(S_n)}{N_n\, (D_n)^2} = L_n = L\,(1 + o(1)). \tag{MC3}$$

In particular, $\sigma_n \to_{n\to\infty} \sigma$ and $L_n \to_{n\to\infty} L$.

In the method of cumulants, the hypotheses (MC2) and (MC3) are usually proved by an *ad hoc* computation of the first moments of S_n, and by using the identities

$$\kappa^{(2)}(X) = \mathrm{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}\left[\overline{X}^2\right];$$

$$\kappa^{(3)}(X) = \mathbb{E}[X^3] - 3\,\mathbb{E}[X^2]\,\mathbb{E}[X] + 2(\mathbb{E}[X])^3 = \mathbb{E}\left[\overline{X}^3\right],$$

where $\overline{X} = X - \mathbb{E}[X]$. On the other hand, we shall explain in a moment how to obtain uniform bounds on cumulants from *sparse dependency graphs*. Let us first detail the probabilistic consequences of the method of cumulants. Recall that the Kolmogorov distance between two random variables (or their distribution) is $d_{\mathrm{Kol}}(X, Y) = \sup_{t \in \mathbb{R}} |F_X(t) - F_Y(t)|$, where F_X and F_Y are the cumulative distribution functions of X and Y.

Theorem 3 (Estimates from the method of cumulants). *Let $(S_n)_{n \in \mathbb{N}}$ be a sequence of random variables that satisfies the hypotheses of the method of cumulants, with parameters (D_n, N_n, A) and limits (σ^2, L). We suppose that $\sigma^2 > 0$, and we set:*

$$X_n = \frac{S_n - \mathbb{E}[S_n]}{(N_n)^{1/3}\,(D_n)^{2/3}} \qquad ; \qquad Y_n = \frac{S_n - \mathbb{E}[S_n]}{\sigma_n\,\sqrt{N_n\,D_n}} = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{var}(S_n)}}.$$

(0) The random variables $(X_n)_{n \in \mathbb{N}}$ converge in the mod-Gaussian sense, with parameters $t_n = (\sigma_n)^2(\frac{N_n}{D_n})^{1/3}$ and limit $\psi(z) = \exp(\frac{Lz^3}{6})$.

(1) Central limit theorem: we have the convergence in distribution $Y_n \rightharpoonup_{n \to \infty} \mathcal{N}_{\mathbb{R}}(0, 1)$.

(2) Extended zone of normality: if $y_n = o((\frac{N_n}{D_n})^{1/6})$, then

$$\mathbb{P}[Y_n \geq y_n] = \mathbb{P}[\mathcal{N}_{\mathbb{R}}(0, 1) \geq y_n]\,(1 + o(1)).$$

(3) Moderate deviations (see [FMN16, Equation (9.10)]): for any sequence $(y_n)_{n \in \mathbb{N}}$ that tends to $+\infty$ but is a $o((\frac{N_n}{D_n})^{1/4})$,

$$\mathbb{P}[Y_n \geq y_n] = \frac{e^{-\frac{(y_n)^2}{2}}}{y_n\sqrt{2\pi}}\,\exp\left(\frac{L}{6\sigma^3}\sqrt{\frac{D_n}{N_n}}\,(y_n)^3\right)(1 + o(1)).$$

By symmetry, we have similar results for the negative deviations.

(4) The Kolmogorov distance between Y_n and the standard Gaussian law is bounded by

$$d_{\mathrm{Kol}}(Y_n, \mathcal{N}_{\mathbb{R}}(0, 1)) \leq \frac{76.36\,A^3}{(\sigma_n)^3}\sqrt{\frac{D_n}{N_n}},$$

see [FMN19, Corollary 30].

(5) We have the following local limit theorem: for any exponent $\delta \in (0, \frac{1}{2})$, any $y \in \mathbb{R}$ and any Jordan measurable set B with Lebesgue measure $m(B) > 0$,

$$\lim_{n \to \infty}\left(\frac{N_n}{D_n}\right)^{\delta}\mathbb{P}\left[Y_n - y \in \left(\frac{D_n}{N_n}\right)^{\delta}B\right] = \frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,m(B),$$

see [BMN19, Theorem 9 and Proposition 20].

It is important for the sequel to note that we did not assume $\sigma^2 > 0$ in Definition 2, but that the positivity is needed for the probabilistic estimates. On the other hand, all the results but

the moderate deviation estimates stay true if one assumes only the convergences $\sigma_n \to \sigma$ and $L_n \to L$ (instead of (MC2)). The stronger hypothesis

$$(\sigma_n)^2 = \sigma^2 \left(1 + o\left(\left(\frac{D_n}{N_n} \right)^{1/3} \right) \right)$$

is only needed in order to get the moderate deviation estimates when $y_n = o((\frac{D_n}{N_n})^{\frac{1}{6}+\varepsilon})$ with $\varepsilon \in (0, \frac{1}{12})$ — when $\varepsilon = 0$, the convergence $\sigma_n \to \sigma$ is sufficient.

Theorem 3 shows that the method of cumulants yields much more information than a classical central limit theorem. Besides, in [FMN16, Chapters 9-11] and [FMN19, Sections 4-5], we developed combinatorial tools in order to prove the uniform bounds on cumulants given by Equation (MC1).

Definition 4 (Dependency graphs). *Let $S = \sum_{v \in V} A_v$ be a finite sum of real random variables. We say that a (undirected, simple) graph $G = (V, E)$ is a dependency graph for the family of random variables $(A_v)_{v \in V}$ if the following condition holds: given two disjoint subsets $V_1, V_2 \subset V$, if there is no edge $e = (v, w) \in E$ such that $v \in V_1$ and $w \in V_2$, then $(A_v)_{v \in V_1}$ and $(A_w)_{w \in V_2}$ are independent vectors.*

By abuse of language, we often say that G is a dependency graph for the sum $S = \sum_{v \in V} A_v$ (instead of a dependency graph for the family $(A_v)_{v \in V}$). The notion of dependency graphs leads to the following criterion for the method of cumulants; see [FMN16, Theorems 9.1.7 and 9.1.8].

Theorem 5 (Uniform bounds on cumulants from a dependency graph). *Let $S = \sum_{v \in V} A_v$ be a sum of random variables with*

$$|A_v| \leq A \quad almost \ surely$$

for any $v \in V$. We assume that S admits a dependency graph, and we denote $N = |V|$ and $D = 1 + \max_{v \in V}(\deg v)$. Then, we have the following bound on the cumulants of S:

$$\forall r \geq 1, \quad |\kappa^{(r)}(S)| \leq N (2D)^{r-1} r^{r-2} A^r.$$

Moreover, this upper bound actually holds with a constant $N = \frac{1}{A} \sum_{v \in V} \mathbb{E}[|A_v|]$, which is in general smaller than $|V|$.

As a consequence, if $(S_n)_{n \in \mathbb{N}}$ is a sequence of sums of bounded random variables, and if each sum S_n admits a dependency graph G_n with parameters D_n, N_n and A, then the uniform bound (MC1) holds, and the condition $\lim_{n \to \infty} \frac{D_n}{N_n} = 0$ amounts to the fact that the graphs G_n are sparse. We shall see in this article that certain natural observables of models of random graphs/permutations/partitions can be related to sums of dependent random variables with sparse dependency graphs, and therefore that these models typically exhibit mod-Gaussian convergence. This implies new asymptotic results for these models, for instance speed of convergence estimates and moderate deviations; to the best of our knowledge, all these results are new.

Remark. Moderate deviations and Kolmogorov distance estimates under bounds of cumulants were first obtained by Saulis and Statulevičius in [SS91, Section 2], though under a less explicit form. Bounds on the Kolmogorov distance in the context of dependency graphs can alternatively be obtained through Stein's method, see *e.g.* [Rin94].

Remark (Weighted dependency graphs). In [FMN19, Section 5], we developed a more general method in order to prove uniform bounds on cumulants, which relied on the notion of weighted dependency graphs. Here, we shall not need to use these more complex arguments.

1.2. **Random graphs, random permutations and random partitions.** Let us now present informally the combinatorial random models that we shall study in this paper. We start with random graphs and the theory of graphons. In the following, unless stated otherwise, a *graph* will always be a finite undirected simple graph, that is to say a pair $G = (V_G, E_G)$ with V_G finite set of *vertices*, and E_G subset of the set $\mathfrak{P}_2(V_G)$ of pairs of vertices. Thus, E_G is a finite set of pairs $\{v_1, v_2\}$ with $v_1, v_2 \in V_G$ and $v_1 \neq v_2$. These pairs are the *edges* of the graph.
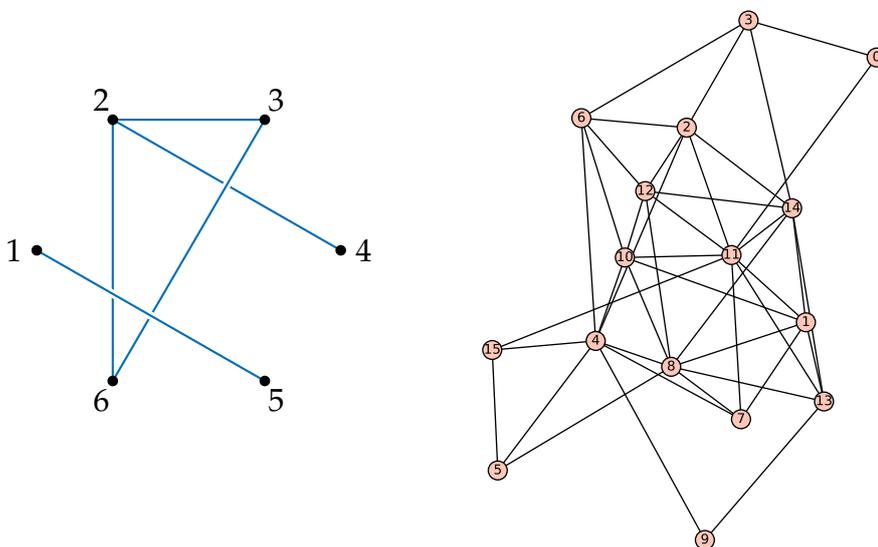

FIGURE 1. Two graphs G having each a density $\frac{1}{20}$ of triangles.

When looking at a large (random) graph G, a way to speak of convergence is to look at the numbers of subgraphs that appear in G. In the left picture of Figure 1, there is only one triangle, formed by the vertices $\{2, 3, 6\}$. As the number of possible triangles in a graph with 6 vertices is $\binom{6}{3} = 20$, the *density of triangles* $t_0(K_3, G)$ in G is $\frac{1}{20}$. The right picture of Figure 1 shows a larger graph, with 16 vertices and again a density of triangles equal to $\frac{1}{20}$ (there are 28 triangles, in this graph, among the $\binom{16}{3} = 560$ possible ones).

We will be interested in sequences $(G_n)_{n \in \mathbb{N}}$ of graphs, for which the density of triangles $(t_0(K_3, G_n))_{n \in \mathbb{N}}$ has a limit in $[0, 1]$. More generally, for any finite graph F of size k, we can look at the density $t(F, G)$ of the subgraph F in G, which is defined by the number of occurrences of F in G, divided by the total number of possibilities for F to appear as a subgraph of G. The precise definition of this density involves the notion of morphism of graphs, and it will be given in Section 4; one can actually define two slightly different densities $t_0(F, G)$ and $t(F, G)$, but for the moment let us forget this subtlety. A convergent

sequence of graphs $(G_n)_{n\in\mathbb{N}}$ will then be a sequence of graphs such that for any F finite graph, $(t(F,G_n))_{n\in\mathbb{N}}$ has a limit in $[0,1]$.

This notion of convergence can also be *realized* as a convergence in some metric space. This construction has been performed by Lovász and Szegedy in [LS06], and it leads to the theory of graphons. Namely, there exists a compact metrisable space $\mathscr{G}$ called the *space of graphons*, in which one can embed any finite graph G, and such that a sequence of finite graphs $(G_n)_{n\in\mathbb{N}}$ has convergent densities $t(F,G_n)$ for any finite graph F if and only if there exists some parameter $\gamma \in \mathscr{G}$ such that $G_n \to \gamma$ in the space of graphons. Moreover, graphs are dense in the space $\mathscr{G}$; this is shown by constructing, for each graphon γ, a model of random graphs $(G_n(\gamma))_{n\in\mathbb{N}}$ that converges almost surely to γ [LS06, Section 2.6].

In the present paper, we prove that, for any graphon γ and fixed graph F, the observable $t(F,G_n(\gamma))$ are generically mod-Gaussian convergent after an appropriate renormalisation. In the previous sentence, the word "generically" means that there exists a universal renormalisation $\alpha_{n,F}\, t(F,G_n(\gamma))$ of the density $t(F,G_n(\gamma))$ which depends only on F, and such that for any graphon $\gamma \in \mathscr{G}$, we have a convergence in law

$$\alpha_{n,F}\, t(F,G_n(\gamma)) \rightharpoonup_{n\to\infty} \mathcal{N}_{\mathbb{R}}(0,\sigma^2(F,\gamma)).$$

When $\sigma^2(F,\gamma) > 0$, we have mod-Gaussian convergence through uniform bounds of cumulants and all the precise probabilistic estimates listed in Theorem 3. Then, the only pairs of parameters (F,γ) for which this asymptotic normality does not sit in the framework of mod-Gaussian convergence are those for which $\sigma^2(F,\gamma) = 0$. Notice that this singular case does not prohibit the existence of another renormalisation $\alpha'_{n,F}\, t(F,G_n(\gamma))$ which falls in the framework of Theorem 3. The theory that we shall develop will allow us:

- to describe the asymptotic fluctuations of random models $(G_n(\gamma))_{n\in\mathbb{N}}$ when these fluctuations have a typical size;

- to identify the parameters γ such that $\sigma^2(F,\gamma) = 0$ for all graphs F. These parameters γ will be called *globally singular* and they seem to exhibit some structure.

For instance, we shall see that the Erdös–Rényi random graphs are singular models with respect to any density of a subgraph; see the remark on singular points on page 50. Our theory is concerned with models of random graphs which are less symmetric and therefore more generic. It enables then a fine understanding of their asymptotic behavior.

An approach analogous to the theory of graphons can be used in order to deal with models of random permutations. Recall that a *permutation* of size n is a bijection $\sigma : [\![1,n]\!] \to [\![1,n]\!]$. The set of all permutations of size n is the *symmetric group* of order n, denoted $\mathfrak{S}(n)$, and with cardinality $n!$. If $\tau \in \mathfrak{S}(k)$ and $\sigma \in \mathfrak{S}(n)$ with $k \leq n$, we say that τ is a *pattern* in σ if there exists a subset $\{a_1 < a_2 < \cdots < a_k\} \subset [\![1,n]\!]$ such that $\sigma(a_i) < \sigma(a_j)$ if and only if $\tau(i) < \tau(j)$. This definition is better understood on a picture: if one draws the diagram of σ (graph of σ viewed as a function), then one can isolate points with abscissa $a_1 < a_2 < \cdots < a_k$ such that the restriction of the diagram of σ to these points is the diagram of the permutation τ; see Figure 2. In the previous example, there are 2 sets $\{a_1 < a_2 < a_3\}$ that make appear $\tau = 213$ as a pattern in $\sigma = 245361$, namely, $\{2,4,5\}$ and $\{3,4,5\}$. Therefore, it is natural to say that $\tau = 213$ has a pattern density $t(\tau,\sigma)$ equal to $\frac{2}{\binom{6}{3}} = \frac{1}{10}$ in the permutation σ.

We can extend the definition to any pattern τ, and exactly as for graphs, we then say that a sequence $(\sigma_n)_{n\in\mathbb{N}}$ of permutations converges if, for any finite permutation τ, the sequence of densities $(t(\tau,\sigma_n))_{n\in\mathbb{N}}$ has a limit in $[0,1]$. Again, this notion of convergence leads to
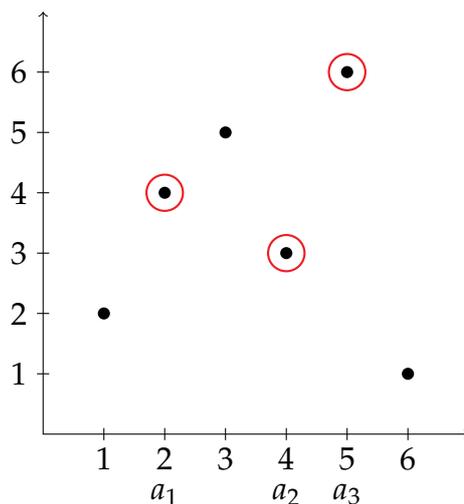
FIGURE 2. The permutation 213 is a pattern in $\sigma = 245361$.

the construction of a compact metrisable set $\mathscr{S}$ called the *space of permutons*, in which one can embed any finite permutation σ, and such that a sequence of permutations $(\sigma_n)_{n\in\mathbb{N}}$ has convergent densities $t(\tau, \sigma_n) \to t(\tau)$ for any pattern τ if and only if there exists some parameter $\pi \in \mathscr{S}$ such that $\sigma_n \to \pi$ in the space of permutons. The space of permutons was introduced by Hoppen *et al.* in [Hop+11; Hop+13]. Again, these authors prove that permutations are dense in $\mathscr{S}$ by constructing, for each permuton π, a model of random permutations $(\sigma_n(\pi))_{n\in\mathbb{N}}$ such that $\sigma_n(\pi)$ tends to π almost surely [Hop+13, Section 4].

In this paper, we shall prove that for any pattern τ and permuton $\pi \in \mathscr{S}$, the observable $t(\tau, \sigma_n(\pi))$ are generically mod-Gaussian convergent after an appropriate renormalisation. Again, this statement implies a central limit theorem for the densities of patterns $t(\tau, \sigma_n(\pi))$, and additional results such as a bound on the speed of convergence or a concentration inequality (see next section).

A third class of models consists in random integer partitions which stem from random permutations obtained by shuffle of cards, and which are related to the representation theory of the infinite symmetric group $\mathfrak{S}(\infty)$ (see [Mél12] and [Mél17, Chapter 12]). The mod-Gaussian convergence of some observables of these models was established in [FMN16, Chapter 11], but we take here a different approach to emphasize similarities with graphons and permutons.

Recall that an *integer partition* of size n is a non-increasing sequence $\lambda = (\lambda_1 \geq \cdots \geq \lambda_r)$ of positive integers with $\lambda_1 + \lambda_2 + \cdots + \lambda_r = n$. The set of all integer partitions of size n is denoted $\mathfrak{P}(n)$, and an integer partition λ is usually represented by its Young diagram, which is the array of boxes with λ_1 cells on its first bottom row, λ_2 cells on its second row, *etc.*

The *Frobenius coordinates* of λ are the half-integers $(a_1, a_2, \ldots, a_d; b_1, b_2, \ldots, b_d)$ defined as follows: d is the size of the diagonal of the Young diagram of λ, and

$$a_i = \lambda_i - i + \frac{1}{2} \qquad ; \qquad b_i = \lambda'_i - i + \frac{1}{2},$$

where λ' is the partition obtained from λ by symetrising the Young diagram with respect to its diagonal (see *e.g.* [Mél17, Section 7.2]). For instance, when $\lambda = (5, 4, 2)$, one obtains as
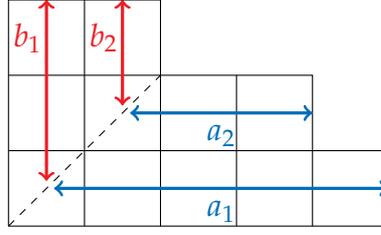
FIGURE 3. The Young diagram and the Frobenius coordinates of the integer partition $(5, 4, 2)$ of size $n = 11$.

Frobenius coordinates

$$a_1 = \frac{9}{2}, \ a_2 = \frac{5}{2} \qquad ; \qquad b_1 = \frac{5}{2}, \ b_2 = \frac{3}{2}.$$

We refer to Figure 3 for a geometric interpretation of the Frobenius coordinates. They allow one to see any integer partition as an element of the *Thoma simplex*, which is the bi-infinite dimensional simplex

$$\mathscr{P} = \left\{ (\alpha, \beta) = ((\alpha_1 \geq \alpha_2 \geq \cdots \geq 0), (\beta_1 \geq \beta_2 \geq \cdots \geq 0)) \mid \sum_{i=1}^{\infty} (\alpha_i + \beta_i) \leq 1 \right\}.$$

Indeed, one can associate with $\lambda \in \mathfrak{P}(n)$ the two sequences

$$\alpha(\lambda) = \left(\frac{a_1(\lambda)}{n}, \frac{a_2(\lambda)}{n}, \ldots, \frac{a_d(\lambda)}{n}, 0, 0, \ldots \right) ;$$

$$\beta(\lambda) = \left(\frac{b_1(\lambda)}{n}, \frac{b_2(\lambda)}{n}, \ldots, \frac{b_d(\lambda)}{n}, 0, 0, \ldots \right),$$

and the pair $\omega(\lambda) = (\alpha(\lambda), \beta(\lambda))$ belongs to $\mathscr{P}$. We equip the Thoma simplex $\mathscr{P}$ with the pointwise convergence topology, *i.e.* a sequence $(\alpha^{(n)}, \beta^{(n)})$ converges if for each $i \geq 1$, the coordinates $\alpha_i^{(n)}$ and $\beta_i^{(n)}$ converge. As we will see later in Proposition 18, this is equivalent to the convergence of the following observables, to which we will refer as *Frobenius moments*: for $k \geq 2$, let

$$t(k, \omega = (\alpha, \beta)) := \sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1} \sum_{i=1}^{\infty} (\beta_i)^k.$$

As for graphons and permutons, any parameter $\omega \in \mathscr{P}$ defines, for each $n \geq 1$, a model of random integer partitions of size n, denoted $(\lambda_n(\omega))_{n \in \mathbb{N}}$. Namely, the distribution of $\lambda_n(\omega)$ is given by

$$\mathbb{P}[\lambda_n(\omega) = \lambda] = (\dim \lambda) \, s_\lambda(\omega),$$

where $\dim \lambda$ is the number of standard tableaux with shape λ, and $s_\lambda(\omega)$ is some speciali-sation of the Schur function s_λ associated with the parameter ω; see Section 3.3 for details on the construction of the random partitions $\lambda_n(\omega)$ and for their combinatorial interpre-tation. Kerov and Vershik [KV81] showed that for any $\omega = (\alpha, \beta)$ in $\mathscr{P}$, one has conver-gence in probability $\lambda_n(\omega) \to \omega$ in the Thoma simplex. We shall prove that the observables $t(k, \lambda_n(\omega))$ are generically mod-Gaussian convergent after an appropriate renormalisation. This completes a previous result of [FMN16, Chapter 11], where the same result was proven for the random character values $\chi^{\lambda_n(\omega)}(c_k)$.

1.3. **Concentration inequalities.** The theory of graphons and permutons developed in the papers [LS06; LS07; Bor+08; Hop+11; Hop+13] relies on certain concentration inequalities for the observables of the random graphs or permutations; see for instance [Bor+08, Lemma 4.4] and [Hop+11, Theorem 4.2]. One of the objective of this paper is to show that these inequalities are in fact easy consequences of the mod-Gaussian structure. Indeed, under the hypotheses of the method of cumulants, we can state several concentration inequalities which show that the random variables considered are uniformly sub-Gaussian. We start with an inequality which only involves the parameters (D_n, N_n, A) of the method of cumulants:

Proposition 6 (Concentration inequality). *Let $(S_n)_{n \in N}$ be a sequence of random variables such that* (MC1) *holds for some parameters (D_n, N_n, A). We also assume that $|S_n| \leq N_n A$ almost surely (this is for instance the case if S_n is a sum of bounded random variables with a dependency graph with parameters (D_n, N_n, A)). For any $x > 0$,*

$$\mathbb{P}[|S_n - \mathbb{E}[S_n]| \geq x] \leq 2 \exp\left(-\frac{x^2}{9 \, D_n N_n A^2}\right).$$

We can also write a concentration inequality which involves explicitly the variance of S_n, and which is more precise for small fluctuations:

Proposition 7 (Concentration inequality involving the variance). *With the exact same assumptions as in Proposition 6, we have for any $x > 0$*

$$\mathbb{P}[|S_n - \mathbb{E}[S_n]| \geq x] \leq 2 \exp\left(-\frac{x^2}{2\left(\text{var}(S_n) + 2e \, D_n N_n A^2 \sqrt{\frac{x}{N_n A}}\right)}\right).$$

Both inequalities are proved in Section 2.1 by using a Chernoff bound. We shall see in Theorems 21 and 26 below that the aforementioned concentration inequalities for graphon and permuton models follow immediately from Proposition 6, though with different constants than in their original form. Besides, we shall obtain new concentration inequalities for the observables of random integer partitions (Theorem 12).

In the case where the bounds on cumulants are obtained through a (sparse) dependency graph, a better bound than Proposition 6 has been given by Janson in [Jan04], with instead of D_n the (fractional) chromatic number $\chi^*(G_n)$ of the dependency graph G_n of S_n. Proposition 6 is therefore particularly interesting in the cases where there is no underlying sparse dependency graph. An example of this is Theorem 12 below for random partitions under central measures: indeed, there is no standard dependency graph in this setting, but one in a noncommutative probability space, to which Janson's result does not apply (see [FMN16, Chapter11] for the construction of this graph). Another example without underlying (sparse) dependency graph is presented in Section 2.2: we give a concentration inequality for the magnetization in the Ising model, based on the bounds on cumulants given in [DIS74] and [FMN19, Section 5].

1.4. **Main results and outline of the paper.** Let us now detail the content of this article. We start by stating some new results on the models that were informally introduced in Section 1.2. These theorems can serve as a guideline for the reader, and they give a good idea of the kind of results that one can obtain in the framework of mod-Gaussian convergence. Moreover, these asymptotic results seem to be somehow universal for random combinatorial models with a concentration property. We start with the results on random graphs:

Theorem 8 (Speed of convergence for graphon models). *Let $\gamma \in \mathscr{G}$ be a graphon, and $G_n(\gamma)$ be the random graph of size n associated with it (see Section 3.1 for details). If F is a finite graph with size k, then the rescaled subgraph density*

$$Y_n(F, \gamma) = \frac{t(F, G_n(\gamma)) - \mathbb{E}[t(F, G_n(\gamma))]}{\sqrt{\operatorname{var}(t(F, G_n(\gamma)))}}$$

satisfies

$$d_{\mathrm{Kol}}(Y_n(F, \gamma), \mathcal{N}_{\mathbb{R}}(0, 1)) = O\left(\frac{1}{(n \operatorname{var}(t(F, G_n(\gamma))))^{3/2}} \frac{k^4}{\sqrt{n}}\right),$$

with a constant in the $O(\cdot)$ that is universal (it does not depend on γ or on F). The quantity $n \operatorname{var}(t(F, G_n(\gamma)))$ always admits a finite limit (see Proposition 20 for the computation of this limit), and if $\lim_{n\to\infty} n \operatorname{var}(t(F, G_n(\gamma))) > 0$, then the estimate above shows that $Y_n(F, \gamma)$ converges to a standard Gaussian distribution.

Theorem 9 (Moderate deviations for graphon models). *We use the same notation as in Theorem 8 and also assume*

$$\lim_{n\to\infty} n \operatorname{var}(t(F, G_n(\gamma))) > 0.$$

Then, then there is a real number $l(F, \gamma)$ such that, if $y_n = o(n^{1/4})$ and $y_n \to +\infty$, we have

$$\mathbb{P}[Y_n(F, \gamma) \geq y_n] = \frac{\mathrm{e}^{-\frac{(y_n)^2}{2}}}{y_n \sqrt{2\pi}} \exp\left(l(F, \gamma) \frac{(y_n)^3}{\sqrt{n}}\right) (1 + o(1)).$$

In particular, if $y_n = o(n^{1/6})$, we have $\mathbb{P}[Y_n(F, \gamma) \geq y_n] = \mathbb{P}[Z \geq y_n](1 + o(1))$, where Z is a standard Gaussian random variable, and $1/6$ is the maximal exponent such that this happens. We say that the zone of normality of $Y_n(F, \gamma)$ is $o(n^{1/6})$.

In these theorems, the condition $\lim_{n\to\infty} n \operatorname{var}(t(F, G_n(\gamma))) > 0$ will be easy to check, by evaluating a certain observable $\kappa_2(F, F)$ on the graphon γ. At first sight, Theorem 8 seems to imply a stronger result, namely, that one has convergence of $Y_n(F, \gamma)$ to a Gaussian distribution if $\lim_{n\to\infty} n^{4/3} \operatorname{var}(t(F, G_n(\gamma))) = \infty$. However, we shall see that $\operatorname{var}(t(F, G_n(\gamma)))$ can be expanded in integer powers of $\frac{1}{n}$, so the two aforementioned conditions are actually equivalent. The same remark will hold for the two other kinds of random models.

We have exact analogues of Theorems 8 and 9 for the models of random permutations:

Theorem 10 (Speed of convergence for random permutations). *Let $\pi \in \mathscr{S}$ be a permuton, and $\sigma_n(\pi)$ be the random permutation of size n associated with it (see Section 3.2 for details). If τ is a finite permutation with size k, then the rescaled pattern density*

$$Y_n(\tau, \pi) = \frac{t(\tau, \sigma_n(\pi)) - \mathbb{E}[t(\tau, \sigma_n(\pi))]}{\sqrt{\operatorname{var}(t(\tau, \sigma_n(\pi)))}}$$

satisfies

$$d_{\mathrm{Kol}}(Y_n(\tau, \pi), \mathcal{N}_{\mathbb{R}}(0, 1)) = O\left(\frac{1}{(n \operatorname{var}(t(\tau, \sigma_n(\pi))))^{3/2}} \frac{k^4}{\sqrt{n}}\right),$$

with a constant in the $O(\cdot)$ that is universal. The quantity $n \operatorname{var}(t(\tau, \sigma_n(\pi)))$ always admits a finite limite (see Proposition 25 for the computation of this limit), and if $\lim_{n\to\infty} n \operatorname{var}(t(\tau, \sigma_n(\pi))) > 0$, then $Y_n(\tau, \pi)$ converges to a standard Gaussian distribution.

Theorem 11 (Moderate deviations for random permutations). *We use the same notation as in Theorem 10 and also assume*

$$\lim_{n \to \infty} n \operatorname{var}(t(\tau, \sigma_n(\pi))) > 0.$$

Then there is a real number $l(\tau, \pi)$ which one can compute exactly, such that, if $y_n = o(n^{1/4})$ and $y_n \to +\infty$, we have

$$\mathbb{P}[Y_n(\tau, \pi) \geq y_n] = \frac{e^{-\frac{(y_n)^2}{2}}}{y_n \sqrt{2\pi}} \exp\left(l(\tau, \pi) \frac{(y_n)^3}{\sqrt{n}}\right) (1 + o(1)).$$

In particular, if $y_n = o(n^{1/6})$, we have $\mathbb{P}[Y_n(\tau, \pi) \geq y_n] = \mathbb{P}[Z \geq y_n](1 + o(1))$, where Z is a standard Gaussian random variable, and $1/6$ is the maximal exponent such that this happens. We say that the zone of normality *of $Y_n(\tau, \pi)$ is $o(n^{1/6})$.*

Last, let us state some new concentration results for the central measures on integer partitions:

Theorem 12 (Concentration inequalities for random integer partitions). *Let $\lambda_n(\omega)$ be the random integer partition with size n chosen under the central measure associated with a parameter $\omega = (\alpha, \beta)$ of the Thoma simplex (see Section 3.3 for details on the definition of these models). Given an integer k, we consider the k-th Frobenius moments*

$$t(k, \lambda_n(\omega)) = \frac{1}{n^k} \left(\sum_{i=1}^{d} (a_i(\lambda_n(\omega)))^k + (-1)^{k-1} \sum_{i=1}^{d} (b_i(\lambda_n(\omega)))^k\right);$$

$$t(k, \omega) = \sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1} \sum_{i=1}^{\infty} (\beta_i)^k.$$

We have

$$\mathbb{P}[|t(k, \lambda_n(\omega)) - t(k, \omega)| \geq x] \leq 4 \exp\left(-\frac{nx^2}{9k^2}\right).$$

We do not give concentration results for graphons and permutons since those are already present in the literature. Similarly, we chose not to state speed of convergence and moderate deviation results for random partitions; see [FMN16, Section 11] for results of this kind, though the observables considered there are different.

All the theorems stated previously fall in the framework of mod-Gaussian convergence, and as a consequence, their proofs are very similar. This paper is written in a way which emphasizes these similarities. In Section 3, we present the three compact spaces of parameters which correspond to classes of models of random graphs, of random permutations and of random integer partitions. We explain in each case how to build a random combinatorial object $O_n(m)$ with size n when m is a parameter in the compact space $\mathcal{M}$. We also explain how to see $O_n(m)$ as a random element $M_n(m)$ in $\mathcal{M}$. The asymptotic concentration of the models $(O_n(m))_{n \in \mathbb{N}}$ amounts then to the convergence in probability $M_n(m) \to m$ for any $m \in \mathcal{M}$.

In Section 4, we introduce for each space of parameters $\mathcal{M}$ an algebra $\mathcal{O}_{\mathfrak{M}}$ of continuous functions on $\mathcal{M}$ (observables), which has the following properties:

- The convergence in $\mathcal{M}$ is equivalent to the convergence of all the observables in $\mathcal{O}_{\mathfrak{M}}$.

- The fluctuations of the observables of the random models $O_n(m)$ are mod-Gaussian after an appropriate renormalisation, and the limiting parameters σ^2 and L of these mod-Gaussian sequences can be computed by using operations in the algebra $\mathscr{O}_{\mathfrak{M}}$.

It is quite remarkable that the topology of each class of objects can always be metrised by a nice algebraic structure, and that this structure also appears in the calculations of the fluctuations.

We prove the mod-Gaussian convergence of the observables in Section 5 (Theorems 21, 26 and 29), most of the time by exhibiting a dependency graph for the random variables under study. Though the random combinatorial objects and their constructions are very different, the dependency structures as well as the formulas for the limiting variances and limiting third cumulants are almost the same for the three cases (graphs, permutations and partitions). For instance, the variance of a graph density $t(F, G_n(\gamma))$ will be estimated by using a graph observable

$$\kappa_2(F, F) = \frac{1}{|F|^2} \sum_{1 \le a,b \le |F|} ((F \bowtie F)(a,b) - F \sqcup F),$$

whereas the variance of a pattern density $t(\tau, \sigma_n(\pi))$ will be estimated by using a permutation observable

$$\kappa_2(\tau, \tau) = \frac{1}{|\tau|^2} \sum_{1 \le a,b \le |\tau|} ((\tau \bowtie \tau)(a,b) - \tau \times \tau).$$

The notations used throughout the paper will make it easy to see these analogies. In our last Section 6, we discuss these similarities and we introduce a notion of *mod-Gaussian moduli space* for a pair $(\mathscr{M}, \mathscr{O}_{\mathfrak{M}})$. Our definition yields a rigorous meaning to genericity when studying the fluctuations of random combinatorial objects in a certain class. In particular, the random models with additional symmetries are usually non-generic, and they appear as singular points of the mod-Gaussian moduli spaces. To the best of our knowledge, this geometric approach of the study of classes of random models is a new viewpoint in probability theory.

2. CONCENTRATION INEQUALITIES

This section focuses on the concentration inequalities discussed in Section 1.3. We first prove the announced inequalities, and then give an application to the d-dimensional Ising model.

2.1. **Proofs of the concentration inequalities.**

Proof of Proposition 6. Setting $X = S_n - \mathbb{E}[S_n]$, we use Chernov's bound $\mathbb{P}[X \ge x] \le \frac{\mathbb{E}[e^{tX}]}{e^{tx}}$ and we shall choose later the optimal parameter $t > 0$. By assumption,

$$\log \mathbb{E}[e^{tX}] = \sum_{r=2}^{\infty} \frac{\kappa^{(r)}(S_n)}{r!} t^r \le \frac{N_n}{2D_n} \sum_{r=2}^{\infty} \frac{r^{r-2}}{r!} (2D_n At)^r$$

$$\le \frac{N_n}{2\sqrt{2\pi} D_n} \sum_{r=2}^{\infty} \frac{1}{r^{5/2}} (2eD_n At)^r$$

by using Stirling's bound $r! \geq \left(\frac{r}{e}\right)^r \sqrt{2\pi r}$. For $r \geq 2$, $r^{5/2} \geq \frac{\sqrt{27}}{2} r(r-1)$, so

$$\log \mathbb{E}[e^{tX}] \leq \frac{N_n}{\sqrt{54\pi}\, D_n} \sum_{r=2}^{\infty} \frac{\theta^r}{r(r-1)} = \frac{N_n}{\sqrt{54\pi}\, D_n} \left((1-\theta)\log(1-\theta) + \theta\right)$$

with $\theta = 2eD_n A t$, and assuming $\theta < 1$. We only lose a tiny bit with the Stirling approximation, and it allows one to keep explicit constants. If one substracts $tx = \frac{\theta\, x}{2eD_n A}$ from the previous upper bound and choose the optimal $\theta = 1 - e^{-\frac{\sqrt{54\pi}}{2e} \frac{x}{N_n A}}$, one obtains:

$$\log \mathbb{P}[X \geq x] \leq \frac{N_n}{\sqrt{54\pi}\, D_n} \left(\left(1 - e^{-\frac{\sqrt{54\pi}}{2e} \frac{x}{N_n A}}\right) - \frac{\sqrt{54\pi}}{2e} \frac{x}{N_n A} \right)$$

for any positive x. Notice now that the function $f(u) = \frac{1 - e^{-u} - u}{u^2}$ is negative and increasing. Moreover, in the previous bound, we can assume without loss of generality that $x \leq N_n A$, because $|X| \leq N_n A$ almost surely by hypothesis. Therefore, we can replace the previous bound by

$$\log \mathbb{P}[X \geq x] \leq \frac{N_n}{\sqrt{54\pi}\, D_n} f\left(\frac{\sqrt{54\pi}}{2e}\right) \left(\frac{\sqrt{54\pi}}{2e} \frac{x}{N_n A}\right)^2 \leq \frac{C\, x^2}{D_n N_n A^2},$$

where

$$C = \frac{\sqrt{54\pi}}{4e^2} f\left(\frac{\sqrt{54\pi}}{2e}\right) \leq -0.114,$$

which it is convenient to approximate from above by $-\frac{1}{9}$. Finally, since we can use the same bound for the variable $-S_n$ instead of S_n, we can bound $\mathbb{P}[|X| \geq x]$ by twice the bound on $\mathbb{P}[X \geq x]$. □

Proof of Proposition 7. With the same notations as in the proof of Proposition 6, we obtain by isolating the term of order $r = 2$ in the log-generating function:

$$\log \mathbb{E}[e^{tX}] - tx \leq \frac{N_n}{D_n} \left(\theta^3 \left(\sum_{r=3}^{\infty} \frac{r^{r-2}}{2e^r\, r!} \right) + \frac{\theta^2}{8e^2} \left(\frac{\sigma_n}{A}\right)^2 - \frac{\theta}{2e} \frac{x}{N_n A} \right)$$

with $(\sigma_n)^2 = \frac{\text{var}(S_n)}{N_n D_n}$ and $\theta = 2eD_n A t$, assuming $\theta \leq 1$. The convergent series $\alpha = \sum_{r=3}^{\infty} \frac{r^{r-2}}{2e^r\, r!}$ is smaller than 0.0323, and we have $6\,\alpha e \leq 1$. We therefore have the polynomial bound

$$\log \mathbb{E}[e^{tX}] - tx \leq \frac{N_n}{D_n} \left(\theta^3 \alpha + \frac{\theta^2}{8e^2} \left(\frac{\sigma_n}{A}\right)^2 - \frac{\theta}{2e} \frac{x}{N_n A} \right)$$

$$\leq \frac{N_n}{eD_n} \left(\theta^3 \alpha e + \frac{\theta^2}{8e} \left(\frac{\sigma_n}{A}\right)^2 - \frac{\theta}{2} \frac{x}{N_n A} \right)$$

$$\leq \frac{N_n}{eD_n} \left(\frac{\theta^3}{6} + \frac{\theta^2}{8e} \left(\frac{\sigma_n}{A}\right)^2 - \frac{\theta}{2} \frac{x}{N_n A} \right).$$

for $c \geq 1$. The optimal θ is the positive solution of $\theta^2 + \frac{\theta}{2e} \left(\frac{\sigma_n}{A}\right)^2 - \frac{x}{N_n A} = 0$, that is

$$\theta = \sqrt{\left(\frac{1}{4e} \left(\frac{\sigma_n}{A}\right)^2\right)^2 + \frac{x}{N_n A}} - \frac{1}{4e} \left(\frac{\sigma_n}{A}\right)^2 \leq \sqrt{\frac{x}{N_n A}} \leq 1.$$

Thus, this value of θ is indeed smaller than 1, and by using this value, we obtain:

$$
\log \mathbb{P}[X \geq x] \leq -\frac{N_n}{4\mathrm{e}D_n}\left(\frac{4\,\theta^3}{3} + \frac{\theta^2}{2\mathrm{e}}\left(\frac{\sigma_n}{A}\right)^2\right) \leq -\frac{N_n}{4\mathrm{e}D_n}\left(\theta^3 + \frac{\theta^2}{2\mathrm{e}}\left(\frac{\sigma_n}{A}\right)^2\right) = -\frac{\theta\,x}{4\mathrm{e}D_n A}
$$

$$
\leq -\frac{1}{\left(\frac{\sigma_n}{A}\right)^2 + \sqrt{\left(\frac{\sigma_n}{A}\right)^4 + \frac{16\mathrm{e}^2\,x}{N_n A}}}\,\frac{x^2}{N_n D_n A^2}
$$

$$
\leq -\frac{x^2}{2\left(\mathrm{var}(S_n) + 2\mathrm{e}N_n D_n A^2 \sqrt{\frac{x}{N_n A}}\right)}
$$

by using on the last line the inequality $a + \sqrt{a^2 + b} \leq 2a + \sqrt{b}$ for $a, b \geq 0$. As in the previous proof, the case of $\mathbb{P}[X \leq -x]$ can be treated with the same tools. $\qquad\square$

2.2. **Application to the d-dimensional Ising model.** We consider the d-dimensional Ising model with inverse temperature β and magnetic field h. We place ourselves either at very high temperature (i.e. $h = 0$, $\beta < \beta_0(d)$, for some threshold $\beta_0(d)$ smaller than the critical inverse temperature $\beta_c(d)$), or in presence of a magnetic field $h \neq 0$). A generic reference on the Ising model is, *e.g.*, [FV17, Chapter 3]. We consider the model on the full lattice $\mathbb{Z}^d$ and study the magnetization in a box Δ (which we will make grow to $\mathbb{Z}^d$), *i.e.*

$$
M_\Delta := \sum_{x \in \Delta} \sigma(x).
$$

As explained in [FMN19, Section 5.3] (see also [DIS74]), M_Δ admits uniform bounds on cumulants with parameters $(C_1, C_2\,|\Delta|, C_3)$, for some constant C_1, C_2, C_3 (these constants depend on the parameters (β, h), but not on the box $|\Delta|$). Applying Proposition 6 gives the following concentration inequality: for any $x > 0$, any $n \geq 1$, any $d \geq 2$ and any $\beta < \beta_0(d)$, we have

$$
\mathbb{P}\left[\frac{|M_\Delta - \mathbb{E}[M_\Delta]|}{\sqrt{|\Delta|}} \geq x\right] \leq 2\exp(-C(d,\beta)\,x^2),
$$

for some positive constant $C(d, \beta) > 0$. We recover the concentration inequalities for Gibbs measures from [Kül03, Theorem 1] and [Cha+07; CCR17]. Such inequalities hold in the Dobrushin uniqueness regime, for instance with $h \neq 0$ and at very high temperature.

3. The three spaces of parameters

In this section, we present three compact spaces of parameters which label models of random graphs, of random permutations and of random integer partitions. We try to use similar notations for each class of models, and we survey their theory which is relatively new for graphons and permutons.

3.1. **The space of graphons.** In this paragraph, we follow the discussion of [LS06; LS07; Bor+08; Bor+11]; we also refer to [Lov12] for a comprehensive treatment of the theory of graph homomorphisms and of limits of sequences of dense or spare graphs. We call *graph function* a measurable function $g : [0,1]^2 \to [0,1]$ that is symmetric: $g(x,y) = g(y,x)$ for almost any (x,y) (relatively to the Lebesgue measure on the square $[0,1]^2$). The set of graph functions will be denoted $\mathscr{F} \subset \mathrm{L}^\infty([0,1]^2, dx\,dy)$. Given a finite graph G with vertex set $[n]$,

we can associate with it a canonical graph function g. Thus, $g = g(G)$ is the function on the square that takes its values in $\{0, 1\}$, and is such that

$$g(x, y) = 1 \text{ if } x \in \left(\frac{i-1}{n}, \frac{i}{n}\right], \; y \in \left(\frac{j-1}{n}, \frac{j}{n}\right] \text{ and } \{i, j\} \in E_G,$$

and 0 otherwise. We refer to Figure 4 for an example. Note that the graph function depends on the labeling of the vertices of G, *i.e.* two isomorphic graphs G and G' may have different graph functions g and g'. The graph functions g and g' of isomorphic graphs are however related as follows: there exists a measure-preserving bijection $\sigma : [0, 1] \to [0, 1]$ such that $g' = g^\sigma$, where $g^\sigma(x, y) = g(\sigma(x), \sigma(y))$.
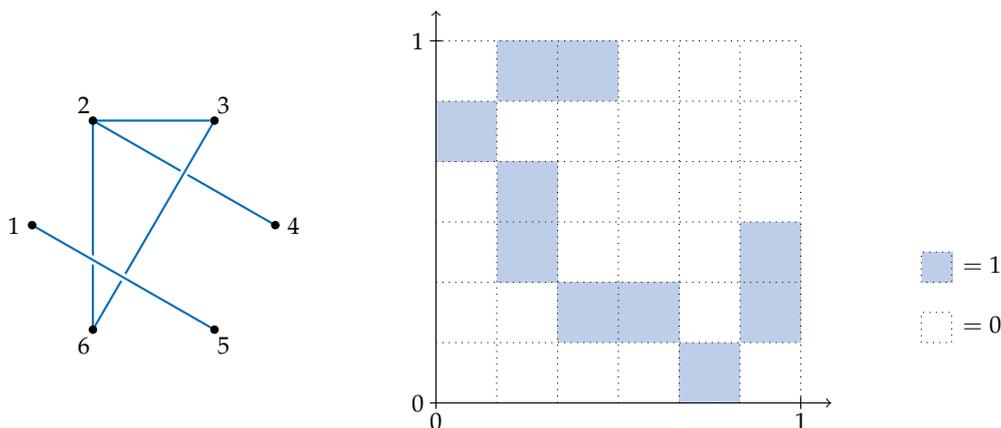

FIGURE 4. A graph and its associated graph function.

Given $g \in L^\infty([0, 1]^2, dx\,dy)$, we set

$$\|g\|_\square = \sup_{S, T \subset [0,1]} \left| \int_{S \times T} g(x, y) \, dx \, dy \right|,$$

where the supremum runs over pairs of measurable subsets (S, T) of $[0, 1]$. This is a norm on the space $L^\infty([0, 1]^2)$ which is equivalent to the norm of operator $\| \cdot \|_{L^\infty([0,1]) \to L^1([0,1])}$. The *cut-metric* on graph functions $g \in \mathcal{F}$ (see [Bor+08, Section 3.4]) is defined by

$$d_\square(g, g') = \inf_\sigma \|g^\sigma - g'\|_\square,$$

where the infimum runs over measure-preserving bijections σ of the interval $[0, 1]$. Notice that $d_\square(g, g')$ is also the infimum $\inf_{\sigma, \tau} \|g^\sigma - (g')^\tau\|_\square$ over pairs of Lebesgue isomorphisms; as a consequence, $d_\square$ satisfies the triangular inequality. We define an equivalence relation on $\mathcal{F}$ by

$$g \sim g' \iff d_\square(g, g') = 0.$$

If γ and γ' are the equivalence classes of the graph functions g and g', then the quotient space $\mathscr{G} = \mathcal{F}/\sim$ is endowed with the distance $\delta_\square(\gamma, \gamma') = d_\square(g, g')$. We call *graphon* an equivalence class of graph functions in $\mathscr{G}$, and the space of graphons $(\mathscr{G}, \delta_\square)$ is a compact metric space: see [Bor+08, Proposition 3.6].

If $\gamma \in \mathscr{G}$ and g is a graph function in this equivalence class, we can associate with g a sequence of random graphs $(G_n(\gamma))_{n \in \mathbb{N}}$ with $|G_n(\gamma)| = n$. The construction of $G_n(\gamma)$ ensures that its law is independent of the representative g chosen in the equivalence class γ, which justifies the notation $G_n(\gamma)$ instead of $G_n(g)$. To construct $G_n(\gamma)$, we first draw n

independent random variables $X_1, \ldots, X_n$ uniformly in $[0,1]$, and then, $\binom{n}{2}$ Bernoulli random variables $B_{i,j}$ that are independent conditionally to $(X_1, \ldots, X_n)$, and such that

$$\mathbb{P}[B_{i,j} = 1|(X_1, \ldots, X_n)] = g(X_i, X_j)$$

for any $1 \leq i < j \leq n$. The random graph $G_n(\gamma)$ is then the graph on n vertices $1, 2, \ldots, n$, with i connected to j if and only if $B_{i,j} = 1$. We have drawn in Figure 5 two examples of such random graphs, when γ is the representative of the function $g(x,y) = (x+y)/2$ or of the function $g(x,y) = xy$.
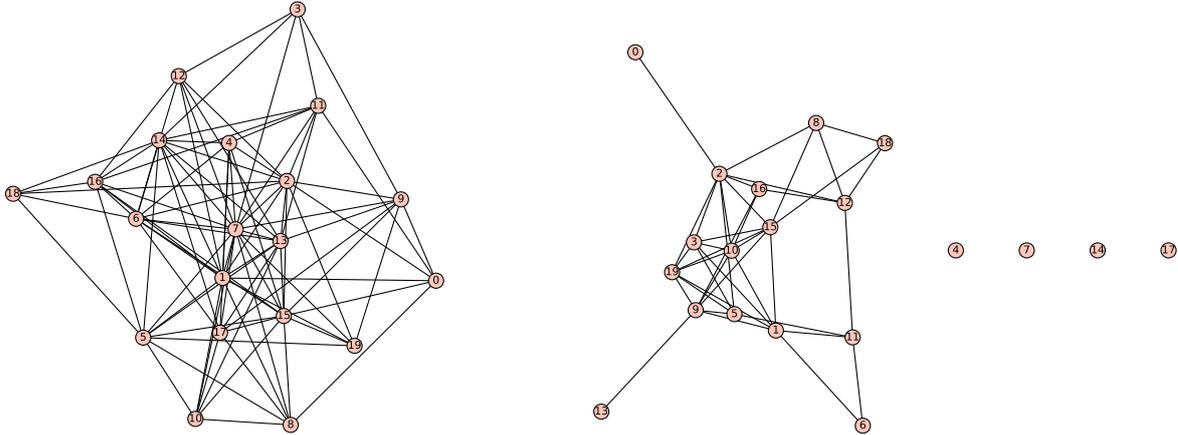

FIGURE 5. Two random graphs of size $n = 20$ associated with the graph functions $g(x,y) = \frac{x+y}{2}$ and $g'(x,y) = xy$.

We denote $\Gamma_n(\gamma)$ the equivalence class in $\mathscr{G}$ of the graph function $g(G_n(\gamma))$ that is canonically associated with the random graph $G_n(\gamma)$.

Proposition 13 (Corollary 2.6 in [LS06]). *For any graphon $\gamma \in \mathscr{G}$, the sequence of random graphons $(\Gamma_n(\gamma))_{n \in \mathbb{N}}$ converges in probability towards γ.*

Thus, the space of graphons $\mathscr{G}$ parametrises certain models of random graphs which have a property of asymptotic concentration. We shall recall in Section 4.1 that this framework allows one to deal with models that converge with respect to the notion of subgraph density, and we shall prove later that the models $(G_n(\gamma))_{n \in \mathbb{N}}$ have subgraph densities which are mod-Gaussian convergent.

3.2. **The space of permutons.** We now present an analogous construction with random permutations instead of random graphs, and which involves the notion of *permutons*; we refer to [Hop+11; Hop+13; Gle+15] for the origins of this notion. We call *permuton* a Borel probability measure π on the square $[0,1]^2$ whose marginal laws are uniform: namely, if $p_1, p_2 : [0,1]^2 \to [0,1]$ are the two coordinate projections and if $p_{1,*}, p_{2,*} : \mathscr{M}^1([0,1]^2) \to \mathscr{M}^1([0,1])$ are the two induced applications between the spaces of probability measures, then we require that

$$p_{1,*}(\pi) = p_{2,*}(\pi) = \text{Lebesgue measure on } [0,1].$$

In other words, a permuton is a coupling of two uniform laws on $[0,1]$. We denote $\mathscr{S}$ the space of permutons; these objects are also called Markov operators or doubly stochastic measures in [Pec59; Bro65; Bro66]; polymorphisms in [Ver05]; and copulas or joinings in other places of the literature. If the space of probability measures $\mathscr{M}^1([0,1]^2)$ is endowed

with the weak topology of convergence in law, then $\mathscr{M}^1([0,1]^2)$ is a compact metrisable space (see for instance [Bil69, Chapter 1]), and $\mathscr{S}$ is a closed subset of it, hence a compact metrisable space itself. Given a finite permutation $\sigma \in \mathfrak{S}(n)$, one associate with it a canonical permuton $\pi = \pi(\sigma)$, which is the probability measure on $[0,1]^2$ with density

$$f(\sigma; x, y) = n \, 1_{\sigma(\lceil nx \rceil) = \lceil ny \rceil}.$$

It is easily checked that this density yields uniform marginal laws; we refer to Figure 6 for an example.

Conversely, given a permuton $\pi \in \mathscr{S}$, we associate with π a sequence of random permutations $(\sigma_n(\pi))_{n \in \mathbb{N}}$ with $\sigma_n(\pi) \in \mathfrak{S}(n)$ for any n. If $(x_1, y_1), \ldots, (x_n, y_n)$ is a family of points in the square $[0,1]^2$, we say that these points are in a general configuration if all the x_i's are distinct, and if all the y_i's are also distinct. With a general family of n points, we associate a unique permutation $\sigma \in \mathfrak{S}(n)$ with the following property: if $\psi_1 : \{x_1, \ldots, x_n\} \to [\![1, n]\!]$ and $\psi_2 : \{y_1, \ldots, y_n\} \to [\![1, n]\!]$ are increasing bijections, then, for $i \le n$

$$\sigma(\psi_1(x_i)) = \psi_2(y_i).$$

for any $i \in [\![1, n]\!]$. We then say that σ is the *configuration* of the set of points; and we denote $\sigma = \mathrm{conf}((x_1, y_1), \ldots, (x_n, y_n))$. Intuitively, this means that the family of points $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ is the diagram of the permutation σ, up to an increasing reparametrisation of the two axes. Now, for $\pi \in \mathscr{S}$, a family of independent points $(X_1, Y_1), \ldots, (X_n, Y_n)$ under $\pi^{\otimes n}$ is in general configuration with probability 1, since the marginal laws of the coordinates are uniform. We can therefore define a random permutation

$$\sigma_n(\pi) = \mathrm{conf}((X_1, Y_1), \ldots, (X_n, Y_n)).$$

We refer to Figure 7 for an example, with a random permutation associated with the permuton π that is supported on the disc inscribed in the square $[0,1]^2$, and that has a density proportional to $\frac{1}{\sqrt{1-4r^2}}$, where r is the distance to the center $(\frac{1}{2}, \frac{1}{2})$ of this disc (this ensures that the marginal laws are uniform).
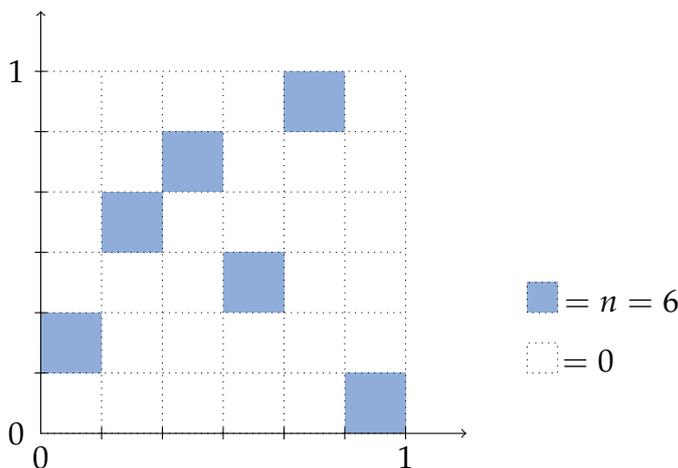

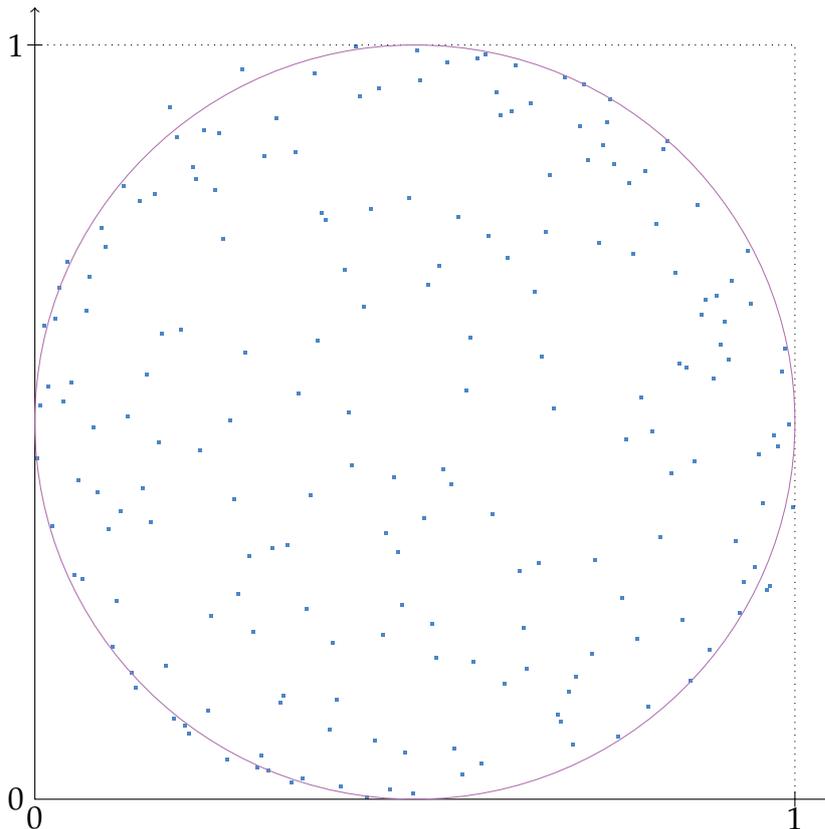FIGURE 6. The density of the permuton $\pi(\sigma)$ associated with the permutation $\sigma = 245361$.

FIGURE 7. Diagram of a random permutation with size $n = 200$ associated with the permuton π which is the projection on the xy-plane of the uniform law on a three dimensional sphere with radius $\frac{1}{2}$ and center $(\frac{1}{2}, \frac{1}{2}, 0)$.

We denote $\Pi_n(\pi)$ the permuton in $\mathscr{S}$ associated with the random permutation $\sigma_n(\pi)$. In Figure 7, it appears that for n large, the random permuton $\Pi_n(\pi)$ looks a lot like the empirical measure of n independent points under π. Indeed, the reparametrisation of $[0,1]^2$ associated with the order statistics ψ_1 and ψ_2 that were introduced in the definition of $\sigma_n(\pi) = \mathrm{conf}((X_1, Y_1), \ldots, (X_n, Y_n))$ can be shown to be very close to the identity map. This implies the following limiting result:

Proposition 14 (Theorem 1.6.(ii) in [Hop+13]). *For any permuton $\pi \in \mathscr{S}$, the sequence of random permutons $(\Pi_n(\pi))_{n \in \mathbb{N}}$ converges in probability towards π.*

Thus, the space of permutons $\mathscr{S}$ parametrises certain models of random permutations whose diagrams have a property of asymptotic concentration. This property of concentration can be shown to be equivalent to the convergence of the densities of patterns (see Section 4.2), and we shall prove later that the pattern densities of the models $(\sigma_n(\pi))_{n \in \mathbb{N}}$ are mod-Gaussian convergent.

3.3. **The Thoma simplex.** An analogous construction exists in the setting of random integer partitions, and the underlying theory goes back to the works of Kerov and Vershik in the 80's, see in particular [KV77; KV81]. The space of parameters corresponding to these models is related to the classification of the totally positive sequences and of the positive specialisations of the Schur functions; see [Ais+51; Tho64]. We recall from the introduction

that the *Thoma simplex* is the set of pairs of sequences

$$\mathscr{P} = \left\{ \omega = (\alpha, \beta) = ((\alpha_1 \geq \alpha_2 \geq \cdots \geq 0), (\beta_1 \geq \beta_2 \geq \cdots \geq 0)) \mid \sum_{i=1}^{\infty} (\alpha_i + \beta_i) \leq 1 \right\}.$$

In the following we shall denote $\gamma = 1 - \sum_{i=1}^{\infty}(\alpha_i + \beta_i)$. As explained in Section 1.2, the system of Frobenius coordinates allows one to see any integer partition $\lambda \in \mathfrak{P}(n)$ as an element $\omega(\lambda) \in \mathscr{P}$, with $\gamma = 0$ if $n \geq 1$. Conversely, fix $\omega \in \mathscr{P}$, and an integer $n \geq 1$. It is known from [Tho64; KV81] that the parameter ω corresponds to a function $\chi^{\omega} : \mathfrak{S}(\infty) \to \mathbb{C}$ which is

- tracial: $\chi^{\omega}(\sigma_1 \sigma_2) = \chi^{\omega}(\sigma_2 \sigma_1)$;
- normalised: $\chi^{\omega}(\mathrm{id}_{\mathbb{N}^*}) = 1$;
- non-negative definite: the matrix $(\chi^{\omega}(\sigma_i(\sigma_j)^{-1}))_{1 \leq i,j \leq N}$ is Hermitian and non-negative for any finite family of permutations in $\mathfrak{S}(\infty)$.

The formula for $\chi^{\omega}(\sigma)$ is

$$\chi^{\omega}(\sigma) = \prod_{\substack{c \text{ cycle of } \sigma \\ \text{with length } k \geq 2}} t(k, \omega) = \prod_{\substack{c \text{ cycle of } \sigma \\ \text{with length } k \geq 2}} \left(\sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1} \sum_{i=1}^{\infty} (\beta_i)^k \right).$$

It is convenient to set $t(1, \omega) = 1$ for any $\omega \in \mathscr{P}$, so that one can take into account the fixed points of a permutation σ in the previous formula. Thoma's theorem shows that $\omega \mapsto \chi^{\omega}$ is a bijection from $\mathscr{P}$ to the set of extremal points in the compact convex set of non-negative definite normalised tracial functions on $\mathfrak{S}(\infty)$. Consider now the restriction of χ^{ω} to a finite symmetric group $\mathfrak{S}(n)$: it is still tracial normalised non-negative, but it is not extremal anymore, and we have a decomposition

$$(\chi^{\omega})_{|\mathfrak{S}(n)} = \sum_{\lambda \in \mathfrak{P}(n)} \mathbb{P}_{n,\omega}[\lambda] \chi^{\lambda}(\sigma),$$

where the χ^{λ}'s are the characters of the irreducible Specht representations of $\mathfrak{S}(n)$, and $\mathbb{P}_{n,\omega}[\cdot]$ is a probability measure on $\mathfrak{P}(n)$. We denote in the sequel $\lambda_n(\omega) \in \mathfrak{P}(n)$ a random integer partition chosen according to this probability measure $\mathbb{P}_{n,\omega}$, and $\Omega_n(\omega) \in \mathscr{P}$ the corresponding random parameter of the Thoma simplex. We refer to Figure 8 for an example of random partition associated with a parameter $\omega \in \mathscr{P}$.

The law of the random integer partition $\lambda_n(\omega)$ can be computed by using the algebra of symmetric functions Sym (*cf.* [Mac95] for the theory of this Hopf algebra, which we shall consider here over the field of real numbers). More precisely, denote s_{λ} the Schur function with label λ, and if $k \geq 1$, denote p_k the k-th power sum function; we recall that Sym $= \mathbb{R}[p_1, p_2, \ldots]$, and that the family of Schur functions $(s_{\lambda})_{\lambda \in \bigsqcup_{n \in \mathbb{N}} \mathfrak{P}(n)}$ form a linear basis of Sym. We define for any $\omega \in \mathscr{P}$ a morphism of algebras Sym $\to \mathbb{R}$ by setting $p_k(\omega) = t(k, \omega)$. Since $(p_k)_{k \geq 1}$ is a transcendence basis of Sym over $\mathbb{R}$, this rule allows one to define $s_{\lambda}(\omega)$ for any λ integer partition. By using the Frobenius–Schur formula relatinf power sums to Schur functions, one can then show that

$$\mathbb{P}_{n,\omega}[\lambda] = (\dim \lambda) s_{\lambda}(\omega),$$

where $\dim \lambda$ is the number of standard tableaux with shape λ. These measures are called central measures in [KV81], and one has the following limiting result, which is quite clear in Figure 8.
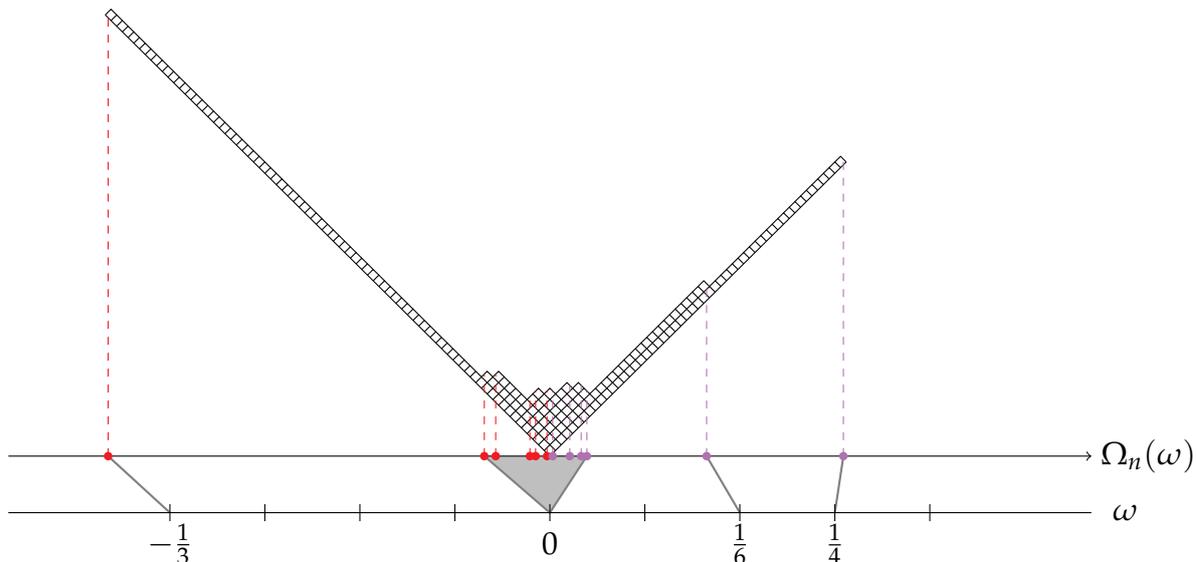
FIGURE 8. Thoma parameter $\Omega_n(\omega)$ of a random partition $\lambda_n(\omega)$ with $n = 200$ and $\omega = ((\frac{1}{4}, \frac{1}{6}, 0, \ldots), (\frac{1}{3}, 0, \ldots))$.

Proposition 15 (Section 5 in [KV81]). *For any parameter $\omega \in \mathscr{P}$ in the Thoma simplex, the sequence of random parameters $(\Omega_n(\omega))_{n \in \mathbb{N}}$ converges coordinate-wise and in probability towards ω.*

This result implies that for any $\omega \in \mathscr{P}$ and any $i \geq 1$,

$$\frac{\lambda_{n,i}(\omega)}{n} \to_{\mathbb{P}} \alpha_i \qquad ; \qquad \frac{\lambda'_{n,i}(\omega)}{n} \to_{\mathbb{P}} \beta_i.$$

The convergence coordinate by coordinate in $\mathscr{P}$ can be reinterpreted as the convergence of moments of certain probability measures on $[-1, 1]$ associated with the parameters of the Thoma simplex, and we shall prove in Section 5 that these moments are generically mod-Gaussian convergent after an appropriate renormalisation.

Remark. The reader might wonder how one can draw at random an integer partition $\lambda_n(\omega)$ as in Figure 8. The easiest way to do it is to use the Robinson–Schensted–Knuth algorithm in order the relate the central measures on integer partitions to certain models of random permutations obtained by shuffles; see [Ful02] and [Mél17, Section 12.2].

4. THE ALGEBRAS OF OBSERVABLES

In the previous section, we introduced three compact spaces of parameters $\mathscr{G}$, $\mathscr{S}$ and $\mathscr{P}$ which parametrise models of random graphs, of random permutations and of random integer partitions; and all these models have an asymptotic concentration property (Propositions 13, 14 and 15). The purpose of this section is to reinterpret the convergence in terms of observables of the combinatorial objects. The interest of these observables is that they will allow us to speak of fluctuations of models in Section 5, and to prove an underlying mod-Gaussian convergence. We shall thus complete the following table:

space	parameters	random combinatorial object	random parameter	observables	algebra of observables
$\mathscr{G}$	γ	graph $G_n(\gamma)$	$\Gamma_n(\gamma)$	subgraph densities	algebra of graphs $\mathscr{O}_G$
$\mathscr{S}$	π	permutation $\sigma_n(\pi)$	$\Pi_n(\pi)$	pattern densities	algebra of permutations $\mathscr{O}_{\mathfrak{S}}$
$\mathscr{P}$	$\omega = (\alpha, \beta)$	partition $\lambda_n(\omega)$	$\Omega_n(\omega)$	Frobenius moments	algebra of partitions $\mathscr{O}_{\mathfrak{P}}$

For each space of parameters $\mathscr{M}$, we shall exhibit a combinatorial algebra $\mathscr{O}_{\mathfrak{M}}$, endowed with a morphism of algebras Ψ from $\mathscr{O}_{\mathfrak{M}}$ to $\mathscr{C}(\mathscr{M})$, the algebra of continuous functions on $\mathscr{M}$. These morphisms have the property that the convergence $m_n \to m$ in the compact metrisable space $\mathscr{M}$ is equivalent to the convergence of all the observables $\Psi(f)(m_n) \to \Psi(f)(m)$ for any $f \in \mathscr{O}_{\mathfrak{M}}$. Equivalently (through the Stone–Weierstrass theorem), the image of Ψ is a dense subalgebra of $\mathscr{C}(\mathscr{M})$. The combinatorics of the observables are easier to understand inside an abstract algebra $\mathscr{O}_{\mathfrak{M}}$, instead of directly inside the algebra of continuous functions $\mathscr{C}(\mathscr{M})$; this is one of the reasons why we use this point of view.

4.1. **Subgraph counts.** Let $G(n)$ denote the set of isomorphism classes of simple graphs with n vertices. For instance, $G(4)$ is the set that consists in the 11 following graphs:
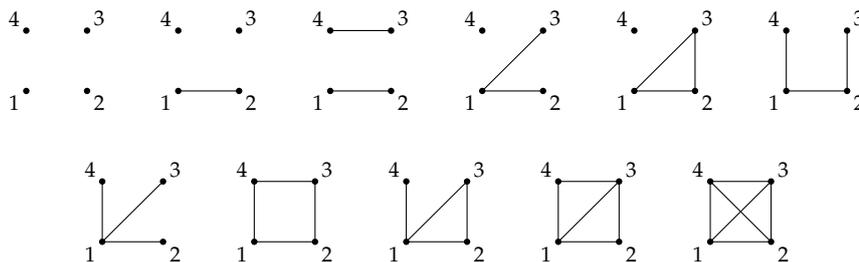
The *algebra of graphs* $\mathscr{O}_G$ is the (commutative) algebra over the set of real numbers whose combinatorial basis consists in the elements of $G = \bigsqcup_{n \in \mathbb{N}} G(n)$, and whose product is defined by

$$F_1 \times F_2 = F_1 \sqcup F_2,$$

where the right-hand side of the formula stands for the isomorphism class of the disjoint union of two graphs in the classes F_1 and F_2. The algebra $\mathscr{O}_G$ is graded by $\deg F = |V_F| = n$ if $F \in G(n)$.

One can evaluate an element of $\mathscr{O}_G$ on a graph G or on a graphon γ by using the notion of subgraph count and of subgraph density. Let $F = (V_F, E_F)$ and $G = (V_G, E_G)$ be two finite graphs. A *morphism* from F to G is a map $\phi : V_F \to V_G$ such that, if $\{v_1, v_2\} \in E_F$, then $\{\phi(v_1), \phi(v_2)\} \in E_G$. We denote $\hom(F, G)$ the set of morphisms from F to G. The *subgraph count* of F in G is the cardinality $|\hom(F, G)|$, and the *subgraph density* of F in G is defined by

$$t(F, G) = \frac{|\hom(F, G)|}{|V_G|^{|V_F|}}.$$

This density is a real number between 0 and 1.

Given a finite graph F and a graph function g, we can also define a density of F in g by the following formula:

$$t(F,g) = \int_{[0,1]^k} \left(\prod_{e=\{i,j\}\in E_F} g(x_i, x_j) \right) dx_1 \, dx_2 \cdots dx_k,$$

where V_F is identified with $[\![1,k]\!]$ if $k = |V_F|$. For instance, if F is the graph of Figure 4, then

$$t(F,g) = \int_{[0,1]^6} g(x_1, x_5) g(x_2, x_3) g(x_2, x_4) g(x_2, x_6) g(x_3, x_6) \, dx_1 \, dx_2 \cdots dx_6.$$

Let us describe some easy properties of this functions. First, it is easily seen that $t(F,g)$ only depends on the equivalence class γ of g in $\mathscr{G}$, so $t(F,\gamma) = t(F,g)$ is well defined for any graphon $\gamma \in \mathscr{G}$. Moreover, if γ is the graphon associated with a finite graph G, then

$$t(F,\gamma) = t(F,G)$$

for any finite graph F, see [Bor+08, Equation (3.2)]. Lastly, if F and F' are two finite graphs, then for any graphon γ, we have $t(F \sqcup F', \gamma) = t(F, \gamma) \, t(F', \gamma)$. In other words, the map $\Psi : \mathscr{O}_G \to \mathbb{R}^{\mathscr{G}}$ defined by $\Psi(F) = t(F, \cdot)$ is a morphism of algebras. The next statement connects this morphism to the topology of $\mathscr{G}$ defined by the cut-metric.

Proposition 16 (Theorem 5.1 in [LS07] and Theorems 2.6 and 3.8 in [Bor+08]). *A sequence of graphons $(\gamma_n)_{n\in\mathbb{N}}$ converges in $(\mathscr{G}, \delta_\square)$ to a graphon γ if and only if, for any finite graph F, $t(F, \gamma_n) \to t(F, \gamma)$.*

Equivalently, the range of the morphism of algebras $\Psi : \mathscr{O}_G \to \mathbb{R}^{\mathscr{G}}$ defined by $\Psi(F) = t(F, \cdot)$ is included in $\mathscr{C}(\mathscr{G})$, and it is a dense subalgebra of this algebra of continuous functions. This result can be used to prove Proposition 13: indeed, one shows easily that for any finite graph F and any graphon $\gamma \in \mathscr{G}$,

$$|\mathbb{E}[t(F, \Gamma_n(\gamma))] - t(F, \gamma)| \leq \frac{|V_F|^2}{2n} \qquad ; \qquad \mathrm{var}(t(F, \Gamma_n(\gamma))) \leq \frac{3\,|V_F|^2}{n}$$

see [LS06, Lemma 2.4]; whence the asymptotic concentration by using the Bienaymé–Chebyshev inequality. In Section 5.1, we shall prove that $t(F, \Gamma_n(\gamma))$ is actually mod-Gaussian after an appropriate scaling.

Remark (Kernel of the morphism $\mathscr{O}_G \to \mathscr{C}(\mathscr{G})$). The morphism of algebras $\Psi : \mathscr{O}_G \to \mathscr{C}(\mathscr{G})$ is not injective, since the one-point graph $\bullet$ has density $t(\bullet, \gamma) = 1$ for any graphon γ. In [Whi32; ELS79], it is shown that the graph densities $t(F, \cdot)$ are algebraically independent over $\mathbb{R}$ when F runs over the set of isomorphism classes of *connected* finite graphs. Therefore, the kernel of Ψ is actually the ideal of $\mathscr{O}_G$ generated by the difference of graphs $\bullet - \varnothing$. We refer to [Bor+06] for a general survey of the properties of enumeration of graph homomorphisms.

Remark. One could also work with *embeddings* of F into G, that is morphisms that are injective maps $V_F \to V_G$. Set

$$t_0(F,G) = \frac{|\mathrm{emb}(F,G)|}{|V_G|^{\downarrow|V_F|}},$$

where $\mathrm{emb}(F,G)$ is the set of embeddings of F into G, and $n^{\downarrow k}$ denotes the falling factorial $n(n-1)(n-2)\cdots(n-k+1)$ — thus, $|V_G|^{\downarrow|V_F|}$ is the number of injective maps from V_F to

V_G. The two quantities $t(F, G)$ and $t_0(F, G)$ are close when G is sufficiently large:

$$|t(F, G) - t_0(F, G)| \leq \frac{1}{|V_G|} \binom{|V_F|}{2},$$

see [LS06, Lemma 2.1]. However, t_0 cannot be extended an algebra morphism from $\mathscr{O}_G$ to $\mathscr{C}(\mathscr{G})$. Therefore we prefer to work with the first definition of the density $t(F, G)$.

4.2. **Permutation patterns.** As before, $\mathfrak{S}(n)$ is the symmetric group of order n, and we shall denote $\mathbb{R}\mathfrak{S}(n)$ the real vector space that is spanned by the permutations of size n. We introduce a product $\mathbb{R}\mathfrak{S}(m) \times \mathbb{R}\mathfrak{S}(n) \to \mathbb{R}\mathfrak{S}(m + n)$ which we call the *graphical shuffle product*; this operation was already considered, also in connection with pattern occurrences, in [Var14] and [Bas+18, Section 4].

If $\sigma \in \mathfrak{S}(m)$ and $\tau \in \mathfrak{S}(n)$, consider two parts A and B of $[\![1, m + n]\!]$ with cardinality $|A| = |B| = m$. The (A, B)-shuffle product of σ and τ, denoted $\sigma \overset{B}{\underset{A}{\times}} \tau$, is the unique permutation $\rho \in \mathfrak{S}(m + n)$ such that

- ρ maps the subset A onto B (and hence its complement $\overline{A}$ onto $\overline{B}$);

- the patterns induced by ρ on the set A and $\overline{A}$ are σ and τ respectively.

Explicitly, if we denote $\psi_{m,A}$ and $\psi_{m,B}$ (respectively, $\psi_{n,\overline{A}}$ and $\psi_{n,\overline{B}}$) the two increasing bijections from $[\![1, m]\!]$ to A and to B (respectively, from $[\![1, n]\!]$ to the complement subsets $\overline{A}$ and $\overline{B}$), then ρ is given by

$$\rho(k) = \begin{cases} \psi_{m,B} \circ \sigma \circ \psi_{m,A}^{-1}(k) & \text{if } k \in A, \\ \psi_{n,\overline{B}} \circ \tau \circ \psi_{n,\overline{A}}^{-1}(k) & \text{if } k \in \overline{A}. \end{cases}$$

An example is given on Figure 9. Informally we take the diagrams of the permutations σ of τ and merge them by shuffling independently the x-coordinates and the y-coordinates of the dots, which explain the name of graphical shuffle product.
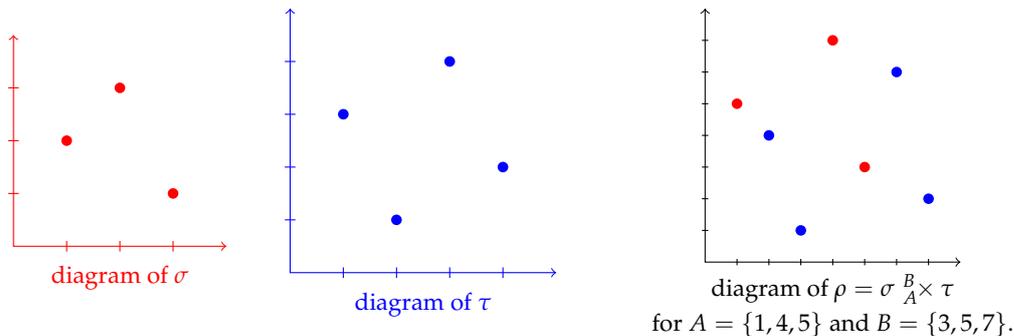

diagram of σ

diagram of τ

diagram of $\rho = \sigma \overset{B}{\underset{A}{\times}} \tau$
for $A = \{1, 4, 5\}$ and $B = \{3, 5, 7\}$.

FIGURE 9. Example of (A, B)-shuffle of permutations.

Now, the graphical shuffle product $\sigma \times \tau$ in $\mathbb{R}\mathfrak{S}(m + n)$ is defined as a linear combination of all the possible (A, B)-shuffle products, A and B being arbitrary subsets with cardinality m in $[\![1, m + n]\!]$:

$$\sigma \times \tau = \frac{m! \, n!}{(m + n)!} \sum_{|A| = |B| = m} \sigma \overset{B}{\underset{A}{\times}} \tau.$$

Notice that the number of terms in the sum of the right-hand side is $(\frac{(m+n)!}{m! \, n!})^2$, and that some permutations $\rho \in \mathfrak{S}(m + n)$ may appear with multiplicity larger than 1 in this sum.

For instance, if $\sigma = [12]$ and $\tau = [21]$, then their graphical shuffle product is the linear combination

$$[12] \times [21] = \frac{1}{6}([1243] + [1324] + [2134] + [2413] + [3142] + [3421] + [4231] + [4312])$$

$$+ \frac{1}{3}([1342] + [1423] + [2314] + [2431] + [3124] + [3241] + [4132] + [4213])$$

$$+ \frac{1}{2}([1432] + [2341] + [3214] + [4123]).$$

The *algebra of permutations* is the real algebra $\mathscr{O}_{\mathfrak{S}}$ which as a vector space is equal to the direct sum $\bigoplus_{n=0}^{\infty} \mathbb{R}\mathfrak{S}(n)$, and which is endowed with the graphical shuffle product. Since the graphical shuffle product $\sigma \times \tau$ encodes all the ways of mixing graphically the two patterns σ and τ, this operation is clearly commutative, and one sees readily that it is also associative. Thus, $\mathscr{O}_{\mathfrak{S}}$ is a commutative algebra whose combinatorial basis consists in all finite permutations $\sigma \in \bigsqcup_{n \in \mathbb{N}} \mathfrak{S}(n)$. It is graded by $\deg \sigma = |\sigma| = n$ if $\sigma \in \mathfrak{S}(n)$.

Recall from the introduction that if $\tau \in \mathfrak{S}(k)$ and $\sigma \in \mathfrak{S}(n)$, then τ is a pattern of σ if there exists a subset $\{a_1 < a_2 < \cdots < a_k\}$ of $[\![1, n]\!]$ such that $\sigma(a_i) < \sigma(a_j)$ if and only if $\tau(i) < \tau(j)$. As for graphs, we can then define the *pattern density* of τ in σ by the ratio

$$t(\tau, \sigma) = \frac{\mathrm{occ}(\tau, \sigma)}{\binom{n}{k}},$$

where the numerator of this fraction is the number of *occurrences* of τ in σ, that is the number of subsets $\{a_1 < a_2 < \cdots < a_k\} \subset [\![1, n]\!]$ that make appear τ as a pattern of σ. On the other hand, if π is a permuton, we can also define the density of $\tau \in \mathfrak{S}(k)$ in π by the following formula:

$$t(\tau, \pi) = \int_{([0,1]^2)^k} 1_{\mathrm{conf}((x_1, y_1), \ldots, (x_k, y_k)) = \tau}\, \pi(dx_1\, dy_1) \cdots \pi(dx_k\, dy_k)$$

$$= \mathbb{P}_{\pi^{\otimes k}}[\mathrm{conf}((X_1, Y_1), \ldots, (X_k, Y_k)) = \tau].$$

If σ is a permutation and if $\pi(\sigma)$ is the associated permuton, then we may have $t(\tau, \sigma) \neq t(\tau, \pi(\sigma))$. The difference is however small for σ sufficiently large. More precisely, if $\tau \in \mathfrak{S}(k)$ and $\sigma \in \mathfrak{S}(n)$, then

$$|t(\tau, \sigma) - t(\tau, \pi(\sigma))| \leq \frac{1}{n}\binom{k}{2} = \frac{1}{|\sigma|}\binom{|\tau|}{2};$$

see [Hop+13, Lemma 3.5]. Therefore, in order to evaluate a pattern density in a permutation σ, we can use either $t(\tau, \sigma)$ or $t(\tau, \pi(\sigma))$. This is an important difference between the permutons and the two other classes of models, and it will lead us to state two distinct sets of limiting results for the models of random permutations. The second choice of observables $t(\tau, \pi(\sigma))$ is the better one when looking at random permutations and fluctuations thereof, for the following reason:

Proposition 17 (Graphical shuffle products of permutations). *Consider the linear map* $\Psi : \mathscr{O}_{\mathfrak{S}} \to \mathbb{R}^{\mathscr{S}}$ *which is defined by* $\Psi(\tau) = t(\tau, \cdot)$, *the right-hand of this formula being a function on permutons. This map is a morphism of algebras, and its range is a dense subalgebra of the algebra of continuous functions* $\mathscr{C}(\mathscr{S})$.

Proof. The second part of the proposition is contained in [Hop+13, Theorem 1.8], which proves that a sequence of permutons $(\pi_n)_{n \in \mathbb{N}}$ converges with respect to the weak topology on probability measures if and only if the sequences of observables $(t(\tau, \pi_n))_{n \in \mathbb{N}}$ converge

for any finite permutation τ. Therefore, each observable $t(\tau, \cdot)$ is continuous on $\mathscr{S}$, and these observables separate the permutons. Hence, by Stone–Weierstrass, $\Psi(\mathscr{O}_{\mathfrak{S}})$ is a dense subalgebra of $\mathscr{C}(\mathscr{S})$.

Let us now prove that for any permutations τ_1 and τ_2 with sizes k_1 and k_2, and for any permuton π,

$$t(\tau_1, \pi)\, t(\tau_2, \pi) = \frac{k_1!\, k_2!}{(k_1 + k_2)!} \sum_{|A|=|B|=k_1} t(\tau_1 \,{}^{B}_{A}\!\times\, \tau_2, \pi), \tag{1}$$

where the sum runs over pairs of subsets of $[\![1, k_1 + k_2]\!]$ with cardinality k_1. This amounts to the property of morphism of algebras that remains to be proved. In [Bas+18, Proposition 4.5], an analogue result is given for the observables of permutations instead of permutons: for any $n \geq 0$ and any $\sigma \in \mathfrak{S}(n)$,

$$t(\tau_1, \sigma)\, t(\tau_2, \sigma) = \frac{k_1!\, k_2!}{(k_1 + k_2)!} \sum_{|A|=|B|=k_1} t(\tau_1 \,{}^{B}_{A}\!\times\, \tau_2, \sigma) + O\!\left(\frac{1}{n}\right), \tag{2}$$

where the constant in the $O(\cdot)$ depends on τ_1 and τ_2 but does not depend on n and σ. Since this formula is true for any σ, it is also true for the random permutations $\sigma_n(\pi)$ attached to a permuton $\pi \in \mathscr{S}$. Since $t(\tau, \sigma_n(\pi)) = t(\tau, \Pi_n(\pi)) + O(\frac{1}{n})$ and since $\lim_{n\to\infty} \Pi_n(\pi) = \pi$ by Proposition 14, Equation (1) follows by applying Equation (2) to $\sigma_n(\pi)$, and then by taking the limit of both sides when n goes to infinity. □

In Section 5.2, we shall prove that the sequences of random variables $(t(\tau, \sigma_n(\pi)))_{n\in\mathbb{N}}$ and $(t(\tau, \Pi_n(\pi)))_{n\in\mathbb{N}}$ are mod-Gaussian convergent after an appropriate scaling.

Remark (Kernel of the morphism $\mathscr{O}_{\mathfrak{S}} \to \mathscr{C}(\mathscr{S})$). As for graphons, the morphism of algebras $\Psi : \mathscr{O}_{\mathfrak{S}} \to \mathscr{C}(\mathscr{S})$ is not injective, since the image of the permutation $\tau = [1]$ of degree 1 is the constant function 1 on $\mathscr{S}$. This implies identities such as

$$t([21], \cdot) = t([21], \cdot) \times t([1], \cdot)$$
$$= t([321], \cdot) + \frac{2}{3}\left(t([231], \cdot) + t([312], \cdot)\right) + \frac{1}{3}\left(t([132], \cdot) + t([213], \cdot)\right).$$

Thus, the kernel of the morphism Ψ contains the ideal of $\mathscr{O}_{\mathfrak{S}}$ generated by $[1] - [\ ]$, where $[\ ]$ denotes the empty permutation of size 0. We do not know whether this ideal is the full kernel of Ψ.

4.3. **Observables of partitions.** We finally construct an algebra $\mathscr{O}_{\mathfrak{P}}$ of observables of partitions and of parameters of the Thoma simplex, which will play the same role as the role of $\mathscr{O}_G$ and $\mathscr{O}_{\mathfrak{S}}$ for graphons and permutons. This algebra is closely related to the Kerov–Olshanski algebra of polynomial functions on partitions which was introduced in [KO94], and whose combinatorics are detailed in [IO02] and in [Mél17, Part III]. If λ and μ are two integer partitions of size m and n, we denote $\lambda \times \mu$ their disjoint union, that is the integer partition with size $m + n$ and whose parts are those of λ and those of μ, reordered in order to get a non-increasing sequence. For instance,

$$(4, 4, 1) \times (5, 3, 1) = (5, 4, 4, 3, 1, 1).$$

The *algebra of partitions* $\mathscr{O}_{\mathfrak{P}}$ is the real algebra spanned by the partitions in $\mathfrak{P} = \bigsqcup_{n\in\mathbb{N}} \mathfrak{P}(n)$, endowed with the product defined above. It is commutative and graded by $\deg \lambda = |\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_r$.

Given an element $\omega \in \mathscr{P}$ and an integer $k \geq 1$, we recall the definition

$$t(k, \omega) = p_k(\omega) = \begin{cases} 1 & \text{if } k = 1 \\ \sum_{i=1}^{\infty} (\alpha_i)^k + (-1)^{k-1} \sum_{i=1}^{\infty} (\beta_i)^k & \text{if } k \geq 2. \end{cases}$$

We extend this definition to integer partitions ρ by setting

$$t(\rho, \omega) = t(\rho_1, \omega) \, t(\rho_2, \omega) \cdots t(\rho_r, \omega)$$

if $\rho = (\rho_1, \rho_2, \dots, \rho_r)$. We thus obtain a map $\Psi : \mathscr{O}_{\mathfrak{P}} \to \mathbb{R}^{\mathscr{P}}$ which is clearly a morphism of algebras. The following result is a bit less trivial. We endow $\mathscr{P}$ with the topology of convergence of all coordinates: a sequence $(\omega_n)_{n \in \mathbb{N}} = ((\alpha_{n,1}, \alpha_{n,2}, \dots), (\beta_{n,1}, \beta_{n,2}, \dots))_{n \in \mathbb{N}}$ converges towards a parameter $\omega = (\alpha, \beta)$ if and only if, for any $i \geq 1$, $\alpha_{n,i} \to \alpha_i$ and $\beta_{n,i} \to \beta_i$. This makes $\mathscr{P}$ into a metrisable compact space.

Proposition 18 (Convergence in the Thoma simplex). *A sequence of Thoma parameters $(\omega_n)_{n \in \mathbb{N}}$ converges towards a parameter $\omega \in \mathscr{P}$ if and only if $t(\rho, \omega_n) \to t(\rho, \omega)$ for any integer partition $\rho \in \mathfrak{P}$. Equivalenty, the range of the morphism of algebras $\Psi : \mathscr{O}_{\mathfrak{P}} \to \mathbb{R}^{\mathscr{P}}$ is a dense subalgebra of $\mathscr{C}(\mathscr{P})$.*

Proof. To any parameter $\omega \in \mathscr{P}$, we associate a probability measure θ_ω on $[-1, 1]$:

$$\theta_\omega = \sum_{i=1}^{\infty} \alpha_i \, \delta_{\alpha_i} + \sum_{i=1}^{\infty} \beta_i \, \delta_{-\beta_i} + \gamma \, \delta_0.$$

Notice then that $t(k, \omega)$ is the $(k-1)$-th moment of this measure:

$$t(k, \omega) = \int_{-1}^{1} s^{k-1} \, \theta_\omega(ds).$$

Therefore, the convergence of all the observables $t(\rho, \omega_n)$ is equivalent to the convergence of all the moments of the measures θ_{ω_n}. Since these measures are compactly supported, we have:

$$\left(\forall \rho \in \mathfrak{P}, \ t(\rho, \omega_n) \to t(\rho, \omega) \right) \quad \Longleftrightarrow \quad \left(\theta_{\omega_n} \rightharpoonup \theta_\omega \right),$$

where the convergence on the right-hand side is with respect to the weak topology on probability measures on $[-1, 1]$. It is then shown in [Mél17, proof of Theorem 12.19] that the restriction of the weak topology of $\mathscr{M}^1([-1, 1])$ to the set of measures $\{\theta_\omega, \ \omega \in \mathscr{P}\}$ is the topology of convergence of all the coordinates of the sequences. $\square$

In Section 5.3, we establish the mod-Gaussian convergence of the sequences of observables $(t(\rho, \Omega_n(\omega)))_{n \in \mathbb{N}}$, for any $\rho \in \mathfrak{P}$.

Remark (Kernel of the morphism $\mathscr{O}_{\mathfrak{P}} \to \mathscr{C}(\mathscr{P})$). As before, the morphism of algebras $\Psi : \mathscr{O}_{\mathfrak{P}} \to \mathscr{C}(\mathscr{P})$ is not injective, since $t(1, \omega) = 1$ for any parameter $\omega \in \mathscr{P}$. The kernel of Ψ is actually the ideal spanned by $1 - \varnothing$, which is equivalent to saying that the observables $t(k \geq 2, \cdot)$ are algebraically independent in $\mathbb{R}^{\mathscr{P}}$. To prove this result, we set for $\mu \in \mathfrak{P}(n)$ and $k \geq 1$:

$$p_k(\mu) = \sum_{i=1}^{d} (a_i(\mu))^k + (-1)^{k-1} \sum_{i=1}^{d} (b_i(\mu))^k = n^k \, t(k, \omega(\mu)).$$

We extend this definition to partitions ρ by setting $p_\rho(\mu) = \prod_{i=1}^{r} p_{\rho_i}(\mu) = n^{|\rho|} \, t(\rho, \omega(\mu))$. Suppose then that one has a relation

$$\sum_{\rho} c_\rho \, t(\rho, \cdot) = 0$$

where the integer partitions ρ appearing in the sum have all their parts larger than 2 (we also allow the empty partition). Let d be the maximal size of a partition ρ appearing in the sum, and λ be an arbitrary integer partition of size n. By evaluating the vanishing linear combination on λ and by multiplying by n^d, we obtain

$$\sum_\rho c_\rho \, p_{\rho \times 1^{d-|\rho|}}(\lambda) = 0$$

since $p_1(\lambda) = n$. However, by [IO02, Proposition 1.5], the functions $p_{k \geq 1}$ on $\mathfrak{P}$ are algebraically independent. Therefore, all the coefficients c_ρ vanish, whence the result.

5. DEPENDENCY STRUCTURES AND MOD-GAUSSIAN CONVERGENCE

We can finally study the fluctuations of the random models $(\Gamma_n(\gamma))_{n \in \mathbb{N}}$, $(\Pi_n(\pi))_{n \in \mathbb{N}}$ and $(\Omega_n(\omega))_{n \in \mathbb{N}}$. The easiest case to deal with is by far the one of graphs, because of the identity of observables $t(F, G) = t(F, \gamma(G))$, and because of the clear dependency structure in the random variable $n^{|F|} t(F, G_n(\gamma))$. The two other models will require a bit more work, but the proofs all follow the same lines. On the other hand, we shall see that the computation of the limiting variances σ^2 and of the limiting third cumulants L amounts to making certain operations in the algebras of observables $\mathscr{O}_G$, $\mathscr{O}_{\mathfrak{S}}$ and $\mathscr{O}_{\mathfrak{P}}$; this is the reason why we discussed these algebraic structures.

5.1. **Fluctuations of random graphs and graphons.** In this section, we fix a graphon $\gamma \in \mathscr{G}$, a graph function g in the equivalence class γ, and a graph F with k vertices. We set $S_n(F, \gamma) = n^{|F|} t(F, G_n(\gamma))$, and we are going to prove that $(S_n(F, \gamma))_{n \in \mathbb{N}}$ satisfies the hypotheses (MC1), (MC2) and (MC3). Given two graphs F and G with vertex-set $[\![1, k]\!]$ and $[\![1, n]\!]$, we can decompose the function $n^k t(F, G) = |\hom(F, G)|$ as a sum of n^k functions:

$$|\hom(F, G)| = \sum_{\psi : [\![1,k]\!] \to [\![1,n]\!]} A_{F,\psi}(G), \quad \text{with } A_{F,\psi}(G) = \begin{cases} 1 & \text{if } \psi \text{ is a morphism from } F \text{ to } G, \\ 0 & \text{otherwise.} \end{cases}$$

(3)

The functions $A_{F,\psi}$ in turns factorize as

$$A_{F,\psi} = \prod_{\{i,j\} \in E_F} B_{\psi(i),\psi(j)},$$

(4)

where $B_{i,j}(G)$ is 1 if there is an edge between i and j in G, and 0 otherwise.

We are now interested to specialize this to the random graph $G = G_n(\gamma)$. By construction, the law of the random variables $B_{i,j} = B_{i,j}(G_n(\gamma))$ is then characterized by

$$\mathbb{P}[B_{i,j} = 1|(X_1, \dots, X_n)] = g(X_i, X_j),$$

where the X_i are independent uniform random variables in $[0, 1]$. A way to construct these random variables $B_{i,j}$ so that they are independent conditionally to $(X_1, \dots, X_n)$ is as follows: we introduce other independent uniform random variables $U_{i,j} \in [0, 1]$ for any pair $\{i, j\}$, and then set

$$B_{i,j} = 1_{(U_{i,j} \leq g(X_i, X_j))}.$$

In this setting, $B_{i,j}$ is measurable with respect to the σ-field generated by $(U_{i,j}, X_i, X_j)$, so if $\{i,j\} \cap \{i',j'\} = \varnothing$, then $B_{i,j}$ and $B_{i',j'}$ are independent. Summing up and using the short notation $A_{F,\psi} = A_{F,\psi}(G_n(\gamma))$, we have

$$S_n(F, \gamma) = \sum_{\psi : [\![1,k]\!] \to [\![1,n]\!]} A_{F,\psi},$$

with the following dependency structure.

Lemma 19. *The graph G_n*

- *with vertex set $V_n = \{\psi : [\![1,k]\!] \to [\![1,n]\!]\}$,*
- *and with an edge between ψ and ϕ if $\psi(i) = \phi(j)$ for some indices $i, j \in [\![1,k]\!]$*

is a dependency graph for the family of random variables $(A_{F,\psi})_{\psi \in V_n}$ involved in the expansion of $S_n(F, \gamma)$.

Proof. Consider two families $\{\psi_c\}_{c \in C}$ and $\{\psi_d\}_{d \in D}$ without any edge between C and D. This means that the sets $\mathbb{C} = \bigcup_{c \in C} \psi_c([\![1,k]\!])$ and $\mathbb{D} = \bigcup_{d \in D} \psi_d([\![1,k]\!])$ do not intersect. However, the random vector $(A_{F,\psi_c})_{c \in C}$ (respectively, $(A_{F,\psi_d})_{d \in D}$) is measurable with respect to the σ-algebra generated by the variables

$$\{X_i, U_{i,j}\}_{i,j \in \mathbb{C}} \quad (\text{respectively, } \{X_i, U_{i,j}\}_{i,j \in \mathbb{D}}).$$

Therefore, these two random vectors are independent. $\square$

The parameters of the dependency graph G_n can be chosen as follows. Obviously, $|A_{F,\psi}| \le 1$ almost surely, so one can take $A = 1$. The total number of vertices is $N_n = N_{n,k} = n^k$, and the number of maps ϕ connected or equal to another map ψ is bounded by $k^2 n^{k-1}$, so we can take $D_n = D_{n,k} = k^2 n^{k-1}$. Hence, the previous lemma and Theorem 5 imply the hypothesis (MC1) for $S_n(F, \gamma)$ with parameters $(D_{n,k}, N_{n,k}, A) = (k^2 n^{k-1}, n^k, 1)$. To complete this analysis, we need to compute

$$\sigma^2(F, \gamma) = \lim_{n \to \infty} \frac{\mathrm{var}(S_n(F, \gamma))}{k^2 n^{2k-1}} \quad ; \quad L(F, \gamma) = \lim_{n \to \infty} \frac{\kappa^{(3)}(S_n(F, \gamma))}{k^4 n^{3k-2}}.$$

It is a bit clearer to work with distinct graphs F, G, H with k vertices, and to evaluate the rescaled *joint cumulants*

$$\frac{\kappa(S_n(F, \gamma), S_n(G, \gamma))}{k^2 n^{2k-1}} \quad \text{and} \quad \frac{\kappa(S_n(F, \gamma), S_n(G, \gamma), S_n(H, \gamma))}{k^4 n^{3k-2}}.$$

We refer to [FMN16, Section 9.2] for details on the notion of joint cumulants, which generalise the cumulants introduced in our Definition 2. For our purpose, it is sufficient to know that

$$\kappa^{(r)}(X) = \kappa(\underbrace{X, X, \ldots, X}_{r \text{ occurrences}})$$

and that

$$\kappa(X, Y) = \mathrm{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y];$$
$$\kappa(X, Y, Z) = \mathbb{E}[XYZ] - \mathbb{E}[XY]\mathbb{E}[Z] - \mathbb{E}[XZ]\mathbb{E}[Y] - \mathbb{E}[YZ]\mathbb{E}[X] + 2\,\mathbb{E}[X]\mathbb{E}[Y]\mathbb{E}[Z].$$

Proposition 20 (Limiting first cumulants of subgraph counts). *Denote $\mathscr{O}_{G,k}$ the vector space spanned by the graphs with k vertices in $\mathscr{O}_G$. There exists two linear maps*

$$\kappa_2 : \mathscr{O}_{G,k} \otimes \mathscr{O}_{G,k} \to \mathscr{O}_G;$$
$$\kappa_3 : \mathscr{O}_{G,k} \otimes \mathscr{O}_{G,k} \otimes \mathscr{O}_{G,k} \to \mathscr{O}_G$$

such that, for any finite graphs F, G and H with $k \geq 1$ vertices, and for any graphon γ,

$$\frac{\kappa(S_n(F,\gamma), S_n(G,\gamma))}{k^2 \, n^{2k-1}} = \kappa_2(F,G)(\gamma) + O(n^{-1});$$

$$\frac{\kappa(S_n(F,\gamma), S_n(G,\gamma), S_n(H,\gamma))}{k^4 \, n^{3k-2}} = \kappa_3(F,G,H)(\gamma) + O(n^{-1}),$$

with constants in the $O(\cdot)$'s which depend only on k.

Because of this description, in order to compute the limiting first cumulants of the random variables $S_n(F,\gamma)$, it will suffice to evaluate two observables $\kappa_2(F,F)$ and $\kappa_3(F,F,F)$ on the graphon γ (here we make a slight abuse of notation by writing $f(\gamma)$ instead of $\Psi(f)(\gamma)$ for $f \in \mathscr{O}_G$). This is the main reason why we introduced the graded algebra of observables $\mathscr{O}_G$, and moreover, we shall provide hereafter a combinatorial description of the maps κ_2 and κ_3.

Proof. We start with the covariances and we expand $\mathrm{cov}(S_n(F,\gamma), S_n(G,\gamma))$ by bilinearity:

$$\mathrm{cov}(S_n(F,\gamma), S_n(G,\gamma)) = \sum_{\psi,\phi} \mathrm{cov}(A_{F,\psi}, A_{G,\phi}),$$

where the sum runs over pairs of maps (ψ, ϕ) from $[\![1,k]\!]$ to $[\![1,n]\!]$. In this sum, if $\psi([\![1,k]\!]) \cap \phi([\![1,k]\!]) = \varnothing$, then the corresponding covariance vanishes, because of the dependency graph identified in Lemma 19. Thus, we can restrict the sum to the cases where $|\psi([\![1,k]\!]) \cap \phi([\![1,k]\!])| \geq 1$, which implies that $|\psi([\![1,k]\!]) \cup \phi([\![1,k]\!])| \leq 2k-1$. If $|\psi([\![1,k]\!]) \cup \phi([\![1,k]\!])| \leq 2k-2$, then the maps ψ and ϕ can be described:

- by specifying all the identities $\psi(a) = \phi(b)$ or $\psi(a) = \psi(b)$ or $\phi(a) = \phi(b)$; there is a finite number of configurations of such identities, which depends only on k.

- and then by specifying less than $2k-2$ values in $[\![1,n]\!]$; there are less than n^{2k-2} possible choices.

As a consequence, since $\mathrm{cov}(A_{F,\psi}, A_{G,\phi}) \leq 1$ for any maps ψ and ϕ, the maps ψ and ϕ such that

$$|\psi([\![1,k]\!]) \cap \phi([\![1,k]\!])| \geq 1 \qquad ; \qquad |\psi([\![1,k]\!]) \cup \phi([\![1,k]\!])| \leq 2k-2$$

will have a contribution to $\mathrm{cov}(S_n(F,\gamma), S_n(G,\gamma))$ that is a $O(n^{2k-2})$, with a constant in the $O(\cdot)$ that depends only on k. This contribution becomes a $O(n^{-1})$ when divided by $k^2 \, n^{2k-1}$, so

$$\frac{\kappa(S_n(F,\gamma), S_n(G,\gamma))}{k^2 \, n^{2k-1}} = \frac{1}{k^2 \, n^{2k-1}} \sum_{\substack{|\psi([\![1,k]\!]) \cap \phi([\![1,k]\!])|=1 \\ |\psi([\![1,k]\!]) \cup \phi([\![1,k]\!])|=2k-1}} \mathrm{cov}(A_{F,\psi}, A_{G,\phi}) + O(n^{-1}).$$

The two identities $|\psi([\![1,k]\!]) \cap \phi([\![1,k]\!])| = 1$ and $|\psi([\![1,k]\!]) \cup \phi([\![1,k]\!])| = 2k-1$ imply that ψ and ϕ have images of size k, so they are injective maps. Therefore, there are exactly $k^2 \, n^{\downarrow 2k-1}$ pairs of maps (ψ, ϕ) that satisfy these two conditions: indeed, to construct such a pair,

- one chooses two indices $a, b \in [\![1,k]\!]$ such that $\psi(a) = \phi(b)$ (k^2 possibilities);

- then, there are $2k-1$ distinct values in $[\![1,n]\!]$ to choose, hence $n^{\downarrow 2k-1}$ choices.

If a and b are two indices in $[\![1, k]\!]$, denote $(F \bowtie G)(a, b)$ the graph on $2k - 1$ vertices obtained by identifying in $F \sqcup G$ the vertex a in V_F with the vertex b in V_G. This is better understood with an example, see Figure 10.
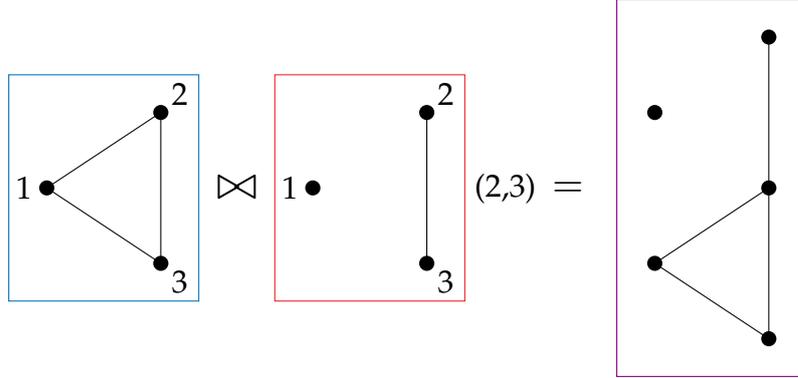

FIGURE 10. The junction $(F \bowtie G)(a, b)$ of two graphs along a pair of points $(a \in V_F, b \in V_G)$.

The graph $(F \bowtie G)(a, b)$ is involved in the computation of the covariance of $A_{F,\psi}$ and $A_{G,\phi}$ if $\psi(a) = \phi(b)$ and $|\psi([\![1, k]\!]) \cup \phi([\![1, k]\!])| = 2k - 1$. Indeed, if ψ is an injective map, then

$$
\mathbb{E}[A_{F,\psi}] = \int_{[0,1]^n} \left(\prod_{\{i,j\} \in E_F} g(x_{\psi(i)}, x_{\psi(j)}) \right) dx_1 \cdots dx_n
$$

$$
= \int_{[0,1]^k} \left(\prod_{\{i,j\} \in E_F} g(x_i, x_j) \right) dx_1 \cdots dx_k = t(F, g) = t(F, \gamma).
$$

When ψ and ψ are two injective maps whose images intersect at exactly one point $\psi(a) = \phi(b)$, a similar calculation yields

$$
\mathbb{E}[A_{F,\psi} A_{G,\phi}] = \int_{[0,1]^n} \left(\prod_{\{i,j\} \in E_F} g(x_{\psi(i)}, x_{\psi(j)}) \right) \left(\prod_{\{i,j\} \in E_G} g(x_{\phi(i)}, x_{\phi(j)}) \right) dx_1 \cdots dx_n
$$

$$
= \int_{[0,1]^n} \left(\prod_{\{i,j\} \in E_{(F \bowtie G)(a,b)}} g(x_{\theta(i)}, x_{\theta(j)}) \right) dx_1 \cdots dx_n = t((F \bowtie G)(a, b), \gamma)
$$

where θ is the unique injective map $\theta : V_{(F \bowtie G)(a,b)} \to [\![1, n]\!]$ whose restriction to V_F is ψ, and whose restriction to V_G is ϕ.

Let $F \bowtie G$ be the multiset of all the k^2 graphs $(F \bowtie G)(a, b)$, with $a, b \in [\![1, k]\!]$. We can finally compute the limiting covariance:

$$
\frac{\kappa(S_n(F, \gamma), S_n(G, \gamma))}{k^2 \, n^{2k-1}} = \frac{n^{\downarrow 2k-1}}{k^2 \, n^{2k-1}} \sum_{a,b=1}^{k} t((F \bowtie G)(a, b), \gamma) - t(F, \gamma) \, t(G, \gamma) + O(n^{-1})
$$

$$
= \psi \left(\frac{1}{k^2} \sum_{H \in F \bowtie G} (H - F \times G) \right) (\gamma) + O(n^{-1}).
$$

This proves the first part of our proposition, with the map κ_2 given by

$$\kappa_2(F,G) = \frac{1}{k^2} \sum_{H \in F \bowtie G} (H - F \times G) = \frac{1}{k^2} \sum_{1 \le a,b \le k} ((F \bowtie F)(a,b) - F \times G).$$

A similar discussion allows one to compute the asymptotics of the third cumulants. Note first that the joint cumulants have the property that if one can split a family $(X_1, \ldots, X_k)$ into two non-empty families that are independent, then $\kappa(X_1, \ldots, X_k) = 0$. Now, in order to compute $\kappa(S_n(F, \gamma), S_n(G, \gamma), S_n(H, \gamma))$, we expand this joint cumulant by multilinearity, and we are reduced to the computation of the third cumulant $\kappa(A_{F,\psi}, A_{G,\phi}, A_{H,\theta})$, where ψ, ϕ, θ are maps $[\![1,k]\!] \to [\![1,n]\!]$. By using the vanishing property of joint cumulants on independent vectors, one can get rid of all the triples (ψ, ϕ, θ) such that one of the three maps has an image disjoint from the other two. On the other hand, if the cardinality of the union $\psi([\![1,k]\!]) \cup \phi([\![1,k]\!]) \cup \theta([\![1,k]\!])$ is smaller than or equal to $3k - 3$, then the corresponding joint cumulant is a $O(1)$, and all these contributions form a $O(n^{3k-3})$. Hence,

$$\kappa(S_n(F,\gamma), S_n(G,\gamma), S_n(H,\gamma)) = \sum_{\psi,\phi,\theta} \kappa(A_{F,\psi}, A_{G,\phi}, A_{H,\theta}) + O(n^{3k-3}),$$

where the constant hidden in the $O(\cdot)$ only depends on k, and where the sum is on triples of functions $\psi, \phi, \theta : [\![1,k]\!] \to [\![1,n]\!]$ which are injective, such that

$$|\psi([\![1,k]\!]) \cup \phi([\![1,k]\!]) \cup \theta([\![1,k]\!])| = 3k - 2.$$

This implies that one of the following situations occur:

(a) there are indices $a, b, c \in [\![1,k]\!]$ such that $\psi(a) = \phi(b) = \theta(c)$, and there are no other identities between the images of ψ, ϕ, θ;

(b) or, there are indices $a, b \ne c, d \in [\![1,k]\!]$ such that $\psi(a) = \phi(b)$ and $\phi(c) = \theta(d)$, and there are no other identities between the images of ψ, ϕ, θ;

(c) or, one has the same configuration as (b), up to a cyclic permutation of the roles played by ψ, ϕ and θ.

In the first case, a, b, c being fixed, there are exactly $n^{\downarrow 3k-2}$ possibilities for ψ, ϕ, θ, and the joint cumulant of $A_{F,\psi}, A_{G,\phi}$ and $A_{H,\theta}$ is in this case

$$t((F \bowtie G \bowtie H)(a,b,c), \gamma) - t((F \bowtie G)(a,b) \sqcup H, \gamma) - t((F \bowtie H)(a,c) \sqcup G, \gamma)$$
$$- t((G \bowtie H)(b,c) \sqcup F, \gamma) + 2\,t(F \sqcup G \sqcup H, \gamma),$$

where $(F \bowtie G \bowtie H)(a,b,c)$ denotes the graph on $3k - 2$ vertices obtained by identifying in $F \sqcup G \sqcup H$ the vertex a in V_F, the vertex b in V_G and the vertex c in V_H. In the second case, $a, b \ne c, d$ being fixed, there are again $n^{\downarrow 3k-2}$ possibilities for ψ, ϕ, θ, and the joint cumulant of $A_{F,\psi}, A_{G,\phi}$ and $A_{H,\theta}$ is then

$$t((F \bowtie G \bowtie H)(a,b;c,d), \gamma) + t(F \sqcup G \sqcup H, \gamma)$$
$$- t((F \bowtie G)(a,b) \sqcup H, \gamma) - t((G \bowtie H)(c,d) \sqcup F, \gamma),$$

where $(F \bowtie G \bowtie H)(a,b;c,d)$ denotes the graph on $3k - 2$ vertices obtained by identifying in $F \sqcup G \sqcup H$ the vertex a in V_F with the vertex b in V_G, and the vertex d in V_H with the vertex c in V_G. These computations of "elementary" joint cumulants follow from the same argument

as for the computation of the elementary covariance $\text{cov}(A_{F,\psi}, A_{G,\phi})$. We conclude that

$$\lim_{n\to\infty} \frac{\kappa(S_n(F,\gamma), S_n(G,\gamma), S_n(H,\gamma))}{n^{3k-2}}$$

$$= \sum_{1\leq a,b,c\leq k} t((F\bowtie G\bowtie H)(a,b,c),\gamma) + 2\,t(F\sqcup G\sqcup H,\gamma) - t((F\bowtie G)(a,b)\sqcup H,\gamma)$$
$$- t((G\bowtie H)(b,c)\sqcup F,\gamma) - t((F\bowtie H)(a,c)\sqcup G,\gamma)$$

$$+ \sum_{\text{cyc}}\sum_{1\leq a,b\neq c,d\leq k} t((F\bowtie G\bowtie H)(a,b;c,d),\gamma) + t(F\sqcup G\sqcup H,\gamma)$$
$$- t((F\bowtie G)(a,b)\sqcup H,\gamma) - t((G\bowtie H)(c,d)\sqcup F,\gamma)$$

where the symbol $\sum_{\text{cyc}}$ means that one permutes cyclically the roles played by F, G and H (thus, the sum runs over three permutations). Moreover, this limit is attained at speed $O(n^{-1})$. This proves the second part of the proposition, with

$$\kappa_3(F,G,H) = \frac{1}{k^4} \sum_{1\leq a,b,c\leq k} \begin{pmatrix} (F\bowtie G\bowtie H)(a,b,c)+2\,F\times G\times H - F\times(G\bowtie H)(b,c) \\ -G\times(F\bowtie H)(a,c)-H\times(F\bowtie G)(a,b) \end{pmatrix}$$

$$+ \frac{1}{k^4} \sum_{\text{cyc}}\sum_{1\leq a,b\neq c,d\leq k} \begin{pmatrix} (F\bowtie G\bowtie H)(a,b;c,d)+F\times G\times H \\ -H\times(F\bowtie G)(a,b)-F\times(G\bowtie H)(c,d) \end{pmatrix}. \qquad \square$$

Example. Let us illustrate the computation of the maps κ_2 and κ_3 on an example. If $F = G = K_3 = \triangleleft$ is the triangle, then $H = \bowtie$ is the only isomorphism class in $F\bowtie G$, so

$$\kappa_2(K_3, K_3) = H - K_3 \times K_3.$$

For the computation of the limiting third cumulant, the only possible junction of three triangles at a common point is

$$I = \text{[figure: three triangles meeting at a common central point]},$$
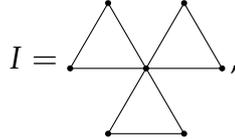
whereas a junction of three triangles $(K_3\bowtie K_3\bowtie K_3)(a,b;c,d)$ with $b\neq c$ is always isomorphic to

$$J = \text{[figure: three triangles in a chain]}.$$
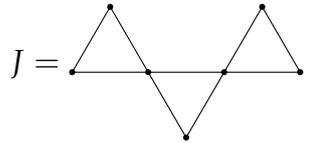
Therefore,

$$\kappa_3(K_3, K_3, K_3) = \frac{1}{3}\,I + 2\,J + \frac{8}{3}\,K_3 \times K_3 \times K_3 - 5\,K_3 \times H.$$

Theorem 21 (Mod-Gaussian convergence of graphons). *Let $\gamma \in \mathscr{G}$ be a graphon, and F be a finite graph with k vertices.*

(1) *The random variable $S_n(F,\gamma) = n^k\, t(F, \Gamma_n(\gamma))$ satisfies the hypotheses of the method of cumulants with parameters $D_n = k^2\, n^{k-1}$, $N_n = n^k$ and $A = 1$, and with limits $\sigma^2 = \kappa_2(F,F)(\gamma)$ and $L = \kappa_3(F,F,F)(\gamma)$.*

(2) If $\kappa_2(F, F)(\gamma) > 0$, then the random variables

$$Y_n(F, \gamma) = \frac{t(F, \Gamma_n(\gamma)) - t(F, \gamma)}{\sqrt{\mathrm{var}(t(F, G_n(\gamma)))}}$$

satisfy all the limiting results from Theorem 3.

(3) Besides, we have the uniform concentration inequality

$$\mathbb{P}[|t(F, \Gamma_n(\gamma)) - \mathbb{E}[t(F, \Gamma_n(\gamma))]| \geq x] \leq 2 \exp\left(-\frac{nx^2}{9k^2}\right).$$

Indeed, we have checked the three hypotheses of the method of cumulants. For the concentration inequality, a slightly better result was proven in [Bor+08], namely,

$$\mathbb{P}[|t(F, \Gamma_n(\gamma)) - \mathbb{E}[t(F, \Gamma_n(\gamma))]| \geq x] \leq 2 \exp\left(-\frac{nx^2}{4k^2}\right).$$

The original proof of this result relied on martingale techniques; here we have shown that up to a constant, one gets similar inequalities from the mod-Gaussian structure. On the other hand, the mod-Gaussian convergence implies the speed of convergence estimate stated in Theorem 8, noticing that

$$\lim_{n\to\infty} n \, \mathrm{var}(t(F, \Gamma_n(\gamma))) = k^2 \, \kappa_2(F, F)(\gamma).$$

Similarly, we get immediately the moderate deviation estimates stated in Theorem 9, with

$$l(F, \gamma) = \frac{k}{6} \frac{\kappa_3(F, F, F)(\gamma)}{(\kappa_2(F, F)(\gamma))^{3/2}}.$$

Example. Fix a graphon γ, and suppose that $t(H, \gamma) \neq (t(K_3, \gamma))^2$, where H is a before the "bowtie" graph. The densities of triangles satisfy the central limit theorem

$$Y_n = \frac{t(K_3, G_n(\gamma)) - t(K_3, \gamma)}{\sqrt{\mathrm{var}(t(K_3, G_n(\gamma)))}} \rightharpoonup \mathcal{N}_\mathbb{R}(0, 1),$$

and the left-hand side of this formula can be replaced by the more explicit random variable

$$\widetilde{Y}_n = \sqrt{n} \, \frac{t(K_3, G_n(\gamma)) - t(K_3, \gamma)}{3 \sqrt{t(H, \gamma) - t(K_3 \times K_3, \gamma)}}$$

since

$$\mathbb{E}[t(K_3, G_n(\gamma))] = t(K_3, \gamma) + O\left(\frac{1}{n}\right);$$

$$\mathrm{var}(t(K_3, G_n(\gamma))) = \frac{9}{n}(t(H, \gamma) - t(K_3 \times K_3, \gamma)) + O\left(\frac{1}{n^2}\right).$$

For instance, if γ is the graphon of the graph function $g(x, y) = xy$, then one can compute $t(K_3, \gamma) = \frac{1}{27}$, and one has the central limit theorem

$$\widetilde{Y}_n = \sqrt{\frac{5n}{4}} \left(\frac{27 \, t(K_3, G_n((x, y) \mapsto xy)) - 1}{3}\right) \rightharpoonup \mathcal{N}_\mathbb{R}(0, 1).$$

The zone of normality of $(Y_n)_{n \in \mathbb{N}}$ or of $(\widetilde{Y}_n)_{n \in \mathbb{N}}$ is a $o(n^{1/6})$. At the edge of this zone, one enters the regime of moderate deviations, and

$$\mathbb{P}[Y_n \geq n^{1/6} x] = \mathbb{P}\left[\widetilde{Y}_n \geq n^{1/6} x\right] (1 + o(1))$$

$$= \frac{e^{-\frac{n^{1/3} x^2}{2}}}{n^{1/6} x \sqrt{2\pi}} \exp\left(\frac{t(I, \gamma) + 6\, t(J, \gamma) + 8\, (t(K_3, \gamma))^3 - 15\, t(H \times K_3, \gamma)}{6\, (t(H, \gamma) - t(K_3 \times K_3, \gamma))^{3/2}} x^3\right) (1 + o(1))$$

for $x > 0$ fixed. Last, for n large enough,

$$d_{\mathrm{Kol}}(Y_n, \mathcal{N}_{\mathbb{R}}(0,1)) \leq \frac{230}{((t(H, \gamma) - t(K_3 \times K_3, \gamma))^{3/2}} \frac{1}{\sqrt{n}}.$$

As far as we know, all these results for the fluctuations of densities of subgraphs in graphon models are new.

5.2. **Fluctuations of random permutations and permutons.**
We have the same kind of results for the fluctuations of the models $(\sigma_n(\pi))_{n \in \mathbb{N}}$ of random permutations, and the main difference between this theory and the theory for graphon models is that one can consider the observables $t(\tau, \sigma_n(\pi))$ or the observables $t(\tau, \Pi_n(\pi))$. The difference between these observables is bounded by $\frac{|\tau|^2}{2n}$, so a central limit theorem for one of these quantities will imply a central limit theorem for the other quantity. Unfortunately, if one wants to get more precise results (*e.g.* a speed of convergence estimate), then one needs to be more careful and to identify a dependency structure for each random variable $t(\tau, \sigma_n(\pi))$ or $t(\tau, \Pi_n(\pi))$. The first case is almost identical to the case of graphs, whereas the second case requires a bit more work (see Lemma 23).

In the sequel, we fix a permuton $\pi \in \mathscr{S}$ and a permutation τ of size k. We set $S_n(\tau, \pi) = \binom{n}{k} t(\tau, \sigma_n(\pi))$ and $\widetilde{S}_n(\tau, \pi) = n^k t(\tau, \Pi_n(\pi))$; beware of the slightly different scalings involved in these formulas. We shall expand the variables $S_n(\tau, \pi)$ and $\widetilde{S}_n(\tau, \pi)$ as sums of bounded random variables with sparse dependency graphs. For $S_n(\tau, \pi)$, if we introduce independent random points $(X_1, Y_1), \ldots, (X_n, Y_n)$ with law π and such that $\sigma_n(\pi) = \mathrm{conf}((X_1, Y_1), \ldots, (X_n, Y_n))$, then we have:

$$S_n(\tau, \pi) = \mathrm{occ}(\tau, \sigma_n(\pi)) = \sum_{\substack{L \subset \llbracket 1, n \rrbracket \\ |L| = k}} A_{\tau, L},$$

with

$$A_{\tau, L} = \begin{cases} 1 & \text{if } \mathrm{conf}(\{(X_l, Y_l),\ l \in L\}) = \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 22. *The graph G_n*
- *with vertex set $V_n = \{L \subset \llbracket 1, n \rrbracket,\ |L| = k\}$,*
- *and with an edge between L and M if $L \cap M \neq \emptyset$*

is a dependency graph for the family of random variables $(A_{\tau, L})_{L \in V_n}$ involved in the expansion of $S_n(\tau, \pi)$.

Proof. Denote $Z_i = (X_i, Y_i)$. Given two families $(A_{\tau,L})_{L \in W_1}$ and $(A_{\tau,M})_{M \in W_2}$ that are not connected in this graph, if $U_1 = \bigcup_{L \in W_1} L$ and $U_2 = \bigcup_{M \in W_2} M$, then we have

$$\left(\bigcup_{L \in W_1} L \right) \cap \left(\bigcup_{M \in W_2} M \right) = U_1 \cap U_2 = \emptyset.$$

Then, $(A_{\tau,L})_{L \in W_1}$ is measurable with respect $\sigma(\{Z_i, i \in U_1\})$, and $(A_{\tau,M})_{M \in W_2}$ is measurable with respect to $\sigma(\{Z_i, i \in U_2\})$. As the random points Z_i are independent, we conclude that $(A_{\tau,L})_{L \in W_1}$ and $(A_{\tau,M})_{M \in W_2}$ are independent families $\square$

The parameters of the dependency graph G_n in Lemma 22 are $D_n = D_{n,k} = k\binom{n-1}{k-1}$, $N_n = N_{n,k} = \binom{n}{k}$ and $A = 1$. The existence of this dependency graph implies the hypothesis (MC1) for $S_n(\tau, \pi)$. For $\widetilde{S}_n(\tau, \pi)$, we have a similar decomposition

$$\widetilde{S}_n(\tau, \pi) = \sum_{1 \leq l_1, \ldots, l_k \leq n} \widetilde{A}_{\tau,(l_1,\ldots,l_k)},$$

but the definition of the elementary variables $\widetilde{A}_{\tau,(l_1,\ldots,l_k)}$ is more complicated than before. It relies on the following:

Lemma 23. *Let τ be a permutation of size $k \leq n$. Given a family of points $(z_1, \ldots, z_n)$ in $[0,1]^2$ in a general configuration, we set*

$$\widehat{t}(\tau, (z_1, \ldots, z_n)) = t(\tau, \pi(\mathrm{conf}(z_1, \ldots, z_n))).$$

There exists a measurable function $F_\tau : ([0,1]^2)^k \to [0,1]$ such that, for any sequence $(z_1, \ldots, z_n)$ in a general configuration,

$$\widehat{t}(\tau, (z_1, \ldots, z_n)) = \sum_{1 \leq l_1, \ldots, l_k \leq n} \frac{1}{n^k} F_\tau(z_{l_1}, \ldots, z_{l_k}).$$

Remark (Values of F_τ on general configurations). It will follow from the proof of the lemma that when $(t_1, \ldots, t_k)$ is a family in general configuration, one has simply $F_\tau(t_1, \ldots, t_k) = 1_{\mathrm{conf}(t_1,\ldots,t_k)=\tau}$. However, in the expansion above, the families $(z_{l_1}, \ldots, z_{l_k})$ are usually not in general configuration, since they might contain points z_l with multiplicity larger than 2 (for instance if $l_1 = l_2$). Thus, our lemma extends in a measurable way the domain of the function $(t_1, \ldots, t_k) \mapsto 1_{\mathrm{conf}(t_1,\ldots,t_k)=\tau}$ to any k-tuple of points.

Proof. By definition, if $\sigma = \mathrm{conf}(z_1, \ldots, z_n)$ and $W_1, \ldots, W_k$ are independent points with law $\pi(\sigma)$, then

$$\widehat{t}(\tau, (z_1, \ldots, z_n)) = \mathbb{P}[\mathrm{conf}(W_1, \ldots, W_k) = \tau].$$

We construct the random variables $W_1, \ldots, W_k$ as follows. If $z_l = (x_l, y_l)$, we denote $\psi_1 : \{x_1, \ldots, x_n\} \to [\![1, n]\!]$ and $\psi_2 : \{y_1, \ldots, y_n\} \to [\![1, n]\!]$ the two increasing bijections; then, the configuration σ is defined by the identity $\sigma(\psi_1(x_l)) = \psi_2(y_l)$. Let us introduce independent random variables $(S_i)_{i \in [\![1,k]\!]}$ uniformly distributed over the square $[0,1]^2$, and also independent discrete random variables $L_1, \ldots, L_k$ uniformly distributed in $[\![1, n]\!]$. We then set:

$$W_i = \left(\frac{\psi_1(x_{L_i})}{n}, \frac{\psi_2(y_{L_i})}{n} \right) - \frac{S_i}{n}.$$

Each W_i has law $\pi(\sigma)$, and on the other hand, since the L_i's and the S_i's are all independent, the random variables $W_1, \ldots, W_k$ are independent. Now, let us construct other random variables $V_1, \ldots, V_k$ such that

$$\mathrm{conf}(V_1, \ldots, V_k) = \mathrm{conf}(W_1, \ldots, W_k) \quad \text{almost surely.}$$

The idea is that instead of choosing random points in squares of size $\frac{1}{n}$ attached to the points $\left(\frac{\psi_1(x_l)}{n}, \frac{\psi_2(y_l)}{n}\right)$, one can choose random points in small squares of size ε attached to the points z_l, and this without changing the configuration. Since $(z_1, \ldots, z_n)$ is in general configuration, we can find an $\varepsilon > 0$ such that $x_l \geq \varepsilon$ and $y_l \geq \varepsilon$ for any l, and such that $|x_l - x_m| > \varepsilon$ and $|y_l - y_m| > \varepsilon$ if $l \neq m$. We then set:

$$V_i = z_{L_i} - \varepsilon\, S_i.$$

It is easy to convince oneself on a diagram that the points $V_1, \ldots, V_k$ can be obtained from the points $W_1, \ldots, W_k$ by a bijection that is increasing in both coordinates; see Figure 11. Therefore, in particular, the configuration is unchanged.
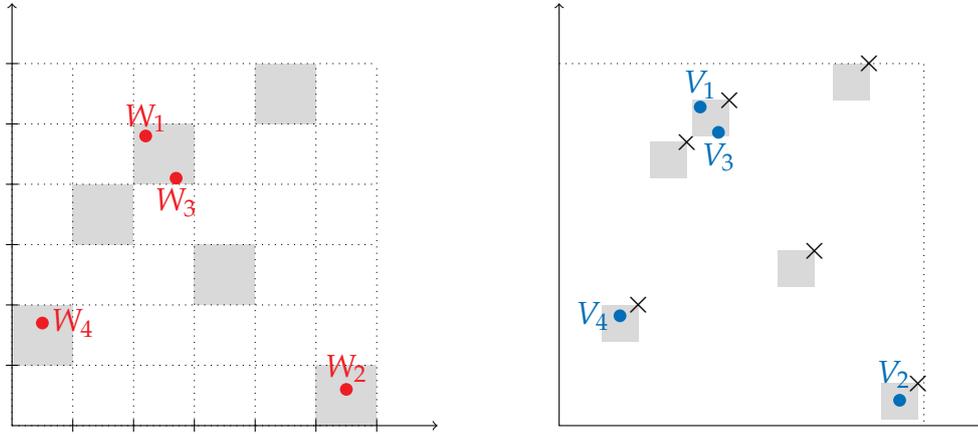

FIGURE 11. The random points $W_1, \ldots, W_k$ and $V_1, \ldots, V_k$ have the same configuration.

We now have:

$$\widehat{t}(\tau, (z_1, \ldots, z_n)) = \mathbb{P}[\mathrm{conf}(W_1, \ldots, W_k) = \tau] = \mathbb{P}[\mathrm{conf}(V_1, \ldots, V_k) = \tau]$$

$$= \sum_{1 \leq l_1, \ldots, l_k \leq n} \frac{1}{n^k}\, \mathbb{P}[\mathrm{conf}(z_{l_1} - \varepsilon\, S_1, \ldots, z_{l_k} - \varepsilon\, S_k) = \tau]$$

Given points $t_1, \ldots, t_k$ in the square $[0,1]^2$, we define

$$F_\tau(t_1, \ldots, t_k) = \lim_{\varepsilon \to 0} \left(\mathbb{P}[\mathrm{conf}(t_1 - \varepsilon\, S_1, \ldots, t_k - \varepsilon\, S_k) = \tau] \right),$$

where $(S_1, \ldots, S_k)$ is a set of independent uniform random variables on $[0,1]^2$. This defines a measurable function of $(t_1, \ldots, t_k) \in ([0,1]^2)^k$, and the limit when ε goes to zero is stationary for any k-tuple of points (this is the same argument of "rescaling of small squares" as before). Since $\widehat{t}(\tau, (z_1, \ldots, z_n)) = \sum_{1 \leq l_1, \ldots, l_k \leq n} \frac{1}{n^k} \mathbb{P}[\mathrm{conf}(z_{l_1} - \varepsilon\, S_1, \ldots, z_{l_k} - \varepsilon\, S_k) = \tau]$ for ε small enough, by taking the limit, one obtains the identity claimed. $\qquad\square$

Given a family of independent random points $Z_1, \ldots, Z_n$ with law π, we can now write:

$$\widetilde{S}_n(\tau, \pi) = n^k \, t(\tau, \pi(\mathrm{conf}(Z_1, \ldots, Z_n))) = n^k \, \widehat{t}(\tau, (Z_1, \ldots, Z_n))$$
$$= \sum_{1 \leq l_1, \ldots, l_k \leq n} F_\tau(Z_{l_1}, \ldots, Z_{l_n}),$$

so if we set $\widetilde{A}_{\tau, (l_1, \ldots, l_k)} = F_\tau(Z_{l_1}, \ldots, Z_{l_n})$, then we have an expansion of $\widetilde{S}_n(\tau, \pi)$ with the following property:

Lemma 24. *The graph $\widetilde{G}_n$*

- *with vertex set $\widetilde{V}_n = (\llbracket 1, n \rrbracket)^k$,*
- *and with an edge between $(l_1, \ldots, l_k)$ and $(m_1, \ldots, m_k)$ if $l_a = m_b$ for some indices a and b*

is a dependency graph for the family of random variables $(\widetilde{A}_{\tau, (l_1, \ldots, l_k)})_{(l_1, \ldots, l_k) \in \widetilde{V}_n}$ involved in the expansion of $\widetilde{S}_n(\tau, \pi)$.

One sees at once that the parameters of the dependency graph $\widetilde{G}_n$ in Lemma 24 can be taken equal to $\widetilde{D}_{n,k} = k^2 \, n^{k-1}$, $\widetilde{N}_{n,k} = n^k$ and $A = 1$. Thus, the sequence of random variables $(\widetilde{S}_n(\tau, \pi))_{n \in \mathbb{N}}$ satisfies the hypothesis (MC1) with respect to these parameters.

We now turn to the computation of the limiting variances and third cumulants. It turns out that one obtains the same result in both cases:

Proposition 25 (Limiting first cumulants of pattern occurrences). *Denote $\mathscr{O}_{\mathfrak{S},k}$ the vector space spanned by the permutations of size k in $\mathscr{O}_{\mathfrak{S}}$. There exists two linear maps*

$$\kappa_2 : \mathscr{O}_{\mathfrak{S},k} \otimes \mathscr{O}_{\mathfrak{S},k} \to \mathscr{O}_{\mathfrak{S}};$$
$$\kappa_3 : \mathscr{O}_{\mathfrak{S},k} \otimes \mathscr{O}_{\mathfrak{S},k} \otimes \mathscr{O}_{\mathfrak{S},k} \to \mathscr{O}_{\mathfrak{S}}$$

such that, for any permutations τ, ρ and μ in $\mathfrak{S}(k)$, and for any permuton π,

$$\lim_{n \to \infty} \frac{\kappa(S_n(\tau, \pi), S_n(\rho, \pi))}{k \binom{n}{k} \binom{n-1}{k-1}} = \lim_{n \to \infty} \frac{\kappa(\widetilde{S}_n(\tau, \pi), \widetilde{S}_n(\rho, \pi))}{k^2 \, n^{2k-1}} = \kappa_2(\tau, \rho)(\pi);$$

$$\lim_{n \to \infty} \frac{\kappa(S_n(\tau, \pi), S_n(\rho, \pi), S_n(\mu, \pi))}{k^2 \binom{n}{k} \binom{n-1}{k-1}^2} = \lim_{n \to \infty} \frac{\kappa(\widetilde{S}_n(\tau, \pi), \widetilde{S}_n(\rho, \pi), \widetilde{S}_n(\mu, \pi))}{k^4 \, n^{3k-2}} = \kappa_3(\tau, \rho, \mu)(\pi).$$

Moreover, in each case, the limit is attained at speed $O(n^{-1})$, with a constant in the $O(\cdot)$ that depends only on k.

Proof. As before, $(Z_n)_{n \in \mathbb{N}}$ will be a sequence of independent random points in $[0,1]^2$ with law π. As in the case of graphs, we can expand the covariance $\mathrm{cov}(S_n(\tau, \pi), S_n(\rho, \pi))$ by bilinearity:

$$\mathrm{cov}(S_n(\tau, \pi), S_n(\rho, \pi)) = \sum_{|L| = |M| = k} \mathrm{cov}(A_{\tau,L}, A_{\rho,M}),$$

the sum running over pairs of subsets with size k in $\llbracket 1, n \rrbracket$. Each variable $A_{\tau,L}$ is equal to $1_{\mathrm{conf}(\{Z_l, \, l \in L\}) = \tau}$; and similarly for the variables $A_{\rho,M}$. We can restrict the sum to pairs of subsets that intersect, since otherwise the covariance vanishes. On the other hand, the pairs (L, M) where $|L \cap M| \geq 2$ yield a total contribution which is a $O(n^{2k-2})$, and which becomes

a $O(n^{-1})$ when divided by $k\binom{n}{k}\binom{n-1}{k-1}$. Hence, we can restrict our attention to pairs (L, M) where $|L \cap M| = 1$:

$$\frac{\kappa(S_n(\tau, \pi), S_n(\rho, \pi))}{k\binom{n}{k}\binom{n-1}{k-1}} = \frac{1}{k\binom{n}{k}\binom{n-1}{k-1}} \sum_{\substack{|L|=|M|=k \\ |L\cap M|=1}} \mathrm{cov}(A_{\tau,L}, A_{\rho,M}) + O(n^{-1}).$$

The remaining covariances can be computed by using the *amalgamated graphical shuffle products* $(\tau \bowtie \rho)(a, b)$ of the permutations τ and ρ. Fix two integers $a, b \in [\![1, k]\!]$. The set $(\tau \boxtimes \rho)(a, b)$ is the set of permutations σ of size $2k - 1$, such that there exists a partition of $[\![1, 2k - 1]\!]$ in three parts

$$I_- = \{i_1 < i_2 < \cdots < i_k\} \setminus \{i_a\};$$
$$J_- = \{j_1 < j_2 < \cdots < j_k\} \setminus \{j_b\};$$
$$K = \{i_a = j_b\}$$

with $I = I_- \sqcup K$ that makes appear τ as a pattern of σ, and $J = J_- \sqcup K$ that makes appear ρ as a pattern of σ. Notice that there can exist several partitions $[\![1, 2k - 1]\!] = I_- \sqcup J_- \sqcup K$ with this property for σ; in this case, we count σ several times. In other words, $(\tau \boxtimes \rho)(a, b)$ is defined as a multiset. For instance, if $k = 3$, $\tau = 312$, $\rho = 132$, $a = 2$ and $b = 3$, then there are 9 permutations in $(\tau \boxtimes \rho)(a, b)$, given by the diagrams of Figure 12.

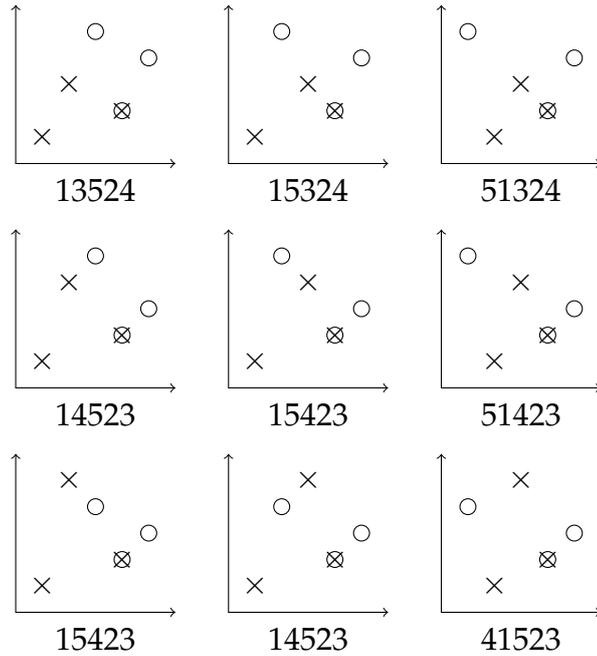FIGURE 12. The permutations involved in the amalgamated graphical shuffle product of $\tau = 312$ and $\rho = 132$ at $(a, b) = (2, 3)$.

We now define the amalgamated graphical shuffle product of τ and ρ at the points a and b by the following operation in the algebra of permutations $\mathscr{O}$:

$$(\tau \bowtie \rho)(a, b) = \frac{(k!)^2}{(2k - 1)!} \sum_{\sigma \in (\rho \boxtimes \tau)(a, b)} \sigma.$$

The combinatorial factor $\frac{(k!)^2}{(2k-1)!}$ is akin to the combinatorial factor $\frac{k_1! k_2!}{k!}$ involved in the usual graphical shuffle product.

Let L and M be two subsets of $[\![1, n]\!]$ which meet at one point and which have cardinality k. We have

$$\mathrm{cov}(A_{\tau, L}, A_{\rho, M}) = \mathbb{P}[\mathrm{conf}(\{Z_l, \ l \in L\}) = \tau \text{ and } \mathrm{conf}(\{Z_m, \ m \in M\}) = \rho]$$
$$- \mathbb{P}[\mathrm{conf}(\{Z_l, \ l \in L\}) = \tau] \, \mathbb{P}[\mathrm{conf}(\{Z_m, \ m \in M\}) = \rho],$$

and since the Z_n's are independent and with same law, this quantity does not depend on the subsets L and M. It is for instance equal to

$$\mathbb{P}[\mathrm{conf}(Z_1, Z_2, \dots, Z_k) = \tau \text{ and } \mathrm{conf}(Z_k, Z_{k+1}, \dots, Z_{2k-1}) = \rho] - t(\tau, \pi) \, t(\rho, \pi).$$

Let us then analyse the event $E = \{\mathrm{conf}(Z_1, \dots, Z_k) = \tau \text{ and } \mathrm{conf}(Z_k, \dots, Z_{2k-1}) = \rho\}$. If one writes the multiset $(\tau \boxtimes \rho)(a, b)$ as a set of pairs $(\sigma, I_- \sqcup J_- \sqcup K)$, and if ψ is the unique increasing bijection $\{X_1, \dots, X_{2k-1}\} \to [\![1, 2k - 1]\!]$, then

$$E = \bigsqcup_{\substack{1 \le a, b \le k \\ (\sigma, I_- \sqcup J_- \sqcup K) \in (\tau \boxtimes \rho)(a,b)}} \left\{ \begin{array}{c} \mathrm{conf}(Z_1, \dots, Z_{2k-1}) = \sigma \\ \psi(\{X_1, \dots, X_{k-1}\} \sqcup \{X_k\} \sqcup \{X_{k+1}, \dots, X_{2k-1}\}) = I_- \sqcup K \sqcup J_- \end{array} \right\}.$$

However, for any sequence $Z_1, \dots, Z_{2k-1}$ of independent points with law π, the relabeling map ψ is independent from the configuration σ of the points $Z_1, \dots, Z_{2k-1}$, therefore,

$$\mathbb{P}[E] = \sum_{\substack{1 \le a, b \le k \\ (\sigma, I_- \sqcup J_- \sqcup K)}} t(\sigma, \pi) \, \mathbb{P}[\psi(\{X_1, \dots, X_{k-1}\} \sqcup \{X_k\} \sqcup \{X_{k+1}, \dots, X_{2k-1}\}) = I_- \sqcup K \sqcup J_-]$$

$$= \frac{((k-1)!)^2}{(2k-1)!} \sum_{\substack{1 \le a, b \le k \\ (\sigma, I_- \sqcup J_- \sqcup K)}} t(\sigma, \pi) = \frac{1}{k^2} \sum_{1 \le a, b \le k} ((\tau \bowtie \rho)(a, b))(\pi).$$

We conclude that

$$\frac{\kappa(S_n(\tau, \pi), S_n(\rho, \pi))}{k \binom{n}{k} \binom{n-1}{k-1}} = \frac{1}{k^2} \sum_{1 \le a, b \le k} (((\tau \bowtie \rho)(a, b))(\pi) - t(\tau, \pi) \, t(\rho, \pi)) + O(n^{-1}).$$

This proves the existence of a map κ_2, as in the Lemma, where κ_2 is defined by

$$\kappa_2(\tau, \rho) = \frac{1}{k^2} \sum_{1 \le a, b \le b} ((\tau \bowtie \rho)(a, b) - \tau \times \rho).$$

For the estimation of the covariances of the quantities $\widetilde{S}_n$, we can write:

$$\mathrm{cov}(\widetilde{S}_n(\tau, \pi), \widetilde{S}_n(\rho, \pi)) = \sum_{\substack{(l_1, \dots, l_k) \in [\![1, n]\!]^k \\ (m_1, \dots, m_k) \in [\![1, n]\!]^k}} \mathrm{cov}(\widetilde{A}_{\tau, (l_1, \dots, l_k)}, \widetilde{A}_{\rho, (m_1, \dots, m_k)})$$

$$= \sum_{1 \le a, b \le k^2} \sum_{\substack{(l_1 \ne \dots \ne l_k) \in [\![1, n]\!]^k \\ (m_1 \ne \dots \ne m_k) \in [\![1, n]\!]^k \\ l_a = m_b}} \mathrm{cov}(\widetilde{A}_{\tau, (l_1, \dots, l_k)}, \widetilde{A}_{\rho, (m_1, \dots, m_k)}) + O(n^{2k-2})$$

where on the second line the sum is over pairs of k-arrangements with only one equality of indices $l_a = m_b$. However, by a previous remark, if $(l_1 \ne \dots \ne l_k)$ is an arrangement, then

$$\widetilde{A}_{\tau, (l_1, \dots, l_k)} = F_\tau(Z_{l_1}, \dots, Z_{l_k}) = 1_{\mathrm{conf}(Z_{l_1}, \dots, Z_{l_k}) = \tau} = A_{\tau, \{l_1, \dots, l_k\}},$$

because in this case $(Z_{l_1}, \ldots, Z_{l_k})$ is in a general configuration with probability 1. This implies immediately that $\frac{\mathrm{cov}(\widetilde{S}_n(\tau,\pi),\widetilde{S}_n(\rho,\pi))}{k^2\,n^{2k-1}}$ has the same asymptotics as $\frac{\kappa(S_n(\tau,\pi),S_n(\rho,\pi))}{k\binom{n}{k}\binom{n-1}{k-1}}$.

The calculation of the asymptotics of the third cumulants follows the same lines. In order to evaluate

$$\frac{1}{k^2\binom{n}{k}\binom{n}{k-1}^2}\,\kappa\left(\widetilde{S}_n(\tau,\pi),\widetilde{S}_n(\rho,\pi),\widetilde{S}_n(\mu,\pi)\right),$$

we need to compute $\kappa(A_{\tau,L}, A_{\rho,M}, A_{\mu,N})$ when $L = \{l_1 < l_2 < \cdots < l_k\}$, $M = \{m_1 < \cdots < m_k\}$ and $N = \{n_1 < \cdots < n_k\}$ are three k-subsets of $[\![1,n]\!]$ that meet one of the following conditions:

(a) either $L \cap M = M \cap N = L \cap N = \{l_a = m_b = n_c\}$ and there are no other equality of indices;

(b) or, $L \cap M = \{l_a = m_b\}$ and $M \cap N = \{m_c = n_d\}$ with $b \neq c$, and there are no other equality of indices;

(c) or, one has the same configuration as (b), up to a cyclic permutation of the roles played by L, M, N.

In the first case, the cumulant $\kappa(A_{\tau,L}, A_{\rho,M}, A_{\mu,N})$ does not depend on L, M, N, and it is equal to

$$\mathbb{P}[\mathrm{conf}(Z_0, Z_1, \ldots, Z_{k-1}) = \tau,\ \mathrm{conf}(Z_0, Z_k, \ldots, Z_{2k-2}) = \rho,\ \mathrm{conf}(Z_0, Z_{2k-1}, \ldots, Z_{3k-3}) = \mu]$$

$$-\frac{1}{k^3}\sum_{1\leq a,b,c\leq k}\left((\tau\bowtie\rho)(a,b)(\pi)\,\mu(\pi) + (\rho\bowtie\mu)(b,c)(\pi)\,\tau(\pi) + (\tau\bowtie\mu)(a,c)(\pi)\,\rho(\pi)\right)$$

$$+2\,\tau(\pi)\,\rho(\pi)\,\mu(\pi).$$

In this formula, we have written $\sigma(\pi)$ for the evaluation of a pattern density, instead of $t(\sigma,\pi)$. To compute the first term $\mathbb{P}[E]$, we introduce the multiset $(\tau\boxtimes\rho\boxtimes\mu)(a,b,c)$, which is the set of permutations $\sigma\in\mathfrak{S}(3k-2)$ such that there exists a partition of $[\![1,3k-2]\!]$ in four parts

$$H_- = \{h_1 < h_2 < \cdots < h_k\}\setminus\{h_a\};$$
$$I_- = \{i_1 < i_2 < \cdots < i_k\}\setminus\{i_b\};$$
$$J_- = \{j_1 < j_2 < \cdots < j_k\}\setminus\{j_c\};$$
$$K = \{h_a = i_b = j_c\}$$

with $H = H_-\sqcup K$ that makes appear τ as a pattern of σ, and similarly for $(I_-\sqcup K,\rho)$ and $(J_-\sqcup K,\mu)$. Then, if

$$(\tau\bowtie\rho\bowtie\mu)(a,b,c) = \frac{(k!)^3}{(3k-2)!}\sum_{\sigma\in(\tau\boxtimes\rho\boxtimes\mu)(a,b,c)}\sigma,$$

the same arguments as for the computation of the limiting covariance yields

$$\mathbb{P}[E] = \frac{1}{k^3}\sum_{1\leq a,b,c\leq k}((\tau\bowtie\rho\bowtie\mu)(a,b,c))(\pi).$$

On the other hand, there are

$$\binom{n}{3k-2} \frac{(3k-2)!}{((k-1)!)^3}$$

parts A, B, C that meet the condition (a). Hence, the contribution to the limit of $\widetilde{d}_n$ of the parts A, B, C satisfying (a) is given by evaluating on the permuton π the observable

$$\frac{1}{k^4} \sum_{1 \le a,b,c \le k} \binom{(\tau \bowtie \rho \bowtie \mu)(a,b,c) + 2\,\tau \times \rho \times \mu - (\tau \bowtie \rho)(a,b) \times \mu}{-(\rho \bowtie \mu)(b,c) \times \tau - (\tau \bowtie \mu)(a,c) \times \rho}.$$

To treat the other cases (b) or (c), we introduce the multiset $(\tau \boxtimes \rho \boxtimes \mu)(a,b;c,d)$, which is the set of permutations $\sigma \in \mathfrak{S}(3k-2)$ such that there exists a partition of $[\![1, 3k-2]\!]$ in five parts

$$
\begin{aligned}
H_- &= \{h_1 < h_2 < \cdots < h_k\} \setminus \{h_a\}; \\
I_- &= \{i_1 < i_2 < \cdots < i_k\} \setminus \{i_b, i_c\}; \\
J_- &= \{j_1 < j_2 < \cdots < j_k\} \setminus \{j_d\}; \\
K &= \{h_a = i_b\}; \\
L &= \{i_c = j_d\}
\end{aligned}
$$

with $H = H_- \sqcup K$ that makes appear τ as a pattern of σ, and similarly for the two pairs $(I_- \sqcup K \sqcup L, \rho)$ and $(J_- \sqcup L, \mu)$. As usual, we count a permutation σ several times if several partitions satisfy these conditions. If we define

$$(\tau \bowtie \rho \bowtie \mu)(a,b;c,d) = \frac{(k!)^3}{(3k-2)!} \sum_{\sigma \in (\tau \boxtimes \rho \boxtimes \mu)(a,b;c,d)} \sigma,$$

then for any subsets L, M, N that satisfy (b), the cumulant $\kappa(A_{\tau,L}, A_{\rho,M}, E_{\mu,N})$ is obtained by evaluating the observable

$$\frac{1}{k^3(k-1)} \sum_{1 \le a,b \ne c,d \le k} (\tau \bowtie \rho \bowtie \mu)(a,b;c,d) + \tau \times \rho \times \mu - (\tau \bowtie \rho)(a,b) \times \mu - (\rho \bowtie \mu)(c,d) \times \tau$$

on π. Since the number of parts L, M, N which satisfy (b) is

$$\binom{n}{3k-2} \frac{(3k-2)!}{((k-1)!)^2 (k-2)!},$$

we conclude that the contribution of the case (b) to the limit of the rescaled joint cumulant of $S_n(\tau, \pi)$, $S_n(\rho, \pi)$ and $S_n(\mu, \pi)$ is given by the observable

$$\frac{1}{k^4} \sum_{1 \le a,b \ne c,d \le k} (\tau \bowtie \rho \bowtie \mu)(a,b;c,d) + \tau \times \rho \times \mu - (\tau \bowtie \rho)(a,b) \times \mu - (\rho \bowtie \mu)(c,d) \times \tau.$$

Hence, we conclude that if κ_3 is the linear map $(\mathscr{O}_{\mathfrak{S},k})^{\otimes 3} \to \mathscr{O}_{\mathfrak{S}, 3k-2}$ defined by

$$\kappa_3(\tau, \rho, \mu) = \frac{1}{k^4} \sum_{1 \le a,b,c \le k} \binom{(\tau \bowtie \rho \bowtie \mu)(a,b,c) + 2\,\tau \times \rho \times \mu - (\tau \bowtie \rho)(a,b) \times \mu}{-(\rho \bowtie \mu)(b,c) \times \tau - (\tau \bowtie \mu)(a,c) \times \rho}$$

$$+ \frac{1}{k^4} \sum_{\text{cyc}} \sum_{1 \le a,b \ne c,d \le k} \binom{(\tau \bowtie \rho \bowtie \mu)(a,b;c,d) + \tau \times \rho \times \mu}{-(\tau \bowtie \rho)(a,b) \times \mu - (\rho \bowtie \mu)(c,d) \times \tau}$$

then

$$\frac{\kappa(S_n(\tau, \pi), S_n(\rho, \pi), S_n(\mu, \pi))}{k^2 \binom{n}{k} \binom{n-1}{k-1}^2} = \kappa_3(\tau, \rho, \mu)(\gamma) + O(n^{-1}).$$

Finally, the asymptotics of the rescaled joint cumulant $\frac{\kappa(\widetilde{S}_n(\tau,\pi),\widetilde{S}_n(\rho,\pi),\widetilde{S}_n(\mu,\pi))}{k^4\,n^{3k-2}}$ are the same, for the same reason as in the case of covariances: when summing over k-tuples $(l_1,\ldots,l_k)$, a random variable $\widetilde{A}_{\tau,(l_1,\ldots,l_k)}$ is equal to $A_{\tau,\{l_1,\ldots,l_k\}}$ as soon as $(l_1,\ldots,l_k)$ is an arrangement, and this case is the main contribution when computing the joint cumulant of order 3. □

Remark (Cardinality of an amalgamated graphical shuffle product). Let us give a general formula for the cardinality of the multiset $(\tau \boxtimes \rho)(a,b)$ involved in the definition of $(\tau \bowtie \rho)(a,b)$ and in the computation of the limiting covariances. Knowing a and b, in order to construct σ in $(\tau \boxtimes \rho)(a,b)$, we first need to decide which indices among those smaller than $i_a = j_b$ will fall in I_-, and which indices among those larger than $i_a = j_b$ will fall in I_-. If one draws the diagram of τ with circles and the diagram of ρ with crosses as in Figure 12, then this amounts to choose the *horizontal* positions of the circles and crosses. These choices determine the partition $I_- \sqcup J_- \sqcup K$ of $[\![1, 2k-1]\!]$, and there are

$$\binom{a+b-2}{a-1}\binom{2k-a-b}{k-a}$$

possibilities. We then also have to choose the *vertical* positions of the circles and crosses, and this enumeration is the same as before, but with a and b replaced by $\tau(a)$ and $\rho(b)$. Hence,

$$|(\tau \boxtimes \rho)(a,b)| = \binom{a+b-2}{a-1}\binom{2k-a-b}{k-a}\binom{\tau(a)+\rho(b)-2}{\tau(a)-1}\binom{2k-\tau(a)-\rho(b)}{k-\tau(a)}.$$

We have thus checked all the hypotheses for the following theorem:

Theorem 26 (Mod-Gaussian convergence of permutons). *Let $\pi \in \mathscr{S}$ be a permuton, and τ be a finite permutation in $\mathfrak{S}(k)$.*

(1) *The random variable $S_n(\tau, \pi) = \mathrm{occ}(\tau, \sigma_n(\pi)) = \binom{n}{k} t(\tau, \sigma_n(\pi))$ satisfies the hypotheses of the method of cumulants with parameters $D_n = k\binom{n-1}{k-1}$, $N_n = \binom{n}{k}$ and $A = 1$, and with limits $\sigma^2 = \kappa_2(\tau, \tau)(\pi)$ and $L = \kappa_3(\tau, \tau, \tau)(\pi)$. The random variable $\widetilde{S}_n(\tau, \pi) = n^k t(\tau, \Pi_n(\pi))$ satisfies the hypotheses of the method of cumulants with parameters $\widetilde{D}_n = n^2\,n^{k-1}$, $\widetilde{N}_n = n^k$ and $A = 1$, and with the same limits as for $S_n(\tau, \pi)$.*

(2) *If $\kappa_2(\tau, \tau)(\pi) > 0$, then the random variables*

$$Y_n(\tau, \pi) = \frac{t(\tau, \sigma_n(\pi)) - t(\tau, \pi)}{\sqrt{\mathrm{var}(t(\tau, \sigma_n(\pi)))}};$$

$$\widetilde{Y}_n(\tau, \pi) = \frac{t(\tau, \Pi_n(\pi)) - \mathbb{E}[t(\tau, \Pi_n(\pi))]}{\sqrt{\mathrm{var}(t(\tau, \Pi_n(\pi)))}}$$

satisfy all the limiting results from Theorem 3.

(3) *We have the concentration inequalities*

$$\mathbb{P}[|t(\tau, \sigma_n(\pi)) - t(\tau, \pi)| \geq x] \leq 2 \exp\left(-\frac{nx^2}{9k^2}\right);$$

$$\mathbb{P}[|t(\tau, \Pi_n(\pi)) - \mathbb{E}[t(\tau, \Pi_n(\pi))]| \geq x] \leq 2 \exp\left(-\frac{nx^2}{9k^2}\right).$$

Regarding the concentration inequalities, the first one appears in [Hop+11, Theorem 4.2] with a constant $\frac{1}{2}$ instead of $\frac{1}{9}$; the second one seems new. On the other hand, the second part of the theorem implies immediately the estimates of Theorems 10 and 11, with

$$\lim_{n\to\infty} n \operatorname{var}(t(\tau, \sigma_n(\pi))) = \lim_{n\to\infty} n \operatorname{var}(t(\tau, \Pi_n(\pi))) = k^2 \kappa_2(\tau, \tau)(\pi);$$

$$l(\tau, \pi) = \frac{k}{6} \frac{\kappa_3(\tau, \tau, \tau)(\pi)}{(\kappa_2(\tau, \tau)(\pi))^{3/2}}.$$

Besides, we have similar estimates for the random variables $\widetilde{Y}_n(\tau, \pi)$ instead of the variables $Y_n(\tau, \pi)$.

Example. Let $\tau = 21$. The quantity $S_n(\tau, \pi)$ is the number of inversions of the random permutation $\sigma_n(\pi)$. One computes

$$\kappa_2(\tau, \tau) = 321 + \frac{1}{3}(312 + 231) - 4321 - \frac{2}{3}(3412 + 3421 + 4231 + 4312)$$

$$- \frac{1}{3}(2143 + 2413 + 2431 + 3142 + 3241 + 4132 + 4213).$$

To fix the ideas, let us consider the case when π is the uniform measure on $[0,1]^2$. Then, $t(\tau, \pi) = \frac{1}{k!}$ for any permutation τ of size k, therefore, $\kappa_2(\tau, \tau)(\pi) = \frac{1}{36}$. Theorem 10 yields then

$$d_{\mathrm{Kol}}\left(3\sqrt{n}\left(\frac{\operatorname{inv}(\sigma_n(\pi))}{\binom{n}{2}} - \frac{1}{2}\right), \mathcal{N}_{\mathbb{R}}(0,1)\right) \leq \frac{33000}{\sqrt{n}}$$

for n large enough, where $\operatorname{inv}(\sigma)$ is the number of inversions of a permutation σ. This recovers, up to the value of the constant, a result of Fulman [Ful04]. Similar estimates of the speed of convergence in the central limit theorem of the number of occurences of other patterns (and even vincular patterns) in uniform random permutations have recently been given in [Hof17]. In the general case of a random permutation $\sigma_n(\pi)$, such estimates seem however to be new.

Remark (on hypothesis $\kappa_2(\tau, \tau) > 0$). The most studied case is of course the case of the uniform permuton: *i.e.* π is the Lebesgue measure on $[0,1]^2$, which implies that $\sigma_n(\pi)$ is a uniform random permutation of size n. Then one can prove that the hypothesis $\kappa_2(\tau, \tau) > 0$ is satisfied for all patterns τ: see [JNZ15; Hof17].

5.3. **Fluctuations of random integer partitions.** The case of the models of random partitions $(\lambda_n(\omega))_{n\in\mathbb{N}}$ is very similar to the two other cases, and most of the arguments that we shall use come from [FMN16, Chapter 11]. The main new ingredient will be a change of basis argument in the Kerov–Olshanski algebra of polynomial functions on Young diagrams, which is used in the proof of:

Proposition 27 (Uniform bounds for the cumulants of the Frobenius moments). *Let $\omega \in \mathscr{P}$ be a parameter of the Thoma simplex, and ρ an integer partition with size k. The random variable $S_n(\rho, \omega) = p_\rho(\lambda_n(\omega)) = n^k t(\rho, \Omega_n(\omega))$ satisfies the uniform bound on cumulants (MC1) with parameters $D_n = k^2 n^{k-1}$, $N_n = n^k$ and $A = 1$.*

Proof. Given two integer partitions λ and ρ with arbitrary sizes, we define the renormalised character value $\Sigma_\rho(\mu)$ by the following formula:

$$\Sigma_\rho(\mu) = \begin{cases} |\mu|(|\mu| - 1) \cdots (|\mu| - |\rho| + 1) \chi^\mu(\sigma_\rho) & \text{if } |\mu| \geq |\rho|, \\ 0 & \text{if } |\mu| < |\rho|, \end{cases}$$

where σ_ρ is a permutation with cycle-type ρ, and $\chi^\mu(\sigma) = \frac{\operatorname{tr} \rho^\mu(\sigma)}{\operatorname{tr} \rho^\mu(1)}$ is the normalised irreducible character of the symmetric group $\mathfrak{S}(|\mu|)$ that is associated with the integer partition μ. We refer to [IO02] and [Mél17, Section 7.3] for a presentation of these functions, and to *loc. cit.* or to [Sag01] for generalities on the representations of the symmetric groups. The functions Σ_ρ with $\rho \in \mathfrak{P}$ span an algebra called the Kerov–Olshanski algebra of polynomial functions on Young diagrams [KO94]. The use of the renormalised character values in the study of the central measures is natural since these measures $\mathbb{P}_{n,\omega}$ were given in Section 3.3 a representation theoretic definition.

In [FMN16, Section 11], we proved the mod-Gaussian convergence of the random character values

$$\Sigma_\rho(\lambda_n(\omega)) = n^{\downarrow k} \chi^{\lambda_n(\omega)}(\sigma_\rho) \quad \text{for } n \geq k = |\rho|$$

after an appropriate scaling. Remark 11.4.2 in *loc. cit.* ensures that for any integer partitions $\rho^{(1)}, \ldots, \rho^{(r)}$,

$$\left| \kappa\left(\Sigma_{\rho^{(1)}}(\lambda_n(\omega)), \Sigma_{\rho^{(2)}}(\lambda_n(\omega)), \ldots, \Sigma_{\rho^{(r)}}(\lambda_n(\omega)) \right) \right| \leq \left(\frac{2k^2}{n} \right)^{r-1} r^{r-2} n^{\downarrow k_1} n^{\downarrow k_2} \cdots n^{\downarrow k_r}$$

where $k_i = |\rho^{(i)}|$, and $k = \max(k_1, \ldots, k_r)$. On the other hand, the functions Σ_ρ are related to the observables of integer partitions and Thoma parameters by a change of basis formula

$$p_\rho = \sum_{|\mu| \leq |\rho|} c_{\rho,\mu} \Sigma_\mu,$$

where the coefficients $c_{\rho,\mu}$ are positive rational numbers, and where the only non-zero coefficient $c_{\rho,\mu}$ with $|\mu| = |\rho|$ is $c_{\rho,\rho} = 1$. For instance,

$$p_6 = \Sigma_6 + 6\,\Sigma_{(3,2)} + 6\,\Sigma_{(4,1)} + \frac{95}{4}\,\Sigma_4 + 15\,\Sigma_{(2,1,1)} + 35\,\Sigma_{(2,1)} + \frac{91}{16}\,\Sigma_2.$$

We refer to [Was81, §3.6] and [IO02, Section 3] for a description of the relations between the observables p_ρ and the observables Σ_ρ; see also [Mél17, Section 7.3]. It is not needed to know exactly what are the coefficients $c_{\rho,\mu}$; we shall only use the fact that

$$n^{|\rho|} \geq \prod_{i=1}^r \left(\left(n - \frac{1}{2} \right)^{\rho_i} - \left(-\frac{1}{2} \right)^{\rho_i} \right) = p_\rho(1^n) = \sum_{|\mu| \leq |\rho|} c_{\rho,\mu} \Sigma_\mu(1^n) = \sum_{|\mu| \leq |\rho|} c_{\rho,\mu} n^{\downarrow |\mu|}.$$

We can expand the r-th cumulant of $p_\rho(\lambda_n(\omega))$ by multilinearity:

$$\left| \kappa^{(r)}(p_\rho(\lambda_n(\omega))) \right| \leq \sum_{\mu^{(1)}, \ldots, \mu^{(r)}} c_{\rho,\mu^{(1)}} \cdots c_{\rho,\mu^{(r)}} \left| \kappa\left(\Sigma_{\mu^{(1)}}(\lambda_n(\omega)), \ldots, \Sigma_{\mu^{(r)}}(\lambda_n(\omega)) \right) \right|$$

$$\leq \left(\frac{2k^2}{n} \right)^{r-1} r^{r-2} \sum_{\mu^{(1)}, \ldots, \mu^{(r)}} c_{\rho,\mu^{(1)}} \cdots c_{\rho,\mu^{(r)}} n^{\downarrow |\mu^{(1)}|} \cdots n^{\downarrow |\mu^{(r)}|}$$

$$\leq \left(\frac{2k^2}{n} \right)^{r-1} r^{r-2} n^{kr} = n^k (2k^2 n^{k-1})^{r-1} r^{r-2}. \qquad \square$$

Remark (Dependency graphs for Frobenius moments (absence of)). Proposition 27 is the only case in this paper where we obtain a uniform bound on cumulants without specifying a dependency graph for the underlying random variables. On the other hand, the bound on the joint cumulants of the variables $\Sigma_\rho(\lambda_n(\omega))$ follows indeed from the theory of dependency graphs, but in a setting of non-commutative probability theory: see [FMN16, Section 11.3].

We now introduce the maps κ_2 and κ_3 on the algebra of partitions. Given two integer partitions $\rho = (\rho_1, \ldots, \rho_{\ell(\rho)})$ and $\mu = (\mu_1, \ldots, \mu_{\ell(\mu)})$ with size k, for $a \in [\![1, \ell(\rho)]\!]$ and $b \in [\![1, \ell(\mu)]\!]$, we set

$$(\rho \bowtie \mu)(a, b) = (\rho \setminus \{\rho_a\}) \sqcup (\mu \setminus \{\mu_b\}) \sqcup \{\rho_a + \mu_b - 1\},$$

the right-hand side being reordered in order to obtain an integer partition. For instance, $((3, 2, 2) \bowtie (4, 1))(2, 2) = (4, 3, 2, 2)$. The map $\kappa_2 : \mathscr{O}_{\mathfrak{P},k} \otimes \mathscr{O}_{\mathfrak{P},k} \to \mathscr{O}_{\mathfrak{P}}$ is then defined by

$$\kappa_2(\rho, \mu) = \frac{1}{k^2} \sum_{a=1}^{\ell(\rho)} \sum_{b=1}^{\ell(\mu)} \rho_a \mu_b \left((\rho \bowtie \mu)(a, b) - \rho \times \mu \right).$$

Similarly, given three integer partitions ρ, μ and ν in $\mathfrak{P}(k)$, we define

$$(\rho \bowtie \mu \bowtie \nu)(a, b, c) = (\rho \setminus \{\rho_a\}) \sqcup (\mu \setminus \{\mu_b\}) \sqcup (\nu \setminus \{\nu_c\}) \sqcup \{\rho_a + \mu_b + \nu_c - 2\};$$

$$(\rho \bowtie \mu \bowtie \nu)(a, b; c, d) = (\rho \setminus \{\rho_a\}) \sqcup (\mu \setminus \{\mu_b, \mu_c\}) \sqcup (\nu \setminus \{\nu_d\}) \sqcup \{\rho_a + \mu_b - 1, \mu_c + \nu_d - 1\}$$

where in the first case we have $a \in [\![1, \ell(\rho)]\!]$, $b \in [\![1, \ell(\mu)]\!]$ and $c \in [\![1, \ell(\nu)]\!]$; and in the second case we have $a \in [\![1, \ell(\rho)]\!]$, $b \neq c \in [\![1, \ell(\mu)]\!]$ and $d \in [\![1, \ell(\nu)]\!]$. We then set

$$\kappa_3(\rho, \mu, \nu) = \frac{1}{k^4} \sum_{a=1}^{\ell(\rho)} \sum_{b=1}^{\ell(\mu)} \sum_{c=1}^{\ell(\nu)} \rho_a \mu_b \nu_c \left(\begin{smallmatrix} (\rho \bowtie \mu \bowtie \nu)(a,b,c) + 2\,\rho \times \mu \times \nu - (\rho \bowtie \mu)(a,b) \times \nu \\ - (\mu \bowtie \nu)(b,c) \times \rho - (\rho \bowtie \nu)(a,c) \times \mu \end{smallmatrix} \right)$$

$$+ \frac{1}{k^4} \sum_{\text{cyc}} \sum_{a=1}^{\ell(\rho)} \sum_{b=1}^{\ell(\mu)} \sum_{c=1}^{\ell(\nu)} \rho_a \mu_b (\mu_b - 1) \nu_c \left(\begin{smallmatrix} \rho \times \mu \times \nu + (\rho \bowtie \mu \bowtie \nu)(a,b,c) \\ - (\rho \bowtie \mu)(a,b) \times \nu - (\mu \bowtie \nu)(b,c) \times \rho \end{smallmatrix} \right)$$

$$+ \frac{1}{k^4} \sum_{\text{cyc}} \sum_{a=1}^{\ell(\rho)} \sum_{(b \neq c)=1}^{\ell(\mu)} \sum_{d=1}^{\ell(\nu)} \rho_a \mu_b \mu_c \nu_d \left(\begin{smallmatrix} \rho \times \mu \times \nu + (\rho \bowtie \mu \bowtie \mu)(a,b;c,d) \\ - (\rho \bowtie \mu)(a,b) \times \nu - (\mu \bowtie \nu)(c,d) \times \rho \end{smallmatrix} \right)$$

with the cyclic sums meaning that we permute cyclically the roles played by ρ, μ and ν. We then have:

Proposition 28 (Limiting first cumulants of Frobenius moments). *For any partitions ρ, μ and ν with size k and any Thoma parameter $\omega \in \mathscr{P}$,*

$$\lim_{n \to \infty} \frac{\kappa(\Sigma_\rho(\lambda_n(\omega)), \Sigma_\mu(\lambda_n(\omega)))}{k^2 \, n^{2k-1}} = \lim_{n \to \infty} \frac{\kappa(S_n(\rho, \omega), S_n(\mu, \omega))}{k^2 \, n^{2k-1}} = \kappa_2(\rho, \mu)(\omega);$$

and

$$\lim_{n \to \infty} \frac{\kappa(\Sigma_\rho(\lambda_n(\omega)), \Sigma_\mu(\lambda_n(\omega)), \Sigma_\nu(\lambda_n(\omega)))}{k^4 \, n^{3k-2}} = \lim_{n \to \infty} \frac{\kappa(S_n(\rho, \omega), S_n(\mu, \omega), S_n(\nu, \omega))}{k^4 \, n^{3k-2}}$$

$$= \kappa_3(\rho, \mu, \nu)(\omega),$$

In each case, the limit is attained at speed $O(n^{-1})$, with a constant in the $O(\cdot)$ that depends only on k.

Proof. The case of the random character values is the content of [FMN16, Propositions 11.4.3 and 11.4.4], and the case of the observables $S_n(\rho, \omega) = p_\rho(\lambda_n(\omega))$ follows by using the expansion $p_\rho = \Sigma_\rho +$ terms with lower degree in the algebra of polynomial functions on Young diagrams. $\qquad\square$

Theorem 29 (Mod-Gaussian convergence of random integer partitions). *Fix $\omega \in \mathscr{P}$ and $\rho \in \mathfrak{P}(k)$.*

(1) *The random variable $S_n(\rho, \omega) = n^k\, t(\rho, \Omega_n(\omega))$ satisfies the hypotheses of the method of cumulants with parameters $D_n = k^2\, n^{k-1}$, $N_n = n^k$ and $A = 1$, and with limits $\sigma^2 = \kappa_2(\rho, \rho)(\omega)$ and $L = \kappa_3(\rho, \rho, \rho)(\omega)$.*

(2) *If $\kappa_2(\rho, \rho)(\omega) > 0$, then the random variables*

$$Y_n(\rho, \omega) = \frac{t(\rho, \Omega_n(\omega)) - \mathbb{E}[t(\rho, \Omega_n(\omega))]}{\sqrt{\mathrm{var}(t(\rho, \Omega_n(\omega)))}}$$

satisfy all the limiting results from Theorem 3. Moreover, $|\mathbb{E}[t(\rho, \Omega_n(\omega))] - t(\rho, \omega)| \leq \frac{2k^2}{n}$.

(3) *We have the concentration inequalities*

$$\mathbb{P}\big[|t(\rho, \Omega_n(\omega)) - \mathbb{E}[t(\rho, \Omega_n(\omega))]| \geq x\big] \leq 2\exp\left(-\frac{nx^2}{9k^2}\right);$$

$$\mathbb{P}\big[|t(\rho, \Omega_n(\omega)) - t(\rho, \omega)| \geq x\big] \leq 4\exp\left(-\frac{nx^2}{9k^2}\right).$$

Proof. The previous discussion proves the first part of the theorem, and the asymptotic results on $Y_n(\rho, \omega)$ follow immediately by Theorem 3. For the bound on $\mathbb{E}[t(\rho, \Omega_n(\omega))] - t(\rho, \omega)$, we use the identity $\mathbb{E}[\Sigma_\rho(\lambda_n(\omega))] = n^{\downarrow k}\, t(\rho, \omega)$, and the expansion

$$p_\rho = \sum_{|\mu| \leq |\rho|} c_{\rho,\mu}\, \Sigma_\mu = \Sigma_\rho + \sum_{|\mu| < |\rho|} c_{\rho,\mu}\, \Sigma_\mu.$$

Thus,

$$|\mathbb{E}[t(\rho, \Omega_n(\omega))] - t(\rho, \omega)| \leq \frac{1}{n^k}\left(\sum_{|\mu| < |\rho|} c_{\rho,\mu}\,|t(\mu, \omega)| + (n^k - n^{\downarrow k})|t(\rho, \omega)|\right)$$

$$\leq \frac{2}{n^k}\left(n^k - n^{\downarrow k}\right) \leq \frac{2k^2}{n}.$$

The first concentration inequality follows then from Proposition 6, and the second concentration inequality is obtained by the following computation. One can assume without loss of generality $x \leq 1$, and then

$$\mathbb{P}\big[|t(\rho, \Omega_n(\omega)) - t(\rho, \omega)| \geq x\big] \leq \mathbb{P}\left[|t(\rho, \Omega_n(\omega)) - \mathbb{E}[t(\rho, \Omega_n(\omega))]| \geq \max\left(0, x - \frac{2k^2}{n}\right)\right]$$

$$\leq 2\exp\left(-\frac{n}{9k^2}\max\left(0, x - \frac{2k^2}{n}\right)^2\right)$$

$$\leq 2\exp\left(-\frac{n}{9k^2}\left(x^2 - \frac{4k^2 x}{n}\right)\right)$$

$$\leq 2\,e^{\frac{4}{9}}\exp\left(-\frac{nx^2}{9k^2}\right) \leq 4\exp\left(-\frac{nx^2}{9k^2}\right). \qquad \square$$

We recover in particular the exponential concentration of the central measures stated in Theorem 12; this result completes the law of large numbers obtained by Kerov and Vershik in [KV81], and the central limit theorem obtained independently by Bufetov and Méliot in [Buf12; Mél12] (see also [Mél17, Section 12.3]).

6. MOD-GAUSSIAN MODULI SPACES

The similarities that we encountered when studying subgraphs in random graphs, patterns in random permutations and Frobenius moments of random partitions makes it tempting to try to formalise this common structure by the following definition.

Definition 30 (Mod-Gaussian moduli space). *Let $\mathcal{M}$ be a compact metrisable space, and $\mathcal{O}_{\mathfrak{M}} = \bigoplus_{k \in \mathbb{N}} \mathcal{O}_{\mathfrak{M},k}$ be a graded algebra. We say that the pair $(\mathcal{M}, \mathcal{O}_{\mathfrak{M}})$ can be endowed with a structure of mod-Gaussian moduli space if the following conditions are satisfied:*

(MS1) *The pair is endowed with a morphism of algebras $\Psi : \mathcal{O}_{\mathfrak{M}} \to \mathscr{C}(\mathcal{M})$ whose image is a dense subalgebra of $\mathscr{C}(\mathcal{M})$. In particular, by Stone–Weierstrass theorem,*
$$\big(m_n \to_{n \to \infty} m \in \mathcal{M}\big) \iff \big(\forall f \in \mathcal{O}, \ \Psi(f)(m_n) \to_{n \to \infty} \Psi(f)(m)\big).$$
In the sequel we simply denote $\Psi(f)(m) = f(m)$.

(MS2) *The space $\mathcal{M}$ is also endowed with a construction of random objects which have a property of asymptotic concentration. Namely, for each parameter $m \in \mathcal{M}$, there is a family of random variables $(M_n(m))_{n \in \mathbb{N}}$ in $\mathcal{M}$, such that $M_n(m) \to_{n \to \infty} m$ in probability.*

(MS3) *For any $k \geq 0$, there exist two linear maps*
$$\kappa_2 : \mathcal{O}_{\mathfrak{M},k} \otimes \mathcal{O}_{\mathfrak{M},k} \to \mathcal{O}_{\mathfrak{M}},$$
$$\kappa_3 : \mathcal{O}_{\mathfrak{M},k} \otimes \mathcal{O}_{\mathfrak{M},k} \otimes \mathcal{O}_{\mathfrak{M},k} \to \mathcal{O}_{\mathfrak{M}};$$
two sequences $N_{n,k} \to \infty$ and $D_{n,k} = o(N_{n,k})$; and a distinguished linear basis $\mathfrak{M}(k)$ of $\mathcal{O}_k$ such that for any $f \in \mathfrak{M}(k)$, the sequence of random variables
$$S_n(f, m) = N_{n,k} \, f(M_n(m))$$
satisfies the hypotheses of the method of cumulants with parameters $(D_{n,k}, N_{n,k}, 1)$, and with limiting variance $\sigma^2(m) = \kappa_2(f,f)(m)$ and third cumulant $L(m) = \kappa_3(f,f,f)(m)$.

Rigorously, a structure of mod-Gaussian moduli space is given by a family
$$(\mathcal{M}, \mathcal{O}_{\mathfrak{M}}; \Psi; (M_n(m))_{m \in \mathcal{M}, \, n \in \mathbb{N}}; \kappa_2, \kappa_3; (\mathfrak{M}(k))_{k \in \mathbb{N}}; (N_{n,k})_{n,k \in \mathbb{N}}, (D_{n,k})_{n,k \in \mathbb{N}}).$$
However, in the sequel, we shall just speak of the pair $(\mathcal{M}, \mathcal{O}_{\mathfrak{M}})$, being understood that this means that there is a canonical (and interesting!) associated family
$$(\Psi; (M_n(m))_{m \in \mathcal{M}, \, n \in \mathbb{N}}; \kappa_2, \kappa_3; (\mathfrak{M}(k))_{k \in \mathbb{N}}; (N_{n,k})_{n,k \in \mathbb{N}}, (D_{n,k})_{n,k \in \mathbb{N}}).$$
All the results of this article can be summarised by the following statement: the pairs $(\mathcal{G}, \mathcal{O}_G)$, $(\mathcal{S}, \mathcal{O}_{\mathfrak{S}})$ and $(\mathcal{P}, \mathcal{O}_{\mathfrak{P}})$ are three mod-Gaussian moduli spaces, each time with $N_{n,k} = n^k$ and $D_{n,k} = k^2 n^{k-1}$. In addition to the concision of this statement, the interest of the notion of mod-Gaussian moduli spaces lies in the following remarks, by which we conclude this paper.

Remark (Moduli spaces and convergence in probability). Given a mod-Gaussian moduli space $(\mathcal{M}, \mathcal{O}_{\mathfrak{M}})$, since $\mathcal{M}$ is metrisable, there is a distance d on $\mathcal{M}$ which corresponds to the topology:
$$\big(m_n \to_{n \to \infty} m \in \mathcal{M}\big) \iff \big(d(m_n, m) \to 0\big).$$
However, in many cases, $\mathcal{M}$ is a part of an infinite-dimensional vector space, and the distances corresponding to its topology are not very practical to deal with. For instance, in the space of graphons, the cut-metric indeed corresponds to the topology of $\mathcal{G}$, but given two graphs G and H with arbitrary sizes, their distance $\delta_\square(G, H)$ is not easy to manipulate or

even compute, in particular if $|G| \neq |H|$ (see for instance [Bor+08, Section 2.3]). The same remark applies to the two other cases of moduli spaces studied in the article, or to the space of probability measures evoked hereafter. This is the reason why it is better to control the topology by an algebra of observables $\mathscr{O}_{\mathfrak{M}}$, in particular for the study of random models that are convergent in probability. Indeed, the convergence in probability of a sequence of random variables $(M_n)_{n \in \mathbb{N}}$ towards m, which is defined by

$$\forall \varepsilon > 0, \ \mathbb{P}[d(M_n, m) \geq \varepsilon] \to_{n \to \infty} 0,$$

can then be characterised by

$$\forall \varepsilon > 0, \ \forall f \in \mathscr{O}_{\mathfrak{M}}, \ \mathbb{P}[|f(M_n) - f(m)| \geq \varepsilon] \to_{n \to \infty} 0,$$

and this second condition is usually much easier to check (assuming that one has chosen an adequate algebra of observables). In the mod-Gaussian framework, we saw that these probabilities $\mathbb{P}[|f(M_n) - f(m)| \geq \varepsilon]$ are exponentially small.

Remark (Classical theorems of probability theory). The framework of mod-Gaussian and mod-ϕ convergent sequences that was developed in particular in [JKN11; DKN15; FMN16; FMN19] can be seen as a far reaching generalisation of the classical theorems of probability theory: law of large numbers, central limit theorem, Cramér's large deviation estimates, Berry–Esseen estimates of the speed of convergence, *etc.* for the scaled sums of i.i.d. random variables. Let us explain how to summarise all these results by constructing an adequate mod-Gaussian moduli space. Let X be a compact metrisable space, and $\mathscr{M}^1(X)$ be the space of Borel probability measures on X, endowed with the topology of convergence in law. It is well known that $\mathscr{M}^1(X)$ is again compact metrisable, see [Bil69, Chapter 1, Section 5]. Let $\mathscr{A}$ be a dense subalgebra of the separable algebra $\mathscr{C}(X)$, endowed with a distinguished countable basis $\mathfrak{A}$ (thus, $\mathscr{A} = \mathrm{Span}_{\mathbb{R}}(\mathfrak{A})$). We can assume without loss of generality that $\|f\|_\infty = 1$ for any $f \in \mathfrak{A}$. A graded algebra of observables for $\mathscr{M}^1(X)$ is the symmetric algebra

$$\mathscr{O}_{\mathfrak{A}} = S(\mathscr{A}) = \bigoplus_{k=0}^{\infty} S^k(\mathscr{A}),$$

where $S^k(\mathscr{A})$ is the quotient of $\mathscr{A}^{\otimes k}$ by the relations $t^\sigma = t$ for any k-tensor t and any permutation $\sigma \in \mathfrak{S}(k)$. We evaluate a tensor $t = f_1 \otimes f_2 \otimes \cdots \otimes f_k$ on a probability measure $\pi \in \mathscr{M}^1(X)$ by

$$t(\pi) = f_1(\pi) \, f_2(\pi) \cdots f_k(\pi) = \int_{X^k} f_1(x_1) \cdots f_k(x_k) \, \pi^{\otimes k}(dx_1 \cdots dx_k).$$

It is easy to see that the convergence in law in $\mathscr{M}^1(X)$ amounts to the convergence of all the observables $t \in \mathscr{O}_{\mathfrak{A}}$. On the other hand, for any $\pi \in \mathscr{M}^1(X)$, a way to approximate π by a discrete "combinatorial" object is by using the empirical measures

$$\Pi_n(\pi) = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i},$$

where the x_i's are independent identically distributed variables with law π. By the classical law of large numbers, $\Pi_n(\pi) \to \pi$ in probability and for any probability measure $\pi \in \mathscr{M}^1(X)$. It is then easy to prove by using the theory of dependency graphs that one has in

fact a structure of mod-Gaussian moduli space on $(\mathscr{M}^1(X), \mathscr{O}_\mathfrak{A})$. This structure is associated with the parameters $N_{n,k} = n^k$ and $D_{n,k} = k^2\, n^{k-1}$ and to the maps

$$\kappa_2(f_1 \otimes \cdots \otimes f_k, g_1 \otimes \cdots \otimes g_k)$$

$$= \frac{1}{k^2} \sum_{1 \le a,b \le k} (f_a g_b - f_a \otimes g_b) \otimes \left(\bigotimes_{a' \ne a} f_{a'} \right) \otimes \left(\bigotimes_{b' \ne b} g_{b'} \right);$$

$$\kappa_3(f_1 \otimes \cdots \otimes f_k, g_1 \otimes \cdots \otimes g_k, h_1 \otimes \cdots \otimes h_k)$$

$$= \frac{1}{k^4} \sum_{1 \le a,b,c \le k} \kappa_3(f_a, g_b, h_c) \otimes \left(\bigotimes_{a' \ne a} f_{a'} \right) \otimes \left(\bigotimes_{b' \ne b} g_{b'} \right) \otimes \left(\bigotimes_{c' \ne c} h_{c'} \right),$$

where for κ_3,

$$\kappa_3(f_a, g_b, h_c) = f_a g_b h_c - (f_a g_b) \otimes h_c - (f_a h_c) \otimes g_b - (g_b h_c) \otimes f_a + 2\, f_a \otimes g_b \otimes h_c.$$

Thus, this basic setting in classical probability theory can be integrated in our Definition 30. We think that many other classes of random models give rise to a mod-Gaussian moduli space, in particular when one tries to approximate an object m by a random sequence $(M_n(m))_{n \in \mathbb{N}}$ of "discrete" objects. For instance, De Catelan and Méliot have proved recently that the space of metric measured spaces introduced in [Gro07, Section $3\frac{1}{2}$] and in [GPW09] gives rise to such a structure, and that the singular points of this mod-Gaussian moduli space (see the next remarks) are exactly the compact homogeneous spaces $X = G/K$; see [CM19].

Remark (Moduli spaces as fields of Gaussian fluctuations). Let us explain the intuition behind our definition of mod-Gaussian moduli space. It is convenient to represent a mod-Gaussian moduli space by a surface as in Figure 13.
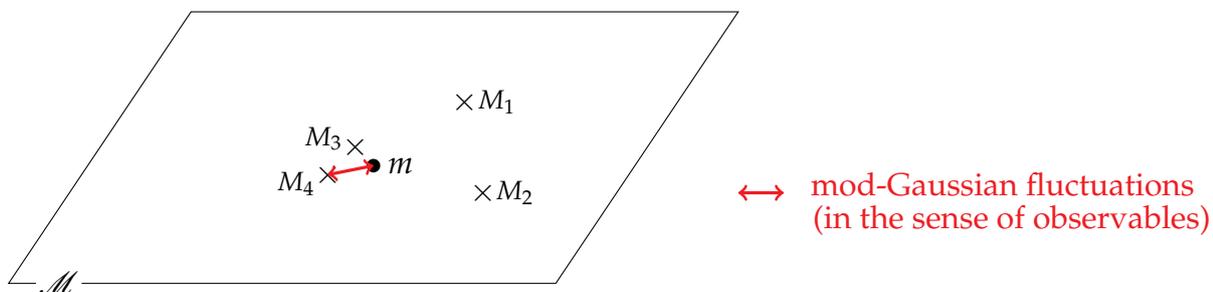

FIGURE 13. Mod-Gaussian moduli spaces as fields of fluctuations.

For any parameter $m \in \mathscr{M}$, the random model $M_n(m)$ can be seen as a canonical way to construct random perturbations $M_n(m)$ of the parameter m. These random objects are perturbations, because by Hypothesis (MS2), as n goes to infinity, $M_n(m)$ converges back to the parameter m. In all the examples that we looked at, each random variable $M_n(m)$ corresponds to a random combinatorial object of size n: graph with n vertices, permutation on n points, integer partition with size n. The parameter m drives then the asymptotics of the random model $(M_n(m))_{n \in \mathbb{N}}$. The *mod-Gaussian* part of the definition amounts to ask that the fluctuations $f(M_n(m)) - \mathbb{E}[f(M_n(m))]$ are mod-Gaussian convergent after an appropriate renormalisation, for any function f in a dense subalgebra $\Psi(\mathscr{O}_\mathfrak{M})$ of $\mathscr{C}(\mathscr{M})$. The exact

renormalisation required may depend on the observable f, but only through a gradation of the algebra of observables $\mathscr{O}_{\mathfrak{M}}$. On the other hand, the asymptotics of the second and third cumulants of the observables can be encoded by simple linear functions $(\mathscr{O}_{\mathfrak{M},k})^{\otimes 2} \to \mathscr{O}$ and $(\mathscr{O}_{\mathfrak{M},k})^{\otimes 3} \to \mathscr{O}$.

So, we can think of a mod-Gaussian moduli space as a space whose elements m have a way to generate random perturbations of themselves, which correspond to random combinatorial models, and which are asymptotically Gaussian, and even mod-Gaussian. The interest of the definition is that if one can show that some space $\mathscr{M}$ is a mod-Gaussian moduli space, then one gets at once numerous precise probabilistic estimates (moderate deviations, bounds of Berry–Esseen type, *etc.*), and this for an extremely large family of random variables.

Remark (Singularities of a mod-Gaussian moduli space). Given a mod-Gaussian moduli space $(\mathscr{M}, \mathscr{O}_{\mathfrak{M}})$, if f is an element of $\mathfrak{M}(k)$ and if $m \in \mathscr{M}$, assuming that $\kappa_2(f, f)(m) > 0$, all the results from Theorem 3 apply to the random variable

$$Y_n(f, m) = \frac{f(M_n(m)) - \mathbb{E}[f(M_n(m))]}{\sqrt{\mathrm{var}(f(M_n(m)))}}.$$

We did not discuss yet the case when the limiting variance $\kappa_2(f, f)(m)$ vanishes. In this situation, we cannot apply Theorem 3, and the generic renormalisation of $f(M_n(m))$ does not yield Gaussian fluctuations (or to be more precise, it yields random variables that converge in probability to 0). We then say that m is a *singular point* or *singularity* of the space $\mathscr{M}$ (with respect to the observable f). The idea is that certain parameters $m \in \mathscr{M}$ correspond to models with smaller variances than usual (again, with respect to some observables f). A typical example is when $\mathscr{M} = \mathscr{G}$ and when γ is the graphon corresponding to a constant graph function $g = p$, with $p \in [0, 1]$. The corresponding random graphs $G_n(p)$ are the Erdös–Rényi random graphs with parameter p. It is easy to see that for any finite graph F,

$$t(F, p) = p^{|E_F|}.$$

Since the number of edges in a product $F \times G$ and in a joined graph $(F \bowtie G)(a, b)$ is the same, we have $\kappa_2(F, F)(p) = 0$ for any graph F. Thus, the graphon p is a *global singularity* of the moduli space of graphons $\mathscr{G}$ (*i.e.*, it is a singularity with respect to any graph observable F). In this particular case, one can show that the variance of $t(F, G_n(p))$ is generically of order $\frac{1}{n^2}$ instead of $\frac{1}{n}$; and one can even show the mod-Gaussian convergence of an appropriate (non-generic) renormalisation of $t(F, G_n(p))$, see [FMN16, Chapter 10].

Another example of a globally singular point is with $\mathscr{M} = \mathscr{P}$ and $\omega = \omega_0 = (0, 0)$. This point corresponds to the celebrated Plancherel measures on integer partitions, and for any integer partition ρ, one has

$$t(\rho, \omega_0) = \begin{cases} 1 & \text{if } \rho = 1^k, \\ 0 & \text{otherwise.} \end{cases}$$

This implies the vanishing of the limiting variance $\kappa_2(\rho, \rho)(\omega_0)$ for any integer partition ρ. Again, in this particular case, the random variables $\Sigma_\rho(\lambda_n(\omega_0))$ can be shown to satisfy a central limit theorem after an appropriate non-generic renormalisation; see *e.g.* [IO02] for the details. It is also conjectured that this singular point corresponds to mod-Gaussian fluctuations, but with different parameters $N_{n,k}$ and $D_{n,k}$ than in the generic case.

Remark (Link with the geometric notion of moduli space). In algebraic geometry, the *moduli spaces* are geometric spaces whose points parametrise (isomorphism classes of) geometric objects of some fixed kind. For instance, the smooth projective complex curves of genus g and with n marked points are parametrised by the moduli space $\mathcal{M}_{g,n}$, which is itself a geometric space of dimension $3g - 3 + n$ (to be precise, $\mathcal{M}_{g,n}$ is an algebraic stack). In particular, the elliptic curves (smooth projective curves of genus 1 with 1 marked point) are classified by the moduli space $\mathcal{M}_{1,1}$, which is also the one-dimensional complex orbifold $\mathrm{SL}(2,\mathbb{Z})\backslash\backslash\mathbb{H}$ (see *e.g.* [Hai08]). This space allows one to consider smooth families of elliptic curves (it is the universal solution of a classification problem), and on the other hand, the singularities of $\mathcal{M}_{1,1}$ correspond to elliptic curves with special symmetries, namely, those associated with the complex lattices with fundamental domains

On the other hand, our *mod-Gaussian moduli spaces* are compact topological spaces whose points m parametrise certain random models $(M_n(m))_{n\in\mathbb{N}}$: random graphs, random permutations, *etc.* All these models have asymptotic properties (central limit theorem, local limit theorem, *etc.*) that vary continuously with the parameter m. Thus, the introduction of the mod-Gaussian moduli space $(\mathcal{M}, \mathcal{O}_{\mathfrak{M}})$ allows one to understand generic properties of these models, and to consider smooth families of such models. Moreover, the singularities of $\mathcal{M}$ correspond to random models with special symmetries that force the variances to be smaller than for the other random models. As evoked in the previous remark, the study of these singularities can be extremely interesting: indeed, when m is a singular parameter, one can try to establish the mod-Gaussian behavior of an adequate non-generic renormalisation of the observables of the models $M_n(m)$. The fact that a parameter $m \in \mathcal{M}$ with additional symmetries can correspond to vanishing variances and to singularities in the fluctuations of the model $(M_n(m))_{n\in\mathbb{N}}$ is to be compared with the fact that a singular point of a moduli space in algebraic geometry usually parametrises a manifold with additional (non-generic) symmetries. This analogy explains why we borrowed the terminology of moduli space from algebraic geometry.

References

[Ais+51]   M. Aissen, A. Edrei, I. J. Schoenberg, and A. Whitney. "On the generating functions of totally positive sequences." In: *Proc. Nat. Acad. Sci. USA* 37.5 (1951), pp. 303–307.

[Bas+18]   F. Bassino, M. Bouvel, V. Féray, L. Gerin, and A. Pierrot. "The Brownian limit of separable permutations". In: *Ann. Probab.* 46.4 (2018), pp. 2134–2189.

[Bil69]   P. Billingsley. *Convergence of Probability Measures*. John Wiley and Sons, 1969.

[BMN19]   M. Dal Borgo, P.-L. Méliot, and A. Nikeghbali. "Local limit theorems and mod-ϕ convergence". In: *Latin American Journal of Probability and Mathematical Statistics* 16.1 (2019), pp. 817–853.

[Bor+06]   C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. "Counting graph homomorphisms". In: *Topics in discrete mathematics*. Vol. 26. Algorithms Combin. Springer-Verlag, 2006, pp. 315–371.

[Bor+08] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. "Convergent sequences of dense graphs I. Subgraph frequencies, metric properties and testing". In: *Adv. Math.* 219.6 (2008), pp. 1801–1851.

[Bor+11] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. "Limits of randomly grown graph sequences". In: *European Journal of Combinatorics* 32 (2011), pp. 985–999.

[Bro65] J. R. Brown. "Doubly stochastic measures and Markov operators". In: *Michigan Math. J.* 12.3 (1965), pp. 367–375.

[Bro66] J. R. Brown. "Approximation theorems for Markov operators". In: *Pacific J. Math.* 16.1 (1966), pp. 13–23.

[Buf12] A. Bufetov. "A central limit theorem for extremal characters of the infinite symmetric group". In: *Functional Analysis and its Applications* 46.2 (2012), pp. 83–93.

[CM19] J. De Catelan and P.-L. Méliot. *Fluctuations of the Gromov–Prohorov sample model*. 2019. arXiv: 1912.00787 [math.PR].

[Cha+07] J.-R. Chazottes, P. Collet, C. Külske, and F. Redig. "Concentration inequalities for random fields via coupling". In: *Probab. Th. Rel. Fields* 137 (2007), pp. 201–225.

[CCR17] J.-R. Chazottes, P. Collet, and F. Redig. "On concentration inequalities and their applications for Gibbs measures in lattice systems". In: *J. Stat. Phys.* 169 (2017), pp. 504–506.

[DKN15] F. Delbaen, E. Kowalski, and A. Nikeghbali. "Mod-ϕ convergence". In: *Int. Math. Res. Notices* 2015 (2015), pp. 3445–3485.

[DIS74] M. Duneau, D. Iagolnitzer, and B. Souillard. "Strong cluster properties for classical systems with finite range interaction". In: *Comm. Math. Phys.* 35 (1974), pp. 307–320.

[ELS79] P. Erdös, L. Lovász, and J. Spencer. "Strong independence of graphcopy functions". In: *Graph Theory and Related Topics*. Academic Press, 1979, pp. 165–172.

[FMN16] V. Féray, P.-L. Méliot, and A. Nikeghbali. *Mod-ϕ convergence: Normality zones and precise deviations*. SpringerBriefs in Probability and Mathematical Statistics. Springer–Verlag, 2016.

[FMN19] V. Féray, P.-L. Méliot, and A. Nikeghbali. "Mod-ϕ convergence, II: Estimates on the speed of convergence". In: *Séminaire de Probabilités L*. Vol. 2252. Lecture Notes in Mathematics. 2019.

[FV17] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: a Concrete Mathematical Introduction*. Cambridge University press, 2017.

[Ful02] J. Fulman. "Applications of Symmetric Functions to Cycle and Increasing Subsequence Structure After Shuffles". In: *Journal of Algebraic Combinatorics* 16 (2002), pp. 165–194.

[Ful04] J. Fulman. "Stein's method and non-reversible Markov chains". In: *Stein's Method: Expository Lectures and Applications*. Ed. by P. Diaconis and S. Holmes. Institute of Mathematical Statistics, 2004, pp. 66–74.

[Gle+15] R. Glebov, A. Grzesik, T. Klimosová, and D. Král'. "Finitely forcible graphons and permutons". In: *Journal of Combinatorial Theory, Series B* 110 (2015), pp. 112–135.

[GPW09] A. Greven, P. Pfaffelhuber, and A. Winter. "Convergence in distribution of random metric measure spaces (The Λ-coalescent measure tree)". In: *Probab. Th. Rel. Fields* 145.1 (2009), pp. 285–322.

[Gro07] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. Modern Birkhäuser Classics. Birkhäuser, 2007.

[Hai08] R. Hain. *Lectures on moduli spaces of elliptic curves*. 2008. arXiv: 0812:1803 [math.AG].

[Hof17] L. Hofer. "A Central Limit Theorem for Vincular Permutation Patterns". preprint arXiv:1704.00650. 2017.

[Hop+13] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Ráth, and R. M. Sampaio. "Limits of permutation sequences". In: *Journal of Combinatorial Theory, Series B* 103.1 (2013), pp. 93–113.

[Hop+11] C. Hoppen, Y. Kohayakawa, C. G. Moreira, and R. M. Sampaio. *Limits of permutation sequences through permutation regularity*. 2011. arXiv: 1106.1663 [math.CO].

[IO02] V. Ivanov and G. Olshanski. "Kerov's central limit theorem for the Plancherel measure on Young diagrams". In: *Symmetric Functions 2001*. Vol. 74. NATO Science Series II. Mathematics, Physics and Chemistry. 2002, pp. 93–151.

[JKN11] J. Jacod, E. Kowalski, and A. Nikeghbali. "Mod-Gaussian convergence: new limit theorems in Probability and Number Theory". In: *Forum Mathematicum* 23.4 (2011), pp. 835–873.

[Jan04] S. Janson. "Large deviations for sums of partly dependent random variables". In: *Random Structures & Algorithms* 24.3 (2004), pp. 234–248.

[JNZ15] S. Janson, B. Nakamura, and D. Zeilberger. "On the asymptotic statistics of the number of occurrences of multiple permutation patterns". In: *J. Combin.* 6.1-2 (2015), pp. 117–143.

[KO94] S. V. Kerov and G. Olshanski. "Polynomial functions on the set of Young diagrams". In: *Comptes Rendus Acad. Sci. Paris Série I* 319 (1994), pp. 121–126.

[KV77] S. V. Kerov and A. M. Vershik. "Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux". In: *Soviet Mathematics Doklady* 18 (1977), pp. 527–531.

[KV81] S. V. Kerov and A. M. Vershik. "Asymptotic theory of characters of the symmetric group". In: *Funct. Anal. Appl.* 15.4 (1981), pp. 246–255.

[Kül03] C. Külske. "Concentration inequalities for functions of Gibbs fields with application to diffraction and random Gibbs measures". In: *Comm. Math. Phys.* 239 (2003), pp. 29–51.

[Lov12] L. Lovász. *Large Networks and Graph Limits*. Vol. 60. Colloquium Publications. American Mathematical Society, Providence, 2012.

[LS06] L. Lovász and B. Szegedy. "Limits of dense graph sequences". In: *Journal of Combinatorial Theory, Series B* 96 (2006), pp. 933–957.

[LS07]   L. Lovász and B. Szegedy. "Szemerédi's lemma for the analyst". In: *Geom. Func. Anal.* 17 (2007), pp. 252–270.

[Mac95]   I. G. Macdonald. *Hall Polynomials and Symmetric Functions*. 2nd. Oxford University Press, 1995.

[Mél12]   P.-L. Méliot. "Fluctuations of central measures on partitions". In: *Proceedings of the 24th International Conference on Formal Power Series and Algebraic Combinatorics (Nagoya, Japan)*. 2012, pp. 387–398.

[Mél17]   P.-L. Méliot. *Representation Theory of Symmetric Groups*. CRC Press, 2017.

[MN15]   P.-L. Méliot and A. Nikeghbali. "Mod-Gaussian convergence and its applications for models of statistical mechanics". In: *In Memoriam Marc Yor - Séminaire de Probabilités XLVII*. Vol. 2137. Lecture Notes in Mathematics. Springer-Verlag, 2015, pp. 369–425.

[Pec59]   J. E. L. Peck. "Doubly stochastic measures". In: *Michigan Math. J.* 6.3 (1959), pp. 217–220.

[Rin94]   Y. Rinott. "On normal approximation rates for certain sums of dependent random variables". In: *J. Comp. Appl. Math.* 55.2 (1994), pp. 135–143.

[Sag01]   B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms and Symmetric Functions*. 2nd. Vol. 203. Graduate Texts in Mathematics. Springer-Verlag, 2001.

[SS91]   L. Saulis and V. A. Statulevičius. *Limit theorems for large deviations*. Vol. 73. Mathematics and its Applications (Soviet Series). Kluwer Academic Publishers, 1991.

[Tho64]   E. Thoma. "Die unzerlegbaren, positive-definiten Klassenfunktionen der abzählbar unendlichen symmetrischen Gruppe". In: *Math. Zeitschrift* 85 (1964), pp. 40–61.

[Var14]   Y. Vargas. "Hopf algebra of permutation pattern functions". In: *DMTCS proceedings* AT (2014), pp. 839–850.

[Ver05]   A. M. Vershik. "Polymorphisms, Markov processes, and quasi-similarity". In: *Discrete and continuous dynamical systems* 13.5 (2005), pp. 1305–1324.

[Was81]   A. J. Wassermann. "Automorphic actions of compact groups on operator algebras". PhD thesis. University of Pennsylvania, 1981.

[Whi32]   H. Whitney. "The coloring of graphs". In: *Ann. Math.* 33 (1932), pp. 688–718.

UNIVERSITÄT ZÜRICH, INSTITUT FÜR MATHEMATIK, WINTERTHURERSTR. 190, CH-8057 ZÜRICH
Email address: valentin.feray@math.uzh.ch, ashkan.nikeghbali@math.uzh.ch

UNIVERSITÉ PARIS-SUD, DÉPARTEMENT DE MATHÉMATIQUES, BÂTIMENT 307, FACULTÉ DES SCIENCES D'ORSAY, F-91405 ORSAY CEDEX
Email address: pierre-loic.meliot@math.u-psud.fr