# Quantitative CLTs for Geometric Statistics of Dependent Marked Point Processes

Tianshu Cong*
*Jilin University*
Aihua Xia†
*University of Melbourne*
J. E. Yukich‡
*Lehigh University*

June 29, 2026

**Abstract**

Given a geometric statistic expressible as a sum of scores which depend on local data, Błaszczyszyn, Yogeshwaran and Yukich [4] established central limit theorems for centered and normalized versions of these statistics, subject to the underlying point process having fast decay of correlations and also subject to a variance growth condition. Building on this, Cong and Xia [10] derived rates of normal approximation, as measured by the Wasserstein distance, for statistics of point processes exhibiting fast decay of dependence. Here we go further and establish weak mixing conditions yielding rates of normal convergence for geometric statistics of marked point processes. The mixing conditions, which are in terms of verifiable geometric criteria, provide rates of normal approximation in the Wasserstein distance for statistics of a large class of point processes with dependent marks. Examples include statistics of determinantal point processes, unevenly spaced time series, continuum percolation, interacting diffusions on spatial random graphs, as well as local $U$-statistics.

*School of Mathematics, Jilin University, JL 130012, China. Email: congtianshu@jlu.edu.cn. Work supported by National Natural Science Foundation of China (Grant No. 12401184), and Jilin Province Postdoctoral Research Start-up Fund.

†School of Mathematics and Statistics, The University of Melbourne, VIC 3010, Australia. E-mail: aihuaxia@unimelb.edu.au. Work supported in part by Australian Research Council Grant No DP220102666.

‡Department of Mathematics, Lehigh University, Bethlehem, PA 18015, USA. E-mail: joseph.yukich@lehigh.edu. Work supported in part by a Simons Research grant.

*Key words and phrases:* Wasserstein distance; Stein's method; fast decay of dependence; determinantal point process; unevenly spaced time series; stabilization; spatial random graphs; interacting diffusions; U-statistics.

*AMS 2020 Subject Classification:* Primary 60F05, Secondary 60D05; 60G55; 62E20; 05C80.

# 1 Introduction and main results

Consider a random spatial model, possibly time-dependent, and driven by an underlying ground point process $\mathcal{P}$ in $\mathbb{R}^d$, which we always take to be stationary. The point process $\mathcal{P}$ may have spatial dependencies and it may carry marks, which may depend on each other and on the ground process; we shall write $\hat{\mathcal{P}}$ to denote such marked point processes. We are interested in the normal approximation of statistics of spatial random models driven by $\hat{\mathcal{P}}$, including basic counting statistics for $\hat{\mathcal{P}}$, statistics of time series driven by $\hat{\mathcal{P}}$, as well as local $U$-statistics and statistics of interacting diffusions on random graphs defined on $\hat{\mathcal{P}}$.

In all cases we assume that the statistics take the form

$$H_\lambda := \sum_{x \in \mathcal{P} \cap \Gamma_\lambda} \xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda),$$

where $\Gamma_\lambda \subset \mathbb{R}^d$, $\lambda \in (0, \infty)$, are observation windows, the marks $M_x$ are random elements in a mark space, and the *score* $\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)$ describes the interaction between a marked point $(x, M_x) \in \hat{\mathcal{P}}$ and the process $\hat{\mathcal{P}}$ itself, with possible additional dependence on $\Gamma_\lambda$. The scores $\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)$, $x \in \mathcal{P}$, are in general spatially dependent. The statistics $H_\lambda$ typically depend on the geometry of the point process $\mathcal{P}$ and the underlying model and hence are called *geometric statistics.* If $\xi \equiv 1$ then $H_\lambda$ reduces to the counting statistic over the window $\Gamma_\lambda$.

The study of geometric statistics has a long history and has roots in the works of Beardwood, Halton and Hammersley [1] and Steele [41]. Bickel and Breiman [3] were among the first to establish a central limit theorem for a geometric statistic, namely they showed that the standardized and centered sum of edge lengths of the $k$-nearest neighbor graph on an i.i.d. random sample converges to a normal random variable as the sample size tends to infinity. In these seminal papers, as well as in the others that followed, the central limit theory for $(H_\lambda)_{\lambda \in (0,\infty)}$ focused on statistics consisting of sums of score functions on Poisson or binomial input, possibly with independent marks, and such that the score functions $\xi$ depended on local data in the sense that they satisfy stopping set stabilization criteria. In this paper we do not restrict to Poisson or binomial input, nor do we restrict to independent marks or to score functions with stopping sets.

When $\mathcal{P}$ has fast decay of correlations and when $\xi$ satisfies stopping set exponential stabilization, then $((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda})_{\lambda\in(0,\infty)}$ satisfies a qualitative central limit theorem as in [4]. For $\mathcal{P}$ satisfying an exponential mixing condition in the total variation distance (known as EDD; see Section 1.4) then convergence rates for $((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda})_{\lambda\in(0,\infty)}$ to the standard normal under the Wasserstein distance were obtained in [10]. The class of point processes was further enlarged in Błaszczyszyn, Yogeshwaran and Yukich [5], allowing qualitative central limit theorems for $((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda})_{\lambda\in(0,\infty)}$ with $\mathcal{P}$ replaced by a marked point process $\hat{\mathcal{P}}$ *in which points and marks may be dependent*, and with $\xi$ required to satisfy a localization condition weaker than stabilization and allowing for long-range interactions, namely with $\xi$ required to be *distributionally close* to a family of short-range scores. Short-range scores are those depending only on data within a fixed deterministic radius.

The purpose of this paper is to establish *quantitative* central limit theorems in this more general set-up, namely for statistics $H_\lambda$ driven by a dependently marked point process $\hat{\mathcal{P}}$, where $\hat{\mathcal{P}}$ has fast decay of dependence, its ground process $\mathcal{P}$ is stationary, and the score functions are closely approximated by a family of short-range scores. Our approach extends Schreiber and Yukich [37], Xia and Yukich [45], and Hirsch, Otto and Svane [20] in three ways: we do not restrict to input which is a rarefied Gibbsian point process, we allow the input to have dependent marks, and we require that the score function satisfy localization criteria which are weaker than the stabilization criteria required in these papers. We extend Cong and Xia [11] by allowing for dependent marks and by relaxing the stabilization conditions on the score function in that paper. Finally, our work shows that some of the qualitative central limit theorems in [5] can be upgraded to quantitative central limit theorems when the ground point process $\mathcal{P}$ is EDD. In the EDD setting for $\mathcal{P}$, we treat a broad class of score functions with fewer Palm requirements than those considered in [5]. The moment conditions on scores are less stringent and require neither the existence of high-order Palm distributions nor integrability of the score with respect to these distributions. On the other hand, the score functions considered here are required to satisfy stabilization criteria with respect to a local estimator, either in a probabilistic sense (cf. Definition 1.10) or in an $L^p$ sense (cf. Definition 1.13). We refer to Section 1.4 for further remarks comparing our conditions with the existing ones in the literature.

We illustrate the versatility of our approach by obtaining quantitative central limit theorems for a range of functionals, some of which fall outside the scope of statistics encountered in classical stochastic geometric models: (i) counting statistics of determinantal point processes with kernels decaying exponentially fast, (ii) statistics of unevenly spaced time series, (iii) statistics of continuum percolation , (iv) statistics of interacting diffusions on spatial random graphs with EDD input, and (v) certain local U-statistics of EDD marked point processes.

We deduce these quantitative central limit theorems from our main finding, which is that when the sums $H_\lambda$ satisfy a mixing condition on well-separated disjoint sets, then one obtains rates of normal approximation for $(H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}$, where the rate is with respect to the Wasserstein distance. The mixing condition stipulates that the covariance of functions of sums of scores on disjoint sets decays rapidly with respect to the distance between the sets, uniformly over a class of test functions and test sets. Mixing is shown to hold whenever $\mathcal{P}$ satisfies EDD and whenever the score $\xi$ is well enough approximated by a family of short-range scores, a geometric condition which is fulfilled whenever $\xi$ satisfies the standard stopping set stabilization criterion or even a weaker criterion whereby $\xi$ is well approximated by a family of short-range scores either in the probabilistic or $L^p$ sense.

## 1.1 Terminology

We consider a ground point process $\mathcal{P}$ on $\mathbb{R}^d$, $d \geq 1$, and take it to be a simple stationary point process with intensity $\mu \in (0, \infty)$. We let $\hat{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, M_x)}$ be a *simple, marked point process* on $\mathbb{R}^d \times \mathbb{M}$, where $\mathbb{M}$ is a measurable space, $\delta$ is the Dirac measure, $(x, M_x)$ are random elements in $\mathbb{R}^d \times \mathbb{M}$, and for any $x \in \mathbb{R}^d$, the set $\{x\} \times \mathbb{M}$ contains at most one element of $\hat{\mathcal{P}}$. Elements $x \in \mathcal{P}$ are referred to as *points*, while $M_x$ and $\mathbb{M}$ are the *marks* and the *mark space*, respectively. We do not require independence of marks. The space $\mathbb{M}$ is general and, for example, could be the space of time-evolving functions, which would be the case when each $M_x, x \in \mathcal{P}$, is a stochastic process representing the state of an interacting particle system or the sample path of an interacting diffusion. The projection of $\hat{\mathcal{P}}$ on $\mathbb{R}^d$ is the ground process $\mathcal{P}$. The marked point process $\hat{\mathcal{P}}$ is a random element in the space $\mathcal{N}_{\mathbb{R}^d \times \mathbb{M}}$ of locally finite subsets of $\mathbb{R}^d \times \mathbb{M}$ endowed with the evaluation $\sigma$-algebra. Let $\Gamma_\lambda := [-0.5\lambda^{1/d}, 0.5\lambda^{1/d}]^d,\ \lambda \in (0, \infty)$, denote observation windows.

A score is a measurable function $\xi : (\mathbb{R}^d \times \mathbb{M}) \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{M}} \times \mathcal{B}(\mathbb{R}^d) \to \mathbb{R}$ where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel $\sigma$-algebra on $\mathbb{R}^d$. The value of $\xi((x, m_x), \hat{\mathcal{X}}, A)$, where $\hat{\mathcal{X}} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{M}}$ and $A \in \mathcal{B}(\mathbb{R}^d)$, is relevant only for $(x, m_x) \in \hat{\mathcal{X}} \cap (A \times \mathbb{M})$ and by convention we set it to be 0 when $(x, m_x) \notin \hat{\mathcal{X}} \cap (A \times \mathbb{M})$. We put

$$H_\lambda := \sum_{x \in \mathcal{P} \cap \Gamma_\lambda} \xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda),\ \ \sigma_\lambda^2 := \operatorname{Var} H_\lambda,\ \ Z_\lambda := \frac{H_\lambda - \mathbb{E}H_\lambda}{\sigma_\lambda}.$$

We wish to establish rates of convergence of $(Z_\lambda)_{\lambda \in (0,\infty)}$ to the normal. When $\xi((x, m_x), \hat{\mathcal{X}}, A)$ does not depend on $A \in \mathcal{B}(\mathbb{R}^d)$, then the analysis of $(H_\lambda)_{\lambda \in (0,\infty)}$ is easier because it is unaffected by the position of $x$ relative to the boundary of $\Gamma_\lambda$. We include the third parameter because, in some applications, the region $\Gamma_\lambda$ may influence the value of the score function as would be the case for statistics of the Voronoi tessellation [11, Section 3.2]. While many of our examples hold for the canonical choices of $\xi$ given by $\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)$ and $\xi((x, M_x), \hat{\mathcal{P}}) := \xi((x, M_x), \hat{\mathcal{P}}, \mathbb{R}^d)$, we shorten our presentation and typically state results for only one of the two scores. When marks are not present, all statements remain valid by viewing the ground process as a marked point process with degenerate marks. In this case, we write the score as $\xi(x, \mathcal{P}, \Gamma_\lambda)$ and all other notation remains unchanged.

For any real-valued random variable $X$ and $p \in [1, \infty)$ we put $\|X\|_p = (\mathbb{E}(|X|^p))^{1/p}$. For any full-dimensional $B \in \mathcal{B}(\mathbb{R}^d)$, let $\mathrm{Vol}(B)$ denote its volume and for $\mathcal{X} \in \mathcal{N}_{\mathbb{R}^d}$, let $\mathrm{card}(\mathcal{X})$ be its cardinality.

**Definition 1.1.** *($p_1$-moment on the ground process) The ground point process $\mathcal{P}$ has a $p_1$-moment, $p_1 \in [1, \infty)$, if*

$$b_1 := b_1(p_1) := \|\mathrm{card}(\mathcal{P} \cap \Gamma_1)\|_{p_1} < \infty. \tag{1.1}$$

**Remark 1.2.** *The moment condition* (1.1) *holds for all $p_1 \in [1, \infty)$ for homogeneous Poisson point processes and for $\alpha-$determinantal point processes whose kernel decays exponentially fast, provided $-1/\alpha \in \mathbb{N}$, as established in Section 2.2.2 and Section 2.1 (Remark (i)) of [4]. For a Gibbs point process with Papangelou intensity bounded above by a locally integrable function $h : \mathbb{R}^d \to \mathbb{R}$, iterating the GNZ equation gives the $n$-th factorial moment $\mathbb{E}[(\mathcal{P}(B))^{(n)}] \leq (\int_B h(x)\,\mathrm{d}x)^n$ for every bounded Borel set $B$; see, e.g., Dereudre [13, Section 2], hence* $\mathrm{card}(\mathcal{P} \cap \Gamma_1)$ *has finite moments of all orders. This includes many commonly used Gibbs point processes; for example, those whose interaction range is smaller than the critical value of continuum percolation of an associated Poisson point process [20, Section 4.1].*

The ground point process $\mathcal{P}$ is of a general nature, and one may require conditions on the higher-order Palm distributions of $\mathcal{P}$ when proving limit theorems. However, the class of score functions satisfying conditions based on higher-order Palm distributions is narrower than the class based on first-order Palm distributions (see the examples in Remark 1.12 (iv) and Remark 3.12). While the first-order Palm distributions of ground point processes $\mathcal{P}$ are often well-studied, their higher-order Palm distributions are less well-known and typically cumbersome. Except for Poisson and related point processes, obtaining closed forms or analyzing the properties of these distributions is often more challenging than estimating higher-order moments of the score function $\xi$ with respect to the first-order Palm distributions of $\mathcal{P}$. Thus, we circumvent the use of higher-order Palm distributions by instead assuming only higher moment conditions on the ground point process $\mathcal{P}$ and on the score function $\xi$ with respect to the first-order Palm distributions; see [10, Lemma 5.5].

For any fixed distinct $x_1, ..., x_q \in \mathbb{R}^d$, we denote by $\mathbb{P}_{x_1,...,x_q}$ the Palm probability given points $x_1, ..., x_q$; likewise $\mathbb{E}_{x_1,...,x_q}$ and $\mathcal{L}_{x_1,...,x_q}$ denote the Palm expectation and Palm distribution, respectively.

**Definition 1.3.** *($p_2$-moment on the score) The score $\xi$ has a $p_2$-moment, $p_2 \in [1, \infty)$, if*

$$b_2 := b_2(p_2) := \sup_{\lambda \in (0,\infty)} \sup_{x \in \Gamma_\lambda} \left( \mathbb{E}_x |\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)|^{p_2} \right)^{1/p_2} < \infty. \tag{1.2}$$

Given $A \in \mathcal{B}(\mathbb{R}^d)$, the sum of the score functions on the point set $\mathcal{P} \cap A$ is

$$H_{\lambda,A} := \sum_{x \in \mathcal{P} \cap A} \xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda), \quad H_\lambda := H_{\lambda,\Gamma_\lambda} \tag{1.3}$$

and its centered version is

$$S_A := S_{\lambda,A} := H_{\lambda,A} - \mathbb{E}H_{\lambda,A}. \tag{1.4}$$

For $A, B \in \mathcal{B}(\mathbb{R}^d)$, let $\operatorname{diam}(A) = \sup\{d(x,y) : \ x, y \in A\}$ be the diameter of $A$, $d(A,B) = \inf\{d(x,y) : \ x \in A, y \in B\}$, $d(x,B) = d(B,x) := d(\{x\}, B)$, where $d$ is the Euclidean distance on $\mathbb{R}^d$. Given $\rho \in (0,\infty)$ let $B_\rho(A) = \{y : \ d(y,A) < \rho\}$ be the $\rho$-enlargement of $A$. Define the class of test functions

$$\mathcal{F}_b := \{g : \mathbb{R} \to \mathbb{R}; \ \|g\|_\infty \vee \|g'\|_\infty := \sup_{x \in \mathbb{R}} |g(x)| \vee \sup_{x \in \mathbb{R}} |g'(x)| \leq 1\}.$$

The following assumption, similar to the exponential decay of dependence in [10, (2.1)], specifies a weak mixing condition on the score sums $\{S_A\}_{A \in \mathcal{B}(\mathbb{R}^d)}$, which, together with moment assumptions on the ground point process and the scores, yields proximity bounds between $(H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}$ and the standard normal in terms of the Wasserstein distance. The mixing condition requires that the covariance of $S_A$ and $f(S_B)$, as well as the covariance of $S_A^2$ and $f(S_B)$, decays fast with respect to the distance between $A$ and $B$, uniformly for $f$ in a class of test functions and uniformly for $A, B$ belonging to a class of test sets.

**Definition 1.4.** *(mixing of score sums) Let $\mathcal{H}_1$ be the class of open hyperrectangles in $\mathbb{R}^d$ with sides at least unit length and with faces parallel to coordinate planes in $\mathbb{R}^d$, and let $\mathcal{H}_2$ be the class of unions of two incident hyperrectangles in $\mathcal{H}_1$. Let $\mathcal{F}_1$ denote the test function class $\mathcal{F}_b$ together with the identity $g(x) = x$ and let $\mathcal{F}_2 = \mathcal{F}_b$. The score sums satisfy exponential mixing if there exists $\lambda_0 \in (0,\infty)$ and positive constants $\{\gamma_i\}_{i=0}^3$ such that for $\lambda \geq \lambda_0$, $j \in \{1,2\}$ and all $g \in \mathcal{F}_j$, $A \in \mathcal{H}_j$, and all sets $B \in \mathcal{B}(\Gamma_\lambda)$ that are the finite union of disjoint hyperrectangles in $\mathcal{H}_1$ satisfying the separation condition $d(A,B) \geq \gamma_3[\log(\operatorname{diam}(A) \vee \operatorname{diam}(B)) \vee 1]$, we have*

$$\begin{aligned} &\left|\mathbb{E}(S_A^j g(S_B)) - \mathbb{E}(S_A^j)\mathbb{E}(g(S_B))\right| = \left|\operatorname{Cov}(S_A^j, g(S_B))\right| \\ &\leq \gamma_1(\operatorname{diam}(A)^{\gamma_0} \vee 1)(\operatorname{diam}(B)^{\gamma_0} \vee 1)\mathrm{e}^{-\gamma_2 d(A,B)} =: U(\boldsymbol{\gamma}, A, B), \end{aligned} \tag{1.5}$$

*where $\boldsymbol{\gamma} := \{\gamma_i\}_{i=0}^3$. The score sums satisfy polynomial mixing if* (1.5) *holds with $U(\boldsymbol{\gamma}, A, B)$ replaced by*

$$U'(\boldsymbol{\gamma}, A, B) := \gamma_1(\operatorname{diam}(A)^{\gamma_0} \vee 1)(\operatorname{diam}(B)^{\gamma_0} \vee 1)d(A,B)^{-\gamma_2}, \tag{1.6}$$

*where the parameters $\boldsymbol{\gamma}$ in* (1.5) *and* (1.6) *are different in general.*

In Section 1.3 we determine conditions on the ground point process $\mathcal{P}$ and on the scores $\xi$ which imply the mixing bounds (1.5) and (1.6).

**Remark 1.5.** *(i) (the covariance conditions* (1.5) *and* (1.6)*) The mixing of score sums implies covariance conditions of a more general form for a larger class of sets as in Lemma 2.4, which plays an essential role in proving Theorem 1.6, our main result. If either* (1.5) *or* (1.6) *hold for all pairs of disjoint sets $A$ and $B$, then there is no need to specify the parameter $\gamma_3$ controlling the separation condition.*

*(ii) (comparison with mixing of score sums without centering) The mixing properties* (1.5) *and* (1.6) *are equivalent, up to increasing the exponent $\gamma_0$ by an additive term $d$, to the corresponding properties obtained by replacing $S_A$ and $S_B$ with $H_{\lambda,A}$ and $H_{\lambda,B}$, respectively. See Lemma 2.5.*

*(iii) (comparison with the $(BL, \psi, \theta)$ dependence condition) The $(BL, \psi, \theta)$ dependence condition is an assumption imposed on covariances involving a specified class of test functions satisfying certain bounded Lipschitz conditions as introduced in the context of random field theory; cf. Bulinski and Shashkin [8, p. 93–95] and the references therein. This assumption is used to obtain a rate of convergence for the normal approximation error of the partial sum of a centered $(BL, \psi, \theta)$-dependent random field under the Kolmogorov distance via Stein's method, see [8, p. 182–188] for details. Our framework requires verifying a mixing property for a smaller class of test sets and a different class of test functions.*

## 1.2 A general quantitative CLT under mixing of score sums

We use the Wasserstein distance to quantify the difference between two real-valued random variables $X_1$ and $X_2$, where

$$d_W(X_1, X_2) := \sup_{h \in \mathcal{F}_{\mathrm{Lip}}} |\mathbb{E}h(X_1) - \mathbb{E}h(X_2)|,$$

where $\mathcal{F}_{\mathrm{Lip}} := \{f : \mathbb{R} \to \mathbb{R}, \quad |f(x) - f(y)| \leq |x - y|, \text{ for all } x, y \in \mathbb{R}\}$. We quantify the $d_W$ distance between $(H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}$ and the standard normal $N(0, 1)$, here denoted by $Z$. The rates of normal approximation in the following general result as well as in the corollaries and applications assume that $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ for some $\nu > 2/3$. We regard the verification of this condition as a separate issue and do not systematically address this.

**Theorem 1.6.** *(mixing of score sums implies a quantitative CLT) Assume that the point process $\mathcal{P}$ and score $\xi$ respectively satisfy the moment assumptions* (1.1) *and* (1.2) *for $p_1$ and $p_2$ satisfying*

$$\frac{p_1 p_2}{p_1 + p_2 - 1} \geq 3.$$

*Assume that* $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some* $\nu > 2/3$.

*(a) If the score sums satisfy exponential mixing* (1.5), *then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left((\log\lambda)^{3d}\lambda^{-1.5\nu+1}\right). \tag{1.7}$$

*(b) If polynomial mixing* (1.6) *holds and*

$$\gamma_2 > \max\left\{\frac{3(-\nu d + 2d + 2\gamma_0)}{3\nu - 2}, \frac{6(-\nu d + d + \gamma_0)}{3\nu - 2} + \gamma_0\right\}, \tag{1.8}$$

*or equivalently*

$$\tau(\gamma_0, \gamma_2) := \frac{3}{2}\left(\frac{2\nu d + 2\gamma_0}{3d + \gamma_2} \vee \frac{\nu d + 2\gamma_0}{3d + \gamma_2 - \gamma_0}\right) < \frac{3}{2}\nu - 1, \tag{1.9}$$

*then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(\gamma_0,\gamma_2)}\right) = o(1). \tag{1.10}$$

**Remark 1.7.** *(i) (the optimality of rates under exponential mixing) It is not uncommon that* $\nu = 1$, *cf. [4]. In this case the* $O((\log\lambda)^{3d}\lambda^{-0.5})$ *bound* (1.7) *is nearly optimal, save for the logarithmic factor. In general one cannot do better than establishing a rate* $O(\lambda^{-0.5})$ *as seen in Section 2.3 of Schulte and Yukich [38], which treats Poisson input.*

*(ii) (the rate of convergence under polynomial mixing) In the case of polynomial mixing, i.e., under condition* (1.6), *as* $\gamma_2 \to \infty$, *the bound* (1.10) *approaches* $O(\lambda^{-1.5\nu+1})$, *which, in general, is unimprovable when* $\nu = 1$.

*(iii) (moment conditions for the ground point process and for the scores) If the ground point process, respectively the score, has a* $p = (3+\delta)$*-moment for some positive* $\delta$, *then the score, respectively ground point process, must have a* $3(p-1)/(p-3)$ *moment. For the point processes discussed in Remark 1.2, one can take* $p_1$ *in Definition 1.1 to be arbitrarily large. Consequently, it suffices that the score satisfies* (1.2) *for some* $p_2 > 3$. *Likewise, if a score has moments of all orders then* $\mathcal{P}$ *need only have a moment of order* $p_1 > 3$. *When compared with [10] this gives normal approximation results holding under slightly more flexible moment assumptions on the point process and the score.*

*(iv) (ground point process with infinite second moment) One may easily construct a simple stationary point process* $\mathcal{P}$ *on* $\mathbb{R}^d$ *such that* $\|\operatorname{card}(\mathcal{P} \cap \Gamma_1)\|_2 = \infty$. *For example, start with a unit intensity Poisson point process* $\Psi$ *on* $\mathbb{R}^d$ *representing locations of parents. Next, independently of other parents, each parent produces* $\eta$ *children with distribution*

$$\mathbb{P}(\eta = i) = \frac{4}{i(i+1)(i+2)}, \ i \in \mathbb{N}.$$

*Assuming the locations of the children of a parent at $x$ are independent and identically distributed within the open ball centered at $x$ with radius $1$, and independently of the children of other parents, then the point process $\mathcal{P}$ formed by the locations of the children of $\Psi$ is a simple stationary point process on $\mathbb{R}^d$ with intensity $2$ and satisfies $\mathbb{E}(\mathrm{card}(\mathcal{P} \cap \Gamma_1)^2) = \infty$.*

## 1.3 Corollaries

We give conditions on the marked point process $\hat{\mathcal{P}}$ and on the score $\xi$ which together ensure mixing of score sums as in Definition 1.4. Our first condition is adapted from [10].

**Definition 1.8.** *(EDD marked point processes) The marked point process $\hat{\mathcal{P}}$ exhibits exponential decay of dependence (EDD) if there exist constants $\theta_0 \in [0, \infty)$, $\theta_i \in (0, \infty)$ for $1 \le i \le 3$, such that for any $A$, $B \in \mathcal{B}(\mathbb{R}^d)$ with $d(A, B) \ge \theta_3[\log(\mathrm{diam}(A) \vee \mathrm{diam}(B)) \vee 1]$, we have*

$$\beta_{A,B} := d_{TV}\left(\mathcal{L}(\hat{\mathcal{P}}|_{A \cup B}), \mathcal{L}(\hat{\mathcal{P}}|_A \cup \tilde{\hat{\mathcal{P}}}|_B)\right) \le \theta_1(\mathrm{diam}(A)^{\theta_0} \vee 1)(\mathrm{diam}(B)^{\theta_0} \vee 1)\mathrm{e}^{-\theta_2 d(A,B)}, \tag{1.11}$$

*where $\tilde{\hat{\mathcal{P}}}$ is an independent copy of $\hat{\mathcal{P}}$. A point process $\mathcal{P}$ is EDD if it satisfies (1.11) with degenerate marks.*

**Remark 1.9.** *By definition, if a marked point process $\hat{\mathcal{P}}$ is EDD, then so is the ground process $\mathcal{P}$. EDD point processes include Gibbs point processes whose interaction range is smaller than the critical value of continuum percolation of an associated Poisson point process as well as determinantal point processes whose kernel satisfies*

$$|K(x, y)| \le C_1 \mathrm{e}^{-C_2 \mathrm{e}^{C_3|x-y|}}, \quad x, y \in \mathbb{R}^d, \tag{1.12}$$

*for some positive constants $C_1, C_2$ and $C_3$, see [20, Appendix A] and [10, Section 3.2]. Disagreement-coupling methods related to those in [20] have recently been extended to marked Euclidean spaces and to certain repulsive pair-potential models in Hirsch, Otto and Svane [21]. The strong spatial mixing theorem of Michelen and Perkins [29] also provides a route to verifying EDD for finite-range repulsive pair-potential Gibbs processes with activity below $\mathrm{e}/\Delta_\phi$. On the determinantal side, if $\alpha$ satisfies $-1/\alpha \in \mathbb{N}$ and the kernel $K$ is Hermitian with $0 \le K \le -1/\alpha$, then $\alpha$-determinantal point processes exist, as shown in Shirai and Takahashi [39, Theorem 1.2 and the proof of Lemma 3.3], and can be interpreted as the superposition of $-1/\alpha$ i.i.d. determinantal point processes with kernel $-\alpha K$. If the kernel satisfies (1.12), the resulting $\alpha$-determinantal point process is still EDD, since it is a finite superposition of i.i.d. EDD point processes [10, Lemma 3.5].*

If the score function is trivial, i.e., $\xi \equiv 1$, and the EDD point process has a $p_1$-moment for some $p_1 > 3$, then $H_\lambda$ is the counting statistic and fulfils the exponential mixing condition (1.5) with

$$\gamma_0 = 3d + \theta_0 \frac{p_1 - 3}{p_1}, \qquad \gamma_2 = \theta_2 \frac{p_1 - 3}{p_1}, \qquad \gamma_3 = \theta_3,$$

and with $\gamma_1$ a finite constant depending only on the moment and EDD constants. This follows by applying Lemmas 2.1 and 2.5 together with Hölder's inequality.

Next we consider conditions on general score functions. One possibility is to require that $\xi$ satisfies exponential stopping set stabilization as in Lachièze-Rey, Schulte and Yukich [24] and [38], but this condition is sometimes restrictive.

Instead, for each $r \in (0, \infty)$ we construct a short-range score function $\hat{\xi}^{[r]}$ based solely on the point pattern within radius $r$, and such that $\hat{\xi}^{[r]}$ approximates $\xi$ with sufficient accuracy, and then use this approximation for quantifying the error in approximating $H_\lambda$ by the corresponding sum of short-range scores. Given a score $\xi$ and $r \in (0, \infty)$, *a short-range estimator*

$$\hat{\xi}^{[r]} : (\mathbb{R}^d \times \mathbb{M}) \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{M}} \times \mathcal{B}(\mathbb{R}^d) \to \mathbb{R}$$

is defined to be a measurable function of the marked point process that depends solely on the input within an open ball of radius $r$ around the location of interest, i.e., for almost all realizations $\hat{\mathcal{X}}$ of the marked point process $\hat{\mathcal{P}}$, for all locally finite sets $\hat{\mathcal{Y}} \subset (\mathbb{R}^d \backslash B(x,r)) \times \mathbb{M}$, and for all $(x, m_x) \in \hat{\mathcal{X}}$ the following holds:

$$\hat{\xi}^{[r]} \left( (x, m_x), \left[ \hat{\mathcal{X}} \cap (B(x,r) \times \mathbb{M}) \right] \cup \hat{\mathcal{Y}}, \Gamma_\lambda \right) = \hat{\xi}^{[r]} \left( (x, m_x), \hat{\mathcal{X}} \cap (B(x,r) \times \mathbb{M}), \Gamma_\lambda \right). \tag{1.13}$$

In other words $\hat{\xi}^{[r]}$ *has a stopping set which is an open ball of radius* $r$. The restricted score

$$\hat{\xi}^{[r]} \left( (x,m), \hat{\mathcal{X}}, \Gamma_\lambda \right) := \xi \left( (x,m), \hat{\mathcal{X}} \cap (B(x,r) \times \mathbb{M}), \Gamma_\lambda \right), \;\; r \in (0, \infty), \tag{1.14}$$

is a special case of the short-range estimator. We emphasize that while we could choose $\hat{\xi}^{[r]}$ to be the restriction of $\xi$ to the radius $r$ ball as in (1.14), it is advantageous to allow for other choices of $\hat{\xi}^{[r]}$, as when considering statistics of continuum percolation models as in Lo and Xia [27] and statistics of Laguerre tessellations as in Trauthwein and Yukich [43]. Short-range estimators $\hat{\xi}^{[r]}$ are used in [5] where $\hat{\xi}^{[r]}$ denotes the restricted score and it plays a central role in establishing optimal rates of normal approximation on the Poisson space in [43]. The benefits are also discussed in Section 3.4.

**Definition 1.10.** *(probabilistic stabilization) The score $\xi$ satisfies exponential probabilistic stabilization if there is a family of short-range scores $(\hat{\xi}^{[r]})_{r \in (0,\infty)}$ and positive constants $C_1$ and $C_2$ such that*

$$\varphi^{(Pr)}(r) := \sup_{\lambda \in (0,\infty)} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x \left( \xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \right)$$

$$\leq C_1 \exp(-C_2 r), \quad r \in (0,\infty). \tag{1.15}$$

*If there exist positive constants $C$ and $\beta$ such that $\varphi^{(Pr)}(r) \leq Cr^{-\beta}$, $r \in (0,\infty)$, then $\xi$ is said to satisfy polynomial probabilistic stabilization with order $\beta$.*

Probabilistic stabilization of scores is a localization criterion similar to, but weaker than, the standard notion of stopping-set stabilization. It does not require the existence of a stabilizing stopping set, nor does it require comparing the restricted version in (1.14) with the original score function. Instead, it only requires the existence of an estimator based on local information that is sufficiently close to the original score function; see Section 1.4 (i).

As soon as the scores $\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)$ have enough short-range structure, in the sense that they are close to a family of short-range scores $\hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)$ as in Definition 1.10, uniformly over all $x \in \Gamma_\lambda, r \in (0,\infty)$, then $H_\lambda$ is well-approximated by the normal, provided $\hat{\mathcal{P}}$ is EDD.

**Corollary 1.11.** *(rates of normal convergence for sums of scores satisfying probabilistic stabilization) Assume that $\mathcal{P}$ satisfies the $p_1$-moment condition* (1.1) *and that the marked point process $\hat{\mathcal{P}}$ satisfies exponential decay of dependence (1.11). Assume that the scores satisfy the $p_2$-moment condition* (1.2) *with*

$$\underline{p}_0 := \frac{p_1 p_2}{p_1 + p_2 - 1} \geq 3.$$

*Lastly, assume* $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some* $\nu > 2/3$.

*(a) If $\xi$ satisfies exponential probabilistic stabilization* (1.15) *for the given family of short-range scores $(\hat{\xi}^{[r]})_{r \in (0,\infty)}$ then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left((\log \lambda)^{3d} \lambda^{-1.5\nu+1}\right).$$

*(b) If $\xi$ satisfies polynomial probabilistic stabilization for the given family of short-range scores $(\hat{\xi}^{[r]})_{r \in (0,\infty)}$ then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(\gamma_0,\gamma_2)}\right),$$

*with $\gamma_2 = \beta(1 - 2(p_1 + p_2 - 1)/(p_1 p_2)) = \beta(1 - 2/\underline{p}_0) =: \beta \underline{p}_1$, $\gamma_0 = d(2 + \underline{p}_1)$ and, as in* (1.9), *$\tau(\gamma_0, \gamma_2) = \frac{3}{2}\left(\frac{2\nu d + 2\gamma_0}{3d + \gamma_2} \vee \frac{\nu d + 2\gamma_0}{3d + \gamma_2 - \gamma_0}\right)$.*

**Remark 1.12.** *(i) (on the rates) The assertions of parts (ii) and (iii) of Remark 1.7 also apply to Corollary 1.11. If the score function has a finite dependence range, then it satisfies exponential probabilistic stabilization: indeed, taking $\hat{\xi}^{[r]} = \xi$ for* *all $r$ larger than the diameter of the dependence range gives $\varphi^{(Pr)}(r) = 0$ for all sufficiently large $r$, and hence the result in part (a) applies. Moreover, if $\varphi^{(Pr)}$ in Definition 1.10 decays sub-exponentially, i.e., $\varphi^{(Pr)}(r) \leq C_\beta r^{-\beta}$ for all positive $\beta$ with some positive constant $C_\beta$, then $d_W((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}, Z) = O(\lambda^{-p})$ for all $p < 3\nu/2 - 1$.*

*(ii) (on the moments) The trade-off between a $p = (3+\delta)$-moment on the ground point process, respectively score, and a $3(p-1)/(p-3)$ moment on the score, respectively ground point process, as described in Remark 1.7 (iii), applies here as well.*

*(iii) (comparison with bounded-Lipschitz localization and the results based on it) Another localization assumption, called* bounded-Lipschitz localization (BL-localization), *is used in Sections 4-7 of [5] to establish a* qualitative *central limit theorem for* $(H_\lambda)_{\lambda \in (0,\infty)}$. *This condition requires for every $r \in (0, \infty)$, $q \in \mathbb{N}$, that the $q$-tuples of scores satisfy*

$$d_{BL}\Big(\mathcal{L}_{\mathbf{x}}\Big(\big(\xi((x_i, M_{x_i}), \hat{\mathcal{P}}, \Gamma_\lambda)\big)_{i=1,\ldots,q}\Big), \mathcal{L}_{\mathbf{x}}\Big(\big(\hat{\xi}^{[r]}((x_i, M_{x_i}), \hat{\mathcal{P}}, \Gamma_\lambda)\big)_{i=1,\ldots,q}\Big)\Big) \leq \varphi_q(r), \tag{1.16}$$

*uniformly in $\lambda \in (0, \infty)$ and over all distinct $x_1, \ldots, x_q \in \Gamma_\lambda$. Here, $d_{BL}$ denotes the bounded-Lipschitz metric, for each $q \in \mathbb{N}$, $\mathcal{L}_{\mathbf{x}} := \mathcal{L}_{x_1,\ldots,x_q}$ denotes the conditional law given $x_1, \ldots, x_q \in \mathcal{P}$, and $\varphi_q$ is a fast decreasing function, that is, one which decays faster than any polynomial power. When $q = 1$, the assumption in* (1.16) *is weaker than probabilistic stabilization, since $d_{BL}$ is a weaker metric and the condition only requires closeness in distribution between the score functions and their estimators, rather than closeness in probability, which potentially offers applications to a wider class of scores. On the other hand,* (1.16) *is a more complex requirement than* (1.15)*, as it requires closeness of* vectors of scores *and their localized versions with respect to higher-order Palm distributions (see Remark 1.12 (iv) and Remark 3.12). By contrast, our localization condition only involves the case $q = 1$, but with $d_{BL}$ replaced by the stronger probabilistic assumption. We also require* $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some $\nu > 2/3$, whereas [5] only requires $\nu > 0$.*

*(iv) (on the difference between conditions on higher-order and first-order Palm distributions) In general, when it comes to applications, one wishes to impose assumptions involving Palm distributions of the lowest possible order. We illustrate this through a simple example involving nearest neighbor distances. Though the example is based on exponentially decaying functions $(\varphi_q)_{q \in \mathbb{N}}$ in* (1.16)*, it can be easily adapted to the sub-exponential setting. Let $\mathcal{P}$ be a homogeneous Poisson point process in $\mathbb{R}^d$ with unit intensity. For each point $x \in \mathcal{P}$, define*

$$\xi(x, \mathcal{P}, \Gamma_\lambda) = \xi(x, \mathcal{P}) = \operatorname{card}(\mathcal{P} \cap B(x, [-\ln d(x, \mathcal{P} \backslash \{x\})] \vee 1)).$$

*Then one can easily verify* (1.15) *by setting*

$$\hat{\xi}^{[r]}(x, \mathcal{P}, \Gamma_\lambda) = \operatorname{card}(\mathcal{P} \cap B(x, r \wedge \{[-\ln d(x, \mathcal{P} \backslash \{x\})] \vee 1\})),$$

*since, for $r > 1$,*

$$\begin{aligned}&\mathbb{P}_x\left(\xi(x,\mathcal{P},\Gamma_\lambda) \neq \hat{\xi}^{[r]}(x,\mathcal{P},\Gamma_\lambda)\right) \leq \mathbb{P}_x\left(-\ln d(x,\mathcal{P}\backslash\{x\}) > r\right)\\ &= \mathbb{P}_x\left(d(x,\mathcal{P}\backslash\{x\}) < \mathrm{e}^{-r}\right) = \mathbb{P}\left(\mathrm{card}(\mathcal{P}\cap B(\mathbf{0},\mathrm{e}^{-r})) \geq 1\right) =: \varphi(r),\end{aligned}$$

*and $\varphi(r)$ decays exponentially fast. However, for any given $r$, by taking $x_1$ and $x_2$ sufficiently close to each other, the left-hand side of (1.16) with $q = 2$ can be made arbitrarily close to 1, since the dependence range of $\mathcal{L}_{x_1,x_2}(\xi(x,\mathcal{P},\Gamma_\lambda))$ becomes arbitrarily large as $d(x_1,x_2) \to 0$. This example not only illustrates the difference between assumptions based on first-order Palm distributions and those based on higher-order Palm distributions, but can also be adapted to show that each time assumptions are strengthened from the $q$-th to the $(q+1)$-th order Palm distributions, some potential applications may be lost. This can be achieved by replacing $d(x,\mathcal{P}\backslash\{x\})$ with the distance between the $(q+1)$-th nearest neighbor of $x$ in $\mathcal{P}$ and the set consisting of $x$ together with its first $q$ nearest neighbors.*

In addition to probabilistic stabilization, it is useful to consider stabilization in the $L^p$ sense.

**Definition 1.13.** *($L^{p_3}$ stabilization) Given $p_3 \in (0,\infty)$, $\xi$ satisfies exponential stabilization in the $L^{p_3}$ metric, abbreviated exponential $L^{p_3}$-stabilization, if there is a family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$ and positive constants $C_1 := C_1(p_3)$ and $C_2 := C_2(p_3)$ such that*

$$\begin{aligned}\varphi^{(L^{p_3})}(r) &:= \sup_{\lambda\in(0,\infty)}\sup_{x\in\Gamma_\lambda}\left(\mathbb{E}_x\left|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda) - \hat{\xi}^{[r]}((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)\right|^{p_3}\right)^{1/p_3}\\ &\leq C_1\exp(-C_2 r), \quad r\in(0,\infty).\end{aligned}$$

*If there exist positive constants $C := C(p_3)$ and $\beta := \beta(p_3)$ such that $\varphi^{(L^{p_3})}(r) \leq Cr^{-\beta}, r \in (0,\infty)$, then $\xi$ is said to satisfy polynomial $L^{p_3}$-stabilization with order $\beta$.*

Since the moments of the short-range scores cannot, in general, be controlled by (1.2), we impose a separate moment condition on the short-range scores themselves. The family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$ satisfies *a moment condition of order $p_2' \in [1,\infty)$* if

$$b_2' := b_2'(p_2') := \sup_{\lambda\in(0,\infty)}\sup_{x\in\Gamma_\lambda}\sup_{r\in(0,\infty)}\left(\mathbb{E}_x|\hat{\xi}^{[r]}((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{p_2'}\right)^{1/p_2'} < \infty. \tag{1.17}$$

Under suitable moment assumptions (1.2) and (1.17) on score functions and their estimators, the probabilistic stabilization implies $L^p$ stabilization with different parameters, but in some applications (for example, in Section 1.4 (i)) it is more direct to verify probabilistic stabilization instead of $L^p$ stabilization. Also, in some other applications, the $L^p$ stabilization is useful for estimators that are not easily seen to satisfy probabilistic stabilization; see Sections 3.4 and 3.5.

**Corollary 1.14.** *(rates of normal convergence for sums of scores satisfying $L^{p_3}$-stabilization) Assume that $\mathcal{P}$ has a $p_1$-moment as in* (1.1) *and that the marked point process $\hat{\mathcal{P}}$ satisfies exponential decay of dependence (1.11). Assume that the score $\xi$ has a $p_2$-moment as in* (1.2) *and that there is a family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$ which satisfy the $p_2'$-moment condition* (1.17) *with*

$$\underline{p} = \frac{p_1(p_2' \wedge p_2)}{p_1 + (p_2' \wedge p_2) - 1} \geq 3.$$

*Lastly, assume* $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some $\nu > 2/3$.*

*(a) If $\xi$ satisfies exponential $L^{p_3}$-stabilization for the given family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$, with $p_3 \geq \underline{p}/(\underline{p}-2)$, then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left((\log\lambda)^{3d}\lambda^{-1.5\nu+1}\right).$$

*(b) If $\xi$ satisfies polynomial $L^{p_3}$-stabilization for the given family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$, with $p_3 \geq \underline{p}/(\underline{p}-2)$, then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(2d,\beta)}\right),$$

*where* $\tau(2d,\beta) := \frac{3d}{2}\left(\frac{2\nu+4}{3d+\beta} \vee \frac{\nu+4}{d+\beta}\right)$.

**Remark 1.15.** *(i) If* $\operatorname{card}(\mathcal{P} \cap \Gamma_1)$ *has finite moments of all orders, i.e., assumption* (1.1) *holds for all $p_1 \in (0,\infty)$, then it suffices that $p_2 \wedge p_2' > 3$.*

*(ii) For statement (b), the displayed bound converges to* 0 *as $\lambda$ goes to infinity whenever $p_3 \in (0,\infty)$ (instead of $p_3 \geq \underline{p}/(\underline{p}-2)$) provided $\beta$ is sufficiently large (see Corollary 2.7 (b)).*

*(iii) As in (ii), if $d$ and $\nu$ are fixed and polynomial $L^{p_3}$-stabilization holds for some $p_3 \in (0,\infty)$, then the bound for $d_W((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}, Z)$ approaches $O(\lambda^{-1.5\nu+1})$ as $\beta \to \infty$. Furthermore, if $\varphi^{(L^{p_3})}(r)$ in Definition 1.13 decays sub-exponentially for some $p_3 \in (0,\infty)$, then $d_W((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}, Z) = O(\lambda^{-\kappa})$ for all $\kappa < 1.5\nu - 1$.*

*(iv) $L^{p_3}$-stabilization is stronger than the bounded-Lipschitz localization condition* (1.16) *when $q = 1$. However, unlike* (1.16)*, it requires neither estimates for vectors of scores nor estimates under higher-order Palm distributions, cf. Remark 1.12 (iii).*

## 1.4 Comparison with the literature

(i) (Comparison with [10]) As seen below, standard exponential stopping set stabilization for score functions implies exponential probabilistic stabilization and therefore the main results in [10] are a special case of part (a) of Theorem 1.6. Under exponential stopping set stabilization, [10] obtains a rate of convergence with a factor of $(\log \lambda)^{5d}$ instead of the factor $(\log \lambda)^{3d}$ given here. Polynomial stabilization of scores is not considered in [10].

**Definition 1.16.** *(stopping set stabilization) The score $\xi$ is exponentially stabilizing (resp. polynomially stabilizing with order $\beta$) with respect to $\mathcal{L}(\hat{\mathcal{P}})$ if for all $\lambda \in (0, \infty)$, $x \in \Gamma_\lambda$, and almost all realizations $\hat{\mathcal{X}} := \{(x, m_x)\}_{x \in \mathcal{X}}$ of the law $\mathcal{L}_x(\hat{\mathcal{P}})$, where $\mathcal{X}$ is the ground configuration of $\hat{\mathcal{X}}$, there exists a radius of stabilization*

$$R := R^\xi := R((x, m_x), \hat{\mathcal{X}}, \Gamma_\lambda) \in (0, \infty),$$

*such that for all locally finite $\hat{\mathcal{Y}} \subset (\mathbb{R}^d \backslash B(x, R)) \times \mathbb{M}$, we have*

$$R\left((x, m_x), \left[\hat{\mathcal{X}} \cap (B(x, R) \times \mathbb{M})\right] \cup \hat{\mathcal{Y}}, \Gamma_\lambda\right) = R\left((x, m_x), \hat{\mathcal{X}} \cap (B(x, R) \times \mathbb{M}), \Gamma_\lambda\right),$$
$$\xi\left((x, m_x), \left[\hat{\mathcal{X}} \cap (B(x, R) \times \mathbb{M})\right] \cup \hat{\mathcal{Y}}, \Gamma_\lambda\right) = \xi\left((x, m_x), \hat{\mathcal{X}} \cap (B(x, R) \times \mathbb{M}), \Gamma_\lambda\right),$$

*and there exist constants $C_1$ and $C_2$ such that the tail probability*

$$\tau(t) := \sup_{\lambda \in (0,\infty)} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x\left(R((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \geq t\right)$$

*satisfies $\tau(t) \leq C_1 \mathrm{e}^{-C_2 t}$ (respectively $\tau(t) \leq C_1 t^{-\beta}$) for all $t \in (0, \infty)$.*

Next, take the short-range score $\hat{\xi}^{[r]}$ to be the restriction of $\xi$ to the radius $r$ ball, as in (1.14). For this choice of $\hat{\xi}^{[r]}$ it follows that exponential stabilization implies exponential probabilistic stabilization and polynomial stabilization implies polynomial probabilistic stabilization with the same rate, since

$$\mathbb{P}_x\left(\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)\right) \leq \mathbb{P}_x\left(R((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \geq r\right)$$

for all $r \in (0, \infty)$. Hence, if the score stabilizes at a given rate, then it satisfies probabilistic stabilization at the same rate. Together with Corollary 1.11 (b), this establishes a rate of normal convergence when the score function is polynomially stabilizing.

(ii) (Related quantitative CLTs) Given Gibbsian input whose potential satisfies a finite range condition and given a large enough inverse temperature parameter, Theorem 2.3 of [37] establishes $d_K((H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\mathrm{Var}\, H_\lambda}, Z) = O\left((\log \lambda)^{3d} \lambda^{-0.5}\right)$ provided that the variance is of volume order, and where $d_K$ denotes the Kolmogorov distance. Later, this bound was sharpened to $O\left((\log \lambda)^{2d} \lambda^{-0.5}\right)$ for a larger class of Gibbsian input in [20]. When the input $\mathcal{P}$ is a Poisson point process, one can improve the rates of convergence in our main results and establish presumably optimal rates. Notably, if the scores satisfy exponential probabilistic stabilization and a fifth moment condition, then the rates of normal convergence for $(H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\mathrm{Var}\, H_\lambda}$ can be improved to $O(\lambda^{-0.5})$ provided the variance grows like $\Omega(\lambda)$. These Berry-Esseen rates of normal approximation hold in the

Wasserstein and Kolmogorov distances, as shown in [43] and in fact hold under *bounded Lipschitz localization of scores*, where the closeness of scores in the probabilistic sense in Definition 1.10 is replaced by a closeness of scores in the bounded Lipschitz distance.

(iii) (Related qualitative CLT) It is shown in [5] that if the input has fast decay of correlations, if the score has a $(2+\delta)$-moment, and if the score satisfies bounded Lipschitz localization and $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ for some $\nu > 0$ then one obtains a qualitative CLT for $(H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}$, but with no rate of normal convergence.

# 2 Auxiliary results and the proof of the main result

The proof of Theorem 1.6 is preceded by several lemmas.

**Lemma 2.1.** *(easy moment bounds for $\mathcal{P}$) If $\mathcal{P}$ satisfies the $p_1$-moment condition* (1.1)*, then for any $B \in \mathcal{B}(\mathbb{R}^d)$ that can be covered by $L_B$ unit cubes, we have*

$$\|\operatorname{card}(B \cap \mathcal{P})\|_{p_1} \le L_B b_1,$$

*where $b_1$ is defined in* (1.1)*. In particular, if $B$ is the union of $m$ hyperrectangles with side lengths at least 1, then*

$$\|\operatorname{card}(B \cap \mathcal{P})\|_{p_1} \le 2^d m b_1 \operatorname{Vol}(B).$$

*Furthermore, if these $m$ hyperrectangles are disjoint, then*

$$\|\operatorname{card}(B \cap \mathcal{P})\|_{p_1} \le 2^d b_1 \operatorname{Vol}(B).$$

*Also, if $B \in \mathcal{B}(\mathbb{R}^d)$ is a ball with diameter at least 1, then*

$$\|\operatorname{card}(B \cap \mathcal{P})\|_{p_1} \le 4^d \Gamma(0.5d + 1)\pi^{-0.5d} b_1 \operatorname{Vol}(B)$$

*where $\Gamma(s)$ denotes the gamma function $\int_0^\infty t^{s-1}\mathrm{e}^{-t}\mathrm{d}t$ for $s \in (0, \infty)$.*

*Proof.* The first inequality follows from the Minkowski inequality. If $B$ is a hyperrectangle with side lengths $l_i \ge 1$, $1 \le i \le d$, then it can be covered by $L_B := \Pi_{i=1}^d \lceil l_i \rceil \le 2^d \Pi_{i=1}^d l_i = 2^d \operatorname{Vol}(B)$ unit cubes. If $B$ is the union of $m$ such hyperrectangles, say $B_i$, $1 \le i \le m$, then $L_B \le \sum_{i=1}^m L_{B_i} \le 2^d m \operatorname{Vol}(B)$. Furthermore, if these hyperrectangles are disjoint, then the upper bound becomes $L_B \le \sum_{i=1}^m L_{B_i} \le \sum_{i=1}^m (2^d \operatorname{Vol}(B_i)) = 2^d \operatorname{Vol}(B)$, and hence $\|\operatorname{card}(B \cap \mathcal{P})\|_{p_1} \le 2^d b_1 \operatorname{Vol}(B)$. If $B$ is a ball, we can cover it with a cube whose side length equals the diameter of the ball, and the statement follows from the ratio between the volumes of the cube and the ball. $\square$

Recall the definition of $H_{\lambda,A}$ and $S_{\lambda,A}$ in (1.3) and (1.4), respectively, and recall that $\mu$ is the constant intensity of $\mathcal{P}$.

**Lemma 2.2.** *(moment bounds for $H_{\lambda,B}$ and $S_{\lambda,B}$) If $\mathcal{P}$ satisfies the $p_1$-moment condition* (1.1) *and if $\xi$ satisfies the $p_2$-moment condition* (1.2)*, then there exists a constant $C := C(b_1, b_2, \mu)$ independent of $\lambda$ such that for any set $B \in \mathcal{H}_2$,*

$$\|H_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} \le C\,\mathrm{Vol}(B), \tag{2.1}$$

$$\|S_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} \le C\,\mathrm{Vol}(B). \tag{2.2}$$

*Proof.* For $s_1,\ s_2 \in [1,\infty)$, let $t_1,\ t_2$ denote their respective conjugate indices, i.e., $1/s_1 + 1/t_1 = 1$ and $1/s_2 + 1/t_2 = 1$. Repeatedly applying Hölder's inequality, we have

$$\begin{aligned}
\|H_{\lambda,B}\|_{s_1} &\le \left\{\mathbb{E}\left[\left(\sum_{x\in\mathcal{P}\cap B}|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|\right)^{s_1}\right]\right\}^{1/s_1}\\
&\le \left\{\mathbb{E}\left[\sum_{x\in\mathcal{P}\cap B}|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{s_1}\mathrm{card}(B\cap\mathcal{P})^{s_1/t_1}\right]\right\}^{1/s_1}\\
&\le \left\{\mathbb{E}\left[\left(\sum_{x\in\mathcal{P}\cap B}|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{s_1s_2}\right)^{1/s_2}\left(\sum_{x\in\mathcal{P}\cap B}\mathrm{card}(B\cap\mathcal{P})^{s_1t_2/t_1}\right)^{1/t_2}\right]\right\}^{1/s_1}\\
&= \left\{\mathbb{E}\left[\left(\sum_{x\in\mathcal{P}\cap B}|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{s_1s_2}\right)^{1/s_2}\left(\mathrm{card}(B\cap\mathcal{P})^{1+s_1t_2/t_1}\right)^{1/t_2}\right]\right\}^{1/s_1}\\
&\le \left\{\left[\mathbb{E}\sum_{x\in\mathcal{P}\cap B}|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{s_1s_2}\right]^{1/s_2}\left[\mathbb{E}\left(\mathrm{card}(B\cap\mathcal{P})^{1+s_1t_2/t_1}\right)\right]^{1/t_2}\right\}^{1/s_1}\\
&= \left(\int_B \mathbb{E}_x(|\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)|^{s_1s_2})\mu\mathrm{d}x\right)^{1/(s_1s_2)}\|\mathrm{card}(B\cap\mathcal{P})\|_{1+s_1t_2/t_1}^{1-1/(s_1s_2)}.
\end{aligned} \tag{2.3}$$

Setting

$$s_1 = \frac{p_1p_2}{p_1+p_2-1},\ \ s_2 = \frac{p_1+p_2-1}{p_1}$$

and using the $p_2$-moment condition (1.2) and Lemma 2.1 in (2.3), we obtain

$$\|H_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} \le b_2(\mu\,\mathrm{Vol}(B))^{1/p_2}\|\mathrm{card}(B\cap\mathcal{P})\|_{p_1}^{1-1/p_2} \le 0.5C\,\mathrm{Vol}(B),$$

where $b_2$ is defined in (1.2). This implies (2.1). For (2.2), we apply the Minkowski inequality followed by the Lyapunov inequality, which yields

$$\begin{aligned}
&\|S_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} \le \|H_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} + \|\mathbb{E}(H_{\lambda,B})\|_{p_1p_2/(p_1+p_2-1)}\\
&\le 2\|H_{\lambda,B}\|_{p_1p_2/(p_1+p_2-1)} \le C\,\mathrm{Vol}(B).
\end{aligned}$$

□

Using the same idea as in the proof of Lemma 2.2, we establish an upper bound for the $L^p$ norm of

$$H^{[r]}_{\lambda,B} = \sum_{x\in\mathcal{P}\cap B}\left[\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda) - \hat{\xi}^{[r]}\left((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda\right)\right],\ \ B\in\mathcal{H}_2,$$

for a given family of short-range scores $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$.

**Lemma 2.3.** *(moment bounds for $H^{[r]}_{\lambda,B}$) Assume $\mathcal{P}$ has a $p_1$-moment as in* (1.1)*.*

*(a) If $\xi$ satisfies exponential $L^{p_3}$-stabilization for some $p_3 \in (0,\infty)$ and if $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$ is the associated family of short-range scores, then there exist constants $C_1' := C_1'(b_1,\mu)$ and $C_2' := C_2'(b_1,\mu)$ independent of $\lambda$ such that for any $B \in \mathcal{H}_2$ and all $r \in (0,\infty)$*

$$\left\| H^{[r]}_{\lambda,B} \right\|_{p_1p_3/(p_1+p_3-1)} \le C_1' \operatorname{Vol}(B) \mathrm{e}^{-C_2' r}. \tag{2.4}$$

*(b) If $\xi$ satisfies polynomial $L^{p_3}$-stabilization for some $p_3 \in (0,\infty)$ and if $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$ is the associated family of short-range scores, then there exists a constant $C_3' := C_3'(b_1,\mu)$ independent of $\lambda$ such that for any set $B \in \mathcal{H}_2$ and all $r \in (0,\infty)$*

$$\left\| H^{[r]}_{\lambda,B} \right\|_{p_1p_3/(p_1+p_3-1)} \le C_3' \operatorname{Vol}(B) r^{-\beta}. \tag{2.5}$$

*Proof.* The inequalities follow from a step-by-step repetition of the proof of Lemma 2.2, replacing $H_{\lambda,\cdot}$, $\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)$, $p_2$, and $b_2$ with $H^{[r]}_{\lambda,\cdot}$, $\xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda) - \hat{\xi}^{[r]}((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda)$, $p_3$, and $\varphi^{(L^{p_3})}(r)$, respectively. This yields the bound

$$\left\| H^{[r]}_{\lambda,B} \right\|_{p_1p_3/(p_1+p_3-1)} \le \varphi^{(L^{p_3})}(r) (\mu \operatorname{Vol}(B))^{1/p_3} \| \operatorname{card}(B \cap \mathcal{P}) \|_{p_1}^{1-(1/p_3)},$$

which establishes (2.4) and (2.5) under the respective assumptions of exponential and polynomial $L^{p_3}$-stabilization. □

The proof of our main result relies on blocking arguments over big blocks and small blocks, given respectively by the rectangles $A_2$ and $A_1$ in the next lemma, which we will eventually take to have edge lengths proportional to $\log \lambda$ and a constant, respectively. The next lemma is designed to handle error terms arising in these blocking arguments.

**Lemma 2.4.** *(consequence of mixing for the set-function class corresponding to $j = 2$)*

*(a) Assume the exponential mixing condition* (1.5) *holds for the set-function class corresponding to $j = 2$. For all hyperrectangles $A_1$, $A_2$ in $\mathcal{H}_1$ with $A_1 \subseteq A_2$, and for all $B \in \mathcal{B}(\Gamma_\lambda)$ which are finite unions of disjoint hyperrectangles in $\mathcal{H}_1$ with $d(A_2, B) \ge \gamma_3[\log(\operatorname{diam}(A_2) \vee \operatorname{diam}(B)) \vee 1]$, we have for $g \in \mathcal{F}_2$*

$$|\mathbb{E}(S_{A_1}S_{A_2}g(S_B)) - \mathbb{E}(S_{A_1}S_{A_2})\,\mathbb{E}(g(S_B))| \le 1.5(2d+1)U(\boldsymbol{\gamma}, A_2, B). \tag{2.6}$$

*(b) If the polynomial mixing condition* (1.6) *holds for the set-function class corresponding to $j = 2$ then the right-hand side of* (2.6) *can be replaced by* $1.5(2d+1)U'(\boldsymbol{\gamma}, A_2, B)$*.*

*Proof.* We will prove (a) only, as the proof of (b) is nearly a word-for-word repetition. We write $A_{2,0} = A_1$, partition $A_2 \setminus A_1$ into at most $2d$ hyperrectangles incident to $A_1$, and write them as $A_{2,i}$, $1 \le i \le C \le 2d$. Since $S_{A_2} = \sum_{i=0}^{C} S_{A_{2,i}}$, we can rewrite the left-hand side of (2.6) as

$$\begin{aligned}
&\mathbb{E}\left(S_{A_1}S_{A_2}g(S_B)\right) - \mathbb{E}\left(S_{A_1}S_{A_2}\right)\mathbb{E}\left(g(S_B)\right) \\
&= \sum_{i=0}^{C}\left\{\mathbb{E}\left(S_{A_1}S_{A_{2,i}}g(S_B)\right) - \mathbb{E}\left(S_{A_1}S_{A_{2,i}}\right)\mathbb{E}\left(g(S_B)\right)\right\}. \qquad (2.7)
\end{aligned}$$

We write $S_{A_1}S_{A_{2,i}}$ as a sum of squares so that we may apply mixing (1.5) for the set-function class corresponding to $j = 2$. Writing $S_{A_1}S_{A_{2,i}} = 0.5\{S^2_{A_1\cup A_{2,i}} - S^2_{A_1} - S^2_{A_{2,i}}\}$, $1 \le i \le C$, we have

$$\begin{aligned}
&\mathbb{E}(S_{A_1}S_{A_{2,i}}g(S_B)) - \mathbb{E}(S_{A_1}S_{A_{2,i}})\mathbb{E}\left(g(S_B)\right) \\
&= 0.5\left\{\left[\mathbb{E}\left(S^2_{A_1\cup A_{2,i}}g(S_B)\right) - \mathbb{E}(S^2_{A_1\cup A_{2,i}})\mathbb{E}\left(g(S_B)\right)\right]\right. \\
&\quad - \left[\mathbb{E}\left(S^2_{A_1}g(S_B)\right) - \mathbb{E}\left(S^2_{A_1}\right)\mathbb{E}\left(g(S_B)\right)\right] \\
&\quad \left.- \left[\mathbb{E}\left(S^2_{A_{2,i}}g(S_B)\right) - \mathbb{E}\left(S^2_{A_{2,i}}\right)\mathbb{E}\left(g(S_B)\right)\right]\right\}. \qquad (2.8)
\end{aligned}$$

Applying (1.5) when $j = 2$ for $A = A_1,\ A_{2,i}$ and $A_1 \cup A_{2,i}$ respectively, we have

$$\left|\mathbb{E}\left(S^2_{A_1}g(S_B)\right) - \mathbb{E}\left(S^2_{A_1}\right)\mathbb{E}\left(g(S_B)\right)\right| \le U(\boldsymbol{\gamma}, A_1, B), \qquad (2.9)$$

$$\left|\mathbb{E}\left(S^2_{A_{2,i}}g(S_B)\right) - \mathbb{E}\left(S^2_{A_{2,i}}\right)\mathbb{E}\left(g(S_B)\right)\right| \le U(\boldsymbol{\gamma}, A_{2,i}, B), \qquad (2.10)$$

$$\left|\mathbb{E}\left(S^2_{A_1\cup A_{2,i}}g(S_B)\right) - \mathbb{E}\left(S^2_{A_1\cup A_{2,i}}\right)\mathbb{E}\left(g(S_B)\right)\right| \le U(\boldsymbol{\gamma}, A_1\cup A_{2,i}, B). \qquad (2.11)$$

The diameters of $A_1,\ A_{2,i},\ A_1 \cup A_{2,i},\ 1 \le i \le C \le 2d$, are all bounded above by $\mathrm{diam}(A_2)$, and their distances to $B$ are all bounded below by $d(A_2, B)$. Hence we can combine (2.7)-(2.11) to obtain (2.6). □

Mixing properties of centered sums of scores are equivalent to those of the uncentered sums, up to a change in the exponent $\gamma_0$ by $d$.

**Lemma 2.5.** *(mixing for centered and uncentered score sums) If* $\underline{p}_0 = p_1p_2/(p_1+p_2-1) \ge 1$*, then the mixing properties* (1.5) *and* (1.6) *in Definition 1.4 are equivalent, up to increasing the exponent* $\gamma_0$ *by* $d$*, to the corresponding properties obtained by replacing* $S_A$ *and* $S_B$ *with* $H_{\lambda,A}$ *and* $H_{\lambda,B}$*, respectively. More precisely, if the centered version holds with parameter vector* $\boldsymbol{\gamma} = \{\gamma_i\}_{i=0}^3$*, then the uncentered version holds with parameter vector* $\boldsymbol{\gamma}' = \{\gamma_i'\}_{i=0}^3$ *with* $\gamma_0' = \gamma_0 + d$ *and* $\gamma_i' = \gamma_i$ *for* $i = 2, 3$*. Conversely, if the uncentered version holds with parameter vector* $\boldsymbol{\gamma}'$*, then the centered version holds with parameter vector* $\boldsymbol{\gamma}$ *satisfying the same relations, i.e.* $\gamma_0 = \gamma_0' + d$ *and* $\gamma_i = \gamma_i'$ *for* $i = 2, 3$.

*Proof.* We only prove the assertion for exponential mixing (1.5); the proof for polynomial mixing (1.6) is identical, with $U$ replaced by $U'$. We first show that the centered mixing property follows from the corresponding uncentered one. The converse follows by the same argument, and we omit the details.

Assume that the uncentered score sums satisfy (1.5) with parameter $\boldsymbol{\gamma}' = \{\gamma_i'\}_{i=0}^3$. For $j = 1$ and $g(x) = x$, the statement holds trivially, since centering does not change the covariance. For $j = 1$ and $g \in \mathcal{F}_b$, define $g_B(t) := g(t - \mathbb{E}H_{\lambda,B})$, for all $t \in \mathbb{R}$ and $B \in \mathcal{B}\left(\mathbb{R}^d\right)$. Then $g_B \in \mathcal{F}_b$ and $g_B(H_{\lambda,B}) = g(S_B)$. Hence

$$|\mathbb{E}(S_A g(S_B)) - \mathbb{E}(S_A)\mathbb{E}(g(S_B))| = |\mathrm{Cov}(H_{\lambda,A}, g_B(H_{\lambda,B}))| \le U(\boldsymbol{\gamma}', A, B).$$

For $j = 2$ and $g \in \mathcal{F}_b$, from (2.1), we have

$$\begin{aligned}
&\left|\mathbb{E}(S_A^2 g(S_B)) - \mathbb{E}(S_A^2)\mathbb{E}(g(S_B))\right| \\
&= \left|\mathrm{Cov}\left((H_{\lambda,A} - \mathbb{E}H_{\lambda,A})^2, g_B(H_{\lambda,B})\right)\right| \\
&\le \left|\mathrm{Cov}(H_{\lambda,A}^2, g_B(H_{\lambda,B}))\right| + 2|\mathbb{E}H_{\lambda,A}|\, |\mathrm{Cov}(H_{\lambda,A}, g_B(H_{\lambda,B}))| \\
&\le U(\boldsymbol{\gamma}', A, B) + C(\mathrm{diam}(A)^d \vee 1)U(\boldsymbol{\gamma}', A, B) \\
&\le (C+1)(\mathrm{diam}(A)^d \vee 1)U(\boldsymbol{\gamma}', A, B),
\end{aligned}$$

for some positive constant $C$. Thus the centered score sums satisfy (1.5) with the same $\gamma_2, \gamma_3$, and with $\gamma_0$ replaced by $\gamma_0' + d$. This completes the proof. □

Now we may prove Theorem 1.6.

*Proof of Theorem 1.6.* To facilitate the proof, we replace the Euclidean metric on $\mathbb{R}^d$ with the maximum distance metric

$$d_\infty(x, y) = \max_{1\le i\le d} |x_i - y_i|, \quad x, y \in \mathbb{R}^d,$$

so that balls defined in terms of $d_\infty$ are cubes. This change of metric is localized to this proof and elsewhere we keep the Euclidean metric on $\mathbb{R}^d$. For $\rho \in (0, \infty)$ and $A \in \mathcal{B}(\mathbb{R}^d)$, let $B_\infty(A, \rho) = \{y : \ d_\infty(x, y) < \rho \text{ for some } x \in A\}$.

(a) Without loss, we assume $\lambda_0 \ge 1$ and $\sigma_\lambda \ge 2$ for all $\lambda > \lambda_0$. For $\lambda > \lambda_0$, we subdivide $\Gamma_\lambda$ into a maximal collection of disjoint, open, congruent cubes, each with a volume greater than 1. The boundaries of such cubes have zero measure and do not affect the distribution of the sums, so we ignore them throughout the blocking argument. Let $\mathcal{C} := \{\mathbb{C}_1, \dots, \mathbb{C}_{k_\lambda}\}$ be an enumeration of these cubes, where $\mathbb{C}_i$ are cubes with edge length $\lambda^{1/d}/\lfloor\lambda^{1/d}\rfloor$ for all $1 \le i \le k_\lambda$. This partition ensures that the volume of each cube in the partition is in $[1, 2^d)$ for all admissible $\lambda$. Then $k_\lambda = \Omega(\lambda)$. For a given $\rho \in (0, \infty)$, let $N_{i,\lambda,\rho} = B_\infty(\mathbb{C}_i, \rho) \cap \Gamma_\lambda$ and $N_{i,\lambda,2\rho} = B_\infty(\mathbb{C}_i, 2\rho) \cap \Gamma_\lambda$. We have $N_{i,\lambda,\rho} \subset B_\infty(\mathbb{C}_i, \rho)$ and $N_{i,\lambda,2\rho} \subset B_\infty(\mathbb{C}_i, 2\rho)$, and hence the volumes of $N_{i,\lambda,\rho}$ and $N_{i,\lambda,2\rho}$ are bounded by $O(\rho^d)$, $\rho$ large, for all $1 \le i \le k_\lambda$. Let

$$S_{i,\lambda} = S_{\mathbb{C}_i}/\sigma_\lambda, \quad S_{i,\lambda,\rho} = S_{N_{i,\lambda,\rho}}/\sigma_\lambda, \quad S_{i,\lambda,2\rho} = S_{N_{i,\lambda,2\rho}}/\sigma_\lambda.$$

We use Stein's method for normal approximation and modify the argument in the proof of [10, Theorem 2.1] to establish (1.7). To this end, recall the Stein equation for normal approximation from Chen, Goldstein and Shao [9, p. 15]:

$$f'(w) - wf(w) = h(w) - Nh, \tag{2.12}$$

where $Nh := \mathbb{E}h(Z)$. The solution of (2.12) is

$$f_h(w) = \mathrm{e}^{w^2/2}\int_{-\infty}^{w} \mathrm{e}^{-t^2/2}(h(t) - Nh)\mathrm{d}t = -\mathrm{e}^{w^2/2}\int_{w}^{\infty} \mathrm{e}^{-t^2/2}(h(t) - Nh)\mathrm{d}t,$$

and it satisfies the following properties [9, p. 16]:

$$\|f_h\|_\infty \le 2,\ \|f_h'\|_\infty \le \sqrt{\frac{2}{\pi}},\ \|f_h''\|_\infty \le 2.$$

The Stein equation (2.12) transforms the estimate of $d_W(Z_\lambda, Z)$, with $Z_\lambda = (H_\lambda - \mathbb{E}H_\lambda)/\sqrt{\operatorname{Var} H_\lambda}$, into the study of $f'(Z_\lambda) - Z_\lambda f(Z_\lambda)$:

$$d_W(Z_\lambda, Z) = \sup_{h\in\mathcal{F}_{\mathrm{Lip}}} |\mathbb{E}(h(Z_\lambda) - Nh)| \le \sup_{f\in\mathcal{F}} |\mathbb{E}\left(f'(Z_\lambda) - Z_\lambda f(Z_\lambda)\right)|, \tag{2.13}$$

where $\mathcal{F} := \left\{f:\ \mathbb{R}\to\mathbb{R}, \|f\|_\infty \le 2,\ \|f'\|_\infty \le \sqrt{2/\pi},\ \|f''\|_\infty \le 2\right\}$.

The first term in (2.13) satisfies

$$\begin{aligned}
&\mathbb{E}f'(Z_\lambda)\\
&= \mathbb{E}\left(Z_\lambda^2\right)\mathbb{E}f'\left(Z_\lambda\right)\\
&= \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}f'\left(Z_\lambda\right) + \mathbb{E}\left(Z_\lambda^2 - \sum_{i=1}^{k_\lambda} S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}f'(Z_\lambda)\\
&= \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\right)\left[\mathbb{E}f'(Z_\lambda) - \mathbb{E}f'(Z_\lambda - S_{i,\lambda,2\rho}) + \mathbb{E}f'(Z_\lambda - S_{i,\lambda,2\rho})\right] + \epsilon_1\\
&= \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}\int_0^{S_{i,\lambda,2\rho}} f''(Z_\lambda - x)\mathrm{d}x + \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}f'\left(Z_\lambda - S_{i,\lambda,2\rho}\right) + \epsilon_1\\
&= \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}f'(Z_\lambda - S_{i,\lambda,2\rho})\right) + \epsilon_1 + \epsilon_2 + \epsilon_3,
\end{aligned} \tag{2.14}$$

where, keeping in mind that $Z_\lambda = \sum_{i=1}^{k_\lambda} S_{i,\lambda}$, we have

$$\begin{aligned}
\epsilon_1 =& \mathbb{E}\left(Z_\lambda^2 - \sum_{i=1}^{k_\lambda} S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}f'(Z_\lambda) = \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}\left(Z_\lambda - S_{i,\lambda,\rho}\right)\right)\mathbb{E}f'(Z_\lambda);\\
\epsilon_2 =& \sum_{i=1}^{k_\lambda}\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\right)\mathbb{E}\left(\int_0^{S_{i,\lambda,2\rho}} f''(Z_\lambda - x)\mathrm{d}x\right);
\end{aligned}$$

$$\epsilon_3 = -\sum_{i=1}^{k_\lambda} \mathbb{E}\left[S_{i,\lambda}S_{i,\lambda,\rho}\left(f'(Z_\lambda - S_{i,\lambda,2\rho}) - \mathbb{E}f'(Z_\lambda - S_{i,\lambda,2\rho})\right)\right].$$

The terms $\epsilon_1, \epsilon_2, \epsilon_3$ may be bounded by either using exponential mixing conditions to control score sums on sets separated by $\rho$ or by using moment bounds on score sums over regions with diameter proportional to $\rho$. The first approach produces errors which are decreasing with increasing $\rho$, whereas the second produces errors increasing with $\rho$. A good balance is achieved by taking $\rho$ to be proportional to $\log\lambda$. The details go as follows.

For $\epsilon_1$, using the bound $\|f'\|_\infty \le \sqrt{2/\pi}$ and exponential mixing condition (1.5) with $j = 1$, $g(x) = x$, $x \in \mathbb{R}$, $A = \mathbb{C}_i$, $B = \Gamma_\lambda \backslash N_{i,\lambda,\rho}$, $1 \le i \le k_\lambda$, we have

$$\begin{aligned}|\epsilon_1| &\le \sqrt{\frac{2}{\pi}} k_\lambda \sigma_\lambda^{-2}\gamma_1\left(\left(2\sqrt{d}\right)^{\gamma_0} \vee 1\right)\left(\left(\sqrt{d}\lambda^{1/d}\right)^{\gamma_0} \vee 1\right)\mathrm{e}^{-\gamma_2\rho} \\ &= O\left(\sigma_\lambda^{-2}\lambda^{\gamma_0/d+1}\mathrm{e}^{-\gamma_2\rho}\right) = O\left(\lambda^{-1}\right) \qquad (2.15)\end{aligned}$$

for $\rho > C_1 \log\lambda$, where $C_1$ is a positive constant.

For $\epsilon_2$, using $\|f''\|_\infty \le 2$, Lemma 2.2 and Lyapunov's inequality, we have

$$\begin{aligned}|\epsilon_2| \le& 2\sum_{i=1}^{k_\lambda} \mathbb{E}\left(|S_{i,\lambda}S_{i,\lambda,\rho}|\right)\mathbb{E}\left(|S_{i,\lambda,2\rho}|\right) \\ \le& \sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda}^2 + S_{i,\lambda,\rho}^2\right)\mathbb{E}\left(|S_{i,\lambda,2\rho}|\right) \\ \le& \sum_{i=1}^{k_\lambda}\left(\|S_{i,\lambda}\|_3^2 + \|S_{i,\lambda,\rho}\|_3^2\right)\|S_{i,\lambda,2\rho}\|_3 \\ =& O\left(\sigma_\lambda^{-3}k_\lambda\rho^{3d}\right) = O\left(\lambda^{-1.5\nu+1}\rho^{3d}\right). \qquad (2.16)\end{aligned}$$

To bound $\epsilon_3$, as $\|f'\|_\infty \le \sqrt{2/\pi}$, we apply Lemma 2.4 (a) with $g(x) = f'(x/\sigma_\lambda)$ for $x \in \mathbb{R}$, $A_1 = \mathbb{C}_i$, $A_2 = N_{i,\lambda,\rho}$, $B = \Gamma_\lambda \backslash N_{i,\lambda,2\rho}$, $1 \le i \le k_\lambda$, to obtain

$$\begin{aligned}|\epsilon_3| \le& \sum_{i=1}^{k_\lambda}\left|\mathbb{E}\left(S_{i,\lambda}S_{i,\lambda,\rho}\left(f'(Z_\lambda - S_{i,\lambda,2\rho}) - \mathbb{E}f'(Z_\lambda - S_{i,\lambda,2\rho})\right)\right)\right| \\ \le& 1.5k_\lambda\sigma_\lambda^{-2}\gamma_1(2d+1)\left(\left(\sqrt{d}\,(2+2\rho)\right)^{\gamma_0} \vee 1\right)\left(\left(\sqrt{d}\lambda^{1/d}\right)^{\gamma_0} \vee 1\right)\mathrm{e}^{-\gamma_2\rho} \\ =& O\left(\sigma_\lambda^{-2}\rho^{\gamma_0}\lambda^{\gamma_0/d+1}\mathrm{e}^{-\gamma_2\rho}\right) = O\left(\lambda^{-1}\right), \qquad (2.17)\end{aligned}$$

when $\rho \ge C_2 \log\lambda$ for some positive constant $C_2$, since $\|g(x)\|_\infty \le \sqrt{2/\pi} \le 1$, $\|g'(x)\|_\infty = \|f''(x/\sigma_\lambda)\|_\infty/\sigma_\lambda \le 2/\sigma_\lambda \le 1$, and so $g \in \mathcal{F}_b \subset \mathcal{F}_1$.

For the second term in (2.13), we have

$$
\begin{aligned}
\mathbb{E}\left(Z_\lambda f(Z_\lambda)\right) =& \sum_{i=1}^{k_\lambda} \mathbb{E}\left[S_{i,\lambda} f(Z_\lambda)\right] \\
=& \sum_{i=1}^{k_\lambda} \mathbb{E}\left[S_{i,\lambda}\left(f(Z_\lambda) - f(Z_\lambda - S_{i,\lambda,\rho})\right)\right] \\
&+ \sum_{i=1}^{k_\lambda} \mathbb{E}\left[S_{i,\lambda}\left(f(Z_\lambda - S_{i,\lambda,\rho}) - \mathbb{E}\left(f(Z_\lambda - S_{i,\lambda,\rho})\right)\right)\right] \\
&+ \sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda}\right) \mathbb{E} f\left(Z_\lambda - S_{i,\lambda,\rho}\right) \\
=& \sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda} S_{i,\lambda,\rho} \int_0^1 f'(Z_\lambda - u S_{i,\lambda,\rho}) \mathrm{d}u\right) + \epsilon_4 + \epsilon_5, \qquad (2.18)
\end{aligned}
$$

where

$$
\begin{aligned}
\epsilon_4 =& \sum_{i=1}^{k_\lambda} \left\{\mathbb{E}\left[S_{i,\lambda} f(Z_\lambda - S_{i,\lambda,\rho})\right] - \mathbb{E}\left(S_{i,\lambda}\right) \mathbb{E}\left(f(Z_\lambda - S_{i,\lambda,\rho})\right)\right\}; \\
\epsilon_5 =& \sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda}\right) \mathbb{E}\left(f(Z_\lambda - S_{i,\lambda,\rho})\right) = 0. \qquad (2.19)
\end{aligned}
$$

To bound $\epsilon_4$, since $\|f\|_\infty \le 2$, an application of exponential mixing (1.5) with $g(x) = 0.5 f(x/\sigma_\lambda)$ for $x \in \mathbb{R}$, $j = 1$, $A = \mathbb{C}_i$, $B = \Gamma_\lambda \backslash N_{i,\lambda,\rho}$, $1 \le i \le k_\lambda$, yields

$$
\begin{aligned}
|\epsilon_4| \le& 2 k_\lambda \sigma_\lambda^{-1} \gamma_1 \left((2\sqrt{d})^{\gamma_0} \vee 1\right) \left(\left(\sqrt{d}\lambda^{1/d}\right)^{\gamma_0} \vee 1\right) \mathrm{e}^{-\gamma_2 \rho} \\
=& O\left(\sigma_\lambda^{-1} \lambda^{\gamma_0/d+1} \mathrm{e}^{-\gamma_2 \rho}\right) = O\left(\lambda^{-1}\right), \qquad (2.20)
\end{aligned}
$$

for $\rho > C_3 \log \lambda$ with a sufficiently large $C_3$, since $\|g(x)\|_\infty \le 1$,

$$
\|g'(x)\|_\infty = \|0.5 f'(x/\sigma_\lambda)\|_\infty / \sigma_\lambda \le \sqrt{1/2\pi}/\sigma_\lambda \le 1,
$$

so $g \in \mathcal{F}_b \subset \mathcal{F}_1$.

The remainder of the difference between (2.14) and (2.18) is

$$
\begin{aligned}
|\epsilon_6| :=& \left|\sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda} S_{i,\lambda,\rho} f'(Z_\lambda - S_{i,\lambda,2\rho})\right) - \sum_{i=1}^{k_\lambda} \mathbb{E}\left(S_{i,\lambda} S_{i,\lambda,\rho} \int_0^1 f'(Z_\lambda - u S_{i,\lambda,\rho}) \mathrm{d}u\right)\right| \\
=& \sum_{i=1}^{k_\lambda} \left|\mathbb{E}\left(S_{i,\lambda} S_{i,\lambda,\rho} \int_0^1 \int_0^{S_{i,\lambda,2\rho} - u S_{i,\lambda,\rho}} f''(Z_\lambda - S_{i,\lambda,2\rho} + v) \mathrm{d}v \mathrm{d}u\right)\right| \\
\le& 2 \sum_{i=1}^{k_\lambda} \mathbb{E}\left[|S_{i,\lambda} S_{i,\lambda,\rho}| \left(|S_{i,\lambda,2\rho}| + |S_{i,\lambda,\rho}|\right)\right]
\end{aligned}
$$

$$\leq \frac{2}{3}\sum_{i=1}^{k_\lambda}\mathbb{E}\left(2\left|S_{i,\lambda}\right|^3+3\left|S_{i,\lambda,\rho}\right|^3+\left|S_{i,\lambda,2\rho}\right|^3\right), \tag{2.21}$$

where the penultimate inequality follows from the bound $\|f''\|_\infty \leq 2$ and the last inequality is from the inequality of arithmetic and geometric means, that is, $a_1a_2a_3 \leq (a_1^3+a_2^3+a_3^3)/3$ for nonnegative $a_1, a_2, a_3$, applied to $|S_{i,\lambda}S_{i,\lambda,\rho}S_{i,\lambda,2\rho}|$ and $|S_{i,\lambda}|\,S_{i,\lambda,\rho}^2$. Applying Lemma 2.2 in (2.21), we have

$$|\epsilon_6| = O\left(k_\lambda\sigma_\lambda^{-3}\rho^{3d}\right) = O\left(\lambda^{-1.5\nu+1}\rho^{3d}\right). \tag{2.22}$$

Combining (2.13) - (2.22), we obtain

$$d_W(Z_\lambda, Z) = O(\lambda^{-1.5\nu+1}\rho^{3d})$$

for $\rho > C_4\log\lambda$, where $C_4 = \max\{C_1, C_2, C_3\}$, as claimed.

(b) Setting $N'_{i,\lambda,2\rho} = B_\infty(\mathbb{C}_i, (\gamma_3 \vee 1)(1+2\rho)) \cap \Gamma_\lambda$, we still have the volume of $N'_{i,\lambda,2\rho}$ bounded above by $O(\rho^d)$. By repeating the steps of the proof for part (a), with $N_{i,\lambda,2\rho}$ replaced by $N'_{i,\lambda,2\rho}$, we can see that

$$d_W(Z_\lambda, Z) \leq \sum_{i=1}^{6}\epsilon_i, \tag{2.23}$$

for the same $\epsilon_i$'s as in part (a). Let $\rho \geq C_5\log\lambda$ with a sufficiently large $C_5$. By adjusting the bounds in (2.15), (2.17) and (2.20) - replacing the bounds in exponential mixing condition (1.5) with those in (1.6) - along with (2.16), (2.19) and (2.22), we obtain

$$\begin{aligned}
|\epsilon_1| &= O\left(\lambda^{\gamma_0/d+1-\nu}\rho^{-\gamma_2}\right);\\
|\epsilon_2| &= O\left(\lambda^{-1.5\nu+1}\rho^{3d}\right);\\
|\epsilon_3| &= O\left(\lambda^{\gamma_0/d+1-\nu}\rho^{\gamma_0-\gamma_2}\right);\\
|\epsilon_4| &= O\left(\lambda^{\gamma_0/d+1-\nu/2}\rho^{-\gamma_2}\right);\\
\epsilon_5 &= 0;\\
|\epsilon_6| &= O\left(\lambda^{-1.5\nu+1}\rho^{3d}\right).
\end{aligned}$$

Combining the above bounds with (2.23) and by taking

$$\rho = \Omega\left(\lambda^{\frac{\nu+\gamma_0/d}{3d+\gamma_2}\vee\frac{0.5\nu+\gamma_0/d}{3d+\gamma_2-\gamma_0}}\right),$$

we obtain the bound

$$d_W(Z_\lambda, Z) = O\left(\lambda^{-1.5\nu+1+\tau(\gamma_0,\gamma_2)}\right) = o(1),$$

given the assumptions, where, as in (1.9),

$$\tau(\gamma_0,\gamma_2)=\frac{3}{2}\left(\frac{2\nu d+2\gamma_0}{3d+\gamma_2}\vee\frac{\nu d+2\gamma_0}{3d+\gamma_2-\gamma_0}\right).$$

This concludes the proof of Theorem 1.6. □

*Proof of Corollary 1.11.* We show that the conditions of Theorem 1.6 are fulfilled. The moment condition on the ground point process as well as the moment condition on the score are assumed; thus it remains to verify exponential mixing (1.5) of the score sums for part (a) and polynomial mixing (1.6) for part (b). We first prove part (a).

*Step (i). Coupling.* Let $A,B\in\mathcal{B}(\mathbb{R}^d)$ be disjoint sets, not necessarily hyperrectangles, such that $d(A,B)\geq C_1\log(\mathrm{diam}(A)\vee\mathrm{diam}(B)\vee C_2)$ for some positive constants $C_1$ and $C_2$. We put $r=d(A,B)/3$. Recall that $\tilde{\hat{\mathcal{P}}}$ is an independent copy of $\hat{\mathcal{P}}$. The total variation bound (1.11) characterizing EDD point processes implies that we can construct a *maximal coupling* $(\Psi_1,\Psi_2)$ and $(\Psi_3,\Psi_4)$ in terms of the total variation distance between

$$(\hat{\mathcal{P}}|_{B_r(A)},\hat{\mathcal{P}}|_{B_r(B)})\text{ and }(\hat{\mathcal{P}}|_{B_r(A)},\tilde{\hat{\mathcal{P}}}|_{B_r(B)}).$$

More precisely, the marked point processes $(\Psi_1,\Psi_2)$ and $(\Psi_3,\Psi_4)$ are random elements defined on the same probability space, taking values in $\mathcal{N}_{B_r(A)\times\mathbb{M}}\times\mathcal{N}_{B_r(B)\times\mathbb{M}}$, equipped with the $\sigma$-algebra generated by the vague topology, such that

$$\begin{aligned}(\Psi_1,\Psi_2)&\stackrel{d}{=}(\hat{\mathcal{P}}|_{B_r(A)},\hat{\mathcal{P}}|_{B_r(B)}),\\(\Psi_3,\Psi_4)&\stackrel{d}{=}(\hat{\mathcal{P}}|_{B_r(A)},\tilde{\hat{\mathcal{P}}}|_{B_r(B)}),\end{aligned}$$

and by (1.11) we have

$$\mathbb{P}(E_1):=\mathbb{P}\left((\Psi_1,\Psi_2)\neq(\Psi_3,\Psi_4)\right)\leq C_3(\mathrm{diam}(A)^{C_4}\vee 1)(\mathrm{diam}(B)^{C_4}\vee 1)\mathrm{e}^{-C_5d(A,B)}\quad(2.24)$$

for some positive constants $C_3$, $C_5$, and a non-negative constant $C_4$. From the coupling $(\Psi_3,\Psi_4)$, we can construct two independent marked point processes $\Psi_5$ and $\Psi_6$, each with the same distribution as $\hat{\mathcal{P}}$ such that $\Psi_5|_{B_r(A)}=\Psi_3$, $\Psi_6|_{B_r(B)}=\Psi_4$. Let $\bar{\Psi}_5$ and $\bar{\Psi}_6$ denote their respective ground processes. Define

$$\begin{aligned}H'_{\lambda,A}&:=\sum_{x\in\bar{\Psi}_5\cap A}\xi((x,M_{5,x}),\Psi_5,\Gamma_\lambda),\ \ S'_A:=H'_{\lambda,A}-\mathbb{E}(H'_{\lambda,A}),\\H'_{\lambda,B}&:=\sum_{x\in\bar{\Psi}_6\cap B}\xi((x,M_{6,x}),\Psi_6,\Gamma_\lambda),\ \ S'_B:=H'_{\lambda,B}-\mathbb{E}(H'_{\lambda,B}).\end{aligned}$$

*Step (ii). Approximating score sums by sums of short-range scores.* Recall $r=d(A,B)/3$. We use this value of $r$ to construct short-range versions of $\bar{S}'_A$ and $\bar{S}'_B$ given by

$$\bar{S}_A^{[r]}:=\sum_{x\in\bar{\Psi}_5\cap A}\hat{\xi}^{[r]}((x,M_{5,x}),\Psi_5,\Gamma_\lambda)-\mathbb{E}(H'_{\lambda,A}),\quad(2.25)$$

$$\bar{S}_B^{[r]} := \sum_{x \in \bar{\Psi}_6 \cap B} \hat{\xi}^{[r]}((x, M_{6,x}), \Psi_6, \Gamma_\lambda) - \mathbb{E}(H'_{\lambda,B}). \tag{2.26}$$

By construction, we have

$$S'_A \stackrel{d}{=} S_A,\ S'_B \stackrel{d}{=} S_B,\ \bar{S}_A^{[r]} \in \sigma(\Psi_5|_{B_r(A)}) = \sigma(\Psi_3), \bar{S}_B^{[r]} \in \sigma(\Psi_6|_{B_r(B)}) = \sigma(\Psi_4),$$

and

$$S'_A \perp\!\!\!\perp S'_B,\ \bar{S}_A^{[r]} \perp\!\!\!\perp \bar{S}_B^{[r]},$$

where $\perp\!\!\!\perp$ denotes independence.

Let $E_2 := \{S'_A \neq \bar{S}_A^{[r]}\}$ and $E_3 := \{S'_B \neq \bar{S}_B^{[r]}\}$. We compute

$$\begin{aligned}
\mathbb{P}(E_2) &\le \mathbb{E} \int_{x \in A} \mathbf{1}_{[\xi((x, M_{5,x}), \Psi_5, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_{5,x}), \Psi_5), \Gamma_\lambda)]} \bar{\Psi}_5(\mathrm{d}x) \\
&= \int_{x \in A} \mathbb{P}_x(\xi((x, M_{5,x}), \Psi_5, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_{5,x}), \Psi_5, \Gamma_\lambda)) \mu \mathrm{d}x \\
&\le \mu \varphi^{(Pr)}(r) \operatorname{Vol}(A),
\end{aligned} \tag{2.27}$$

and

$$\begin{aligned}
\mathbb{P}(E_3) &\le \mathbb{E} \int_{x \in B} \mathbf{1}_{[\xi((x, M_{6,x}), \Psi_6, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_{6,x}), \Psi_6, \Gamma_\lambda)]} \bar{\Psi}_6(\mathrm{d}x) \\
&= \int_{x \in B} \mathbb{P}_x(\xi((x, M_{6,x}), \Psi_6, \Gamma_\lambda) \neq \hat{\xi}^{[r]}((x, M_{6,x}), \Psi_6, \Gamma_\lambda)) \mu \mathrm{d}x \\
&\le \mu \varphi^{(Pr)}(r) \operatorname{Vol}(B).
\end{aligned} \tag{2.28}$$

*Step (iii). Using short-range score sums to show $S_A = S'_A$ on a high probability event.* Let $E = (E_1 \cup E_2 \cup E_3)^c$. On the coupling space, $S_A$ and $S_B$ are realized on $\hat{\mathcal{P}}$, whose restrictions to $B_r(A)$ and $B_r(B)$ are $\Psi_1$ and $\Psi_2$, respectively. Since the localized estimators depend only on these restrictions, on $E_1^c$, the localized sums from the two coupled copies coincide. Combining (1.15), (2.24), (2.27) and (2.28), we find on the event $E$,

$$S_A = S'_A = \bar{S}_A^{[r]},\ S_B = S'_B = \bar{S}_B^{[r]}.$$

Recalling that $\varphi^{(Pr)}(\cdot)$ decays exponentially fast and recalling $r = d(A, B)/3$, the probability of the complement satisfies

$$\mathbb{P}(E^c) \le C_6(\operatorname{diam}(A)^{C_7} \vee 1)(\operatorname{diam}(B)^{C_7} \vee 1)\mathrm{e}^{-C_8 d(A,B)}$$

for some positive constants $C_6$, $C_7$ and $C_8$. Here, $\bar{S}_A^{[r]}$ and $\bar{S}_B^{[r]}$ are only needed to construct a coupling between $(S_A, S_B)$ and $(S'_A, S'_B)$.

*Step (iv). Conclusion.* Applying Hölder's inequality and Lemma 2.2, we obtain for all $j \in \{1, 2\}$ and $g \in \mathcal{F}_j$, $A \in \mathcal{H}_j$, and all admissible $B \subset \Gamma_\lambda$ satisfying the separation condition in Definition 1.4,

$$\left|\mathbb{E}(S_A^j g(S_B)) - \mathbb{E}(S_A^j)\mathbb{E}\left(g(S_B)\right)\right|$$

$$
\begin{aligned}
&= \left|\mathbb{E}\left((S'_A)^j g(S'_B)\right) - \mathbb{E}\left((S'_A)^j g(S'_B)\mathbf{1}_{E^c}\right) + \mathbb{E}\left(S_A^j g(S_B)\mathbf{1}_{E^c}\right) - \mathbb{E}\left((S'_A)^j\right)\mathbb{E}\left(g(S'_B)\right)\right| \\
&\le \left|\mathbb{E}\left(S_A^j g(S_B)\mathbf{1}_{E^c}\right)\right| + \left|\mathbb{E}\left((S'_A)^j g(S'_B)\mathbf{1}_{E^c}\right)\right| && (2.29)\\
&\le U(\boldsymbol{\gamma}, A, B) && (2.30)
\end{aligned}
$$

for some $\boldsymbol{\gamma}$, which confirms that exponential mixing (1.5) holds. Hence, the result follows by direct application of Theorem 1.6 (a).

(b) This is a step-by-step repetition of the proof of (a), with the primary modification being the bounds in (2.27) and (2.28), which, using $\varphi^{(Pr)}(r) \le Cr^{-\beta}$, now take the form

$$
\mathbb{P}(E_2) \le C_9 \operatorname{Vol}(A) d(A,B)^{-\beta}, \quad \mathbb{P}(E_3) \le C_9 \operatorname{Vol}(B) d(A,B)^{-\beta} \tag{2.31}
$$

for some constant $C_9 \in (0,\infty)$. Combining (2.24) and (2.31), and choosing a suitable constant $\gamma_3$ for (1.6), we obtain

$$
\mathbb{P}(E^c) \le C_{10}(\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1) d(A,B)^{-\beta}
$$

for some constant $C_{10} \in (0,\infty)$. Recall that

$$
\underline{p}_1 = 1 - \frac{2}{\underline{p}_0} = 1 - \frac{2(p_1 + p_2 - 1)}{p_1 p_2}.
$$

Then, applying (2.29), Hölder's inequality, and Lemma 2.2, we obtain the following bounds.

When $j = 1$ and $g \in \mathcal{F}_1$, i.e., $g(x) = x$ or $g \in \mathcal{F}_b$, we have

$$
\begin{aligned}
&|\mathbb{E}\left(S_A g(S_B)\right) - \mathbb{E}\left(S_A\right)\mathbb{E}\left(g(S_B)\right)| \\
&\le |\mathbb{E}\left(S_A g(S_B)\mathbf{1}_{E^c}\right)| + |\mathbb{E}\left(S'_A g(S'_B)\mathbf{1}_{E^c}\right)| \\
&\le 2\|S_A\|_{p_1p_2/(p_1+p_2-1)} \|g(S_B)\|_{p_1p_2/(p_1+p_2-1)} \mathbb{P}(E^c)^{\underline{p}_1} \\
&\le C_{11}(\operatorname{diam}(A)^{d(1+\underline{p}_1)} \vee 1)(\operatorname{diam}(B)^{d(1+\underline{p}_1)} \vee 1) d(A,B)^{-\beta \underline{p}_1}
\end{aligned}
$$

for some constant $C_{11} \in (0,\infty)$.

When $j = 2$ and $g \in \mathcal{F}_b$, we similarly obtain

$$
\begin{aligned}
&\left|\mathbb{E}\left(S_A^2 g(S_B)\right) - \mathbb{E}\left(S_A^2\right)\mathbb{E}\left(g(S_B)\right)\right| \\
&\le \left|\mathbb{E}\left(S_A^2 \mathbf{1}_{E^c}\right)\right| + \left|\mathbb{E}\left((S'_A)^2 \mathbf{1}_{E^c}\right)\right| \\
&\le 2\|S_A\|^2_{p_1p_2/(p_1+p_2-1)} \mathbb{P}(E^c)^{\underline{p}_1} \\
&\le C_{12}(\operatorname{diam}(A)^{d(2+\underline{p}_1)} \vee 1)(\operatorname{diam}(B)^{d(2+\underline{p}_1)} \vee 1) d(A,B)^{-\beta \underline{p}_1}
\end{aligned}
$$

for some constant $C_{12} \in (0,\infty)$. Thus, polynomial mixing holds with parameters $\gamma_0 = d(2 + \underline{p}_1)$, $\gamma_2 = \beta \underline{p}_1$, and the proof is complete by applying Theorem 1.6 (b). □

Before going into the proof of Corollary 1.14, we state a simple lemma.

**Lemma 2.6.** *If $X$ is a random variable such that $\|X\|_{m_2}$ exists for some positive constant $m_2$, then for all $m_1 < m_3 < m_2$,*

$$\|X\|_{m_3} \leq \|X\|_{m_1}^{m_1(m_2-m_3)/(m_3(m_2-m_1))} \|X\|_{m_2}^{m_2(m_3-m_1)/(m_3(m_2-m_1))}.$$

*Proof.* For any $m_3 \in (m_1, m_2)$, Hölder's inequality implies

$$\begin{aligned}\mathbb{E}\left(|X|^{m_3}\right) =& \mathbb{E}\left(|X|^{m_2(m_3-m_1)/(m_2-m_1)} |X|^{m_1(m_2-m_3)/(m_2-m_1)}\right) \\ \leq& \left\| |X|^{m_2(m_3-m_1)/(m_2-m_1)} \right\|_{(m_2-m_1)/(m_3-m_1)} \left\| |X|^{m_1(m_2-m_3)/(m_2-m_1)} \right\|_{(m_2-m_1)/(m_2-m_3)} \\ =& \mathbb{E}\left(|X|^{m_1}\right)^{(m_2-m_3)/(m_2-m_1)} \mathbb{E}\left(|X|^{m_2}\right)^{(m_3-m_1)/(m_2-m_1)}.\end{aligned}$$

The statement holds by taking $m_3$-th roots on both sides. □

**Corollary 2.7.** *Assume that $\xi$ satisfies the $p_2$-moment condition* (1.2) *and $\hat{\xi}^{[r]}$ satisfies the $p_2'$-moment condition* (1.17)*.*

*(a) If exponential $L^p$ stabilization holds for $p = p_3$, then it also holds for all $p \leq p_3 \vee (p_2 \wedge p_2')$.*

*(b) If polynomial $L^p$ stabilization holds for $p = p_3$ with $\beta = C_2$ for some positive constant $C_2$, then it also holds for any $p \in (p_3, p_2 \wedge p_2')$ with $\beta = C_2 p_3[(p_2 \wedge p_2') - p]/\{p[(p_2 \wedge p_2') - p_3]\}$.*

*Proof.* (a) The statement holds trivially for all $p < p_3$ by Lyapunov's inequality. For any $p \in (p_3, p_2 \wedge p_2')$, the statement holds by applying Lemma 2.6 with $X \sim \mathcal{L}_x\left(\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) - \hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)\right)$, $m_1 = p_3$ and $m_2 = p_2 \wedge p_2'$ once we note that $\|X\|_{m_1}$ is uniformly bounded by Minkowski's inequality.
(b) Statement (b) holds by taking $X \sim \mathcal{L}_x\left(\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) - \hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)\right)$, $m_1 = p_3$ and $m_2 = p_2 \wedge p_2'$ in Lemma 2.6. □

*Proof of Corollary 1.14.* The moment conditions on the ground point process and on the score are already given; it remains to verify exponential mixing (1.5) for part (a) and polynomial mixing (1.6) for part (b) and to apply Theorem 1.6. Recall the definitions of the sums of short-range scores in (2.25) and (2.26) where we have taken $r = d(A, B)/3$. The independence condition $\bar{S}_A^{[r]} \perp\!\!\!\perp \bar{S}_B^{[r]}$ implies for any $j \in \{1, 2\}$ and $g \in \mathcal{F}_j$, $A \in \mathcal{H}_j$,

$$\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j g\left(\bar{S}_B^{[r]}\right)\right) - \mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j\right) \mathbb{E}\left(g\left(\bar{S}_B^{[r]}\right)\right) = 0.$$

This yields the following bound for any $j \in \{1, 2\}$ and $g \in \mathcal{F}_j$, $A \in \mathcal{H}_j$,

$$\begin{aligned}\epsilon(r) :=& \left|\mathbb{E}\left(S_A^j g(S_B)\right) - \mathbb{E}\left(S_A^j\right) \mathbb{E}\left(g(S_B)\right)\right| \\ =& \left|\left[\mathbb{E}\left(S_A^j g(S_B)\right) - \mathbb{E}\left(S_A^j\right) \mathbb{E}\left(g(S_B)\right)\right] - \left[\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j g\left(\bar{S}_B^{[r]}\right)\right) - \mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j\right) \mathbb{E}\left(g\left(\bar{S}_B^{[r]}\right)\right)\right]\right|\end{aligned}$$

$$
\begin{aligned}
&\leq \left|\mathbb{E}\left(\left[S_A^j - \left(\bar{S}_A^{[r]}\right)^j\right] g(S_B)\right)\right| + \left|\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j \left[g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right)\right]\right| \\
&\quad + \left|\mathbb{E}\left(S_A^j - \left(\bar{S}_A^{[r]}\right)^j\right) \mathbb{E}(g(S_B))\right| + \left|\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^j\right) \mathbb{E}\left(g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right)\right| \\
&=: \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4. \qquad (2.32)
\end{aligned}
$$

By Lemma 2.2, with $\xi$, $p_2$, and $b_2$ replaced by $\hat{\xi}^{[r]}$, $p_2'$, and $b_2'$, respectively, we obtain

$$
\|\bar{S}_B^{[r]}\|_3 \leq \|\bar{S}_B^{[r]}\|_{p_1 p_2'/(p_1+p_2'-1)} \leq C_1 \operatorname{Vol}(B) \qquad (2.33)
$$

for some constant $C_1 \in (0,\infty)$ and all $r \in (0,\infty)$. To guarantee that both $p_1p_2'/(p_1+p_2'-1) \geq 3$ and $p_1p_2/(p_1+p_2-1) \geq 3$, it suffices to require $\underline{p} = p_1(p_2 \wedge p_2')/(p_1+p_2\wedge p_2'-1) \geq 3$.

Moreover, for any random variable $X$, any conjugate pair $p, q \geq 1$ satisfying $1/p + 1/q = 1$, and for all functions $g \in \mathcal{F}_1$, namely for all $g(x) = x$ or $g(x) \in \mathcal{F}_b$, Hölder's inequality yields:

$$
\left|\mathbb{E}\left\{X\left[g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right]\right\}\right| \leq \|X\|_p \left\|g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right\|_q \leq \|X\|_p \|g'\|_\infty \|S_B - \bar{S}_B^{[r]}\|_q.
$$

In particular, we obtain the following useful inequalities:

$$
\begin{aligned}
&\left|\mathbb{E}\left\{X\left[g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right]\right\}\right| \leq \|X\|_p \|H_{\lambda,B}^{[r]}\|_q, \\
&\|g(X)\|_p \leq \|X\|_p \vee 1.
\end{aligned}
$$

These inequalities will be used repeatedly in the proofs.

(a) The assumption $\underline{p} \geq 3$ implies $p_2 \wedge p_2' > 3 \geq \underline{p}/(\underline{p}-2) > \underline{p}/(\underline{p}-1)$. Then from Corollary 2.7 (a), the exponential $L^p$ stabilization also holds for $p$ taking $\underline{p}/(\underline{p}-1)$ and $\underline{p}/(\underline{p}-2)$. When $j=1$ and either $g(x) = x$ or $g(x) \in \mathcal{F}_b$, then by (2.33), Lemma 2.2, Lemma 2.3, and Hölder's inequality, the four terms on the right-hand side of (2.32) satisfy the bounds

$$
\begin{aligned}
\epsilon_1 &= \left|\mathbb{E}\left(H_{\lambda,A}^{[r]} g(S_B)\right)\right| \leq \|H_{\lambda,A}^{[r]}\|_2 \|g(S_B)\|_2 = O\left((\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1) \mathrm{e}^{-C_2' r}\right), \\
\epsilon_2 &\leq \|\bar{S}_A^{[r]}\|_2 \|H_{\lambda,B}^{[r]}\|_2 = O\left((\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1) \mathrm{e}^{-C_2' r}\right), \\
\epsilon_3 &\leq \|H_{\lambda,A}^{[r]}\|_1 \|g(S_B)\|_1 = O\left((\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1) \mathrm{e}^{-C_2' r}\right), \\
\epsilon_4 &\leq \|\bar{S}_A^{[r]}\|_1 \|H_{\lambda,B}^{[r]}\|_1 = O\left((\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1) \mathrm{e}^{-C_2' r}\right).
\end{aligned}
\qquad (2.34)
$$

Therefore, each $\epsilon_i, 1 \leq i \leq 4$, is bounded above by $O\left((\operatorname{Vol}(A) \vee 1)(\operatorname{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right)$ for some constant $C_2' \in (0,\infty)$.

Similarly, when $j = 2$ and $g(x) \in \mathcal{F}_b$, the four terms on the right-hand side of (2.32) satisfy the following bounds:

$$
\epsilon_1 = \left|\mathbb{E}\left(\left(S_A - \bar{S}_A^{[r]}\right)\left(S_A + \bar{S}_A^{[r]}\right) g(S_B)\right)\right| \leq \|H_{\lambda,A}^{[r]}\|_{\underline{p}/(\underline{p}-1)} (\|S_A\|_{\underline{p}} + \|\bar{S}_A^{[r]}\|_{\underline{p}})
$$

$$
\begin{aligned}
&= O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right), \\
\epsilon_2 &\leq \left|\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^2 \left|g(S_B) - g\left(\bar{S}_B^{[r]}\right)\right|\right)\right| \leq \left\|\left(\bar{S}_A^{[r]}\right)^2\right\|_{\underline{p}/2} \|H_{\lambda,B}^{[r]}\|_{\underline{p}/(\underline{p}-2)} = \|\bar{S}_A^{[r]}\|_{\underline{p}}^2 \|H_{\lambda,B}^{[r]}\|_{\underline{p}/(\underline{p}-2)} \\
&= O\left(\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right), \\
\epsilon_3 &\leq \left|\mathbb{E}\left(H_{\lambda,A}^{[r]}\left(S_A + \bar{S}_A^{[r]}\right)\right)\right| \leq \|H_{\lambda,A}^{[r]}\|_2 \left(\|S_A\|_2 + \|\bar{S}_A^{[r]}\|_2\right) \\
&= O\left(\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right), \\
\epsilon_4 &\leq \left|\mathbb{E}\left(\left(\bar{S}_A^{[r]}\right)^2\right)\right| \|H_{\lambda,B}^{[r]}\|_1 \leq \|\bar{S}_A^{[r]}\|_2^2 \|H_{\lambda,B}^{[r]}\|_1 \\
&= O\left(\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1))\mathrm{e}^{-C_2' r}\right).
\end{aligned}
$$

Hence,

$$\max_{i \leq 4} \epsilon_i = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right) \tag{2.35}$$

for some constant $C_2' \in (0, \infty)$. Combining the bounds in (2.34) and (2.35) with (2.32), we conclude that

$$\epsilon(r) = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)\mathrm{e}^{-C_2' r}\right).$$

This verifies the exponential mixing condition (1.5) with parameters $\gamma_0 = 2d$, $\gamma_2 = C_2'/3$ and any $\gamma_3 > 0$. The conclusion follows by Theorem 1.6 (a).

(b) As in the proof of part (a), we obtain the following similar bounds. For $j = 1$ and either $g(x) = x$ or $g(x) \in \mathcal{F}_b$, the four terms on the right-hand side of (2.32) satisfy

$$
\begin{aligned}
\epsilon_1 &\leq \|H_{\lambda,A}^{[r]}\|_2 \|g(S_B)\|_2 = O\left((\mathrm{Vol}(A) \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_2 &\leq \|\bar{S}_A^{[r]}\|_2 \|H_{\lambda,B}^{[r]}\|_2 = O\left((\mathrm{Vol}(A) \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_3 &\leq \|H_{\lambda,A}^{[r]}\|_1 \|g(S_B)\|_1 = O\left((\mathrm{Vol}(A) \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_4 &\leq \|\bar{S}_A^{[r]}\|_1 \|H_{\lambda,B}^{[r]}\|_1 = O\left((\mathrm{Vol}(A) \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right).
\end{aligned}
\tag{2.36}
$$

For $j = 2$ and $g(x) \in \mathcal{F}_b$, we have

$$
\begin{aligned}
\epsilon_1 &\leq \|H_{\lambda,A}^{[r]}\|_{\underline{p}/(\underline{p}-1)} (\|S_A\|_{\underline{p}} + \|\bar{S}_A^{[r]}\|_{\underline{p}}) = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_2 &\leq \left\|\bar{S}_A^{[r]}\right\|_{\underline{p}}^2 \|H_{\lambda,B}^{[r]}\|_{\underline{p}/(\underline{p}-2)} = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_3 &\leq \|H_{\lambda,A}^{[r]}\|_2 \left(\|S_A\|_2 + \|\bar{S}_A^{[r]}\|_2\right) = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right), \\
\epsilon_4 &\leq \|\bar{S}_A^{[r]}\|_2^2 \|H_{\lambda,B}^{[r]}\|_1 = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right).
\end{aligned}
\tag{2.37}
$$

Combining the bounds in (2.36) and (2.37) with (2.32), we obtain

$$\epsilon(r) = O\left((\mathrm{Vol}(A)^2 \vee 1)(\mathrm{Vol}(B) \vee 1)r^{-\beta}\right).$$

This implies polynomial mixing of scores with parameters $\gamma_0 = 2d$, $\gamma_2 = \beta$ and any $\gamma_3 > 0$, and the conclusion follows by Theorem 1.6 (b). $\square$

# 3 Applications

This section provides examples illustrating the applicability of Theorem 1.6, Corollary 1.11 and Corollary 1.14. The order of presentation parallels that of the main results in Section 1.

## 3.1 Marked determinantal point processes

We show that determinantal point processes having an exponentially decaying kernel and equipped with i.i.d. marks $\{M_x\}_{x\in\mathcal{P}}$ give rise to statistics $H_\lambda := \sum_{x\in\mathcal{P}\cap\Gamma_\lambda} M_x$ which satisfy the conditions of Theorem 1.6. In the degenerate case $M_x \equiv 1$, the statistic $H_\lambda$ reduces to the counting statistic. We deduce that these counting statistics satisfy a quantitative CLT with a nearly optimal rate. Local (and weaker) qualitative central limit theorems for the point counts of unmarked determinantal point processes were established in Soshnikov [40] and Forrester and Lebowitz [16]. However, for the counting statistics and localized marked sums considered here under merely exponential decay of the kernel, comparable nearly optimal Wasserstein bounds do not appear to be available. Closely related results include the Wasserstein bound of Wiroonsri [44], which has a rate slower than $\lambda^{-1/4}$; the quantitative CLT of [10], which assumes the *super-exponential* decay condition (1.12) and thereby excludes commonly used determinantal point processes, including the Ginibre point process; and the Kolmogorov distance bound of Dinh et al. [14] for linear statistics of unmarked determinantal point processes.

The key idea is that for these statistics, the mixing condition (1.5) can be upper bounded by sums of expectations expressed with respect to first- and second-order Palm distributions of $\mathcal{P}$. By employing the couplings in Møller and O'Reilly [30], we find that the discrepancies between the distribution of $\mathcal{P}$ and its first- and second-order Palm distributions are localized near the reference points of the Palm distributions. More precisely, under the conditioning that one or two reference points lie in $A$, the configuration in $B$ changes only slightly, in the sense that under the relevant coupling, the probability that the original and Palm distribution configurations differ by at least one point decays exponentially fast as $d(A,B)\to\infty$, and the total discrepancy between these configurations in the set $B$ is bounded by two points. This, together with some additional estimates, yields mixing of score sums.

We focus on a stationary determinantal point process $\mathcal{P}$ on $\mathbb{R}^d$ with Lebesgue measure as the reference measure and a Hermitian kernel $K:\mathbb{R}^d\times\mathbb{R}^d\to\mathbf{C}$, namely $K(x,y) := K_0(x-y) = \overline{K_0(y-x)}$. Let $\hat{K}_0(x) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i x\cdot t} K_0(t)\mathrm{d}t$ be the Fourier transform of $K_0$, where $x\cdot t$ is the dot product. We assume:

**Condition 1.** The function $K_0$ is continuous and complex-valued on $\mathbb{R}^d$, $K_0 \not\equiv 0$, and $|K_0(x)| = O(\mathrm{e}^{-|x|})$ as $|x|\to\infty$.

**Condition 2.** $0 \le \hat{K}_0 \le 1$.

The determinantal point process $\mathcal{P}$ is simple with its distribution determined by

$$\mathbb{E}\left[\prod_{i=1}^{m} \mathcal{P}(A_i)\right] = \int_{A_1\times A_2\times\cdots\times A_m} \det\left[K_0(u_i - u_j)\right]_{i,j=1}^{m} du_1 \ldots \mathrm{d}u_m,$$

for all mutually disjoint bounded sets $A_1, \ldots, A_m \in \mathcal{B}(\mathbb{R}^d)$, $m \in \mathbb{N}$. The function $\rho_m : (\mathbb{R}^d)^m \to [0,\infty)$ given by $\rho_m(x_1,\ldots,x_m) := \det\left[K_0(x_i - x_j)\right]_{1\le i,j\le m}$ is the $m$-th order correlation function of $\mathcal{P}$.

**Theorem 3.1.** *(normal approximation for sum of marks on a determinantal point process) For the determinantal point process $\mathcal{P}$ with kernel satisfying Conditions 1 and 2, let $\hat{\mathcal{P}}$ be a marked point process equipped with i.i.d. marks $\{M_x\}_{x\in\mathcal{P}}$. Assume that $0 < \|M_{\mathbf{0}}\|_{p_0} < \infty$ for some $p_0 > 3$. Then with*

$$H_\lambda = \sum_{x\in\mathcal{P}\cap\Gamma_\lambda} M_x, \quad Z_\lambda := \frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\mathrm{Var}H_\lambda}},$$

*we have*

$$d_W(Z_\lambda, Z) = O\left((\log\lambda)^{3d}\lambda^{-0.5}\right).$$

**Remark 3.2.** *(comparison with related literature) The assumptions imply that $\mathcal{P}$ has exponentially decaying correlations and thus by Theorem 5.2 in [5], $(Z_\lambda)_{\lambda\ge 1}$ converges to the standard normal in distribution. Theorem 3.1 establishes a rate of convergence for the case of i.i.d. marks, which also covers the unmarked case. This result also extends the quantitative CLT in [10], since the determinantal point processes satisfying Conditions 1 and 2 are not currently known to be EDD. Furthermore, following the same idea as in the proof, we can show that the statement still holds if the marks $M_x$ depend on the location $x$, while remaining independent of one another and of all other points.*

It is not known whether $\mathcal{P}$ is EDD. To prove Theorem 3.1, we verify directly that the conditions of Theorem 1.6 are fulfilled.

**Lemma 3.3.** *If Conditions 1 and 2 hold, then $K_0(\mathbf{0}) > \int_{\mathbb{R}^d} |K_0(x)|^2 \mathrm{d}x := \||K_0\||_2^2$.*

*Proof.* We prove the statement by contradiction. Assume that $K_0(\mathbf{0}) = \int_{\mathbb{R}^d} |K_0(x)|^2 \mathrm{d}x$. The Fourier inversion theorem states that

$$K_0(x) = \int_{\mathbb{R}^d} \mathrm{e}^{2\pi i x\cdot t}\hat{K}_0(t)\mathrm{d}t. \tag{3.1}$$

Evaluating at $x = \mathbf{0}$, we get

$$K_0(\mathbf{0}) = \int_{\mathbb{R}^d} \hat{K}_0(t)dt \ge \int_{\mathbb{R}^d} (\hat{K}_0(t))^2 \mathrm{d}t = \int_{\mathbb{R}^d} |K_0(x)|^2 \mathrm{d}x. \tag{3.2}$$

The inequality in (3.2) follows from Condition 2 whereas the last equality follows from the Plancherel theorem. Since $K_0(\mathbf{0}) = \int_{\mathbb{R}^d} |K_0(x)|^2 \mathrm{d}x$, the inequality (3.2) is an equality and Condition 2 ensures that $\{t \in \mathbb{R}^d : \hat{K}_0(t) \notin \{0,1\}\}$ has Lebesgue measure 0.

On the other hand, Condition 1 ensures that $K_0$ is both integrable and square integrable with respect to the Lebesgue measure, which implies that $\hat{K}_0$ must be continuous. This leads to the conclusion that $\hat{K}_0$ must be constant, either $\hat{K}_0 \equiv 0$ or $\hat{K}_0 \equiv 1$. If $\hat{K}_0 \equiv 0$, then $K_0 \equiv 0$, which contradicts $K_0 \not\equiv 0$. If $\hat{K}_0 \equiv 1$, then $K_0$ is not square integrable with respect to the Lebesgue measure, which contradicts Condition 1. Thus, we have reached a contradiction, and the original assumption must be false. □

*Proof of Theorem 3.1.* For all $x \in \mathcal{P}, \lambda \geq 1$, put $\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) = M_x$. To apply part (a) of Theorem 1.6, it is enough to check conditions (1.1), (1.2), exponential mixing of score sums as in Definition 1.4, and that $\sigma_\lambda^2 = \mathrm{Var} H_\lambda = \Omega(\lambda)$. From Section 2.2.2 and Section 2.1, Remark (i) of [4], $\mathbb{E}\left(\mathcal{P}(B)^k\right)$ is finite for all bounded $B \in \mathcal{B}(\mathbb{R}^d)$ and all $k \in \mathbb{N}$, which ensures (1.1). This, together with the independence of marks, $\|M_{\mathbf{0}}\|_{p_0} < \infty$ for some $p_0 > 3$ and Remark 1.7 (iii), also ensures condition (1.2) in Theorem 1.6. By combining Theorem 1.1.2 (ii) and Remark (iii) related to Theorem 1.1.3 of [4] and Lemma 3.3, we conclude that $\sigma_\lambda^2 = \Omega(\lambda)$. It remains to show exponential mixing of score sums (1.4). This will necessitate coupling arguments and some additional background material.

Let $\mathcal{P}_x^!$ be the reduced Palm process of $\mathcal{P}$ at point $x$, that is, $\mathcal{P}_x^! + \delta_x \sim \mathcal{L}_x(\mathcal{P})$. Recall the assumptions required for the existence of a coupling of the determinantal point process with kernel function $K(u,v)$ and its reduced Palm distributions in [30]:

**Assumption 1.** The kernel $K$ is Hermitian, that is, $K(u,v) = \overline{K(v,u)}$ for all $u, v \in \mathbb{R}^d$.

**Assumption 2.** For any compact set $S \subset \mathbb{R}^d$, $\iint_{S^2} |K(u,v)| dudv < \infty$, which means that $K$ is locally square integrable.

**Assumption 3.** For any compact set $S \subset \mathbb{R}^d$, $\int_S |K(u,u)| \mathrm{d}u < \infty$, which means that $K$ is of local trace class.

**Assumption 4.** For any compact set $S \subset \mathbb{R}^d$, all eigenvalues of $K$ on $L^2(S)$, the space of square-integrable complex functions with respect to the Lebesgue measure restricted to $S$, lie within the interval $[0,1]$.

It is stated in [30] that Assumption 4 is equivalent to the existence of the determinantal point process given that Assumptions $1-3$ are satisfied, and the determinantal point process is unique, see, e.g., Hough et al. [22, Theorem 4.5.5 and Lemma 4.2.6].

For the determinantal point process $\mathcal{P}$, $K(u,v) = K_0(u-v)$. We claim that Conditions 1 and 2 imply these four properties (i) $K_0$ is Hermitian, (ii) $K_0(\mathbf{0}) \geq |||K_0|||_2^2$, (iii) $K_0(\mathbf{0}) > |K_0(x)|$, for all $x \neq \mathbf{0}$, and (iv) all eigenvalues lie within $[0,1]$. Property (i) can be directly verified from (3.1), and Property (ii) follows from (3.2). Property (iv) is a direct consequence of Condition 2 and Lavancier, Møller and Rubak [26, Proposition 3.1]. Therefore, it remains to prove Property (iii). To this end, write $b_1 = \int_{\mathbb{R}^d} \cos^2(2\pi x{\cdot}t)\hat{K}_0(t)\mathrm{d}t$ and $b_2 = \int_{\mathbb{R}^d} \sin^2(2\pi x{\cdot}t)\hat{K}_0(t)\mathrm{d}t$. Since $b_1 + b_2 = K_0(\mathbf{0}) \in (0,\infty)$, at least one of $b_1$ and $b_2$ is positive, so without loss, we assume that $b_1 \in (0,\infty)$. Then

$$\frac{|\cos(2\pi x{\cdot}t)|}{\sqrt{b_1}} \cdot \frac{1}{\sqrt{K_0(\mathbf{0})}} \leq \frac{1}{2}\left(\frac{\cos^2(2\pi x{\cdot}t)}{b_1} + \frac{1}{K_0(\mathbf{0})}\right), \tag{3.3}$$

and

$$\begin{aligned} S_x &:= \left\{t:\ \frac{|\cos(2\pi x{\cdot}t)|}{\sqrt{b_1}} \cdot \frac{1}{\sqrt{K_0(\mathbf{0})}} = \frac{1}{2}\left(\frac{\cos^2(2\pi x{\cdot}t)}{b_1} + \frac{1}{K_0(\mathbf{0})}\right)\right\} \\ &= \left\{t:\ |\cos(2\pi x{\cdot}t)| = \sqrt{\frac{b_1}{K_0(\mathbf{0})}}\right\} \end{aligned}$$

has Lebesgue measure 0. Multiplying both sides of (3.3) by $\hat{K}_0(t)$ and integrating in terms of $t$, we obtain

$$\int_{\mathbb{R}^d} \frac{|\cos(2\pi x{\cdot}t)|}{\sqrt{b_1}} \cdot \frac{1}{\sqrt{K_0(\mathbf{0})}}\hat{K}_0(t)\mathrm{d}t < 1,$$

which is equivalent to

$$\int_{\mathbb{R}^d} |\cos(2\pi x{\cdot}t)|\hat{K}_0(t)\mathrm{d}t < \sqrt{b_1 K_0(\mathbf{0})}. \tag{3.4}$$

Now, applying (3.4) for the term with cos and the Cauchy-Schwarz inequality for the term with sin below, we get

$$\begin{aligned} |K_0(x)|^2 &= \left(\int_{\mathbb{R}^d} \cos(2\pi x{\cdot}t)\hat{K}_0(t)\mathrm{d}t\right)^2 + \left(\int_{\mathbb{R}^d} \sin(2\pi x{\cdot}t)\hat{K}_0(t)\mathrm{d}t\right)^2 \\ &< b_1 K_0(\mathbf{0}) + b_2 K_0(\mathbf{0}) = K_0(\mathbf{0})^2, \end{aligned} \tag{3.5}$$

hence Property (iii) follows.

Now, Properties (i)–(iv) show that $K(u,v)$ satisfies Assumptions 1-4. According to [30, Theorem 1], $\mathcal{P}_x^!$ is also a determinantal point process with kernel

$$K^x(u,v) = K_0(u-v) - \frac{K_0(u-x)K_0(x-v)}{K_0(\mathbf{0})},$$

and there exists a coupling of $\mathcal{P}$ and $\mathcal{P}_x^!$ such that, almost surely, $\mathcal{P}_x^! \subseteq \mathcal{P}$ and $\mathcal{P}^{\Delta,x} = \mathcal{P}\backslash\mathcal{P}_x^!$ consists of at most one point, with

$$p_x := \mathbb{P}(\mathcal{P}^{\Delta,x} \neq \emptyset) = \frac{\|\!|K_0\|\!|_2^2}{K_0(\mathbf{0})},$$

and conditional on $\mathcal{P}^{\Delta,x} \neq \emptyset$, the point in $\mathcal{P}^{\Delta,x}$ has intensity decaying exponentially fast with the distance to $x$:

$$f_x(v) = \frac{|K_0(v-x)|^2}{|||K_0|||_2^2}.$$

Furthermore, one can verify that $K^x(u,v)$ also satisfies Assumptions 1-3. For distinct $x_1, x_2 \in \mathbb{R}^d$, the second order intensity function at $x_1$ and $x_2$ is

$$\rho(x_1,x_2) = K_0(\mathbf{0})^2 - |K_0(x_2-x_1)|^2 \in \left(0, K_0(\mathbf{0})^2\right].$$

Since the reduced Palm process $\mathcal{P}^!_{x_1}$ is also a determinantal point process, we can apply the result from [30] to analyze the properties of the second-order reduced Palm process $\mathcal{P}^!_{x_1,x_2}$ of $\mathcal{P}$ at the distinct points $x_1$ and $x_2$. This process is the same as the reduced Palm distribution of $\mathcal{P}^!_{x_1}$ at point $x_2$, provided that $\mathcal{P}^!_{x_1}$ satisfies Assumptions 1-4. We thank J. Møller for pointing this out in a personal communication.

The Hermitian assumption can be directly verified using the Hermitian property of $K_0$. Assumptions 2 and 3 follow from the compactness of the set $S$ and the boundedness of the kernel as shown in (3.5). Assumption 4 follows from the existence of the Palm process.

Thus we have shown that Conditions 1 and 2 imply Assumptions 1-4. Now we can finally establish the pertinent coupling. Applying [30, Theorem 1] again, there exists a coupling of $\mathcal{P}^!_{x_1,x_2}$ and $\mathcal{P}^!_{x_1}$ such that, almost surely, $\mathcal{P}^!_{x_1,x_2} \subseteq \mathcal{P}^!_{x_1}$, and $\mathcal{P}^{\Delta,x_1,x_2} := \mathcal{P}^!_{x_1} \backslash \mathcal{P}^!_{x_1,x_2}$ consists of at most one point. The probability of this event is given by

$$p_{x_1,x_2} := \mathbb{P}(\mathcal{P}^{\Delta,x_1,x_2} \neq \emptyset) = \frac{|||K^{x_1}(x_2,\cdot)|||_2^2}{K^{x_1}(x_2,x_2)},$$

and conditional on $\mathcal{P}^{\Delta,x_1,x_2} \neq \emptyset$, the intensity of the point in $\mathcal{P}^{\Delta,x_1,x_2}$ is

$$f_{x_1,x_2}(v) = |K^{x_1}(x_2,v)|^2/|||K^{x_1}(x_2,\cdot)|||_2^2.$$

In summary, we can find a coupling of $\mathcal{P}^!_{x_1,x_2}$, $\mathcal{P}^!_{x_1}$ and $\mathcal{P}$ such that $\mathcal{P}^!_{x_1,x_2} \subseteq \mathcal{P}^!_{x_1} \subseteq \mathcal{P}$, and $\mathcal{P}^{\Delta,x_1,x_2} = \mathcal{P}^!_{x_1} \backslash \mathcal{P}^!_{x_1,x_2}$ and $\mathcal{P}^{\Delta,x_1} = \mathcal{P} \backslash \mathcal{P}^!_{x_1}$ both contain at most one point and the intensity $f_{x_1,x_2}$ decays exponentially fast with respect to $|x_1 - v|$ and $|x_2 - v|$. Since the marks are independent, we automatically obtain a coupling of the corresponding marked point processes.

From Lemma 2.5, to show the exponential mixing (1.5), it is equivalent to show the corresponding property for the sums of scores without centering. We start with $j = 2$. Since $H^2_{\lambda,A} = \iint_{A^2} M_{x_1} M_{x_2} \mathcal{P}(dx_1)\mathcal{P}(dx_2)$, using the second order Palm distribution and the independence of the marks conditional on the ground point process, we have for any disjoint $A$ and $B$ in $\mathcal{B}(\mathbb{R}^d)$

$$\left|\mathbb{E}\left(H^2_{\lambda,A} g(H_{\lambda,B})\right) - \mathbb{E}\left(H^2_{\lambda,A}\right)\mathbb{E}\left(g(H_{\lambda,B})\right)\right|$$

$$
\begin{aligned}
&= \left|\mathbb{E}\left(\iint_{A^2} M_{x_1} M_{x_2}\left[g\left(\int_B M_y \mathcal{P}(\mathrm{d}y)\right) - g\left(\int_B \tilde{M}_y \mathcal{P}'(\mathrm{d}y)\right)\right] \mathcal{P}(dx_1)\mathcal{P}(dx_2)\right)\right| \\
&\leq \left|\iint_{A^{2,\neq}} \rho(x_1,x_2)\mathbb{E}\left(M_{\mathbf{0}}\right)^2 \mathbb{E}\left[g\left(\int_B M_y \mathcal{P}^!_{x_1,x_2}(\mathrm{d}y)\right) - g\left(\int_B \tilde{M}_y \mathcal{P}'(\mathrm{d}y)\right)\right] dx_1 dx_2\right| \\
&\quad + \left|\int_A K_0(\mathbf{0})\mathbb{E}\left(M_{\mathbf{0}}^2\right)\mathbb{E}\left(\left[g\left(\int_B M_y \mathcal{P}^!_x(\mathrm{d}y)\right) - g\left(\int_B \tilde{M}_y \mathcal{P}'(\mathrm{d}y)\right)\right]\right)\mathrm{d}x\right|, \qquad (3.6)
\end{aligned}
$$

where here and in the following, $\mathcal{P}'$, $\tilde{M}_x$ denote independent copies of $\mathcal{P}$ and $M_x$, respectively. The assumption of independence between the marks and the ground process enables us to couple $\hat{\mathcal{P}}$ and its corresponding Palm process via $\mathcal{P}$ and its Palm process, resulting in the clean explicit form in (3.6).

By condition 1, there are positive constants $C_1$ and $C_2$ such that $|K_0(x)| \leq C_1 \mathrm{e}^{-C_2|x|}$ for all $x \in \mathbb{R}^d$. For any $x \in A$ we bound $K_0(x-v)$ by $C_1\mathrm{e}^{-C_2 d(A,B)}$. We find, for $x_1 \neq x_2$,

$$
\begin{aligned}
&\mathbb{P}\left(\left(\mathcal{P}\backslash\mathcal{P}^!_{x_1,x_2}\right)(B) \neq 0\right) = \mathbb{P}\left(\left(\mathcal{P}^{\Delta,x_1,x_2} \cup \mathcal{P}^{\Delta,x_1}\right)(B) \neq 0\right) \\
&\leq \int_B (p_{x_1} f_{x_1}(v) + p_{x_1,x_2} f_{x_1,x_2}(v))\mathrm{d}v \\
&= \int_B \left(\frac{|K_0(v-x_1)|^2}{K_0(\mathbf{0})} + \frac{|K^{x_1}(x_2,v)|^2}{K^{x_1}(x_2,x_2)}\right)\mathrm{d}v \\
&= \int_B \frac{1}{K_0(\mathbf{0})}\left(|K_0(v-x_1)|^2 + \frac{|K_0(x_2-v)K_0(\mathbf{0}) - K_0(x_2-x_1)K_0(x_1-v)|^2}{K_0(\mathbf{0})^2 - |K_0(x_2-x_1)|^2}\right)\mathrm{d}v \\
&\leq \frac{1}{K_0(\mathbf{0})}\mathrm{Vol}(B)\left[C_1^2\mathrm{e}^{-2C_2 d(A,B)} + \frac{2(C_1^4 \vee 1)\mathrm{e}^{-2C_2 d(A,B)}}{K_0(\mathbf{0})^2 - |K_0(x_2-x_1)|^2}\right] \\
&= \frac{1}{K_0(\mathbf{0})}\mathrm{Vol}(B)\left[C_1^2\mathrm{e}^{-2C_2 d(A,B)} + \frac{2(C_1^4 \vee 1)\mathrm{e}^{-2C_2 d(A,B)}}{\rho(x_1,x_2)}\right].
\end{aligned}
$$

By independence and the bound $\|M_{\mathbf{0}}\|_1 \leq \|M_{\mathbf{0}}\|_{p_0} < \infty$, the first term on the right-hand side of (3.6) is bounded above by

$$
\begin{aligned}
&2K_0(\mathbf{0})^2\mathrm{Vol}(A)^2\mathrm{Vol}(B)\mathbb{E}(M_{\mathbf{0}})^2\frac{1}{K_0(\mathbf{0})}\left[C_1^4\mathrm{e}^{-2C_2 d(A,B)} + (C_1^4 \vee 1)\mathrm{e}^{-2C_2 d(A,B)}\right] \\
&=: \mathrm{Vol}(A)^2\mathrm{Vol}(B)C_3\mathrm{e}^{-C_4 d(A,B)}. \qquad (3.7)
\end{aligned}
$$

Similarly, for the second term on the right-hand side of (3.6), we have

$$
\mathbb{P}\left(\left(\mathcal{P}\backslash\mathcal{P}^!_x\right)(B) \neq 0\right) = \int_B \frac{|K_0(v-x_1)|^2}{K_0(\mathbf{0})}\mathrm{d}v \leq \frac{1}{K_0(\mathbf{0})}\mathrm{Vol}(B)C_1^2\mathrm{e}^{-2C_2 d(A,B)}, \qquad (3.8)
$$

which, together with the boundedness of $g$ and the bound $\|M_{\mathbf{0}}\|_2 \leq \|M_{\mathbf{0}}\|_{p_0} < \infty$, ensures that the second term is bounded above by

$$
2\mathrm{Vol}(A)\mathrm{Vol}(B)\mathbb{E}(|M_{\mathbf{0}}|)C_1^2\mathrm{e}^{-2C_2 d(A,B)} := \mathrm{Vol}(A)\mathrm{Vol}(B)C_5\mathrm{e}^{-C_4 d(A,B)}. \qquad (3.9)
$$

Combining (3.6), (3.7) and (3.9) gives

$$\left|\mathbb{E}\left(H_{\lambda,A}^2 g(H_{\lambda,B})\right) - \mathbb{E}\left(H_{\lambda,A}^2\right)\mathbb{E}\left(g(H_{\lambda,B})\right)\right| \le (\mathrm{Vol}(A)^2 \vee 1)\mathrm{Vol}(B)(C_3 + C_5)\mathrm{e}^{-C_4 d(A,B)}. \tag{3.10}$$

For $j = 1$ and $g(x) = x$, we have

$$\begin{aligned}
&|\mathbb{E}(H_{\lambda,A}H_{\lambda,B}) - \mathbb{E}(H_{\lambda,A})\mathbb{E}(H_{\lambda,B})| \\
&= \left|\mathbb{E}\iint_{A\times B} M_{x_1}M_{x_2}\mathcal{P}(\mathrm{d}x_1)\mathcal{P}(\mathrm{d}x_2) - \left(\mathbb{E}\int_A M_x\mathcal{P}(\mathrm{d}x)\right)\left(\mathbb{E}\int_B M_y\mathcal{P}(\mathrm{d}y)\right)\right| \\
&= \left|\int_A K_0(\mathbf{0})\mathbb{E}(M_{\mathbf{0}})\mathbb{E}\left(\int_B M_y\mathcal{P}_x^!(\mathrm{d}y) - \int_B \tilde{M}_y\mathcal{P}'(\mathrm{d}y)\right)dx\right| \\
&\le 2\mathbb{E}(M_{\mathbf{0}})^2\,\mathrm{Vol}(A)\,\mathrm{Vol}(B)C_1^2\mathrm{e}^{-2C_2 d(A,B)},
\end{aligned} \tag{3.11}$$

where the last inequality follows from (3.8). If $g \in \mathcal{F}_b$ instead, replacing $M_{x_2}$ in (3.11) by $g(M_{x_2})$ yields

$$\begin{aligned}
&|\mathbb{E}(H_{\lambda,A}g(H_{\lambda,B})) - \mathbb{E}(H_{\lambda,A})\mathbb{E}(g(H_{\lambda,B}))| \\
&= \left|\int_A K_0(\mathbf{0})\mathbb{E}(M_{\mathbf{0}})\mathbb{E}\left(\int_B g(M_y)\mathcal{P}_x^!(\mathrm{d}y) - \int_B g(\tilde{M}_y)\mathcal{P}'(\mathrm{d}y)\right)dx\right| \\
&\le 2|\mathbb{E}(M_{\mathbf{0}})|\,\mathrm{Vol}(A)\,\mathrm{Vol}(B)C_1^2\mathrm{e}^{-2C_2 d(A,B)}.
\end{aligned} \tag{3.12}$$

Combining (3.11), (3.12), and the moment properties of $M_{\mathbf{0}}$ discussed above, we obtain:

$$|\mathbb{E}\left(H_{\lambda,A}g(H_{\lambda,B})\right) - \mathbb{E}\left(H_{\lambda,A}\right)\mathbb{E}\left(g(H_{\lambda,B})\right)| \le \mathrm{Vol}(A)\mathrm{Vol}(B)C_6\mathrm{e}^{-C_4 d(A,B)}.$$

Together with (3.10) and Lemma 2.5, this implies the exponential mixing of score sums (1.5) for all disjoint $A, B$ without requiring that they satisfy a separation condition. We have $\gamma_0 = \gamma_0' + d = 2d + d = 3d$, $\gamma_2 = \gamma_2' = C_4$, and any $\gamma_3 > 0$, since $\mathrm{Vol}(D) \le (\mathrm{diam}(D))^d$ for all $D \in \mathcal{B}(\mathbb{R}^d)$. The proof now follows from Theorem 1.6 (a). □

## 3.2 Unevenly spaced time series

Consider an unevenly spaced time series, namely a sequence of random events $\{(T_i, U_i)\}_{i\in\mathbb{Z}:=\{\dots,-2,-1,0,1,2,\dots\}}$, where the time intervals between successive recording times $\{T_i\}_{i\in\mathbb{Z}}$ are not constant. A common approach to analyzing such data involves transforming it into equally spaced observations through interpolation (see Tong [42], Brockwell and Davis [7], and references therein). We take a different approach and work with the untransformed data.

In this section, we use Theorem 1.6 (b) to establish a quantitative CLT for sums of score functions driven by the sequence $\{(T_i, U_i)\}_{i\in\mathbb{Z}}$, where $\{T_i\}_{i\in\mathbb{Z}}$ is an increasing sequence forming a stationary renewal point process, with $T_1$ being the first renewal after time 0, and $\{U_i\}_{i\in\mathbb{Z}}$ a strictly stationary sequence of random variables independent of $\{T_i\}_{i\in\mathbb{Z}}$. The process $\{(T_i, U_i)\}_{i\in\mathbb{Z}}$ is a marked point process, with the stationary renewal point process as the underlying ground point process $\mathcal{P} = \sum_{i\geq 1} \delta_{T_i}$ and with marks $U_i$ associated with the points $T_i$ taking values in a given measurable mark space $\mathbb{M}$. We do not assume that $\mathcal{P}$ satisfies the EDD property.

The behavior of a renewal point process depends on the properties of its renewal times. The renewal point process is called *non-lattice* if there do not exist constants $a \in \mathbb{R}$ and $b \in (0,\infty)$ such that the support of the inter-arrival times is a subset of $a+b\mathbb{N}$. It is called *spread-out* if there exists a positive integer $i$ such that the $i$-th convolution of the distribution function, $G$, of the inter-arrival times has an absolutely continuous component. That is, $G^{*i} = c_1G_a+c_2G_s$ holds for two constants $c_1 \in (0,\infty),\ c_2 = 1-c_1 \geq 0$, where $G_a$ is an absolutely continuous distribution and $G_s$ is another distribution. If a renewal point process is spread-out, then it is non-lattice. However, the reverse is not true: if the inter-arrival distribution is discrete and takes all positive rational values with positive probability, then the process is non-lattice but not spread-out.

The dependence structure of the marks is as follows. Fix $k_0 \in \mathbb{N}$. Each mark $U_i$ depends on the values of the preceding $k_0$ marks, with additional dependence on $R_i$, where $\{R_i\}_{i\in\mathbb{Z}}$ is a sequence of *i.i.d.* random variables taking values in some measurable space $(\mathcal{E}, \mathscr{E})$. We assume

$$U_i = f(\{U_j\}_{i-k_0\leq j<i}, R_i),\ i \in \mathbb{Z}, \tag{3.13}$$

where $f : \mathbb{M}^{k_0} \times \mathcal{E} \to \mathbb{M}$ is measurable, and $R_i$ is independent of both $\{T_j\}_{j\in\mathbb{Z}}$ and $\{U_j\}_{j\leq i-1}$. This assumption includes the autoregressive process of order $k_0$ as a special case and guarantees that $\{U_i\}_{i\in\mathbb{Z}}$ is strictly stationary. Under this setting, the marked point process $\hat{\mathcal{P}} = \sum_{i\in\mathbb{Z}} \delta_{(T_i,U_i)}$ on $\mathbb{R}\times\mathbb{M}$ reduces to a point-process representation of an AR($k_0$) model when $T_i = i$ for $i \in \mathbb{Z}$, that is, when the renewal point process is lattice-based.

Define $\xi : (\mathbb{R}\times\mathbb{M}) \times \mathcal{N}_{\mathbb{R}\times\mathbb{M}} \to \mathbb{R}$ by

$$\xi((T_i, U_i), \hat{\mathcal{P}}) = \xi_{loc}\left((T_i, U_i), \hat{\mathcal{P}}|_{[T_{i-j_0}\wedge(T_i-t_0), T_{i+j_0}\vee(T_i+t_0)]}\right) \tag{3.14}$$

for all $i \in \mathbb{Z}$, where $j_0 \in \mathbb{N}\cup\{0\}$ and $t_0 \geq 0$ are constants, and where $\xi_{loc} : (\mathbb{R}\times\mathbb{M}) \times \mathcal{N}_{\mathbb{R}\times\mathbb{M}} \to \mathbb{R}$ characterises the interaction between a renewal point (together with its mark) and its surrounding window in time, as measured by $t_0$ and/or the number of renewals $j_0$. We define $H_\lambda = \sum_{T_i\in\mathcal{P}\cap\Gamma_\lambda} \xi((T_i, U_i), \hat{\mathcal{P}})$. Statistics $H_\lambda$ of interest are generated by choosing $\xi$ as follows:

(i) $\xi((T_i, U_i), \hat{\mathcal{P}}) = U_i$, in which case $H_\lambda$ gives the total sum of a strictly stationary sequence of random variables counted by the renewal process;

(ii) $\xi((T_i, U_i), \hat{\mathcal{P}}) = (U_i - \mathbb{E}U_i)(U_{i-j} - \mathbb{E}U_{i-j})$, in which case $H_\lambda$ becomes, up to a multiplicative constant, the estimator of the $j$-th autocovariance in time series, cf. Hamilton [17, Section 3.1];

(iii) $\xi((T_i, U_i), \hat{\mathcal{P}}) = g(T_i - T_{i-1})$ for some function $g : \mathbb{R} \to \mathbb{R}$, when one studies the renewal-reward process (cf. Ross [36, Section 7.4]); and

(iv) $\xi((T_i, U_i), \hat{\mathcal{P}}) = h(U_i)$ for some function $h : \mathbb{M} \to \mathbb{R}$; when $k_0 = 1$ and $\{T_i\}_{i\in\mathbb{Z}}$ reduces to the lattice $\mathbb{Z}$, one obtains the discrete version of the additive Markov functionals, cf. Revuz [33].

Since the ground process and the marks arise from an iterative approach, we impose a decay-of-dependence condition on the marks. This is achieved through the *$\phi$-mixing coefficient* between two $\sigma$-algebras $\mathscr{A}$ and $\mathscr{B}$, defined as

$$\phi(\mathscr{A}, \mathscr{B}) := \sup_{E\in\mathscr{A}, E'\in\mathscr{B}, \mathbb{P}(E)>0} |\mathbb{P}(E'|E) - \mathbb{P}(E')| .$$

We define $H_\lambda = \sum_{T_i\in\mathcal{P}\cap\Gamma_\lambda} \xi((T_i, U_i), \hat{\mathcal{P}})$, $\sigma_\lambda^2 = \operatorname{Var} H_\lambda$, and $Z_\lambda = (H_\lambda - \mathbb{E}H_\lambda)/\sigma_\lambda$, where $\Gamma_\lambda = [-0.5\lambda, 0.5\lambda]$. The main result of this section is as follows:

**Theorem 3.4.** *(normal approximation for statistics of a spread-out renewal point process) Given a marked stationary spread-out renewal point process, assume that the inter-arrival times have a finite $\alpha$-th moment for some $\alpha > 2$, $\mathbb{E}_{\mathbf{0}}\left(|\xi((\mathbf{0}, U_{\mathbf{0}}), \hat{\mathcal{P}})|^p\right) < \infty$ for some constant $p > p_0 \geq 3$, and*

$$\phi(\sigma(U_1, U_2, \ldots, U_{k_0}), \sigma(U_{n_0+1}, U_{n_0+2}, \ldots, U_{n_0+k_0})) < \frac{1}{2} \tag{3.15}$$

*for some $n_0 \in \mathbb{N}, n_0 > k_0$. If $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ for some $\nu > 2/3$, and*

$$\frac{(36 - 6\nu) \vee 28}{3\nu - 2} < \frac{\alpha(p_0 - 2)}{p_0},$$

*then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(\gamma_0,\gamma_2)}\right) = O\left(\lambda^{-1.5\nu+1+3p_0\left(\frac{2\nu+4}{6p_0+\alpha(p_0-2)} \vee \frac{\nu+4}{2p_0+\alpha(p_0-2)}\right)}\right)$$
$$= o(1),$$

*where $Z \sim N(0, 1)$, $\gamma_0 = 2$, $\gamma_2 = \alpha(p_0 - 2)/(2p_0)$ and $\tau$ is defined in (1.9).*

**Remark 3.5.** *When the ground process $\mathcal{P}$ is a homogeneous Poisson point process, we can apply Theorem 1.6 (a) instead of (b) and use the same approach to derive a better bound. We omit the details.*

The outline of the proof is as follows. We apply Theorem 1.6 (b), where the main challenge is to establish the polynomial decay property in Definition 1.4. Since $d = 1$, $A$ is an interval and $B$ lies on both sides of $A$ with $d(A, B) = \rho$. The estimate of $\operatorname{Cov}(S_A, g(S_B))$ reduces to accounting for the cost of replacing $S_B$ by independent copies of $S_{B_1}$ and $S_{B_2}$, where $B_1 \subseteq B$ is the interval to the left of $A$ and $B_2 \subseteq B$ is the interval to the right.

To carry out this replacement, we must also account for the number of renewals between $A$ and $B_1$ and also between $A$ and $B_2$. Lemma 3.6 shows that the number of renewals is roughly linear in the length of the interval of interest, while Lemma 3.7 shows that the cost of replacing each piece in the ground process by an independent copy can be controlled using the renewal point process property: given sufficient time, the renewal point process returns to stationarity, allowing us to apply the same argument one interval at a time. The cost of replacing the marks is also controllable by considering the lifted mark process $\{\mathbf{U}_i\}_{i\in\mathbb{Z}}$, where $\mathbf{U}_i := (U_i, U_{i+1}, \ldots, U_{i+k_0-1})$, as a Markov process and using the $\phi$-mixing property of the Markov process, as in Bradley [6], to control $\phi$-mixing on $\sigma$-algebras distant $\rho$ from each other. The final step is to combine these two lemmas; the complexity arises from the fact that each of $A$, $B_1$ and $B_2$ requires two-sided neighborhoods to allow for the dependence of the marks.

Now we consider the asymptotic linearity of the number of renewals in an interval.

**Lemma 3.6.** *(asymptotic linearity of the number of renewals) Given a stationary non-lattice renewal point process $\{T_i\}$, assume the inter-arrival times $\tau_i := T_{i+1} - T_i$ satisfy $\mathbb{E}(\tau_i^\alpha) < \infty$ for some $\alpha > 2$. Then for any constant $C \in (0, (2\mathbb{E}\tau_1)^{-1})$, there exist positive constants $C'$ and $C''$ depending on $C$ such that the number of renewals $N_t$ in the time interval $[0,t]$ satisfies*

$$\mathbb{P}(N_t \le Ct) \le C' t^{-\alpha/2} \text{ for } t \ge C''.$$

*Proof.* Let $G$ be the distribution of the inter-arrival times with mean $\mu_G = \mathbb{E}(\tau_1)$, and let $N_{o,t} + 1$ represent the number of renewals in the interval $[0,t]$, conditioned on a renewal at 0. Let $F_{T_1}$ be the distribution function of $T_1$ and let $n_t = \lfloor Ct \rfloor$. Then

$$\begin{aligned}
\mathbb{P}(N_t \le Ct) &= \int_0^\infty \mathbb{P}(N_t \le n_t | T_1 = s)\, \mathrm{d}F_{T_1}(s) \\
&\le \int_0^{t/2} \mathbb{P}(N_{o,t-s} + 1 \le n_t)\, \mathrm{d}F_{T_1}(s) + \mathbb{P}\left(T_1 > \frac{t}{2}\right) \\
&\le \mathbb{P}\left(N_{o,t/2} \le n_t - 1\right) + \mathbb{P}\left(T_1 > \frac{t}{2}\right). \qquad (3.16)
\end{aligned}$$

Now, since $\mathbb{E}(T_1^{\alpha-1}) = \mathbb{E}(\tau_1^\alpha)/(\alpha\mu_G) < \infty$ and $(t/2)^{\alpha-1}\mathbb{P}(T_1 > t/2) \le \int_{t/2}^\infty s^{\alpha-1}\mathrm{d}F_{T_1}(s) \to 0$ as $t \to \infty$, we obtain

$$\mathbb{P}\left(T_1 > \frac{t}{2}\right) = o(t^{-(\alpha-1)}). \qquad (3.17)$$

For the first term in (3.16), let $Y_i = \tau_i - \mu_G$, $S_n := \sum_{i=1}^n \tau_i$, and $\zeta_n := \sum_{i=1}^n Y_i$. The probability $\mathbb{P}(N_{o,t/2} \le n_t - 1)$ is related to the sum $S_{n_t}$, which can be expressed as follows:

$$\mathbb{P}\left(N_{o,t/2} \le n_t - 1\right) = \mathbb{P}\left(S_{n_t} > \frac{t}{2}\right) = \mathbb{P}\left(\zeta_{n_t} > \frac{t}{2} - n_t\mu_G\right). \qquad (3.18)$$

Next, we define $w_t = t/2 - n_t\mu_G > 0$, and apply Chebyshev's inequality followed by a Rosenthal inequality in Rosenthal [35, p. 279], which states that given $\alpha > 2$, there exists a positive constant $C_\alpha$ depending only on $\alpha$ such that for any sequence of independent random variables $Y_i$ in the space $L^\alpha$ with $\mathbb{E}Y_i = 0$, $1 \le i \le n$, we have

$$\left\|\sum_{i=1}^n Y_i\right\|_\alpha \le C_\alpha \max\left\{\left(\sum_{i=1}^n \mathbb{E}\left(|Y_i|^\alpha\right)\right)^{1/\alpha}, \left(\sum_{i=1}^n \mathbb{E}\left(Y_i^2\right)\right)^{1/2}\right\}.$$

This gives

$$\mathbb{P}(\zeta_{n_t} > w_t) \le \frac{\mathbb{E}(|\zeta_{n_t}|^\alpha)}{w_t^\alpha} \le c_\alpha w_t^{-\alpha} \max\left\{n_t\mathbb{E}(|Y_1|^\alpha), n_t^{\alpha/2}(\mathbb{E}(Y_1^2))^{\alpha/2}\right\},$$

where $c_\alpha := C_\alpha^\alpha$ is a positive constant and hence

$$\mathbb{P}(\zeta_{n_t} > w_t) = O\left(\frac{n_t^{\alpha/2}}{w_t^\alpha}\right) = O(t^{-\alpha/2}). \tag{3.19}$$

Combining (3.16) - (3.19) completes the proof. □

**Lemma 3.7.** *(the cost of replacing renewals on intervals by independent copies) Let $\{T_i\}$ be a stationary, spread-out renewal point process, where the inter-arrival times have a finite $\alpha$-th moment for some $\alpha > 2$. Let $I_i = [a_i, b_i]$ for $i = 1, \ldots, m$, where $m \ge 2$ and $-\infty \le a_1 < b_1 < a_2 < b_2 < \cdots < a_m < b_m \le \infty$, with the convention $(a, \infty] = (a, \infty)$ and $[-\infty, b] = (-\infty, b]$. Let $\mathscr{R}_i$ be the restriction of the renewal point process to the interval $I_i$, and let $\mathscr{R}_i'$ be a copy of $\mathscr{R}_i$, such that $\{\mathscr{R}_i' : 1 \le i \le m\}$ are independent of each other. For any constant $C \in (0, \infty)$, if the minimum separation distance between intervals satisfies $\min_{1\le i\le m-1}(a_{i+1} - b_i) \ge Ct$, then we have as $t \to \infty$*

$$d_{TV}\left((\mathscr{R}_1, \ldots, \mathscr{R}_m), (\mathscr{R}_1', \ldots, \mathscr{R}_m')\right) = o\left(t^{-(\alpha-1)}\right).$$

*Proof.* Without loss of generality, we assume that $\{\mathscr{R}_i\}$ and $\{\mathscr{R}_i'\}$ are defined on the same probability space and are independent. Let $\mathcal{R}_i = \left(\mathscr{R}_1, \ldots, \mathscr{R}_i, \mathscr{R}_{i+1}', \ldots, \mathscr{R}_m'\right)$ for $1 \le i \le m$ with the convention $\mathcal{R}_m = (\mathscr{R}_1, \ldots, \mathscr{R}_m)$. The triangle inequality gives

$$d_{TV}\left((\mathscr{R}_1, \ldots, \mathscr{R}_m), (\mathscr{R}_1', \ldots, \mathscr{R}_m')\right) \le \sum_{i=2}^m d_{TV}\left(\mathcal{R}_i, \mathcal{R}_{i-1}\right),$$

and hence it suffices to show that

$$d_{TV}\left(\mathcal{R}_i, \mathcal{R}_{i-1}\right) = o\left(t^{-(\alpha-1)}\right)$$

for all $2 \le i \le m$.

To this end, let $\Omega$ denote all locally finite subsets of $\cup_{j=1}^m I_j$, representing all possible configurations of the renewal point processes restricted to $\cup_{j=1}^m I_j$, and let $\Omega_i$ represent all locally finite subsets of $\left(\cup_{j=1}^m I_j\right)\backslash I_i$, representing all possible configurations of the renewal point processes restricted to $\left(\cup_{j=1}^m I_j\right)\backslash I_i$. Each configuration can be viewed as a finite integer-valued measure, and we equip both $\Omega$ and $\Omega_i$ with the weak topology. Let $\mathcal{F}$ and $\mathcal{F}_i$ be the $\sigma$-algebras generated on $\Omega$ and $\Omega_i$, respectively. For ease of notation, let $\mathcal{R}_{i-} = \left(\mathscr{R}_1, \ldots, \mathscr{R}_{i-1}, \mathscr{R}'_{i+1}, \ldots, \mathscr{R}'_m\right)$. Without loss of generality, assume $b_{i-1} = 0$, so we can use the notation $T_1$ to denote the first renewal time after $b_{i-1}$. Then, we have

$$
\begin{aligned}
&d_{TV}\left(\mathcal{R}_i, \mathcal{R}_{i-1}\right) \\
&= \sup_{A\in\mathcal{F}} \left|\mathbb{P}\left(\mathcal{R}_i \in A\right) - \mathbb{P}\left(\mathcal{R}_{i-1} \in A\right)\right| \\
&\le \sup_{A\in\mathcal{F}} \int_{\Omega_i}\int_0^\infty \left|\mathbb{P}\left(\mathcal{R}_i \in A | \mathcal{R}_{i-} = \xi, T_1 = s\right) - \mathbb{P}\left(\mathcal{R}_{i-1} \in A | \mathcal{R}_{i-} = \xi, T_1 = s\right)\right| \\
&\qquad \times \mathbb{P}\left(\mathcal{R}_{i-} \in \mathrm{d}\xi, T_1 \in \mathrm{d}s\right) \\
&\le \int_{\Omega_i}\int_0^\infty d_{TV}\left(\mathcal{L}(\mathscr{R}_i | \mathcal{R}_{i-} = \xi, T_1 = s), \mathcal{L}(\mathscr{R}_i)\right) \mathbb{P}\left(\mathcal{R}_{i-} \in \mathrm{d}\xi, T_1 \in \mathrm{d}s\right) \\
&\le \int_{\Omega_i}\int_0^{0.5Ct} d_{TV}\left(\mathcal{L}(\mathscr{R}_i | \mathcal{R}_{i-} = \xi, T_1 = s), \mathcal{L}(\mathscr{R}_i)\right) \mathbb{P}\left(\mathcal{R}_{i-} \in \mathrm{d}\xi, T_1 \in \mathrm{d}s\right) \\
&\quad + \mathbb{P}\left(T_1 > 0.5Ct\right).
\end{aligned} \tag{3.20}
$$

Conditional on $T_1 = s$, the distribution of $\mathscr{R}_i$ does not depend on $\mathcal{R}_{i-}$, and hence the bound (3.20) simplifies to

$$
\begin{aligned}
&d_{TV}\left(\mathcal{R}_i, \mathcal{R}_{i-1}\right) \\
&\le \int_0^{0.5Ct} d_{TV}\left(\mathcal{L}(\mathscr{R}_i | T_1 = s), \mathcal{L}(\mathscr{R}_i)\right) \mathbb{P}\left(T_1 \in \mathrm{d}s\right) + \mathbb{P}\left(T_1 > 0.5Ct\right).
\end{aligned} \tag{3.21}
$$

Now, we use the coupling defined in Hernández Ruíz [18, Section 11.2] to complete the proof. Let $\{T_{o,i}\}$, with $T_{o,0} = 0$, be the zero-delayed renewal point process where the inter-renewal times have the same distribution as that of $\tau_1$. From the proof in [18, p. 435–436], we know that there exists a positive random variable $\mathscr{T}$ and a coupling $\left(\{T'_{o,i}\}, \{T'_i\}\right)$ such that (i) $\mathcal{L}\left(\{T'_{o,i}\}\right) = \mathcal{L}\left(\{T_{o,i}\}\right)$ and $\mathcal{L}\left(\{T'_i\}\right) = \mathcal{L}\left(\{T_i\}\right)$, (ii) the renewal times of $\{T'_{o,i}\}$ coincide with those of $\{T'_i\}$ at and after $\mathscr{T}$, (iii) $\mathbb{E}\left[\mathscr{T}^{\alpha-1}\right] < \infty$ [18, Lemma 3.1]. Using this coupling, we have for all $s \in [0, 0.5Ct]$,

$$
\begin{aligned}
&d_{TV}\left(\mathcal{L}(\mathscr{R}_i | T_1 = s), \mathcal{L}(\mathscr{R}_i)\right) \\
&\le \mathbb{P}(\{T'_{o,i}\} \cap [a_i - s, b_i - s] \neq \{T'_i\} \cap [a_i - s, b_i - s]) \\
&\le \mathbb{P}(\mathscr{T} > a_i - s).
\end{aligned} \tag{3.22}
$$

Thus, using the same proof as that for (3.17), from (3.21) and (3.22), we obtain as $t \to \infty$

$$
d_{TV}\left(\mathcal{R}_i, \mathcal{R}_{i-1}\right) \le \mathbb{P}(\mathscr{T} > a_i - 0.5Ct) + \mathbb{P}\left(T_1 > 0.5Ct\right) = o\left(t^{-(\alpha-1)}\right),
$$

since $\mathbb{E}\left(\mathscr{T}^{\alpha-1}\right) < \infty$ and $a_i - 0.5Ct \ge 0.5Ct$, where we have assumed that $b_{i-1} = 0$. This completes the proof. $\square$

*Proof of Theorem 3.4.* We deduce this from Theorem 1.6 (b). The proof has multiple steps.

Step (i). Showing moment assumptions on ground point process and on the score.

We begin by verifying the moment assumptions (1.1), (1.2). Denote by $N_t$ and $N_{o,t}+1$ the numbers of renewals of the process $\{T_i\}_{i\in\mathbb{Z}}$ and its 0-delayed counterpart $\{T_{o,i}\}$, respectively, in the time interval $[0,t]$. From the construction, it follows that $N_t$ is stochastically dominated by $N_{o,t}+1$ for all $t$.

To verify the moment condition on the ground process (1.1), it is sufficient to show that $N_{o,1}$ has finite moments of arbitrary order. Since the inter-arrival times $\tau_i = T_{i+1}-T_i$ are positive with a positive probability, there exists a positive $x_0$ such that $\mathbb{P}(\tau_1 \geq x_0) > p'$ for some constant $p' > 0$. Therefore, $N_{o,1}$ is stochastically dominated by a negative-binomial random variable with parameters $r = \lceil 1/x_0 \rceil + 1$ and $p = p'$, which has finite moments of arbitrary order. Thus, assumption (1.1) holds for all large $p_1$.

The assumed $p$-th moment condition with $p > 3$ together with arbitrarily large $p_1$ ensures the moment conditions in Theorem 1.6 (b) according to Remark 1.7 (iii).

Step (ii). Constructing a partition of $\Gamma_\lambda$ and a coupling on this partition.

We aim to verify polynomial mixing of score sums in (1.6) with $A = [a,b]$ in (1.6) and $\Gamma_\lambda = [-0.5\lambda, 0.5\lambda]$. Without loss, put $B = \Gamma_\lambda \setminus B_\rho(A)$ for some positive $\rho$. If $B \subset \Gamma_\lambda \backslash B_\rho(A)$, the argument works with the same coupling. Since $d = 1$, $A$ is an interval and $B$ lies on both sides of $A$. As mentioned earlier, the estimate of $\mathrm{Cov}(S_A, g(S_B))$ reduces to accounting for the cost of replacing $S_B$ by independent copies of $S_{B_1}$ and $S_{B_2}$, where $B_1$ is the subset of $B$ to the left of $A$ and $B_2$ is the subset to the right. To do this, we introduce a partition to help control dependencies of renewals and dependencies of marks, which goes as follows.

Without loss of generality, we can assume that $\max\{|a-\rho|, |b+\rho|\} < 0.5\lambda$ and $\rho \geq 3t_0$, which, together with Remark 1.5 (i), also ensure that $A$ and $B$ satisfy the condition in Definition 1.4 with $\gamma_3 = 3t_0$. These conditions ensure that $B = [-0.5\lambda, a-\rho] \cup [b+\rho, 0.5\lambda] =: B_1 \cup B_2$ lies on both sides of $A$, and in the following construction, the neighborhoods of intervals $A$, $B_1$ and $B_2$ all have length greater than or equal to $t_0$. When $\rho$ is sufficiently large, with high probability, the score functions evaluated at renewals in $A$ are determined by renewal points in the $\rho/3$-enlarged interval $B_{\rho/3}(A)$ given by

$$A_e := \left[a - \frac{\rho}{3},\ b + \frac{\rho}{3}\right]$$

together with their marks, while the score functions with centers in $B$ are determined by the renewal points in an enlarged interval around $B$, namely

$$\left(-\infty,\ a - \frac{2\rho}{3}\right] \cup \left[b + \frac{2\rho}{3},\ \infty\right) =: B_{1,e} \cup B_{2,e}$$

together with their marks. The gap between these two sets,

$$D := \left(a - \frac{2\rho}{3},\ a - \frac{\rho}{3}\right) \cup \left(b + \frac{\rho}{3},\ b + \frac{2\rho}{3}\right) =: D_1 \cup D_2,$$

is used with high probability to break the dependence between the sums of score functions with arrival times in sets $A$ and $B$. Since $t_0$ also plays a role in the dependence range of the score function and since it extends the dependence range even further, we further define the following intervals:

$$B_{1,re} = \left(a - \rho, a - \frac{2\rho}{3}\right], \quad B_{2,le} = \left[b + \frac{2\rho}{3}, b + \rho - t_0\right),$$
$$A_{le} = \left[a - \frac{\rho}{3}, a - t_0\right), \quad A_{re} = \left(b, b + \frac{\rho}{3}\right].$$

The above intervals are illustrated in Figure 1.

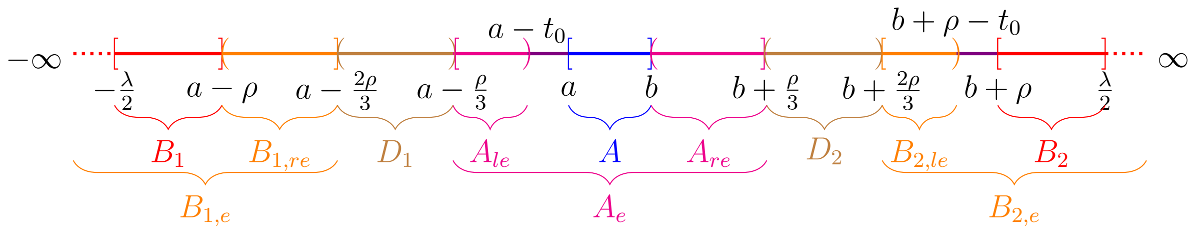


**Figure 1:** Interval partition of $\Gamma_\lambda$ for polynomial mixing

The renewal points in the stationary renewal point process $\mathcal{P}$ on a given interval $I$ are simply the points of the process $\mathcal{P}$ that fall within the interval $I$, expressed as $\mathcal{P}|_I$. From Lemma 3.7 (with $t$ and $C$ replaced by $\rho/3$ and 1, respectively) and extending the probability space if necessary, we can construct, on the same probability space, a copy $\{\mathcal{P}'|_{B_{1,e}}, \mathcal{P}'|_{B_{2,e}}, \mathcal{P}'|_{A_e}\}$ of $\{\mathcal{P}|_{B_{1,e}}, \mathcal{P}|_{B_{2,e}}, \mathcal{P}|_{A_e}\}$, such that $\mathcal{P}'|_{B_{1,e}}, \mathcal{P}'|_{B_{2,e}}, \mathcal{P}'|_{A_e}$ are independent and

$$E_1 := \left\{(\mathcal{P}|_{B_{1,e}}, \mathcal{P}|_{B_{2,e}}, \mathcal{P}|_{A_e}) = (\mathcal{P}'|_{B_{1,e}}, \mathcal{P}'|_{B_{2,e}}, \mathcal{P}'|_{A_e})\right\}$$

satisfies

$$\mathbb{P}(E_1^c) \leq o\left(\rho^{-(\alpha-1)}\right).$$

To fill the gap of $\mathcal{P}'$ on $D$ and to complete the definition of $\mathcal{P}'$, we define $\mathcal{P}'|_D = \mathcal{P}|_D$. It is important to note that the inter-arrival times of $\mathcal{P}'$ are generally no longer independent and identically distributed.

Step (iii). Breaking the direct dependence of the score function outside the neighborhood and controlling the dependence between marks within the neighborhoods.

To break the direct dependence in the marks, we require a sufficient number of renewal points in each of the three sub-intervals between $B_1$ and $A$ and in each of the three sub-intervals between $A$ and $B_2$; hence we set

$$E_2 = \{\min\{\mathcal{P}(B_{1,re}), \mathcal{P}(B_{2,le}), \mathcal{P}(A_{le}), \mathcal{P}(A_{re}), \mathcal{P}(D_1), \mathcal{P}(D_2)\} \geq C_1\rho\},$$

where $C_1 := (4\mathbb{E}\tau_1)^{-1} \in (0, (2\mathbb{E}\tau_1)^{-1})$ as in Lemma 3.6, and let $E_2'$ be the corresponding event for $\mathcal{P}'$. Lemma 3.6 ensures that

$$\mathbb{P}(E_2^c) = \mathbb{P}((E_2')^c) = O\left(\rho^{-\alpha/2}\right).$$

Without loss of generality, we assume that $\rho \geq (k_0 + j_0)/C_1$ (in addition to $\rho \geq 3t_0$) for $k_0$ and $j_0$ specified in (3.13) and (3.14). This condition ensures that we can rule out any direct dependence of the score functions in $A$, $B_1$, and $B_2$ on the marks $U$ outside their respective neighborhoods $A_e$, $B_{1,e}$ and $B_{2,e}$ on event $E_2$. The requirement on the number of points in $D_1$ and $D_2$ also guarantees that the dependence of the marks within these neighborhoods is controllable on event $E_2$, which will be detailed later in (3.28).

Step (iv). Constructing the marked point process $\widehat{\mathcal{P}'}\Big|_{\Gamma_\lambda}$ given $\widehat{\mathcal{P}}\Big|_{\Gamma_\lambda}$.

To equate $\widehat{\mathcal{P}}|_{\Gamma_\lambda}$ with its marked counterpart $\widehat{\mathcal{P}'}\Big|_{\Gamma_\lambda}$ with high probability, we reuse the marks associated with $\mathcal{P}|_{\Gamma_\lambda}$, say $\{U_l, U_{l+1}, \dots\}$, and assign them in the same order to the points of $\mathcal{P}'|_{\Gamma_\lambda}$ (for example, $U_l$ goes to the smallest point of $\mathcal{P}'|_{\Gamma_\lambda}$, $U_{l+1}$ goes to the next smallest, and so on). The reason for constructing $\widehat{\mathcal{P}'}\Big|_{\Gamma_\lambda}$ is that the marks on the intervals $B_{1,e} \cup B_{2,e} \cup A_e$ are determined by the configuration of renewals beyond these intervals, as stated in (3.13). The advantage of $\widehat{\mathcal{P}'}|_{\Gamma_\lambda}$ is that it permits us to use the independent restrictions of $\mathcal{P}'|_{\Gamma_\lambda}$ to eliminate the dependence of the marks. From this construction, we observe that the corresponding $S_A'$ and $S_B'$ of $\widehat{\mathcal{P}'}$ satisfy the following equalities:

$$S_A \mathbf{1}_{E_1 \cap E_2} = S_A' \mathbf{1}_{E_1 \cap E_2'}, \quad S_B \mathbf{1}_{E_1 \cap E_2} = S_B' \mathbf{1}_{E_1 \cap E_2'}.$$

Step (v). Showing that mixing of score sums $S_A'$, $S_B'$ implies mixing of score sums for the function-set class corresponding to $j = 1$ in Definition 1.4.

We aim to replace $S_A$ and $S_B$ with $S_A'$ and $S_B'$ on $E_3 := E_1 \cap E_2$ via the following inequality:

$$\begin{aligned}
&|\mathbb{E}(S_A S_B) - \mathbb{E}(S_A)\mathbb{E}(S_B)| \\
&= \left|\mathbb{E}(S_A(\mathbf{1}_{E_3} + \mathbf{1}_{E_3^c})S_B) - \mathbb{E}(S_A(\mathbf{1}_{E_3} + \mathbf{1}_{E_3^c}))\mathbb{E}(S_B)\right| \\
&\leq |\mathbb{E}(S_A \mathbf{1}_{E_3} S_B) - \mathbb{E}(S_A \mathbf{1}_{E_3})\mathbb{E}(S_B)| + \left|\mathbb{E}(S_A \mathbf{1}_{E_3^c} S_B)\right| + \left|\mathbb{E}(S_A \mathbf{1}_{E_3^c})\mathbb{E}(S_B)\right|. \qquad (3.23)
\end{aligned}$$

Additionally, by Lemma 2.2, since $p_1$ in assumption (1.1) can be arbitrarily large, both $S_A$ and $S_B$ have finite $L^{p_0}$ norms for $p_0$ in the statement of Theorem 3.4, which are bounded above by $C_2\mathrm{Vol}(A)$ and $C_2\mathrm{Vol}(B)$, respectively, for some positive constant $C_2$.

Hölder's inequality yields

$$\begin{aligned}
\left|\mathbb{E}(S_A \mathbf{1}_{E_3^c} S_B)\right| &\leq \|S_A\|_{p_0} \|S_B\|_{p_0} \mathbb{P}(E_3^c)^{(p_0-2)/p_0} \\
&= O\left(\mathrm{Vol}(A)\mathrm{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right), \qquad (3.24)
\end{aligned}$$

and

$$\begin{aligned}\left|\mathbb{E}(S_A\mathbf{1}_{E_3^c})\mathbb{E}(S_B)\right| &\leq \|S_A\|_{p_0}\|S_B\|_{p_0}\mathbb{P}(E_3^c)^{(p_0-2)/p_0}\\ &= O\left(\mathrm{Vol}(A)\mathrm{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right). \end{aligned}\tag{3.25}$$

Combining (3.23), (3.24) and (3.25), we obtain

$$\begin{aligned}&|\mathbb{E}(S_AS_B) - \mathbb{E}(S_A)\mathbb{E}(S_B)|\\ &\leq |\mathbb{E}(S_A\mathbf{1}_{E_3}S_B) - \mathbb{E}(S_A\mathbf{1}_{E_3})\mathbb{E}(S_B)| + O\left(\mathrm{Vol}(A)\mathrm{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right).\end{aligned}\tag{3.26}$$

Each term involving $E_3^c$ incurs a cost of at most $O\left(\mathrm{Vol}(A)\mathrm{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right)$. Therefore, we focus on the leading term that does not involve $E_3^c$, and instead add the error due to the manipulation required to transition from one form to the other. We use the notation $\beta_1 \lhd \beta_2$ to denote that $\beta_1 \leq \beta_2 + O\left(\mathrm{Vol}(A)\mathrm{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right)$. With this, we can obtain from (3.26) the following expression:

$$\begin{aligned}&|\mathbb{E}(S_AS_B) - \mathbb{E}(S_A)\mathbb{E}(S_B)|\\ \lhd\, &|\mathbb{E}(S_AS_B\mathbf{1}_{E_3}) - \mathbb{E}(S_A\mathbf{1}_{E_3})\mathbb{E}(S_B\mathbf{1}_{E_3})|\\ =\, &\left|\mathbb{E}(S'_AS'_B\mathbf{1}_{E'_3}) - \mathbb{E}(S_A\mathbf{1}_{E_3})\mathbb{E}(S_B\mathbf{1}_{E_3})\right|\\ \lhd\, &|\mathbb{E}(S'_AS'_B) - \mathbb{E}(S'_A)\mathbb{E}(S'_B)|,\end{aligned}\tag{3.27}$$

where $E'_3 := E_1 \cap E'_2 = E_3$. To see this, note that by construction, the coupled processes $\mathcal{P}$ and $\mathcal{P}'$ agree on $D_1$ and $D_2$ and on the event $E_1$ they also agree on the relevant neighborhoods of $A$ and $B$ and hence the events $E_2$ and $E'_2$ coincide on $E_1$.

Step (vi). Showing the mixing of score sums $S'_A$, $S'_B$ using classical $\phi$-mixing bounds.

When we lift the marks to higher-dimensional space by defining

$$\mathbf{U}_i = (U_i, U_{i+1}, \ldots, U_{i+k_0-1}) \in \mathbb{M}^{k_0},$$

then $\{\mathbf{U}_i\}_{i\in\mathbb{Z}}$ forms a strictly stationary $k_0$-dimensional Markov process. For each $i$, the mark $U_i$ is the first component of the lifted vector $\mathbf{U}_i$. Let $\mathscr{A}'$ denote the $\sigma$-algebra generated by all the marked points in $\widehat{\mathcal{P}'}$ with the lifted marks $\mathbf{U}_i$, where at least one component is used in $S'_A$. From the construction, we have $S'_A \in \mathscr{A}' \subset \sigma\left(\widehat{\mathcal{P}'}_{A_e}\right)$, $S'_B \in \sigma\left(\widehat{\mathcal{P}'}_{B_{1,e}\cup B_{2,e}}\right)$. From [6, Theorem 3.3 (2)], which states that when (3.15) holds, the $\phi$-mixing coefficient converges to 0 exponentially fast as $n_0$ goes to infinity, along with the fact that $\{\mathcal{P}'_{B_{1,e}}, \mathcal{P}'_{B_{2,e}}, \mathcal{P}'_{A_e}\}$ are independent, the $\phi$-mixing coefficient satisfies

$$\phi(\mathscr{A}', \sigma(S'_B)) = O\left(\mathrm{e}^{-C_3\rho}\right),\tag{3.28}$$

for some positive constant $C_3$.

Moreover, by Davydov [12, Lemma 2.1 and its Corollary], for any two $\sigma$-fields $\mathscr{A}$ and $\mathscr{B}$ with the $\alpha$-coefficient $\alpha(\mathscr{A},\mathscr{B}) = \sup_{E\in\mathscr{A},E'\in\mathscr{B}} |\mathbb{P}(E\cap E') - \mathbb{P}(E)\mathbb{P}(E')|$, we have for any $p,q \in (1,\infty]$ such that $1/p+1/q<1$ and random variables $X\in L^p(\mathscr{A})$, $Y \in L^q(\mathscr{B})$,

$$|\operatorname{Cov}(X,Y)| \le 12\alpha(\mathscr{A},\mathscr{B})^{1-1/p-1/q}\|X\|_p\|Y\|_q.$$

Together with the fact that $\alpha(\mathscr{A},\mathscr{B}) \le \phi(\mathscr{A},\mathscr{B})$ from the definition, this gives

$$|\operatorname{Cov}(X,Y)| \le 12\phi(\mathscr{A},\mathscr{B})^{1-1/p-1/q}\|X\|_p\|Y\|_q. \tag{3.29}$$

Although a bound with a sharper constant is available in Rio [34, (1.3)], we do not pursue this refinement here, since our focus is on the rate of the approximation error. Since $S'_A$ is $\mathcal{A}'$-measurable, applying the covariance inequality (3.29) with $p=q=3$ together with (3.28) and the moment bounds $\|S'_A\|_3 = O(\operatorname{Vol}(A))$ and $\|S'_B\|_3 = O(\operatorname{Vol}(B))$ gives

$$\begin{aligned} |\mathbb{E}(S'_AS'_B) - \mathbb{E}(S'_A)\mathbb{E}(S'_B)| &= |\operatorname{Cov}(S'_A,S'_B)| \le 12\phi(\mathcal{A}',\sigma(S'_B))^{1/3}\|S'_A\|_3\|S'_B\|_3 \\ &= O\left(\operatorname{Vol}(A)\operatorname{Vol}(B)\mathrm{e}^{-C_3\rho/3}\right). \end{aligned} \tag{3.30}$$

Combining (3.27) and (3.30), we obtain

$$|\mathbb{E}(S_AS_B) - \mathbb{E}(S_A)\mathbb{E}(S_B)| = O\left(\operatorname{Vol}(A)\operatorname{Vol}(B)\rho^{-\alpha(p_0-2)/(2p_0)}\right). \tag{3.31}$$

To show mixing of score sums for the function-set class corresponding to $j=2$ in Definition 1.4, we use a similar approach. Abusing notation, let $\beta_1 \lhd \beta_2$ signify that $\beta_1 \le \beta_2 + O\left((\operatorname{Vol}(A))^2\rho^{-\alpha(p_0-2)/(2p_0)}\right)$. Specifically, we have:

$$\begin{aligned} &\left|\mathbb{E}\left(S_A^2 g(S_B)\right) - \mathbb{E}\left(S_A^2\right)\mathbb{E}\left(g(S_B)\right)\right| \\ \lhd &\left|\mathbb{E}\left(S_A^2 g(S_B)\mathbf{1}_{E_3}\right) - \mathbb{E}\left(S_A^2\mathbf{1}_{E_3}\right)\mathbb{E}\left(g(S_B)\mathbf{1}_{E_3}\right)\right| \\ = &\left|\mathbb{E}\left((S'_A)^2 g(S'_B)\mathbf{1}_{E'_3}\right) - \mathbb{E}\left((S'_A)^2\mathbf{1}_{E'_3}\right)\mathbb{E}\left(g(S'_B)\mathbf{1}_{E'_3}\right)\right| \\ \lhd &\left|\mathbb{E}\left((S'_A)^2 g(S'_B)\right) - \mathbb{E}\left((S'_A)^2\right)\mathbb{E}\left(g(S'_B)\right)\right|. \end{aligned} \tag{3.32}$$

Applying the covariance inequality (3.29) with $p=3/2$ and $q=\infty$ gives

$$\begin{aligned} \left|\mathbb{E}((S'_A)^2 g(S'_B)) - \mathbb{E}((S'_A)^2)\mathbb{E}(g(S'_B))\right| &= \operatorname{Cov}\left((S'_A)^2, g(S'_B)\right) \\ &\le 12\phi(\mathcal{A}',\sigma(S'_B))^{1/3}\|(S'_A)^2\|_{3/2}\|g(S'_B)\|_\infty \\ &= O\left(\operatorname{Vol}(A)^2\mathrm{e}^{-C_3\rho/3}\right), \end{aligned}$$

which, together with (3.32), guarantees

$$\left|\mathbb{E}\left(S_A^2 g(S_B)\right) - \mathbb{E}\left(S_A^2\right)\mathbb{E}\left(g(S_B)\right)\right| = O\left(\operatorname{Vol}(A)^2(\operatorname{Vol}(B)\vee 1)\rho^{-\alpha(p_0-2)/(2p_0)}\right). \tag{3.33}$$

Combining (3.31) and (3.33) implies that polynomial mixing of score sums (1.6) is satisfied by taking $\gamma_0 = 2$, letting $\gamma_1$ be the constant implicit in the '$O$' bound, and setting $\gamma_2 = \alpha(p_0 - 2)/(2p_0)$, noting that $\mathrm{Vol}(A) \leq \mathrm{diam}(A)$ for all sets $A$ when $d = 1$. Recall $\gamma_3 = 3t_0$. Therefore, condition (1.8) with $d = 1$ becomes

$$2\gamma_2 = \frac{\alpha(p_0 - 2)}{p_0} > \frac{(36 - 6\nu) \vee 28}{3\nu - 2},$$

as assumed. Theorem 3.4 follows by applying Theorem 1.6 (b). □

## 3.3 Continuum percolation models

For a locally finite point set $V$ in $\mathbb{R}^d$, $d \geq 2$, and $r \geq 1$, by $\mathcal{G}(V, r)$ we mean the graph with vertex set $V$ and edge set $E := \{(u, v); \|u - v\| \leq r,\ u, v \in V^{2,\neq}\}$, where $V^{2,\neq} := \{(u, v); u \neq v,\ u, v \in V\}$. When $V$ is finite, we define $C(V, i) := C(V, i, 1)$ and $L(V, i) := |C(V, i)|$ as the $i$-th largest component of the graph $\mathcal{G}(V, 1)$ and its cardinality, respectively. By convention, we set $C(V, i) = \emptyset$ and $L(V, i) = 0$ if the number of components in $\mathcal{G}(V, 1)$ is less than $i$. Components with the same size are ordered by the lexicographical order of the smallest vertices in each component.

Let $\mathcal{P}_\mu$ be a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity $\mu$. We have $\mathcal{G}(\mathcal{P}_\mu, r) = r^{1/d}\mathcal{G}(r^{-1/d}\mathcal{P}_\mu, 1)$, where multiplying a graph by $r^{1/d}$ means scaling all vertex coordinates by $r^{1/d}$ while keeping the edge structure unchanged, and equality means that both sides have the same vertices and the same edges. Thus, without loss of generality, we focus on $\mathcal{G}(V, r)$ with $r = 1$.

For a point process $\mathcal{P}$ on $\mathbb{R}^d$, $\mathcal{G}(\mathcal{P}, r)$ is the *random geometric graph* generated by $\mathcal{P}$ with radius $r$, a model for continuum percolation. We consider $\mathcal{G}(\mathcal{P} \cap \Gamma_\lambda, 1)$ for a point process $\mathcal{P}$ and a positive constant $\lambda$. The model can be cast in terms of urban agglomeration, and therefore we may refer to each component $C(\mathcal{P} \cap \Gamma_\lambda, i)$ as a *community*.

For a fixed positive integer $l_0 \in \mathbb{N}$, the component/community $C(\mathcal{P} \cap \Gamma_\lambda, i)$ is called *$l_0$-isolated* if its cardinality satisfies $L(\mathcal{P} \cap \Gamma_\lambda, i) \leq l_0$. We let $I_0$ be the indices such that $i \in I_0$ implies that $C(\mathcal{P} \cap \Gamma_\lambda, i)$ is $l_0$-isolated. Each component $C(\mathcal{P} \cap \Gamma_\lambda, i), i \geq 1$, has a range of influence given by $f(L(\mathcal{P} \cap \Gamma_\lambda, i))$, where $f : \mathbb{N} \to [0, \infty)$ is a non-decreasing function.

An $l_0$-isolated community $C(\mathcal{P} \cap \Gamma_\lambda, i), i \in I_0$, is said to be *remote* if its distance to every non-isolated community $C(\mathcal{P} \cap \Gamma_\lambda, j)$, if such communities exist, is greater than $f(L(\mathcal{P} \cap \Gamma_\lambda, j))$. We let $I_0^{rem} \subseteq I_0$ be the indices of the remote components.

We service each remote community by connecting it to its nearest non-isolated community $C(\mathcal{P} \cap \Gamma_\lambda, j), j \notin I_0$. This ensures that all communities are serviced, as we assume that non-remote communities are already serviced.

This construction gives rise to the graph $\mathcal{G}^{iso}(\mathcal{P},\lambda)$ thus obtained by connecting each remote community $C(\mathcal{P}\cap\Gamma_\lambda,i),\ i\in I_0$, to the collection of non-isolated communities $\mathcal{N}^{ni}(\mathcal{P}\cap\Gamma_\lambda):=\{C(\mathcal{P}\cap\Gamma_\lambda,i)\}_{i\notin I_0}$ by adding an edge between $C(\mathcal{P}\cap\Gamma_\lambda,i),\ i\in I_0$, and the nearest point in $\mathcal{N}^{ni}(\mathcal{P}\cap\Gamma_\lambda)$. For a remote community $C(\mathcal{P}\cap\Gamma_\lambda,i)$, if there exist multiple pairs of nearest points between it and $\mathcal{N}^{ni}(\mathcal{P}\cap\Gamma_\lambda)$ at the same distance, one pair is chosen uniformly at random. The existence of isolated communities depends on the configuration of all points. The long-range dependence structure precludes letting the short-range estimator $\hat{\xi}^{[r]}$ be a restricted score function, which is often the natural choice for tackling problems such as this one, and which is customarily employed to define stabilization, cf. [10] and references therein. Instead we will use a different choice for the short-range estimator $\hat{\xi}^{[r]}$. As will be shown in the proof, our result does not depend on the particular method used to connect points in remote communities to those in non-isolated communities.

Given a graph $G$ with vertex set $V(G)$, the *graph distance* $d_G(x,y)$ between $x,y\in V(G)$ is the minimum number of edges connecting $x$ and $y$. By convention, $d_G(x,y)=\infty$ if there is no path between $x$ and $y$. The graph ball centered at vertex $v_0\in V(G)$ and with radius $r\in\mathbb{N}$ is

$$B^G(v_0,r):=G_{\{v;\ d_G(v,v_0)\le r\}},$$

where $G_U$ denotes the subgraph $G$ induced by the vertex set $U$. We now consider the case where the score function

$$\xi(x,\mathcal{P},\Gamma_\lambda)=h(B^{\mathcal{G}^{iso}(\mathcal{P},\lambda)}(x,r_0))$$

is given by a function $h$ defined on the graph ball in $\mathcal{G}^{iso}(\mathcal{P},\lambda)$ of graph radius $r_0$. This type of model can be used to represent transportation costs or urban influence and is particularly relevant in the context of spatial growth and development. While traffic congestion and the reliability of transportation networks are key factors, another critical aspect of development is ensuring logistical connectivity between remote communities and well-serviced urban centres.

Define

$$H_\lambda:=\sum_{x\in\mathcal{P}\cap\Gamma_\lambda}\xi(x,\mathcal{P},\Gamma_\lambda).$$

We note that the score function $\xi$ is typically not stabilizing, as its evaluation may require information about the entire configuration $\mathcal{P}\cap\Gamma_\lambda$ in order to determine the existence of remote communities. This statistic can be used to analyze commonly studied problems in the literature, including counting subgraphs that are isomorphic to a given graph or counting components of size $r_0$ for some $r_0\in\mathbb{N}$. With a simple adjustment, the same idea can also be used to analyze these problems for random geometric graphs in the continuum percolation model.

Let $\mathcal{P}_\mu^{\mathbf{0}}$ denote $\mathcal{P}_\mu\cup\{\mathbf{0}\}$, where $\mathbf{0}$ is the origin in $\mathbb{R}^d$. The probability $p_\infty(\mu)$ represents the probability that the origin $\mathbf{0}$ is contained in an infinite component of $\mathcal{G}\left(\mathcal{P}_\mu^{\mathbf{0}},1\right)$, and is the *percolation probability* (or *percolation function*). The threshold intensity for the phase transition giving the existence of an infinite component, known as the *critical intensity* $\mu_c$, is defined by

$$\mu_c := \inf\{\mu \in (0, \infty);\ p_\infty(\mu) > 0\}.$$

A fundamental result in continuum percolation theory states that $\mu_c$ is finite when $d \geq 2$; see Meester and Roy [28, Chapter 3]. Recall that a point process $\mathcal{Q}_1$ stochastically dominates $\mathcal{Q}_2$ on the same space if there exists a coupling of them, $(\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2)$, such that $\bar{\mathcal{Q}}_1 \stackrel{d}{=} \mathcal{Q}_1$, $\bar{\mathcal{Q}}_2 \stackrel{d}{=} \mathcal{Q}_2$, and $\bar{\mathcal{Q}}_2 \subset \bar{\mathcal{Q}}_1$ almost surely.

With this setup, we have the following theorem.

**Theorem 3.8.** *Let $d \geq 2$. We assume:*

- *the point process $\mathcal{P}$ satisfies the moment condition* (1.1) *for some positive constant $p_1$,*
- *$\mathcal{P}$ exhibits exponential decay of dependence as in (1.11), and stochastically dominates $\mathcal{P}_{\mu_0}$ for some $\mu_0 > \mu_c$,*
- *the function $f$ that determines the remote communities satisfies $f(s) > \sqrt{d}s$ for all $s \in \mathbb{N}$,*
- *the score function $\xi$ is a function of the graph ball in $\mathcal{G}^{iso}(\mathcal{P}, \lambda)$ with graph radius $r_0 \in \mathbb{N}$ and satisfies* (1.2) *for some positive constant $p_2$ such that $p_1p_2/(p_1 + p_2 - 1) \geq 3$,*
- $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some $\nu > 2/3$.*

*Then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left((\log \lambda)^{3d}\lambda^{-1.5\nu+1}\right),$$

*where $Z \sim N(0, 1)$.*

Since the conditions of Theorem 3.8 are insufficient to imply that the Palm process of $\mathcal{P}$ stochastically dominates a Poisson point process with a supercritical intensity, and since there is no existing result establishing that every Palm process of $\mathcal{P}$ possesses a giant component with probability converging to 1 as $\lambda$ grows, we cannot establish (1.15) directly. Because (1.15) is formulated in terms of the Palm distribution, Corollary 1.11 therefore cannot be applied directly in this case.

*Proof of Theorem 3.8.* The existence of remote communities depends on the entire configuration of $\mathcal{P} \cap \Gamma_\lambda$. It follows that $\xi$ is not stabilizing, hence the usual stabilization methods based on restricted score functions do not work. However, we can construct an estimate $\hat{\xi}^{[r]}$ of $\xi$ using the local information about the continuum percolation and then deduce Theorem 3.8 from Theorem 1.6.

The moment assumptions (1.1) and (1.2) are assumed. We thus only need to show exponential mixing of the score sums in (1.4). For this purpose, we follow the same steps as in the proof of Corollary 1.11 (a) to construct the corresponding point processes $\Psi_i$, for $1 \le i \le 6$, and to define the event $E_1$, as well as the random variables $S'_A$, $S'_B$, $\bar{S}_A^{[r]}$ and $\bar{S}_B^{[r]}$, in the same manner as in the proof of Corollary 1.11 with

$$\hat{\xi}^{[r]}(x, \mathcal{P}, \Gamma_\lambda) := \hat{\xi}^{[r_0]}(x, \mathcal{P}, \Gamma_\lambda) := h(B^{\mathcal{G}(\mathcal{P}\cap\Gamma_\lambda,1)}(x, r_0)),$$

for $A, B \in \mathcal{B}(\mathbb{R}^d)$ and $r = d(A, B)/3 \ge r_0$. We will, however, need to modify the definitions of the events $E_2$ and $E_3$ used in the proof of Corollary 1.11. To this end, we adopt the terminology from Penrose and Pisztora [31], where a connected component is said to be *crossing* if the union of balls of radius 1 centered at the vertices of the component intersects all $2d$ faces of the boundary of $\Gamma_\lambda$. We then define $E_2$ (resp. $E_3$) as the event that $C(\Psi_5 \cap \Gamma_\lambda, 1)$ (resp. $C(\Psi_6 \cap \Gamma_\lambda, 1)$) is *not* crossing. Since $\mathcal{P}$ stochastically dominates $\mathcal{P}_{\mu_0}$, it follows from [31, Proposition 2] with $\phi_\lambda = 0.5\lambda$ that there is a constant $C_6 := C_6(\mu_0)$ such that

$$\mathbb{P}(E_2) \vee \mathbb{P}(E_3) \le \exp(-C_6\lambda).$$

Whenever $C(\Psi_5 \cap \Gamma_\lambda, 1)$ and $C(\Psi_6 \cap \Gamma_\lambda, 1)$ are crossing, there must be components of size at least $\lambda^{1/d}$ and this, together with the condition $f(s) \ge \sqrt{d}s$ ensures that on $E_2^c \cap E_3^c$ there are no remote communities in $\Psi_5 \cap \Gamma_\lambda$ and $\Psi_6 \cap \Gamma_\lambda$, and consequently no edges are added between remote communities and isolated communities. Therefore, $\mathcal{G}^{iso}(\mathcal{P}, \lambda)$ and $\mathcal{G}(\mathcal{P} \cap \Gamma_\lambda, 1)$ coincide and thus the graph ball of radius $r$ in the first graph is a subset of the Euclidean ball of radius $r$ in the second graph. Under the event $E = (E_1 \cup E_2 \cup E_3)^c$, when $d(A, B) \ge 3r_0$, we have $S_A = S'_A = \bar{S}_A^{[r]}$, $S_B = S'_B = \bar{S}_B^{[r]}$. Hence, exponential mixing of the score sums in (1.4) follows from the same argument for (2.30) in Step (iv) of the proof of Corollary 1.11. The desired statement then follows directly by applying Theorem 1.6 (a). □

## 3.4 Interacting diffusions on spatial random graphs

In this section we obtain quantitative CLTs for statistics of interacting diffusion models on spatial random graphs driven by EDD input. The results are quantitative counterparts of the qualitative CLTs established in [5, Section 9] and complement the Berry–Esseen rates of normal approximation for such statistics under Poisson input in [43].

Before introducing the specific model, we describe the underlying *spatial random graphs*. For ease of reading, this description is adapted from [5], where a more complete treatment is provided and where, for example, the relevant measurability issues are addressed.

Given a locally finite point set $\mathcal{X} \subset \Gamma_\lambda$, let $\mathcal{G}(\mathcal{X}, \sim) =: \mathcal{G}(\mathcal{X})$ denote a graph with vertex set $\mathcal{X}$, where edges are determined by a deterministic rule $\sim$ depending on the entire configuration. That is, $\mathcal{X} \mapsto \mathcal{G}(\mathcal{X}, \sim)$ is a measurable mapping from the space $\mathcal{N}_{\Gamma_\lambda}$ of locally finite subsets of $\Gamma_\lambda$ to the space of finite labelled graphs on $\Gamma_\lambda$, equipped with the discrete $\sigma$-algebra.

A function $S_\lambda : \Gamma_\lambda \times \mathcal{N}_{\Gamma_\lambda} \to [0, \infty)$, $\lambda \in (0, \infty]$, with the convention $\Gamma_\infty = \mathbb{R}^d$, is called an *interaction range* with respect to the graph operator $\mathcal{G}(\cdot, \sim)$ if it satisfies:

(i) for all $\mathcal{X} \in \mathcal{N}_{\Gamma_\lambda}$ and $x \in \mathcal{X}$, the neighborhood of $x$ is determined by $\mathcal{X} \cap B(x, S_\lambda(x, \mathcal{X}))$, i.e., $y \sim x$ in $\mathcal{G}(\mathcal{X}, \sim)$ iff $y \sim x$ in $\mathcal{G}(\mathcal{X} \cap B(x, S_\lambda(x, \mathcal{X})), \sim)$,

(ii) the ball $B(x, S_\lambda(x, \mathcal{X}))$ is a stopping set, in the sense that

$$S_\lambda(x, \mathcal{X}) = S_\lambda\left(x, [\mathcal{X} \cap B(x, S_\lambda(x, \mathcal{X}))] \cup \mathcal{Y}\right)$$

for any $\mathcal{Y} \in \mathcal{N}_{\mathbb{R}^d \setminus B(x, S_\lambda(x, \mathcal{X}))}$.

By construction, $y \sim x$ in $\mathcal{G}(\mathcal{X}, \sim)$ only if $|x - y| \leq \min\{S_\lambda(x, \mathcal{X}), S_\lambda(y, \mathcal{X})\}$.

The following stabilization property for the graph $\mathcal{G}$ is inspired by a similar, though slightly stronger property in [5, Definition 8.1 and display (9.2)]; see Remark 3.12 for the differences.

**Definition 3.9.** *(stabilizing property of the graph operator $\mathcal{G}(\cdot, \sim)$ with respect to $\mathcal{P}$) We say that $\mathcal{G}$ stabilizes with respect to the point process $\mathcal{P}$ if there exists a family of interaction ranges $(S_\lambda)_{\lambda \in (0,\infty]}$ such that*

$$\sup_{\lambda \in (0,\infty]} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x\left(S_\lambda(x, \mathcal{P} \cap \Gamma_\lambda) > r\right) \leq \varphi(r), \qquad r \in (0, \infty),$$

*for some fast decreasing function $\varphi$, i.e., for $\varphi$ satisfying*

$$\limsup_{r \to \infty} r^m \varphi(r) = 0, \qquad \forall m \geq 1.$$

We further assume that the graph operator is translation invariant, that is,

$$x \sim y \text{ in } \mathcal{G}(\mathcal{X}, \sim) \text{ iff } x + z \sim y + z \text{ in } \mathcal{G}(\mathcal{X} + z, \sim), \quad \forall x, y \in \mathcal{X},\ z \in \mathbb{R}^d.$$

**Definition 3.10.** *(admissible graphs) The graph operator $\mathcal{G}$ is admissible on $\mathcal{P}$ if it is translation invariant and stabilizes with respect to $\mathcal{P}$ as in Definition 3.9.*

As shown in [5, Appendix A, Examples A.1–A.4], if $\mathcal{P}$ is a homogeneous Poisson point process or a stationary determinantal point process with an exponentially decaying kernel, then admissible graphs on $\mathcal{P}$ include the undirected $k$-nearest neighbors graph, the sphere of influence graph, and the Delaunay graph. Throughout this section, we work with such point processes $\mathcal{P}$ that are also EDD, together with the corresponding graph operators $\mathcal{G}(\cdot, \sim)$.

It is shown in [5, Section 9] that statistics of interacting diffusions on spatial random graphs satisfy asymptotic normality, provided the underlying point process has fast decay of correlations. The key to their results is that the relevant statistics could be expressed as a sum of score functions which satisfy exponential stabilization in the $L^2$ distance. That work made no attempt to find rates of normal convergence. Here we indicate that our general results yield rates of normal convergence when the input is a class of EDD point processes and when the initial diffusion displacements are independent and uniformly bounded. We recall the set-up and refer to [5, Section 9] for full details.

Given a countable $V \subset \mathbb{R}^d$, let $\mathcal{G} := \mathcal{G}(V)$ be an admissible interaction graph on $V$. Let $N_v$ be the neighborhood of $v \in V$. Fix a time horizon $t_0 \in (1, \infty)$. Consider a system of interacting $\mathbb{R}^{d'}$-valued diffusions $M(v,t) := M^G(v,t), v \in V, t \in [0, t_0]$, defined by

$$dM(v,t) = b(t, M(v,t), M(N_v,t))dt + \sigma(t, M(v,t), M(N_v,t))dZ_v(t), \tag{3.34}$$

where $Z_v(\cdot), v \in V$, are i.i.d. standard Brownian motions in $\mathbb{R}^{d'}, d' \in \mathbb{N}$, and where $M(N_v,t) = \{M(v',t)\}_{v' \sim v}$, where $v' \sim v$ means that $v'$ is a neighbor of $v$. Denote by $M(v) := M(v,0) \in \mathbb{R}^{d'}, v \in V$, the initial states, assumed to be almost surely uniformly bounded, i.e., $\sup_{v \in V} |M(v,0)| \leq L$ a.s. for a constant $L \in (0,\infty)$. These processes take values in the space $\mathbb{C}^{(0)}(\mathbb{R}^{d'})$ of continuous paths in $\mathbb{R}^{d'}$, with the functions $b$ and $\sigma$ being the *drift* and *diffusion* coefficients, respectively. Here $b$ is $\mathbb{R}^{d'}$-valued whereas the diffusion matrix $\sigma$ is $\mathbb{R}^{d'} \times \mathbb{R}^{d'}$-valued; see [5, Definition 9.2] for full details. The particles interact directly only with their (finite) graph neighbors, in contrast to the mean field model, where particles interact with all other particles.

We further assume Lipschitz conditions on $b$ and $\sigma$, as spelled out in Lacker, Ramanan and Wu [25] and [5, Definition 9.2]. When $\mathcal{G}$ is admissible, the Lipschitz assumption implies the existence of a strong solution for the system of interacting diffusions; c.f. [25, Theorem 3.1].

When the processes are defined on the entirety of the graph $\mathcal{G}$, we denote them simply as $M := M^{\mathcal{G}}$. The paths (trajectories) of the processes up to time $t$ are denoted by $M[v,t]$. Let $\hat{\mathcal{P}} := \{(x, M(x), Z_x)\}_{x \in \mathcal{P}}$ be a marked point process on $\mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{C}^{(0)}(\mathbb{R}^{d'})$, with $M(x), x \in \mathcal{P}$, denoting the independent initial states.

To cast the statistics of interacting diffusions in the framework of our general results, we consider, *for fixed* $t \in (0,\infty)$, real-valued score functions of the trajectories

$$\xi((x, M(x), Z_x), \hat{\mathcal{P}}) := h(M^{\mathcal{G}(\mathcal{P})}[x,t]) \tag{3.35}$$

where $h : \mathbb{C}^{(0)}(\mathbb{R}^{d'})(t) \to \mathbb{R}$ is a Lipschitz(1) function with respect to the sup-norm, $\|\cdot\|_\infty$, on $\mathbb{C}^{(0)}(\mathbb{R}^{d'})(t)$, the space of trajectories up to time $t$. The measurability of $\xi$ with respect to $\hat{\mathcal{P}}|_{\Gamma_\lambda}$ is established in the proof of Theorem 7.8 of [25] and also in Section 9 of [5]. Such scores give rise to the diffusion statistics

$$H_\lambda := H_\lambda^\xi := \sum_{x \in \mathcal{P} \cap \Gamma_\lambda} \xi((x, M(x), Z_x), \hat{\mathcal{P}}).$$

Finally, for $k \in \mathbb{N}$ denote by $\mathcal{G}_x^k(\mathcal{X})$ the subgraph of $\mathcal{G}(\mathcal{X})$ induced by the graph ball $B^{\mathcal{G}(\mathcal{X})}(x,k)$ centered at $x \in \mathcal{X}$ with graph radius $k$. This gives rise to the corresponding interacting diffusion processes $\left\{M^{\mathcal{G}_x^k(\mathcal{X})}(y,t)\right\}$ for $t \in [0, t_0]$ on $y \in B^{\mathcal{G}(\mathcal{X})}(x,k)$, which have the same coefficients $(b, \sigma)$, are driven by the same Brownian motions, and share the same initial conditions $M^{\mathcal{G}_x^k(\mathcal{X})}(y,0) = M(y,0)$ for $y \in B^{\mathcal{G}(\mathcal{X})}(x,k)$. Let $\mathcal{P}_x$ denote the Palm process of $\mathcal{P}$ at $x$, that is, $\mathcal{P}_x \sim \mathcal{L}_x(\mathcal{P})$.

**Corollary 3.11.** *(rates of normal convergence for statistics of interacting diffusions on EDD input) Let $\hat{\mathcal{P}} = \{(x, M(x), Z_x)\}$ be as above, with $M(x), x \in \mathcal{P}$, independent initial states which are a.s. uniformly bounded, and $\mathcal{G} = \mathcal{G}(\cdot, \sim)$ an admissible graph on $\mathcal{P}$. Let $M = M^{\mathcal{G}(\mathcal{P})}$ be the system of interacting diffusions as in* (3.34) *on $\mathcal{G}$ with diffusion coefficients $b, \sigma$ satisfying the Lipschitz assumption of [25]. Fix $t_0 \in (0, \infty)$ and assume the score function $\xi$ in* (3.35) *and the corresponding function $h$ satisfy the following moment conditions:*

$$\begin{gathered} \sup_{x \in \mathbb{R}^d} \mathbb{E}[\max(1, |h(M^{\mathcal{G}(\mathcal{P}_x)}[x, t_0])|^p)] < \infty, \\ \sup_{x \in \mathbb{R}^d, k \in \mathbb{N}} \mathbb{E}\left[\max\left(1, \left|h\left(M^{\mathcal{G}_x^k(\mathcal{P}_x)}[x, t_0]\right)\right|\right)^{p'}\right] < \infty, \end{gathered} \tag{3.36}$$

*for some positive constants $p$, $p'$ such that $\min\{p, p'\} > 4$. If $\operatorname{Var} H_\lambda = \Omega(\lambda^\nu)$ for some $\nu > 2/3$, then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-\kappa}\right),$$

*for all $\kappa < 1.5\nu - 1$.*

*Proof.* The proof is adapted from that of [5, Theorem 9.3] to our setting, which imposes fewer Palm requirements on the graph structure $\mathcal{G}$. We take the short-range estimator of the score function $\xi$ in (1.13) to be

$$\hat{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}}) = h\left(M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P})}[x, t_0]\right) \mathbf{1}_{E_{x,r}},$$

where $E_{x,r} := \{S_\infty(y, \mathcal{P}) \le \lfloor\sqrt{r}\rfloor \text{ for all } y \in \mathcal{P} \cap B\left(x, r - \lfloor\sqrt{r}\rfloor\right)\}$. By construction, $\hat{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}})$ is determined by $\hat{\mathcal{P}}|_{B(x,r)}$. Moreover, by Remark 1.7 (iii) and the moment assumptions, the power $p_1$ in (1.1) can be chosen arbitrarily large, while $p_2$ and $p_2'$ can be chosen as $p$ and $p'$ in (3.36), respectively. Thus, by taking sufficiently large $p_1$, in Corollary 1.14 we take $\underline{p} := (p \wedge p')/2 + 2 > 4$, and set $p_3 := \underline{p}/(\underline{p} - 2) < 2$. It suffices to show that $\xi$ satisfies polynomial $L^{p_3}$-stabilization in Definition 1.13 with arbitrarily large $\beta$, as discussed in Remark 1.15 (iii). Set

$$\tilde{\xi}^{[r]}\left((x, M(x), Z_x), \hat{\mathcal{P}}\right) = h\left(M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P})}[x, t_0]\right).$$

By the moment assumption in (3.36), we have $\|\tilde{\xi}^{[r]}\left((x, M(x), Z_x), \hat{\mathcal{P}}_x\right)\|_{p'} < \infty$ uniformly in $x$ and $r$, where $\hat{\mathcal{P}}_x$ is the Palm process of $\hat{\mathcal{P}}$ at $x$, i.e., $\hat{\mathcal{P}}_x \sim \mathcal{L}_x(\hat{\mathcal{P}})$. By the Minkowski inequality, Hölder's inequality, the Lipschitz(1) assumption of $h$ and Lemmas 9.6 and 9.7 in [5], whose bounds are stated for second moments and hence are used below after taking square roots, we have

$$\begin{aligned}
&\left\|\xi((x, M(x), Z_x), \hat{\mathcal{P}}_x) - \hat{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}}_x)\right\|_{p_3} \\
&\leq \left\|\xi((x, M(x), Z_x), \hat{\mathcal{P}}_x) - \tilde{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}}_x)\right\|_{p_3} \\
&\quad + \left\|\hat{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}}_x) - \tilde{\xi}^{[r]}((x, M(x), Z_x), \hat{\mathcal{P}}_x)\right\|_{p_3} \\
&\leq \left\|\left\|M^{\mathcal{G}(\mathcal{P}_x)}[x, t_0] - M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right\|_\infty\right\|_{p_3} + \left\|\mathbf{1}_{E_{x,r}^c}\left(|h(\mathbf{0}(\cdot))| + \left\|M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right\|_\infty\right)\right\|_{p_3} \\
&\leq \left(\frac{C^{\lfloor\sqrt{r}\rfloor}}{\lfloor\sqrt{r}\rfloor!}\right)^{1/2} + |h(\mathbf{0}(\cdot))|\mathbb{P}\left(E_{x,r}^c\right)^{1/p_3} + \mathbb{P}\left(E_{x,r}^c\right)^{(2-p_3)/2p_3}\left((C')^{1/2} + \left(\frac{C^{\lfloor\sqrt{r}\rfloor}}{\lfloor\sqrt{r}\rfloor!}\right)^{1/2}\right),
\end{aligned} \tag{3.37}$$

where $\mathbf{0}(\cdot)$ denotes the zero process on the time interval $[0, t]$, and $C$ and $C'$ are the positive constants appearing in Lemmas 9.7 and 9.6, respectively. Here, the second inequality follows from the Lipschitz property of $h$, which gives

$$\left|h\left(M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right)\right| \leq |h(\mathbf{0}(\cdot))| + \left\|M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right\|_\infty,$$

and the third inequality follows from Minkowski's inequality, Hölder's inequality, and the fact that

$$\left\|\left\|M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right\|_\infty\right\|_2 \leq \left\|\left\|M^{\mathcal{G}(\mathcal{P}_x)}[x, t_0] - M^{\mathcal{G}_x^{\lfloor\sqrt{r}\rfloor}(\mathcal{P}_x)}[x, t_0]\right\|_\infty\right\|_2 + \left\|\left\|M^{\mathcal{G}(\mathcal{P}_x)}[x, t_0]\right\|_\infty\right\|_2.$$

The first term on the right-hand side of (3.37) decays faster than $r^{-\beta}$ for every $\beta \in (0, \infty)$ as $r$ goes to infinity. Furthermore,

$$\begin{aligned}
\mathbb{P}_x(E_{x,r}^c) &\leq \int_{B\left(x, r - \lfloor\sqrt{r}\rfloor\right)} \mathbb{P}_y\left(S_\infty(y, \mathcal{P}) > \left\lfloor\sqrt{r}\right\rfloor\right)\mu \mathrm{d}y \\
&\leq \mu \operatorname{Vol}\left(B\left(x, r - \left\lfloor\sqrt{r}\right\rfloor\right)\right) C'_\beta r^{-\beta'} \\
&\leq C''_\beta r^{d-\beta'}
\end{aligned} \tag{3.38}$$

for some positive constants $C'_\beta$ and $C''_\beta$ for all $\beta' \in (0, \infty)$, where $\mu$ is the intensity of $\mathcal{P}$ and the last inequality follows from the stabilizing assumption in Definition 3.9. Consequently, the second and third terms on the right-hand side of (3.37) also decay faster than $r^{-\beta}$ for all $\beta \in (0, \infty)$ by taking $\beta' = 2\beta p_3/(2 - p_3) + d$ in (3.38). Hence the estimates above verify polynomial $L^{p_3}$-stabilization in Definition 1.13 with arbitrarily large polynomial decay exponent. Together with the moment condition for the short-range estimators verified above, Corollary 1.14 and Remark 1.15 (iii) yield the desired statement. □

**Remark 3.12.** *(cf. Remark 1.12 (iv)) In [5], the definition of a spatial random graph is based on the supremum of conditional probabilities given that $\mathcal{P}$ contains $x$ together with any other finite set $\{x_1, \ldots, x_p\} \subset \Gamma_\lambda \backslash \{x\}$, for $p \in \mathbb{N}$, which is a potentially restrictive condition as it is based on higher-order Palm distributions. Our results are instead based on the first-order Palm distribution, leading to a slightly broader class of graphs for the same underlying point process $\mathcal{P}$. For example, for any configuration $\mathcal{X} \in \mathcal{N}_{\mathbb{R}^d}$, we define the graph $\mathcal{G}(\mathcal{X}, \sim)$ as follows: we connect each $x \in \mathcal{X}$ to its nearest neighbor in $\mathcal{X}$,*

$$N_x := N_x(\mathcal{X}) := \underset{y \in \mathcal{X} \backslash \{x\}}{\arg\min}\, d(x, y),$$

*provided that $\mathrm{d}(x, N_x) < 1/\mathrm{e}$ and $\mathrm{card}(\mathcal{X} \cap B(x, -\ln d(x, N_x)))$ is even. Then,*

$$S_\lambda(x, \mathcal{X}) := \left[ \sup_{y \in \{z \in \mathcal{X} \cap B(x, 1/\mathrm{e}); N_z = x\} \cup \{x\}} (-\ln d(y, N_y)) \right] \wedge d(x, \partial\Gamma_\lambda) + 1/\mathrm{e}$$

*defines an interaction range with respect to the graph $\mathcal{G}(\cdot, \sim)$. Let $\mathcal{P}$ be a homogeneous Poisson point process in $\mathbb{R}^d$ with unit intensity. One readily verifies that*

$$\begin{aligned}
&\sup_{\lambda \in (0,\infty]} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x(S_\lambda(x, \mathcal{P} \cap \Gamma_\lambda) > r) \\
&\le \sup_{\lambda \in (0,\infty]} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x \left( \sup_{y \in \{z \in \mathcal{P} \cap \Gamma_\lambda \cap B(x, 1/\mathrm{e}); N_z = x\} \cup \{x\}} (-\ln d(y, N_y(\mathcal{P} \cap \Gamma_\lambda))) + \frac{1}{e} > r \right) \\
&\le \sup_{\lambda \in (0,\infty]} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x \left( d(x, N_x(\mathcal{P} \cap \Gamma_\lambda)) < \mathrm{e}^{\mathrm{e}^{-1} - r} \right) \\
&\le \sup_{\lambda \in (0,\infty]} \sup_{x \in \Gamma_\lambda} \mathbb{P}_x \left( \mathrm{card} \left( \mathcal{P} \cap (B(x, \mathrm{e}^{\mathrm{e}^{-1} - r})) \backslash \{x\} \right) \ge 1 \right) \\
&= \mathbb{P} \left( \mathrm{card} \left( \mathcal{P} \cap B \left( \mathbf{0}, \mathrm{e}^{\mathrm{e}^{-1} - r} \right) \right) \ge 1 \right)
\end{aligned}$$

*for all $r \ge 1$, where the last equality follows from the translation invariance of the homogeneous Poisson point process and the fact that its reduced Palm distribution coincides with its original distribution. The right-hand side decays exponentially fast. However,*

$$\sup_{\lambda \in (0,\infty]} \sup_{x_1, x_2 \in \Gamma_\lambda, x_1 \neq x_2} \mathbb{P}_{x_1, x_2}(S_\lambda(x_1, \mathcal{P} \cap \Gamma_\lambda) > r) = 1$$

*for all $r \ge 1$, since when $x_2$ is sufficiently close to $x_1$ and $\mathrm{d}(x_1, \partial\Gamma_\lambda) > r$, the dependence range of whether $x_1$ is connected to its nearest neighbor is always greater than $r$, which* *shows that the decay condition for the tails of $S_\lambda$ as given by [5, Definition 8.1] does not hold when $p = 2$. Graphs such as these, admittedly not common, illustrate that when we only require assumptions based on one-point Palm distributions then the class of admissible graphs may be broader than the class of graphs satisfying assumptions based on all Palm distributions.*

## 3.5 Local $U$-statistics

The limit theory for local $U$-statistics has attracted considerable attention over the past decades, see, e.g., Eichelsbacher and Thäle [15], Reitzner and Schulte [32] and [4, 5] for statistics with Poisson and non-Poisson input. Given $\mathcal{X} \in \mathcal{N}_{\mathbb{R}^d}$, such statistics have the form $[\sum_{\mathbf{x}\in\mathcal{X}^{k,\neq}} h_0(\mathbf{x})]/k!$, where $h_0$ is a real symmetric measurable function on $(\mathbb{R}^d)^{k,\neq}$.

Limit theorems for local $U$-statistics may be established under a locally dependent setup as in [4, 5], meaning that $h_0(x_1,\dots,x_k)$ vanishes whenever there exists a pair $x_i, x_j$ with $i, j \leq k$ such that $|x_i - x_j| \geq r$. Related qualitative CLTs for certain $U$-statistics of Gibbs particle processes whose dominating Boolean model is subcritical were obtained in [2]. Alternatively, such limit theorems may be derived under the assumption that $\mathcal{P}$ is a Poisson point process on a general metric space, see, e.g., [15, 32].

In this section, we deduce from Corollary 1.14 (a) and (b) rates for the normal approximation error of local $U$-statistics of EDD marked point processes $\hat{\mathcal{P}}$. No independence assumptions are imposed beyond the EDD condition in Definition 1.8. Instead, we let the kernel $h$ be a real symmetric measurable function on the space of the marked $k$-tuples $\{((y_1,m_1),...,(y_k,m_k))\,;\ (y_1,\dots,y_k) \in (\mathbb{R}^d)^{k,\neq},\ m_1,\dots,m_k \in \mathbb{M}\}$. We assume that the conditional $L^{p_0}$ norm of $h((x_1,M_{x_1}),\dots,(x_k,M_{x_k}))$ decays either exponentially or polynomially in the diameter of the argument set, for a given positive $p_0$, where $M_x$ denotes the mark of the point $x \in \mathcal{P}$. We denote by $H_\lambda$ the $U$-statistic generated by $\hat{\mathcal{P}}|_{\Gamma_\lambda}$, namely

$$H_\lambda = \frac{1}{k!}\sum_{\hat{\mathbf{x}}\in(\hat{\mathcal{P}}|_{\Gamma_\lambda})^{k,\neq}} h(\hat{\mathbf{x}}) = \sum_{x\in\mathcal{P}\cap\Gamma_\lambda} \xi((x,M_x),\hat{\mathcal{P}},\Gamma_\lambda), \tag{3.39}$$

where

$$\xi((x,m_x),\hat{\mathcal{X}},\Gamma_\lambda) := \frac{1}{k!}\sum_{\hat{\mathbf{x}}\in(\hat{\mathcal{X}}|_{\Gamma_\lambda\setminus\{x\}})^{k-1,\neq}} h((x,m_x),\hat{\mathbf{x}}). \tag{3.40}$$

**Theorem 3.13.** *(quantitative CLT for local $U$-statistics) Let $\hat{\mathcal{P}}$ be an EDD marked point process such that the ground point process $\mathcal{P}$ is either (i) a stationary Gibbs point process whose interaction range is smaller than the critical value of continuum percolation of an associated Poisson point process, or (ii) an $\alpha$-determinantal point process with kernel $K(x,y) := K_0(x-y) = \overline{K_0(y-x)}$ satisfying condition* (1.12)*, and $\alpha$ such that $-1/\alpha \in \mathbb{N}$. Let $H_\lambda$ be as in* (3.39) *and assume* $\mathrm{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some constant $\nu > 2/3$.*

(a) *If the conditional $L^{p_0}$ norm of $h$ decays exponentially fast for some $p_0 > 3$, i.e., if*

$$\mathbb{E}_{x_1,\ldots,x_k}\left(|h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))|^{p_0}\right)^{1/p_0} \leq C_1 \mathrm{e}^{-C_2 \operatorname{diam}(\{x_1,\ldots,x_k\})}$$

*for some positive constants $C_1$ and $C_2$ for all distinct $x_1, \ldots, x_k \in \mathbb{R}^d$, we have*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left((\log \lambda)^{3d} \lambda^{-1.5\nu+1}\right),$$

*where $Z \sim N(0,1)$.*

(b) *If the conditional $L^{p_0}$ norm of $h$ decays polynomially fast for some $p_0 > 3$, i.e., if*

$$\mathbb{E}_{x_1,\ldots,x_k}\left(|h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))|^{p_0}\right)^{1/p_0} \leq C_3 \left(1 \vee \operatorname{diam}(\{x_1, \ldots, x_k\})\right)^{-C_4}$$

*for some positive constants $C_3$ and $C_4$ for all distinct $x_1, \ldots, x_k \in \mathbb{R}^d$, with*

$$C_4 > \frac{d}{3\nu - 2}((18 - 3\nu) \vee 14) + d(k-1) + 1,$$

*we have*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(2d,\beta)}\right) = o(1),$$

*for $\beta = C_4 - d(k-1) - 1$ where $Z \sim N(0,1)$ and where $\tau$ is defined in* (1.9).

**Remark 3.14.** *(i) Theorem 3.13 can be easily extended to the case where the score is a function of $\hat{\mathcal{P}}$ instead of $\hat{\mathcal{P}}|_{\Gamma_\lambda}$, i.e., for the case in which $H_\lambda = \sum_{x \in \mathcal{P} \cap \Gamma_\lambda} \xi((x, m_x), \hat{\mathcal{P}})$, with minor adjustments.*

*(ii) With minor modifications, all results also apply to EDD marked point processes $\hat{\mathcal{P}}$ whose ground process satisfies the bound* (3.41) *below for sufficiently large $k_0$. The same conclusions remain valid if the above conditional $L^{p_0}$ bounds are assumed to hold uniformly in the essential supremum sense. For sake of brevity we omit the details.*

Before proceeding with the proof of the theorem, we state moment bounds for the considered point processes.

**Lemma 3.15.** *For each point process $\mathcal{P}$ in Theorem 3.13 and each $k_0 \in (0, \infty)$, there exists a constant $D_{k_0} \in (0, \infty)$, depending only on $\mathcal{L}(\mathcal{P})$ and $k_0$, such that for all punctured balls $B^\circ(x, r) := B(x, r) \backslash \{x\}$,*

$$\sup_{x \in \mathbb{R}^d} \mathbb{E}_x(\mathcal{P}(B^\circ(x, r))^{k_0})^{1/k_0} \leq D_{k_0}(1 \vee r)^d, \quad r \in (0, \infty). \tag{3.41}$$

*Proof.* When $\mathcal{P}$ is a Gibbs point process whose interaction range is smaller than the critical value of continuum percolation of an associated Poisson point process, it follows from [20, Section 4.1] that it is stochastically dominated by a homogeneous Poisson process $\mathcal{P}_\tau$ with a positive intensity $\tau$. Let $\mu$ be the intensity of $\mathcal{P}$. Fix $0 < \delta_1 < \delta_2 < r$, and define the spherical shell $B_{x,\delta_2} = B(x,r) \setminus B(x,\delta_2)$. Then

$$\frac{\mathbb{E}\left(\mathcal{P}(B_{x,\delta_2})^{k_0}\mathcal{P}(B(x,\delta_1))\right)}{\mathbb{E}\left(\mathcal{P}(B(x,\delta_1))\right)} \le \frac{\mathbb{E}\left(\mathcal{P}_\tau(B_{x,\delta_2})^{k_0}\mathcal{P}_\tau(B(x,\delta_1))\right)}{\mu\,\mathrm{Vol}(B(x,\delta_1))} = \frac{\tau}{\mu}\mathbb{E}\left(\mathcal{P}_\tau(B_{x,\delta_2})^{k_0}\right). \quad (3.42)$$

First taking the limit as $\delta_1 \to 0$, and then as $\delta_2 \to 0$ in (3.42), the definition of the Palm distribution in Kallenberg [23, (10.3)] gives

$$\mathbb{E}_x\left(\mathcal{P}(B^\circ(x,r))^{k_0}\right) \le \frac{\tau}{\mu}\mathbb{E}\left(\mathcal{P}_\tau(B^\circ(x,r))^{k_0}\right).$$

This implies inequality (3.41).

When $\mathcal{P}$ is a determinantal point process, condition (1.12) together with the Hermitian property ensures that its kernel satisfies Assumptions 1-3, while Assumption 4 is equivalent to the existence of the process given the first three assumptions (cf. the proof of Theorem 3.1). According to [30, Theorem 1], for all $x \in \mathbb{R}^d$, there exists a coupling of $\mathcal{P}$ and its reduced Palm version $\mathcal{P}_x^!$ such that $\mathcal{P}_x^! \subseteq \mathcal{P}$ almost surely. This property, together with Remark 1.2 and Lemma 2.1, guarantees the desired conclusion. □

*Proof of Theorem 3.13.* We will deduce this from Corollary 1.14. The point process $\mathcal{P}$ satisfies the EDD condition as shown in [10, Section 3]. By Remarks 1.2 and 1.15 (i), it suffices to show that assumptions (1.2), (1.17) hold and that exponential $L^{p_3}$-stabilization for part (a) and polynomial $L^{p_3}$-stabilization for part (b) hold for a family of short-range score functions $(\hat{\xi}^{[r]})_{r\in(0,\infty)}$, for $p_2 = p_2' = p_0 > 3$ and $p_3 = 3$. Define $\hat{\xi}^{[r]}$ to be the restriction of $\xi$ to $\hat{\mathcal{P}}|_{B(x,r)}$ as in (1.14), that is,

$$\hat{\xi}^{[r]}((x,m_x),\hat{\mathcal{X}},\Gamma_\lambda) = \frac{1}{k!}\sum_{\hat{\mathbf{x}}\in(\hat{\mathcal{X}}|_{B(x,r)^\circ\cap\Gamma_\lambda})^{k-1,\neq}} h((x,m_x),\hat{\mathbf{x}}).$$

Define for $\mathbf{y} = (y_1,\dots,y_n) \in \left(\mathbb{R}^d\right)^{n,\neq}$ and $x \in \mathbb{R}^d$

$$d_s(x,\mathbf{y}) := d_s(x,\{y_1,\dots,y_n\}) := \sup\{d(x,y_j) :\ 1 \le j \le n\}.$$

Since $d_s(x,\{y_1,\dots,y_n\}) \le \mathrm{diam}(\{x,y_1,\dots,y_n\})$, the fast-decay assumptions on $h$ in parts (a) and (b) imply, respectively, that

$$\mathbb{E}_{x,y_1\dots,y_{k-1}}\left(|h((x,M_x),(y_1,M_{y_1})\dots,(y_{k-1},M_{y_{k-1}}))|^{p_0}\right)^{1/p_0} \le C_1\mathrm{e}^{-C_2 d_s(x,\{y_1,\dots,y_{k-1}\})} \quad \text{and,}$$

$$\mathbb{E}_{x,y_1\dots,y_{k-1}}\left(|h((x,M_x),(y_1,M_{y_1})\dots,(y_{k-1},M_{y_{k-1}}))|^{p_0}\right)^{1/p_0} \le C_3\left(1 \vee d_s(x,\{y_1,\dots,y_{k-1}\})\right)^{-C_4}.$$

We then set

$$\mathrm{UB}(\varrho) = C_1 \mathrm{e}^{-C_2 \varrho} \quad \text{for (a)}, \qquad \mathrm{UB}(\varrho) = C_3(1 \vee \varrho)^{-C_4} \quad \text{for (b)}.$$

For any marked configuration $\hat{\mathcal{X}}$ with ground configuration $\mathcal{X}$, and for $(x, m_x) \in \hat{\mathcal{X}}$, we decompose the sum in (3.40) according to the spherical shell $B(x,j) \setminus B(x, j-1)$ which contains among the ground coordinates of $\hat{\mathbf{y}} \in (\hat{\mathcal{X}})^{k-1,\neq}$ the point furthest from $x$. Precisely, for $j \in \mathbb{N}$, we define

$$\begin{aligned} S_j\left(x, \hat{\mathcal{X}}\right) &:= \sum_{\hat{\mathbf{y}} \in (\hat{\mathcal{X}})^{k-1,\neq},\ d_s(x,\mathbf{y}) \in [j-1,j)} h((x, m_x), \hat{\mathbf{y}}) \\ &= \sum_{\mathbf{y} \in (\mathcal{X}|_{B^\circ(x,j)})^{k-1,\neq} \setminus (\mathcal{X}|_{B^\circ(x,j-1)})^{k-1,\neq}} h((x, m_x), \hat{\mathbf{y}}), \end{aligned}$$

where $\hat{\mathbf{y}} = ((y_1, m_{y_1}), \ldots, (y_n, m_{y_n}))$ for $\mathbf{y} = (y_1, \ldots, y_n) \in \mathcal{X}^{n,\neq}$. The number of admissible $(k-1)$-tuples contributing to the $j$-th shell term is bounded by the $(k-1)$-st power of the number of ground points in the corresponding shell. Thus, for any $j \in \mathbb{N}$ we obtain via Lemma 3.15

$$\begin{aligned} \mathbb{E}_x(|S_j\left(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}\right)|^3)^{1/3} &\le \mathbb{E}_x(|S_j(x, \hat{\mathcal{P}}|_{\Gamma_\lambda})|^{p_0})^{1/p_0} \\ &= \mathbb{E}_x\left(\mathbb{E}\left(\left|S_j\left(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}\right)\right|^{p_0} \middle| \mathcal{P} \cap (B(x,j) \setminus B(x, j-1))\right)\right)^{1/p_0} \\ &\le \mathrm{UB}(j-1)^{p_0} \mathbb{E}_x\left(\mathcal{P}(B(x,j) \setminus B(x,j-1))^{(k-1)p_0}\right)^{1/p_0} \\ &\le D_{(k-1)p_0}^{k-1} \mathrm{UB}(j-1) j^{d(k-1)}, \end{aligned} \tag{3.43}$$

where the penultimate inequality follows from Minkowski's inequality.

Rewrite $\xi$ and $\hat{\xi}^{[r]}$ via the additive representations

$$\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) = \frac{1}{k!} \sum_{j \in \mathbb{N}} S_j(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}) \text{ and } \hat{\xi}^{[r]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) = \frac{1}{k!} \sum_{j \in \mathbb{N}} S_j(x, \hat{\mathcal{P}}|_{B(x,r) \cap \Gamma_\lambda}).$$

We first show that the moment assumptions (1.2) and (1.17) hold for $p_2 = p_2' = p_0 > 3$. The Minkowski inequality and (3.43) imply that in cases (a) and (b), we have

$$\begin{aligned} \left(\mathbb{E}_x \left|\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)\right|^{p_0}\right)^{1/p_0} &\le \frac{1}{k!} \sum_{j=1}^{\infty} \left(\mathbb{E}_x \left|S_j\left(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}\right)\right|^{p_0}\right)^{1/p_0} \\ &\le \frac{D_{(k-1)p_0}^{k-1}}{k!} \sum_{j=1}^{\infty} \mathrm{UB}(j-1)^{d(k-1)} < \infty, \end{aligned}$$

where, by assumption, $C_4 > d(k-1)+1$. Moreover, the assumption (1.17) also holds for $p_2' = p_0$ by replacing $\xi$ and $S_j\left(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}\right)$ with $\hat{\xi}^{[r]}$ and $S_j\left(x, \hat{\mathcal{P}}|_{B(x,r)\cap\Gamma_\lambda}\right)$, respectively.

Next, we show that exponential $L^{p_3}$-stabilization, as in case (a), and polynomial $L^{p_3}$-stabilization, as in case (b), hold for $p_3 = 3$. For any given $r_0 > 1$, we have

$$\begin{aligned}
&\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) - \hat{\xi}^{[r_0]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) \\
&= \frac{1}{k!}\left[\sum_{j=\lceil r_0\rceil+1}^{\infty} S_j\left(x, \hat{\mathcal{P}}|_{\Gamma_\lambda}\right) + \sum_{\hat{\mathbf{y}}\in\left(\hat{\mathcal{P}}|_{B^\circ(x,\lceil r_0\rceil)}\right)^{k-1,\neq}\setminus\left(\hat{\mathcal{P}}|_{B^\circ(x,r_0)}\right)^{k-1,\neq}} h\left((x, M_x), \hat{\mathbf{y}}\right)\right]. \quad (3.44)
\end{aligned}$$

As the first term on the right-hand side in (3.44) is the leading term, we can assume, without loss of generality, that $r_0$ is an integer so that the second sum vanishes. Then, applying the Minkowski inequality and using (3.43) again, we get

$$\begin{aligned}
&\left(\mathbb{E}_x\left|\xi((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda) - \hat{\xi}^{[r_0]}((x, M_x), \hat{\mathcal{P}}, \Gamma_\lambda)\right|^3\right)^{1/3} \\
&\le \frac{1}{k!}\sum_{j=r_0+1}^{\infty}\left(\mathbb{E}_x\left|S_j(x, \hat{\mathcal{P}}|_{\Gamma_\lambda})\right|^3\right)^{1/3} \\
&\le \frac{D_{(k-1)p_0}^{k-1}}{k!}\sum_{j=r_0+1}^{\infty}\mathrm{UB}(j-1)j^{d(k-1)}
\end{aligned}$$

implying cases (a) and (b) provided $C_4 > d(k-1)+1$. Indeed, the right-hand side is bounded above by $C_5\mathrm{e}^{-C_6 r_0}$ for some positive constants $C_5$ and $C_6$ in case (a), and by $C_7 r_0^{d(k-1)+1-C_4}$ for some positive constant $C_7$ in case (b). Hence, statements (a) and (b) follow from Corollary 1.14 (a) and (b), respectively. □

### 3.5.1 Subgraph counting in the random connection model

We study the problem of counting subgraphs within a class of random connection graphs inspired by the *age-dependent random connection model* (cf. Hirsch, Lachièze-Rey and Owada [19] and references therein), where vertices are equipped with *i.i.d.* marks and the existence of an edge between two vertices depends on both their Euclidean distance $d$ and their respective marks.

Assume that $\mathcal{P}$ is one of the point processes introduced in Theorem 3.13, and attach to each $x \in \mathcal{P}$ an independent mark $M_x \sim U[0,1]$, with the marks being independent of $\mathcal{P}$. Given a function $\psi : (0,\infty) \to [0,1]$ and a symmetric function $g : [0,1]\times[0,1] \to [0,1]$, we construct a random graph $\mathcal{G}(\mathcal{P}) := \mathcal{G}(\mathcal{P}, g, \psi)$ with vertex set $\mathcal{P}$ by connecting points $x, y \in \mathcal{P}$ whenever $g(M_x, M_y) \le \psi(d(x,y))$.

For example, taking $\psi$ to be a constant $p \in [0,1]$ leads to the connection rule that we connect vertices $x, y$ whenever their marks satisfy $g(M_x, M_y) \le p$, that is, all pairs of points are connected with the same probability; the probabilities are dependent when a common vertex is involved and thus this model does not precisely coincide with the age-dependent random connection model. On the other hand, if $\psi(t) = \mathbf{1}_{(0\le t\le r)}$ and $g \not\equiv 0$, then we obtain a version of the random geometric graph with parameter $r$. Fixing $p \geq 0$, taking $g(s,t) = \max\{s,t\}$ and $\psi(r) = Cr^{-p}, r \in (0,\infty)$, gives rise to the graph defined by the connection rule that we connect $x, y$ whenever

$$\max\{U_x, U_y\} \leq C(d(x,y))^{-p},$$

which says that we connect points $x, y$ as soon as the balls centered at $x$ and $y$ with respective radii $C^{1/p}U_x^{-1/p}, C^{1/p}U_y^{-1/p}$ contain both $x, y$. When $\mathcal{P}$ is replaced by the lattice $\mathbb{Z}^d$, there is a regime of $p$-values such that the resulting graph is scale-free; see Yukich [46].

Let $\mathcal{G}_\lambda(\mathcal{P})$ denote the restriction of $\mathcal{G}(\mathcal{P})$ to the vertex set $\mathcal{P} \cap \Gamma_\lambda$. Let $H_\lambda := H_\lambda(F)$ denote the number of subgraphs of $\mathcal{G}_\lambda(\mathcal{P})$ that are isomorphic to a given connected simple graph $F$ with $k$ vertices, where *connected* means that there exists a path between each pair of vertices in $F$, and *simple* means it is an unweighted, undirected graph containing neither self-loops nor multiple edges.

**Theorem 3.16.** *(quantitative CLT for the subgraph count in a type of age-dependent random connection model) Assume that the point process $\mathcal{P}$ satisfies the assumptions in Theorem 3.13,* $\mathrm{Var} H_\lambda = \Omega(\lambda^\nu)$ *for some positive constant $\nu > 2/3$, and $g : [0,1] \times [0,1] \to [0,1]$ is a symmetric function which satisfies $g(s,t) \geq \max\{s,t\}$. Let $Z \sim N(0,1)$.*

*(a) If $\psi$ decays exponentially fast, i.e., there exist positive constants $C_1$ and $C_2$ such that $\psi(r) \leq C_1 \mathrm{e}^{-C_2 r}$ for all $r \in (0,\infty)$, then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\mathrm{Var}\, H_\lambda}}, Z\right) = O\left((\log\lambda)^{3d}\lambda^{-1.5\nu+1}\right).$$

*(b) If there exist constants $C_3 \in (0,\infty)$ and*

$$C_4 > 3\left(\frac{d}{3\nu-2}((18-3\nu)\vee 14) + d(k-1) + 1\right)$$

*such that $\psi(r) \leq C_3(1 \vee r)^{-C_4}$ for all $r \in (0,\infty)$, then*

$$d_W\left(\frac{H_\lambda - \mathbb{E}H_\lambda}{\sqrt{\mathrm{Var}\, H_\lambda}}, Z\right) = O\left(\lambda^{-1.5\nu+1+\tau(2d,\beta)+\epsilon}\right),$$

*for all $\epsilon \in (0,\infty)$, where $\beta = C_4/3 - d(k-1) - 1$ and where $\tau$ is defined in* (1.9). *The above bound is $o(1)$ when $\epsilon$ is small enough.*

*Proof.* Define $h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))$ to be the number of subgraphs of $\mathcal{G}(\mathcal{P})$ that are isomorphic to $F$, with vertex set $\{x_1, \ldots, x_k\} \in \mathcal{P}^{k,\neq}$ and associated marks $\{M_{x_1}, \ldots, M_{x_k}\}$. Without loss of generality, assume that the vertex set of $F$ is $\{v_1, \ldots, v_k\}$, and write $i \overset{F}{\sim} j$

to indicate that there exists an edge between $v_i$ and $v_j$ in $F$ for $i, j \in \{1, \ldots, k\}$. Let $\mathrm{aut}(F)$ denote the number of automorphisms of the graph $F$. If we take

$$h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k})) = \frac{1}{\mathrm{aut}(F)} \sum_{\boldsymbol{\pi} \in \mathbf{P}\{1,\ldots,k\}} \prod_{i \overset{F}{\sim} j} \mathbf{1}_{g(M_{x_{\boldsymbol{\pi}(i)}}, M_{x_{\boldsymbol{\pi}(j)}}) \le \psi(d(x_{\boldsymbol{\pi}(i)}, x_{\boldsymbol{\pi}(j)}))},$$

where $\mathbf{P}\{1, \ldots, k\}$ denotes the set of all permutations of $\{1, \ldots, k\}$, then $H_\lambda$ can be represented as a local $U$-statistic as in (3.39).

Next, we upper bound the $L^p$ norm of $h$, $p \in (0, \infty)$. From the construction, for any $\boldsymbol{\pi} \in \mathbf{P}\{1, \ldots, k\}$, the product

$$\prod_{i \overset{F}{\sim} j} \mathbf{1}_{g(M_{x_{\boldsymbol{\pi}(i)}}, M_{x_{\boldsymbol{\pi}(j)}}) \le \psi(d(x_{\boldsymbol{\pi}(i)}, x_{\boldsymbol{\pi}(j)}))}$$

is the indicator of the event that an isomorphism of $F$, in which vertex $v_i$ is mapped to $x_{\boldsymbol{\pi}(i)}$, $1 \le i \le k$, forms a subgraph of $\mathcal{G}(\mathcal{P})$ given that $\{x_1, \ldots, x_k\} \subset \mathcal{P}$. Because $F$ is connected, any such isomorphism must contain a path of the form $(x_{i_1}, \ldots, x_{i_m})$ for some $2 \le m \le k$, such that $d(x_{i_1}, x_{i_m}) = \mathrm{diam}(\{x_1, \ldots, x_k\})$, i.e., a path connecting the most distant points among $\{x_1, \ldots, x_k\}$. Therefore, for any $\boldsymbol{\pi} \in \mathbf{P}\{1, \ldots, k\}$ and $\{x_1, \ldots, x_k\} \in \mathcal{P}^{k,\neq}$, and recalling $g(s, t) \ge \max\{s, t\}$, we estimate the $L^p$-norm of the indicator function as follows:

$$\begin{aligned}
\left\| \prod_{i \overset{F}{\sim} j} \mathbf{1}_{g(M_{x_{\boldsymbol{\pi}(i)}}, M_{x_{\boldsymbol{\pi}(j)}}) \le \psi(d(x_{\boldsymbol{\pi}(i)}, x_{\boldsymbol{\pi}(j)}))} \right\|_p &\le \left\| \prod_{j=1}^{m-1} \mathbf{1}_{\max\{M_{x_{i_j}}, M_{x_{i_{j+1}}}\} \le \psi(d(x_{i_j}, x_{i_{j+1}}))} \right\|_p \\
&\le \prod_{j=1}^{m-1} \mathbb{P}\left( M_{x_{i_j}} \le \psi(d(x_{i_j}, x_{i_{j+1}})) \right)^{1/p}.
\end{aligned}$$

In case (a), where $\psi$ decays exponentially fast, we have

$$\|h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))\|_p \le \frac{k!}{\mathrm{aut}(F)} (C_1 \vee 1)^{(k-1)/p} \mathrm{e}^{-C_2 \mathrm{diam}(\{x_1,\ldots,x_k\})/p}. \tag{3.45}$$

When $\psi$ decays polynomially fast as in case (b), we have

$$\|h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))\|_p \le \frac{k!}{\mathrm{aut}(F)} (C_3 \vee 1)^{1/p} \left( \frac{\mathrm{diam}(\{x_1, \ldots, x_k\})}{k-1} \right)^{-C_4/p}, \tag{3.46}$$

noting that the existence of a path $(x_{i_{\boldsymbol{\pi},1}}, \ldots, x_{i_{\boldsymbol{\pi},m}})$ for some $2 \le m \le k$ implies that there is at least one edge in the path with length at least $d(x_{i_{\boldsymbol{\pi},1}}, x_{i_{\boldsymbol{\pi},m}})/(k-1)$. Take $p = 3 + \epsilon_0$ for some $\epsilon_0 \in (0, \infty)$. Then (3.46) shows that

$$\|h((x_1, M_{x_1}), \ldots, (x_k, M_{x_k}))\|_{3+\epsilon_0} \le C_5 \left(\mathrm{diam}(\{x_1, \ldots, x_k\})\right)^{-C_4/(3+\epsilon_0)}, \tag{3.47}$$

i.e., the $L^{3+\epsilon_0}$-norm of the kernel $h$ decays polynomially fast with exponent $C_4/(3+\epsilon_0)$. Since the marks are independent, the conditional $L^p$-norm of $h$ given $\{x_1,\dots,x_k\} \subset \mathcal{P}$, coincides with the ordinary $L^p$-norm under Palm conditioning. The assertion in case (a) follows from Theorem 3.13 and (3.45). For case (b), choose $\epsilon_0 > 0$ sufficiently small so that

$$\frac{C_4}{3+\epsilon_0} > \frac{d}{3\nu - 2}\left((18-3\nu) \vee 14\right) + d(k-1) + 1.$$

Then (3.47) implies the polynomial-decay condition of Theorem 3.13 (b) with decay exponent $C_4/(3+\epsilon_0)$, and hence the conclusion follows from Theorem 3.13 (b). □

**Acknowledgements**. We thank Anne Marie Svane for constructive and insightful comments on an earlier version of this paper. J. Yukich gratefully acknowledges the kind hospitality and generous financial support from the University of Melbourne, where the writing of this paper was initiated.

# References

[1] Beardwood, J., Halton, J. and Hammersley, J. (1959). The shortest path through many points. *Math. Proc. Camb. Philos. Soc.* **55**, 299–327.

[2] Beneš, V., Hofer-Temmel, C., Last, G. and Večeřa, J. (2020). Decorrelation of a class of Gibbs particle processes and asymptotic properties of $U$-statistics. *J. Appl. Prob.* **57**, 928–955.

[3] Bickel, P. and Breiman, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems, and a goodness of fit test. *Ann. Prob.* **11**, 184-214.

[4] 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.

[5] Błaszczyszyn, B., Yogeshwaran, D. and Yukich, J. E. (2026). Limit theory for Lipschitz-localized statistics in random geometric models. arXiv:2605.28430.

[6] Bradley, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Prob. Surveys* **2**, 107–144.

[7] Brockwell, P. J. and Davis, R. A. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Springer.

[8] Bulinski, A. and Shashkin, A. (2007). *Limit Theorems for Associated Random Fields and Related Systems.* Vol. 10. World Scientific.

[9] Chen, L. H. Y., Goldstein, L. and Shao, Q. M. (2011). *Normal Approximation By Stein's Method*. Springer-Verlag.

[10] Cong, T. and Xia, A. (2023). Convergence rate for geometric statistics of point processes having fast decay of dependence. *Electron. J. Prob.* **28**, 1–35.

[11] Cong, T. and Xia, A. (2024). Normal approximation in total variation for statistics in geometric probability. *Adv. Appl. Prob.* **56**, 106–155.

[12] Davydov, Y. A. (1968). Convergence of Distributions Generated by Stationary Stochastic Processes. *Theory Probab. Its Appl.* **13**, 691–696.

[13] Dereudre, D. (2019). Introduction to the theory of Gibbs point processes. In *Stochastic Geometry*, Lect. Notes in Math., **2237**, 181–229. Springer.

[14] Dinh, T.C., Ghosh, S., Tran, H.S. and Tran, M.H. (2025). Gaussian fluctuations for spin systems and point processes: near-optimal rates via quantitative Marcinkiewicz's theorem. arXiv:2107.08469v3.

[15] Eichelsbacher, P. and Thäle, C. (2014). New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.*, **19**, 1–25.

[16] Forrester, P. J. and Lebowitz, J. L. (2014). Local central limit theorem for determinantal point processes. *J. Stat. Phys.* **157**, 60–69.

[17] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

[18] Hernández Ruíz, L. I. (2025). Results for convergence rates associated with renewal processes. In *Séminaire de Probabilités LII,* Lect. Notes in Math., **2363**, 427–450. Springer.

[19] Hirsch, C., Lachièze-Rey, R. and Owada, T. (2025). Normal approximation for subgraph counts in age-dependent random connection models, arXiv:2505.09318.

[20] Hirsch, C., Otto, M. and Svane, A. M. (2025). Normal approximation for Gibbs processes via disagreement couplings. *Electron. J. Prob.* **30**, 1–56.

[21] Hirsch, C., Otto, M. and Svane, A. M. (2026). Normal and Poisson approximation for Gibbs point processes with pair potentials. arXiv:2601.18695.

[22] Hough, J. B., Krishnapur, M., Peres, Y. and Viràg, B. (2009). *Zeros of Gaussian analytic functions and determinantal point processes.* American Mathematical Society, Providence, RI.

[23] Kallenberg, O. (1983). *Random Measures.* Academic Press, London.

[24] Lachièze-Rey, R., Schulte, M. and Yukich, J. E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Prob.* **29**, 931–993.

[25] Lacker, D., Ramanan, K. and Wu, R. (2023). Local weak convergence for sparse networks of interacting processes. *Ann. Appl. Prob.* **33**, 843–888.

[26] Lavancier, F., Møller, J. and Rubak, E. (2015). Determinantal point process models and statistical inference. *J. R. Stat. Soc., B: Stat. Methodol.* **77**, 853–877.

[27] Lo, T. Y. Y. and Xia, A. (2024). On the rate of normal approximation for Poisson continuum percolation. *Stat. Probab. Lett.* **210**, Article 110110.

[28] Meester, R. and Roy, R. (1996). *Continuum Percolation.* Oxford University Press.

[29] Michelen, M. and Perkins, W. (2022). Strong spatial mixing for repulsive point processes. *J. Stat. Phys.* **189**, Art. No. 9.

[30] Møller, J. and O'Reilly, E. (2021). Couplings for determinantal point processes and their reduced Palm distributions with a view to quantifying repulsiveness. *J. Appl. Prob.* **58**, 469–483.

[31] Penrose, M. D. and Pisztora, A. (1996). Large deviations for discrete and continuous percolation. *Adv. Appl. Prob.* **28**, 29–52.

[32] Reitzner, M. and Schulte, M. (2013). Central limit theorems for $U$-statistics of Poisson point processes. *Ann. Prob.* **41**, 3879–3909.

[33] Revuz, D. (1970). Mesures associées aux fonctionnelles additives de Markov. I. *Trans. Amer. Math. Soc.* **148**, 501–531.

[34] Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Prob. Statist.* **29**, 587–597.

[35] Rosenthal, H. P. (1970). On the subspaces of $L_p(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8**, 273–303.

[36] Ross, S. (2010). *Introduction to Probability Models (Tenth Edition).* Academic Press.

[37] Schreiber, T. and Yukich, J. E. (2013). Limit theorems for geometric functionals of Gibbs point processes. *Ann. Inst. H. Poincaré Prob. Statist.* **49**, 1158–1182.

[38] Schulte, M. and Yukich, J. E. (2023). Rates of multivariate normal approximation for statistics in geometric probability. *Ann. Appl. Prob.*, **33**, 507–548.

[39] 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.

[40] Soshnikov, A. (2002). Gaussian limit for determinantal random point fields. *Ann. Prob.* **30**, 171–187.

[41] Steele, J. M. (1981). Subadditive Euclidean functionals and nonlinear growth in geometric probability. *Ann. Prob.* **9**, 365–376.

[42] Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press.

[43] Trauthwein, T. and Yukich, J. E. (2026). Second-order Poincaré inequalities and localization on the Poisson space. arXiv:2605.23292.

[44] Wiroonsri, N. (2019). Normal approximation for associated point processes via Stein's method with applications to determinantal point processes. *J. Math. Anal. Appl.* **480**, 123396.

[45] Xia, A. and Yukich, J. E. (2015). Normal approximation for statistics of Gibbsian input in geometric probability. *Adv. Appl. Prob.* **47**, 934–972.

[46] Yukich, J. E. (2006). Ultra-small scale-free geometric networks, *J. Appl. Prob.*, **43**, 665–677.