Check for updates

J. Appl. Probab. **62,** 1406–1428 (2025)
*doi:*10.1017/jpr.2025.13

MODERATE DEVIATIONS FOR WEIGHT-DEPENDENT RANDOM CONNECTION MODELS

NILS HEERTEN,* *Ruhr University Bochum*
CHRISTIAN HIRSCH,** *Aarhus University*
MORITZ OTTO,*** *Leiden University*

Abstract

In this paper we derive cumulant bounds for subgraph counts and power-weighted edge lengths in a class of spatial random networks known as weight-dependent random connection models. These bounds give rise to different probabilistic results, from which we mainly focus on moderate deviations of the respective statistics, but also show a concentration inequality and a normal approximation result. This involves dealing with long-range spatial correlations induced by the profile function and the weight distribution. We start by deriving the bounds for the classical case of a Poisson vertex set, and then provide extensions to α-determinantal processes.

Keywords: Moderate deviations; α-determinantal process

2010 Mathematics Subject Classification: Primary 60F10
Secondary 60G55; 60D05

1. Introduction

In the random geometric graph, the vertex set is a random set of points in Euclidean space, and any two vertices are connected by an edge if their distance is smaller than a fixed threshold. Such a random geometric graph is a model in a wide variety of applications such as wireless networks. Therefore its asymptotic behavior for a large number of nodes has been studied extensively in the literature [16]. However, with the rise of network science, it has become apparent that often this model is not flexible enough to capture more complex structures. Driven by this need, the class of *weight-dependent random connection models* (WRCMs) has recently appeared as a promising approach for describing phenomena in large complex networks [7, 8, 13]. Instead of a fixed connection threshold, the WRCMs quantify the connectivity via a *profile function* and a *weight distribution*. However, due to the complexity of the models, the analysis of their asymptotic behavior is still in its infancy. In this paper we study the question of asymptotic normality of suitable functionals on the WRCM. This is relevant in order to put the analysis of data via WRCMs on a firm statistical foundation.

One of the main tools in the analysis of asymptotic normality is the method of cumulants [3, 19]. Its popularity stems from the level of generality where it is applicable and the variety of implications it draws. Moreover, one obtains specific non-asymptotic bounds for the rate

Received 21 March 2024; accepted 15 January 2025.
* Email address: nils.heerten@rub.de
** Email address: hirsch@math.au.dk
*** Email address: m.f.p.otto@math.leidenuniv.nl

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

of convergence to the normal distribution. However, the challenge in applying the method of cumulants is that it requires the computation of higher-order integrals of the functional under consideration. This is particularly challenging in settings involving long-range spatial correlations. Important progress was recently achieved by Liu and Privault [15] and Schulte and Thäle [20], who derived a general formula for the cumulants of U-statistics based on Poisson processes; see e.g. [20, Theorem 3.2]. This formula relies on the computation of higher-order integrals, which are particularly accessible for random geometric graphs, where the corresponding integrands are bounded and of bounded support. This is no longer true in the WRCM where the range of the support depends on the profile function and the weight distribution.

While Poisson processes are certainly the simplest model for a random point pattern in Euclidean space, they are inappropriate in settings with substantial spatial correlations. In order to overcome this limitation, one considers determinantal and permanental point processes that are based on a spatial kernel function [11]. These classes of point processes are particularly well-suited to cumulant-based methods since their cumulant density can be expressed in closed form in terms of the kernel function [9]. Recently, asymptotic normality for U-statistics of determinantal point processes associated with spectral projection kernels on the unit sphere has been obtained in [5]. One of the key observations in [10] and [12] is that the cumulants of U-statistics can again be bounded in terms of higher-order integrals. These integrals are reminiscent of those appearing in [20, Theorem 3.2] but are more complex due to the presence of the kernel function. We show that the results in the Poisson case can be extended to this more complex setting at the cost of weaker bounds on the convergence rate. In this context we provide a lower bound on the variance of the functional under consideration, which may be of independent interest.

The main contributions of this paper are as follows.

(i) Show asymptotic normality for power-weighted edge lengths and subgraph counts in the WRCM. One of the key achievements of our result is that our bounds are completely transparent in the moment assumptions on the weight distribution and the profile function.

(ii) While [20] focuses on the case of Poisson processes, we extend the results to α-determinantal processes. This relies on expansions derived in [9], [10], and [12].

(iii) Explain how variance lower bounds can be established for the considered functionals.

The rest of the manuscript is organized as follows. In Section 2 we introduce the model and state our main results. Then we discuss certain preliminary results concerning general cumulant computations in Section 5. Sections 6, 7, and 8 contain the proofs of the main results. Finally, Section 4 contains some considerations regarding the lower variance bounds.

2. Model and main results

The main goal of this work is to establish moderate deviations in the setting of WRCM. To do so, we will rely on the so-called method of cumulants (see Section 5.2 below for more details). One of the core strengths of this method is that it gives rise not only to moderate deviations but to a variety of different limit results as elaborated in [3], and [19]. To make our presentation self-contained, we reproduce these consequences here.

Let $(X_n)_{n\geq 1}$ be a sequence of square-integrable random variables, let $\gamma \geq 0$ be a constant, and let $(\Delta_n)_{n\geq 1}$ be a positive real-valued sequence. We say that $(X_n)_{n\geq 1}$ satisfies the following.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

- **MDP** $(\gamma, (\Delta_n)_{n\geq 1})$ if for any positive real-valued sequence $(a_n)_{n\geq 1}$ with

$$\lim_{n\to\infty} a_n = \infty \quad \text{and} \quad \lim_{n\to\infty} \frac{a_n}{\Delta_n^{1/(1+2\gamma)}} = 0$$

the rescaled random variables

$$(\widehat{X}_n)_{n\geq 1} := (a_n^{-1}(X_n - \mathbb{E}[X_n])/\sqrt{\mathrm{Var}(X_n)})_{n\geq 1}$$

satisfy a *moderate deviation principle* (MDP) with speed a_n^2 and good rate function $\mathcal{I}(z) = z^2/2$. That is, for all Borel sets A,

$$\begin{aligned} -\inf_{z\in \mathrm{int}(A)} \mathcal{I}(z) &\leq \liminf_{n\to\infty} a_n^{-2} \log \mathbb{P}(\widehat{X}_n \in A) \\ &\leq \limsup_{n\to\infty} a_n^{-2} \log \mathbb{P}(\widehat{X}_n \in A) \\ &\leq -\inf_{z\in \mathrm{cl}(A)} \mathcal{I}(z), \end{aligned}$$

where $\mathrm{int}(A)$ and $\mathrm{cl}(A)$, respectively, stand for the interior and the closure of A [2]. This might look quite familiar to a large deviation principle. However, since in our setting the speed a_n is of order smaller than $\sqrt{\mathrm{Var}(X_n)}$, the term MDP is more appropriate.

- **CI** $(\gamma, (\Delta_n)_{n\geq 1})$ if the *Bernstein-type concentration inequality*

$$\mathbb{P}\big(|X_n - \mathbb{E}[X_n]| \geq z\sqrt{\mathrm{Var}(X_n)}\big) \leq 2\exp\left(-\frac{1}{4}\min\left\{\frac{z^2}{2^{1+\gamma}}, (z\Delta_n)^{1/(1+\gamma)}\right\}\right)$$

holds for all $n \geq 1$ and $z \geq 0$,

- **NACC** $(\gamma, (\Delta_n)_{n\geq 1})$ if a *normal approximation bound with Crámer correction* holds, that is, if there exist constants $c_0, c_1, c_2 > 0$ only depending on γ such that, for all $n \geq 1$ and $z \in \left[0, c_0\Delta_n^{1/(1+2\gamma)}\right]$,

$$\mathbb{P}\big(|X_n - \mathbb{E}[X_n]| \geq z\sqrt{\mathrm{Var}(X_n)}\big) = \mathrm{e}^{L^+_{n,z}}(1-\Phi(z))\left(1 + c_1\theta^+_{n,z}\frac{1+z}{\Delta_n^{1/(1+2\gamma)}}\right)$$

and

$$\mathbb{P}\big(|X_n - \mathbb{E}[X_n]| \leq -z\sqrt{\mathrm{Var}(X_n)}\big) = \mathrm{e}^{L^-_{n,z}}(1-\Phi(z))\left(1 + c_1\theta^-_{n,z}\frac{1+z}{\Delta_n^{1/(1+2\gamma)}}\right)$$

with $\theta^+_{n,z}, \theta^-_{n,z} \in [-1, 1]$ and

$$L^+_{n,z}, L^-_{n,z} \in \big(-c_2 z^3/\Delta_n^{1/(1+2\gamma)}, c_2 z^3/\Delta_n^{1/(1+2\gamma)}\big),$$

where Φ is the distribution function of a standard normal random variable.

We now provide a formal definition of the WRCM considered in this paper; see [7], [8], and [13]. Let $\mathcal{P}_n := \{P_i\}_{i\geq 1} := \{(X_i, U_i)\}_{i\geq 1}$ be a Poisson point process on the space $W \times [0, \infty)$ with intensity measure $t_n \mathrm{d}x \otimes \mathbb{P}_U(\mathrm{d}u)$ for some compact set $W \subseteq \mathbb{R}^d$ with non-empty interior, some $t_n > 0$, and $\mathbb{P}_U$ is a probability distribution. Then, in the WRCM, between any pair of points P_i, P_j we place an edge independently with probability

$$\varphi_n(X_i - X_j, U_i, U_j) := \varphi\left(\frac{|B_{\|X_i - X_j\|}|}{\nu_n \kappa(U_i, U_j)}\right),$$

where we now explain the individual components in detail.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

(i) $\varphi\colon [0,\infty)\to[0,1]$ is the *profile function*, which is non-increasing and for which we assume the normalization $\int \varphi(t)\,\mathrm{d}t = 1$.

(ii) $|B_{\|X_i - X_j\|}|$ is the volume of the Euclidean ball of radius $\|X_i - X_j\|$ in $\mathbb{R}^d$. In the literature, the scaling $\|X_i - X_j\|^d$ is sometimes used. Since the difference is only in a constant multiple, the two parameterizations are equivalent. For our analysis, the parameterization with the ball-based scaling will substantially simplify the statements and proofs of our main results.

(iii) $(\nu_n)_{n\geq 1}$ is a positive real-valued sequence that is bounded from above. It can be thought of as the volume of a ball around a typical node within which most neighbors of this node are located.

(iv) Finally, $\kappa\colon [0,\infty)\times[0,\infty)\to[0,\infty)$ is the *interaction kernel*, which is non-decreasing in both arguments. We assume that κ is such that $1\leq\kappa(u,v)\leq uv$ for all $u,v\geq 0$.

In a broad sense, our cumulant bounds are similar to those considered earlier in the setting of standard random geometric graphs [20]. However, in the setting of WRCM the convergence will depend sensitively on moment estimates for both the profile function and the weight distribution. To make this precise, we now introduce specific conditions on these quantities.

We assume that $U\geq 1$ a.s. and that there are constants $c_{U,1}>0$ and $c_{U,2}\leq 1$ such that for all $x\geq 1$,

$$M_U(x) := \mathbb{E}[U^x] \leq c_{U,1}^x \Gamma(1+x)^{c_{U,2}}, \tag{MU}$$

where

$$\Gamma(y) := \int_0^\infty t^{y-1}\mathrm{e}^{-t}\,\mathrm{d}t$$

is the standard Gamma function. For instance, if X is exponentially distributed with some parameter $\lambda>0$, then $U := X+1$ satisfies (MU) with $c_{U,2}=1$. Moreover, if X is a standard normal random variable and $U\sim X\mid\{X>1\}$, then (MU) holds with $c_{U,2}=1/2$. As we will see in Theorem 2.1 below, the value of the constant $c_{U,2}$ is much more important than that of $c_{U,1}$.

The first functional of interest is the subgraph count, which was also considered in [20], and we recall the definition here. Let G be a fixed connected graph on the vertices $\{1,\dots,q\}$. Then we let $S_n(G)$ denote $1/q!$ times the number of q-tuples $(P_1,\dots,P_q)\in\mathcal{P}_n^q$ such that for every edge $\{i,j\}$ in G, the WRCM has an edge between P_i and P_j.

Theorem 2.1. (Cumulant bounds for subgraph count.) *Let $d\geq 1$ and G be a fixed connected graph on q vertices. Assume that $U\geq 1$ a.s. and that condition* (MU) *holds, and set $A_{\mathrm{SG}} := (q-1)(1+c_{U,2})$. Moreover, assume that there exists $v>0$ such that for all $n\geq 1$,*

$$\mathrm{Var}(S_n(G)) \geq v|W|t_n^q\nu_n^{q-1}(1\vee t_n\nu_n)^{q-1}. \tag{VSG}$$

Then $S_n(G)$ satisfies **MDP** $(A_{\mathrm{SG}},(\beta_n)_n)$, **CI** $(A_{\mathrm{SG}},(\beta_n)_n)$, *and* **NACC** $(A_{\mathrm{SG}},(\beta_n)_n)$ *with*

$$\beta_n := \frac{\sqrt{v|W|t_n(1\wedge t_n\nu_n)^{q-1}}}{q^{3q}\big(b_{\mathrm{SG}}\vee b_{\mathrm{SG}}^3/v\big)}, \quad \text{where } b_{\mathrm{SG}} := ((1\vee c_{U,1}^2)(q-1)^{c_{U,2}})^{q-1},\ n\geq 1,$$

Second, for fixed $\tau\geq 0$, we consider the τ-power-weighted edge length given by

$$S_n^{(\tau)} := \frac{1}{2}\sum_{P_i\sim P_j\in\mathcal{P}_n}\|X_i - X_j\|^\tau,$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

where $P_i \sim P_j$ indicates the existence of an edge in the WRCM. Here we additionally need to impose similar assumptions concerning φ as we did for the weights in (MU). More precisely, we assume that there are constants $c_{\varphi,1} > 0$ and $c_{\varphi,2} \leq 1$ such that for all $x \geq 0$,

$$\begin{aligned} M_\varphi(x) &:= \int_0^\infty \varphi(t) t^x \, \mathrm{d}t \leq c_{\varphi,1}^x \Gamma(1+x)^{c_{\varphi,2}} \quad \text{and} \\ M'_\varphi(x) &:= \sup_{t>0} \varphi(t) t^x \leq c_{\varphi,1}^x \Gamma(1+x)^{c_{\varphi,2}}. \end{aligned} \tag{MPHI}$$

For instance, (MPHI) holds if φ has compact support. Moreover, similarly as for the case of weights, one can also consider the cases of exponential tails where $c_{\varphi,2} = 1$ and Gaussian tails where $c_{\varphi,2} = 1/2$.

Theorem 2.2. (Cumulant bounds for power-weighted edge lengths.) *Let $d \geq 1$ and $\tau \geq 0$. Assume that $U \geq 1$ a.s. and that conditions* (MU) *and* (MPHI) *hold, and set $A_{\mathrm{EL}} := 1 + 2c_{U,2}(\tau_d + 1) + c_{\varphi,2}\tau_d$, where $\tau_d := \tau/d$. Moreover, assume that there exists $v > 0$ such that for all $n \geq 1$,*

$$\mathrm{Var}(S_n^{(\tau)}) \geq v|W|t_n^2 v_n^{2\tau_d+1}(1 \vee t_n v_n). \tag{VEL}$$

Then the power-weighted edge length $S_n^{(\tau)}$ satisfies **MDP** $(A_{\mathrm{EL}}, (\beta_n)_n)$, **CI** $(A_{\mathrm{EL}}, (\beta_n)_n)$, *and* **NACC** $(A_{\mathrm{EL}}, (\beta_n)_n)$ *with*

$$\beta_n := \frac{\sqrt{v|W|t_n(1 \wedge t_n v_n)}}{64\big(b_{\mathrm{EL}} \vee b_{\mathrm{EL}}^3/v\big)}, \quad \text{where } b_{\mathrm{EL}} := (c_{\varphi,1}(\tau_d \vee 1)^{c_{\varphi,2}})^{\tau_d}(c_{U,1}(\tau_d+1)^{c_{U,2}})^{2(\tau_d+1)}$$

for $n \geq 1$.

Note that for $U = 1$ a.s. and $\varphi(\cdot) = \mathbf{1}_{[0,1]}$, the weight-dependent random connection model reduces to the classical random geometric graph. In this situation, [20, Section 4.2] establishes the analogue of Theorem 2.1. Since we can choose $a = c_{U,1} = 1$, $c_{U,2} = 0$ in this case, Theorem 2.1 generalizes [20, Corollary 4.2] to the weight-dependent situation.

We also note that for many real-world datasets one can encounter the setting where the weights only have a finite number of moments. In such situations our results do not apply since the condition (MU) is violated.

Finally, in Theorems 2.3 and 2.4 below, we replace the Poisson point process with an α-determinantal point process (α-DPP) as input. We note that while a homogeneous Poisson point process is a special case of an α-DPP, the convergence rates derived for α-DPPs are less sharp than that for Poisson input. Hence the results in Theorems 2.3 and 2.4 do not supersede those in Theorems 2.1 and 2.2.

Recall that for $\alpha \in \mathbb{R}$ a simple point process is called *α-determinantal point process* (α-DPP) with *covariance kernel* $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ if, for all $k \geq 1$, its kth product density ρ_k is given by

$$\rho_k(x_1, \ldots, x_k) = \sum_{\pi \in \mathrm{Per}(k)} \alpha^{k-n(\pi)} \prod_{i=1}^k K(x_i, x_{\pi(i)}) \tag{2.1}$$

for all $x_1, \ldots, x_k \in \mathbb{R}^d$, where $\mathrm{Per}(k)$ is the set of all permutations of $\{1, \ldots, k\}$ and $n(\pi)$ is the number of cycles in the permutation π. Important special cases are *determinantal processes* ($\alpha = -1$) and *permanental processes* ($\alpha = 1$). Following [9, Section 3],

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

we assume throughout that the kernel is stationary, i.e. $K(x, y) = K(o, y - x) = K_0(y - x)$ for all $x, y \in \mathbb{R}^d$, where $K_0(x)$ is a continuous, complex-valued, positive semi-definite function, i.e. $\sum_{i,j=1}^n z_i K_0(x_i - x_j)\overline{z_j} \geq 0$ for all $x_1, \dots, x_n \in \mathbb{R}^d$, $z_1, \dots, z_n \in \mathbb{C}$. This implies that $K_0(o) \geq 0$, $K_0(-x) = \overline{K_0(x)}$, $|K_0(x)| \leq K_0(o)$ for all $x \in \mathbb{R}^d$ and that the α-DPP with kernel K is stationary. Moreover, we assume that

$$K_0(o) \in [0, 1] \quad \text{and} \quad \|K_0\|_1 := \int_{\mathbb{R}^d} |K_0(x)| \, \mathrm{d}x < \infty. \tag{DPP}$$

Under these assumptions, it is shown in [9, Theorem 2] that the α-DPP with kernel K is Brillinger-mixing, which says that for all $k \geq 1$, its kth cumulant density is absolutely integrable.

Since there is no natural analogue to an increasing intensity parameter for α-DPPs, we will instead consider the scenario of an independent $\mathbb{P}_U$-marking $\mathcal{P}$ of an α-DPP on $W_n := n^{1/d}W$. This is as in [23]. Note that due to the scaling invariance property of a Poisson process, Theorems 2.1 and 2.2 could also be formulated in this setup.

Theorem 2.3. (Cumulant bounds for subgraph count for α-DPPs.) *Let $d \geq 1$ and let $\mathcal{P}$ be an independent $\mathbb{P}_U$-marking of an α-DPP for some $\alpha \in [-1, 1]$. Assume that the stationary covariance kernel K satisfies the conditions* (DPP)*. Let G be a fixed connected graph with q vertices and consider the subgraph count $S_n(G)$. Assume that $U \geq 1$ a.s. and that condition* (MU) *holds and set $A_{\mathrm{SG}'} := 2q + (q-1)c_{U,2} - 1$. Moreover, assume that there exists $v > 0$ such that for all $n \geq 1$,*

$$\mathrm{Var}(S_n(G)) \geq v|W_n|v_n^{q-1}(1 \vee v_n^{q-1}). \tag{VSG$'$}$$

Then $S_n(G)$ satisfies **MDP** $(A_{\mathrm{SG}'}, (\beta_n)_n)$, **CI** $(A_{\mathrm{SG}'}, (\beta_n)_n)$, *and* **NACC** $(A_{\mathrm{SG}'}, (\beta_n)_n)$ *with*

$$\beta_n := \frac{\sqrt{v|W_n|(1 \wedge v_n)^{q-1}}}{(2q^2)^{3q}\big(b_{\mathrm{SG}'} \vee b_{\mathrm{SG}'}^3/v\big)}, \quad \text{where } b_{\mathrm{SG}'} := (\|K_0\|_1 \vee 1)^{q-1} b_{\mathrm{SG}}, \ n \geq 1,$$

with b_{SG} from Theorem 2.1.

The next result is the analogue of Theorem 2.2 for α-determinantal input.

Theorem 2.4. (Cumulant bounds for power-weighted edge lengths for α-DPPs.) *Let $d \geq 1$ and let $\mathcal{P}$ be an independent $\mathbb{P}_U$-marking of an α-DPP for some $\alpha \in [-1, 1]$. Assume that the stationary covariance kernel K satisfies the conditions* (DPP)*. Consider the τ-power-weighted edge length $S_n^{(\tau)}$. Assume that $U \geq 1$ a.s. and that conditions* (MU) *and* (MPHI) *hold, and set $A_{\mathrm{EL}'} := 3 + c_{\varphi,2}\tau_d + 2c_{U,2}(\tau_d + 1)$, where $\tau_d := \tau/d$. Moreover, assume that there exists $v > 0$ such that for all $n \geq 1$,*

$$\mathrm{Var}(S_n^{(\tau)}) \geq v|W_n|v_n^{2\tau_d+1}(1 \vee v_n). \tag{VEL$'$}$$

Then the power-weighted edge lengths $S_n^{(\tau)}$ satisfy **MDP** $(A_{\mathrm{EL}'}, (\beta_n)_n)$, **CI** $(A_{\mathrm{EL}'}, (\beta_n)_n)$, *and* **NACC** $(A_{\mathrm{EL}'}, (\beta_n)_n)$ *with*

$$\beta_n := \frac{\sqrt{v|W_n|(1 \wedge v_n)}}{64\big(b_{\mathrm{EL}'} \vee b_{\mathrm{EL}'}^3/v\big)}, \quad \text{where } b_{\mathrm{EL}'} := (\|K_0\|_1 \vee 1) b_{\mathrm{EL}}, \ n \geq 1,$$

with b_{EL} from Theorem 2.2.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

Remark 2.1. We stress that the existence of an α-DPP with a given covariance kernel K is not guaranteed. Hence we say that some $\alpha \in \mathbb{R}$ is *admissible* if there exists a point process whose k product intensity is given by (2.1). By [21, Theorem 1.2], it is known that all $\alpha \in \{\pm 1/m\colon m \in \mathbb{N}\}$ are admissible.

Remark 2.2. In Propositions 4.1 and 4.3 below we show that the variance lower bounds assumed in Theorems 2.1 and 2.2 hold for an appropriate sequence $(v_n)_{n\geq 1}$ and give a possible choice for the constant $v > 0$. Note that variance lower bounds are also required in Theorems 2.3 and 2.4 for $\tau \geq 0$. Here, due to the different scenario of an increasing sampling window, we rather work with a sequence where $(v_n/n)_{n\geq 1}$ is bounded. In Remark 4.1 we address the problem of obtaining variance lower bounds for U-statistics of α-determinantal processes if $\alpha < 0$. It is interesting that for U-statistics of determinantal point processes associated with spectral projection kernels, the symmetry of the sphere can be used to derive explicit variance asymptotics (see [5, Theorem 3]).

Remark 2.3. Our techniques also allow us to obtain Berry–Esseen bounds for $S_n(G)$ and $S_n^{(\tau)}$. For details we refer to [3, Theorem 2.4].

3. Applications

In this section we explain how our results apply to well-known random geometric graphs and make the constants in the theorems above more explicit. In the following, let $W \subset \mathbb{R}^d$ be a compact set with non-empty interior.

Example 1. (*Total edge length in the Gilbert graph.*) For $\kappa \equiv 1$ and $\varphi(\cdot) := \mathbf{1}\{\cdot \leq 1\}$ our model reduces to the Gilbert graph, in which two points X_i and X_j are connected whenever their Euclidean distance is at most $(v_n/\kappa_d)^{1/d}$. Then (MU) and (MPHI) are satisfied with $c_{U,1} = c_{\varphi,1} = 1$ and $c_{U,2} = c_{\varphi,2} = 0$. Let $\tau := 1$. Then $S_n^{(1)}$ is the total edge length in the random geometric graph built on the point process $\mathcal{P}$. From Proposition 4.1 below it follows that conditions (VEL) and (VEL′) hold for some $v > 0$ (see Remark 2.2).

If $\mathcal{P}$ is a Poisson process with intensity measure $t_n \mathrm{d}x \otimes \mathbb{P}_U(\mathrm{d}u)$ on $W \times [1, \infty)$, Theorem 2.2 gives **MDP**/**CI**/**NACC** $(1, (\beta_n)_n)$ with

$$\beta_n = \frac{\sqrt{v|W|t_n(1 \wedge t_n v_n)}}{64(1 \vee 1/v)}, \quad n \in \mathbb{N}.$$

Next, let $\mathcal{P}$ be an independent $\mathbb{P}_U$-marking of an α-DPP with $\alpha \in [0, 1]$ on $n^{1/d}W \times [1, \infty)$, where the stationary covariance kernel satisfies condition (DPP). Then Theorem 2.4 yields **MDP**/**CI**/**NACC** $(3, (\beta_n)_n)$ with

$$\beta_n = \frac{\sqrt{vn|W|(1 \wedge v_n)}}{64 \max(1, 1/v, \|K_0\|_1, \|K_0\|_1^3/v)}, \quad n \in \mathbb{N}.$$

Example 3.2. (*Triangle count in the Boolean model.*) Let $\kappa(U_i, U_j) = (U_i^{1/d} + U_j^{1/d})^d$ and $\varphi(\cdot) := \mathbf{1}\{\cdot \leq 1\}$. Then we are in the situation of a Boolean model, where X_i and X_j are connected if the balls with volumes $v_n U_i$ and $v_n U_j$ centered at X_i and X_j, respectively, have non-empty intersection. Let $\lambda > 0$ and assume that $\mathbb{P}_U$ has the probability density $\lambda \mathrm{e}^{-\lambda(u-2)}\mathbf{1}_{[2,\infty)}(u)$. Then assumption (MU) holds with $c_{U,1} = \lambda \mathrm{e}^{2\lambda}$ and $c_{U,2} = 1$. From Proposition 4.3 it follows that conditions (VSG) and (VSG′) are satisfied for some $v > 0$.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

If $\mathcal{P}$ is a Poisson process with intensity measure $t_n \mathrm{d}x \otimes \mathbb{P}_U(\mathrm{d}u)$ on $W \times [1, \infty)$, Theorem 2.1 gives **MDP / CI / NACC** $(4, (\beta_n)_n)$ with

$$\beta_n := \frac{\sqrt{v|W|t_n(1 \wedge t_n v_n)^2}}{3^9\big(b_{\mathrm{SG}} \vee b_{\mathrm{SG}}^3/v\big)}, \quad \text{where } b_{\mathrm{SG}} := 4(1 \vee \lambda^4 \mathrm{e}^{8\lambda}),\ n \geq 1.$$

Next, assume that $\mathcal{P}$ is an independent $\mathbb{P}_U$-marking of an α-DPP with $\alpha \in [0, 1]$ on $n^{1/d}W \times [1, \infty)$, where the stationary covariance kernel satisfies condition (DPP). Then Theorem 2.3 yields **MDP / CI / NACC** $(7, (\beta_n)_n)$ with

$$\beta_n := \frac{\sqrt{vn|W|(1 \wedge v_n)^2}}{18^9\big(b_{\mathrm{SG}'} \vee b_{\mathrm{SG}'}^3/v\big)}, \quad \text{where } b_{\mathrm{SG}'} := 4(1 \vee \lambda^4 \mathrm{e}^{8\lambda})(\|K_0\|_1 \vee 1)^2,\ n \geq 1.$$

4. Variance lower bounds

All of our main results depend on lower bounds for the variance of $S_n(G)$ and $S_n^{(\tau)}$, respectively. Before proving our results in the upcoming sections, we elaborate on how these lower bounds can be established. Although the following propositions are stated for the scenario of a fixed point process restricted to a growing sampling window W_n, they hold analogously in the case of a Poisson process with a growing intensity t_n on a fixed sampling window W.

For some $a > 0$, let

$$W_n^{-a} := \{x \in W_n \colon B_{(a/|B_1|)^{1/d}}(x) \subseteq W_n\}$$

be the set of all $x \in W_n$ such that a ball centered at x with volume a is still contained entirely in W_n.

We start with the lower bound on $\mathrm{Var}(S_n^{(\tau d)})$.

Proposition 4.1. (Variance lower bounds for power-weighted edge count.) *For $\alpha \in [0, 1]$, let $\mathcal{P}$ be a stationary α-DPP. Let $\varepsilon > 0$ and $\tau \geq 0$. Then, for $n \geq 1$,*

$$\mathrm{Var}(S_n^{(\tau d)}) \geq v|W_n|v_n^{2\tau+1}(1 \vee v_n),$$

where

$$v := \inf_{n \geq 1} \frac{|W_n^{-\varepsilon v_n}|}{|W_n|} |B_1|^{-2\tau} K_0(o)^2 \int_0^{\varepsilon} t^{2\tau+1}\varphi(t)\,\mathrm{d}t \left(1 \wedge 2K_0(o) \int_0^{\varepsilon} t^{2\tau+1}\varphi(t)\,\mathrm{d}t\right).$$

Note that if the sequence $(v_n/n)_{n \geq 1}$ is bounded, we have $v > 0$. Indeed, our assumptions (that $\int_0^{\infty} \varphi(t)\,\mathrm{d}t = 1$ and that φ is non-increasing) guarantee that there exists some $\varepsilon > 0$ such that $\varphi(\varepsilon) > 0$ and

$$\inf_{n \geq 1} \frac{|W_n^{-\varepsilon v_n}|}{|W_n|} > 0.$$

Proposition 4.2. (Proof of Proposition 4.1.) *Given an α-DPP $\mathcal{P}$, recall that its kth product density ρ_k is given as in (2.1). Then, expanding the second moment of $S_n^{(\tau d)}$ into three terms,*

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

we find that the variance of $S_n^{(\tau d)}$ *is given by*

$$\begin{aligned}
|B_1|^{2\tau}\operatorname{Var}(S_n^{(\tau d)}) &= |B_1|^{2\tau}\left(\mathbb{E}\left[\left(S_n^{(\tau d)}\right)^2\right] - \mathbb{E}\left[S_n^{(\tau d)}\right]^2\right)\\
&= \int_{W_n^2} |B_{\|x-y\|}|^{2\tau}\mathbb{E}[\varphi_n(x-y,U,V)]\rho_2(x,y)\,\mathrm{d}(x,y)\\
&\quad + 2\int_{W_n^3} |B_{\|x-y\|}|^{\tau}|B_{\|x-y'\|}|^{\tau}\mathbb{E}[\varphi_n(x-y,U,V)\varphi_n(x-y',U,V')]\\
&\qquad \times \rho_3(x,y,y')\,\mathrm{d}(x,y,y')\\
&\quad + \int_{W_n^4} |B_{\|x-y\|}|^{\tau}|B_{\|x'-y'\|}|^{\tau}\mathbb{E}[\varphi_n(x-y,U,V)\varphi_n(x'-y',U',V')]\\
&\qquad \times \rho_4(x,x',y,y'))\,\mathrm{d}(x,y,x',y')\\
&\quad - \int_{W_n^4} |B_{\|x-y\|}|^{\tau}|B_{\|x'-y'\|}|^{\tau}\mathbb{E}[\varphi_n(x-y,U,V)\varphi_n(x'-y',U',V')]\\
&\qquad \times \rho_2(x,y)\rho_2(x',y')\,\mathrm{d}(x,y,x',y'),
\end{aligned}$$

where $U, U', V, V' \sim \mathbb{P}_U$ *are independent. Since for an* α*-DPP with* $\alpha \geq 0$ *all terms on the right-hand side in* (2.1) *are non-negative, we have* $\rho_4(x,x',y,y') \geq \rho_2(x,y)\rho_2(x',y')$ *for all* $x, x', y, y' \in \mathbb{R}^d$*. Therefore,*

$$\begin{aligned}
|B_1|^{2\tau}\operatorname{Var}(S_n^{(\tau d)}) &\geq \int_{W_n^2} |B_{\|x-y\|}|^{2\tau}\mathbb{E}[\varphi_n(x-y,U,V)]\rho_2(x,y)\,\mathrm{d}(x,y)\\
&\quad + 2\int_{W_n^3} |B_{\|x-y\|}|^{\tau}|B_{\|x-y'\|}|^{\tau}\mathbb{E}[\varphi_n(x-y,U,V)\varphi_n(x-y',U,V')]\\
&\qquad \times \rho_3(x,y,y')\,\mathrm{d}(x,y,y').
\end{aligned}$$

We now give lower bounds for the two terms separately, starting with the first one. We use spherical coordinates and exploit that $\kappa(u,v) \geq 1$*. Since* $\rho_2(x,y) \geq K_0(o)^2$ *for all* $x, y \in \mathbb{R}^d$*, this gives*

$$\int_{W_n^2} |B_{\|x-y\|}|^{2\tau}\mathbb{E}[\varphi_n(x-y,U,V)]\rho_2(x,y)\,\mathrm{d}(x,y) \geq K_0(o)^2|W_n^{-\varepsilon\nu_n}|\nu_n^{2\tau+1}\int_0^{\varepsilon} t^{2\tau+1}\varphi(t)\,\mathrm{d}t.$$

Similarly, for the second integral we obtain the lower bound

$$K_0(o)^3|W_n^{-\varepsilon\nu_n}|\nu_n^{2(\tau+1)}\left(\int_0^{\varepsilon} t^{2\tau+1}\varphi(t)\,\mathrm{d}t\right)^2.$$

Hence combining the two bounds yields the assertion.

Next we consider the variance of the subgraph count $S_n(G)$.

Proposition 4.3. (Variance lower bounds for the subgraph count.) *For* $\alpha \in [0,1]$*, let* $\mathcal{P}$ *be a stationary* α*-DPP, and let* G *be a fixed connected graph on* q *vertices. Let* $\varepsilon > 0$*. Then, for* $n \geq 1$*,*

$$\operatorname{Var}(S_n(G)) \geq v|W_n|\nu_n^{q-1}(1 \vee \nu_n^{q-1}),$$

where

$$v := \inf_{n\geq 1}\frac{|W_n^{-\varepsilon\nu_n/2^d}|}{|W_n|}\varphi(\varepsilon)^{2\binom{q}{2}}K_0(o)(\varepsilon K_0(o)/2^d)^{q-1}(1 \wedge q(\varepsilon K_0(o)/2^d)^{q-1}).$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

Proof. Let G be a fixed connected graph on vertices $1, \dots, q$. For $s_i \in \mathbb{X} := W \times [1, \infty) \times [0, 1]^{\mathbb{N}}$, $i = 1, \dots, q$, let $f_n(s_1, \dots, s_q)$ be $1/q!$ times the number of subgraphs in the WRCM built on these points that are isomorphic to G (see the proof of Theorem 2.1). Moreover, for $k \in \mathbb{N}$, let ρ_ℓ be the kth product density of $\mathcal{P}$. Then we have

$$\begin{aligned}\mathrm{Var}(S_n(G)) = \sum_{k=0}^{q} \binom{q}{k} \int_{\mathbb{X}^q} \int_{\mathbb{X}^k} f_n(\mathbf{s}) f_n(\mathbf{s}_{q-k}, \mathbf{t}) \rho_{q+k}(\mathbf{x}, \mathbf{y}) \mu_n^k(\mathrm{d}\mathbf{t}) \mu_n^q(\mathrm{d}\mathbf{s}) \\ - \int_{\mathbb{X}^q} \int_{\mathbb{X}^q} f_n(\mathbf{s}) f_n(\mathbf{t}) \rho_q(\mathbf{x}) \rho_q(\mathbf{y}) \mu_n^q(\mathrm{d}\mathbf{t}) \mu_n^q(\mathrm{d}\mathbf{s}), \end{aligned} \tag{4.1}$$

where $\mathbf{s} := (s_1, \dots, s_q)$ and x_i is the first coordinate of s_i for $i = 1, \dots, q$, and $\mathbf{s}_{q-k} := (s_1, \dots, s_{q-k})$ (similarly for $\mathbf{t}$ and $\mathbf{y}$). Accordingly, we formally put $f_n(\mathbf{s}_0, \mathbf{t}) = f_n(\mathbf{t})$. Since for $\alpha > 0$ all terms on the right-hand side in (2.1) are non-negative, it holds that $\rho_{2q}(\mathbf{x}, \mathbf{y}) \geq \rho_q(\mathbf{x}) \rho_q(\mathbf{y})$ for all $\mathbf{x} \in \mathbb{R}^{qd}$, $\mathbf{y} \in \mathbb{R}^{qd}$. This yields

$$\mathrm{Var}(S_n(G)) \geq \sum_{k=0}^{q-1} \binom{q}{k} \int_{\mathbb{X}^q} \int_{\mathbb{X}^k} f_n(\mathbf{s}) f_n(\mathbf{s}_{q-k}, \mathbf{t}) \rho_{q+k}(\mathbf{x}, \mathbf{y}) \mu_n^k(\mathrm{d}\mathbf{t}) \mu_n^q(\mathrm{d}\mathbf{s}). \tag{4.2}$$

To bound the term with $k = 0$, note that $f_n(\mathbf{x})$ is bounded by the indicator that $x_2, \dots, x_q$ are all contained in a ball with volume $r_n := \varepsilon \nu_n / 2^d$ around x_1 and that the WRCM built on $x_1, \dots, x_q$ is connected. Here $\varepsilon > 0$ is chosen such that $\varphi(\varepsilon) > 0$ and such that for all $x \in W_n^{-r_n}$ and all $n \geq 1$, the ball centered at x with volume r_n is entirely contained in W_n. Such an ε exists since the sequence $(\nu_n / n)_n$ is bounded. Since any two points in a ball with volume r_n are connected with probability at least $p := \varphi(\varepsilon)$ and $\rho_q(\mathbf{x}) \geq K_0(o)^q$ for $\alpha > 0$, this yields the lower bound

$$\int_{\mathbb{X}^q} f_n(\mathbf{s}) \rho_q(\mathbf{x}) \mu_n^q(\mathrm{d}\mathbf{s}) \geq K_0(o)^q |W_n^{-r_n}| p^{\binom{q}{2}} r_n^{q-1}.$$

With an analogous argument for the term with $k = q - 1$ in (4.2), we find that

$$\int_{\mathbb{X}^q} \int_{\mathbb{X}^{q-1}} f_n(\mathbf{s}) f_n(s_1, \mathbf{t}) \rho_{2q-1}(\mathbf{x}, \mathbf{y}) \mu_n^{q-1}(\mathrm{d}\mathbf{t}) \mu_n^q(\mathrm{d}\mathbf{s}) \geq K_0(o)^{2q-1} |W_n^{-r_n}| p^{2\binom{q}{2}} r_n^{2q-2}.$$

Combining these estimates gives the assertion. □

Remark 4.1. It is a notoriously challenging task to obtain lower bounds on the variance of geometric functionals of (stationary) determinantal point processes (see also the discussion in Remark (iii) after Theorem 1.14 in [1]). This is due to the fact that $\rho_{2q}(\mathbf{x}, \mathbf{y}) \leq \rho_q(\mathbf{x}) \rho_q(\mathbf{y})$ for $\tau \leq 0$, which results in positive and negative terms in the expansion of the variance in (4.1). In the case $q = 1$ (where S_n is the number of points in $\mathcal{P} \cap W_n$), we have for $\int |K_0(x)|^2 \, \mathrm{d}x = K_0(o)$ that the variance of S_n grows for $n \to \infty$ as the surface area of W_n, i.e. $\mathrm{Var}(S_n) \sim |W_n|^{(d-1)/d}$ (see [6], (4.15)). For instance, this is the case in the Ginibre process, and similar phenomena also occur for $\int |K_0(x)|^2 \, \mathrm{d}x > K_0(o)$. However, if $\int |K_0(x)|^2 \, \mathrm{d}x < K_0(o)$ we have $\mathrm{Var}(S_n) \sim |W_n|$. If $q \geq 2$, the situation is more complex, since the asymptotic behavior of the variance of $S_n(G)$ is determined by the asymptotic properties of K, $(\nu_n)_{n \geq 1}$ as well as the profile function φ. A concrete scenario in which we can determine the growth rate of $\mathrm{Var}(S_n(G))$ is when $q = 2$, G is the connected graph on two vertices, $\nu_n \to 0$, and the profile

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

function φ has bounded support. In this case we obtain from (4.1) that

$$\begin{aligned}\mathrm{Var}(S_n(G)) = &\int_{W_n^2} \mathbb{E}[\varphi_n(x_1 - x_2, U_1, U_2)]\rho_2(x_1, x_2)\,\mathrm{d}(x_1, x_2)\\ &+ 2\int_{W_n^3} \mathbb{E}[\varphi_n(x_1 - x_2, U_1, U_2)\varphi_n(x_1 - x_3, U_1, U_3)]\rho_3(x_1, x_2, x_3)\,\mathrm{d}(x_1, x_2, x_3)\\ &+ \int_{W_n^4} \mathbb{E}[\varphi_n(x_1 - x_2, U_1, U_2)\varphi_n(x_3 - x_4, U_3, U_4)]\\ &\quad\times(\rho_4(x_1, x_2, x_3, x_4) - \rho_2(x_1, x_2)\rho_2(x_3, x_4))\,\mathrm{d}(x_1, x_2, x_3, x_4).\end{aligned}$$

A detailed analysis of the three different terms that proceeds along the lines of the proof of [10, Theorem 3.2] shows that the first term grows at rate $|W_n|\nu_n$, whereas the second and the third terms grow at a rate that is at most $|W_n|\nu_n^2$. Hence there is some $v > 0$ such that

$$\mathrm{Var}(S_n(G)) \sim v|W_n|\nu_n \quad \text{as } n \to \infty.$$

5. Preliminaries

In this section we set the stage for the proof of our main results and recall some material from [20]. To keep the presentation self-contained, we start in Section 5.1 with elementary facts on some partitions which play a fundamental role in the method of cumulants. Then, in Section 5.2, we recall the method of cumulants, and following this, the main results from [20].

5.1. Partitions

Partitions play a central role in our application of the method of cumulants, and in the present section we reproduce some material from [14, Chapter 12.2]; see also [20, Section 2.4]. However, since we only deal with single U-statistics and not linear combinations of U-statistics, we can simplify the presentation.

Let $m \geq 1$ and let $q \geq 1$. We then define $N = mq$, and for each $\ell \in \{1, \dots, m\}$ we let $J_\ell := \{(\ell - 1)q + 1, \dots, \ell q\}$. As an illustration, one may imagine N numbers arranged in an array of m rows, each having q entries. Then J_ℓ is the ℓth line in this arrangement. A partition σ of $\{1, \dots, N\}$ is a collection $\{B_1, \dots, B_k\}$ of $1 \leq k \leq N$ pairwise disjoint non-empty sets, called blocks, such that $B_1 \cup \cdots \cup B_k = \{1, \dots, N\}$. By $\Pi^m(q)$, we denote the set of all partitions σ such that

$$\max_{B\in\sigma} \max_{\ell\leq m} |B \cap J_\ell| \leq 1,$$

and by $\Pi^m_{\geq 2}(q)$ the set of all $\sigma \in \Pi^m(q)$ with $\min_{B\in\sigma} |B| \geq 2$.

Every partition $\sigma \in \Pi^m(q)$ induces a partition σ^* of $\{1, \dots, m\}$ in the following way: $k, \ell \in \{1, \dots, m\}$ are in the same block of σ^* whenever there is a block $B \in \sigma$ such that $|B \cap J_k| = 1$ and $|B \cap J_\ell| = 1$. Let $\widetilde{\Pi}^m(q)$ be the set of all partitions $\sigma \in \Pi^m(q)$ such that $|\sigma^*| = 1$, and define $\widetilde{\Pi}^m_{\geq 2}(q)$ in the same way as $\Pi^m_{\geq 2}(q)$.

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space. For functions $f\colon \mathbb{X}^q \to \mathbb{R}$, we define their product $f^{\otimes m}\colon \mathbb{X}^N \to \mathbb{R}$ by

$$(f^{\otimes m})(x_1, \dots, x_N) := \prod_{\ell=1}^{m} f(x_{(\ell-1)q+1}, \dots, x_{\ell q}).$$

Notice that for each ℓ the function f in the product takes entries of J_ℓ as its input.

Then, for some $\sigma \in \Pi^m(q)$, the function $(f^{\otimes m})_\sigma\colon \mathbb{X}^{|\sigma|} \to \mathbb{R}$ is obtained by replacing in $f^{\otimes m}$ all variables that belong to the same block of σ by a new common variable. For example,

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

if $m = q = 2$, and for a partition σ on $\{1, \dots, 4\}$ that consists of the blocks $\{2\}$, $\{1, 3\}$, and $\{4\}$, this means

$$(f^{\otimes 2})_\sigma(x_1, x_2, x_3) = f^{\otimes 2}(x_1, x_2, x_1, x_3) = f(x_1, x_2)f(x_1, x_3)$$

(see e.g. [14, Chapter 12.2]). Note that in this way $(f^{\otimes m})_\sigma$ is only defined up to permutations of its arguments. Since in what follows we always integrate with respect to all arguments of $(f^{\otimes m})_\sigma$, this does not cause problems.

5.2. Method of cumulants

In this section we introduce the key tool for the proof of the main results, namely the method of cumulants, whose development has been mainly attributed to V. Statulevičius. The results of earlier articles (see e.g. [18], [22]) are collected in the textbook [19]. We also want to point to the more recent survey [3]. Henceforth, $\operatorname{cum}(X_1, \dots, X_m)$ denotes the joint cumulant of the random variables $X_1, \dots, X_m$. We recall that with i the imaginary unit, this is given as

$$\operatorname{cum}(X_1, \dots, X_m) := (-\mathrm{i})^m \frac{\partial^m}{\partial t_1 \cdots \partial t_m} \log \phi_{X_1,\dots,X_m}(t_1, \dots, t_m)|_{t_1=\cdots=t_m=0},$$

where

$$\phi_{X_1,\dots,X_m}(t_1, \dots, t_m) := \mathbb{E}\left[\exp\left(\mathrm{i} \sum_{i \le m} t_i X_i\right)\right]$$

is the characteristic function of the random vector $(X_1, \dots, X_m)$. We henceforth always assume that all considered random variables have finite moments of all orders. Moreover, we denote the mth-order cumulant of a random variable X by $\operatorname{cum}_m(X)$ and note that $\operatorname{cum}(X, \dots, X) = \operatorname{cum}_m(X)$, where the left-hand side takes an m-tuple as its input.

The following proposition can be regarded as the application of the method of cumulants to Poisson-based U-statistics and is a variant of [20, Theorem 3.2]. In contrast to the latter, we allow for the additional factor $(m!)^p$ in assumption (5.1) below. Although we do not need it for our purposes, we note that the statement could be generalized to linear combinations of U-statistics.

Proposition 5.1. (Variant of Theorem 3.2 in [20].) *Let $(\mathcal{P}_n)_n$ be a family of Poisson processes over σ-finite measure spaces $((\mathbb{X}_n, \mathcal{X}_n, \mu_n))_n$ and let $f_n \colon \mathbb{X}_n^q \to \mathbb{R}$, $n \in \mathbb{N}$, be measurable and satisfy the integrability condition*

$$\int_{\mathbb{X}^i} \left(\int_{\mathbb{X}^{q-i}} f(y_1, \dots, y_i, x_1, \dots, x_{q-i})\, \mu_n^{q-i}(\mathrm{d}(x_1, \dots, x_{q-i})) \right)^2 \mu_n^i(\mathrm{d}(y_1, \dots, y_i)) < \infty$$

for $i \in \{0, \dots, q\}$ and $\sup_{n \ge 1} \|f_n\|^2_{L^2(\mu_n^q)} > 0$. Define for $n \ge 1$ the random variables

$$Z_n := \sum_{(x_1,\dots,x_q) \in \mathcal{P}_{n,\neq}^q} f_n(x_1, \dots, x_q).$$

Assume that there is a positive real-valued sequence $(\beta_n)_{n \ge 1}$ and some $p \ge 0$, such that for any $n \in \mathbb{N}$, for all $\sigma \in \widetilde{\Pi}_m(q)$ and $m \ge 3$,

$$\operatorname{Var}(Z_n)^{-m/2} \left| \int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \, \mathrm{d}\mu^{|\sigma|} \right| \le (m!)^p \beta_n^{-(m-2)}. \tag{5.1}$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

Then

$$\left(\frac{Z_n}{\sqrt{\mathrm{Var}(Z_n)}}\right)_n$$

satisfies **MDP**, **CI**, *and* **NACC**, *all with parameters* $(p+q-1, (q^{-3q}\beta_n)_n)$.

Proof. The main tool of the cumulant method is to find parameters $\gamma \geq 0$ and $\Delta_n > 0$, such that the cumulants of the random variable $Z_n/\sqrt{\mathrm{Var}(Z_n)}$ as in the proposition satisfy the condition

$$\left|\mathrm{cum}_m\left(\frac{Z_n}{\sqrt{\mathrm{Var}(Z_n)}}\right)\right| \leq \frac{(m!)^{1+\gamma}}{\Delta_n^{m-2}}$$

for all $m \geq 3$ and $n \in \mathbb{N}$. This condition is widely known as the Statulevičius condition and implies that **MDP**, **CI**, and **NACC** are all satisfied with parameters $(\gamma, (\Delta_n)_{n\geq 1})$; for more details we refer to [3, Section 2]. To do so, we follow the arguments in the proof of [20, Theorem 3.1]. If we fix some $m \geq 3$, by definition of the cumulant, its multilinearity, and the triangle inequality, we have

$$\left|\mathrm{cum}_m\left(\frac{Z_n}{\sqrt{\mathrm{Var}\,(Z_n)}}\right)\right| = \frac{1}{\mathrm{Var}\,(Z_n)^{m/2}}\,|\,\mathrm{cum}(Z_n,\ldots,Z_n)|.$$

Now, according to [20, Theorem 3.7], we can rewrite

$$\begin{aligned}|\,\mathrm{cum}_m\,(Z_n,\ldots,Z_n)| &= \left|\sum_{\sigma\in\widetilde{\Pi}^m(q)}\int_{\mathbb{X}^{|\sigma|}}(f_n^{\otimes m})_\sigma\,\mathrm{d}\mu^{|\sigma|}\right| \\ &\leq \left|\widetilde{\Pi}^m(q)\right|\sup_{\sigma\in\widetilde{\Pi}^m(q)}\left|\int_{\mathbb{X}^{|\sigma|}}(f_n^{\otimes m})_\sigma\,\mathrm{d}\mu^{|\sigma|}\right|.\end{aligned}$$

Additionally, we observe

$$\left|\widetilde{\Pi}^m(q)\right| \leq |\Pi^m(q)| \leq q^{qm}(m!)^q,$$

where the first inequality is a straightforward consideration by the definitions of $\widetilde{\Pi}^m(q)$ and $\Pi^m(q)$ respectively, and the second one is due to the first part of [20, Proposition 6.1]. Putting this together with assumption (5.1) gives

$$\left|\mathrm{cum}_m\left(\frac{Z_n}{\sqrt{\mathrm{Var}\,(Z_n)}}\right)\right| \leq q^{qm}(m!)^{p+q}\beta_n^{-(m-2)} \leq (m!)^{p+q}(q^{-3q}\beta_n)^{-(m-2)},$$

so that in comparison to the above cumulant bound we observe that $\gamma = p+q-1$ and $\Delta_n = \beta_n q^{-3q}$. □

5.3. Parameter specifications

In this section we prepare the proofs of our main theorems by explaining how we choose the model parameters in Proposition 5.1.

Even if all Poisson points $\{(X_i, U_i)\}_i$ are assumed as given, in the current framework the decision whether to put an edge between (X_i, U_i) and (X_j, U_j) still involves additional randomness. We resolve this technicality in the same way as in [17] (see particularly Section 4 therein). More precisely, we formally enlarge the state space of $\mathcal{P}_n$ from $W \times [1,\infty)$ to $\mathbb{X} := W \times [1,\infty) \times [0,1]^{\mathbb{N}}$ and augment each Poisson point (X_i, U_i) with additional i.i.d. uniform marks $(T_i^{(j)})_{j\geq 1}$. Hence we choose $\mu_n := t_n \mathrm{d}x \otimes \mathbb{P}_U \otimes (\mathrm{Unif}[0,1])^{\mathbb{N}}$. We then index the

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

points of $\mathcal{P}_n$ in lexicographic order and put an edge between P_i and P_j with $i<j$ if and only if

$$T_i^{(j)} \le \varphi_n(X_i - X_j, U_i, U_j) = \varphi\left(\frac{|B_{\|X_i - X_j\|}|}{v_n \kappa(U_i, U_j)}\right),$$

and thereby realize the edge probability $\varphi_n(X_i - X_j, U_i, U_j)$.

6. Proof of Theorem 2.1

Let G be a fixed connected graph on vertices $1, \dots, q$. For

$$p_i = \left(x_i, u_i, \left(t_i^{(j)}\right)_j\right) \in W \times [1,\infty) \times [0,1]^{\mathbb{N}}, \quad i = 1, \dots, q,$$

let $f_n(p_1, \dots, p_q)$ be $1/q!$ times the number of subgraphs in the WRCM built on these points that are isomorphic to G. While now the basic idea is similar to the representation given in [20, (7.1)], where the functionals for the subgraph count in the classical random connection model are given as a linear combination of indicator variables, the presence of long-range correlations through the profile function φ and the weights requires additional work.

Let $\sigma \in \widetilde{\Pi}^m(q)$ be a partition, and note that it induces a connected graph (V, E) with vertex set $V = \{1, \dots, |\sigma|\}$, where the block $B_i \in \sigma$ has been identified with node $i \in V$, $i = 1, \dots, |\sigma|$. There is an edge $(i, j) \in E$ if there are $k_i, k_j \in \{1, \dots, q\}$ and $\ell \in \{0, \dots, m-1\}$, such that $\ell q + k_i \in B_i$, $\ell q + k_j \in B_j$ and there is an edge between k_i and k_j in G. For instance, let G be the graph on vertices $\{1, 2, 3\}$ with edges $(1, 2)$ and $(2, 3)$, let $m = 2$ and $\sigma = \{B_1, \dots, B_5\}$ with $B_1 = \{1, 4\}$, $B_2 = \{2\}$, $B_3 = \{3\}$, $B_4 = \{5\}$, $B_5 = \{6\}$. Then $V = \{1, \dots, 5\}$ and $E = \{(1, 2), (2, 3), (1, 4), (4, 5)\}$. Note, moreover, that in this case

$$(f_n^{\otimes 2})_\sigma(p_1, \dots, p_5) = f_n(p_1, p_2, p_3) f_n(p_1, p_4, p_5), \quad p_1, \dots, p_5 \in \mathbb{X}.$$

In general, we have

$$\int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \, \mathrm{d}\mu_n^{|\sigma|} = t_n^{|\sigma|} \int_{W^{|\sigma|}} \mathbb{E}\left[\prod_{(i,j)\in E} \varphi_n(x_i - x_j, U_i, U_j)\right] \mathrm{d}(x_1, \dots, x_{|\sigma|}). \tag{6.1}$$

We shall give a few more details regarding the above equation. On the left-hand side of (6.1), the integration over $\mathbb{X}^{|\sigma|}$ formally is an integration over $(W \times [1,\infty) \times [0,1]^{\mathbb{N}})^{|\sigma|}$ and the integral is split accordingly. The integrand is then the indicator that for an edge $(i, j) \in E$ (the edge set of the graph induced by σ), there is an edge in the WRCM between the points P_i and P_j. The latter now depends on the marks $T_i^{(j)}$ as described above. The integral over $[0,1]^{|\sigma|}$ is then formally solved by writing the probability that $T_i^{(j)} \le \varphi_n(\cdot)$ as $\varphi_n(\cdot)$. Further, the expectation on the right-hand side of (6.1) represents the integral over $[1,\infty)^{|\sigma|}$ so that the only 'visible' integral is the one with respect to the Lebesgue measure $t_n^{|\sigma|} \mathrm{d}(x_1, \dots, x_{|\sigma|})$.

Next we fix a spanning tree of (V, E) with induced edge set $E_s \subseteq E$. Then we successively integrate the tree E_s, where we use spherical coordinates in each integration step. Since the factors in the product on the right-hand side of (6.1) are at most one, this indeed gives an upper bound.

In more detail, first observe that by Fubini's theorem and the fact that $E_s \subseteq E$, the integral in (6.1) is bounded from above by

$$\mathbb{E}\left[\int_{W^{|\sigma|}} \prod_{(i,j)\in E_s} \varphi_n(x_i, x_j, U_i, U_j) \, \mathrm{d}x_1, \dots, \mathrm{d}x_{|\sigma|}\right].$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

For better illustration, we consider one of the $|\sigma|$-fold integrals separately and let $(i,j) \in E_s$. Then,

$$\int_W \varphi_n(x_i, x_j, U_i, U_j)\,\mathrm{d}x_j \le \int_{\mathbb{R}^d} \varphi\left(\frac{B_{||x_i - x_j||}}{\nu_n \kappa(U_i, U_j)}\right)\mathrm{d}x_j.$$

The right-hand side integral over $\mathbb{R}^d$ can now be rewritten using spherical coordinates around x_i. This gives

$$\nu_n \kappa(U_i, U_j) \int_0^\infty \varphi(s)\,\mathrm{d}s \;=\; \nu_n \kappa(U_i, U_j),$$

where we have used that $\int_0^\infty \varphi(s)\,\mathrm{d}s = 1$ by assumption from Section 2. Now, since the edge set of the spanning tree satisfies $|E_s| = |\sigma| - 1$ by construction, an iteration of the above yields

$$\begin{aligned}\mathbb{E}\left[\int_{W^{|\sigma|}} \prod_{(i,j)\in E_s} \varphi_n(x_i, x_j, U_i, U_j)\,\mathrm{d}x_1, \ldots, \mathrm{d}x_{|\sigma|}\right] \\ \le \int_W \mathrm{d}x_{|\sigma|}\; \nu_n^{|\sigma|-1} \mathbb{E}\left[\prod_{(i,j)\in E_s} \kappa(U_i, U_j)\right].\end{aligned}$$

Since the first factor on the right-hand side equals the volume of W, we have arrived at

$$\int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma\,\mathrm{d}\mu_n^{|\sigma|} \le t_n^{|\sigma|} \nu_n^{|\sigma|-1} |W|\, \mathbb{E}\left[\prod_{(i,j)\in E_s} \kappa(U_i, U_j)\right]. \tag{6.2}$$

So the only thing left to care about is the expectation on the right-hand side of (6.2). Therefore let $\deg_s(i)$ denote the degree of $i \in \{1, \ldots, |\sigma|\}$ in (V, E_s). Then $\kappa(u, v) \le uv$ and (MU) give

$$\begin{aligned}\mathbb{E}\left[\prod_{(i,j)\in E_s} \kappa(U_i, U_j)\right] &\le \prod_{i\le|\sigma|} M_U(\deg_s(i)) \\ &\le \prod_{i\le|\sigma|} c_{U,1}^{\deg_s(i)} (\deg_s(i)!)^{c_{U,2}} \\ &= c_{U,1}^{2(|\sigma|-1)} \prod_{i\le|\sigma|} (\deg_s(i)!)^{c_{U,2}}.\end{aligned}$$

Here the product over $i = 1, \ldots, |\sigma|$ can be expressed as

$$\prod_{i\le|\sigma|} (\deg_s(i)!)^{c_{U,2}} = \exp\left(c_{U,2} \sum_{i\le|\sigma|} \log\Gamma(\deg_s(i) + 1)\right).$$

First we show that the function $f(s_1, \ldots, s_{|\sigma|}) := \sum_{i\le|\sigma|} \log\Gamma(s_i + 1)$ is convex. Indeed, for $n \ge 1$, the $(n+1)$th-order derivative of $\log\Gamma(s)$ is the polygamma function of order n given by

$$\Psi_n(s) := (-1)^{n+1} n! \sum_{k\ge0} \frac{1}{(s+k)^{n+1}}\,.$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

This implies that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\log\Gamma(s)=\Psi_1(s)=\sum_{k\geq 0}\frac{1}{(s+k)^2}>0,$$

which yields convexity of the function $\log\Gamma(s)$. Hence, on the set

$$\{(s_1,\dots,s_{|\sigma|})\in\mathbb{N}^{|\sigma|}\colon s_1+\cdots+s_{|\sigma|}=2(|\sigma|-1)\},$$

the function f takes its maximum at an extremal point. In other words, there is some $i_0\in\{1,\dots,|\sigma|\}$ such that $s_{i_0}=|\sigma|-1$, and $s_i=1$ for $i\neq i_0$. In particular,

$$\prod_{i\leq|\sigma|}(\deg_{\mathrm{s}}(i)!)^{c_{U,2}}\leq((|\sigma|-1)!)^{c_{U,2}}\leq((m(q-1))!)^{c_{U,2}}\leq(q-1)^{m(q-1)c_{U,2}}(m!)^{(q-1)c_{U,2}},$$

where we have used $|\sigma|\leq m(q-1)+1$ and the elementary inequality

$$(ck)!\leq c^{ck}(k!)^c, \tag{6.3}$$

which holds for all $c\geq 1$ and $k\in\mathbb{N}$. Thus,

$$\mathbb{E}\bigg[\prod_{(i,j)\in E_s}\kappa(U_i,U_j)\bigg]\leq c_{U,1}^{2(|\sigma|-1)}(q-1)^{m(q-1)c_{U,2}}(m!)^{(q-1)c_{U,2}}\leq b_{\mathrm{SG}}^m(m!)^{(q-1)c_{U,2}}, \tag{6.4}$$

where we recall that

$$b_{\mathrm{SG}}:=\big((1\vee c_{U,1}^2)(q-1)^{c_{U,2}}\big)^{q-1}$$

from Theorem 2.1. Hence we arrive at

$$\int_{\mathbb{X}^{|\sigma|}}(f_n^{\otimes m})_\sigma\,\mathrm{d}\mu_n^{|\sigma|}\leq t_n^{|\sigma|}\nu_n^{|\sigma|-1}|W|\,b_{\mathrm{SG}}^m(m!)^{(q-1)c_{U,2}}.$$

Now, by condition (VSG), there exists some constant $v>0$ such that

$$\operatorname{Var}(S_n(G))\geq v|W|t_n^q\nu_n^{q-1}\big(1\vee t_n^{q-1}\nu_n^{q-1}\big).$$

Next we distinguish the cases $t_n\nu_n\leq 1$ and $t_n\nu_n>1$. If $t_n\nu_n\leq 1$, we use $|\sigma|\geq q$ and find with $a:=(q-1)c_{U,2}$ that

$$\begin{aligned}(m!)^{-a}(\operatorname{Var}S_n(G))^{-m/2}\int_{\mathbb{X}^{|\sigma|}}(f_n^{\otimes m})_\sigma\,\mathrm{d}\mu_n^{|\sigma|}&\leq\frac{|W|t_n^qb_{\mathrm{SG}}^m\nu_n^{q-1}}{|W|^{m/2}v^{m/2}t_n^{mq/2}\nu_n^{m(q-1)/2}}\\&=\frac{b_{\mathrm{SG}}^m|W|^{-(m-2)/2}}{v^{m/2}t_n^{(m-2)q/2}\nu_n^{(m-2)(q-1)/2}}\\&\leq(1\vee b_{\mathrm{SG}}^2/v)^{m-2}\big(b_{\mathrm{SG}}/\sqrt{|W|vt_n^q\nu_n^{q-1}}\big)^{m-2}.\end{aligned}$$

If $t_n\nu_n\geq 1$, we use the bound $|\sigma|\leq m(q-1)+1$ and obtain

$$\begin{aligned}(m!)^{-a}(\operatorname{Var}S_n(G))^{-m/2}\int_{\mathbb{X}^{|\sigma|}}(f_n^{\otimes m})_\sigma\,\mathrm{d}\mu_n^{|\sigma|}&\leq\frac{|W|t_n^{m(q-1)+1}b_{\mathrm{SG}}^m\nu_n^{m(q-1)}}{|W|^{m/2}v^{m/2}t_n^{mq-m/2}\nu_n^{m(q-1)}}\\&=\frac{|W|^{-(m-2)/2}t_n^{1-m/2}b_{\mathrm{SG}}^m}{v^{m/2}}\\&\leq\big(1\vee b_{\mathrm{SG}}^2/v\big)^{m-2}\big(b_{\mathrm{SG}}/\sqrt{|W|vt_n}\big)^{m-2}.\end{aligned}$$

Therefore a cumulant bound as in (5.1) is satisfied and the result follows from Proposition 5.1.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

7. Proof of Theorem 2.2

In this section we prove Theorem 2.2. We apply Proposition 5.1 with state space $\mathbb{X} := W \times [1, \infty) \times [0, 1]^{\mathbb{N}}$, $q = 2$ and

$$f_n(P_i, P_j) := |B_{\|X_i - X_j\|}|^\tau \mathbf{1}\{T_i^{(j)} \le \varphi_n(X_i - X_j, U_i, U_j)\}.$$

In the proof, we replace the τ-power-weighted edge length $\|X_i - X_j\|^\tau$ by

$$|B_{\|X_i - X_j\|}|^\tau = |B_1|^\tau \|X_i - X_j\|^{d\tau}.$$

Note that the linear scaling of $S_n^{(\tau)}$ by $|B_1|^\tau$ has no influence on the fraction $S_n^{(\tau)}/\sqrt{\mathrm{Var}(S_n^\tau)}$ and therefore does not affect our bounds. As we will see below, this reparameterization will make the integral computations substantially more accessible.

As in Section 6, we use the fact that the partition σ induces a connected graph (V, E) on the vertex set $V = \{1, \dots, |\sigma|\}$. In fact, in the present setting this graph is far easier to comprehend since $q = 2$. Moreover, for $(i, j) \in E$ we henceforth let m_{ij} be the multiplicity of (i, j) in σ and note that $\sum_{(i,j)\in E} m_{ij} = m$. Then,

$$\int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \, \mathrm{d}\mu_n^{|\sigma|} = t_n^{|\sigma|} \mathbb{E}\left[\int_{W^{|\sigma|}} \prod_{(i,j)\in E} |B_{\|x_i - x_j\|}|^{\tau m_{ij}} \varphi_n(x_i - x_j, U_i, U_j) \, \mathrm{d}(x_1, \dots, x_{|\sigma|}) \right]. \tag{7.1}$$

Again, we fix the edge set $E_s \subseteq E$ of a spanning tree and rewrite the inner integral in the above as

$$\begin{aligned}
&\int_{W^{|\sigma|}} \left[\prod_{(i,j)\in E_s} \varphi_n(x_i - x_j, U_i, U_j)) |B_{\|x_i - x_j\|}|^{\tau m_{ij}} \right] \\
&\qquad \times \left[\prod_{(i,j)\in E\setminus E_s} \varphi_n(x_i - x_j, U_i, U_j) |B_{\|x_i - x_j\|}|^{\tau m_{ij}} \right] \mathrm{d}(x_1, \dots, x_{|\sigma|}) \\
&\le \left[\prod_{(i,j)\in E\setminus E_s} (\nu_n \kappa(U_i, U_j))^{\tau m_{ij}} M'_\varphi(\tau m_{ij}) \right] \\
&\qquad \times \int_{W^{|\sigma|}} \prod_{(i,j)\in E_s} \varphi_n(x_i - x_j, U_i, U_j) |B_{\|x_i - x_j\|}|^{\tau m_{ij}} \, \mathrm{d}(x_1, \dots, x_{|\sigma|}),
\end{aligned}$$

where M'_φ is given at (MPHI). Next, we bound the integral over $x_1, \dots, x_{|\sigma|}$. Again, successively integrating out the contributions of the spanning tree leaves gives the bound

$$\begin{aligned}
&\int_{W^{|\sigma|}} \prod_{(i,j)\in E_s} \varphi_n(x_i - x_j, U_i, U_j) |B_{\|x_i - x_j\|}|^{\tau m_{ij}} \, \mathrm{d}(x_1, \dots, x_{|\sigma|}) \\
&\qquad \le |W| \prod_{(i,j)\in E_s} [(\nu_n \kappa(U_i, U_j))^{\tau m_{ij}+1} M_\varphi(\tau m_{ij})].
\end{aligned}$$

Now, assumption (MPHI) gives that

$$\prod_{(i,j)\in E\setminus E_s} M'_\varphi(\tau m_{ij}) \prod_{(i,j)\in E_s} M_\varphi(\tau m_{ij}) \le c_{\varphi,1}^{\tau m} \Gamma(1 + \tau m)^{c_{\varphi,2}}.$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

Note that for $d_1, \dots, d_k \geq 0$ and identically distributed random variables $Y_1, \dots, Y_k$, Hölder's inequality gives

$$\mathbb{E}[Y_1^{d_1} \dots Y_k^{d_k}] \leq \mathbb{E}[Y_1^{d_1+\dots+d_k}].$$

Using this, $|\sigma| \leq m+1$, and (6.3) shows that (7.1) is bounded by

$$\begin{aligned}
&t_n^{|\sigma|} |W| c_{\varphi,1}^{\tau m} \Gamma(1+\tau m)^{c_{\varphi,2}} \mathbb{E}\Bigg[\prod_{(i,j)\in E_s} (v_n \kappa(U_i, U_j))^{\tau m_{ij}+1} \prod_{(i,j)\in E\setminus E_s} (v_n \kappa(U_i, U_j))^{\tau m_{ij}}\Bigg] \\
&\quad = t_n^{|\sigma|} |W| c_{\varphi,1}^{\tau m} \Gamma(1+\tau m)^{c_{\varphi,2}} \mathbb{E}\Bigg[\prod_{(i,j)\in E} (v_n \kappa(U_i, U_j))^{\tau m_{ij}+\mathbf{1}\{(i,j)\in E_s\}}\Bigg] \\
&\quad \leq t_n^{|\sigma|} |W| c_{\varphi,1}^{\tau m} \Gamma(1+\tau m)^{c_{\varphi,2}} v_n^{\tau m+|\sigma|-1} \mathbb{E}\big[\kappa(U_1, U_2)^{(\tau+1)m}\big] \\
&\quad \leq t_n^{|\sigma|} |W| (c_{\varphi,1}(\tau \vee 1)^{c_{\varphi,2}})^{\tau m} (m!)^{c_{\varphi,2}\tau} v_n^{\tau m+|\sigma|-1} \mathbb{E}\big[\kappa(U_1, U_2)^{(\tau+1)m}\big],
\end{aligned}$$

where we have used that $\Gamma(1+\tau m) \leq (\tau \vee 1)^{\tau m}(m!)^{\tau}$ in the last inequality, which follows from convexity for $\tau \leq 1$ and from (6.3) for $\tau > 1$. Hence it remains to bound the expected value on the right-hand side. To do so, since $\kappa(u, v) \leq uv$, using condition (MU) we find with the bound (6.3) that

$$\begin{aligned}
\mathbb{E}\big[\kappa(U_1, U_2)^{(\tau+1)m}\big]^{1/2} &\leq c_{U,1}^{(\tau+1)m} \Gamma(1+(\tau+1)m)^{c_{U,2}} \\
&\leq (c_{U,1}(\tau+1)^{c_{U,2}})^{(\tau+1)m} (m!)^{c_{U,2}(\tau+1)}. \qquad (7.2)
\end{aligned}$$

Moreover, by condition (VEL), there exists some $v > 0$ such that

$$\operatorname{Var}(S_n^{(\tau d)}) \geq v |W| t_n^2 v_n^{2\tau+1} (1 \vee t_n v_n).$$

Now we combine these bounds and consider the cases $t_n v_n \geq 1$ and $t_n v_n < 1$ separately. Recall that $b_{\mathrm{EL}} := (c_{\varphi,1}(\tau \vee 1)^{c_{\varphi,2}})^{\tau} (c_{U,1}(\tau+1)^{c_{U,2}})^{2(\tau+1)}$ and $a := 2c_{U,2}(\tau+1) + c_{\varphi,2}\tau$. If $t_n v_n < 1$, we use $|\sigma| \geq 2$ and obtain

$$\begin{aligned}
(m!)^{-a} \operatorname{Var}(S_n^{(\tau d)})^{-m/2} \int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \, \mathrm{d}\mu_n^{|\sigma|} &\leq \frac{|W| t_n^2 v_n^{\tau m} b_{\mathrm{EL}}^m v_n}{|W|^{m/2} v^{m/2} v_n^{\tau m} t_n^m v_n^{m/2}} \\
&= \frac{b_{\mathrm{EL}}^m}{|W|^{(m-2)/2} v^{m/2} t_n^{m-2} v_n^{(m-2)/2}} \\
&\leq \big(1 \vee b_{\mathrm{EL}}^2/v\big)^{m-2} \big(b_{\mathrm{EL}}/\sqrt{|W| v t_n^2 v_n}\big)^{m-2}.
\end{aligned}$$

If $t_n v_n \geq 1$, because of $|\sigma| \leq m+1$, we have

$$\begin{aligned}
(m!)^{-a} \operatorname{Var}(S_n^{(\tau d)})^{-m/2} \int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \, \mathrm{d}\mu_n^{|\sigma|} &\leq \frac{|W| t_n^{m+1} v_n^{\tau m} b_{\mathrm{EL}}^m v_n^m}{|W|^{m/2} v^{m/2} v_n^{\tau m} t_n^{2m-m/2} v_n^m} \\
&= \frac{b_{\mathrm{EL}}^m}{|W|^{(m-2)/2} t_n^{(m-2)/2} v^{m/2}} \\
&\leq \big(1 \vee b_{\mathrm{EL}}^2/v\big)^{m-2} \big(b_{\mathrm{EL}}/\sqrt{|W| v t_n}\big)^{m-2}. \qquad \square
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

8. Proof of Theorems 2.3 and 2.4

Before giving the proof of Theorem 2.3, we provide a formula from the Appendix of [10] (or see Lemma 1 in [5]) for the mth cumulant of a U-statistic of an α-DPP, and explain how this formula can be interpreted graphically. This will allow us to proceed similarly as in the proof of Theorem 2.1, with the difference that edges in the graph induced by σ can now have two different meanings.

Given a partition $\sigma \in \Pi^m(q)$ of $\{1, \dots, mq\}$, let ρ be a partition of $\{1, \dots, |\sigma|\}$. Recalling the definition of J_t from Section 5.1, we call the pair (σ, ρ) *decomposable* if there are a decomposition $\{T_1, T_2\}$ of $\{1, \dots, q\}$ and decompositions $\sigma = \sigma^{(1)} \cup \sigma^{(2)}$ and $\rho = \rho^{(1)} \cup \rho^{(2)}$ such that $\sigma^{(1)}$ is a partition of $\bigcup_{t\in T_1} J_t$, $\sigma^{(2)}$ is a partition of $\bigcup_{t\in T_2} J_t$, $\rho^{(1)}$ is a partition of T_1, and $\rho^{(2)}$ is a partition of T_2. Otherwise (σ, ρ) is *indecomposable*. Given $\sigma \in \Pi^m(q)$, let $\Pi(q;\sigma)$ be the set of all partitions ρ such that the pair (σ, ρ) is indecomposable. This notion of indecomposability is compatible with the notion of indecomposable integrals used in [10] in the following sense. An integral is indecomposable in the sense of [10] if it cannot be defined with respect to a decomposable partition. Hence the notion of indecomposable partitions is less restrictive than indecomposable integrals.

Therefore, with f_n defined as in the proof of Theorem 2.1, we have

$$|\operatorname{cum}_m(S_n(G))| \le \sum_{\sigma\in\Pi^m(q)} \sum_{\rho\in\Pi(q;\sigma)} \int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \prod_{Q\in\rho} |c^{(|Q|)}| \, \mathrm{d}\mu_n^{|\sigma|}, \tag{8.1}$$

where

$$c^{(k)}(x_1, \dots, x_k) = \alpha^{k-1} \sum_{\pi\in\mathrm{Per}(k)} K(x_1, x_{\pi(2)}) K(x_{\pi(2)}, x_{\pi(3)}) \cdots K(x_{\pi(k)}, x_1) \tag{8.2}$$

is the kth cumulant density of the α-DPP and we recall that $\mathrm{Per}(k)$ is the set of all permutations of $1, \dots, k$.

As above, we note that the choice of partitions $\Pi^m(q)$ and $\Pi(q;\sigma)$ induces a graph on $\{1, \dots, |\sigma|\}$. However, in contrast to the Poisson setting, an edge in this graph can now have two different meanings, and therefore we refer to them as type I or type II edges. First, an edge of type I is put between nodes i and j if the product of the functions f_n imposes an edge. This is just as in Theorems 2.1 and 2.2. However, secondly, an edge of type II is now also if i and j are members of the same element in $\rho \in \Pi(q;\sigma)$. This is new in the determinantal case. Then the indecomposability assumption means that the so-constructed graph is connected.

Proof of Theorem 2.3. First we bound the number of partitions involved in the double sum in (8.1). From [20, Proposition 6.1] we have $|\Pi^m(q)| \le q^{qm}(m!)^q$. To bound the sum over the partitions $\Pi(q;\sigma)$, we observe from (8.2) that when expanding the kth-order cumulant measure, we obtain a sum consisting of $(m-1)! \le m!$ summands. For $|\rho| = r$, let $\rho_1, \dots, \rho_r$ denote the r blocks of ρ. Hence, given $\sigma \in \Pi^m(q)$, we obtain the bound

$$\begin{aligned}
\sum_{\rho\in\Pi(q;\sigma)} \prod_{Q\in\rho} |\mathrm{Per}(|Q|)| &= \sum_{r=1}^{|\sigma|} \sum_{\substack{\rho\in\Pi(q;\sigma)\\ |\rho|=r}} |\rho_1|! \cdots |\rho_r|! \\
&\le \sum_{r=1}^{|\sigma|} \frac{2^{|\sigma|-1}|\sigma|!}{r!} \\
&\le 2^{|\sigma|}|\sigma|! \\
&\le (2q)^{qm}(m!)^q,
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

where the upper bound of the inner sum holds by [4, Lemma 3.5] and we have used (6.3) in the last step. Since $|\alpha| \le 1$, this gives that $|\operatorname{cum}(S_n(G))|$ is bounded by

$$(2q^2)^{qm}(m!)^{2q} \sup_{\substack{\sigma\in\Pi^m(q)\\ \rho\in\Pi(q;\sigma)}} \int_{\mathbb{X}^{|\sigma|}} (f_n^{\otimes m})_\sigma \prod_{Q\in\rho} \sup_{\pi\in\mathrm{Per}(|Q|)} |k_{1,\pi(2)}k_{\pi(2),\pi(3)}\cdots k_{\pi(|Q|),1}|\, \mu_n^{|\sigma|}(\mathrm{d}\mathbf{x}), \tag{8.3}$$

where $k_{i,j} := K(x_i, x_j)$ for $i, j = 1, \dots, |Q|$.

Hence it remains to bound the individual integrands in (8.3), which consist of a product of the functions $f(\cdot)$ and products of the kernels $K(\cdot,\cdot)$.

We construct a spanning tree such that for points x_i, x_j in the same element of the partition ρ, we place a type II edge (x_i, x_j) in the spanning tree if the factor $K(x_i, x_j)$ is present in the integral. We put type I edges in the spanning tree that connect the different $Q \in \rho$. To bound the integral (8.3), we integrate over the edges of the spanning tree. If we integrate over a type I edge $((x, u), (y, v))$, then we obtain a contribution of $K_0(o)v_n\kappa(u, v)$. This is analogous to the proof of Theorem 2.1. Second, each time we integrate over a type II edge, we obtain a contribution bounded by $\|K_0\|_1$. This follows from the integrability of the covariance kernel K. Let $E_{s,1}$ denote the set of type I in the spanning tree and let $E_{s,2}$ denote the set of type II edges in the spanning tree, respectively. Since $\sup_{x,y\in\mathbb{R}^d}|K(x, y)| \le K_0(o) \le 1$, we deduce from (6.4) the bound

$$\begin{aligned}
&\mathbb{E}\left[\int_{W_n^{|\sigma|}} \prod_{(i,j)\in E_{s,1}} \varphi_n(x_i - x_j, U_i, U_j) \prod_{(i,j)\in E_{s,2}} |K(x_i, x_j)|\, \mathrm{d}(x_1, \dots, x_{|\sigma|})\right]\\
&\quad\le |W_n| v_n^{|E_{s,1}|}\mathbb{E}\left[\prod_{(i,j)\in E_{s,1}} \kappa(U_i, U_j)\right]\|K_0\|_1^{|E_{s,2}|}\\
&\quad\le |W_n| v_n^{|E_{s,1}|}(m!)^{(q-1)c_{U,2}} b_{\mathrm{SG}}^m \|K_0\|_1^{|E_{s,2}|},
\end{aligned}$$

where b_{SG} is given in Theorem 2.1. Since $q - 1 \le |E_{s,1}| \le m(q - 1)$ and $|E_{s,2}| \le m - 1$, we can in general conclude that

$$|\operatorname{cum}_n(S_n(G))| \le |W_n|(1 \vee v_n)^{m(q-1)}((\|K_0\|_1 \vee 1)(2q^2)^q b_{\mathrm{SG}})^m (m!)^{2q+a}.$$

From here, the assertion follows as in the proof of Theorem 2.1. □

Proof of Theorem 2.4. To prove Theorem 2.4, we combine arguments used in the proofs of Theorems 2.2 and 2.3. Again, we replace the τ-power-weighted edge length $\|X_i - X_j\|^\tau$ by $|B_{\|X_i - X_j\|}|^\tau$ to ease the presentation. For $m \ge 1$, let $\sigma \in \Pi^m(2)$ and $\rho \in \Pi(2;\sigma)$. Let (V, E) be the connected graph induced by the pair (σ, ρ) on the vertex set $V = \{1, \dots, |\sigma|\}$. As above, we can partition E into a set E_1 of type I edges and a set E_2 of type II edges. For $(i, j) \in E_1$, let m_{ij} be the multiplicity of (i, j) in (V, E) and note that $\sum_{(i,j)\in E_1} m_{ij} = m$. Let $E_{s,2} \subseteq E_2$ be such that the induced subgraph of (V, E) with edge set $E_1 \cup E_{s,2}$ is connected. Since $\sup_{x,y\in\mathbb{R}^d}|K(x, y)| \le K_0(o) \le 1$ for all $x, y \in W_n$, we obtain analogously to (8.3) that $|\operatorname{cum}_m(S_n^{(\tau d)})|$ is bounded by

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

$$64^m(m!)^4\sup_{\sigma\in\Pi^m(q)}\sup_{\rho\in\Pi(q;\sigma)}\mathbb{E}\Bigg[\int_{W_n^{|\sigma|}}\prod_{(i,j)\in E_1}\varphi_n(x_i-x_j,U_i,U_j)|B_{\|x_i-x_j\|}|^\tau \times\prod_{(i,j)\in E_{s,2}}|K(x_i,x_j)|\,\mathrm{d}(x_1,\dots,x_{|\sigma|})\Bigg].$$

Now we fix some $E_{s,1}\subseteq E_1$ such that $E_{s,1}\cup E_{s,2}$ is the edge set of a spanning tree of (V,E). Then we find as in the proof of Theorem 2.2 that $|\operatorname{cum}_m(S_n^{(\tau d)})|$ is bounded by

$$64^m(m!)^4\mathbb{E}\Bigg[\prod_{(i,j)\in E_1\setminus E_{s,1}}(\nu_n\kappa(U_i,U_j))^{\tau m_{ij}}M'_\varphi(\tau m_{ij}) \times\int_{W_n^{|\sigma|}}\prod_{(i,j)\in E_{s,1}}\varphi_n(x_i-x_j,U_i,U_j)|B_{\|x_i-x_j\|}|^\tau\prod_{(i,j)\in E_{s,2}}|K(x_i,x_j)|\,\mathrm{d}(x_1,\dots,x_{|\sigma|})\Bigg], \tag{8.4}$$

where M'_φ is given at (MPHI). Next, we bound the integral over $x_1,\dots,x_{|\sigma|}$. Again, successively integrating out the contributions of the leaves implies that the expectation above is bounded by

$$|W_n|\|K_0\|_1^{|E_{s,2}|}\prod_{(i,j)\in E_{s,1}}\big[(\nu_n\kappa(U_i,U_j))^{\tau m_{ij}+1}M_\varphi(\tau m_{ij})\big].$$

Now assumption (MPHI) and $\sum_{(i,j)\in E_1}m_{ij}=m$ give that

$$\prod_{(i,j)\in E_1\setminus E_{s,1}}M'_\varphi(\tau m_{ij})\prod_{(i,j)\in E_{s,1}}M_\varphi(\tau m_{ij})\le c_{\varphi,1}^{\tau m}\Gamma(1+\tau m)^{c_{\varphi,2}}.$$

Moreover, analogously to the proof of Theorem 2.1, by the Hölder inequality,

$$\mathbb{E}\Bigg[\prod_{(i,j)\in E_{s,1}}(\nu_n\kappa(U_i,U_j))^{\tau m_{ij}+1}\prod_{(i,j)\in E_1\setminus E_{s,1}}(\nu_n\kappa(U_i,U_j))^{\tau m_{ij}}\Bigg] \le\nu_n^{\tau m+|E_{s,1}|}\mathbb{E}\big[\kappa(U_1,U_2)^{(\tau+1)m}\big].$$

Thus (8.4) is bounded by

$$64^m(m!)^4|W_n|\|K_0\|_1^{|E_{s,2}|}c_{\varphi,1}^{\tau m}\Gamma(1+\tau m)^{c_{\varphi,2}}\nu_n^{\tau m+|E_{s,1}|}\mathbb{E}\big[\kappa(U_1,U_2)^{(\tau+1)m}\big].$$

Using

$$\mathbb{E}\big[\kappa(U_1,U_2)^{(\tau+1)m}\big]^{1/2}\le(c_{U,1}(\tau+1)^{c_{U,2}})^{(\tau+1)m}(m!)^{c_{U,2}(\tau+1)},$$

by (7.2) and the bounds $q-1\le|E_{s,1}|\le m(q-1)$ and $|E_{s,2}|\le m-1$, we arrive at the estimate

$$|\operatorname{cum}_m(S_n^{(\tau d)})|\le|W_n|((1\vee\|K_0\|_1)64b_{\mathrm{EL}})^m\nu_n^{\tau m}(1\vee\nu_n)^{m(q-1)}(m!)^{4+c_{\varphi,2}\tau+2c_{U,2}(\tau+1)},$$

where

$$b_{\mathrm{EL}}:=(c_{\varphi,1}(\tau\vee1)^{c_{\varphi,2}})^\tau(c_{U,1}(\tau+1)^{c_{U,2}})^{2(\tau+1)}.$$

Now the claim follows as in the proofs above. □

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

Acknowledgements

We would like to thank Friedrich Götze for pointing us to the recent work [5]. We are also grateful to two anonymous reviewers for their thoughtful comments which improved our article.

Funding information

This work was supported by a visit grant by the Danish Data Science Academy (DDSA-V-2023-009), the VILLUM FONDEN (VIL69126), and by the NWO Gravitation project NETWORKS (024.002.003).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Błaszczyszyn, B., Yogeshwaran, D. and Yukich, J. E. (2019). Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Prob.* **47**, 835–895.
[2] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.
[3] Döring, H., Jansen, S. and Schubert, K. (2021). The method of cumulants for the normal approximation. *Prob. Surveys.* **19**, 185–270.
[4] Eichelsbacher, P., Raič, M. and Schreiber, T. (2015). Moderate deviations for stabilizing functionals in geometric probability. *Ann. Inst. H. Poincaré Prob. Statist.* **51**, 89–128.
[5] Feng, R., Götze, F. and Yao, D. (2024). Determinantal point processes on spheres: Multivariate linear statistics. *J. Funct. Anal.* **287**, 110493.
[6] Forrester, P. J. and Honner, G. (1999). Exact statistical properties of the zeros of complex random polynomials. *J. Phys. A.* **32**, 2961–2981.
[7] Gracar, P., Grauer, A., Lüchtrath, L. and Mörters, P. (2019). The age-dependent random connection model. *Queueing Systems* **93**, 309–331.
[8] Gracar, P., Heydenreich, M., Mönch, C. and Mörters, P. (2022). Recurrence versus transience for weight-dependent random connection models. *Electron. J. Prob.* **27**, 60.
[9] Heinrich, L. (2016). On the strong Brillinger-mixing property of α-determinantal point processes and some applications. *Appl. Math.* **61**, 443–461.
[10] Heinrich, L. and Klein, S. (2014). Central limit theorems for empirical product densities of stationary point processes. *Stat. Inference Stoch. Process.* **17**, 121–138.
[11] Hough, J. B., Krishnapur, M., Peres, Y. and Virág, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society, Providence.
[12] Jolivet, E. (1981). Central limit theorem and convergence of empirical processes for stationary point processes. In *Point Processes and Queuing Problems*. North-Holland, Amsterdam & New York.
[13] Komjáthy, J. and Lodewijks, B. (2020). Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs. *Stoch. Process. Appl.* **130**, 1309–1367.
[14] Last, G. and Penrose, M. D. (2018). *Lectures on the Poisson Process*. Cambridge University Press, Cambridge.
[15] Liu, Q. and Privault, N. (2024). Normal approximation of subgraph counts in the random-connection model. *Bernoulli* **30**, 3224–3250.
[16] Penrose, M. D. (2003). *Random Geometric Graphs*. Oxford University Press, Oxford.
[17] Penrose, M. D. (2018). Inhomogeneous random graphs, isolated vertices, and Poisson approximation. *J. Appl. Prob.* **55**, 112–136.
[18] Rudzkis, R., Saulis, L. and Statulevičius, V. (1978). A general lemma on probabilities of large deviations. *Litovsk. Mat. Sb.* **18**, 99–116, 217.
[19] Saulis, L. and Statulevičius, V. A. (1991). *Limit Theorems for Large Deviations*. Kluwer Academic, Dordrecht.
[20] Schulte, M. and Thäle, C. (2024). Moderate deviations on Poisson chaos. *Electron. J. Prob.* **29**, 146.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press

[21] SHIRAI, T. AND TAKAHASHI, Y. (2003). Random point fields associated with certain Fredholm determinants I: Fermion, Poisson and Boson point processes. *J. Funct. Anal.* **205**, 414–463.
[22] STATULEVIČIUS, V. (1966). On large deviations. *Z. Wahrscheinlichkeitsth.* **6**, 133–144.
[23] YOGESHWARAN, D. AND ADLER, R. (2015). On the topology of random complexes built over stationary point processes. *Ann. Appl. Prob.* **15**, 3338–3380.

https://doi.org/10.1017/jpr.2025.13 Published online by Cambridge University Press