

# Limit theory for statistics of Lipschitz-localized stochastic processes in spatial random models

B. Błaszczyszyn<sup>†</sup> D. Yogeshwaran<sup>‡</sup> and J. E. Yukich<sup>§</sup>

2026-02-12

## Abstract

We consider spatial models represented by marked point processes, which incorporate randomness through a random collection of points (referred to as *sites*) in  $\mathbb{R}^d$  and random marks (referred to as *states*) taking values in a general Polish space. Generally, the locations and states are assumed to be dependent, and the states may also evolve over time, introducing spatially dependent stochastic processes at these locations.

We focus on statistics  $(H_n)_{n \in \mathbb{N}}$  of these models, which take the form of sums of locally dependent *score functions* of sites and states observed within expanding windows  $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d \subset \mathbb{R}^d$ . As  $n \rightarrow \infty$ , we establish Gaussian fluctuations for the centered and normalized statistics of  $H_n$ , which need not necessarily be linear functionals of the underlying marked point process. Under the additional assumption of stationarity, we establish the asymptotics of the mean and variance of these statistics, leading to a multivariate central limit theorem for vectors of such statistics.

To achieve these general limit results, we introduce a foundational approach that captures asymptotic independence in random spatial models via a new *mixing condition on the correlations of marked point processes*. We also provide a more general stabilization paradigm, termed *BL-localization*, for marks, which reinforces this new mixing condition. Both concepts rely on the bounded Lipschitz metric applied to the Palm distributions of marked point processes. Notably, our localization criterion is weaker than the classical stopping-set stabilization—allowing for score functions whose classical stabilizing radii need not be bounded—but strong enough to yield the asymptotic normality of  $(H_n)_{n \in \mathbb{N}}$  via factorial moment expansions and the cumulant method. Our approach does not require the scores to be defined on the infinite window. In the stationary setting, this also enables the derivation of expectation and variance asymptotics by showing existence of distributional limits of scores in the infinite windows.

Spatial models falling within the scope of our general theorems include spin systems on spatial random graphs, sparse networks of interacting diffusions, and particle systems with finite-time horizon. Interacting particle systems considered here include continuum versions of exclusion

---

<sup>†</sup>Inria/ENS, 48 rue Barrault 75013, Paris, France. Email : blaszczy@ens.fr

<sup>‡</sup>Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, Bangalore. Email : d.yogesh@isibang.ac.in

<sup>§</sup>Department of Mathematics, Lehigh University, Bethlehem, PA 18015. Email : jey0@lehigh.edu

models, generalized random sequential adsorption and ballistic deposition models, finite time-horizon epidemic and voter models, and majority models; the generality of the applications to particle systems addresses some open questions raised by Penrose (2008). We also consider statistics of empirical random fields and geostatistical models defined on spatial random graphs.

*Keywords* : Interacting particle systems, interacting diffusions, spin systems, empirical random fields, geostatistical models, central limit theorem, marked point processes, asymptotic de-correlation, BL-localization, stabilizing statistics, factorial moment expansion, cumulant method.

*AMS MSc Class*: 60G55, 60K35, 60F05, 60D05, 60J76, 82B20, 60J60.

## Contents

	<b>2</b>
<b>I Introduction</b>	<b>5</b>
<b>1 Overview and contributions</b>	<b>5</b>
1.1 Contributions of this paper . . . . .	8
1.1.1 Mixing correlations of marked point processes . . . . .	9
1.1.2 BL-localization of marks and limit theory . . . . .	10
1.1.3 BL-localization and stabilization in applications . . . . .	11
1.2 Existing literature . . . . .	12
1.3 Organization of the paper . . . . .	14
<b>2 Notation, terminology and definitions</b>	<b>15</b>
2.1 List of notation . . . . .	15
2.2 Formal definitions: marked point processes . . . . .	17
<b>II Theory and Foundations</b>	<b>21</b>
<b>3 Mixing correlations of marked point processes</b>	<b>21</b>
3.1 Mixing correlations and asymptotic normality . . . . .	21
3.2 Mixing correlations of points, marks, and marked processes . . . . .	24
<b>4 Functionally constructed marks and their localization</b>	<b>26</b>
4.1 Bounded Lipschitz localization . . . . .	27
4.2 Bounded Lipschitz cluster-localization . . . . .	30
4.3 Stabilization via stopping sets . . . . .	31
4.4 Iterated functionally constructed marks . . . . .	33
<b>5 Limit theory for sums of BL-localizing marks</b>	<b>35</b>
5.1 Asymptotic normality for sums of BL-localizing marks . . . . .	36
5.2 Limit theory for sums of stationary BL-localizing marks . . . . .	37
5.3 Asymptotics for sums of scores on the infinite window . . . . .	41
5.4 Remarks . . . . .	41
<b>6 Proofs of results on mixing correlations—Theorem 4.9</b>	<b>44</b>
6.1 Factorial moment expansions of marked point processes . . . . .	45
6.2 Proof of mixing correlation properties of point processes with iterated marks . . . . .	48
6.2.1 Proof of Theorem 4.9 for restricted score functions . . . . .	49
6.2.2 Proof of Theorem 4.9 for general score functions . . . . .	54

<b>7</b>	<b>Proofs of the limit theorems</b>	<b>57</b>
7.1	Cumulant method for asymptotic normality of purely atomic random signed measures . . . . .	57
7.2	Proof of umbrella CLT in Section 3.1 . . . . .	60
7.3	Distributional limits for BL-localizing score functions . . . . .	64
7.4	Proofs of expectation and variance asymptotics in Section 5.2 . . . . .	70
7.5	Proof of multivariate central limit theorem in Section 5.2 . . . . .	72
7.6	Proof of limit theory for scores on infinite window in Section 5.3 . . . . .	76
 <b>III Applications</b>		 <b>77</b>
<b>8</b>	<b>Spin systems on spatial random graphs</b>	<b>77</b>
8.1	The spin model . . . . .	79
8.1.1	Stabilizing interaction graphs . . . . .	80
8.1.2	Averaged weak spatial mixing of spins . . . . .	81
8.2	Main result for spin models . . . . .	83
8.2.1	Proof of CLT for spin models . . . . .	85
8.2.2	Auxiliary statements for spin systems . . . . .	93
8.3	Averaged weak spatial mixing models via disagreement percolation . . . . .	95
8.3.1	Controlling disagreement percolation probabilities via averaged connective constant . . . . .	97
8.3.2	Controlling spin correlations via disagreement percolation and sharp phase transition . . . . .	100
8.3.3	Spin systems on the Poisson-Boolean model . . . . .	104
8.4	Strong spatial mixing for the hard-core model via combinatorial techniques . . . . .	107
<b>9</b>	<b>Interacting diffusions on spatial random graphs</b>	<b>114</b>
9.1	The diffusion model . . . . .	114
9.1.1	Lipschitz assumption on coefficients . . . . .	116
9.1.2	Stabilizing random graphs . . . . .	117
9.2	Main results for the diffusion model . . . . .	118
9.3	Auxiliary statements for diffusions on graphs . . . . .	122
<b>10</b>	<b>Interacting particle systems on spatial random graphs</b>	<b>124</b>
10.1	General model assumptions . . . . .	125
10.1.1	Sites and states of particles . . . . .	126
10.1.2	Update rules and the evolution of states . . . . .	126
10.1.3	Stabilizing interaction graphs . . . . .	128
10.2	Asymptotic normality for statistics of particle systems . . . . .	128
10.2.1	Proofs . . . . .	132
10.3	Markovian interacting particle systems . . . . .	139
10.3.1	Cooperative and random sequential adsorption (CSA and RSA) . . . . .	141
10.3.2	Epidemic spread and voter models . . . . .	143
10.3.3	Majority dynamics in the continuum . . . . .	144
10.3.4	Ballistic deposition . . . . .	145
10.3.5	Further models . . . . .	147
10.4	Discrete-time interacting particle systems . . . . .	149
<b>11</b>	<b>Empirical random fields and geostatistical models</b>	<b>152</b>
11.1	Limit theory for empirical random fields . . . . .	153
11.2	Limit theory for covariograms . . . . .	155
11.3	Geostatistical Boolean models . . . . .	157
11.4	Mixing correlations of random fields with a covariate point process . . . . .	160
<b>12</b>	<b>Further applications and directions</b>	<b>161</b>
<b>Acknowledgements</b>		<b>164</b>
<b>Appendices</b>		<b>165</b>

<b>A</b>	<b>Mixing point processes and interaction graphs in our framework</b>	<b>165</b>
A.1	Mixing point processes . . . . .	165
A.2	Stabilizing interaction graphs . . . . .	166
<b>B</b>	<b>Some consequences of stabilization via stopping sets</b>	<b>171</b>
<b>C</b>	<b>Some consequences of <math>L^1</math>-stabilization</b>	<b>176</b>
	<b>References</b>	<b>180</b>

## Part I

# Introduction

This introductory part presents key examples motivating our study of statistics of marked point processes and spatial random models. The rest of Section 1 describes the new foundational aspects underpinning the general limit theory developed here, provides a concise review of the relevant literature, and outlines the overall structure of our work. Lastly, in Section 2, we provide the basic terminology and definitions used throughout.

## 1 Overview and contributions

This paper considers sums of statistics of spatial random models which are asymptotically de-correlated over spatial domains. We establish the limit theory of these statistics as the spatial domain increases up to  $\mathbb{R}^d$ . This encompasses Gaussian fluctuations over expanding observation windows, as well as asymptotic results for the mean and variance in the stationary case of models. The models involve multiple sources of randomness, namely the random collection of sites in  $\mathbb{R}^d$ , their random initial states, and, in some cases, additional randomness stemming from the evolution of these states. Our set-up is general and yields the limit theory for statistics of stochastic geometric structures such as continuum spin systems, interacting diffusions, interacting particle systems, geostatistical models and empirical random fields. To illustrate, we present six models with differing interaction mechanisms, but nonetheless statistics of these models may be handled with a common framework.

The common input in our representative examples is a point process  $\mathcal{P} = \{x_i\}$ , namely a countable set of random spatial locations in  $\mathbb{R}^d$ , also called *sites* in the case of particle systems. We may, for example, take  $\mathcal{P}$  to be either the Poisson point process on  $\mathbb{R}^d$ , a Neyman-Scott or Matérn cluster process on  $\mathbb{R}^d$ , the Ginibre point process on  $\mathbb{R}^2$ , or a stationary alpha-determinantal point process with an exponentially decreasing kernel. Given the sequence of windows  $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$ ,  $n \in \mathbb{N}$ , we denote the restriction of  $\mathcal{P}$  to  $W_n$  by  $\mathcal{P}_n := \mathcal{P} \cap W_n$ . In many of our models there is an underlying spatial random graph  $\mathcal{G}(\mathcal{P})$  or  $\mathcal{G}(\mathcal{P}_n)$  defined on  $\mathcal{P}$  or  $\mathcal{P}_n$ , respectively, and called the *interaction graph*. For example, one may take this to be the  $k$ -nearest neighbor graph or the Delaunay graph or the Gilbert graph, i.e., the random geometric graph. More generally, the spatial random graphs  $\mathcal{G}(\mathcal{P})$  are in the ‘sparse regime’ (i.e., finite average degree) and the edges are defined by local rules but are not necessarily uniformly bounded.

**Example 1.1** (Spin systems; Section 8). Consider a Gibbs random field defined on an interaction graph  $\mathcal{G}(\mathcal{P}_n)$ , or equivalently a spin system  $\{V_x\}_{x \in \mathcal{P}_n}$  defined with respect to an adjacency relation in the graph  $\mathcal{G}(\mathcal{P}_n)$ . Assuming the spin system satisfies an averaged version of weak spatial mixing, we show that the total spin  $\sum_{x \in \mathcal{P}_n} V_x$  is asymptotically normal as  $n \rightarrow \infty$ . We use both combinatorial and disagreement percolation methods to deduce that models such as the hard-core and Widom-Rowlinson spin models on certain proximity graphs satisfy this version of weak spatial mixing, and consequently, establish that their total spin  $H_n := \sum_{x \in \mathcal{P}_n} V_x$  is asymptotically normal.

**Example 1.2** (Interacting diffusions; Section 9). Consider a system of interacting  $\mathbb{R}^d$ -valued diffusions  $M(x, t) := M(x, t, \mathcal{P}), t > 0, x \in \mathcal{P}$ , defined on the vertices of a locally finite graph  $\mathcal{G}(\mathcal{P})$ . The diffusion at  $x \in \mathcal{P}$  interacts directly with only the diffusions at its neighbors in the graph. Assuming that diffusions start inside a bounded ball in  $\mathbb{R}^d$  and the drift and diffusion coefficients satisfy Lipschitz conditions, we establish the limit theory for statistics of the diffusions such as  $H_n := \sum_{x \in \mathcal{P}_n} \|M(x, t)\|$ , where  $\|\cdot\|$  is the sup-norm for functions in the interval  $[0, t_0]$ .

**Example 1.3** (Cooperative sequential adsorption on graphs; Section 10.3.1). With  $\mathcal{P}$  as above, we encode the interactions between the sites of  $\mathcal{P}_n$  by a proximity graph, such as the  $k$ -nearest neighbor or Delaunay graph. The dynamics of the model go as follows. Let the initial state be  $M(x, 0) = M(x, 0, \mathcal{P}_n) = 0, x \in \mathcal{P}_n$ . Each site is equipped with a unit rate Poisson clock. When the clock at site  $x$  rings, a particle arrives at site  $x$ . The particle is accepted at site  $x$  with a probability depending on the configuration of occupied neighbors of  $x$  in the interaction graph. Once accepted the particle stays forever. We let  $M(x, t) := M(x, t, \mathcal{P}_n)$  be the state of the site  $x$  at time  $t$ . Thus  $M(x, t) = 1$  if a particle is accepted at site  $x$  by time  $t$ , otherwise put  $M(x, t) = 0$ . Setting  $H_n = \sum_{x \in \mathcal{P}_n} M(x, t_0)$  gives the total number of accepted particles at time  $t_0$ . We establish the limit theory as  $n \rightarrow \infty$  for  $(H_n)_{n \in \mathbb{N}}$  and related statistics of the evolved system up to time  $t_0 \in (0, \infty)$ ; the general framework shows there is a thermodynamic limit for the spatial average  $n^{-1}H_n$  and that its fluctuations are Gaussian.

**Example 1.4** (Continuum interacting particle systems; Section 10). As in Example 1.3, let  $\mathcal{P}_n$  be equipped with a local interaction graph (e.g.,  $k$ -nearest neighbor or Delaunay graph), and assume that each site carries a unit rate Poisson clock. Unlike cooperative sequential adsorption, the dynamics now allows particles to *move* between neighboring sites according to an exclusion-type rule that may depend on the local configuration of occupied neighbors. Initial states may be dependent. The system is evolved on  $W_n$  up to a fixed time horizon  $t_0 \in (0, \infty)$ . In the limit  $n \rightarrow \infty$ , we establish limit theorems for additive spatial statistics at time  $t_0$ , such as the fraction of occupied sites, or the fraction of occupied sites having a neighbor within a fixed distance  $\rho$ .

This framework covers a broad class of continuum interacting particle systems, including exclusion processes, ballistic deposition, majority dynamics, epidemic models (contact process, SIR, voter models), and toppling dynamics (e.g., divisible sandpiles).

Problems in spatial statistics where our general theory can be applied include the following, each involving two independent point processes.

**Example 1.5** (Empirical random fields; Sections 11–11.2). Let  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  be a stationary real-valued random field on  $\mathbb{R}^d$ , and let  $\mathcal{P}$  be a stationary point process on  $\mathbb{R}^d$ , assumed to be independent of  $M$ . The *empirical random field* of  $M$  sampled at the points of  $\mathcal{P}$  over a window  $W_n$  is defined by  $\{(x, M(x)) : x \in \mathcal{P}_n\}$ . Under suitable assumptions on the spatial correlation structure of  $\mathcal{P}$  and  $M$ , we establish asymptotics for the expectation and Gaussian fluctuations of this empirical random field as  $n \rightarrow \infty$ . As an immediate consequence, we obtain limit theorems for empirical distribution-type statistics of the real valued random field  $M \in \mathbb{R}$ , such as  $n^{-1} \sum_{x \in \mathcal{P}_n} \mathbf{1}(M(x) < \tau), \tau \in \mathbb{R}$ .

More involved statistics can also be considered. In particular, one may study a *covariogram* estimator of the random field  $M$  of the form

$$\frac{1}{n} \sum_{\substack{x,y \in \mathcal{P}_n \\ x-y \in B_\delta(\pm h)}} |M(x) - M(y)|^2,$$

where pairs of points are selected whose displacement lies in a ball of radius  $\delta > 0$  around  $\pm h \in \mathbb{R}^d$ . Under appropriate spatial correlation assumptions on  $\mathcal{P}$  and on the random field  $M$ , we derive the asymptotic behavior of the expectation and variance of this estimator, as well as Gaussian confidence intervals for this estimator of the covariogram of  $M$ .

**Example 1.6** (Geostatistical Gilbert disc models; Section 11.3). Consider a generalization of the Gilbert disc model, in which balls are positioned at the realization of a stationary point process  $\mathcal{P}$  in  $\mathbb{R}^d$  with bounded but dependent radii determined by an underlying stationary real valued random field  $\{M(x)\}_{x \in \mathbb{R}^d}$  which is independent of  $\mathcal{P}$ . In this variant of the Boolean model, two points communicate with one another, i.e., are joined with an edge, if the balls at both points intersect, thereby giving rise to a random graph  $\mathcal{G}(\mathcal{P})$  on  $\mathcal{P}$  with spatially correlated edge lengths. For example the field value  $M(x)$ ,  $x \in \mathbb{R}^d$ , may be determined by the position of  $x$  with respect to an independent Poisson-Voronoi tessellation of  $\mathbb{R}^d$ . The total edge length in the random graph  $\mathcal{G}(\mathcal{P} \cap W_n)$  exhibits Gaussian fluctuations and, after scaling, its average converges to an asymptotic limit as  $n \rightarrow \infty$ .

This paper studies the afore-mentioned systems and rigorously establishes the above statements. This is achieved by providing a general framework for establishing the limit theory for statistics of a broad class of geometric structures, one encompassing the above examples among others.

The question of limit theory common to the above examples can be framed in terms of establishing asymptotics for the sum of scores

$$H_n = H_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n), \quad n \rightarrow \infty, \quad (1.1)$$

where  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ , assigned to the sites  $x \in \mathcal{P}_n$ , is a family of real-valued characteristics. These characteristics are referred to as *scores*, though they may also be called *marks* or *states*, depending on the context. Here,  $\tilde{\mathcal{P}}_n = \{(x, U(x))\}_{x \in \mathcal{P}_n}$  represents the sites  $x$ , which may also be associated with *initial states* or *pre-marks*  $U(x)$ , collectively denoted as  $\tilde{x} = (x, U(x))$ . The configuration of sites and their pre-marks (if present) induces the real-valued characteristics  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  in a potentially complex yet localized manner. For instance, by making suitable choices of the pre-marks  $U$ , we may define  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  to coincide with  $V_x$  in the case of spin systems, with  $\|M(x, t)\|$  in the setting of interacting diffusions, or with  $M(x, t_0)$  in the context of cooperative sequential adsorption. Analogously,  $\xi$  can be defined in other classes of models as well.

More general than asymptotics for  $H_n$  as in (1.1), we seek limit theory for linear statistics of the associated (possibly) signed measures

$$\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d}x}, \quad n \rightarrow \infty. \quad (1.2)$$

Questions about spatial averages can be cast in terms of establishing expectation asymptotics as  $n \rightarrow \infty$  for (1.1) and (1.2).

The set-up taken here allows for the following sources of dependencies, each of which brings extra spatial correlations:

- the sites  $\mathcal{P}$  may have spatial correlations and the initial states  $U_x, x \in \mathcal{P}$ , may be dependent;
- the scores at sites depend on pre-marks (initial states) through an interaction graph whose edges may not be uniformly bounded;
- the scores and pre-marks may also influence the scores of neighboring sites, introducing additional dependencies among them. In interacting diffusions or particle systems, for example, updates at a site may also modify the states at all neighboring sites; this locally synchronous updating condition is sufficient to cover many models. This is also the case for spin systems.
- in case of interacting diffusions or particle systems, the updates at a site  $x$  need not be Markovian, but may be a function of the entire time-evolved history at  $x$  and at nearby sites.

Furthermore, as mentioned in Section 12, there are potential applications to statistics of random graphs and networks with dynamically changing geometry, shot-noise models on spatially correlated input, dynamic Markov random fields in the continuum, as well as to population genetics models. We do not develop these areas but leave them for future research.

## 1.1 Contributions of this paper

The contributions of this paper are three-fold. In addition to the already indicated diverse applications, we put forward two new foundational aspects underlying the forthcoming general limit theorems for the statistics (1.1) and (1.2):

- we introduce the new notion of mixing involving *correlations for marked point processes*, which uses geometric criteria to systematically capture the asymptotic independence exhibited in various spatial random models
- we provide a more general localization paradigm, *termed BL-localization* (with BL standing for bounded, Lipschitz), for marks, which broadly reinforces this new type of mixing and which is well suited for quantitatively describing dependence of scores on local data even in the presence of unbounded model dependencies. A key mechanism connecting the localization and mixing of correlations of marks is the systematic use of the bounded Lipschitz metric with respect to the Palm distributions of these marks. In examples such as spin systems, we use a variant termed *BL cluster-localization*.

In the following sections, we also shed light on the proof ideas behind some of the main results.

### 1.1.1 Mixing correlations of marked point processes

Dynamic geometric models have multiple sources of dependencies with each bringing in extra correlations. To capture these spatial dependencies, we introduce the concept of *mixing correlations of points (sites) and marks (states)*, which quantifies the asymptotic decorrelation of marked point processes  $\{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))\}_{x \in \tilde{\mathcal{P}}_n}$  and, as such, undergirds the entire paper. Using the shorthand  $\xi_{i,n} := \xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$ , this new mixing property says that the expected value of the product of any two bounded Lipschitz (BL) test functions  $f, g$ , when respectively evaluated on the marks of any two point sets  $\{x_1, \dots, x_p\}$  and  $\{x_{p+1}, \dots, x_{p+q}\}$  in  $W_n$ , and denoted by

$$\mathbb{E}_{[x]_1^{p+q}} [f(\xi_{1,n}, \dots, \xi_{p,n})g(\xi_{p+1,n}, \dots, \xi_{p+q,n})] \rho^{(p+q)}(x_1, \dots, x_{p+q}),$$

approximately factorizes into the product of two expectations  $\mathbb{E}_{[x]_1^{p+1}} [f(\dots)] \rho^{(p)}(\dots)$  and  $\mathbb{E}_{[x]_1^{p+1,p+q}} [g(\dots)] \rho^{(q)}(\dots)$  calculated separately on these two groups of points, where  $\mathbb{E}[\dots]$  stands for the Palm formalism involving expectations  $\mathbb{E}_{[x]_1^{p+q}}$  of  $\mathcal{P}_n$  given the considered locations  $x_1, \dots, x_{p+q}$  and where  $\rho^{(p+q)}$  denotes their correlation functions. We require that the additive error of this approximate factorization decays to zero faster than any power of the separation distance

$$s := \inf_{i \in \{1, \dots, p\}, j \in \{p+1, \dots, p+q\}} |x_i - x_j|$$

between these two group of sites, uniform over the windows  $(W_n)_{n \in \mathbb{N}}$ . Instead of the BL test functions one could use the class  $\mathbb{B}$  of bounded test functions. This gives rise to two mixing conditions, one for each of the two classes BL and  $\mathbb{B}$ , formally described in Definitions 3.1 and 3.6, respectively.

These two new types of mixing for marked point processes (called BL-mixing correlations in case the test functions are bounded Lipschitz and  $\mathbb{B}$ -mixing correlations for all bounded functions) are inspired by the notion of ‘clustering of correlation functions’  $\rho^{(p+q)}(x_1, \dots, x_{p+q})$  of *unmarked* point processes introduced in statistical physics by Martin and Yalcin [105] and reprised in Nazarov and Sodin [110], Błaszczyszyn et al. [21], and Fenzl [57]. The former, involving BL functions, can also be interpreted as a variant of the (BL,  $\theta$ )-dependence for random fields indexed by  $\mathbb{Z}^d$ , as discussed in Bulinski and Spodarev [27, Section 10.1.6], albeit with less restrictive conditions on the decay constants but stronger requirements on the decay rate. Our mixing property presented here serves as an alternative to the classical mixing conditions considered by Heinrich [71], Heinrich et al. [73], Heinrich and Pawlas [72], Ivanoff [78], Jenish and Prucha [82], among others. Unlike these classical approaches, our criterion is formulated through *geometric* conditions, which may in some cases be simpler to verify.

We prove in Theorem 3.3 that BL-mixing correlations of marked point processes along with moment conditions suffice to prove asymptotic normality of the associated  $\xi$ -weighted measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (1.2). More precisely, for bounded measurable test functions  $f$  on  $W_1$ , the integrals  $\mu_n^\xi(f) := \int f d\mu_n^\xi$  centered and suitably normalized converge to standard normal distribution. This is because BL-mixing of correlations of the point process marked by  $\xi$  implies the Brillinger mixing of the random measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  provided  $\xi$  satisfies certain moment conditions.

This approach is applicable for general point processes without stationarity assumptions on  $\mathcal{P}$ . The approach rests on establishing Theorem 7.2, a central limit theorem for a sequence of purely

atomic random signed measures. This theorem, whose proof depends on the cumulant method, appears to have independent interest as it unifies and generalizes some of the earlier abstract central limit theorems for statistics of point processes.

Mixing correlations of marked point processes arises whenever the points and marks separately possess a decorrelation structure; see Proposition 3.7. Moreover, Theorem 4.9 shows that this approach extends to the analysis of more complex marks constructed functionally from simpler ones (such as  $\xi_{i,n} = \xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  mentioned above in (1.1)), provided that the initial marked process  $\tilde{\mathcal{P}}$  satisfies stronger  $\mathbb{B}$ -mixing correlations, in particular with the additive error decaying exponentially to zero. For this, we use factorial moment expansion from Błaszczyszyn [17], Błaszczyszyn et al. [20], together with localization or stabilization of scores (discussed in the next section) and moment conditions on  $\xi$ .

### 1.1.2 BL-localization of marks and limit theory

We introduce a geometric localization criterion, here called fast BL-*localization*, for marks (scores) via the bounded Lipschitz metric on the space of probability distributions, here denoted  $d_{\text{BL}}$ . Specifically, fast BL-*localization* means that for any localization ball  $B_r(x)$  of radius  $r$ , under the Palm expectation  $\mathbb{E}_x$ , the  $d_{\text{BL}}$  distance between  $\xi(x, \tilde{\mathcal{P}}_n \cap B_r(x))$  and  $\xi(x, \tilde{\mathcal{P}}_n)$  decreases faster than any inverse power of  $r$ , uniformly in  $n$ . This criterion appears to be new even in the setting of scores  $\xi(x, \mathcal{P})$  on unmarked point processes  $\mathcal{P}$ . Similar to the classical notion of stabilization via stopping sets, BL-localization measures the dependence of marks on local (marked) data. However, this dependence is quantified using the distributional metric  $d_{\text{BL}}$ , rather than through stopping sets. Although weaker than classical geometric criteria such as the existence of stopping sets or even  $L^q$ -stabilization, the multi-site version of BL-localization—where for each  $p \in \mathbb{N}$ , the joint distribution of the scores  $\xi(x_1, \tilde{\mathcal{P}}_n), \dots, \xi(x_p, \tilde{\mathcal{P}}_n)$  under  $\mathbb{P}_{[\omega]_1^p}$  depends only on the marked configuration inside the balls  $B_r(x_i)$ —is still strong enough to ensure that the family of marked point processes  $\{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))\}_{x \in \mathcal{P}_n}$  exhibits fast BL-mixing correlations uniformly in  $n$ , provided the base marked process  $\tilde{\mathcal{P}}$  satisfies summable exponential  $\mathbb{B}$ -mixing correlations, as shown in Theorem 4.9.

Importantly, such BL-mixing of the newly constructed marking is a key condition in Theorem 3.3 for proving the asymptotic normality of the empirical measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$ . Therefore, BL-localization or BL cluster-localization serve as a practical and verifiable criteria to guarantee asymptotic normality for both  $(H_n)_{n \in \mathbb{N}}$  and  $(\mu_n^\xi)_{n \in \mathbb{N}}$ , as long as the score function  $\xi$  satisfies suitable moment assumptions; see Theorem 5.2.

Theorem 5.2 is a qualitative central limit theorem and we make no attempt here to assess the proximity between  $(\frac{H_n - \mathbb{E}H_n}{\sqrt{\text{Var}H_n}})_{n \in \mathbb{N}}$  and the standard normal. In the special case that  $(H_n)_{n \in \mathbb{N}}$  are functionals of a Poisson point process  $\mathcal{P}$  on a metric measure space, then Trauthwein and Yukich [149] show that the Malliavin-Stein calculus, as developed by Last et al. [98], Trauthwein [148], yields presumably optimal rates of convergence of  $(\frac{H_n - \mathbb{E}H_n}{\sqrt{\text{Var}H_n}})_{n \in \mathbb{N}}$  to the standard normal, assuming only fast BL-localization of the score  $\xi$ , together with a fifth moment condition on  $\xi$ . The rates of convergence, given in terms of the Kolmogorov and Wassertstein distances, improve upon Lachièze-Rey et al. [93] which assumes that scores satisfy a stronger stopping set exponential stabilization criterion and which assumes the

underlying metric measure space is either Euclidean space or a space satisfying a growth assumption on the measure of spheres.

The BL-localization criterion does *not* require any pathwise or realization-wise relationship between the scores and their truncated versions. Instead, it is formulated purely in terms of their probability distributions. This allows us to establish closeness (in the  $d_{\text{BL}}$  metric) between the laws of the random variables  $\xi(x, \tilde{\mathcal{P}}_n \cap B_r(x))$  and  $\xi(x, \tilde{\mathcal{P}}_n)$ , even when their realizations are pathwise non-comparable. This is why we use the term ‘localization’, in contrast to ‘stabilization’, which involves comparing specific realizations of these random variables—either exactly (via stopping sets) or approximately (in  $L^q$ ). This flexibility in BL-localization plays a key role in establishing asymptotic normality for statistics of spin models and interacting diffusions on spatial random graphs.

Theorem 5.2 has wide applicability and in particular covers the case when the sites form a point process having fast mixing correlations and the marks  $\xi$  exhibit exponentially fast BL-localization or BL cluster-localization. The approach requires neither that  $\xi$  satisfy growth conditions nor does  $\xi$  need to be defined on infinite point processes on all of  $\mathbb{R}^d$ ; thus even in the setting of unmarked point processes and stopping-set stabilization, we improve upon existing results; cf. Section 5.4.

Subject to the stationarity of  $\mathcal{P}$  and the fast BL-localization of  $\xi$ , we also provide expectation and variance asymptotics of  $(H_n)_{n \in \mathbb{N}}$  and  $(\mu_n^\xi)_{n \in \mathbb{N}}$ ; cf. Propositions 5.3 and 5.5. Here again, unlike previous limit results for  $(H_n)_{n \in \mathbb{N}}$  and  $(\mu_n^\xi)_{n \in \mathbb{N}}$ , we require neither existence of  $\xi$  nor ‘some limiting version’ of  $\xi$  on the infinite point process but rather prove and use the existence of distributional limits via Palm theory, which resembles the approach via local weak convergence.

### 1.1.3 BL-localization and stabilization in applications

To apply the general limit theorems for (1.1) and (1.2) in applications as described in Examples 1.1-1.6, one needs to verify suitable localization and moment conditions for appropriate  $\xi$  and thereby establish asymptotic decorrelation of  $(\mu_n^\xi)_{n \in \mathbb{N}}$ . This constitutes a significant part of the article as these applications are always not well-explored in spatial settings and both the framework and verification of localization needs to be tailored to fit within our general theory described above.

In case of spin systems as in Example 1.1, the BL cluster-localization of spins follows from an averaged weak spatial mixing assumption of the spin system and good stabilizing properties of the underlying interaction graphs. For bounded degree graphs, we rely upon bounds derived via the combinatorial approach in Sinclair et al. [143] and disagreement percolation methods of van den Berg and Maes [151]. In case of unbounded degree graphs, we also use recent sharp phase transition results from continuum percolation theory Ziesche [159] along with disagreement percolation bounds to show that weak spatial mixing holds with high probability and this suffices for our central limit theorem. Spin systems are an example where the full reach of BL cluster-localization is exploited and where stronger notions of stabilization do not apply.

For interacting diffusions as in Example 1.2, we use Doob’s inequality and the Itô isometry to establish a decorrelation bound with respect to the graph distance as in [95]. Using again the stabilizing properties of the interaction graph, we obtain  $L^2$ -stabilization for the scores or summands comprising  $H_n$ , which is stronger than BL-localization.

In the case of interacting particle systems as in Examples 1.3 and 1.4, we establish classical stopping set stabilization for functionals  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) := h(M(x, t, \tilde{\mathcal{P}}_n)_{t \in [0, t_0]})$  that depend on the time evolution of particle states, where  $h$  is a suitable function of the state  $M(x, t, \tilde{\mathcal{P}}_n)$  at time  $t$ . This involves a careful analysis of a graphical construction that captures the system’s full dependency structure. For each site  $x$ , the history of a particle’s state is determined by a *backwards in time cluster*, defined via chains in an oriented space-time graph. The diameters of these clusters exhibit exponentially decaying tails, which yields the required asymptotic decorrelation. While the stabilization used here is classical, the generality of the framework makes the results non-trivial to prove—even when site locations form a Poisson process or the initial states are i.i.d.

Our general theory aligns naturally with empirical random fields or geostatistical marking as in Examples 1.5 or 1.6, respectively, as the two point processes in those examples are independent. The approach is advantageous as it only requires BL-localization of the random field  $\zeta$  with respect to the underlying point process  $\tilde{\mathcal{P}}$  for the limit theory to hold.

## 1.2 Existing literature

Many spatial random models have been investigated either in the mean-field regime or for Erdős-Rényi-type random graphs; we do not detail these contributions here, as there is little overlap and still less relevance. We shall focus instead on overlap with existing central limit theorems for models on lattice-based or other discrete sparse graphs.

**General limit theorems for marked and unmarked point processes:** Up to now research involving spatial random models has focussed primarily on *static cases involving independent initial states*, i.e. the study of (1.1) and (1.2) with  $U(x)$  being independent marks (states) of points  $x$  (sites) given the ground process  $\mathcal{P}$ . When the collection of sites  $\mathcal{P}$  is a *Poisson point process with independent initial states* (marks), the limit theory for the static random measures at (1.2) is established in Baryshnikov and Yukich [10], Penrose [115], and Penrose and Yukich [121, 122]. The seminal work of Last et al. [98] employs the Malliavin-Stein calculus to establish quantitative central limit theorems for general functionals on the space of Poisson input. Their work is extended and refined by Trauthwein [148]. Additional works establishing quantitative central limit theorems for the functionals  $(H_n^\xi)_{n \in \mathbb{N}}$  on both marked and unmarked Poisson input include Lachièze-Rey et al. [93], Lachièze-Rey [90], Schulte and Yukich [139], Lachièze-Rey et al. [94], Schulte and Yukich [140]; these papers assume that the scores  $\xi$  satisfy stabilization criteria.

The case of Gibbsian point processes is considered in Schreiber and Yukich [135], Xia and Yukich [157] and Hirsch et al. [74], whereas the case of point processes satisfying classical mixing conditions is handled in Heinrich [71], Heinrich et al. [73], Heinrich and Pawlas [72], Ivanoff [78], Jenish and Prucha [82]. In Błaszczyszyn et al. [21], the limit theory for the static measures in (1.2) is further extended to include a ground process  $\mathcal{P}$  (of sites) having fast mixing correlations; see also Fenzl [57] and Cong and Xia [34] for independent (static) marks given such a ground process  $\mathcal{P}$ . Recent work of Chen et al. [32] uses Stein couplings to obtain Berry-Esseen types of bounds on normal approximation of statistics of general random measures, but not necessarily those expressed as sums of marks.

The fast mixing correlations of ground processes is exhibited by point processes which are either stationary  $\alpha$ -determinantal or permanental point processes having a fast decreasing kernel, the point process given by the zero set of a Gaussian entire function, certain rarified Gibbsian point processes or some Cox point processes (see [21]). The general results of [21] extend those of Soshnikov [145] and Nazarov and Sodin [110], who restrict attention to  $\xi \equiv c$ ,  $c$  a constant, that is to say restrict to the case when the random measures at (1.2) are counting statistics. The work of Beneš et al. [11] gives sufficient conditions for fast mixing correlations of a Gibbs particle process (a point process on the space of compact sets of  $\mathbb{R}^d$ ) and uses them to prove a central limit theorem for  $U$ -statistics of such processes.

**Spin Systems:** The above-mentioned central limit theorems and ours have precursors in central limit theorems for discrete random fields (i.e., indexed by graphs or lattices) which involve some notion of mixing or quasi-association. Such theorems have a long history and we refer the reader to Dobrushin and Tirozzi [44], Holley and Stroock [76], Künsch [89], Bolthausen [24], Bulinski and Spodarev [27], Reddy et al. [128] among others. In particular, these theorems were also applied for spin systems on fixed graphs such as lattices and again work in high temperature or low activity regimes. In comparison, our notion of BL-mixing accounts for the randomness of the underlying graphs.

The specific form of weak spatial mixing which we require is taken from Martinelli and Olivieri [106], itself a modification of the well-known ‘Dobrushin-Shlosman uniqueness’ condition [43]; such a condition is crucial when proving uniqueness, concentration inequalities, and fast mixing of Glauber dynamics for spin systems, among other properties; see for example references in [43, 150, 151, 31, 143]. Many existing central limit theorems use mixing in one form or another, but as far as we understand, weak spatial mixing appears weaker than many of these notions and we are unaware of central limit theorems in the discrete setting under such an assumption. There is considerable literature devoted to existence and phase-transitions in the infinite-volume limit for spin systems on lattices [60, 48] but we bypass these aspects here.

To the best of our knowledge, there is relatively little literature on the central limit theorem for spin systems on sparse random graphs. Giardinà et al. [65, 66] and Can [28] have proved asymptotic normality for Ising models on locally tree-like sparse random graphs. They prove a random quenched central limit theorem and annealed central limit theorem for rather general models of sparse random graphs but for an averaged quenched central limit theorem, which is what we study, they require restrictive assumptions on the underlying graphs. Their proofs proceed by evaluating asymptotics for second derivative of the cumulant-generating function explicitly and also yields explicit variance asymptotics at volume-order scale. This proof strategy is similar to that of Ellis [53, Section V7], who uses it to prove a central limit theorem on lattices. In contrast, our approach for spin systems uses coarser information on the graph geometry and applies to many graphs and spin models, though is subject to lower bounds on the growth of the variance, a topic which we do not systematically investigate.

An approach via a quantitative Marcinkiewicz type of theorem was recently used in Dinh et al. [41] to prove normal approximation for continuous spin systems with a non-trivial external field on

lattices. This approach requires showing zero-freeness of the characteristic function of total spins which relies upon zero-freeness of the partition function via the Lee-Yang theory. Zero-freeness of the partition function is known to imply (or is equivalent to) strong spatial mixing in lattice models [42] and some other examples [62, 101, 130] and hence it is in general stronger than the notion of weak spatial mixing employed in our proofs.

**Interacting particle systems:** A significant part of the literature, see Liggett [100], is devoted to existence, long-time behavior, and phase-transitions in particle systems on lattices or deterministic graphs. Our investigation is focused on the evolution in finite-time horizons, as inspired by [117]. The latter also considers existence of particle systems on the infinite Poisson point process, whereas we consider only particle systems on large windows and hence sidestep the question of existence and long-term behavior. By restricting attention to large windows, this allows us to consider more general particle system models than those in [117] or [100]. Understanding the existence in the infinite volume limit as well as understanding the long-time behavior of our models on general point processes is an intriguing question lying beyond the scope of this article, though the techniques of [117, 108] combined with some of our derivations could be helpful in this direction.

Central limit theorems for interacting particle systems on lattices and transitive graphs have a long history, going back to Holley and Stroock [76, 75], Brox and Rost [26], Rost [131] and, more recently Doukhan et al. [46]. These works assume finite-range interactions apart from other particular assumptions on either the state space, the initial states or Markovian dynamics. Specific models in the continuum, including random sequential adsorption, are treated in Penrose and Yukich [120], but we are not aware of central limit theorems for general continuum interacting particle systems beyond those of [117]. The papers Qi [125] and Onaran et al. [111, 112] give functional central limit theorems for statistics of certain spatial birth death processes (and the latter incorporates diffusive dynamics as well), albeit under different assumptions, whereas [117] establishes a functional central limit theorem for more general time-indexed interacting particle systems, subject to finite range conditions and certain assumptions on the birth and death rates. For a more detailed comparison of proof techniques, see Remark 10.2. The underlying point process in these works is the Poisson process and we are not aware of any results on general point processes.

Concerning our other applications such as interacting diffusions or geostatistical marking, we are not aware of existing central limit theorems. Central limit theorems for empirical random fields under strong mixing conditions are given in [113].

### 1.3 Organization of the paper

After introducing some notation and definitions in the next section (Section 2), the rest of the paper is divided into two parts. Part II lays out the general framework of mixing correlations of point processes and localization of scores, the main theoretical results on expectation, variance asymptotics and normal convergence as well as their proofs. We introduce the twin notions of BL-mixing correlations and BL-localizing score functions and use these key concepts to formulate our main limit theorems. The applications of our general theorems to the afore-mentioned spatial random models are not immediate

and we reserve the entirety of Part III to separately and thoroughly investigate each model. While these applications can be read independently of each other, the interactions in these models are commonly described via a stabilizing interaction graph.

## 2 Notation, terminology and definitions

We frame our approach in the setting of marked point processes. Relevant notation is introduced below, and additional background—including key definitions from Palm theory—is provided in the following subsection. For more on point processes and random measures, we refer the reader to [8, 83, 97]. Section 2.1 lists the commonly used notation whereas Section 2.2 elaborates upon the definitions and terminology.

### 2.1 List of notation

- The collection of random sites  $\mathcal{P}$  is a simple point process on  $\mathbb{R}^d$ .
- $\mathbb{K}, \mathbb{K}', \mathbb{M}$  and  $\mathbb{L}$  are Polish spaces, i.e. separable completely metrizable spaces.
- For integers  $1 \leq a < b$ , denote by  $[z]_a^b$  the vector  $(z_a, \dots, z_b)$  with elements  $z_i$  in  $\mathbb{R}^d, \mathbb{K}, \mathbb{N}^0 := \{0, 1, \dots\}$ , or  $\mathbb{R}^d \times \mathbb{K}$ , depending on the context. Denote the concatenation of such tuples by  $[z]_a^b \sqcup [z']_c^d := (z_a, \dots, z_b, z'_c, \dots, z'_d)$ . We also write  $[k] := \{1, 2, \dots, k\}$  for  $k \in \mathbb{N}$ .
- We consider a compatible metric  $d_{\mathbb{K}}$  on the Polish space  $\mathbb{K}$  and we consider 1-product metric  $d_{\mathbb{K}^p}([u]_1^p, [v]_1^p) = \sum_{i=1}^p d_{\mathbb{K}}(u_i, v_i)$  for  $[u]_1^p, [v]_1^p \in \mathbb{K}^p$ . The same applies to other Polish spaces.
- For a vector  $[x]_1^p \in (\mathbb{R}^d)^p$ , we abbreviate  $B_r([x]_1^p) = \bigcup_{i=1}^p B_r(x_i)$ , where  $B_r(x)$  is a ball of radius  $r > 0$  centered at  $x \in \mathbb{R}^d$ .
- $\mathbb{B}(\mathbb{K}^p)$ —the class of real-valued, measurable functions  $f$  on  $\mathbb{K}^p$  with  $|f| \leq 1, p \in \mathbb{N}$ .
- $\mathcal{B}_b = \mathcal{B}_b(\mathbb{R}^d)$ —the class of bounded Borel subsets of  $\mathbb{R}^d$ .
- $\text{BL}(\mathbb{K}^p)$ —the class of Lipschitz(1) functions  $f : \mathbb{K}^p \rightarrow \mathbb{R}$  with supremum norm bounded by 1. This is the class of  $f : \mathbb{K}^p \rightarrow \mathbb{R}$  such that

$$\sup_{[x]_1^p, [y]_1^p \in \mathbb{K}^p} \frac{|f([x]_1^p) - f([y]_1^p)|}{d_{\mathbb{K}^p}([x]_1^p, [y]_1^p)} \leq 1, \quad \sup_{[x]_1^p \in \mathbb{K}^p} |f([x]_1^p)| \leq 1,$$

where  $d_{\mathbb{K}^p}$  is the 1-product metric on  $\mathbb{K}^p$  with respect to the given metric  $d_{\mathbb{K}}$  on  $\mathbb{K}$  (i.e.,  $d_{\mathbb{K}^p}([x]_1^p, [y]_1^p) = \sum_{i=1}^p d_{\mathbb{K}}(x_i, y_i)$ ) These functions are sometimes referred to as bounded (or double-bounded) Lipschitz functions, where both the Lipschitz constant and the function values are bounded. The underlying domain  $\mathbb{K}$  will be omitted whenever it is implicit.

- Given  $\mathbb{K}^p$ -valued random vectors  $[X]_1^p$  and  $[Y]_1^p$ , defined on possibly different probability spaces, the *bounded Lipschitz distance* between the laws of  $[X]_1^p$  and  $[Y]_1^p$  is

$$d_{\text{BL}}(\mathcal{L}([X]_1^p), \mathcal{L}([Y]_1^p)) = \sup_{f \in \text{BL}(\mathbb{K}^p)} |\mathbb{E}f([X]_1^p) - \mathbb{E}f([Y]_1^p)|.$$

Abusing notation we often write  $d_{\text{BL}}([X]_1^p, [Y]_1^p)$  for  $d_{\text{BL}}(\mathcal{L}([X]_1^p), \mathcal{L}([Y]_1^p))$ . Since  $\mathbb{K}$  (and hence  $\mathbb{K}^p$ ) is a separable metric space, the bounded Lipschitz distance  $d_{\text{BL}}$  metrizes convergence of probability measures on  $\mathbb{K}^p$  (see Bogachev [23, Theorem 8.3.2]). Moreover, if  $d_{\mathbb{K}}$  (and hence  $d_{\mathbb{K}^p}$ ) is complete, then this metrization is also complete (see [23, Theorem 8.10.43]). The bounded Lipschitz distance  $d_{\text{BL}}$  is also equivalent to the Fortet–Mourier distance (see Villani [153]).

- Given  $\mathbb{K}$ -valued random variables  $X, Y$ , defined on possibly different probability spaces, the *total variation distance* between the laws of  $X$  and  $Y$  is

$$d_{\text{TV}}(\mathcal{L}(X), \mathcal{L}(Y)) := \frac{1}{2} \sup_{f \in \mathbb{B}(\mathbb{K})} |\mathbb{E}f(X) - \mathbb{E}f(Y)|.$$

When  $\mathbb{K}$  is finite, we have

$$d_{\text{TV}}(\mathcal{L}(X), \mathcal{L}(Y)) = \frac{1}{2} \sum_{k \in \mathbb{K}} |\mathbb{P}(X = k) - \mathbb{P}(Y = k)|.$$

- We let  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{\tilde{x}}$  be a *simple, marked point process* on  $\mathbb{R}^d \times \mathbb{K}$ , where  $\tilde{x} := (x, U(x))$  are random elements in  $\mathbb{R}^d \times \mathbb{K}$  and for any  $x \in \mathbb{R}^d$ , the set  $\{x\} \times \mathbb{K}$  contains at most one element of  $\tilde{\mathcal{P}}$ . Elements  $x \in \mathbb{R}^d$  are interchangeably referred to as *points* or *sites*, while  $U(x)$  are referred to as *pre-marks* or *initial states*. The projection of  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d$  is the collection of sites  $\mathcal{P}$ , also called the *ground process*.
- By  $\mathcal{P}_n := \mathcal{P} \cap W_n, n \in \mathbb{N}$ , we denote the restriction of the ground process  $\mathcal{P}$  to  $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$ , and, with a slight abuse of notation,  $\tilde{\mathcal{P}}_n := \tilde{\mathcal{P}} \cap W_n$  means the restriction of the marked point process  $\tilde{\mathcal{P}}$  whose ground points belong to  $W_n$ . In certain contexts, we write  $\mathcal{P}_\infty := \mathcal{P}$  and  $\tilde{\mathcal{P}}_\infty = \tilde{\mathcal{P}}$ .
- Let  $\mathcal{N}_{\mathbb{R}^d}$  and  $\mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  denote the spaces of locally finite subsets of  $\mathbb{R}^d$  and  $\mathbb{R}^d \times \mathbb{K}$ , respectively, equipped with suitable topologies and  $\sigma$ -algebras. Let  $\hat{\mathcal{N}}_{\mathbb{R}^d}$  and  $\hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}}$  be the space of *finite* subsets of  $\mathbb{R}^d$  and  $\mathbb{R}^d \times \mathbb{K}$  respectively.  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{K}}$  denotes the collection of finite subsets of  $[0, t_0] \times \mathbb{K}$ .
- $\rho^{(p)}([x]_1^p)$ —the *p*th order correlation function of the ground process  $\mathcal{P}$ ; for a simple point process  $\mathcal{P}$  it vanishes whenever  $x_i = x_j$  for any  $i \neq j$ .
- $\tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p)$ —the *p*th order factorial moment measure of  $\tilde{\mathcal{P}}$ , where  $[\tilde{x}]_1^p = [x, u]_1^p \in (\mathbb{R}^d \times \mathbb{K})^p$ .
- $\mathcal{M}_{[u]_1^p}(d[u]_1^p)$ —the *Palm probability distribution* on  $\mathbb{K}^p$  of the random marks  $[U]_1^p = (U_1, \dots, U_p) = (U(x_1), \dots, U(x_p))$  respectively at the fixed locations  $[x]_1^p = (x_1, \dots, x_p)$ .

- $\mathbb{P}_{[\tilde{x}]_1^p}, \mathbb{E}_{[\tilde{x}]_1^p}$ —the *Palm probability distribution* and the corresponding Palm expectation of the marked point process  $\tilde{\mathcal{P}}$  given points  $[x]_1^p$  and their marks  $[u]_1^p$  where  $[\tilde{x}]_1^p = [x, u]_1^p \in (\mathbb{R}^d \times \mathbb{K})^p$ .
- $\mathbb{P}_{[\tilde{x}]_1^p}^!, \mathbb{E}_{[\tilde{x}]_1^p}^!$ —the *reduced Palm versions* and the corresponding expectation are such that the conditioning points  $[x]_1^p$  and marks  $[u]_1^p$  are removed from  $\tilde{\mathcal{P}}$ .
- $\mathbb{E}_{[\tilde{x}]_1^p}[\dots] := \int_{\mathbb{K}^p} \mathbb{E}_{[\tilde{x}]_1^p}[\dots] \mathcal{M}_{[\tilde{x}]_1^p}(d[u]_1^p)$ —the Palm expectation of  $\tilde{\mathcal{P}}$  given points  $[x]_1^p$  without fixing their marks  $[u]_1^p$ .
- $\tilde{\alpha}_{[\tilde{x}]_1^p}^{(l)}(d[\tilde{g}]_1^l)$ —the *Palm factorial moment measures* of  $\tilde{\mathcal{P}}$  of order  $l$  on  $(\mathbb{R} \times \mathbb{K})^l$  under  $\mathbb{P}_{[\tilde{x}]_1^p}^!$ , for fixed  $[\tilde{x}]_1^p$ .
- $\rho_{[\tilde{x}]_1^p}^{(p)}$ —the  $p$ th order correlation functions of the ground process  $\mathcal{P}$  under its reduced Palm versions  $\mathbb{P}_{[\tilde{x}]_1^p}^!$  (without fixing marks).
- $G(\mathcal{X})$  and  $\mathcal{G}(\mu)$  denote graphs on the vertex set  $\mathcal{X}$  and the counting measure  $\mu$  on  $\mathbb{R}^d$ , respectively. The graph metric is denoted by  $d_{G(\mathcal{X})}$  or  $d_{\mathcal{G}(\mu)}$  and the graph boundary for  $\mathcal{X}' \subset \mathcal{X}$  is denoted by  $\partial\mathcal{X}'$ .
- $\mathfrak{B}_k(\mathcal{X}')$  denotes the  $k$ -ball in the graph distance around  $\mathcal{X}' \subset \mathcal{X}$ .
- $G_{\mathcal{X}'}^{[k]}(\mathcal{X}) := G(\mathfrak{B}_k(\mathcal{X}'))$  denotes the induced sub-graph of  $G(\mathcal{X})$  on  $\mathfrak{B}_k(\mathcal{X}')$ . Similar notation is used for  $\mathcal{G}(\mu)$ .
- For  $\mathcal{X} \subset \mathbb{R}^d$ ,  $\mathcal{G}_{\mathcal{X}'}^{(r)}(\mathcal{X}) := \mathcal{G}(\mathcal{X} \cap B_r(\mathcal{X}'))$  denotes the induced sub-graph of  $\mathcal{G}(\mathcal{X})$  on  $\mathcal{X} \cap B_r(\mathcal{X}')$ , the Euclidean  $r$ -ball around  $\mathcal{X}' \subset \mathbb{R}^d$ .
- $Z$  denotes a standard normal random variable.
- A function  $\phi : [0, \infty) \rightarrow [0, \infty)$  is said to be *fast decreasing* if

$$\limsup_{r \rightarrow \infty} r^m \phi(r) \rightarrow 0, \quad \forall m \in \mathbb{N}. \quad (2.1)$$

Further,  $\phi$  is *exponentially decreasing* if there exists  $b \in (0, \infty)$  such that  $\limsup_{r \rightarrow \infty} r^{-b} \log \phi(r) < 0$ .

## 2.2 Formal definitions: marked point processes

- The collection of random sites  $\mathcal{P}$  is a simple point process on  $\mathbb{R}^d$ ; it is a random element in the space  $\mathcal{N}_{\mathbb{R}^d}$  of locally finite subsets of  $\mathbb{R}^d$  endowed with the evaluation  $\sigma$ -algebra, i.e., the  $\sigma$ -algebra generated by the maps  $\mu \mapsto \mu(B)$  for all Borel subsets  $B$ . Let  $\hat{\mathcal{N}}_{\mathbb{R}^d}$  denote the subspace of  $\mathcal{N}_{\mathbb{R}^d}$  consisting of all finite subsets of  $\mathbb{R}^d$ . We note that  $\mathcal{N}_{\mathbb{R}^d}$  is a Polish space under the vague topology ([83, Lemma 4.6]) and that  $\hat{\mathcal{N}}_{\mathbb{R}^d}$  is likewise a Polish space under the vague topology, which is equivalent to the weak topology [83, Chapter 4].

- We let  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{\tilde{x}}$  be a *simple, marked point process*, where  $\tilde{x} := (x, U(x))$ , with points  $x \in \mathbb{R}^d$  and where the marks  $U(x)$  are random elements in  $\mathbb{K}$ , a Polish space. The  $x$  are random points and we note that  $\tilde{x}$  are random elements in  $\mathbb{R}^d \times \mathbb{K}$  and that  $\tilde{\mathcal{P}}$  is a point process on  $\mathbb{R}^d \times \mathbb{K}$ ; i.e., it is a random element in the space  $\mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  of locally finite subsets of  $\mathbb{R}^d \times \mathbb{K}$  (interpreted as counting measures) and endowed with the evaluation  $\sigma$ -algebra, with any set  $\{x\} \times \mathbb{K}$ ,  $x \in \mathbb{R}^d$ , containing at most one element of  $\tilde{\mathcal{P}}$ . Recall that the  $U(x)$  are interchangeably referred to as marks or initial states. The projection of  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d$  is the collection of sites  $\mathcal{P}$ , also called the *ground process*. As above, we take  $\hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}}$  to be the space of *finite* subsets of  $\mathbb{R}^d \times \mathbb{K}$  and this again is a Polish space under the weak topology, while  $\mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  is also a Polish space under the vague topology [83, resp. Lemma 4.5 and Lemma 4.6].
- Let  $\mathcal{P}^{(p)}$  be the  $p$ -th factorial power of the simple point process  $\mathcal{P}$ , i.e., the collection of tuples  $(x_1, \dots, x_p)$  of distinct points of  $\mathcal{P}$ . It forms a point process on  $(\mathbb{R}^d)^p$  with the mean measure  $\mathbb{E}[\mathcal{P}^{(p)}(\cdot)]$  known as the *factorial moment measure* of order  $p$  of  $\mathcal{P}$ . The  $p$ -th order correlation function  $\rho^{(p)}$  is the Radon-Nikodym density (if it exists) of the factorial moment measure with respect to the Lebesgue measure on  $(\mathbb{R}^d)^p$ . For  $\mathcal{P}$  simple this ensures  $\rho^{(p)}([x]_1^p) = 0$  when  $x_i = x_j$  for any  $i \neq j$ . For any non-negative measurable function  $f : (\mathbb{R}^d)^p \rightarrow [0, \infty)$ , we have

$$\mathbb{E}\left[\sum_{(x_1, \dots, x_p) \in \mathcal{P}^{(p)}} f([x]_1^p)\right] = \int_{(\mathbb{R}^d)^p} f([x]_1^p) \rho^{(p)}([x]_1^p) d([x]_1^p).$$

We slightly abuse notation as the term  $[x]_1^p$  on the left-hand side is a random  $p$ -tuple whereas the term  $[x]_1^p$  on the right is deterministic. We assume throughout that the correlation functions  $\rho^{(p)}$ ,  $p \in \mathbb{N}$ , exist and are bounded. We also write  $\rho(x) := \rho^{(1)}(x)$ ,  $x \in \mathbb{R}^d$ .

- Extending this approach, we define the factorial power  $\tilde{\mathcal{P}}^{(p)}$  of order  $p$  for the simple marked point process  $\tilde{\mathcal{P}}$  as the collection of  $p$ -tuples  $[\tilde{x}]_1^p = [x, u]_1^p$  of distinct marked points of  $\tilde{\mathcal{P}}$ . Its mean measure,  $\tilde{\alpha}^{(p)}(\cdot) := \mathbb{E}[\tilde{\mathcal{P}}^{(p)}(\cdot)]$ , is the  *$p$ -th order factorial moment measure* of  $\tilde{\mathcal{P}}$ , considered on  $(\mathbb{R}^d \times \mathbb{K})^p$ . Consequently, for any non-negative measurable function  $f : (\mathbb{R}^d \times \mathbb{K})^p \rightarrow [0, \infty)$ , we have

$$\mathbb{E}\left[\sum_{(\tilde{x}_1, \dots, \tilde{x}_p) \in \tilde{\mathcal{P}}^{(p)}} f([\tilde{x}]_1^p)\right] = \mathbb{E}\left[\sum_{(\tilde{x}_1, \dots, \tilde{x}_p) \in \tilde{\mathcal{P}}^{(p)}} f([x, U]_1^p)\right] = \int_{(\mathbb{R}^d \times \mathbb{K})^p} f([\tilde{x}]_1^p) \tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p),$$

where in the middle expression we adopt the convention that  $U_i = U(x_i)$  for all  $i \in \mathbb{N}$ ,  $[x, U]_1^p = [(x, U)]_1^p = ((x_1, U_1), \dots, (x_p, U_p))$  and where again on the left-hand side  $[\tilde{x}]_1^p = [x, U]_1^p$  is a random  $p$ -tuple of distinct marked points, whereas the term  $[\tilde{x}]_1^p = [x, u]_1^p$  on the right is deterministic.

- For a fixed vector of distinct locations  $[x]_1^p = (x_1, \dots, x_p) \in (\mathbb{R}^d)^p$ , the *(marks-)Palm probability distribution*  $\mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p)$  on  $\mathbb{K}^p$  describes the distribution of the random marks  $[U]_1^p = (U_1, \dots, U_p)$  at these locations. It is defined by

$$\tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p) = \mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p) d[x]_1^p,$$

where  $\tilde{\alpha}^{(p)}$  is the  $p$ th order factorial moment measure of  $\tilde{\mathcal{P}}$ , and  $[\tilde{x}]_1^p = [x, u]_1^p = ((x_1, u_1), \dots, (x_p, u_p))$ . This disintegration holds since the boundedness of the correlation function  $\rho^{(p)}$  ensures the  $\sigma$ -finiteness of  $\tilde{\alpha}^{(p)}$  on the space  $\mathbb{K}$  that is Polish; see [8, Theorem 14.D.10]. Thus, the Palm probability distribution  $\mathcal{M}_{[\tilde{x}]_1^p}^{(p)}(d[u]_1^p)$  exists for almost every  $[x]_1^p$  and is only meaningful for distinct locations  $x_i$ .

- For fixed distinct marked locations  $[\tilde{x}]_1^p = [x, u]_1^p \in (\mathbb{R}^d \times \mathbb{K})^p$ , we denote by  $\mathbb{P}_{[\tilde{x}]_1^p}$  the corresponding *Palm probability distribution* of the (entire) process  $\tilde{\mathcal{P}}$  given points  $[x]_1^p$  and their marks  $[u]_1^p$ . These are again defined via a disintegration of the higher order Campbell's measure  $C^{(p)}(\cdot \times \cdot)$  on  $(\mathbb{R}^d \times \mathbb{K})^p \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  given by

$$C^{(p)}(B \times L) := C_{\tilde{\mathcal{P}}}^{(p)}(B \times L) := \mathbb{E} \left[ \int_{(\mathbb{R}^d \times \mathbb{K})^p} \mathbf{1}([\tilde{x}]_1^p \in B) \mathbf{1}(\tilde{\mathcal{P}} \in L) \tilde{\mathcal{P}}^{(p)}(d[\tilde{x}]_1^p) \right]$$

with respect to the  $p$ th order factorial moment measure  $\tilde{\alpha}^{(p)}$  of  $\tilde{\mathcal{P}}$  (from which  $C^{(p)}$  inherits  $\sigma$ -finiteness):

$$C^{(p)}(d([\tilde{x}]_1^p, \tilde{\mu})) = \mathbb{P}_{[\tilde{x}]_1^p}(d\tilde{\mu}) \tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p). \quad (2.2)$$

The crucial existence of the disintegration in (2.2) follows again from [8, Theorem 14.D.10] and the observation that  $\mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  is Polish. By  $\mathbb{E}_{[\tilde{x}]_1^p}$ , we denote the corresponding expectation with respect to the Palm probability distribution  $\mathbb{P}_{[\tilde{x}]_1^p}$ . These again are defined only for  $\tilde{\alpha}^{(p)}$  a.e.  $[\tilde{x}]_1^p$  and only meaningful for distinct marked locations.

From the above definition, it follows that the Palm expectation satisfies the following identity known as the Campbell-Little-Mecke formula. For a measurable  $f : (\mathbb{R}^d \times \mathbb{K})^p \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow [0, \infty)$ , we have

$$\mathbb{E} \left[ \sum_{(\tilde{x}_1, \dots, \tilde{x}_p) \in \tilde{\mathcal{P}}^{(p)}} f([\tilde{x}]_1^p, \tilde{\mathcal{P}}) \right] = \int_{(\mathbb{R}^d \times \mathbb{K})^p} \mathbb{E}_{[\tilde{x}]_1^p} [f([\tilde{x}]_1^p, \tilde{\mathcal{P}})] \tilde{\alpha}^{(p)}(d[x, u]_1^p). \quad (2.3)$$

- The *reduced Palm versions*  $\mathbb{P}_{[\tilde{x}]_1^p}^!$  and their expectations  $\mathbb{E}_{[\tilde{x}]_1^p}^!$  are such that the conditioning points  $[x]_1^p$  and marks  $[u]_1^p$  are removed from  $\mathcal{P}$ . These versions, which may be formally defined via

$$\mathbb{P}_{[\tilde{x}]_1^p}^! \left( \tilde{\mathcal{P}} \in \cdot \right) = \mathbb{P}_{[\tilde{x}]_1^p} \left( \tilde{\mathcal{P}} - \sum_{i=1}^p \delta_{\tilde{x}_i} \in \cdot \right),$$

satisfy the following modified Campbell-Little-Mecke formula. For a measurable  $f : (\mathbb{R}^d \times \mathbb{K})^p \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow [0, \infty)$ , we have

$$\mathbb{E} \left[ \sum_{(\tilde{x}_1, \dots, \tilde{x}_p) \in \tilde{\mathcal{P}}^{(p)}} f([\tilde{x}]_1^p, \tilde{\mathcal{P}} - \sum_{i=1}^p \delta_{\tilde{x}_i}) \right] = \int_{(\mathbb{R}^d \times \mathbb{K})^p} \mathbb{E}_{[\tilde{x}]_1^p}^! [f([\tilde{x}]_1^p, \tilde{\mathcal{P}})] \tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p). \quad (2.4)$$

- Let  $\mathbb{E}_{[x]_1^p}[\dots] := \int_{\mathbb{K}^p} \mathbb{E}_{[\tilde{x}]_1^p}[\dots] \mathcal{M}_{[w]_1^p}^{(p)}(d[u]_1^p)$  denote the Palm expectation of  $\tilde{\mathcal{P}}$  given points  $[x]_1^p$  without fixing their marks  $[u]_1^p$ . Under this notation, the Campbell-Little-Mecke formula can be re-written as

$$\mathbb{E}\left[\sum_{(\tilde{x}_1, \dots, \tilde{x}_p) \in \tilde{\mathcal{P}}^{(p)}} f([\tilde{x}]_1^p, \tilde{\mathcal{P}})\right] = \int_{(\mathbb{R}^d)^p} \mathbb{E}_{[\tilde{x}]_1^p}[f([\tilde{x}]_1^p, \tilde{\mathcal{P}})] \rho^{(p)}([x]_1^p) d[x]_1^p \quad (2.5)$$

- The *Palm factorial moment measures* of order  $l$  (on  $(\mathbb{R} \times \mathbb{K})^l$ ) of  $\tilde{\mathcal{P}}$  under  $\mathbb{P}_{[\tilde{x}]_1^p}^!$  (for fixed  $[\tilde{x}]_1^p$ ) are denoted by  $\tilde{\alpha}_{[\tilde{x}]_1^p}^{(l)}(d[\tilde{y}]_1^l)$ . The existence of *Palm correlation functions* is guaranteed by the following Palm algebra relation [8, Proposition 3.3.9]

$$\tilde{\alpha}^{(p+l)}(d([\tilde{x}]_1^p \sqcup [\tilde{y}]_1^l)) = \tilde{\alpha}_{[\tilde{x}]_1^p}^{(l)}(d[\tilde{y}]_1^l) \tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p),$$

for  $[\tilde{x}]_1^p \in (\mathbb{R} \times \mathbb{K})^p$ ,  $[\tilde{y}]_1^l \in (\mathbb{R} \times \mathbb{K})^l$ . The above relation also can be re-written in terms of correlation functions and mark distributions as

$$\mathcal{M}_{[x]_1^p \sqcup [y]_1^l}^{(p+l)}(d([u]_1^p \sqcup [w]_1^l)) \rho^{(p+l)}([x]_1^p \sqcup [y]_1^l) = \tilde{\alpha}_{[x, w]_1^p}^{(l)}(d[y, w]_1^l) \mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p), \quad (2.6)$$

where in  $\mathcal{M}_{[x]_1^p \sqcup [y]_1^l}^{(p+l)}(d([u]_1^p \sqcup [w]_1^l))$  we understand  $[u]_1^p, [w]_1^l$  to denote the markings associated to  $[x]_1^p, [y]_1^l$  respectively. Likewise, the Palm distributions  $(\mathbb{P}_{[\tilde{x}]_1^p}^!)_{[\tilde{y}]_1^l}^!$  of  $\mathbb{P}_{[\tilde{x}]_1^p}^!$  exist and satisfy  $(\mathbb{P}_{[\tilde{x}]_1^p}^!)_{[\tilde{y}]_1^l}^! = \mathbb{P}_{[\tilde{x}]_1^p \sqcup [\tilde{y}]_1^l}^!$  for almost all  $[\tilde{x}]_1^p \sqcup [\tilde{y}]_1^l$  with respect to  $\tilde{\alpha}^{(p+l)}(d([\tilde{x}]_1^p \sqcup [\tilde{y}]_1^l))$ .

- The  $l$ th *correlation functions*  $\rho_{[\tilde{x}]_1^p}^{(l)}([y]_1^l)$  of the ground process  $\mathcal{P}$  under its reduced Palm versions  $\mathbb{P}_{[\tilde{x}]_1^p}^!$  (without fixing marks). Formally, these functions are defined as the Radon-Nikodym density of the measure  $\mathbb{E}_{[\tilde{x}]_1^p}^![\mathcal{P}^{(l)}(\cdot)]$  with respect to the Lebesgue measure. Dropping marks in the Palm algebra relation (2.6) we have

$$\rho^{(p+l)}([x]_1^p \sqcup [y]_1^l) = \rho_{[\tilde{x}]_1^p}^{(l)}([y]_1^l) \rho^{(p)}([x]_1^p) \quad (2.7)$$

## Part II

# Theory and Foundations

This part systematically develops the theoretical framework underpinning the paper. In Section 3, we introduce the key concept of asymptotic de-correlation of marked, not necessarily stationary, point processes  $\tilde{\mathcal{P}}$  as well as the notion of BL-mixing marked point processes. We state a general central limit theorem for such point processes, one which implies most of our Gaussian fluctuation results. Section 4 provides a general way to describe and construct asymptotically de-correlated point processes having dependent marks via jointly BL-localizing score functions. Section 5 establishes the general limit theory for statistics of these point processes. This includes central limit theorems, and, under the additional assumption of stationarity, it provides the first and second order limit theory for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ . The proof of the key mixing correlation bound from Section 4 is in Section 6 whereas the proofs of the general limit theorems are in Section 7.

**Assumption 3.1.** *Our standing assumptions are as follows:*

(i)  $\mathcal{P}$  denotes a simple point process on  $\mathbb{R}^d$  such that the correlation functions  $\rho^{(p)}$ ,  $p \in \mathbb{N}$ , exist and  $\kappa_p := \sup_{[x]_1^p \in (\mathbb{R}^d)^p} \rho^{(p)}([x]_1^p) < \infty$ . We put  $\kappa_0 := \max\{\kappa_1, 1\}$ .

(ii) Marked point processes  $\tilde{\mathcal{P}}$  and  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  are assumed to be simple.

## 3 Mixing correlations of marked point processes

We define the main notion of asymptotic de-correlation of marked point processes, called ‘mixing correlations’, allowing us to formulate our fundamental result (Theorem 3.3) concerning Gaussian fluctuations of some macroscopic characteristics of a large class of random spatial models. Next, in Section 3.2 we elaborate upon the properties of mixing correlations, and in particular show that they can be deduced from mixing correlations of points together with mixing correlations of marks. This will be useful when proving that the main mixing hypotheses follow from more elementary assumptions.

### 3.1 Mixing correlations and asymptotic normality

We formulate the main result of this section, a fundamental tool for studying the Gaussian fluctuations of the various spatial models. Following the first lines of the Introduction, consider a *ground point process*  $\mathcal{P} = \{x_i\}$  on  $\mathbb{R}^d$ , and denote  $\mathcal{P}_n := \mathcal{P} \cap W_n$ ,  $n \in \mathbb{N}$ , where  $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$ . Consider a *family of  $\mathbb{R}$ -valued marked point processes*  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}} = (\{(x_i, \xi_{i,n})\}_{x_i \in \mathcal{P}_n})_{n \in \mathbb{N}}$ , where  $\xi_{i,n}$  are real-valued random variables. We are interested in establishing a central limit theorem for the random variables  $\mu_n^\xi(f) := \int f d\mu_n^\xi$  where

$$\mu_n^\xi := \sum_{x_i \in \mathcal{P}_n} \xi_{i,n} \delta_{n^{-1/d}x_i} \quad (3.1)$$

is a (possibly signed) measure on  $W_1$  and where  $f \in \mathbb{B}(W_1)$  is a bounded test function on  $W_1$ .

If  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  are the respective restrictions to  $(W_n)_{n \in \mathbb{N}}$  of a stationary ground point process  $\mathcal{P}$  equipped with i.i.d. marks, i.e.,  $\{\xi_{i,n}\}_{x_i \in \mathcal{P}_n}$  are i.i.d. for each  $n \in \mathbb{N}$ , then, subject to a moment condition on  $\{\xi_{i,n}\}$ , straightforward modifications of [21] yield the asymptotic normality of  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ ,  $f \in \mathbb{B}(W_1)$ , as noted by [57]. The case of dependent marks, which regularly occurs in models of interest, is more delicate.

Instead of imposing a mixing condition on  $\mathcal{P}$  through its sigma-algebra, we do so via its correlation functions. This approach is inspired by the notion of ‘clustering of correlation functions’ of point processes; see Definition 3.4. By  $(\tilde{\mathcal{P}}_l = \{(x, U_l(x))\}_{x \in \mathcal{P}})_{l \in \mathbb{N}}$  we mean a *family* of marked point processes *sharing the same ground point process*  $\mathcal{P}$  in  $\mathbb{R}^d$  and having possibly different marks  $U_l(x) \in \mathbb{K}$ ,  $x \in \mathcal{P}$ , in the same mark space  $\mathbb{K}$ . These marks may depend on the configuration  $\mathcal{P}$  and/or on an auxiliary source of randomness, all represented on a probability space where the marked point processes  $\tilde{\mathcal{P}}_l = \{(x, U_l(x))\}$  are defined.

**Definition 3.1** (BL-mixing correlations of marked point processes).

(i) A marked point process  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$  has BL-mixing correlations if there exists a family  $(\omega_k)_{k \in \mathbb{N}}$  of decreasing functions  $\omega_k : [0, \infty) \rightarrow [0, \infty)$ ,  $\omega_k \leq \omega_{k+1}$  and  $\lim_{s \rightarrow \infty} \omega_k(s) = 0$  for all  $k$ , and such that for all  $p, q \in \mathbb{N}$ ,  $x_1, \dots, x_{p+q} \in \mathbb{R}^d$ , and all  $f \in \text{BL}(\mathbb{K}^p)$ ,  $g \in \text{BL}(\mathbb{K}^q)$ , we have

$$\left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})] \rho^{(p+q)}([x]_1^{p+q}) - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \leq \omega_{p+q}(s), \quad (3.2)$$

where

$$s := d([x]_1^p, [x]_{p+1}^{p+q}) := \inf_{i \in \{1, \dots, p\}, j \in \{p+1, \dots, p+q\}} |x_i - x_j|. \quad (3.3)$$

We refer to  $(\omega_k)_{k \in \mathbb{N}}$  as the correlation decay functions for  $\tilde{\mathcal{P}}$ .

(ii)  $\tilde{\mathcal{P}}$  has fast BL-mixing correlations if in addition to (3.2), for all  $k \in \mathbb{N}$ , the correlation decay function  $\omega_k$  is fast decreasing, in accordance with property (2.1).

(iii) A family of marked point processes  $(\tilde{\mathcal{P}}_l)_{l \in \mathbb{N}}$ , has BL-mixing correlations, if there exists a family  $(\omega_k)_{k \in \mathbb{N}}$  of functions  $\omega_k : [0, \infty) \rightarrow [0, \infty)$  decreasing to zero and such that (3.2) holds uniformly in  $l \in \mathbb{N}$  with respect to the Palm distributions of marks  $\mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})]$  of  $\tilde{\mathcal{P}}_l$ . The family  $(\tilde{\mathcal{P}}_l)_{l \in \mathbb{N}}$  has fast BL-mixing correlations, if the correlation decay functions  $(\omega_k)_{k \in \mathbb{N}}$  are fast decreasing.

Examples of marked point processes  $\tilde{\mathcal{P}}$  having BL-mixing correlations include point processes  $\mathcal{P}$  with mixing correlations and whose marks are generated either by spins, states of interacting diffusions or interacting particle systems on stabilizing interaction graphs on  $\mathcal{P}$ , as detailed in Part III. The upcoming Proposition 3.7 and Corollary 4.10 provide general conditions under which a marked point process  $\tilde{\mathcal{P}}$  has BL-mixing correlations.

**Definition 3.2** (Moment conditions of marks). *The  $\mathbb{R}$ -valued marks of  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}} = (\{(x_i, \xi_{i,n})\}_{x_i \in \mathcal{P}_n})_{n \in \mathbb{N}}$ , namely  $\xi_{i,n} \in \mathbb{R}$ , satisfy a  $p$ -moment condition,  $p \in [1, \infty)$ , if there are constants  $M_p^\xi < \infty$  such that*

$$\sup_{1 \leq n < \infty} \sup_{1 \leq q \leq p} \sup_{x_1, \dots, x_q \in W_n} \mathbb{E}_{[x]_1^q} [\max(1, |\xi_{1,n}|^p)] \leq M_p^\xi < \infty, \quad (3.4)$$

where  $\xi_{1,n}$  is the mark of  $x_1$  under the Palm distribution of  $\tilde{\mathcal{P}}_n$  given the ground points  $x_1, \dots, x_q$ . Here and elsewhere, we adopt the convention that the sup with respect to  $[x]_1^p$  is to be understood as ess sup with respect to the factorial moment measures  $\rho^{(p)}(\mathrm{d}[x]_1^p)$  of the common ground process  $\mathcal{P}$ . Without loss of generality we assume that  $M_p^\xi$  is increasing in  $p$  such that (3.4) holds.

BL-mixing correlations, defined via the class of BL-test functions, are sufficient to establish the central limit theorem for  $\mathbb{R}$ -valued marked point processes.

**Theorem 3.3** (CLT for sums of marks having fast BL-mixing correlations). *Consider the family  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}} := (\{(x_i, \xi_{i,n})\}_{x_i \in \mathcal{P}_n})_{n \in \mathbb{N}}$ ,  $\xi_{i,n} \in \mathbb{R}$ , of  $\mathbb{R}$ -valued marked point process sharing the ground point process  $\mathcal{P}$  in  $\mathbb{R}^d$ . Assume that  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  has fast BL-mixing correlations as in Definition 3.1 (iii) with marks  $\xi_{i,n}$  satisfying the  $p$ -moment condition (3.4) for all  $p \in \mathbb{N}$ . Furthermore, if  $f \in \mathbb{B}(W_1)$  satisfies  $\mathrm{Var} \mu_n^\xi(f) = \Omega(n^\nu)$  for some  $\nu > 0$ , then as  $n \rightarrow \infty$ ,*

$$(\mathrm{Var} \mu_n^\xi(f))^{-1/2} (\mu_n^\xi(f) - \mathbb{E} \mu_n^\xi(f)) \xrightarrow{d} Z, \quad (3.5)$$

where  $Z$  is the standard normal random variable and  $\xrightarrow{d}$  denotes convergence in distribution.

Theorem 3.3, an umbrella central limit theorem, is proved in Section 7.2. It is a central limit theorem for triangular arrays with the summands of  $\mu_n^\xi$  forming the entries of the  $n$ th row of the array,  $n \in \mathbb{N}$ , and it does not require stationarity assumptions. The result encompasses the asymptotic normality of statistics of spin systems (Section 8), interacting diffusions (Section 9), interacting particle systems (Section 10), as well as empirical random fields and geostatistical models (Section 11), with all models considered on windows  $W_n$ ,  $n \rightarrow \infty$ .

Limit theorems for point processes with dependent marks are typically established under mixing assumptions on both the ground process and the marks; see, for example, [71, 73, 72, 78, 82], and [113]. However, such assumptions can be difficult to verify in practice. In contrast, we have assumed an asymptotic independence condition formulated in terms of the correlation structure of these models. While our definition of mixing correlations for marked point processes may appear as yet another form of mixing, it simplifies verification by allowing the correlations of the ground process and the marks to be treated separately (see Proposition 3.7). Furthermore, it supports the analysis of more complex marks constructed functionally from simpler ones satisfying stronger mixing conditions (see Definition 3.6, Theorem 4.9, and Corollary 4.10. In particular, the latter results develop an approach ensuring that the so-defined marks  $\xi_{i,n}$  are such that the family  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  satisfies fast BL-mixing, as defined in Definition 3.1.

The main idea behind the proof is to show, together with moment conditions on the marks  $\xi_{i,n}$ , that the correlation functions defined via Palm expectations  $\mathbb{E}_{[x]_1^k} = \mathbb{E}_{x_1, \dots, x_k}$  by

$$\begin{aligned} m^{[k]_1^{p+q}}([x]_1^{p+q}; n) &= m^{(k_1, \dots, k_{p+q})}(x_1, \dots, x_{p+q}; n) \\ &:= \mathbb{E}_{[x]_1^{p+q}} [(\xi_{1,n})^{k_1} \dots (\xi_{p+q,n})^{k_{p+q}}] \rho^{(p+q)}([x]_1^{p+q}) \end{aligned} \quad (3.6)$$

approximately factorize into  $m^{[k]_1^p}([x]_1^p; n)m^{[k]_{p+1}^{p+q}}([x]_{p+1}^{p+q}; n)$  uniformly in  $n \in \mathbb{N}$  up to an additive error decaying faster than any power of the separation distance  $s$ , defined at (3.3). Here  $x_1, \dots, x_{p+q}$  are distinct points in  $\mathbb{R}^d$  and  $k_1, \dots, k_{p+q} \in \mathbb{N}$ . The approximate factorization (3.6) may be viewed as a *geometric mixing condition* on correlation functions. It implies Brillinger mixing for the (possibly signed) measures  $\sum_{x_i \in \mathcal{P}_n} \xi_{i,n} \delta_x$  (see Lemma 7.1) and hence it implies a central limit theorem for  $(\mu_n^\xi)_{n \in \mathbb{N}}$ ; see Theorem 7.2.

### 3.2 Mixing correlations of points, marks, and marked processes

We recall from [21] the definition of mixing correlation functions for  $\tilde{\mathcal{P}}$ , we introduce a corresponding notion for the marks, and then show that together they imply the mixing correlations of marked point processes as defined in Section 3.1.

**Definition 3.4** (Mixing correlations of the ground process  $\mathcal{P}$ ).

(i) The correlation functions  $\rho^{(p)}$ ,  $p \in \mathbb{N}$ , of  $\mathcal{P}$  are mixing if for all  $k \in \mathbb{N}$  there exists a decreasing function  $\omega_k : [0, \infty) \rightarrow [0, \infty)$  such that  $\lim_{s \rightarrow \infty} \omega_k(s) = 0$  and for all  $p, q \in \mathbb{N}$ ,  $x_1, \dots, x_{p+q} \in \mathbb{R}^d$ , we have

$$|\rho^{(p+q)}([x]_1^{p+q}) - \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q})| \leq \omega_{p+q}(s), \quad (3.7)$$

where  $s := d([x]_1^p, [x]_{p+1}^{p+q})$  is at (3.3). We assume  $\omega_k \leq \omega_{k+1}$  for all  $k \in \mathbb{N}$ . Denote  $C_k := \sup_s \omega_k(s)$  and observe that  $(C_k)_{k \in \mathbb{N}}$  is non-decreasing.

(ii)  $\mathcal{P}$  has fast mixing correlations if  $\omega_k$  is fast decreasing for all  $k \in \mathbb{N}$ .

As seen in [21, Section 2.2] (see also Appendix A), point processes having fast mixing correlations include determinantal and permanental point processes with a fast decreasing kernel, the zero set of a Gaussian entire function, and rarified Gibbs point processes. Recall from Assumption 3.1 that  $\kappa_p := \sup_{[x]_1^p \in (\mathbb{R}^d)^p} \rho^{(p)}([x]_1^p)$ . As noted in [21, (1.12)], fast mixing correlations yields for all  $p \in \mathbb{N}$

$$\kappa_p \leq p C_p \kappa_0^p. \quad (3.8)$$

We formulate the notion of BL-mixing correlations for the marks  $U(x)$  of a point process  $\tilde{\mathcal{P}}$ , without assuming any specific correlation properties of the ground process  $\mathcal{P}$  itself. Recall that each point  $x \in \mathcal{P}$  is associated with a mark  $U(x) \in \mathbb{K}$ , where  $\mathbb{K}$  is a Polish space.

**Definition 3.5** (BL-mixing correlations of marks).

(i) The marks  $U$  of the point process  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{\tilde{x}}$ ,  $\tilde{x} = (x, U)$ ,  $x \in \mathbb{R}^d$ ,  $U \in \mathbb{K}$ , satisfy BL-mixing correlations if there exists a family of decreasing functions  $\omega_k : [0, \infty) \rightarrow [0, 2]$ ,  $k \in \mathbb{N}$ , called correlation decay functions, such that  $\omega_k \leq \omega_{k+1}$  for all  $k \in \mathbb{N}$  and  $\lim_{s \rightarrow \infty} \omega_k(s) = 0$ , and for all  $p, q \in \mathbb{N}$ ,  $x_1, \dots, x_{p+q} \in \mathbb{R}^d$ , and all  $f \in \mathbf{BL}(\mathbb{K}^p)$ , and  $g \in \mathbf{BL}(\mathbb{K}^q)$  we have

$$\left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})] - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \right| \leq \omega_{p+q}(s), \quad (3.9)$$

where  $s := d([x]_1^p, [x]_{p+1}^{p+q})$  is at (3.3).

(ii) The marks  $U$  have fast BL-mixing correlations if  $(\omega_k)_{k \in \mathbb{N}}$  are fast decreasing.

Before formulating the intuitively appealing result that mixing correlations of the points (sites), together with those of the marks (states) as developed in this section, imply the mixing correlations of the marked point process (as in Definition 3.1), we introduce a stronger notion of mixing correlations for the marks— and consequently for the entire process. Specifically, by replacing the class of bounded Lipschitz functions  $\mathbf{BL}(\mathbb{K})$  with the broader class of bounded functions  $\mathbb{B}(\mathbb{K})$ , we obtain the corresponding definitions of  $\mathbb{B}$ -mixing correlations. This stronger form of mixing is instrumental in the applications of Theorem 3.3. Specifically, if the sequence  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  exhibits  $\mathbb{B}$ -mixing correlations that decay to zero exponentially fast (see Definition 4.8 for more precise formulation of this exponential decay), and if the marks  $\xi_{i,n}$  are functionally constructed from this initial marking—i.e.,  $\xi_{i,n} = \xi(x_i, \tilde{\mathcal{P}}_n)$ —then the resulting marked process satisfies BL-mixing correlations, provided that  $\xi$  fulfills certain localization or stabilization conditions; see Theorem 4.9.

**Definition 3.6** ( $\mathbb{B}$ -mixing correlations).

(i) The marks  $U$  of the point process  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{\tilde{x}}$ ,  $\tilde{x} = (x, U)$ ,  $x \in \mathbb{R}^d$ ,  $U \in \mathbb{K}$ , satisfy  $\mathbb{B}$ -mixing correlations if (3.9) holds when BL is replaced by  $\mathbb{B}$ ; i.e., for all  $p, q \in \mathbb{N}$ , and all  $f \in \mathbb{B}(\mathbb{K}^p)$ ,  $g \in \mathbb{B}(\mathbb{K}^q)$ ,

$$\left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})] - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \right| \leq \omega_{p+q}(s), \quad x_1, \dots, x_{p+q} \in \mathbb{R}^d,$$

where  $s = \inf_{i \in \{1, \dots, p\}, j \in \{p+1, \dots, p+q\}} |x_i - x_j|$  and  $\omega_k : [0, \infty) \rightarrow [0, 2]$ ,  $k \in \mathbb{N}$  are decreasing functions such that  $\lim_{r \rightarrow \infty} \omega_k(r) = 0$ . Again, we assume that  $\omega_k \leq \omega_{k+1}$  for all  $k \in \mathbb{N}$ .

(ii) A marked point process  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$  has  $\mathbb{B}$ -mixing correlations if (3.2) holds when BL is replaced by  $\mathbb{B}$  i.e., for all  $p, q \in \mathbb{N}$ , and all  $f \in \mathbb{B}(\mathbb{K}^p)$ ,  $g \in \mathbb{B}(\mathbb{K}^q)$ ,

$$\left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})] \rho^{(p+q)}([x]_1^{p+q}) - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \leq \omega_{p+q}(s), \quad x_1, \dots, x_{p+q} \in \mathbb{R}^d, \quad (3.10)$$

with  $s, \omega_k, k \in \mathbb{N}$  as in Item (i). Similarly, a family of marked point processes  $(\tilde{\mathcal{P}}_l)_{l \in \mathbb{N}}$ , has  $\mathbb{B}$ -mixing correlations if Definition 3.1(iii) holds when BL is replaced by  $\mathbb{B}$ .

(iii) If  $(\omega_k)_{k \in \mathbb{N}}$  are fast decreasing functions, then we respectively speak of the fast  $\mathbb{B}$ -mixing of marks, a marked point process, or family of marked point processes.

Finally, for both BL- and  $\mathbb{B}$ -mixing correlations of marked point processes, we show that these properties can be established for the marks separately, provided the ground process satisfies the corresponding mixing correlations. This opens up a wide spectrum of models, as the marks exhibiting these mixing correlations are numerous. Beyond the obvious examples of independent marks on a ground process exhibiting mixing correlations, and marks depending within a fixed deterministic distance of the points in a Poisson process, we can also handle marks with dependencies governed by graphs on  $\mathcal{P}$  that involve local interactions, which are made precise through the concept of stabilizing interaction graphs, defined in Section 8.1.

**Proposition 3.7** (Joint mixing correlations of points and marks). *Let  $\tilde{\mathcal{P}}$  be a marked point process whose ground process  $\mathcal{P}$  has mixing correlations as in Definition 3.4 with decay function  $\omega$ . If the marks have BL-mixing correlations as in Definition 3.5 with mixing function  $\tilde{\omega}$  (or  $\mathbb{B}$ -mixing correlations as in Definition 3.6 (i) with mixing function  $\tilde{\omega}$ ) then  $\tilde{\mathcal{P}}$  has BL-mixing correlations as in Definition 3.1(i) (respectively,  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii)) with mixing function*

$$\hat{\omega}_k(s) := \omega_k(s) + \tilde{\omega}_k(s) \max(\kappa_p \kappa_q : p + q = k, p, q \in \mathbb{N}).$$

*Proof of Proposition 3.7.* We are given mixing functions  $\omega_k$  and  $\tilde{\omega}_k$  of the ground point process and of the marks, respectively. For all  $[x]_1^p \in (\mathbb{R}^d)^p$ ,  $[x]_{p+1}^{p+q} \in (\mathbb{R}^d)^q$  the triangle inequality gives

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p) g([U]_{p+1}^{p+q})] \rho^{(p+q)}([x]_1^{p+q}) - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \\ & \leq \left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p) g([U]_{p+1}^{p+q})] \left( \rho^{(p+q)}([x]_1^{p+q}) - \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \right) \right| \\ & \quad + \left| \mathbb{E}_{[x]_{p+1}^{p+q}} [f([U]_1^p) g([U]_{p+1}^{p+q})] \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \right. \\ & \quad \left. - \mathbb{E}_{[x]_1^p} [f([U]_1^p)] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \\ & =: I + J. \end{aligned}$$

Now  $f \in \text{BL}(\mathbb{K}^p)$  and  $g \in \text{BL}(\mathbb{K}^q)$  and the bound (3.7) implies that term  $I$  satisfies

$$I \leq |\rho^{(p+q)}([x]_1^{p+q}) - \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q})| \leq \omega_{p+q}(s)$$

where  $s := d([x]_1^p, [x]_{p+1}^{p+q})$  is at (3.3). For term  $J$ , using (3.9) one obtains  $J \leq \tilde{\omega}_{p+q}(s) \kappa_p \kappa_q$ . This proves Proposition 3.7.  $\square$

## 4 Functionally constructed marks and their localization

In this section we assume that the marks of points  $x_i \in \mathcal{P}_n$  (denoted by  $\xi_{i,n}$  in our umbrella Theorem 3.3) are *functionally constructed from*  $\mathcal{P}_n$  and possibly also from pre-existing marks of this input

process. This appeals to the concepts of localization and stabilization, presented, respectively, in Sections 4.1, 4.2 and 4.3: though formulated in two versions, localization appeals to the quantitative distributional convergence of marks whose ground point process is restricted to balls increasing up to  $\mathbb{R}^d$ , whereas the second and stronger version appeals to the stopping-set structure of the marking functions. The first type, which we call BL-localization (since it relies on the bounded Lipschitz metric), and which is central to our approach, allows the marks to be locally dependent; however, their dependence on  $\tilde{\mathcal{P}}$  need not be bounded, and certainly not uniformly bounded. The variant, called BL *cluster-localization*, requires scores to be localized jointly and also requires the localized scores at well-separated points to factorize into ‘asymptotically independent’ marginals.

The main result of this Section, Theorem 4.9 stated in Subsection 4.4, is an important tool linking the umbrella central limit theorem for real-valued marked point processes, namely Theorem 3.3, to the forthcoming central limit theorem given by Theorem 5.2 involving *functionally constructed marks*, and which are localizing on a general, marked input process. The notion of localization is crucial for all limit theorems in Section 5, including the expectation and variance asymptotics.

More precisely, consider a (say ‘pre-marked’) point process  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$ , with marks in some general space  $\mathbb{K}$ . The models considered here allow for a collection of new  $\mathbb{K}'$ -valued marks on  $\tilde{\mathcal{P}}$ , represented as

$$\xi(\tilde{x}, \tilde{\mathcal{P}}), \tilde{x} = (x, u) \in \tilde{\mathcal{P}}.$$

More formally, we give the following definition.

**Definition 4.1** (Marking function). *A marking function is a measurable function  $\xi : (\mathbb{R}^d \times \mathbb{K}) \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  or  $\xi : (\mathbb{R}^d \times \mathbb{K}) \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  where  $\mathbb{K}'$  is a Polish space. The value of the function  $\xi((x, u), \tilde{\mu})$  is relevant only for  $\tilde{x} = (x, u) \in \tilde{\mu}$  and can be defined arbitrarily for  $(x, u) \notin \tilde{\mu}$ .*

## 4.1 Bounded Lipschitz localization

We introduce a form of localization for the marks  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  taking values in a general Polish space  $\mathbb{K}'$ , which is weaker than those used almost ubiquitously in the models considered in the classical theory of stochastic geometry, but which is still strong enough to insure limit theorems for statistics of spin systems, interacting diffusions, and empirical random fields. This localization, termed BL-localization, is understood in terms of *rates of convergence in distribution with respect to the  $d_{\text{BL}}$  metric of the ( $r$ -)restricted version*

$$\xi^{(r)}(\tilde{x}_1, \tilde{\mathcal{P}}) := \xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r(x_1)) \tag{4.1}$$

to  $\xi(\tilde{x}_1, \tilde{\mathcal{P}})$  as  $r \rightarrow \infty$ , provided  $\xi(\tilde{x}_1, \tilde{\mathcal{P}})$  is defined. Here  $B_r(x) := \{y : |y - x| \leq r\}$  denotes the ball of radius  $r$  centered at  $x$  and we shall slightly abuse notation by defining for  $\tilde{\mu} \in \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}}$  and  $B \in \mathcal{B}(\mathbb{R}^d)$  the restriction  $\tilde{\mu} \cap B := \{(x, u) \in \tilde{\mu} : x \in B\}$ .  $B_r^c(x)$  will denote the complement of  $B_r(x)$  in  $\mathbb{R}^d$ .

Unlike classical stabilization, discussed in Subsection 4.3, this localization is well suited when one can only control distributional convergence of marks on balls of increasing radius, and where the balls do not necessarily constitute stopping sets. We use this criterion, here called *bounded Lipschitz*

*localization*, to establish asymptotic normality as well as expectation and variance asymptotics for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ ,  $f \in \mathbb{B}(W_1)$ , under appropriate moment conditions on  $\xi$ ; see Theorem 5.2 and Proposition 5.3.

Recall from Section 2.1 that  $d_{\text{BL}}(X, Y)$  denotes the bounded Lipschitz distance between random vectors  $X$  and  $Y$ . Furthermore, the vector of the values of the score function  $\xi$  evaluated at marked points  $[\tilde{x}]_1^p = [(x, u)]_1^p \subset \tilde{\mu}$ , is denoted by

$$[\xi]_1^p(\tilde{x}, \tilde{\mu}) := (\xi(\tilde{x}_1, \tilde{\mu}), \dots, \xi(\tilde{x}_p, \tilde{\mu})).$$

This additional notation is necessary because the following localization is in terms of convergence rates for the *joint distribution* of

$$[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) = (\xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r(x_1)), \dots, \xi(\tilde{x}_p, \tilde{\mathcal{P}} \cap B_r(x_p)))$$

to  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  as  $r \rightarrow \infty$ .

**Definition 4.2** (BL-localization of marking functions). *Let  $\xi$  be a marking function and, depending on whether  $\tilde{\mathcal{P}}$  is a finite or infinite marked point process,  $\xi$  is defined on either finite or infinite point sets, i.e.,  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  or  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$ .*

(i) *We say that  $\xi$  is a bounded Lipschitz-localizing marking function (BL-localizing for short) on the point process  $\tilde{\mathcal{P}}$  if for all  $p \in \mathbb{N}$  there are functions  $\varphi_p : (0, \infty) \rightarrow [0, 1]$  decreasing to zero such that*

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) \leq 2\varphi_p(r), \quad r > 0, \quad (4.2)$$

where the extra notation  $[x]_1^p$  in  $d_{\text{BL}, [x]_1^p}$  accounts for the distribution of the vector of scores  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  and  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  under the Palm distribution  $\mathbb{P}_{[x]_1^p}$  of  $\tilde{\mathcal{P}}$ .

(ii) *We say that  $\xi$  is BL-localizing on the finite windows of  $\tilde{\mathcal{P}}$  if for all  $p \in \mathbb{N}$  there are functions  $\varphi_p : (0, \infty) \rightarrow [0, 1]$  decreasing to zero such that*

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)) \leq 2\varphi_p(r), \quad r > 0. \quad (4.3)$$

(iii) *We say that  $\xi$  is fast BL-localizing on  $\tilde{\mathcal{P}}$ , or on the finite windows of  $\tilde{\mathcal{P}}$ , if the functions  $(\varphi_p)_{p \in \mathbb{N}}$  are fast decreasing.*

(iv) *Exponential BL-localization signifies that  $\varphi_p$  is exponentially decreasing to zero for all  $p \in \mathbb{N}$ .*

As elsewhere, in (4.2) and (4.3), we adopt the convention that the sup with respect to  $[x]_1^p$  is understood as *ess sup* with respect to the factorial moment measures  $\rho^{(p)}(d[x]_1^p)d[x]_1^p$  of the common ground process  $\mathcal{P}$ . Without loss of generality, we also assume that  $\varphi_p \leq \varphi_{p+1}$  for all  $p \in \mathbb{N}$ .

Note that BL-localization requires neither that the marking function  $\xi$  be bounded nor Lipschitz. Furthermore, the condition  $\varphi_p \in [0, 1]$  is not restrictive since if there exists a non-negative  $\varphi'_p$  satisfying (4.2) or (4.3), then so does  $\varphi_p = \min\{1, \varphi'_p\}$  and the decreasing or fast-decreasing properties of  $\varphi'_p$  are preserved by  $\varphi_p$ .

The BL-localization property, as observed in Remark 4.7 in Section 4.3, follows readily from either classical (stopping-set) stabilization or  $L^q$ -stabilization, and therefore does not restrict the class of models that can be treated. Moreover, the BL-localization criterion captures the closeness of *probability distributions* of  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ ,  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ , and  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  if the latter is defined, which need not be close in a pathwise sense. This weaker, purely distributional notion of localization provides additional flexibility. A variant, introduced in the next section, is exploited in the study of spin systems in Part III, and both are expected to be useful in a broader class of models.

Fundamentally, BL-localization of  $\xi$  on  $\mathcal{P}$  and  $\mathcal{P}_n$  yields

$$[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow{d} [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \quad n \rightarrow \infty,$$

uniformly under all Palm measures  $\mathbb{P}_{[x]_1^p}$ , with a rate of convergence in the  $d_{\text{BL}, [x]_1^p}$ -distance governed by  $4\varphi_p$ . Indeed, put  $X_n := [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ ,  $X_{n,r} := [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ ,  $X_r := [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$ , and  $X := [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  with  $r = d([x]_1^p, \partial W_n) := \min_{i=1, \dots, p} d(x_i, \partial W_n)$  denoting the distance between  $[x]_1^p$  and the boundary of  $W_n$ . Then by (4.2) and (4.3) we have

$$\begin{aligned} d_{\text{BL}, [x]_1^p}(X_n, X) &\leq d_{\text{BL}, [x]_1^p}(X_n, X_{n,r}) + d_{\text{BL}, [x]_1^p}(X_{n,r}, X) \\ &= d_{\text{BL}, [x]_1^p}(X_n, X_{n,r}) + d_{\text{BL}, [x]_1^p}(X_r, X) \\ &\leq 4\varphi_p(d([x]_1^p, \partial W_n)), \end{aligned} \tag{4.4}$$

where equality holds since the random variables  $X_{n,r}$  and  $X_r$  coincide when  $B_r(x_i) \subset W_n$  for all  $i \in \{1, \dots, p\}$ .

This convergence can be extended to the expectations for real-valued  $\xi$  under moment conditions of order  $p = 1 + \epsilon$ ,  $\epsilon > 0$  (as in (3.4) and with  $\xi_{i,n}$  representing  $\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  and  $\xi(\tilde{x}_i, \tilde{\mathcal{P}})$ ) with the rate

$$|\mathbb{E}_x[\xi(\tilde{x}, \tilde{\mathcal{P}}_n)] - \mathbb{E}_x[\xi(\tilde{x}, \tilde{\mathcal{P}})]| \leq \text{Const.} \times \varphi_1(d([x]_1^p, \partial W_n))^{\epsilon/(2+\epsilon)}.$$

See Lemma 7.5 for details.

Last but not least, BL-localization on finite windows (only (4.3), which does not require defining  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  on an infinite marked point process on  $\mathbb{R}^d$ ) allows one to establish *distributional limits* for  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  toward certain *probability kernels* as  $n \rightarrow \infty$ , uniformly under all Palm measures  $\mathbb{P}_{[x]_1^p}$ , with the same rate of convergence, as well as for their expectations. In this setting, these probability kernels play exactly the role of Palm distributions of (possibly non-existent) infinite-volume object  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  in the analysis of expectation and variance asymptotics of the *sums* of these BL-localizing marks, which will be developed further in Sections 5.2, and 7.3–7.4. This relaxation (as well as a variant presented below) makes it possible to study statistics of spin systems on finite windows, as discussed in Section 8, without requiring the existence of an infinite-volume spin system.

## 4.2 Bounded Lipschitz cluster-localization

BL-localization of marks, as formulated in Definition 4.2, is sufficient for the theoretical foundations and applications considered in Part III, except for the spin systems studied in Section 8. These systems involve yet another relaxation of the classical stabilization framework, related to the *irrelevance of the marking function* itself in the notion of BL-localization.

Indeed, BL-localization of marks is a purely distributional notion: it does not refer to a specific marking function  $\xi$ , but rather to *certain laws* of the vectors  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  for  $r > 0$ , which are required to be close to the corresponding laws of  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ , uniformly in  $n$  (and eventually to those of  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$ , if the latter are well defined).

However for Gibbs models there is typically no canonical specification of the vectors  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  for  $r > 0$  i.e., the *joint law of spins* at  $[x]_1^p$  given *their separate local environments*  $\mathcal{P}_n \cap B_r(x_i)$  which involve dependencies arising from overlap of balls  $B_r(x_i)$ . While this does not hinder the proof of BL-localization condition (4.3) for  $p = 1$ , it is a major obstacle in proving BL-localization for  $p \geq 2$ . The following version of BL-localization, called *cluster-localization*, is designed precisely to prove (fast) mixing properties of correlations of  $\xi$  despite the above complications. It considers the vector

$$[\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) = (\xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r([x]_1^p)), \dots, \xi(\tilde{x}_p, \tilde{\mathcal{P}} \cap B_r([x]_1^p))), \quad (4.5)$$

whose components are the values of the score function  $\xi$  computed with respect to the input process  $\tilde{\mathcal{P}}$  restricted to the *union of respective balls*  $B_r([x]_1^p) = \bigcup_{i=1}^p B_r(x_i)$ .

**Definition 4.3** (BL cluster-localization of marks). *In the setting of Definition 4.2, we say:*

(i) *The score function  $\xi$  is BL cluster-localizing on the point process  $\tilde{\mathcal{P}}$  if, for all  $p \in \mathbb{N}$ , there exist functions  $\varphi_p : (0, \infty) \rightarrow [0, 1]$  decreasing to 0 and constants  $\delta_p \in (0, 1]$  such that the following two conditions hold:*

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) \leq 2\varphi_p(r), \quad r > 0, \quad (4.6)$$

*and for all  $p \geq 2$ ,  $l \in \{1, \dots, p-1\}$  and  $x_1, \dots, x_p \in \mathbb{R}^d$  such that  $r < s([x]_1^l, [x]_{l+1}^p)^{\delta_p}$ , we have*

$$d_{\text{BL}, [x]_1^p}([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), ([\xi^{\cup(r)}]_1^l(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{\cup(r)}]_{l+1}^p(\tilde{x}, \tilde{\mathcal{P}}))) \leq 2\varphi_p(r). \quad (4.7)$$

(ii) BL cluster-localization on finite windows, (iii) fast BL cluster-localization, and (iv) exponential BL cluster-localization are defined *mutatis mutandis* as in Definition 4.2.

**Remark 4.4** (BL-localization vs BL cluster-localization). Replacing  $[\xi^{(r)}]_1^p$  in Definition 4.2 by  $[\xi^{\cup(r)}]_1^p$ , which corresponds to retaining only the first condition (4.6) of BL cluster-localization—a property we sometimes refer to as BL *union-localization*—is often easier to verify. Moreover, condition (4.6) alone already implies the distributional limits of  $[\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  to  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  and/or to the probability kernels discussed in the previous section; see the detailed arguments in Section 7.3 and, in particular, Remark 7.6.

However, BL union-localization does allow for cross-dependence between the neighborhoods of distinct points  $x_i$ , since these neighborhoods are taken jointly. As a consequence, union-localization alone *does not rule out long-range interactions* between individual marks and therefore does not, by itself, suffice to establish mixing correlations, nor even variance asymptotics for sums of localized marks; see the counterexample given in Example 7.7 in Section 7.3.

From this perspective, condition (4.6) is complemented in Definition 4.3 of BL cluster-localization by an additional cluster-level localization condition, namely (4.7), which enforces a cut-off of interactions between groups of points separated at scales larger than  $r$ . Taken together, these two conditions imply all the distributional limits developed in Section 5 for sums of BL-localizing marks, including expectation and variance asymptotics in the stationary setting; see the details in Remark 7.8, Section 7.4.

### 4.3 Stabilization via stopping sets

We consider marked point processes with marking functions satisfying stopping set stabilization, versions of which have appeared in [93, 140] in the setting of finite windows and also [21, 116, 123] in the setting of infinite windows.

**Definition 4.5** (Radius of stabilization). *Given a marking function  $\xi : W_n \times \mathbb{K} \times \hat{\mathcal{N}}_{W_n \times \mathbb{K}} \rightarrow \mathbb{K}'$ ,  $n \in \mathbb{N} \cup \{\infty\}$  (as usual for  $n = \infty$ , we take  $W_\infty := \mathbb{R}^d$ ) and  $(x, u) \in \tilde{\mu} \in \mathcal{N}_{W_n \times \mathbb{K}}$  we define the radius of stabilization  $R_{W_n}^\xi((x, u); \tilde{\mu})$  to be the smallest  $r \in \mathbb{N}$  such that*

$$\xi((x, u), \tilde{\mu} \cap B_r(x)) = \xi((x, u), (\tilde{\mu} \cap B_r(x)) \cup (\tilde{\nu} \cap B_r^c(x))) \quad (4.8)$$

for all  $\tilde{\nu} \in \mathcal{N}_{W_n \times \mathbb{K}}$ . The definition is naturally extended to  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  by setting  $R_{W_n}^\xi((x, u); \tilde{\mu}) := R_{W_n}^\xi((x, u); \tilde{\mu} \cap W_n)$ .

For  $n \in \mathbb{N}$  the radius of stabilization satisfies the bound

$$R_{W_n}^\xi((x, u); \tilde{\mu}) \leq \lceil \text{diam}(W_n) \rceil. \quad (4.9)$$

Indeed, (4.8) holds for  $r = \text{diam}(W_n)$ . For  $n = \infty$  we abbreviate  $R_{W_\infty}^\xi$  by  $R^\xi$  and put  $R^\xi := \infty$  if no finite  $r$  satisfies (4.8).

We refer the reader to the discussion below [116, Definition 2.1] for justification of measurability of the radius of stabilization. By definition, whether the radius of stabilization satisfies the bound  $R_{W_n}^\xi((x, u), \tilde{\mu}) \leq r$  is determined by the realization of  $\tilde{\mu} \cap B_r(x)$  and not on the locations of other points in  $W_n$  or their marks. In other words,  $B_{R_{W_n}^\xi((x, u), \tilde{\mu})}(x)$  is a *stopping set*, that is to say for all  $r > 0$ ,

$$\{\tilde{\mu} \in \hat{\mathcal{N}}_{W_n \times \mathbb{K}} : x \in \mu, R_{W_n}^\xi((x, u), \tilde{\mu}) \leq r\} = \{\tilde{\mu} \in \hat{\mathcal{N}}_{W_n \times \mathbb{K}} : x \in \mu, R_{W_n}^\xi((x, u), \tilde{\mu} \cap B_r(x)) \leq r\} \quad (4.10)$$

and analogously for  $W_\infty$  with  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ .

**Definition 4.6** (Stabilizing marking functions via stopping sets).

(i) We say that  $\xi$  is stabilizing on the marked point process  $\tilde{\mathcal{P}}$  if for all  $p \in \mathbb{N}$  there are functions  $\varphi_p : (0, \infty) \rightarrow [0, 1]$  decreasing to zero such that

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{P}_{[\mathbb{Z}_1^p]}(R^\xi(\tilde{x}_1, \tilde{\mathcal{P}}) > r) \leq \varphi_p(r), \quad r > 0. \quad (4.11)$$

(ii) We say that  $\xi$  is stabilizing on finite windows of  $\tilde{\mathcal{P}}$  if for all  $p \in \mathbb{N}$  there are functions  $\varphi_p : (0, \infty) \rightarrow [0, 1]$  decreasing to zero such that

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[\mathbb{Z}_1^p]}(R_{W_n}^\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n) > r) \leq \varphi_p(r), \quad r > 0. \quad (4.12)$$

(iii) We say that  $\xi$  is fast stabilizing on  $\tilde{\mathcal{P}}$ , or on finite windows of  $\tilde{\mathcal{P}}$ , if  $(\varphi_p)_{p \in \mathbb{N}}$  are fast decreasing. Without loss of generality we always assume that  $\varphi_p \leq \varphi_{p+1}$  for all  $p \in \mathbb{N}$ .

**Remark 4.7** (Comparison between BL-localization and stabilization). We justify our earlier remarks that BL-localization is weaker than stabilization via stopping sets and  $L^q$ -stabilization.

(i) (Stabilization via stopping sets is stronger than BL-localization and BL cluster-localization). The definition of the radius of stabilization gives

$$\sup_{f \in \text{BL}((\mathbb{K}')^p)} \left| \mathbb{E}_{[\mathbb{Z}_1^p]}[f([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] - \mathbb{E}_{[\mathbb{Z}_1^p]}[f([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \right| \leq 2 \sum_{i=1}^p \mathbb{P}_{[\mathbb{Z}_1^p]}(R^\xi(\tilde{x}_i, \tilde{\mathcal{P}}) > r).$$

Thus, stabilization implies BL-localization, fast stabilization implies fast BL-localization, and fast stabilization on finite windows implies fast BL-localization on finite windows. The same line of arguments extends to establish (4.6) and (4.7), namely BL cluster-localization.

The stopping set stabilization property, stronger than localization, confers benefits such as the explicit construction of scores on the infinite window  $W_\infty := \mathbb{R}^d$  as well as weaker moment conditions in the multivariate central limit theorem; this is described in Appendix B.

(ii) ( $L^q$ -stabilization is stronger than BL-localization). Given  $q \geq 1$  and  $\mathbb{K}'$ -valued marks  $\xi$ , the marks satisfy  $L^q$ -stabilization with respect to  $\tilde{\mathcal{P}}$  if for all  $p \in \mathbb{N}$

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{E}_{[\mathbb{Z}_1^p]} |\xi(\tilde{x}_1, \tilde{\mathcal{P}}) - \xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r(x_1))|^q \rightarrow 0$$

as  $r \rightarrow \infty$ . It, too, is stronger than BL-localization. Indeed

$$\begin{aligned} & \sup_{f \in \text{BL}((\mathbb{K}')^p)} \left| \mathbb{E}_{[\mathbb{Z}_1^p]}[f([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] - \mathbb{E}_{[\mathbb{Z}_1^p]}[f([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \right| \\ & \leq \sum_{i=1}^p \mathbb{E}_{[\mathbb{Z}_1^p]} [|\xi(\tilde{x}_i, \tilde{\mathcal{P}}) - \xi(\tilde{x}_i, \tilde{\mathcal{P}} \cap B_r(x_i))|] \\ & \leq \sum_{i=1}^p \mathbb{E}_{[\mathbb{Z}_1^p]} [|\xi(\tilde{x}_i, \tilde{\mathcal{P}}) - \xi(\tilde{x}_i, \tilde{\mathcal{P}} \cap B_r(x_i))|^q]^{1/q}, \end{aligned} \quad (4.13)$$

where in the first inequality we used the BL property of  $f$  with respect to the 1-product metric on  $\mathbb{R}^p$ .  $L^4$ -stabilization has been used by Lachièze-Rey [90, (1.8)] to obtain rates of normal convergence with respect to the Kolmogorov distance and variance asymptotics for geometric statistics of spatial shot-noise fields of Poisson processes. In our approach,  $L^1$ -stabilization, which is still weaker than stopping set stabilization, provides a more natural representation of the limits of the mean and variance established with the minimal assumption of BL-localization. See the first remark in Section 5.4 and further developments in Appendix C.

#### 4.4 Iterated functionally constructed marks

We furnish conditions showing how to verify the general assumptions of Theorem 3.3 when the marks  $\xi_{i,n}$  are constructed as a BL-localizing or BL cluster-localizing function of the input. This result is established in Theorem 4.9 below, the main result of Section 4.

Given a marked point process  $\tilde{\mathcal{P}}$  having mixing correlations as in Definition 3.1, we equip  $\mathcal{P}$  with a new set of functionally constructed marks via a marking function  $\xi$  and establish that the new marked point process retains its mixing correlations. To treat this case, we require however the following strengthened version of  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$  (Definition 3.6), one controlling the growth rate in terms of  $s$  and  $k$  of the decay functions  $\omega_k(s)$ . This strengthened version is a natural requirement when using FME expansions and the method of moments in general.

**Definition 4.8** (Summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$  and  $\mathcal{P}$ ).

(i) *The marked point process  $\tilde{\mathcal{P}}$ , is said to have summable exponential  $\mathbb{B}$ -mixing correlations if it has fast  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii) with functions  $\omega_k(s) = C_k\phi(s)$ ,  $k \in \mathbb{N}$ , such that the  $C_k$  are non-decreasing and where the  $\omega_k(s)$  satisfy the twin growth conditions*

$$\sup_{k \in \mathbb{N}} \frac{C_k}{k^{ak}} < \infty \quad (4.14)$$

for some constant  $a \in [0, 1)$  and

$$\limsup_{r \rightarrow \infty} \frac{\log \phi(r)}{r^b} \in [-\infty, 0), \quad (4.15)$$

for some constant  $b \in (0, \infty)$ .

(ii) *Similarly, the family  $(\tilde{\mathcal{P}}_l = \{(x, U_l(x))\}_{x \in \mathcal{P}})_{l \in \mathbb{N}}$  of marked point processes has summable exponential  $\mathbb{B}$ -mixing correlations if (3.10) holds for functions  $\omega_k(s) = C_k\phi(s)$ ,  $k \in \mathbb{N}$ , with non-decreasing constants  $C_k$ , satisfying (4.14) and (4.15) that is to say for all  $l, p, q \in \mathbb{N}$ , and all  $f \in \mathbb{B}(\mathbb{K}^p)$ ,  $g \in \mathbb{B}(\mathbb{K}^q)$ , we have*

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^{p+q}} [f([U]_1^p)g([U]_{p+1}^{p+q})\rho^{(p+q)}([x]_1^{p+q})] - \mathbb{E}_{[x]_1^p} [f([U]_1^p)]\rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([U]_{p+1}^{p+q})]\rho^{(q)}([x]_{p+1}^{p+q}) \right| \\ & \leq C_{p+q}\phi(s), \quad x_1, \dots, x_{p+q} \in \mathbb{R}^d. \end{aligned} \quad (4.16)$$

(iii) *The point process  $\mathcal{P}$  has summable exponential mixing correlations if it has mixing correlations as in Definition 3.4(i) with  $\omega_k(s) = C_k \phi(s)$ ,  $k \in \mathbb{N}$ , and where  $\phi$  and  $(C_k)_{k \in \mathbb{N}}$  satisfy the above conditions.*

The ‘exponential’ in the terminology refers to the condition (4.15), whereas ‘summability’ refers to (4.14) as it implies finite bounds on the sums  $\sum_k C_k w^k / k!$ ,  $w > 0$ . These are technical necessities for our proofs; see Remark 6.6.

Appendix A and [21, Section 2.2] provide examples of point processes  $\mathcal{P}$  satisfying summable exponential mixing correlations. Equipping such  $\mathcal{P}$  with independent marks yields a marked point process  $\tilde{\mathcal{P}}$  having summable exponential  $\mathbb{B}$ -mixing correlations. Such point processes, which form the foundation of the spatial models in Part III, are enough to get us started in the iteration procedure and allows us to construct new point processes having valuable mixing properties. Indeed, if the ground point process  $\mathcal{P}$  exhibits summable exponential mixing correlations—or more generally, if a ‘pre-marked’ process  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{\tilde{x}} = \sum_{x \in \mathcal{P}} \delta_{(x, U(x))}$  exhibits summable exponential  $\mathbb{B}$ -mixing correlations, in the sense defined above—then the process  $\tilde{\tilde{\mathcal{P}}} = \sum_{x \in \mathcal{P}} \delta_{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}))}$ , obtained by equipping points  $x \in \mathcal{P}$  with marks  $\xi(\tilde{x}, \tilde{\mathcal{P}}) \in \mathbb{K}'$  that are fast stabilizing (resp. fast BL-localizing or fast BL cluster-localizing), also exhibits fast  $\mathbb{B}$ -mixing (resp. fast BL-mixing) correlations.

In other words, mixing properties of marked point processes are preserved under the construction of additional marks. The proof of this result and its corollary is deferred to Section 6.2.2.

**Theorem 4.9** (Fast mixing properties of point processes with functionally constructed iterated marks). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  having summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8. Consider the functionally constructed mark  $V(x) := \xi(\tilde{x}, \tilde{\mathcal{P}})$ , with  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  and set  $\tilde{\tilde{\mathcal{P}}} := \sum_{x \in \mathcal{P}} \delta_{(x, V(x))}$ , a marked point process on  $\mathbb{R}^d \times \mathbb{K}'$ . Fix  $p, q \in \mathbb{N}$ .*

(i) *If  $\xi$  satisfies fast BL-localization (4.2) on  $\tilde{\mathcal{P}}$  for  $p, q$  and  $p + q$ , then  $\tilde{\tilde{\mathcal{P}}}$  satisfies the fast BL-mixing correlation condition for  $p$  and  $q$ , that is to say (3.2) holds for  $f \in \text{BL}((\mathbb{K}')^p)$ ,  $g \in \text{BL}((\mathbb{K}')^q)$ , and some fast decreasing function  $\omega_{p+q}$ .*

(ii) *The statement of Item (i) above remains valid if  $\xi$  satisfies fast BL cluster-localization—that is, conditions (4.6) and (4.7)—on  $\tilde{\mathcal{P}}$  for  $p, q$ , and  $p + q$ .*

(iii) *If  $\xi$  is fast stabilizing on  $\tilde{\mathcal{P}}$  as in (4.11) (via stopping sets) for  $p, q$  and  $p + q$ , then  $\tilde{\tilde{\mathcal{P}}}$  has fast  $\mathbb{B}$ -mixing correlations for  $p$  and  $q$ , i.e., (3.2) holds for  $f \in \mathbb{B}((\mathbb{K}')^p)$ ,  $g \in \mathbb{B}((\mathbb{K}')^q)$ , and some fast decreasing function  $\omega_{p+q}$ .*

*In all cases (i), (ii), and (iii), the fast decreasing function  $\omega_{p+q}$  associated to  $\tilde{\tilde{\mathcal{P}}}$  depends on the functions  $\varphi_{p+q}$  and  $\phi$  as well as on the constants  $(C_k)_{k=1}^\infty$  and  $\kappa_{p+q}$  related to  $\xi$  and  $\tilde{\mathcal{P}}$ , respectively. In particular, there exist constants  $C < \infty$ ,  $c > 0$  which depend on  $p, q$ , such that for all  $s > 0$*

$$\omega_{p+q}(s) \leq C \exp(-cs^b) + C \varphi_{p+q}(cs^{b(1-a)/2(p+q+d)}), \quad (4.17)$$

where  $a, b$  are as in (4.14) and (4.15), respectively.

**Corollary 4.10** (Mixing correlations of families of point processes with functionally constructed iterated marks). *Assume that the family  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  of marked point processes in  $\mathbb{R}^d \times \mathbb{K}$  has summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8(ii). Put  $V_n(x) = \xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  and  $\tilde{\tilde{\mathcal{P}}}_n := \sum_{x \in \mathcal{P}_n} \delta_{(x, V_n(x))}$ ,  $n \in \mathbb{N}$  with  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  being a marking function.*

(i) *If  $\xi$  is fast BL-localizing on finite windows of  $\tilde{\mathcal{P}}$ , as in Definition 4.2(ii) and (iii), or fast BL cluster-localizing as in Definition 4.3(i) and (iii), then the family  $(\tilde{\tilde{\mathcal{P}}}_n)_{n \in \mathbb{N}}$  has fast BL-mixing correlations as in Definition 3.1 (i), (ii), and (iii).*

(ii) *If  $\xi$  is fast stabilizing on finite windows of  $\tilde{\mathcal{P}}$  as in Definition 4.6(ii) and (iii), then the family  $(\tilde{\tilde{\mathcal{P}}}_n)_{n \in \mathbb{N}}$  has fast  $\mathbb{B}$ -mixing correlations as in Definition 3.6 (ii) and (iii).*

Corollary 4.10 will help establish the upcoming Gaussian fluctuations in Theorems 5.2 and 5.4 for fast BL-localizing or BL cluster-localizing statistics of point processes with summable exponentially  $\mathbb{B}$ -mixing correlations. However, weaker assumptions suffice for score functions having a bounded interaction range, i.e. those which coincide with a restricted version  $\xi^{(r)}$  for some  $r \in (0, \infty)$ . Score functions having bounded interaction ranges include local U-statistics. Corollaries 6.7 and 6.8 and the discussion therein further develops these remarks.

Finally, the procedure of constructing new marks  $V(x)$  and  $V_n(x)$ , for  $x \in \mathcal{P}$ , from existing pre-marks can in principle be iterated, yielding yet another family of marks on  $\mathcal{P}$ . This procedure could be referred to as *iterated functionally constructed marks*. However, a technical issue arises in this iteration. In the original setting, the pre-marks are required to satisfy *summable exponential  $\mathbb{B}$ -mixing correlations*, encoded through conditions (4.14) and (4.15), with decay functions of the form  $\omega_k(s) = C_k \phi(s)$ . While the *exponential decay* requirement in condition (4.15) can be readily ensured for the newly constructed marks by imposing explicit assumptions on the localizing or stabilizing function  $\varphi_{p+q}$  appearing in (4.17), verifying the corresponding *summability* condition (4.14) for the associated constants  $C_k$  is not always straightforward.

## 5 Limit theory for sums of BL-localizing marks

Having introduced the notions of marked point processes with mixing correlations and BL-localizing marking functions (including the BL-cluster-localization variant), we now establish the limit theory for functionally constructed marks that are localizing (in either sense) on point processes with summable exponentially decaying  $\mathbb{B}$ -mixing correlations. The general limit results established here form the foundation for the models introduced in Section 1 and are further developed through applications in Sections 8–11 of Part III.

In Section 5.1, we first state a central limit theorem for a sequence of statistics of marked point processes constructed as sums of BL-localizing or BL cluster-localizing marks on expanding windows. This result refines the central limit theorem of our general umbrella result, Theorem 3.3, using the localization property of functionally constructed marks to guarantee their BL-mixing correlations. It neither requires stationarity assumptions on the ground point process nor the existence of a marking

function on an infinite marked point process, thus making the result applicable in particular to the study of spin systems and systems of particle systems on growing windows.

When stationarity of the ground point process is assumed, we establish expectation and variance asymptotics for sums of BL-localizing marks in Section 5.2. These results can equally be obtained under the variant of BL cluster-localization—specifically tailored to applications to spin systems—but, for the sake of simplicity of exposition, we state them only under the default BL-localization assumption. Section 5.3 considers a variant of the above results when the marking function is defined on the entire point process, rather than a finite window.

Finally, we comment on the assumptions underlying our results (such as moment conditions and variance bounds) and outline possible extensions in Section 5.4. That section shows that stronger assumptions—such as stopping-set stabilization or  $L^1$ -stabilization—yield strengthened limit theorems. This is in contrast with Sections 5.1–5.3, where the limit theorems are derived under the weaker assumption of BL-localization. This is in contrast with Sections 5.1–5.3, where the limit theorems are derived under the weaker assumption of BL-localization.

## 5.1 Asymptotic normality for sums of BL-localizing marks

As in Section 4, consider a ‘pre-marked’ point process  $\tilde{\mathcal{P}} = \{\tilde{x} := (x, U(x))\}_{x \in \mathcal{P}}$ , with marks  $U(\cdot)$  in a Polish space  $\mathbb{K}$ . For a ground point process  $\mathcal{P}$  on  $\mathbb{R}^d$ , we have set  $\mathcal{P}_n := \mathcal{P} \cap W_n$ ,  $n \in \mathbb{N}$ , where  $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$ . Consider a  $\mathbb{R}$ -valued marking function  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  applied to marked points  $\tilde{x}$  of  $\tilde{\mathcal{P}}_n$ . As in (3.1) consider (possibly signed) measures

$$\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d}x}, \quad (5.1)$$

where the marking function  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  produces the (final) states  $\xi_{i,n}$  in (3.1).

**Definition 5.1** (Moment conditions for marks on finite windows). *We say that  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  satisfies a  $p$ -moment condition with respect to  $\tilde{\mathcal{P}}$  on finite windows if  $\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  satisfy the  $p$ -moment condition (3.4), that is to say*

$$\sup_{1 \leq n < \infty} \sup_{1 \leq q \leq p} \sup_{x_1, \dots, x_q \in W_n} \mathbb{E}_{[\tilde{x}]_1^q} [\max(1, |\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)|^p)] \leq M_p^\xi < \infty. \quad (5.2)$$

We refine the Gaussian fluctuations of Theorem 3.3 for the sequence  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ ,  $f \in \mathbb{B}(W_1)$ .

**Theorem 5.2** (CLT for sums of fast BL-localizing marks). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  having summable exponential  $\mathbb{B}$ -mixing correlations, as in Definition 4.8. Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a marking function satisfying the  $p$ -moment condition on finite windows (5.2) for all  $p > 1$ . Assume that  $\xi$  is fast BL-localizing on finite windows of  $\tilde{\mathcal{P}}$  as in Definition 4.2 (ii) and (iii) or that  $\xi$  is fast BL cluster-localizing on finite windows of  $\tilde{\mathcal{P}}$  as in Definition 4.3(ii) and (iii). If  $f \in \mathbb{B}(W_1)$  satisfies  $\text{Var} \mu_n^\xi(f) = \Omega(n^\nu)$  for some  $\nu > 0$ , then as  $n \rightarrow \infty$ ,  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$  satisfy the central limit theorem (3.5).*

*Proof.* The proof is an immediate corollary of Theorem 3.3 and Theorem 4.9(i) for  $\xi$  BL-localizing or, respectively (ii), for BL cluster-localization. Indeed, our assumptions on  $\tilde{\mathcal{P}}$  and  $\xi$ , together with part (i) or part (ii) of Theorem 4.9 and part (i) of Corollary 4.10 yields that the family of marked point processes  $\tilde{\mathcal{P}}_n = \{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))\}_{x \in \mathcal{P}_n}, n \in \mathbb{N}$ , has fast BL-mixing correlations as in Definition 3.1(i), (ii), and (iii). By the assumed moment condition (5.2) on  $\xi$  and assumed lower bound for  $\text{Var } \mu_n^\xi(f)$ , it follows that  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$  satisfies all assumptions of Theorem 3.3, thus yielding the central limit theorem for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ .  $\square$

Theorem 5.2 extends Theorems 1.13 and 1.14 of [21], the main results of that paper, by allowing for a significantly broader class of input  $\mathcal{P}$  and scores  $\xi$ : (i) the point process  $\mathcal{P}$  need not be stationary (ii)  $\mathcal{P}$  may carry dependent marks, (iii) the score  $\xi$  need not satisfy stopping set stabilization, and (iv)  $\xi$  need not satisfy a power-growth condition. These comments also apply to the upcoming theorems and propositions in this section. Additionally, if we take  $\xi \equiv 1$  in Theorem 5.2 then this gives the main result of [110, Theorem 1.5], which establishes a central limit theorem for the ‘number count’ statistic of the zero set of a Gaussian analytic function. Such statistics are linear statistics of the ground point process. We establish asymptotic normality for linear statistics of the marks themselves, a more complex issue. Our applications usually take  $\mathcal{P}$  to be stationary, but Theorems 3.3 and 5.2 do not require this assumption.

In the context of Theorem 5.2, it is natural to inquire whether one could deduce a central limit theorem under a  $(2 + \varepsilon)$ -moment condition on  $\xi$ , instead of assuming that all moments of  $\xi$  exist. As seen in the upcoming Theorem 5.4, we may find general conditions (stationarity, volume-order variance growth and translation invariance of  $\xi$ ) under which a  $(2 + \varepsilon)$ -moment condition on  $\xi$  would suffice.

## 5.2 Limit theory for sums of stationary BL-localizing marks

Our theorems so far have not used any stationarity assumption on the marked point process, i.e., invariance of the distribution of  $\tilde{\mathcal{P}}$  with respect to the translation of the sites. However, doing so yields expectation and variance asymptotics and thus also a (multivariate) central limit theorem with explicit limiting covariances. We establish the limit of the expectation and variance of linear statistics of the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  defined at (1.2) under the assumption that  $\xi$  is translation-invariant, i.e., for all  $y \in \mathbb{R}^d$  we have  $\xi(\cdot + y, \cdot + y) = \xi(\cdot, \cdot)$ , where “ $\cdot + y$ ” acts only on the locations of the ground process on  $\mathbb{R}^d$ .

The results of Section 5.1 assume that  $\xi$  is well-defined only on finite point processes, as is the case for example in spin systems and interacting particle systems. To develop explicit asymptotics of  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  on  $\tilde{\mathcal{P}}_n = \tilde{\mathcal{P}} \cap W_n$  as  $n \rightarrow \infty$ , one can use stopping set or  $L^1$ -stabilization as in Section 4.3; see the remarks (i)-(iii) in Section 5.4 and also Appendices B and C. However, the weaker framework of BL-localization developed in Section 4.1 suffices to prove the limit of the expectation and variance of linear statistics of the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$ .

Indeed, BL-localization facilitates the derivation of certain *Palm distributional limits* of  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ ,

specifically for almost all  $x \in \mathbb{R}^d$  and for  $\rho^{(2)}$ -almost all  $[x]_1^2 \in (\mathbb{R}^d)^2$ :

$$\xi(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[n \rightarrow \infty]{\text{BL}, x} \Upsilon_{\mathbf{0}} \quad (5.3)$$

$$[\xi]_1^2(\tilde{x}, \tilde{\mathcal{P}}_n) = \left( \xi(\tilde{x}_1, \tilde{\mathcal{P}}_n), \xi(\tilde{x}_2, \tilde{\mathcal{P}}_n) \right) \xrightarrow[n \rightarrow \infty]{\text{BL}, [x]_1^2} \Upsilon_{(\mathbf{0}, z)} = \left( \Upsilon_{(\mathbf{0}, z)}(1), \Upsilon_{(\mathbf{0}, z)}(2) \right), \quad (5.4)$$

where  $z = x_2 - x_1$ ,  $[x]_1^p$  in BL-convergence  $\xrightarrow{\text{BL}, [x]_1^p}$  represents the distributional convergence of the vector of scores  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  under the Palm distribution  $\mathbb{P}_{[x]_1^p}$  of  $\tilde{\mathcal{P}}$ . To demonstrate the existence of these limits, we prove the Cauchy property of the corresponding sequences in the bounded Lipschitz metric and exploit the completeness property of the metric; see Sections 7.3-7.4.

In contrast to stopping set stabilization, which insures a.s. convergence to a limit random variable (cf. Appendix B), BL-localization provides only distributional convergence to asymptotic probability distributions (more precisely *probability kernels*<sup>1</sup>) represented by the random variable  $\Upsilon_{\mathbf{0}}$  and the vector  $(\Upsilon_{(\mathbf{0}, x)}(1), \Upsilon_{(\mathbf{0}, x)}(2))$ . However, these are sufficient to formulate the asymptotics of the mean and variance as presented in the forthcoming Proposition 5.3. More precisely, the two moments of these distributions correspond to the limits of two correlation functions of the marks of  $\xi$  evaluated on finite windows. As mentioned in Remark 4.4, the same limits remain valid when BL cluster-localization is assumed.

Equipped with these asymptotic probability kernel representations of the mean and the correlation function of  $\xi$ , we prove the expectation and variance asymptotics for the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$ . We emphasize that these asymptotics heavily depend on the stationarity of the point processes and the translation invariance of the score functions. We do not seek minimal hypotheses to ensure these asymptotic results; this comment particularly applies to the assumed fast mixing of  $\xi$  and the summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$ . Refinements of these assumptions are discussed in Section 5.4.

Denote by  $\rho \equiv \rho^{(1)}(x)$  the intensity of the (stationary) point process  $\tilde{\mathcal{P}}$ .

**Proposition 5.3** (Asymptotic mean and variance). *Let  $\tilde{\mathcal{P}}$  be a stationary marked point process on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ . Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a translation invariant marking function.*

(i) *Let  $\xi$  be fast BL-localizing on all finite windows of  $\tilde{\mathcal{P}}$  as in (4.3) for  $p = 1$  and assume  $\xi$  also satisfies a  $(1 + \varepsilon)$ -moment condition on finite windows as at (5.2) for some  $\varepsilon > 0$ . Then for almost all  $x \in \mathbb{R}^d$  as  $n \rightarrow \infty$  the Palm distributional limit of  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ , given by  $\Upsilon_{\mathbf{0}}$  in (5.3), exists and for all  $f \in \mathbb{B}(W_1)$*

$$\left| n^{-1} \mathbb{E} \mu_n^\xi(f) - \rho \mathbb{E}[\Upsilon_{\mathbf{0}}] \int_{W_1} f(x) dx \right| = O(n^{-1/d}). \quad (5.5)$$

*If  $\xi$  satisfies (4.3) for  $p = 1$  with  $\varphi_1$  only decreasing to zero, then the right-hand side of (5.5) is  $o(1)$ .*

<sup>1</sup>For a given order  $p \in \mathbb{N}$ ,  $\Upsilon_{[x]_1^p}$  represents a probability kernel from  $(\mathbb{R}^d)^p$  to  $\mathbb{R}^p$ . That is, for fixed  $[x]_1^p \in (\mathbb{R}^d)^p$ , the vector  $\Upsilon_{[x]_1^p}$  represents some probability distributions on  $\mathbb{R}^p$ . Furthermore, for a given measurable function  $f : \mathbb{R}^p \rightarrow \mathbb{R}$ , the mappings  $(\mathbb{R}^d)^p \ni [x]_1^p \mapsto \mathbb{E}[f(\Upsilon_{[x]_1^p})] \in \mathbb{R}$  are measurable. This is a consequence of  $\Upsilon_{[x]_1^p}$  being the limit of the Palm kernels of  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ .

(ii) Assume  $\xi$  is fast BL-localizing on finite windows of  $\tilde{\mathcal{P}}$  as in (4.3) for  $p \in \{1, 2\}$  and assume  $\xi$  satisfies a  $(2 + \varepsilon)$ -moment condition on finite windows as at (5.2) for some  $\varepsilon > 0$ . Then, for  $\rho^{(2)}$ -almost all  $x_1, x_2 \in \mathbb{R}^d$  as  $n \rightarrow \infty$  the Palm distributional limit of  $(\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n), \xi(\tilde{x}_2, \tilde{\mathcal{P}}_n))$  given by  $\Upsilon_{(\mathbf{0}, z)} = (\Upsilon_{(\mathbf{0}, z)}(1), \Upsilon_{(\mathbf{0}, z)}(2))$ ,  $z = x_2 - x_1$ , in (5.4) exists.

Further, if  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8 then for all  $f \in \mathbb{B}(W_1)$

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^\xi(f) = \sigma^2(\Upsilon) \int_{W_1} f(x)^2 dx \in [0, \infty), \quad (5.6)$$

whereas for all  $f, g \in \mathbb{B}(W_1)$

$$\lim_{n \rightarrow \infty} n^{-1} \text{Cov}(\mu_n^\xi(f), \mu_n^\xi(g)) = \sigma^2(\Upsilon) \int_{W_1} f(x)g(x) dz, \quad (5.7)$$

where

$$\sigma^2(\Upsilon) = \rho \mathbb{E}[\Upsilon_0^2] + \int_{\mathbb{R}^d} (\mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1)\Upsilon_{(\mathbf{0}, z)}(2)]\rho^{(2)}(\mathbf{0}, x) - \rho^2 \mathbb{E}[\Upsilon_0^2]) dz \in [0, \infty). \quad (5.8)$$

The proof of the above result is in Section 7.4. Under stronger stabilization or localization assumptions,  $\mathbb{E}_0[\xi(\tilde{\mathbf{O}}, \tilde{\mathcal{P}})]$ ,  $\mathbb{E}_0[\xi^2(\tilde{\mathbf{O}}, \tilde{\mathcal{P}})]$ , and  $\mathbb{E}_{\mathbf{0}, x}[\xi(\tilde{\mathbf{O}}, \tilde{\mathcal{P}})\xi(\tilde{x}, \tilde{\mathcal{P}})]$ , or variants thereof, can be used in place of expressions  $\mathbb{E}[\Upsilon_0]$ ,  $\mathbb{E}[\Upsilon_0^2]$ , and  $\mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1)\Upsilon_{(\mathbf{0}, z)}(2)]$ , respectively. We elaborate in Remarks (i),(ii) and (iii) of Section 5.4.

To establish the multivariate central limit theorem for the vector  $(\mu_n^{\xi_i}(f))_{i=1, \dots, m}$  formed by the statistics of different score functions  $\xi_i$ ,  $i = 1, \dots, m$ , on the same input process  $\tilde{\mathcal{P}}$ , we need *joint BL-localization of these score functions*. This means the vector  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$  must satisfy the respective assumptions of Definition 4.2 (i), replacing  $\xi \in \mathbb{K}'$  therein by  $\boldsymbol{\xi} \in \mathbb{R}^m$ . Specifically, BL-localization of  $\boldsymbol{\xi}$  of (Palm) order  $p$  on finite windows means that condition (4.3) is satisfied by  $[\boldsymbol{\xi}]_1^p \in (\mathbb{R}^m)^p$ , with the bounded Lipschitz functions defining the metric  $d_{\text{BL}, [\cdot]_1^p}$  on the probability measures on  $\mathbb{R}^m$  considered with respect to the  $\ell^1$ -metric on  $(\mathbb{R}^m)^p$ . Recalling that BL-localization involves probability distributions and not specific realizations of random variables, it follows that BL-localization of the marginals  $\xi_i$  does not imply BL-localization of  $\boldsymbol{\xi}$ .

Similar to (5.3) and (5.4), under two orders  $p = 1, 2$  of this joint BL-localization, one can establish the convergence of two Palm kernels of  $[\boldsymbol{\xi}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ :

$$\boldsymbol{\xi}(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[n \rightarrow \infty]{\text{BL}, x} \boldsymbol{\Upsilon}_0 = \left( \Upsilon_0(i) \right)_{i=1, \dots, m}, \quad (5.9)$$

$$[\boldsymbol{\xi}]_1^2(\tilde{x}, \tilde{\mathcal{P}}_n) = \left( \boldsymbol{\xi}(\tilde{x}_1, \tilde{\mathcal{P}}_n), \boldsymbol{\xi}(\tilde{x}_2, \tilde{\mathcal{P}}_n) \right) \xrightarrow[n \rightarrow \infty]{\text{BL}, [\cdot]_1^2} \boldsymbol{\Upsilon}_{(\mathbf{0}, z)} = \left( (\Upsilon_{(\mathbf{0}, z)}(1, i))_{i=1, \dots, m}, (\Upsilon_{(\mathbf{0}, z)}(2, i))_{i=1, \dots, m} \right), \quad (5.10)$$

for almost all  $x$  and for  $\rho^{(2)}$ -almost all  $[x]_1^2$ , where  $z = x_2 - x_1$ .

The following result combines the Gaussian fluctuations established in Theorem 5.2 and the asymptotics of the mean and variance provided in Proposition 5.3 to establish a multivariate central limit

theorem for  $(\mu_n^{\xi_i}(f))_{i=1,\dots,m}$ . Importantly, this result relaxes the moment conditions of the former theorem to some moment  $p > 2$ , following an approach suggested by Matthias Schulte. However, this may result in a degenerate central limit theorem unless the variance is of volume order.

**Theorem 5.4** (Multivariate CLT for statistics of point processes with fast BL-localizing marks under  $(2 + \varepsilon)$ -moment condition). *Consider a stationary marked point process  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ , having summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8. Let  $\xi_i : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  be translation invariant score functions. Assume that the vector  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$  is jointly fast BL-localizing on finite windows of  $\tilde{\mathcal{P}}$  in the sense that for all  $p \in \mathbb{N}$ , condition (4.3) is satisfied by  $[\boldsymbol{\xi}]_1^p \in (\mathbb{R}^m)^p$ . If for some  $\varepsilon > 0$ ,  $\xi_i$ ,  $i = 1, \dots, m$  satisfy the  $(2 + \varepsilon)$ -moment condition (5.2) on finite windows, then for  $f \in \mathbb{B}(W_1)$*

$$n^{-1/2}(\mu_n^{\xi_1}(f) - \mathbb{E}\mu_n^{\xi_1}(f), \dots, \mu_n^{\xi_m}(f) - \mathbb{E}\mu_n^{\xi_m}(f)) \xrightarrow{d} (Z_1, \dots, Z_m), \quad (5.11)$$

where  $(Z_1, \dots, Z_m)$  is a multivariate normal random vector (with possibly zero components) having zero mean and covariance

$$\text{Cov}(Z_i, Z_j) = \sigma^2(\Upsilon(i), \Upsilon(j)) \int_{W_1} f(x)^2 dx \in (-\infty, \infty),$$

where the random matrix  $(\Upsilon(i), \Upsilon(j))_{1 \leq i \leq j \leq m}$  involves also  $\Upsilon_{(\mathbf{0}, z)}$  given in (5.10):

$$\begin{aligned} \sigma^2(\Upsilon(i), \Upsilon(j)) &:= \rho \mathbb{E}[\Upsilon_{\mathbf{0}}(i) \Upsilon_{\mathbf{0}}(j)] \\ &+ \int_{\mathbb{R}^d} (\mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1, i) \Upsilon_{(\mathbf{0}, z)}(2, j)] \rho^{(2)}(\mathbf{0}, z) - \rho^2 \mathbb{E}[\Upsilon_{\mathbf{0}}(i)] \mathbb{E}[\Upsilon_{\mathbf{0}}(j)]) dz \in (-\infty, \infty). \end{aligned} \quad (5.12)$$

The proof is in Section 7.5. Under the stronger assumption of stopping-set stabilization, it suffices to assume stabilization under Palm conditioning only of order two, i.e., we only require that (4.12) holds for  $p = 1, 2$ ; see Corollary B.3.

As will be evident from the proof in Section 7.5, Theorem 5.4 allows for possibly different exponents  $\nu_i$  when considering the variance lower bounds  $\text{Var} \mu_n^{\xi_i}(f) = \Omega(n^{\nu_i})$  for scores  $\xi_i$ ,  $i = 1, \dots, m$ . However such a variant would require that  $\xi_i$ ,  $i = 1, \dots, m$ , satisfy the  $p$ -moment condition (5.2) on finite windows for *all*  $p > 2$ . In particular, if  $\nu_i < 1$  for some  $i$  then this implies a vanishing limit in (5.6). As a consequence, the normal vector  $(Z_1, \dots, Z_m)$  might have null components. If  $\nu_i = 1$  for some  $i$ , then the above result establishes the non-degenerate multivariate central limit theorem under only a  $p$ -moment condition for some  $p > 2$ . Extending the proof approach, we may establish multivariate normal convergence for

$$n^{-\nu/2}(\mu_n^{\xi_1}(f) - \mathbb{E}\mu_n^{\xi_1}(f), \dots, \mu_n^{\xi_m}(f) - \mathbb{E}\mu_n^{\xi_m}(f))$$

for some  $\nu \in (0, \infty)$  provided that for all  $1 \leq i, j \leq m$ ,

$$\lim_{n \rightarrow \infty} n^{-\nu} \text{Cov}(\mu_n^{\xi_i}(f), \mu_n^{\xi_j}(f)) = \sigma_\nu(\xi_i, \xi_j; f) \in [0, \infty).$$

One could likewise establish multivariate normal convergence for

$$n^{-\nu/2}(\mu_n^\xi(f_1) - \mathbb{E}\mu_n^\xi(f_1), \dots, \mu_n^\xi(f_m) - \mathbb{E}\mu_n^\xi(f_m))$$

for some  $\nu \in (0, \infty)$  provided that for all  $1 \leq i, j \leq m$ ,

$$\lim_{n \rightarrow \infty} n^{-\nu} \text{Cov}(\mu_n^\xi(f_i), \mu_n^\xi(f_j)) = \sigma_\nu(\xi; f_i, f_j) \in [0, \infty).$$

### 5.3 Asymptotics for sums of scores on the infinite window

Unlike the previous section we now assume  $\xi$  is defined on infinite input i.e.,  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$ . This assumption may not always hold, in particular for statistics of some spin systems and interacting particle systems, but it does apply to certain interacting diffusions and to empirical random fields considered in Part III. While the measures  $(\mu_n^\xi)_{n \geq 1}$  are defined in terms of the score functions  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ , we now consider the measures induced by *sums of scores*  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  with respect to input on the infinite window, namely

$$\hat{\mu}_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}) \delta_{n^{-1/d}x}. \quad (5.13)$$

**Proposition 5.5** (Limit theory for marked point processes with scores on infinite windows). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  having summable exponential  $\mathbb{B}$ -mixing correlations, as in Definition 4.8 and let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a marking function on  $\tilde{\mathcal{P}}$ .*

(i) *Let  $\xi$  be fast BL-localizing on  $\tilde{\mathcal{P}}$  as in Definition 4.2 (i) and (iii), and assume it satisfies the  $p$ -moment condition*

$$\sup_{1 \leq q \leq p} \sup_{x_1, \dots, x_q \in \mathbb{R}^d} \mathbb{E}_{[x]_1^q}[\max(1, |\xi(\tilde{x}_i, \tilde{\mathcal{P}})|^p)] \leq M_p^\xi < \infty \quad (5.14)$$

for all  $p > 1$ . If  $f \in \mathbb{B}(W_1)$  satisfies  $\text{Var} \hat{\mu}_n^\xi(f) = \Omega(n^\nu)$  for some  $\nu > 0$ , then as  $n \rightarrow \infty$ ,

$$(\text{Var} \hat{\mu}_n^\xi(f))^{-1/2} \left( \hat{\mu}_n^\xi(f) - \mathbb{E} \hat{\mu}_n^\xi(f) \right) \xrightarrow{d} Z, \quad (5.15)$$

where  $Z$  denotes a standard normal random variable.

(ii) *If moreover  $\mathcal{P}$  is stationary and  $\xi$  is translation invariant then the analogs of expectation, variance asymptotics and multivariate central limit theorems (as in Proposition 5.3 and Theorem 5.4) hold for  $(\hat{\mu}_n^\xi)_{n \in \mathbb{N}}$  where  $\mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$ ,  $\mathbb{E}_0[\xi^2(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$  and  $\mathbb{E}_{0,x}[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})\xi(\tilde{x}, \tilde{\mathcal{P}})]$  can be used in place of  $\mathbb{E}[\Upsilon_0]$ ,  $\mathbb{E}[\Upsilon_0^2]$  and  $\mathbb{E}[\Upsilon_{(0,z)}(1)\Upsilon_{(0,z)}(2)]$ , respectively, under a  $p = (1 + \epsilon)$  or  $p = (2 + \epsilon)$ -moment condition (5.14).*

The proof of Proposition 5.5 is in Section 7.6.

### 5.4 Remarks

In this section, we provide some remarks on the results in Sections 5.1–5.3. First, we detail how the ‘minimal’ representations of the limits of the two moments and the correlations of the score function  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ , as  $n \rightarrow \infty$  in Proposition 5.3, available under BL-localization on finite windows, can be replaced by the respective characteristics of some marks considered on the entire space  $\mathbb{R}^d$ . Next, we present more detailed extensions and remarks regarding Theorems 5.2, 5.4 and Propositions 5.3 and 5.5.

(i) (Asymptotic distribution of marks via  $L^1$ -stabilization.) By assuming fast  $L^1$ -stabilization of  $\xi$  on finite windows (see Appendix C), which is stronger than fast BL-localization, one may obtain a

new *distribution* of marks  $\Upsilon(x)$ ,  $x \in \mathcal{P}$ , attached to points of the entire ground process  $\mathcal{P}$ , provided  $\xi$  satisfies a  $p = 1$  moment condition on finite windows. These new marks satisfy the distributional limit of the ground process marked in finite windows by the scores  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ :

$$\sum_{x \in \mathcal{P}_n} \delta_{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))} \xrightarrow[n \rightarrow \infty]{d} \sum_{x \in \mathcal{P}} \delta_{(x, \Upsilon(x))}.$$

It is important to note that these marks  $\Upsilon(x)$  are not obtained by applying some (existing or extended) score function  $\xi$  of the entire process. However, at least two Palm distributions of  $\Upsilon(x)$  correspond to the two kernels  $\Upsilon_x$  and  $\Upsilon_{(x_1, x_2)}$ , which are limits of the Palm kernels of  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  as  $n \rightarrow \infty$ , established using BL-localization. In particular, under respective moment conditions of  $\xi$  on finite windows, their moments can be used in Proposition 5.3:  $\mathbb{E}[\Upsilon_0] = \mathbb{E}_0[\Upsilon(0)]$ ,  $\mathbb{E}[\Upsilon_0^2] = \mathbb{E}_0[\Upsilon(0)^2]$ ,  $\mathbb{E}[\Upsilon_{(0,x)}(1)\Upsilon_{(0,x)}(2)] = \mathbb{E}_{0,x}[\Upsilon(0)\Upsilon(x)]$ ; see the precise assumptions and formulation in Lemma C.1 in Appendix C.

(ii) (Almost sure extension of  $\xi$  via stopping-set stabilization.) In the case that  $\xi$  satisfies fast stabilization with respect to stopping sets, as described in Definition 4.6(ii), the idea consists of extending the function  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  to the infinite domain  $\mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ , denoted by  $\xi_\infty$ , satisfying:

$$\xi_\infty(\tilde{x}, \tilde{\mu}) = \lim_{n \rightarrow \infty} \xi(\tilde{x}, \tilde{\mu}_n), \quad (5.16)$$

provided the limit exists for all  $\tilde{\mu}$  (at least for those admitted by  $\tilde{\mathcal{P}}$  with probability 1) and  $\tilde{x} \in \tilde{\mu}$ . This condition can be met under the additional condition that the stabilization radii are uniformly bounded, i.e., almost surely  $\limsup_{n \rightarrow \infty} R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n) < \infty$ . We develop these arguments further in Appendix B. In particular, under this assumption, the extended score function  $\xi_\infty$  is fast stabilizing (4.11) admitting  $R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}}) := \limsup_{n \rightarrow \infty} R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  as the radius of stabilization and inherits the moment conditions assumed for  $\xi$  in finite windows (see Lemma B.1). Consequently, under the respective moment conditions of  $\xi$  (on finite windows)  $\mathbb{E}_0[\xi_\infty(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$ ,  $\mathbb{E}_0[\xi_\infty^2(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$ , and  $\mathbb{E}_{0,x}[\xi_\infty(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})\xi_\infty(\tilde{x}, \tilde{\mathcal{P}})]$  can be used in place of  $\mathbb{E}[\Upsilon_0]$ ,  $\mathbb{E}[\Upsilon_0^2]$ , and  $\mathbb{E}[\Upsilon_{(0,z)}(1)\Upsilon_{(0,z)}(2)]$ , respectively, in Proposition 5.3.

(iii) (Score function  $\xi$  defined on the infinite window.) This is perhaps a most natural scenario where the initial score function is also well-defined on the entire input process  $\tilde{\mathcal{P}}$ . If  $\xi$  is fast stabilizing on the infinite window as in (4.11), then  $\limsup_{n \rightarrow \infty} R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n) < \infty$  for all  $\tilde{x} \in \tilde{\mathcal{P}}$  and  $\xi_\infty = \xi$  a.s. (see Lemma B.1); this assumption was adopted in [21]. If  $\xi$  is only fast BL-localizing on the infinite window as in (4.2) (this assumption is adopted in Section 5.3) then at least the Palm distributions of  $\xi$  in finite windows converge to it i.e.,  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[r \rightarrow \infty]{\text{BL}, [z]_1^p} [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}) \stackrel{d, [z]_1^p}{=} \Upsilon_{[z]_1^p}$ , where the equality in distribution holds under Palm distribution of  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  given  $[z]_1^p$  (see Lemma 7.4). Consequently, in both cases, under the respective moment conditions of  $\xi$  (at least on finite windows)  $\mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$ ,  $\mathbb{E}_0[\xi^2(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$ , and  $\mathbb{E}_{0,x}[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})\xi(\tilde{x}, \tilde{\mathcal{P}})]$  can be used in place of  $\mathbb{E}[\Upsilon_0]$ ,  $\mathbb{E}[\Upsilon_0^2]$ , and  $\mathbb{E}[\Upsilon_{(0,z)}(1)\Upsilon_{(0,z)}(2)]$ , respectively, in Proposition 5.3.

(iv) (BL cluster-localization) Proposition 5.3, Theorem 5.4, and Proposition 5.5 remain valid when the marking functions satisfy BL cluster-localization instead of BL-localization, in complete analogy with the two admissible choices in Theorem 5.2. More details are provided in Remarks 7.6 and 7.8.

(v) (Local statistics.) In the case of scores which have a bounded interaction range, including local  $U$ -statistics, one can prove limit theorems under weaker mixing correlation assumptions on  $\tilde{\mathcal{P}}$  as in [21]. This point is further elaborated in Corollary 6.8.

(vi) (Quantitative CLTs—Rates of normal convergence.)

(a) *Poisson input.* When the marked point process  $\tilde{\mathcal{P}}$  is a Poisson point process equipped with independent marks and when  $\xi$  is exponentially stabilizing in the classical sense, i.e., when the decay functions  $\varphi_p$  in Definition 4.6 (i) and (4.11) are exponential functions, and when  $\text{Var}\mu_n^\xi(f) = \Theta(n)$ , then in Theorem 3.3, Theorem 5.2, and Proposition 5.5, one obtains a rate of normal approximation in the Kolmogorov distance of  $O(n^{-1/2})$  as shown in Lachièze-Rey et al. [93]. When the scores are fast BL-localizing and satisfy a  $(4 + \varepsilon)$ -moment condition, then subject to the variance growth condition  $\text{Var}\mu_n^\xi(f) = \Theta(n)$ , Trauthwein et al. [149] obtain rates of normal approximation for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$  in the Kolmogorov and Wasserstein distances of order  $O(n^{-1/2})$ . This extends Lachièze-Rey et al. [93] which is restricted to the setting of exponential stabilization of scores on Poisson point processes with i.i.d. marks.

(b) *Non-Poisson input.* In the case of general non-Poisson  $\mathcal{P}$  with i.i.d. marks and when the variance  $\text{Var}\mu_n^\xi(f)$  is of volume order, i.e., when  $\text{Var}\mu_n^\xi(f) = \Theta(n)$ , then Fenzl [57] refines the cumulant techniques of [21] and uses the general results from Saulis and Statulevicius [133] (see also Döring et al. [45]) to deduce rates of normal convergence and also moderate deviations. In their recent paper, Cong and Xia [34] quantify the rate of normal approximation for scores on an independently marked ground point process whenever the ground point process satisfies  $\beta$ -mixing, which is slightly stronger than the fast mixing of Definition 3.4. The work of Austern and Orbanz [5] can also be used to deduce normal approximation in terms of mixing coefficients. Dinh et al. [41] have derived rates of convergence for linear statistics of stationary  $\alpha$ -determinantal point processes via a quantitative Marcinkiewicz theorem and Chen et al. [32] also deduce rates of convergence for ‘locally defined’ statistics using a coupling of the stationary point process with its Palm version. However it is unclear how to apply these latter results for sums of localizing or stabilizing statistics of marked point processes. CITE [74]

(vii) (Moment conditions.) The moment conditions in Definitions 3.2 and 5.1 hold for a range of input and scores. When the input is either determinantal, permanental, or the zero set of a Gaussian entire function we refer to Sections 2.1 and 2.2 of [21] for sufficient conditions insuring the bound (3.4) of Definition 3.2.

(viii) (Variance asymptotics.) The proof of variance and covariance asymptotics in (5.6) and (5.7) actually only requires BL-localization of  $\xi$  for  $p \in \{1, 2\}$  and the following fast BL-mixing correlations of order 2 uniformly on  $\tilde{\mathcal{P}}_n$ ,  $n \in \mathbb{N}$ . Specifically, the marks  $U_n(x_i) := \xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  must satisfy the bound:

$$\left| \mathbb{E}_{[x]_1^2} [h_1(\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n))h_2(\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n))] \rho^{(2)}([x]_1^2) - \mathbb{E}_{x_1} [h_1(\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n))] \mathbb{E}_{x_2} [h_2(\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n))] \rho^2 \right| \leq \omega_2(|x_1 - x_2|), \quad (5.17)$$

for all functions  $h_1, h_2 \in \text{BL}(\mathbb{R})$  and a fast-decreasing function  $\omega_2$ ; see (7.32) below. The summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$ , assumed in Proposition 5.3 for presentational simplicity, ensures the fulfillment of this condition, as per Theorem 4.9(i) and Corollary 4.10.

(ix) (Variance growth rates.) The variance lower bound assumption in Theorem 3.3, Theorem 5.2, and Proposition 5.5 is crucial. If  $\sigma^2(\xi) > 0$ , then  $\nu = 1$  in Theorem 5.2 and Proposition 5.5. If  $\sigma^2(\xi) = 0$ , then  $\text{Var} \mu_n^\xi(f)$  could depend on the properties of  $f$ ; see Nazarov and Sodin [110, Section 1.5]. Even in the simple case  $\xi \equiv 1$ , it is possible that for some  $f$ ,  $\text{Var} \mu_n^\xi(f) \rightarrow 0$  arbitrarily fast; see for example Gabrielli et al. [61], Nazarov and Sodin [110], Mastrilli et al. [107], Lachièze-Rey [91]. However for sets  $W \subset W_1$  of positive measure, one has  $\text{Var} \hat{\mu}_n^\xi(\mathbf{1}_W) = \Omega(n^\nu)$  with  $\nu \in \{1, \frac{d-1}{d}\}$  (see [88, (3.8)]) with explicit asymptotics in case of ‘nice’ sets  $W$ ; see [144, 88, 80, 21]. These provide the necessary variance lower bounds for Proposition 5.5 when  $f = \mathbf{1}_W$ ,  $W \subset W_1$  of positive measure. In the case of Poisson input  $\mathcal{P}$  and exponentially stabilizing score functions, Schulte and Trapp [138] provide conditions under which  $\sigma^2(\xi) > 0$ . They survey previous work which establishes lower bounds for  $\liminf_{n \rightarrow \infty} \text{Var} \mu_n^\xi(f)$  and introduce a condition implying the strict positivity thereof, and which amounts to checking that  $D_z \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n), z \in W_n$ , is positive on a sets with strictly positive probability, with  $D_z$  being the difference operator as defined in (6.1).

(x) (Covariance growth rates.) Observe that (5.7) provides covariance asymptotics only of volume-order scale. If  $\sigma^2(\xi) = 0$ , it is possible to have different scaling for covariance depending on the test functions. Again for indicator functions, one can derive surface-order covariance asymptotics for  $\hat{\mu}_n^\xi$ ; see for example [88, Theorem 3.1], [144, Theorem 1.3] and [80] and the discussions therein. We refer to [88] for more details concerning the possible variance and covariance asymptotics based on the properties of  $f$ .

(xi) (Binomial input.) We have assumed that  $\mathcal{P}_n$  is the restriction of a point process  $\mathcal{P}$  to the window  $W_n$ . This assumption facilitates establishing expectation and variance asymptotics, but it is not necessary when proving central limit theorems, provided that the family of marked point processes  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  satisfy  $\hat{\omega}$ -mixing correlations as in Definition 3.1(iv) and provided the scores are fast BL-localizing on  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  as in Definition 4.2. When  $\mathcal{P}_n$  is the realization of  $n$  i.i.d. uniform random variables on  $W_n$  then the  $k$ -point correlations functions satisfy mixing, though with an extra error of order  $O(1/n)$ , bringing extra error terms in the subsequent analysis. A thorough study of the limit theory for dependently marked binomial input constitutes a separate project.

## 6 Proofs of results on mixing correlations—Theorem 4.9

In this section we prove Theorem 4.9, our main result on asymptotic de-correlation of marked point processes. Section 6.1 introduces the Factorial Moment Expansion (FME) for marked point processes and Section 6.2 applies the FME to prove Theorem 4.9. These results go a long way towards showing the factorization of mixed moments as at (3.6).

## 6.1 Factorial moment expansions of marked point processes

In this subsection, we first provide conditions insuring the factorial moment expansion (FME) [17, 20] for the expectations of functions of point processes with marks in a Polish space. We then apply this general result to obtain the FME of a *bounded* function  $f$  of a  $p$ -tuple of score functions, each of which has a *bounded* radius of stabilization.

Equip  $\mathbb{R}^d$  with a total order  $\prec$  defined using the lexicographical ordering of the polar co-ordinates. A useful property of this order is that  $\{y : y \prec x\} \subset B_{|x|}(\mathbf{0})$  and thus such sets are bounded. For  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  and  $x \in \mathbb{R}^d$ , define the measure  $\tilde{\mu}_{|x}(\cdot) := \tilde{\mu}(\cdot \cap (\{y : y \prec x\} \times \mathbb{K}))$ . Note that since  $\tilde{\mu}$  is a locally finite measure and the ordering is defined via polar co-ordinates,  $\tilde{\mu}_{|x}$  is a finite measure for all  $x \in \mathbb{R}^d$ . Let  $o$  denote the null-measure i.e.,  $o(B) = 0$  for all Borel subsets  $B$  of  $\mathbb{R}^d \times \mathbb{K}$ . For a measurable function  $\psi : \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$ ,  $l \in \mathbb{N}$ , and  $[\tilde{y}]_1^l = [y, w]_1^l$ , with distinct locations  $[y]_1^l \in (\mathbb{R}^d)^l$  and marks  $[w]_1^l \in \mathbb{K}^l$ , we define the FME kernels [17, 20] as follows. For  $l \in \mathbb{N}$ ,

$$\begin{aligned} D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mu}) &= \sum_{i=0}^l (-1)^{l-i} \sum_{J \subset \binom{[l]}{i}} \psi(\tilde{\mu}_{|y_*} + \sum_{j \in J} \delta_{\tilde{y}_j}) \\ &= \sum_{J \subset [l]} (-1)^{l-|J|} \psi(\tilde{\mu}_{|y_*} + \sum_{j \in J} \delta_{\tilde{y}_j}), \end{aligned} \quad (6.1)$$

where  $\binom{[l]}{i}$  denotes the collection of all subsets of  $[l] := \{1, \dots, l\}$  with cardinality  $i$ ,  $y_* := \min\{y_1, \dots, y_l\}$  with the minimum taken with respect to the order  $\prec$ , and  $w_* = w_j$  with (unique)  $j \in \{1, \dots, l\}$  such that  $y_j = y_*$ . For  $l = 0$ , put  $D^0 \psi(\tilde{\mu}) := \psi(o)$ . Note that  $D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mu})$  is invariant with respect to any joint permutation of marked points in  $[\tilde{y}]_1^l$ . For  $y_l \prec y_{l-1} \prec \dots \prec y_1$  the functional  $D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mu})$  is equal to the  $l$ th order iterated difference operator (see [97, Chapter 18]):

$$D_{\tilde{y}_1}^1 \psi(\tilde{\mu}) = \psi(\tilde{\mu}_{|y_1} + \delta_{\tilde{y}_1}) - \psi(\tilde{\mu}_{|y_1}), \quad D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mu}) = D_{\tilde{y}_l}^1 (D_{[\tilde{y}]_1^{l-1}}^{l-1} \psi(\tilde{\mu})).$$

We say that  $\psi$  is  $\prec$ -continuous at  $\infty$  if for all  $\tilde{\mu} \in \mathcal{N}$  we have

$$\lim_{y \uparrow \infty} \psi(\tilde{\mu}_{|y}) = \psi(\tilde{\mu}).$$

We shall require the following expansion for the expectation of functionals of marked point processes. Recall from Section 2 that  $\mathbb{E}_{[\tilde{y}]_1^l}^!$  denotes expectations with respect to reduced Palm distributions of  $\tilde{\mathcal{P}}$  given points  $[y]_1^l$  and their marks  $[w]_1^l$ .

**Theorem 6.1** ([17, 20]; FME for functionals of marked point processes). *Let  $\tilde{\mathcal{P}}$  be a simple, marked point process on  $\mathbb{R}^d \times \mathbb{K}$  and let  $\psi : \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be  $\prec$ -continuous at  $\infty$ . Assume that for all  $l \in \mathbb{N}$*

$$\int_{(\mathbb{R}^d \times \mathbb{K})^l} \mathbb{E}_{[\tilde{y}]_1^l}^! [|D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mathcal{P}})|] \tilde{\alpha}^{(l)}(d[\tilde{y}]_1^l) < \infty \quad (6.2)$$

and

$$\frac{1}{l!} \int_{(\mathbb{R}^d \times \mathbb{K})^l} \mathbb{E}_{[\tilde{y}]_1^l}^! [D_{[\tilde{y}]_1^l}^l \psi(\tilde{\mathcal{P}})] \tilde{\alpha}^{(l)}(d[\tilde{y}]_1^l) \rightarrow 0 \text{ as } l \rightarrow \infty. \quad (6.3)$$

Then  $\mathbb{E}[\psi(\mathcal{P})]$  has the following factorial moment expansion (FME)

$$\mathbb{E}[\psi(\tilde{\mathcal{P}})] = \psi(o) + \sum_{l=1}^{\infty} \frac{1}{l!} \int_{(\mathbb{R}^d \times \mathbb{K})^l} D_{[\tilde{y}]_1^l}^l \psi(o) \tilde{\alpha}^{(l)}(d[\tilde{y}]_1^l). \quad (6.4)$$

The expansion (6.4) is established in [20, Theorem 3.1] for  $\mathbb{K}$  a locally compact, second countable Hausdorff space and even earlier for unmarked point processes in [17, Theorem 3.2]. The assumption on  $\mathbb{K}$  in [20] is restrictive whereas Theorem 6.1 allows  $\mathbb{K}$  to be Polish. Indeed, the extension follows as the proof of [20, Theorem 3.1] depends only on the crucial existence and use of Palm distributions for marked point processes in this more general setting; see Section 2. Hence, one can follow the proof of [20, Theorem 3.1] to obtain Theorem 6.1 and so we shall not provide the proof here.

Consider now a marking function  $\xi$  and a  $p$ -tuple of distinct points  $[x]_1^p \in (\mathbb{R}^d)^p$  with their marks  $[u]_1^p \in \mathbb{K}^p$ . Let  $\xi(\tilde{x}; \tilde{\mu})$  be a marking function with values in  $\mathbb{K}'$  (a Polish space), defined for  $\tilde{x} = (x, u) \in \tilde{\mu}$ , with  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ . The proof of Theorem 4.9 is based on the FME expansion for

$$\mathbb{E}_{[x]_1^p} [f(\xi((x_1, U_1), \tilde{\mathcal{P}}), \dots, \xi((x_p, U_p), \tilde{\mathcal{P}}))] \rho^{(p)}([x]_1^p), \quad (6.5)$$

where  $f \in \mathbb{B}((\mathbb{K}')^p)$ . Under  $\mathbb{P}_{[x]_1^p}$  the point process  $\tilde{\mathcal{P}}$  has atoms at fixed locations  $x_1, \dots, x_p$  with random marks  $U_1, \dots, U_p$ , which complicates the form of its factorial moment measures. We address this by considering these fixed locations and their marks as parameters of the following modified functional

$$\psi_f([\tilde{x}]_1^p; \tilde{\mu}) := f\left(\xi(\tilde{x}_1, \tilde{\mu} + \sum_{i=1}^p \delta_{\tilde{x}_i}), \dots, \xi(\tilde{x}_p, \tilde{\mu} + \sum_{i=1}^p \delta_{\tilde{x}_i})\right) \quad (6.6)$$

and to not count these marked points  $\tilde{x}_1, \dots, \tilde{x}_p$  in  $\tilde{\mathcal{P}}$ , i.e., to consider  $\tilde{\mathcal{P}}$  under the reduced Palm probabilities  $\mathbb{P}_{[\tilde{x}]_1^p}^!$ . Obviously

$$\mathbb{E}_{[\tilde{x}]_1^p} [f(\xi(\tilde{x}_1, \tilde{\mathcal{P}}), \dots, \xi(\tilde{x}_p, \tilde{\mathcal{P}}))] = \mathbb{E}_{[\tilde{x}]_1^p}^! [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}})].$$

To justify convergence of some upcoming series expansions, we shall need to assume that the ground process  $\mathcal{P}$  satisfies

$$\lim_{l \rightarrow \infty} \frac{w^l}{l!} \kappa_{k+l} = 0 \quad (6.7)$$

for certain values of  $k \in \mathbb{N}, w > 0$ .

Now, we state the following important consequence of Theorem 6.1. As usual, by  $\mu$  we denote the natural projection of  $\tilde{\mu}$  on  $\mathbb{R}^d$ . Let  $\theta_d := \pi^{d/2} / \Gamma(d/2 + 1)$  denote the volume of the unit ball in  $\mathbb{R}^d$ . We recall for fixed  $[x]_1^p \in (\mathbb{R}^d)^p$ , the Palm probability distribution  $\mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p)$  on  $\mathbb{K}^p$  of the random marks  $[U]_1^p = (U_1, \dots, U_p)$  respectively at the fixed locations  $[x]_1^p$ , is defined via the identity

$$\tilde{\alpha}^{(p)}(d[\tilde{x}]_1^p) = \mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p) d([x]_1^p),$$

where  $\tilde{\alpha}^{(p)}$  is the  $p$ th order factorial moment measure of  $\tilde{\mathcal{P}}$ . In the following expansion, the distribution  $\mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p)$  incorporates  $\xi$ .

Recall the notation for the union of balls  $B_r([x]_1^p) := \bigcup_{i=1}^p B_r(x_i) \subset \mathbb{R}^d$ , that is, the union of balls of radius  $r > 0$  centered at the points  $[x]_1^p$ .

**Lemma 6.2** (FME of a bounded function of a  $p$ -tuple of union-restricted score functions). *Consider a score function  $\xi$  defined on the process  $\tilde{\mathcal{P}} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  or, possibly only on finite windows  $\tilde{\mathcal{P}} \cap W_n$ . Assume  $\tilde{\mathcal{P}}$  satisfies (6.7) for  $k = p$  and  $w = 2p\theta_d r^d$  with for some  $r > 0$ . Then for any  $f \in \mathbb{B}((\mathbb{K}')^p)$ , the functional  $\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))$  at (6.6) admits for  $\alpha^{(p)}$  a.e.  $[x]_1^p \in (\mathbb{R}^d)^p$  the following FME*

$$\begin{aligned} & \mathbb{E}_{[x]_1^p} [f(\xi((x_1, U_1), \tilde{\mathcal{P}} \cap B_r([x]_1^p)), \dots, \xi((x_p, U_p), \tilde{\mathcal{P}} \cap B_r([x]_1^p)))] \rho^{(p)}([x]_1^p) \\ &= \int_{\mathbb{K}^p} \psi_f([\tilde{x}]_1^p; o) \mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p) \\ &+ \sum_{l=1}^{\infty} \frac{1}{l!} \int_{(B_r([x]_1^p) \times \mathbb{K})^l} D_{[\tilde{y}]_1^l}^l \psi_f([\tilde{x}]_1^p; o) \mathcal{M}_{[x]_1^p \sqcup [\tilde{y}]_1^l}^{(p+l)}(d([u]_1^p \sqcup [w]_1^l)) \rho^{(p+l)}([x]_1^p \sqcup [y]_1^l) d([y]_1^l) \end{aligned} \quad (6.8)$$

where  $[\tilde{x}]_1^p = [x, u]_1^p$ ,  $[\tilde{y}]_1^l = [y, w]_1^l$ .

*Proof.* Throughout we fix a function  $f$  with  $|f| \leq 1$ . The boundedness of  $f$  makes our proof simpler than the corresponding result in the unmarked case, which does not use this assumption; see [21, Lemma 3.2]. We prove Lemma 6.2 by first establishing the validity of the FME at (6.4) for  $\mathbb{E}_{[x]_1^p}^! [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))] \rho^{(p)}([x]_1^p)$ , with  $[\tilde{x}]_1^p = [x, u]_1^p$  fixed. We then use the resulting FME to derive (6.8).

Indeed, restricting  $\xi$  to  $\tilde{\mathcal{P}} \cap B_r([x]_1^p)$  implies  $\psi_f$  is  $\prec$ -continuous at  $\infty$ . Regarding the integral assumptions (6.2) and (6.3) (with  $\psi$  replaced by  $\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p)) \rho^{(p)}([x]_1^p)$ ,  $\mathbb{E}_{[\tilde{y}]_1^l}^!$  replaced by  $(\mathbb{E}_{[\tilde{x}]_1^p \sqcup [\tilde{y}]_1^l}^!)^l = \mathbb{E}_{[x]_1^p \sqcup [y]_1^l}^!$  and  $\tilde{\alpha}^{(l)}(d[\tilde{y}]_1^l)$  by  $\tilde{\alpha}_{[x]_1^p}^{(l)}(d[\tilde{y}]_1^l)$ ), we observe

$$\begin{aligned} \int_{(\mathbb{R} \times \mathbb{K})^l} \mathbb{E}_{[x]_1^p \sqcup [\tilde{y}]_1^l}^! [ |D_{[\tilde{y}]_1^l}^l \psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))| ] \tilde{\alpha}_{[x]_1^p}^{(l)}(d[\tilde{y}]_1^l) \rho^{(p)}([x]_1^p) &\leq \int_{(B_r([x]_1^p) \times \mathbb{K})^l} 2^l \tilde{\alpha}_{[x]_1^p}^{(l)}(d[\tilde{y}]_1^l) \rho^{(p)}([x]_1^p) \\ &= 2^l \int_{B_r([x]_1^p)^l} \rho^{(p+l)}([x]_1^p \sqcup [y]_1^l) d[y]_1^l \\ &\leq (2p\theta_d r^d)^l \kappa_{p+l}. \end{aligned} \quad (6.9)$$

The first inequality is justified since *the difference operator  $D_{[\tilde{y}]_1^l}^l$  vanishes* as soon as  $y_k \notin B_r([x]_1^p) = \cup_{i=1}^p B_r(x_i)$  for some  $k \in \{1, \dots, l\}$  (that is to say  $D_{[\tilde{y}]_1^l}^l \psi_f([\tilde{x}]_1^p; \tilde{\mu}) = 0$ ; this is a well known property of the iterated add-one-cost operator; see, mutatis mutandis, the detailed arguments given in the proof of [21, Lemma 3.2]) while for  $\{y_1, \dots, y_l\} \subset \cup_{i=1}^p B_r(x_i)$ ,  $D_{[\tilde{y}]_1^l}^l \psi_f([\tilde{x}]_1^p; \tilde{\mu}) \leq 2^l$  (using  $|\psi_f([\tilde{x}]_1^p; \tilde{\mu})| \leq 1$  for  $|f| \leq 1$  and (6.1)). The equality follows by noting that integration with respect to  $[w]_1^l$  in (2.6) yields

$$\int_{\mathbb{K}^l} \tilde{\alpha}_{[x]_1^p}^{(l)}(d[\tilde{y}]_1^l) \rho^{(p)}([x]_1^p) = \rho^{(p+l)}([x]_1^p \sqcup [y]_1^l)$$

Finally, the last inequality in (6.9) follows from the bound  $\rho^{(p+l)}([x]_1^p \sqcup [y]_1^l) \leq \kappa_{p+l}$  which holds uniformly in  $[x]_1^p$  and  $[y]_1^l$ , in view of Assumption 3.1.

The bound (6.9), together with the assumed finiteness of  $\kappa_{p+l}$  and the assumed validity of (6.7) for  $k = p$ ,  $w = 2p\theta_d r^d$ , justifies (6.2) and (6.3). Thus, the FME at (6.4) holds for  $\mathbb{E}_{[x]_1^p}^! [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))] \rho^{(p)}([x]_1^p)$ .

Now we establish (6.8). First, observe that

$$\begin{aligned} & \mathbb{E}_{[x]_1^p} [f(\xi((x_1, U_1), \tilde{\mathcal{P}} \cap B_r([x]_1^p)), \dots, \xi((x_p, U_p), \tilde{\mathcal{P}} \cap B_r([x]_1^p))) \rho^{(p)}([x]_1^p)] \\ &= \int_{\mathbb{K}^p} \mathbb{E}_{[x]_1^p}^l [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))] \mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p). \end{aligned} \quad (6.10)$$

The terms in (6.4) for  $\mathbb{E}_{[x]_1^p}^l [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))]$  are integrals with respect to the  $l$ th order factorial moment measures  $\tilde{\alpha}_{[x, u]_1^p}^{(l)}(d[y, w]_1^l)$  of  $\tilde{\mathcal{P}}$  under  $\mathbb{E}_{[x]_1^p}^l$ . The subsequent integration of these terms with respect to  $\mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p)$ , via (2.6), yields the integrals in (6.8) with respect to the measures  $\mathcal{M}_{[x]_1^p \sqcup [u]_1^l}^{(p+l)}$ . Indeed, the term-wise integration of the FME series of  $\mathbb{E}_{[x]_1^p}^l [\psi_f([\tilde{x}]_1^p; \tilde{\mathcal{P}} \cap B_r([x]_1^p))]$  with respect to  $\mathcal{M}_{[x]_1^p}^{(p)}(d[u]_1^p) \rho^{(p)}([x]_1^p)$  is justified by the uniform bound (6.9) with respect to the marks  $[u]_1^p$  and hence the ‘remaining term’ (6.3) of the FME expansion for (6.10) goes to zero by the assumption (6.7). In this manner, we justify (6.8) and complete the proof of Lemma 6.2.  $\square$

## 6.2 Proof of mixing correlation properties of point processes with iterated marks

Our goal in this subsection is to prove Theorem 4.9. Recall that this theorem states that summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$  guarantees in (i) fast BL-mixing (resp., in (iii),  $\mathbb{B}$ -mixing) correlations of  $\tilde{\tilde{\mathcal{P}}} = \{(x, \xi(x, \tilde{\mathcal{P}}))\}_{x \in \mathcal{P}}$  provided the functionally constructed marking function  $\xi$  satisfies fast BL-localization or BL cluster-localization (resp. stopping-set stabilization). The proof has three main steps:

1. First, we establish in Lemma 6.3 a mixing correlation bound (6.12) for union-restricted score functions when  $\tilde{\mathcal{P}}$  has  $\mathbb{B}$ -mixing correlations. For this, we use the FME at (6.8) to prove the mixing correlation bound. As a consequence, assuming that  $\xi$  is localized at the outset, i.e.  $\xi = \xi^{(r)}$  for some  $r \in (0, \infty)$ , and that  $\tilde{\mathcal{P}}$  has  $\mathbb{B}$ -mixing correlations, we show that the marked point process  $\tilde{\tilde{\mathcal{P}}}$  inherits  $\mathbb{B}$ -mixing correlations of the same type. This is established in Proposition 6.4; see (6.13) below. In both statements, the decay constants and the decay function are given explicitly.
2. When the point process  $\tilde{\mathcal{P}}$  exhibits summable exponential  $\mathbb{B}$ -mixing correlations, we demonstrate that  $\tilde{\tilde{\mathcal{P}}}$  also has fast  $\mathbb{B}$ -mixing correlations. The assumption that  $\xi = \xi^{(r)}$  for some  $r \in (0, \infty)$  is relaxed to a fast stopping-set stabilization assumption. To achieve this, we leverage (6.13) in combination with a truncation approach to exploit these stabilization properties of  $\xi$ . This constitutes the proof of part (iii) of Theorem 4.9.
3. If furthermore fast stabilization of  $\xi$  is relaxed to fast BL-localization or fast BL cluster-localization, then  $\tilde{\tilde{\mathcal{P}}}$  exhibits only fast BL-mixing correlations. This establishes the proof of parts (i) and (ii) of Theorem 4.9, where bound (6.12) is crucial for Item (ii).

### 6.2.1 Proof of Theorem 4.9 for restricted score functions

We recall the notation for the *union-restriction operation* introduced in (4.5). Given a marking function  $\xi$  and points  $x_1, \dots, x_p \in \mathbb{R}^d$ , we define

$$[\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) := (\xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r([x]_1^p)), \dots, \xi(\tilde{x}_p, \tilde{\mathcal{P}} \cap B_r([x]_1^p))),$$

where  $B_r([x]_1^p) := \bigcup_{i=1}^p B_r(x_i)$  denotes the union of balls of radius  $r > 0$  centered at the points  $[x]_1^p$ .

**Lemma 6.3** (A mixing correlation bound for union-restricted scores.). *Fix  $p, q \in \mathbb{N}$  and  $r \in (0, \infty)$ . Assume that we are given:*

(i) *a marked point process  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$  having  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii) with correlation decay functions  $\omega_k(s) = C_k \phi(s)$ ,  $k \in \mathbb{N}$ , where  $(C_k)_{k \geq 1}$  is a non-decreasing sequence such that for some fixed  $\epsilon \in (0, 1)$ , we have*

$$\sum_l \frac{(w_0(1+\epsilon)\kappa_0^{1+\epsilon})^l}{l!} C_{l+p+q} < \infty, \quad (6.11)$$

with  $w_0 = 4(p+q)\theta_d r^d$  and where  $\kappa_0 = \max\{1, \sup_{x \in \mathbb{R}^d} \rho^{(1)}(x)\}$  is related to the ground process  $\mathcal{P}$ .

(ii) *a marking function  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$ .*

*Then for all  $f \in \mathbb{B}((\mathbb{K}')^p)$ ,  $g \in \mathbb{B}((\mathbb{K}')^q)$  and for almost all  $[x]_1^{p+q} \in \mathbb{R}^{d(p+q)}$  such that  $s := d([x]_1^p, [x]_{p+1}^{p+q}) > 4r$  we have*

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^{p+q}} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p+q)}([x]_1^{p+q}) \right. \\ & \quad \left. - \mathbb{E}_{[x]_1^p} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \\ & \leq \phi\left(\frac{s}{2}\right) \sum_{l=1}^{\infty} \frac{w_0^l}{l!} C_{l+p+q} =: \phi\left(\frac{s}{2}\right) \hat{C}_{p+q} < \infty. \end{aligned} \quad (6.12)$$

Before proving the above proposition, we derive a simple but useful consequence for marked point processes with restricted scores.

**Proposition 6.4** (Mixing correlations of marked point processes with restricted scores.). *Fix  $p, q \in \mathbb{N}$  and  $r \in (0, \infty)$ . Assume that  $\tilde{\mathcal{P}}$  satisfies condition (i) of Lemma 6.3. Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{K}'$  be a marking function such that  $\xi(\tilde{x}, \tilde{\mu}) = \xi^{(r)}(\tilde{x}, \tilde{\mu}) = \xi(\tilde{x}, \tilde{\mu} \cap B_r(x))$  for all  $(\tilde{x}, \tilde{\mu}) \in \mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ . Consider the marked point process  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, V(x))}$  on  $\mathbb{R}^d \times \mathbb{K}'$  where  $V(x) := \xi(\tilde{x}, \tilde{\mathcal{P}})$ .*

*Then, for all  $f \in \mathbb{B}((\mathbb{K}')^p)$ ,  $g \in \mathbb{B}((\mathbb{K}')^q)$ , for almost all  $[x]_1^{p+q} \in \mathbb{R}^{d(p+q)}$  such that  $s := d([x]_1^p, [x]_{p+1}^{p+q}) > 4r$ , and with  $w_0 = 4(p+q)\theta_d r^d$  we have*

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^{p+q}} [f([V]_1^p) g([V]_{p+1}^{p+q})] \rho^{(p+q)}([x]_1^{p+q}) - \mathbb{E}_{[x]_1^p} [f([V]_1^p)] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g([V]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}) \right| \\ & \leq \phi\left(\frac{s}{2}\right) \sum_{l=1}^{\infty} \frac{w_0^l}{l!} C_{l+p+q} =: \phi\left(\frac{s}{2}\right) \hat{C}_{p+q} < \infty. \end{aligned} \quad (6.13)$$

*Thus the marked point process  $\tilde{\mathcal{P}}$  satisfies the inequality (3.10) in Definition 3.6(ii) for  $s > 4r$ , with  $\omega_{p+q}(s) := \hat{C}_{p+q} \phi(s/2)$  and  $p, q, \phi$  assumed above.*

*Proof.* It is an easy observation that the union-restriction operation is neutral for a restricted score  $\xi = \xi^{(r)}$ , for some  $r > 0$ : for all  $x_1, \dots, x_p \in \mathbb{R}^d$ ,  $[\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) = [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) = [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$ . Therefore, the inequality (6.13) follows directly from (6.12), which completes the proof.  $\square$

We need another immediate extension of the bound (3.2) for the proof of Lemma 6.3.

**Corollary 6.5** (Extended version of  $\mathbb{B}$ -mixing correlations). *If the marked point process  $\tilde{\mathcal{P}}$  satisfies  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii) then the bound (3.10) also holds with the same  $\omega$  for all measurable functions  $f = f([(x, u)]_1^p)$ ,  $g = g([(x, u)]_{p+1}^{p+q})$  in  $\mathbb{B}(\mathbb{R}^d \times \mathbb{K})^p$  and  $\mathbb{B}(\mathbb{R}^d \times \mathbb{K})^q$ , respectively.*

*Proof of Lemma 6.3.* We will need to use the FME expansion of Lemma 6.2. Thus we need to first verify that our assumptions, together with the bound (3.8) for  $\kappa_k$  and the non-decreasing sequence  $(C_k)_{k \geq 1}$ , imply that  $\mathcal{P}$  satisfies (6.7) for all  $k \in \{p, q, p+q\}$ ,  $w = 2k\theta_d r^d$ . This is a straightforward exercise which we leave to the reader.

Now we are ready to apply Lemma 6.2. Again assume that  $r > 1$  without loss of generality. Consider a  $(p+q)$ -tuple  $[\tilde{x}]_1^{p+q} = [x, u]_1^{p+q}$  of distinct points  $[x]_1^{p+q} \in \mathbb{R}^{d(p+q)}$  with marks  $[u]_1^{p+q} \in \mathbb{K}^{p+q}$ . For given functions  $f, g$  on  $\mathbb{K}^p$  and  $\mathbb{K}^q$ , respectively, define functionals of  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ :

$$\psi_f([\tilde{x}]_1^p; \tilde{\mu}) := f([v]_1^p), \quad (6.14)$$

$$\psi_g([\tilde{x}]_p^{p+q}; \tilde{\mu}) := g([v]_{p+1}^{p+q}), \quad (6.15)$$

$$\psi_{fg}([\tilde{x}]_1^{p+q}; \tilde{\mu}) := f([v]_1^p)g([v]_{p+1}^{p+q}), \quad (6.16)$$

where  $v_i := \xi(\tilde{x}_i; (\tilde{\mu} + \sum_{\tilde{x} \in [\tilde{x}]_a^b} \delta_{\tilde{x}}) \cap B_r([x]_a^b))$ , with  $a = 1, b = p$  in (6.14),  $a = p+1, b = p+q$  in (6.15) and  $a = 1, b = p+q$  in (6.16). Note that the assumptions of Lemma 6.2 are satisfied. Hence we can represent each of the above three functionals as an infinite sum of difference operators, as in the FME representation (6.8).

Let  $[x]_1^{p+q}$  be such that  $s := d([x]_1^p, [x]_{p+1}^{p+q}) > 4r$ . We apply the FME (6.8) to  $\psi_{fg}([\tilde{x}]_1^{p+q}; \tilde{\mu})$ . Thus using the definition of the FME kernels (6.1) and that the FME kernels  $D^l$  vanish outside  $\bigcup_{i=1}^{p+q} B_r(x_i)$ , as noted in the proof of Lemma 6.2, we may crucially replace the integration domain  $(\mathbb{R}^d)^l$  in (6.8), by a product of  $B_r([x]_1^p)$  and  $B_r([x]_{p+1}^{p+q})$ .

From the above vanishing property of difference operators, we obtain the following factorization of  $\psi_{fg}$ . Given  $[\tilde{y}]_1^j \subset B_r([x]_1^p)$  and  $[\tilde{y}]_{j+1}^l \subset B_r([x]_{p+1}^{p+q})$ , together with sets  $J_1 \subset [j]$  and  $J_2 \subset [l] \setminus [j]$ , we have the factorizations

$$\psi_{fg}([\tilde{x}]_1^{p+q}; \sum_{i \in J_1 \cup J_2} \delta_{\tilde{y}_i}) = \psi_f([\tilde{x}]_1^p; \sum_{i \in J_1} \delta_{\tilde{y}_i}) \psi_g([\tilde{x}]_{p+1}^{p+q}; \sum_{i \in J_2} \delta_{\tilde{y}_i})$$

and

$$\sum_{J_1 \subset [j], J_2 \subset [l] \setminus [j]} (-1)^{l-|J_1|-|J_2|} \psi_f(\dots) \psi_g(\dots) = \sum_{J_1 \subset [j]} (-1)^{j-|J_1|} \psi_f(\dots) \sum_{J_2 \subset [l] \setminus [j]} (-1)^{l-j-|J_2|} \psi_g(\dots).$$

Further, we use the definition of FME kernels  $D^l$  as in (6.1) to obtain

$$\mathbb{E}_{[\tilde{y}]_1^{p+q}} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p+q)}([x]_1^{p+q})$$

$$\begin{aligned}
&= \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{1}{j!(l-j)!} \int_{(B_r([x]_1^p))^j \times (B_r([x]_{p+1}^{p+q}))^{l-j}} \int_{\mathbb{K}^{p+q+l}} \\
&\quad \times \sum_{J_1 \subset [j]} (-1)^{j-|J_1|} \psi_f([\tilde{x}]_1^p; \sum_{i \in J_1} \delta_{\tilde{y}_i}) \\
&\quad \times \sum_{J_2 \subset [l] \setminus [j]} (-1)^{l-j-|J_2|} \psi_g([\tilde{x}]_{p+1}^{p+q}; \sum_{i \in J_2} \delta_{\tilde{y}_i}) \\
&\quad \times \mathcal{M}_{[x]_1^{p+q} \sqcup [y]_1^l}^{(p+q+l)} (d([u]_1^{p+q} \sqcup [w]_1^l)) \rho^{(p+q+l)}([x]_1^{p+q} \sqcup [y]_1^l) d[y]_1^l, \tag{6.17}
\end{aligned}$$

where  $\tilde{y}_i = (y_i, w_i)$  and where we emphasize that the marks depend also only on  $[x]_1^{p+q} \sqcup [y]_1^l$ . Similarly,

$$\begin{aligned}
&\mathbb{E}_{[x]_1^p} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \mathbb{E}_{[x]_{p+1}^{p+q}} [g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \\
&= \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{1}{j!(l-j)!} \int_{(B_r([x]_1^p))^j \times (B_r([x]_{p+1}^{p+q}))^{l-j}} \int_{\mathbb{K}^{p+j}} \int_{\mathbb{K}^{q+l-j}} \\
&\quad \times \sum_{J_1 \subset [j]} (-1)^{j-|J_1|} \psi_f([\tilde{x}]_1^p; \sum_{i \in J_1} \delta_{\tilde{y}_i}) \\
&\quad \times \sum_{J_2 \subset [l] \setminus [j]} (-1)^{l-j-|J_2|} \psi_g([\tilde{x}]_{p+1}^{p+q}; \sum_{i \in J_2} \delta_{\tilde{y}_i}) \\
&\quad \times \mathcal{M}_{[x]_1^p \sqcup [y]_1^j}^{(p+j)} (d([u]_1^p, [w]_1^j)) \rho^{(p+j)}([x]_1^p, [y]_1^j) \\
&\quad \times \mathcal{M}_{[x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l}^{(q+l-j)} (d([u]_{p+1}^{p+q} \sqcup [w]_{j+1}^l)) \rho^{(q+l-j)}([x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l) d[y]_1^j d[y]_{j+1}^l. \tag{6.18}
\end{aligned}$$

These two derivations are the analogs of [21, (3.20) and the display preceding (3.23)], which were derived for point processes which did not carry dependent marks.

We bound the difference between the FME expressions (6.17) and (6.18) as follows. We match like terms in the two expressions and bound their difference termwise using the decorrelation (3.2) of the underlying marked point process  $\tilde{\mathcal{P}}$ , observing that

$$d([x]_1^p \sqcup [y]_1^j), [x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l \geq d([x]_1^p, [x]_{p+1}^{p+q}) - 2r \geq d([x]_1^p, [x]_{p+1}^{p+q}) - s/2 \geq s/2.$$

The details go as follows. For sets  $J_1, J_2$  in (6.18) consider functions

$$\begin{aligned}
\hat{f} &= \hat{f}_{J_1} = \hat{f}([(x, u)]_1^p \sqcup [(y, w)]_{J_1}) := \psi_f([\tilde{x}]_1^p; \sum_{i \in J_1} \delta_{\tilde{y}_i}) \\
\hat{g} &= \hat{g}_{J_2} = \hat{g}([(x, u)]_{p+1}^{p+q} \sqcup [(y, w)]_{J_2}) := \psi_g([\tilde{x}]_{p+1}^{p+q}; \sum_{i \in J_2} \delta_{\tilde{y}_i}),
\end{aligned}$$

with  $\hat{f}$  defined on  $(\mathbb{R}^d \times \mathbb{K})^{p+|J_1|}$  and evaluated for marks  $[u]_1^p, [w]_{J_1} := (w_i : i \in J_1)$  and points  $[x]_1^p, [y]_{J_1} := (y_i : i \in J_1)$ , respectively, and similarly for  $\hat{g}$ . In the following derivation, the second inequality is due to (3.2) (in association with Corollary 6.5) termwise for  $d([x]_1^p \sqcup [y]_1^j, [x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l) \geq s/2$ . We have

$$\left| \mathbb{E}_{[x]_1^{p+q}} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p+q)}([x]_1^{p+q}) \right.$$

$$\begin{aligned}
& - \mathbb{E}_{[x]_1^p} [f(\xi^{\cup(r)}_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [g(\xi^{\cup(r)}_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(q)}([x]_{p+1}^{p+q}) \Big| \\
& \leq \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{1}{j!(l-j)!} \int_{(B_r([x]_1^p))^j \times (B_r([x]_{p+1}^{p+q}))^{l-j}} \sum_{J_1 \subset [j]} \sum_{J_2 \subset [l] \setminus [j]} \\
& \times \left| \int_{\mathbb{K}^{p+q+l}} \hat{f}([(x, u)]_1^p \sqcup [(y, w)]_{J_1}) \hat{g}([(x, u)]_{p+1}^{p+q} \sqcup [(y, w)]_{J_2}) \right. \\
& \quad \times \mathcal{M}_{[x]_1^{p+q} \sqcup [y]_1^l}^{(p+q+l)}(\mathrm{d}([u]_1^{p+q} \sqcup [w]_1^l)) \rho^{(p+q+l)}([x]_1^{p+q} \sqcup [y]_1^l) \\
& \quad - \int_{\mathbb{K}^{p+j}} \hat{f}([(x, u)]_1^p \sqcup [(y, w)]_{J_1}) \mathcal{M}_{[x]_1^p \sqcup [y]_1^j}^{(p+j)}(\mathrm{d}([u]_1^p \sqcup [w]_1^j)) \rho^{(p+j)}([x]_1^p \sqcup [y]_1^j) \\
& \quad \times \int_{\mathbb{K}^{q+l-j}} \hat{g}([(x, u)]_{p+1}^{p+q} \sqcup [(y, w)]_{J_2}) \mathcal{M}_{[x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l}^{(q+l-j)}(\mathrm{d}([u]_{p+1}^{p+q} \sqcup [w]_{j+1}^l)) \rho^{(q+l-j)}([x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l) \Big| \mathrm{d}[y]_1^l \\
& = \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{1}{j!(l-j)!} \int_{(B_r([x]_1^p))^j \times (B_r([x]_{p+1}^{p+q}))^{l-j}} \sum_{J_1 \subset [j]} \sum_{J_2 \subset [l] \setminus [j]} \\
& \times \left| \mathbb{E}_{[x]_1^{p+q} \sqcup [y]_1^l} \left[ \hat{f}([(x, u)]_1^p \sqcup [(y, w)]_{J_1}) \hat{g}([(x, u)]_{p+1}^{p+q} \sqcup [(y, w)]_{J_2}) \right] \rho^{(p+q+l)}([x]_1^{p+q} \sqcup [y]_1^l) \right. \\
& \quad - \mathbb{E}_{[x]_1^p \sqcup [y]_1^j} \left[ \hat{f}([(x, u)]_1^p \sqcup [(y, w)]_{J_1}) \right] \rho^{(p+j)}([x]_1^p \sqcup [y]_1^j) \\
& \quad \times \mathbb{E}_{[x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l} \left[ \hat{g}([(x, u)]_{p+1}^{p+q} \sqcup [(y, w)]_{J_2}) \right] \rho^{(q+l-j)}([x]_{p+1}^{p+q} \sqcup [y]_{j+1}^l) \Big| \mathrm{d}[y]_1^l \\
& \leq \phi\left(\frac{s}{2}\right) \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{C_{l+p+q}}{j!(l-j)!} \sum_{J_1 \subset [j]} \sum_{J_2 \subset [l] \setminus [j]} \int_{(B_r([x]_1^p))^j \times (B_r([x]_{p+1}^{p+q}))^{l-j}} \mathrm{d}[y]_1^l \\
& \leq \phi\left(\frac{s}{2}\right) \sum_{l=0}^{\infty} \sum_{j=0}^l \frac{C_{l+p+q}}{j!(l-j)!} (2(p+q)\theta_d r^d)^l \\
& = \phi\left(\frac{s}{2}\right) \sum_{l=0}^{\infty} \frac{(4(p+q)\theta_d r^d)^l}{l!} C_{l+p+q},
\end{aligned}$$

where the identity  $\sum_{j=0}^l \frac{1}{j!(l-j)!} = \frac{2^l}{l!}$  justifies the last equality. This establishes inequality (6.13) and concludes the proof of Lemma 6.3.  $\square$

In the next subsection we extend Proposition 6.4 to non-restricted scores, as stated in Theorem 4.9. First, we comment on the weaker assumptions of the former result and its independent interest when establishing limits for real-valued restricted scores.

**Remark 6.6** (Summable exponential  $\mathbb{B}$ -mixing of  $\tilde{\mathcal{P}}$  ensures fast mixing of correlations). The decay constants and decay function of the input process  $\tilde{\mathcal{P}}$  are specified in a minimal way in Lemma 6.3 and Propositions 6.4. However, in most applications of this result we shall assume that  $\tilde{\mathcal{P}}$  exhibits summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8. Indeed, it may be checked that the

summable condition (4.14) of this definition implies the following condition

$$\sum_k \frac{C_k w^k}{k!} < \infty, \quad \forall w > 0 \quad (6.19)$$

and hence (6.11), for the constants  $C_k$  in Lemma 6.3 and Propositions 6.4. In addition, the fast nature of this mixing of correlations is guaranteed by the assumption (4.15) on  $\phi$  in Definition 4.8. The assumption of summable exponential  $\mathbb{B}$ -mixing correlations will be further used to relax the restriction on the score function  $\xi^{(r)}$  in the next section.

For restricted real scores  $\xi = \xi^{(r)}$ —the subject of Proposition 6.4—the limit theory for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$  stated in Section 5.1 can already be established via this proposition under assumption (6.19) and solely fast  $\mathbb{B}$ -mixing correlations of the input process  $\tilde{\mathcal{P}}$  (without necessarily requiring summable exponential-type mixing). Moreover, if  $\xi^{(r)}$  take the form of local  $U$ -statistics, the same conclusions hold under merely fast  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$ , as explained below.

**Corollary 6.7** (Limit theory for restricted scores). *Consider a marked point process  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$  having fast  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii) and (iii), with correlation decay functions  $\omega_k(s) = C_k \phi(s)$ ,  $k \in \mathbb{N}$ , for  $\phi$  fast decreasing, and such that the constants  $(C_k)_{k \in \mathbb{N}}$  satisfy the constraints (6.19). Assume real-valued, restricted score function  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \xi^{(r)}(\tilde{x}, \tilde{\mathcal{P}}_n)$  for some  $r \in (0, \infty)$  and all  $n \in \mathbb{N}$ . Then the marked point process  $\tilde{\mathcal{P}} = \sum \delta_{(x, \xi^{(r)}(\tilde{x}_i, \tilde{\mathcal{P}}))}$  has fast  $\mathbb{B}$ -mixing correlations. Consequently, with  $\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d}x}$ , all statements in Theorems 5.2 and 5.4, as well as in Propositions 5.3 and 5.5, hold true under their respective moment, variance, and stationarity conditions.*

*Proof.* The result follows from Remark 6.6 and, in particular, from the fact that condition (6.19) implies (6.11) for all  $p, q \in \mathbb{N}$ . Consequently, by Proposition 6.4 and fast decreasing  $\phi$ , the process  $\tilde{\mathcal{P}}$  exhibits fast  $\mathbb{B}$ -mixing correlations.

The assumption of summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$ , required in all statements of the results mentioned in this corollary, serves solely to establish fast BL-mixing correlations of the iterated marks  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$ ,  $\boldsymbol{\xi}(\tilde{x}, \tilde{\mathcal{P}}_n)$ , and  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  via Theorem 4.9. In the case when these scores are restricted to balls of some radius  $r \in (0, \infty)$ , condition (6.19) together with fast  $\mathbb{B}$ -mixing correlations of the input process  $\tilde{\mathcal{P}}$  is sufficient, as explained above.  $\square$

**Corollary 6.8** (Limit theory for local  $U$ -statistics). *Consider a marked point process  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$  having fast  $\mathbb{B}$ -mixing correlations as in Definition 3.6(ii) and (iii). Let the score function be a local  $U$ -statistic of the form*

$$\xi(\tilde{x}, \tilde{\mathcal{P}}) = \xi_h(\tilde{x}, \tilde{\mathcal{P}}) := \frac{1}{k!} \sum_{[\tilde{x}]_1^{k-1} \in \tilde{\mathcal{P}}^{(k-1)}} h(\tilde{x}, \tilde{x}_1, \dots, \tilde{x}_{k-1}), \quad (6.20)$$

for some  $k \in \mathbb{N}$  and a symmetric, bounded function  $h : (\mathbb{R}^d \times \mathbb{K})^k \rightarrow \mathbb{R}$ , satisfying  $h(\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_{k-1}) = 0$  whenever either  $\max_{1 \leq i \leq k-1} |x_i - x_0| > r$  for some  $r > 0$ , or  $x_i = x_j$  for some  $i \neq j$ ,  $i, j \in$

$\{0, \dots, k-1\}$ . Then the marked point process  $\tilde{\mathcal{P}} := \sum_{x_i \in \mathcal{P}} \delta_{(\tilde{x}_i, \xi_h(\tilde{x}_i, \tilde{\mathcal{P}}))}$  has fast  $\mathbb{B}$ -mixing correlations. Consequently, with  $\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d}x}$ , all statements in Theorems 5.2 and 5.4, as well as in Propositions 5.3 and 5.5, hold under their respective moment, variance, and stationarity conditions.

*Proof.* This is a marked version of the local  $U$ -statistics model considered in [21]. As observed in that work, the FME expansion of  $U$ -statistics contains only a finite number of terms. Consequently, we do not need the convergence condition (6.11) for the FME, since the series in formulas (6.17) and (6.18) involve only finitely many terms, and the same holds for the series in (6.13). Hence, Proposition 6.4 applies without imposing any mixing condition on  $\tilde{\mathcal{P}}$  beyond fast  $\mathbb{B}$ -mixing.  $\square$

## 6.2.2 Proof of Theorem 4.9 for general score functions

*Proof of Theorem 4.9.* We first prove part (iii) and then adjust the proof to obtain parts (i) and (ii).

*Proof of part (iii) of Theorem 4.9.* For  $p, q \in \mathbb{N}$  consider again a  $(p+q)$ -tuple of distinct points  $[x]_1^{p+q} \in (\mathbb{R}^d)^{(p+q)}$  with marks  $[u]_1^{p+q} \in \mathbb{K}^{p+q}$ . We recall  $s := d([x]_1^p, [x]_{p+1}^{p+q})$ . Without loss of generality, we assume  $s \in (4, \infty)$ .  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations (Definition 4.8) with the exponents  $a \in [0, 1)$  in (4.14) and  $b \in (0, d)$  in (4.15). Set

$$r := r(s) := \left(\frac{s}{4}\right)^{b(1-a)/(2(p+q+d))}. \quad (6.21)$$

Since  $s \in (4, \infty)$  and  $p+q \geq 2$ , we easily have  $r \in (1, s/4)$ .

Define the restricted score

$$\xi^{(r)}(\tilde{x}, \tilde{\mu}) := \xi(\tilde{x}, \tilde{\mu} \cap B_r(x)), \quad \tilde{x} = (x, u) \in \tilde{\mu}. \quad (6.22)$$

For all  $f \in \mathbb{B}((\mathbb{K}')^p)$  and  $g \in \mathbb{B}((\mathbb{K}')^q)$ , we consider the variants  $\psi_f^r = f([v_r]_1^p)$ ,  $\psi_g^r = g([v_r]_{p+1}^{p+q})$ ,  $\psi_{gf}^r = f([v_r]_1^p)g([v_r]_{p+1}^{p+q})$  of the functionals  $\psi_f, \psi_g, \psi_{gf}$  defined in (6.14)–(6.16) with  $\xi$  replaced by  $\xi^{(r)}$  and  $v_{r,i} := \xi^{(r)}(\tilde{x}_i, \tilde{\mu})$ ,  $1 \leq i \leq p+q$ .

We set

$$A := \mathbb{E}_{[x]_1^p} [f([V]_1^p)] \rho^{(p)}([x]_1^p), \quad (6.23)$$

$$B := \mathbb{E}_{[x]_{p+1}^{p+q}} [g([V]_{p+1}^{p+q})] \rho^{(q)}([x]_{p+1}^{p+q}), \quad (6.24)$$

$$C := \mathbb{E}_{[x]_1^{p+q}} [f([V]_1^p)g([V]_{p+1}^{p+q})] \rho^{(p+q)}([x]_1^{p+q}), \quad (6.25)$$

and similarly  $\tilde{A}, \tilde{B}, \tilde{C}$  with  $[V]_1^{p+q}$  replaced by  $[V_r]_1^{p+q}$ , where  $V_{r,i} := \xi^{(r)}(\tilde{x}_i, \tilde{\mathcal{P}})$ ,  $i = 1, \dots, p+q$ .

To prove Theorem 4.9 (i), we need to show fast decay of  $|AB - C|$  as  $s \rightarrow \infty$ . Towards this, we observe first the following (general) inequalities :

$$|AB - C| \leq |\tilde{A}\tilde{B} - \tilde{C}| + |AB - \tilde{A}\tilde{B}| + |C - \tilde{C}|$$

and

$$|AB - \tilde{A}\tilde{B}| \leq |A(B - \tilde{B})| + |(A - \tilde{A})\tilde{B}| \leq (|A| + |\tilde{B}|)(|A - \tilde{A}| + |B - \tilde{B}|).$$

Thus, we obtain

$$|AB - C| \leq |\tilde{A}\tilde{B} - \tilde{C}| + (1 + |A| + |\tilde{B}|)(|A - \tilde{A}| + |B - \tilde{B}| + |C - \tilde{C}|). \quad (6.26)$$

In the following, we shall give bounds on the various expressions appearing in the right-hand side of (6.26). First, note that since  $|f| \leq 1$ ,  $|g| \leq 1$ , we have

$$(1 + |A| + |\tilde{B}|) \leq 1 + \kappa_p + \kappa_q < \infty, \quad (6.27)$$

where we recall  $\kappa_p := \sup_{[x]_1^p \in (\mathbb{R}^d)^p} \rho^{(p)}([x]_1^p) < \infty$ , by Assumption 3.1.

STEP 1: (Bound for  $|\tilde{A}\tilde{B} - \tilde{C}|$ ). The marking function  $\xi^{(r)}$  and  $\tilde{\mathcal{P}}$  together satisfy the assumptions of Proposition 6.4. Consequently, by (6.13) with  $r$  replaced by  $r(s)$  we have

$$\begin{aligned} |\tilde{A}\tilde{B} - \tilde{C}| &\leq \phi\left(\frac{s}{2}\right) \sum_{l=1}^{\infty} \frac{(4(p+q)\theta_d r(s)^d)^l}{l!} C_{l+p+q} \\ &\leq c_1 \exp\left(-\frac{s^b}{c_1}\right) c_2 \exp\left(c_3\left(\frac{s}{4}\right)^{\frac{b}{2}}\right) \end{aligned} \quad (6.28)$$

for some positive, finite constants  $c_1, c_2, c_3$ . Indeed, (4.15) implies  $\phi(s/2) \leq c_1 \exp(-s^b/c_1)$  and the bound (4.14) on  $C_{l+p+q}$ , by Stirling's formula, gives the exponential bound  $c_2 \exp(c_3(s/4)^{b/2})$  on the series in (6.28); for details, we refer the reader to the arguments in [21, display (3.26), page 872].

STEP 2: (Bounds for  $|A - \tilde{A}|$ ,  $|B - \tilde{B}|$ , and  $|C - \tilde{C}|$ ). We use that  $\xi$  satisfies classical stopping set stabilization as in Definition 4.6 with  $r$  replaced by  $r(s)$  and for  $p, q, p+q$ . Specifically, observe that if  $R^\xi(\tilde{x}_i, \tilde{\mathcal{P}}) \leq r(s)$  for all  $1, \dots, p$  then  $\xi(\tilde{x}_i, \tilde{\mathcal{P}}) = \xi^{(r)}(\tilde{x}_i, \tilde{\mathcal{P}})$  and  $A = \tilde{A}$ . Consequently,

$$\begin{aligned} |A - \tilde{A}| &\leq 2\kappa_p \mathbb{P}_{[x]_1^p} \left( \max_{i=1, \dots, p} R^\xi(\tilde{x}_i, \tilde{\mathcal{P}}) > r(s) \right) \\ &\leq 2p\kappa_p \varphi_p(r(s)), \end{aligned} \quad (6.29)$$

which, by (6.21) and the fast decreasing property of  $\varphi_p$ , is a fast decreasing function of  $s$ . The same arguments work for  $|B - \tilde{B}|$ , and  $|C - \tilde{C}|$ .

Collecting all inequalities (6.26), (6.27), (6.28), and (6.29), we observe that the doubly marked point process  $\tilde{\mathcal{P}}$  has fast  $\mathbb{B}$ -mixing correlations as in (3.10) (but for  $\tilde{\mathcal{P}}$ ) with the decay function  $\omega_{p+q}$  depending on  $\phi$ ,  $\{C_k\}_{k=1}^{\infty}$ ,  $\varphi_{p+q}$ , and  $\kappa_{p+q}$ . Here we also use the assumption that  $\max(\varphi_q, \varphi_p) \leq \varphi_{p+q}$  and  $\max(\kappa_q, \kappa_p) \leq \kappa_{p+q}$  for all  $p, q \in \mathbb{N}$ . This completes the proof of part (iii) of Theorem 4.9.

*Proof of part (i) of Theorem 4.9.* We follow the arguments of the previous part, now applied to the BL-localizing score function  $\xi$ , for which we again consider its restricted version  $\xi^{(r)}$  in (6.22) with  $r$  as at (6.21). Since we want to show that the marks  $\xi$  have BL-mixing correlations as in Definition 3.1(i) and (ii), we consider only BL functions  $f, g$  respectively on  $(\mathbb{K}')^p, (\mathbb{K}')^q$ . Given these functions we consider the expressions  $A, B, C$  as in (6.23)–(6.25) and their counterparts  $\tilde{A}, \tilde{B}, \tilde{C}$ . The general bounds yield (6.26). The bound (6.28) holds by Proposition 6.4 since we have assumed  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations. Indeed (6.13) holds for all bounded functions  $f, g$  and in particular for our BL functions  $f, g$ . Thus STEP 1 goes through verbatim. Only in STEP 2 do we

need to modify the proof, since  $\xi$  satisfies (the weaker) BL-localizing property. However, we may use this property to obtain

$$|A - \tilde{A}| \leq \kappa_p \left| \mathbb{E}_{[x]_1^p} [f([V]_1^p)] - \mathbb{E}_{[x]_1^p} [f([V_r]_1^p)] \right| \leq \kappa_p d_{\text{BL}, [x]_1^p}([V]_1^p, [V_r]_1^p) \leq \kappa_p \varphi_p(r(s)), \quad (6.30)$$

where the last inequality follows directly from BL-localization (4.2) of the marking function  $\xi$  for  $p$ . Using BL-localization (4.2) of the marking function  $\xi$  for  $q$  and  $p + q$ , the same argument extends to  $|B - \tilde{B}|$  and  $|C - \tilde{C}|$  respectively, thereby completing the proof, exactly as in part (iii). For  $|C - \tilde{C}|$ , we have also used the  $(\mathbb{K}')^{p+q} \ni v = (v_p, v_q) \mapsto f(v_p)g(v_q)$  is BL if  $f$  and  $g$  are.

*Proof of part (ii) of Theorem 4.9.* We follow the proof approach as for parts (i) and (iii) but now choose

$$\begin{aligned} \tilde{A} &= \mathbb{E}_{[x]_1^p} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p)}([x]_1^p), \\ \tilde{B} &= \mathbb{E}_{[x]_{p+1}^{p+q}} [g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(q)}([x]_{p+1}^{p+q}), \\ \tilde{C} &= \mathbb{E}_{[x]_1^{p+q}} [f([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))g([\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}}))] \rho^{(p+q)}([x]_1^{p+q}), \end{aligned}$$

Also in contrast to (6.21), we set  $r$  to be

$$r := r(s) := \left(\frac{s}{4}\right)^{\delta'_{p+q}}, \quad \delta'_{p+q} := \min\left\{\frac{b(1-a)}{2(p+q+d)}, \delta_{p+q}\right\}, \quad (6.31)$$

where  $\delta_{p+q} < 1$  denotes the exponent governing the separation of the two groups of points, namely  $r < s([x]_1^p, [x]_{p+1}^{p+q})^{\delta_{p+q}}$ , as required by the BL cluster-localization bound (4.7). Recalling that  $s > 4$  and  $b(1-a)/(2(p+q+d)) < 1$ , we again note that  $1 \leq r = \left(\frac{s}{4}\right)^{\delta'_{p+q}} \leq \left(\frac{s}{4}\right)^{\delta_{p+q}} < s/4$ .

First, (6.12) implies that  $|\tilde{A}\tilde{B} - \tilde{C}|$  satisfies the bound (6.28). By (4.6) and following verbatim the argument in (6.30), we have that  $|A - \tilde{A}|$  and  $|B - \tilde{B}|$  are bounded by  $\kappa_p \varphi_p(r(s))$  and  $\kappa_q \varphi_q(r(s))$ , respectively. Second, we derive that

$$\begin{aligned} |C - \tilde{C}| &\leq \kappa_{p+q} d_{\text{BL}, [x]_1^{p+q}} \left( [\xi]_1^{p+q}(\tilde{x}, \tilde{\mathcal{P}}), ([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}})) \right) \\ &\leq \kappa_{p+q} d_{\text{BL}, [x]_1^{p+q}} \left( [\xi]_1^{p+q}(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{\cup(r)}]_1^{p+q}(\tilde{x}, \tilde{\mathcal{P}}) \right) \\ &\quad + \kappa_{p+q} d_{\text{BL}, [x]_1^{p+q}} \left( [\xi^{\cup(r)}]_1^{p+q}(\tilde{x}, \tilde{\mathcal{P}}), ([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{\cup(r)}]_{p+1}^{p+q}(\tilde{x}, \tilde{\mathcal{P}})) \right) \\ &\leq 4\kappa_{p+q} \varphi_{p+q}(r(s)), \end{aligned}$$

where the second inequality follows from the triangle inequality and the third is due to (4.6) and (4.7), in conjunction with the bound  $r(s) = (s/4)^{\delta'_{p+q}} \leq s^{\delta_{p+q}}$ .

Thus, combining the above bounds for  $|A - \tilde{A}|$ ,  $|B - \tilde{B}|$ ,  $|C - \tilde{C}|$  and  $|\tilde{A}\tilde{B} - \tilde{C}|$  as in (6.26) and along with (6.27), we have the requisite fast decreasing bound on  $|AB - C|$ .  $\square$

*Proof of Corollary 4.10.* This proof closely follows that of Theorem 4.9 and makes use of the bound (3.10) describing summable exponential  $\mathbb{B}$  mixing correlations for a family of marked point processes with the family being  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$ . It suffices to note that the functions  $\phi$ ,  $\{\varphi_p\}_{p=1}^{\infty}$ , and the constants  $\{C_k\}_{k=1}^{\infty}$ ,  $\{\kappa_p\}_{p=1}^{\infty}$  related to  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  and  $\xi$  and appearing implicitly in the proof of Theorem 4.9 uniformly satisfy the definitions of the mixing correlations and localization, respectively, stabilization, on the sequence  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$ . Indeed, these functions and constants can be tracked from the bounds in (6.26)–(6.29).  $\square$

## 7 Proofs of the limit theorems

In this section we prove limit results for dependently marked point processes, as given in Sections 3.1 and 5, namely central limit theorems and expectation and variance asymptotics. In particular, the proof of the main central limit theorem (Theorem 3.3) uses the classical cumulant method for integrals  $\mu_n^\xi(f) = \int f d\mu_n^\xi$  of random (possibly signed, atomic) measures defined in (3.1). A key argument of this method is the approximate factorization (or ‘mixing’ in our terminology) of the correlation functions (3.6) of these measures. This method was previously used e.g. in [21, Theorem 1.13], where these measures are constructed via a deterministic score function  $\xi$  on non-marked point process  $\mathcal{P}$ . However, the cumulant method applies to general, purely atomic random signed measure satisfying moment conditions and exhibiting fast mixing correlations. We believe this result is of independent interest and present it in Section 7.1, recollecting some of the arguments of the proof of [21, Theorem 1.13]. In Section 7.2, we use this general central limit theorem to prove our main central limit theorem—Theorem 3.3. Next, in Section 7.3, we show that if scores  $\xi(x, \tilde{\mathcal{P}}_n)$  satisfy BL-localization then they have limits under Palm distributions as the domain  $W_n$  increases up to  $\mathbb{R}^d$ . These limits determine the precise expectation and variance asymptotics given in Theorem 5.3, whose proof is in Section 7.4. Sections 7.5 proves the multivariate central limit theorem (Theorem 5.4) and Section 7.6 proves limit theory for scores on infinite window - Proposition 5.5.

### 7.1 Cumulant method for asymptotic normality of purely atomic random signed measures

We outline the cumulant method for random measures, but do not provide full details of some standard derivations, which may be found in e.g. [21, Theorem 1.13] and [110].

Let  $\mathfrak{m} = \sum_{(x_i, \xi_i) \in \tilde{\mathcal{P}}} \xi_i \delta_{x_i}$  be a purely atomic random signed, (locally finite) measure on  $\mathbb{R}^d$ . In general, it can be represented by a marked point process  $\tilde{\mathcal{P}} = \{(x_i, \xi_i)\}$  with locations of atoms  $\mathcal{P} = \{x_i\} \subset \mathbb{R}^d$  and their real-valued marks  $\xi_i \in \mathbb{R}$ .

For all  $p \in \mathbb{N}$ , let the ground process  $\mathcal{P}$  admit bounded correlation functions  $\rho^{(p)}$  (densities of the factorial moment measures). Then define the (generalized) correlation functions of  $\mathfrak{m}$  for all  $k_1, \dots, k_p \in \mathbb{N}$

$$m^{[k]_1^p}([x]_1^p) := \mathbb{E}_{[x]_1^p} [\xi_1^{k_1} \dots \xi_p^{k_p}] \rho^{(p)}([x]_1^p), \quad (7.1)$$

where  $\xi_i$  are random marks of fixed, distinct locations of atoms  $x_1, \dots, x_p \in \mathbb{R}^d$  of  $\tilde{\mathcal{P}}_n$  under Palm probabilities  $\mathbb{E}_{[x]_1^p}$ , provided the expectations  $\mathbb{E}_{[x]_1^p} [\xi_1^{k_1} \dots \xi_p^{k_p}]$  are well defined. Let  $M^p[\langle g, \mathfrak{m} \rangle] := \mathbb{E}[(\int_{\mathbb{R}^d} g(x) \mathfrak{m}(dx))^p]$  be the moments of integrals with respect to  $\mathfrak{m}$  of bounded real-valued functions  $g$  with bounded support. The correlation functions allow us to write

$$\begin{aligned} M^p[\langle g, \mathfrak{m} \rangle] &= \sum_{\gamma \in \Pi[p]} \int_{\mathbb{R}^{|\gamma|}} \prod_{j=1}^{|\gamma|} g^{|\gamma(j)|}(x_j) m^{(|\gamma(1)|, \dots, |\gamma(|\gamma|)|)}(x_1, \dots, x_{|\gamma|}) dx_1 \dots dx_{|\gamma|} \\ &= \sum_{\gamma \in \Pi[p]} \langle \bigotimes_{i=1}^{|\gamma|} g^{|\gamma(i)|} m^\gamma, \lambda^{|\gamma|} \rangle. \end{aligned} \quad (7.2)$$

Here  $\Pi[p]$  is the set of all unordered partitions of the set  $\{1, \dots, p\}$ , and where for a partition  $\gamma \in \Pi[p]$ ,  $|\gamma|$  denotes the number of partition classes and for convenience, we order its elements  $\gamma = \{\gamma(1), \dots, \gamma(|\gamma|)\}$  by the number of their elements in  $\gamma(i)$ ;  $m^\gamma := m^{(|\gamma(1)|, \dots, |\gamma(|\gamma|)|)}$ ,  $\lambda^l$  denotes the Lebesgue measure on  $(\mathbb{R}^d)^l$ , and  $\otimes$  denotes the tensor product of functions.

Using the correlation functions  $m^{[k]_1^p}$  we define the *Ursell functions* or *truncated correlation functions*  $m_{\top}^{[k]_1^p}$  of  $\mathfrak{m}$  by

$$m_{\top}^{[k]_1^p}([x]_1^p) = \sum_{\gamma \in \Pi[p]} (-1)^{|\gamma|-1} (|\gamma| - 1)! \prod_{i=1}^{|\gamma|} m^{(k_j: j \in \gamma(i))}(x_j : j \in \gamma(i)), \quad (7.3)$$

provided  $m^{(k_j: j \in J)} < \infty$  for all  $J \subset [p] := \{1, \dots, p\}$ ; see [21, Section 4.4.1]. Ursell functions allow one to express the  $p$ -th order cumulants  $S^p$  of the integrals  $\langle g, \mathfrak{m} \rangle$  in terms of integrals of Ursell functions just as moments of  $\langle g, \mathfrak{m} \rangle$  are expressed as integrals with respect to correlation functions  $m^{[k]_1^p}$  as in (7.2). Indeed, for all  $p \in \mathbb{N}$  we have the relation

$$S^p[\langle g, \mathfrak{m} \rangle] = \sum_{\gamma \in \Pi[p]} \langle \bigotimes_{i=1}^{|\gamma|} g^{|\gamma(i)|} m_{\top}^{\gamma}, \lambda^{|\gamma|} \rangle, \quad (7.4)$$

provided  $m^{[k]_1^{p'}} < \infty$  for all  $k_i \in \mathbb{N}$ ,  $k_1 + \dots + k_{p'} \leq p$ ; see [21, Sections 4.3.2] for the cumulants for a random variable  $S^p[Y]$  and [21, (4.36)] for the expression of cumulants  $S^p[\langle g, \mathfrak{m} \rangle]$  in terms of Ursell functions.

Next, we observe that the property of *mixing of correlations functions*  $m^{[k]_1^p}$  is equivalent to a *diameter bound on the Ursell functions* of  $\mathfrak{m}$  of these measures.

**Lemma 7.1** (mixing of generalized correlation functions of  $\mathfrak{m}$  is equivalent to a diameter bound on Ursell functions of  $\mathfrak{m}$ ). *For a fixed collection of integers  $k_1, \dots, k_p \in \mathbb{N}$ ,  $p \geq 2$ , assume that for all  $J \subset [p]$ ,*

$$\sup_{(x_j)_{j \in J} \in \mathbb{R}^{d|J|}} m^{(k_j: j \in J)}(x_j : j \in J) < \infty. \quad (7.5)$$

*The following statements are equivalent:*

(i) *There exist fast decreasing functions  $\tilde{\phi} = \tilde{\phi}_{k_1, \dots, k_p}$  and constants  $\tilde{C} = C(k_1, \dots, k_p)$ , such that for any subset of integers  $J \subset [p]$ ,  $|J| \geq 2$ , a nonempty, proper subset  $I \subsetneq J$ , and all configurations  $\{x_j : j \in J\}$ ,  $x_j \in \mathbb{R}^d$ , of distinct points we have*

$$\left| m^{(k_j: j \in J)}(x_j : j \in J) - m^{(k_j: j \in I)}(x_j : j \in I) m^{(k_j: j \in J \setminus I)}(x_j : j \in J \setminus I) \right| \leq \tilde{C} \tilde{\phi}(s), \quad (7.6)$$

where  $s := d(\{x_j : j \in I\}, \{x_j : j \in J \setminus I\})$ .

(ii) *There exist fast decreasing functions  $\tilde{\phi}_{\top} = \tilde{\phi}_{\top}^{k_1, \dots, k_p}$  and constants  $\tilde{C}^{\top} = \tilde{C}^{\top}(k_1, \dots, k_p)$ , such that for any subset of integers  $J \subset [p]$ ,  $|J| \geq 2$  and all configurations  $x_j \in \mathbb{R}^d$ ,  $j \in J$  of distinct points we have*

$$|m_{\top}^{(k_j: j \in J)}(x_j : j \in J)| \leq \tilde{C}^{\top} \tilde{\phi}_{\top}(\text{diam}(x_j : j \in J)), \quad (7.7)$$

where  $\text{diam}(x_j : j \in J) := \max_{i, j \in J} |x_i - x_j|$ .

Moreover, the bound (7.7) for  $J = [p]$  implies

$$\sup_{x_1 \in \mathbb{R}^d} \int_{(\mathbb{R}^d)^{(p-1)}} |m_{\top}^{[k]_1^p}([x]_1^p)| dx_2 \cdots dx_p < \infty. \quad (7.8)$$

Further,  $\tilde{\phi}_{\top}$  and  $\tilde{C}^{\top}$  in (ii) depend only on  $\tilde{\phi}$ ,  $\tilde{C}$ , the dimension  $d$  and the bound in (7.5) and similarly the  $\tilde{\phi}$  and  $\tilde{C}$  in (i) depend only on  $\tilde{\phi}_{\top}$ ,  $\tilde{C}^{\top}$  and the bound in (7.5).

One may interpret (7.8) as *Brillinger mixing* for the random measure  $m$  and it can be shown to be equivalent to volume order growth of cumulants. As we shall see shortly in Theorem 7.2, this suffices to derive a central limit theorem. Brillinger mixing and its consequences have been studied for point processes (see [78, Section 3.5] and [14, Theorem 3.2]) and our results extend these to purely atomic random signed measures.

*Proof of Lemma 7.1.* Under the (point-wise) assumption  $m^{(k_j: j \in J)} < \infty$  for all  $J \subset [p] = \{1, \dots, p\}$  one establishes the following relation

$$\begin{aligned} m_{\top}^{(k_1, \dots, k_p)}(x_1, \dots, x_p) &= m^{(k_1, \dots, k_p)}(x_1, \dots, x_p) - m^{(k_j: j \in I)}(x_j : j \in I) m^{(k_j: j \in J \setminus I)}(x_j : j \in J \setminus I) \\ &\quad + \sum_{\substack{\gamma \in \Pi[p], |\gamma| > 1 \\ \gamma \text{ mixes } \{I, I^c\}}} \prod_{i=1}^{|\gamma|} m_{\top}^{(k_j: j \in \gamma(i))}(x_j : j \in \gamma(i)), \end{aligned} \quad (7.9)$$

where a partition  $\gamma$  of  $\Pi[p]$  is said to mix  $\{I, I^c := [p] \setminus I\}$  when there exists an element  $\gamma(i) \in \gamma$  such that  $\gamma(i) \cap I \neq \emptyset$  and  $\gamma(i) \cap I^c \neq \emptyset$ . The proof of the equivalence property follows now from the fact that there exists constant  $\tilde{c}_{|J|}^{\top}$  (depending on the dimension  $d$ ) such that for each configuration  $\{x_j : j \in J\}$ ,  $x_j \in \mathbb{R}^d$ , there exists a partition  $\{I, I^c\}$  of  $J$  such that  $d(\{x_j : j \in I\}, \{x_j : j \in I^c\}) \geq \tilde{c}_{|J|}^{\top} \text{diam}(x_1, \dots, x_p)$  and, trivially, for any partition  $\{I, I^c\}$  of  $J$ ,  $d(\{x_j : j \in I\}, \{x_j : j \in I^c\}) \leq \text{diam}(x_1, \dots, x_p)$ . Indeed, for the first term in the right-hand-side of (7.9) we have

$$|m^{(k_1, \dots, k_p)} - m^{(k_j: j \in I)} m^{(k_j: j \in I^c)}| \leq \tilde{C}(k_1, \dots, k_p) \tilde{\phi}_{k_1, \dots, k_p}(\tilde{c}_{|J|}^{\top} \text{diam}(x_1, \dots, x_p)).$$

Also, for each term of the sum  $\sum_{\gamma} \dots$  in (7.9) we consider  $\gamma(i) \in \gamma$  intersecting both  $I$  and  $I^c$  and develop a similar bound by the induction, and using the uniform bounds (7.5) for other  $\gamma(i') \in \gamma$ . This way,  $|\sum_{\gamma} \dots|$  can be bounded by a *finite* sum of expressions  $C_{\gamma} \tilde{\phi}_{\gamma}(\tilde{c}_{\gamma} \text{diam}(x_1, \dots, x_p))$ , with some constants  $\tilde{C}_{\gamma} < \infty$ ,  $\tilde{c}_{\gamma} > 0$  and fast decreasing functions  $\tilde{\phi}_{\gamma}$ . This justifies that (i) implies (ii). The proof of the converse again uses the expression (7.9) in a similar way (but not requiring the induction).

For the proof of (7.8), fix  $x_1 \in \mathbb{R}^d$  and partition  $(\mathbb{R}^d)^{p-1}$  into  $G_l$ ,  $l \geq 0$  defined as follows:

$$\begin{aligned} G_0 &:= \{(x_2, \dots, x_p) \in \mathbb{R}^{d(p-1)} : \text{diam}(x_1, \dots, x_p) \leq 1\} \\ G_l &:= \{(x_2, \dots, x_p) \in \mathbb{R}^{d(p-1)} : 2^{l-1} < \text{diam}(x_1, \dots, x_p) \leq 2^l\}, \quad l \geq 1. \end{aligned}$$

Now use estimate (7.7) to bound the integral on the left-hand side of (7.8) as

$$\sum_{l=0}^{\infty} \int_{G_l} |m_{\top}^{[k]_1^p}([x]_1^p)| dx_2 \cdots dx_p \leq \theta_d^{p-1} \tilde{C}^{\top} + \theta_d^{p-1} \tilde{C}^{\top} \sum_{l=1}^{\infty} 2^{dl(p-1)} \tilde{\phi}_{\top}(2^{l-1}) < \infty,$$

since  $\tilde{\phi}_{\top}$  is fast decreasing. □

**Theorem 7.2** (Central limit theorem for purely atomic random signed measures). *Let  $(\mathbf{m}_n)_{n \in \mathbb{N}}$  be a sequence of purely atomic random signed measures supported, respectively, by  $W_n = [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$ , and assume they admit correlation functions  $m^{[k]_1^p}([x]_1^p; n)$  defined as at (7.1). Assume there exist fast decreasing functions  $\tilde{\phi}_k$ , and constants  $M_k, \tilde{C}_k$ , such that for all  $p \geq 2, k_1, \dots, k_p \in \mathbb{N}, k_1 + \dots + k_p = k$*

$$\sup_{n \in \mathbb{N}} \sup_{[x]_1^p \in \mathbb{R}^{dp}} m^{[k]_1^p}([x]_1^p; n) \leq M_k < \infty \quad (7.10)$$

and statement (i) of Lemma 7.1 holds for the functions  $\tilde{\phi} = \tilde{\phi}_k$  and constants  $\tilde{C} = \tilde{C}_k$  (uniformly in  $n \in \mathbb{N}$ ). If the sequence  $g_n$  of real-valued, measurable functions on  $\mathbb{R}^d$ , uniformly bounded in  $n$ , satisfies  $\text{Var}[\langle g_n, \mathbf{m}_n \rangle] = \Omega(n^\nu)$  for some  $\nu > 0$ , then

$$(\text{Var}[\langle g_n, \mathbf{m}_n \rangle])^{-1/2} (\langle g_n, \mathbf{m}_n \rangle - \mathbb{E}[\langle g_n, \mathbf{m}_n \rangle]) \xrightarrow{d} Z$$

where  $Z$  denotes a standard normal random variable.

*Proof.* By the moment condition (7.10), the assumption of statement (i) in Lemma 7.1 implies the statement (ii) therein. Hence, by (7.4) and (7.8), we obtain the bounds on the cumulants of the integrals of order

$$S^p[\langle g_n, \mathbf{m}_n \rangle] \leq (1 + \|g_n\|_\infty)^p A_p \lambda(W_n) = (1 + \|g_n\|_\infty)^p A_p n,$$

for some constants  $A_p < \infty$  and where  $\|g\|_\infty$  denotes the sup-norm of the function  $g$ . In other words, cumulants have growth  $S^p[\langle g_n, \mathbf{m}_n \rangle] = O(n)$  for all  $p \geq 2$ . Now, from our variance lower bound assumption  $\text{Var}[\langle g_n, \mathbf{m}_n \rangle] = \Omega(n^\nu)$  we further obtain that

$$S^p [\text{Var}[\langle g_n, \mathbf{m}_n \rangle]^{-1/2} (\langle g_n, \mathbf{m}_n \rangle - \mathbb{E}[\langle g_n, \mathbf{m}_n \rangle])] = O(n^{1-p\nu/2}),$$

and hence the cumulants of order  $p \in (2/\nu, \infty)$  in the left-hand-side above go to zero when  $n \rightarrow \infty$ . This suffices to prove the central limit theorem by a classical result due to Marcinkiewicz (see e.g. [146, Lemma 3]) which says that vanishing of cumulants of high-enough order implies normal convergence.  $\square$

## 7.2 Proof of umbrella CLT in Section 3.1

We will use Theorem 7.2 to establish the central limit theorem stated in Theorem 3.3. In this regard, we will first establish the condition (i) of Lemma 7.1 relating to correlation functions  $m^{[k]_1^{p+q}}([x]_1^{p+q}; n)$  defined in (7.1) for the family of atomic measures  $\mathbf{m}_n := \sum_{x \in \mathcal{P}_n} \xi_{i,n} \delta_x$ . More specifically, we seek to show that there exist a family of fast decreasing functions  $(\tilde{\omega}_k)_{k \in \mathbb{N}}$ , which provide the following bounds for  $p, q, k_1, \dots, k_{p+q} \in \mathbb{N}$ :

$$\sup_{n \in \mathbb{N}} \sup_{[x]_1^{p+q} \in (W_n)^{p+q}} |m^{[k]_1^{p+q}}([x]_1^{p+q}; n) - m^{[k]_1^p}([x]_1^p; n) m^{[k]_{p+1}^{p+q}}([x]_{p+1}^{p+q}; n)| \leq \tilde{\omega}_K(s), \quad (7.11)$$

where  $K = k_1 + \dots + k_{p+q}$  and  $s = d([x]_1^p, [x]_{p+1}^{p+q})$  is the distance defined in (3.3). This bound will be shown to be a consequence of the fast BL-mixing correlations (3.2) (with  $\xi_{i,n}$  representing  $U_n(x_i)$  and

$f, g \in BL$ ) under a suitable moment condition. In other words, the existence of the function  $\tilde{\omega}_K(s)$  in (7.11) is guaranteed under finiteness of the following moments

$$\tilde{M}_{K,\epsilon,p+q} := \sup_{n \in \mathbb{N}} \sup_{r \leq p+q} \sup_{[x]_1^r \in (W_n)^r} \mathbb{E}_{[x]_1^r} [\max(1, |\xi_{1,n}|^{K(1+\epsilon)})] \quad (7.12)$$

for some  $\epsilon > 0$  where, we recall that  $\xi_{1,n}$  is the random mark of the fixed point  $x_1$  under the Palm distribution  $\mathbb{P}_{[x]_1^r}$ . The  $p$ -moment condition (3.4) for  $p = K(1 + \epsilon)$  implies finiteness of  $\tilde{M}_{K,\epsilon,p+q}$  as  $\tilde{M}_{K,\epsilon,p+q} \leq M_{K(1+\epsilon)}^\xi$  due to the fact that  $K \geq p + q$ , one can extend the domain from  $\sup_{r \leq p+q}$  to  $\sup_{r \leq K(1+\epsilon)}$  in (7.12).

**Lemma 7.3** (Sufficient conditions implying fast mixing of correlation functions of marked point processes). *Let  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}} = (\{(x_i, \xi_{i,n})\})_{n \in \mathbb{N}}$ ,  $\xi_{i,n} \in \mathbb{R}$ , be a family of  $\mathbb{R}$ -valued marked point processes sharing the same ground point process  $\mathcal{P}$  in  $\mathbb{R}^d$ . Fix  $p, q, k_1, \dots, k_{p+q} \in \mathbb{N}$  with  $K = k_1 + \dots + k_{p+q}$ . Assume that  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  satisfies the fast BL mixing correlation condition (3.2) for  $p, q$  with fast decay function  $\hat{\omega}_{p+q}$  (and replacing  $U(x_i)$  with  $\xi_{i,n}$  in (3.2)) and for all functions  $f \in BL(\mathbb{R}^p)$ ,  $g \in BL(\mathbb{R}^q)$ .*

*For some  $\epsilon > 0$  we assume finiteness of the moment  $\tilde{M}_{K,\epsilon,p+q}$  as in (7.12) associated to the marks  $\xi_{i,n}$ . Then there is a fast decreasing function  $\tilde{\omega}_K$  such that the correlation functions  $m^{[k]_1^{p+q}}([x]_1^{p+q}; n)$  of  $\tilde{\mathcal{P}}_n$  satisfy (7.11) uniformly in  $n \in \mathbb{N}$  and  $[k]_1^{p+q}$  given  $k_1 + \dots + k_{p+q} = K$ . The function  $\tilde{\omega}_K$  depends on the value of  $K$ , the function  $\hat{\omega}_K$ , and the constants  $M_{K+\epsilon}^\xi$ ,  $C_K$ ,  $\kappa_0$  related to  $\xi$  and  $\tilde{\mathcal{P}}$ .*

In applications we shall take  $\xi_{i,n} = \xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  but here we state Lemma 7.3 in its fullest generality.

*Proof.* For  $i \in \{1, \dots, p+q\}$  define the functions  $g_i(u_i) := u_i^{k_i}$  on  $\mathbb{R}$ . Consider the functions  $h_1([u]_1^p) = \prod_{i=1}^p g_i(u_i)$  and  $h_2([u]_{p+1}^{p+q}) = \prod_{i=p+1}^{p+q} g_i(u_i)$ , respectively, on  $\mathbb{R}^p$  and  $\mathbb{R}^q$ . By the definition (3.6) of correlation functions  $m^{[k]_1^{p+q}}$  the left-hand-side of (7.11) corresponds to

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^{p+q}} [h_1(\xi_{1,n}, \dots, \xi_{p,n}) h_2(\xi_{p+1,n}, \dots, \xi_{p+q,n})] \right. \\ & \left. - \mathbb{E}_{[x]_1^p} [h_1(\xi_{1,n}, \dots, \xi_{p,n})] \rho^{(p)}([x]_1^p) \mathbb{E}_{[x]_{p+1}^{p+q}} [h_2(\xi_{p+1,n}, \dots, \xi_{p+q,n})] \rho^{(q)}([x]_{p+1}^{p+q}) \right|. \end{aligned} \quad (7.13)$$

To use (3.2) with  $U(x_i) = \xi_{i,n}$  we have to truncate  $h_1, h_2$  in a Lipschitz way to fulfill the role of  $f, g$  in BL-mixing condition assumed uniformly for all  $\tilde{\mathcal{P}}_n$ ,  $n \in \mathbb{N}$ . For this purpose we choose  $\epsilon > 0$  as in (7.12), define the function  $a(s) = a_{p+q}(s) := \max\{\hat{\omega}_{p+q}(s)^{-(1+\epsilon)/(2+3\epsilon)}, 1\}$ , and truncate the functions  $g_i$ ,  $1 \leq i \leq p$  as follows :

$$\bar{g}_i := \begin{cases} g_i & \text{when } |g_i| \leq a(s)^{1/p} \\ \text{sgn}(g_i) a(s)^{1/p} & \text{when } |g_i| > a(s)^{1/p}. \end{cases} \quad (7.14)$$

Similarly, we truncate the functions  $g_i$ ,  $p+1 \leq i \leq p+q$  as follows :

$$\bar{g}_i := \begin{cases} g_i & \text{when } |g_i| \leq a(s)^{1/q} \\ \text{sgn}(g_i) a(s)^{1/q} & \text{when } |g_i| > a(s)^{1/q}. \end{cases} \quad (7.15)$$

We assert that each  $\bar{g}_i$ ,  $1 \leq i \leq p$ , is Lipschitz function with Lipschitz constant  $k_i \cdot a(s)^{(k_i-1)/(pk_i)}$ . Indeed notice for  $|x|, |y| \leq a(s)^{1/(pk_i)}$  by the mean value theorem

$$|\bar{g}_i(x) - \bar{g}_i(y)| \leq |x - y| \cdot g'_i(a(s)^{1/(pk_i)}) = k_i \cdot a(s)^{(k_i-1)/(pk_i)} |x - y|.$$

On the other hand, if  $|x| \leq a(s)^{1/(pk_i)} \leq |y|$  then

$$|\bar{g}_i(x) - \bar{g}_i(y)| = |x^{k_i} - \text{sgn}(y)a(s)^{1/p}| \leq k_i \cdot a(s)^{(k_i-1)/(pk_i)} |x - y|,$$

which shows the assertion. Likewise, each  $\bar{g}_i$ ,  $p+1 \leq i \leq p+q$ , is Lipschitz function with Lipschitz constant  $k_i \cdot a(s)^{(k_i-1)/(qk_i)}$ .

Put

$$\bar{h}_1 = \prod_{i=1}^p \bar{g}_i(u_i), \quad \bar{h}_2 = \prod_{i=p+1}^{p+q} \bar{g}_i(u_i).$$

Observe,  $\bar{h}_1 \leq a(s)$  is bounded, Lipschitz function (with respect to the  $\ell^1$  norm on  $\mathbb{R}^p$ ) with Lipschitz constant bounded by  $a(s)^{(p-1)/p} \sum_{i=1}^p k_i \cdot a(s)^{(k_i-1)/(pk_i)} = a(s) \sum_{i=1}^p k_i \leq Ka(s)$ .<sup>2</sup> Similarly  $\bar{h}_2 \leq a(s)$ , and is a Lipschitz function with Lipschitz constant bounded by  $Ka(s)$ . Consequently  $\bar{h}_1/(Ka(s))$ ,  $\bar{h}_2/(Ka(s))$  are bounded Lipschitz functions utilisable to (3.2).

In the evaluation of the truncation error we simplify the notation by writing  $H_1 := h_1(\xi_{1,n}, \dots, \xi_{p,n})$ ,  $H_2 := h_2(\xi_{p+1,n}, \dots, \xi_{p+q,n})$  and similarly for  $\bar{H}_1 := \bar{h}_1(\dots)$  and  $\bar{H}_2 := \bar{h}_2(\dots)$ . With this notation, we write the first expectation in (7.13) as the sum of three terms

$$\begin{aligned} T_1 &:= \rho^{(p+q)}([x]_1^{p+q}) \mathbb{E}_{[x]_1^{p+q}} [H_1(H_2 - \bar{H}_2)], \\ T_2 &:= \rho^{(p+q)}([x]_1^{p+q}) \mathbb{E}_{[x]_1^{p+q}} [(H_1 - \bar{H}_1)\bar{H}_2], \\ T_3 &:= \rho^{(p+q)}([x]_1^{p+q}) \mathbb{E}_{[x]_1^{p+q}} [\bar{H}_1\bar{H}_2]. \end{aligned}$$

Write the second expectation in (7.13) as the sum of four terms:

$$\begin{aligned} T_4 &:= \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \mathbb{E}_{[x]_1^p} [H_1] \mathbb{E}_{[x]_{p+1}^{p+q}} [H_2 - \bar{H}_2], \\ T_5 &:= \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \mathbb{E}_{[x]_1^p} [H_1 - \bar{H}_1] \mathbb{E}_{[x]_{p+1}^{p+q}} [\bar{H}_2], \\ T_6 &:= \rho^{(p)}([x]_1^p) \rho^{(q)}([x]_{p+1}^{p+q}) \mathbb{E}_{[x]_1^p} [\bar{H}_1] \mathbb{E}_{[x]_{p+1}^{p+q}} [\bar{H}_2]. \end{aligned}$$

Applying (3.2) to the Lip(1) functions  $\bar{H}_1/(Ka(s))$  and  $\bar{H}_2/(Ka(s))$ , we find that the difference of terms  $T_3$  and  $T_6$  is bounded by

$$|T_3 - T_6| \leq K^2 a(s)^2 \hat{\omega}_{p+q}(s) \leq K^2 \hat{\omega}_{p+q}(s)^{\epsilon/(2+3\epsilon)},$$

by choice of  $a(s)$ . We now control the terms  $T_1, T_2, T_4, T_5$ , as follows. Put  $\widetilde{M}_{K,\epsilon} := \widetilde{M}_{K,\epsilon,p+q}$ . By the Hölder and the Markov inequalities and the bound on  $\kappa_p$  in (3.8) we have

$$|T_1| \leq \kappa_{p+q} \mathbb{E}_{[x]_1^{p+q}} [|H_1 H_2| \mathbf{1}(\exists i \in \{p+1, \dots, p+q\} : |g_i(\xi_{i,n})| > a(s)^{1/q})]$$

<sup>2</sup> Indeed, a  $p$ -fold product of Lipschitz functions on  $\mathbb{R}$  with respective Lip constants  $L_1, \dots, L_p$  and common sup norm bound  $S$  is a Lipschitz function on  $\mathbb{R}^p$  with Lipschitz constant equal to  $\sum_{i=1}^p L_i S^{p-1}$  with respect to the  $\ell^1$  norm on  $\mathbb{R}^p$ .

$$\begin{aligned}
&\leq \sum_{i=p+1}^{p+q} \kappa_{p+q} \mathbb{E}_{[z]_1^{p+q}} [|H_1 H_2|^{1+\epsilon}]^{1/(1+\epsilon)} a(s)^{-\epsilon/(q(1+\epsilon))} \mathbb{E}_{[z]_1^{p+q}} [|g_i(\xi_{i,n})|]^\epsilon / (1+\epsilon) \\
&\leq q \kappa_{p+q} \widetilde{M}_{K,\epsilon}^2 a(s)^{-\epsilon/(q(1+\epsilon))},
\end{aligned}$$

where we use the moment condition (7.12) twice (for  $\mathbb{E}_{[z]_1^{p+q}} [|H_1 H_2|^{1+\epsilon}]$  and  $\mathbb{E}_{[z]_1^{p+q}} [|g_i(\xi_{i,n})|]$ ) in the third inequality as follows. Specifically, we apply the Hölder inequality to obtain

$$\begin{aligned}
\mathbb{E}_{[z]_1^{p+q}} [|H_1 H_2|^{1+\epsilon}] &\leq \mathbb{E}_{[z]_1^{p+q}} \left[ \prod_{i=1}^{p+q} |\xi_{i,n}|^{(1+\epsilon)k_i} \right] \leq \prod_{i=1}^{p+q} \mathbb{E}_{[z]_1^{p+q}} \left[ \max(1, |\xi_{i,n}|)^{(1+\epsilon)K} \right]^{k_i/K} \\
&\leq \prod_{i=1}^{p+q} (\widetilde{M}_{K,\epsilon})^{k_i/K} = \widetilde{M}_{K,\epsilon}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}_{[z]_1^{p+q}} [|g_i(\xi_{i,n})|] &\leq \mathbb{E}_{[z]_1^{p+q}} [\max(1, |\xi_{i,n}^{k_i}|)] \leq \mathbb{E}_{[z]_1^{p+q}} [\max(1, |\xi_{i,n}|)^{(1+\epsilon)K}]^{k_i/(K(1+\epsilon))} \\
&= (\widetilde{M}_{K,\epsilon})^{k_i/(K(1+\epsilon))} \leq \widetilde{M}_{K,\epsilon}.
\end{aligned}$$

By the symmetry (interchanging the functions  $h_1$  and  $h_2$  as well as  $p$  and  $q$  and using the bound  $|\bar{H}_2| \leq |H_2|$ ) we also have

$$|T_2| \leq p \kappa_{p+q} \widetilde{M}_{K,\epsilon}^2 a(s)^{-\epsilon/(q(1+\epsilon))}.$$

Again using the Hölder and Markov inequalities as when bounding  $T_1$ , we obtain

$$\begin{aligned}
|T_4| &\leq \kappa_p \kappa_q \mathbb{E}_{[z]_1^p} [|H_1|] \mathbb{E}_{[z]_{p+1}^{p+q}} [|H_2| \mathbf{1}(\exists i \in \{p+1, \dots, p+q\} : |g_i(u_i)| > a(s)^{1/q})] \\
&\leq q \kappa_p \kappa_q \widetilde{M}_{K,\epsilon}^2 a(s)^{-\epsilon/(q(1+\epsilon))},
\end{aligned}$$

where we used

$$\begin{aligned}
\mathbb{E}_{[z]_1^p} [|H_1|] &= \mathbb{E}_{[z]_1^p} \left[ \prod_{i=1}^p |\xi_{i,n}|^{k_i} \right] \leq \mathbb{E}_{[z]_1^p} \left[ \prod_{i=1}^p |\max(1, |\xi_{i,n}|)|^{(1+\epsilon)k_i} \right] \\
&\leq \prod_{i=1}^p \mathbb{E}_{[z]_1^p} \left[ |\max(1, |\xi_{i,n}|)|^{(1+\epsilon)(k_1+\dots+k_p)} \right]^{\frac{k_i}{k_1+\dots+k_p}} \\
&\leq \prod_{i=1}^p \mathbb{E}_{[z]_1^{p+q}} \left[ |\max(1, |\xi_{i,n}|)|^{(1+\epsilon)K} \right]^{\frac{k_i}{k_1+\dots+k_p}} \\
&\leq \prod_{i=1}^p (\widetilde{M}_{K,\epsilon})^{\frac{k_i}{k_1+\dots+k_p}} = \widetilde{M}_{K,\epsilon},
\end{aligned}$$

and mutatis mutandis for  $\mathbb{E}_{[z]_{p+1}^{p+q}} [|H_2|^{(1+\epsilon)}] \leq \widetilde{M}_{K,\epsilon}$ . By symmetry,

$$|T_5| \leq p \kappa_p \kappa_q \widetilde{M}_{K,\epsilon}^2 a(s)^{-\epsilon/(p(1+\epsilon))}.$$

Observe by (3.8) that  $\kappa_p \kappa_q \leq (p+q)^2 C_{p+q}^2 \kappa_0^{p+q}$ . Collecting all estimates above we the bound of expression in (7.13) and using that  $a(s) \geq 1$  and  $a(s) \geq \hat{\omega}_{p+q}(s)^{-(1+\epsilon)/(2+3\epsilon)}$ , we derive

$$\begin{aligned} & |T_3 - T_6| + |T_1| + |T_2| + |T_4| + |T_5| \\ & \leq (K^2 + 4(p+q)^3 C_{p+q}^2 \kappa_0^{p+q} \widetilde{M}_{K,\epsilon}^2) a(s)^{-\epsilon/(K(1+\epsilon))} \\ & \leq (K^2 + 4K^3 C_K^2 \kappa_0^K \widetilde{M}_{K,\epsilon}^2) \hat{\omega}_{p+q}(s)^{\epsilon/(K(2+3\epsilon))} \end{aligned}$$

where the first inequality uses the fact that  $\widetilde{M}_{K,\epsilon} \geq 1$  and the second uses  $p+q \leq K$ , and the increasing property of  $\hat{\omega}_k(s)$  in the argument  $k$ . The fast decay of  $(K^2 + 4K^3 C_K^2 \kappa_0^K \widetilde{M}_{K,\epsilon}^2) \hat{\omega}_{p+q}(s)^{\epsilon/(K(2+3\epsilon))}$  follows from the fast decay of  $\hat{\omega}$ , completing the proof of Lemma 7.3.  $\square$

*Proof of Theorem 3.3.* We deduce this result from Theorem 7.2 as follows. Recall that the correlation functions  $m^{[k]_1^{p+q}}([x]_1^{p+q}; n)$  of the measures  $\mathfrak{m}_n := \sum_{x \in \mathcal{P}_n} \xi_{i,n} \delta_x$  are defined at (7.1). (In contrast to  $\hat{\mu}_n^\xi$ , the measure  $\mathfrak{m}_n$  is not scaled to  $W_1$ .) Under the moment assumption (3.4), the Hölder inequality implies that the correlation functions  $m^{[k]_1^p}([x]_1^p; n)$  exist. Moreover, the real-valued marks  $\xi_{i,n}$  have finite moments as in (7.12) for any  $\epsilon > 0$  (uniformly for  $\tilde{\mathcal{P}}_n$ ) as  $\widetilde{M}_{K,\epsilon,p+q} \leq M_{K(1+\epsilon)}^\xi$  and that the moment assumption (3.4) holds for all  $p \in \mathbb{N}$ . By assumption, we have that  $(\tilde{\mathcal{P}}_n)_{n \in \mathbb{N}}$  are fast BL-mixing as in Definition 3.1(iii). Hence, applying Lemma 7.3 with  $h_1([t]_1^p) := \prod_{i=1}^p t_i^{k_i}$  and  $h_2([t]_1^q) := \prod_{i=1}^q t_i^{k_i}$  for  $p, q, k_1, \dots, k_{p+q} \in \mathbb{N}, t_i \in \mathbb{R}$ , we obtain the fast decay of the correlation functions  $m^{[k]_1^p}(\cdot; n)$  uniformly in  $n \in \mathbb{N}$ , i.e., for any  $p, q, k_1, \dots, k_{p+q} \in \mathbb{N}$  and  $x_1, \dots, x_{p+q} \in \mathbb{R}^d$ , we have

$$|m^{[k]_1^{p+q}}([x]_1^{p+q}; n) - m^{[k]_1^p}([x]_1^p; n) m^{[k]_{p+1}^{p+q}}([x]_{p+1}^{p+q}; n)| \leq \hat{\omega}_K(s) \quad (7.16)$$

where  $s := d([x]_1^p, [x]_{p+1}^{p+q})$ ,  $K := \sum_{i=1}^{p+q} k_i$ , and where  $\hat{\omega}_K$  is a fast decreasing function for all  $K \in \mathbb{N}$ .

Thus the correlation functions  $m^{[k]_1^p}(\cdot; n)$  are fast mixing, uniformly in  $n \in \mathbb{N}$ , in accordance with statement (i) of Lemma 7.1 by taking  $\tilde{C} = 1$  and  $\phi_{k_1, \dots, k_p}(s) = \hat{\omega}_K(s)$  where  $K = \sum_{i=1}^p k_i$ . Hence, we obtain the central limit theorem (3.5) from Theorem 7.2, using the  $p$ -moment condition (3.4) to verify (7.10) and the assumed lower bound for  $\text{Var} \mu_n^\xi(f)$ .  $\square$

### 7.3 Distributional limits for BL-localizing score functions

In this section, we justify the existence of distributional limits for  $\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)$  as  $n \rightarrow \infty$  under Palm distributions as in (5.9) and (5.10), of which (5.3) and (5.4) are special cases. These are critical when proving expectation and variance asymptotics (Proposition 5.3 and Theorem 5.4) in the next subsections.

We operate under the premise that the score function  $\xi$  exhibits fast BL-localization across all finite windows of  $\tilde{\mathcal{P}}$ , as in Definition 4.2 (ii). We start by presenting key findings concerning the distributional limits resulting from BL-localization. Recall the notation

$$\xi^{(r)}(\tilde{x}_1, \tilde{\mathcal{P}}) = \xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r(x_1)) \text{ and } [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) = \left( \xi(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r(x_1)), \dots, \xi(\tilde{x}_p, \tilde{\mathcal{P}} \cap B_r(x_p)) \right).$$

**Lemma 7.4** (Distributional limits for vectors of BL-localizing scores). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ . Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be an  $\mathbb{R}$ -valued marking function and assume  $\xi$  is BL-localizing on all finite windows of  $\tilde{\mathcal{P}}$  as in (4.3) for a given  $p \in \mathbb{N}$ .*

(i) For a given vector  $[x]_1^p \in (\mathbb{R}^d)^p$ , the random vectors  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  (under Palm distributions  $\mathbb{P}_{[x]_1^p}$  of  $\tilde{\mathcal{P}}$ ) converge in the  $d_{\text{BL}, [x]_1^p}$  metric (and hence in law) as  $r \rightarrow \infty$  to a probability distribution on  $\mathbb{R}^p$ , represented here by the random vector  $\Upsilon_{[x]_1^p} = (\Upsilon_{[x]_1^p}(1), \dots, \Upsilon_{[x]_1^p}(p)) \in \mathbb{R}^p$ , i.e.,

$$[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) \xrightarrow[r \rightarrow \infty]{\text{BL}, [x]_1^p} \Upsilon_{[x]_1^p}$$

and the speed of convergence is bounded by the same decreasing function  $\varphi_p$  as in (4.3), namely

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} d_{\text{BL}, [x]_1^p}([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \Upsilon_{[x]_1^p}) \leq 4\varphi_p(r), \quad r > 0. \quad (7.17)$$

(ii) The probability distributions represented by  $\Upsilon_{[x]_1^p}$  are probability kernels from  $(\mathbb{R}^d)^p$  to  $\mathbb{R}^p$ ; formally, for any measurable function  $f$  on  $\mathbb{R}^p$ , the function  $(\mathbb{R}^d)^p \ni [x]_1^p \mapsto \mathbb{E}[f(\Upsilon_{[x]_1^p})] \in \mathbb{R}$  is measurable.

(iii) The convergence holds also for  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ :

$$[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[n \rightarrow \infty]{\text{BL}, [x]_1^p} \Upsilon_{[x]_1^p} \quad (7.18)$$

with the speed of convergence bounded by the same decreasing function  $\varphi_p$  of the distance  $d([x]_1^p, \partial W_n)$  from  $[x]_1^p \subset W_n$  to the boundary of  $W_n$ :

$$d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), \Upsilon_{[x]_1^p}) \leq 6\varphi_p(d([x]_1^p, \partial W_n)) \quad (7.19)$$

(iv) If  $\tilde{\mathcal{P}}$  is stationary and  $\xi$  is translation invariant, then the distribution of  $\Upsilon$  is translation invariant: i.e., for all  $y \in \mathbb{R}^d$

$$\Upsilon_{[x]_1^p} \stackrel{\text{law}}{=} \Upsilon_{[x]_1^p - y} = \Upsilon_{(x_1 - y, \dots, x_p - y)}.$$

(v) When  $\xi$  is defined on  $\mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  and satisfies BL-localization (4.2) for some  $p \geq 1$ , then for  $[x]_1^p \subset W_n$  we have

$$d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})) \leq 12\varphi_p(d([x]_1^p, \partial W_n)), \quad (7.20)$$

and hence, under  $\mathbb{P}_{[x]_1^p}$ ,  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  has the law of  $\Upsilon_{[x]_1^p}$ .

*Proof.* (i) Taking  $0 < r \leq r'$  and  $n$  large enough such that  $B_{r'}([x]_1^p) := \bigcup_{i=1}^p B_{r'}(x_i) \subset W_n$ , by the triangle inequality, (4.2) and the decreasing property of  $\varphi_p$ , we obtain

$$\begin{aligned} d_{\text{BL}, [x]_1^p}([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{(r')}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) & \\ & \leq d_{\text{BL}, [x]_1^p}([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)) + d_{\text{BL}, [x]_1^p}([\xi^{(r')}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)) \\ & \leq 4\varphi_p(r), \end{aligned} \quad (7.21)$$

where  $\varphi_p(r) \rightarrow 0$  as  $r \rightarrow \infty$ . Under the Palm distributions  $\mathbb{P}_{[x]_1^p}$ , this shows that any subsequence of random vectors  $[\xi^{(r_k)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$ , where  $r_k \rightarrow \infty$ , is a Cauchy sequence in the  $d_{\text{BL}, [x]_1^p}$  distance. This

justifies the first statement in (i). Indeed,  $\mathbb{R}^p$  is a complete separable metric space, and hence the weak convergence of probability measures on it is completely metrizable by the  $d_{\text{BL},[x]_1^p}$  metric (see [23, Theorem 8.10.43]). To bound the speed of convergence, we take  $r, r'$  as in (7.21) and by the triangle inequality we obtain

$$\begin{aligned} & d_{\text{BL},[x]_1^p} \left( [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \Upsilon_{[x]_1^p} \right) \\ & \leq d_{\text{BL},[x]_1^p} \left( [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi^{(r')}]_1^p(\tilde{x}, \tilde{\mathcal{P}}) \right) + d_{\text{BL},[x]_1^p} \left( [\xi^{(r')}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \Upsilon_{[x]_1^p} \right) \\ & \leq 4\varphi_p(r) + d_{\text{BL},[x]_1^p} \left( [\xi^{(r')}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \Upsilon_{[x]_1^p} \right). \end{aligned}$$

When  $r' \rightarrow \infty$ , the last term goes to 0 by the previous statement and we obtain (7.17).

(ii) This follows from the fact that Palm distributions are probability kernels. Indeed, for a bounded measurable function  $f : \mathbb{R}^p \rightarrow \mathbb{R}$ , the mapping  $(\mathbb{R}^d)^p \ni [x]_1^p \mapsto \mathbb{E}_{[x]_1^p} [f([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}))] \in \mathbb{R}^p$  is measurable. Since the law of  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  under  $\mathbb{P}_{[x]_1^p}$  converges in BL, and hence weakly, to  $\Upsilon_{[x]_1^p}$  for all  $[x]_1^p$ , the aforementioned mapping converges point-wise to  $(\mathbb{R}^d)^p \ni [x]_1^p \mapsto \mathbb{E}[f(\Upsilon_{[x]_1^p})] \in \mathbb{R}^p$ . Consequently, the limit is measurable as a pointwise limit of measurable functions. Extensions to unbounded measurable functions  $f$  are straightforward.

(iii) Taking  $r = d([x]_1^p, \partial W_n)$ , by the triangle inequality, (4.3) and (7.17) we obtain

$$\begin{aligned} & d_{\text{BL},[x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), \Upsilon_{[x]_1^p}) \\ & \leq d_{\text{BL},[x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) + d_{\text{BL},[x]_1^p}([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), \Upsilon_{[x]_1^p}) \\ & \leq 6\varphi_p(r), \end{aligned}$$

which gives (7.19) and consequently also proves the statement of the convergence (7.18).

(iv) By the stationarity of  $\tilde{\mathcal{P}}$ , the law of  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  (under  $\mathcal{P}_{[x]_1^p}$ ) is invariant by the translation of the location  $[x]_1^p$  with respect to any vector  $y \in \mathbb{R}^d$ . Hence, the statement of (iv) follows from the weak convergence in Item (i).

(v) Similarly as in the proof of Item (iii), take  $r = d([x]_1^p, \partial W_n)$  and use the triangle inequality

$$\begin{aligned} & d_{\text{BL},[x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})) \\ & \leq d_{\text{BL},[x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), \Upsilon_{[x]_1^p}) + d_{\text{BL},[x]_1^p}(\Upsilon_{[x]_1^p}, [\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}})) + d_{\text{BL},[x]_1^p}([\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}), [\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})). \\ & \leq 12\varphi_p(r), \end{aligned}$$

where the last inequality follows from (7.19), (7.17) and (4.3).  $\square$

Now, we formulate and prove some results regarding the moment conditions. The key point is that fast BL-localization of  $\xi$  together with  $(p + \epsilon)$ -moments implies  $p$ -moments of  $\Upsilon$ . It's important to note this higher moment condition in the BL setting in comparison to the stopping set setting (cf. Lemma B.1 (iii)).

**Lemma 7.5** (Moment bounds on the distributional limits of BL-localizing scores). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ . Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a  $\mathbb{R}$ -valued marking function which is BL-localizing on all finite windows for Palm distributions of some order  $q \in \mathbb{N}$  (that is (4.3) holds true with  $p$  replaced by  $q$ ).*

(i) *If  $\xi$  satisfies the  $(p + \epsilon)$ -moment condition on finite windows (5.2) for some  $p \geq q$  and some  $\epsilon > 0$ , then the  $\Upsilon_{[x]_1^q}$  have uniformly bounded  $p$ -moments for all  $[x]_1^q \in (\mathbb{R}^d)^q$ :*

$$\sup_{x_1, \dots, x_q \in (\mathbb{R}^d)^q} \max_{i=1, \dots, q} \mathbb{E}[\max(1, |\Upsilon_{[x]_1^i}(i)|^p)] \leq M_p^\xi. \quad (7.22)$$

(ii) *Furthermore, under the assumptions of Item (i) above, we also obtain the speed of convergence of moments; namely for  $[x]_1^l \in (\mathbb{R}^d)^q$  we have*

$$\max_{i=1, \dots, q} \left| \mathbb{E}_{[x]_1^q}[(\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n))^p] - \mathbb{E}[(\Upsilon_{[x]_1^i}(i))^p] \right| \leq (2M_{p+\epsilon/2}^\xi + 6p)\varphi_q(d([x]_1^l, \partial W_n))^{\epsilon/(2p+\epsilon)}, \quad (7.23)$$

where  $d([x]_1^l, \partial W_n)$  is the distance between  $[x]_1^q \subset W_n$  and the boundary of  $W_n$ .

(iii) *If  $\xi$  is BL-localizing on all finite windows of Palm order  $q = 2$  as in (4.3) and satisfies the  $(2 + \epsilon)$ -moment condition on finite windows (5.2) for some  $\epsilon > 0$ , then the vector  $\Upsilon_{[x]_1^2}$  has correlations bounded uniformly for  $[x]_1^2 \in (\mathbb{R}^d)^2$ :*

$$\sup_{(x_1, x_2) \in (\mathbb{R}^d)^2} \mathbb{E}[\max(1, |\Upsilon_{[x]_1^2}(1)\Upsilon_{[x]_1^2}(2)|)] \leq M_2^\xi. \quad (7.24)$$

(iv) *Furthermore, under the assumptions of Item (iii) above, we obtain the speed of convergence of the correlations:*

$$\left| \mathbb{E}_{[x]_1^2}[\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_{[x]_1^2}(1)\Upsilon_{[x]_1^2}(2)] \right| \leq (4M_{2+\epsilon/2}^\xi + 12)\varphi_2(d([x]_1^2, \partial W_n))^{\epsilon/(4+\epsilon)}, \quad (7.25)$$

where  $d([x]_1^2, \partial W_n)$  is the distance between  $[x]_1^2 \subset W_n$  and the boundary of  $W_n$ .

*Proof.* (i) It is well known that convergence in distribution of random variables, along with a uniform bound on  $(p + \epsilon)$ -moments, implies the convergence of  $p$ -moments. In our specific case, as stated in Lemma 7.4 (iii), this implies

$$\mathbb{E}[\max(1, |\Upsilon_{[x]_1^l}(i)|^p)] = \lim_{n \rightarrow \infty} \mathbb{E}_{[x]_1^l}[\max(1, |\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)|^p)] \leq M_p^\xi.$$

(ii) We use a truncation as in the proof of Lemma 7.3. For a function  $f_a : \mathbb{R} \rightarrow \mathbb{R}$  where  $a = a(r) \geq 1$  (the specific value of  $a$  will be determined later), consider the following definition

$$f_a(t) := \begin{cases} t^p & \text{when } |t| \leq a \\ \text{sgn}(t^p)a^p & \text{when } |s| > a. \end{cases}$$

The function  $f_a/(pa^p)$  is a bounded, Lipschitz(1) function, and thus by the triangle inequality we have

$$\left| \mathbb{E}_{[x]_1^q}[(\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n))^p] - \mathbb{E}[(\Upsilon_{[x]_1^i}(i))^p] \right|$$

$$\begin{aligned} &\leq \mathbb{E}_{[\kappa]_1^q} [|\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)|^p \mathbf{1}(|\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)| > a)] + \mathbb{E} [|\Upsilon_{[\kappa]_1^q}(i)|^p \mathbf{1}(|\Upsilon_{[\kappa]_1^q}(i)| > a)] \\ &\quad + pa^p d_{\text{BL}, [\kappa]_1^q} \left( \xi(\tilde{x}_i, \tilde{\mathcal{P}}_n), \Upsilon_{[\kappa]_1^q}(i) \right). \end{aligned} \quad (7.26)$$

Using (7.19), the last term is bounded by  $6pa^p \varphi_q(r)$  where  $r = d([\kappa]_1^l, \partial W_n)$ . Next, let's choose  $p' = p + \epsilon/2$ . Both random variables  $\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  and  $\Upsilon_{[\kappa]_1^q}(i)$  have  $p'$ -moments bounded by  $M_{p'}^\xi$  (for the former by assumption and for the latter by the statement in point (i) with  $p$  replaced by  $p'$ ). Therefore, using Hölder's inequality, both truncated moments are bounded by  $M_{p'}^\xi a^{-\epsilon/2}$  and consequently

$$\left| \mathbb{E}_{[\kappa]_1^l} [(\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n))^p] - \mathbb{E} [(\Upsilon_{[\kappa]_1^l}(i))^p] \right| \leq 2M_{p'}^\xi a(r)^{-\epsilon/2} + 6pa(r)^p \varphi_q(r).$$

Setting  $a(r) = \varphi(r)^{-2/(2p+\epsilon)}$ , the right-hand side simplifies to:

$$(2M_{p+\epsilon/2}^\xi + 6p) \varphi_q(d([\kappa]_1^l, \partial W_n))^{\epsilon/(2p+\epsilon)}.$$

This completes the statement of point (ii).

(iii) Using the same reasoning as in the proof of Item (i), the convergence in distribution of the vector  $[\xi]_1^2(\tilde{x}, \tilde{\mathcal{P}}_n)$  to  $\Upsilon_{[\kappa]_1^2}$ , as per the statement of Lemma 7.4 (iii) and the continuous mapping theorem, implies convergence in distribution of the products  $\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)$  to  $\Upsilon_{[\kappa]_1^2}(1)\Upsilon_{[\kappa]_1^2}(2)$ . Along with uniform  $(2 + \epsilon)$ -moments (5.2), by Hölder's inequality, these products possess bounded  $(1 + \epsilon)$ -moments:

$$\sup_{n \in \mathbb{N}} \sup_{[\kappa]_1^2 \in W_n^2} \mathbb{E}_{[\kappa]_1^2} [|\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)|^{1+\epsilon/2}] < \infty.$$

This implies the convergence of expectations  $\mathbb{E}_{[\kappa]_1^2} [|\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)|]$  to  $\mathbb{E} [|\Upsilon_{[\kappa]_1^2}(1)\Upsilon_{[\kappa]_1^2}(2)|]$  as  $n \rightarrow \infty$ . The explicit bound follows from (5.2) with  $p = 1$ , using Hölder's inequality.

(iv) We use the same reasoning as in the proof of Item (ii), with the truncation of the product of functions as in the proof of Lemma 7.3 with  $p = 2$  and  $k_i = 1$ ,  $i = 1, 2$ . Specifically, consider a function defined as:

$$f_a(t) := \begin{cases} t & \text{when } |t| \leq a^{1/2} \\ \text{sgn}(t)a^{1/2} & \text{when } |t| > a^{1/2}. \end{cases}$$

where the specific value of  $a = a(r) \geq 1$  will be determined later. The product  $f_a(t_1)f_a(t_2) : \mathbb{R}^2 \rightarrow [0, \infty)$  is a function bounded by  $a$  and Lipschitz with constant  $2a^{1/2}$  (cf footnote 2). Consequently,  $2af_a(t_1)f_a(t_2)$  is a bounded Lipschitz function applicable to (7.19). This truncation, along with Hölder's inequality used twice and  $(2 + \epsilon)$ -moment for  $\xi$ , leads to the following bound:

$$\left| \mathbb{E}_{[\kappa]_1^2} [\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)] - \mathbb{E}_{[\kappa]_1^2} [f_a(\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n))f_a(\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n))] \right| \leq 2M_{2+\epsilon/2}^\xi a^{-\epsilon/4}.$$

The same approximation by the truncation  $f_a$  holds true for  $|\Upsilon_{[\kappa]_1^2}(1)\Upsilon_{[\kappa]_1^2}(2)|$  because  $\Upsilon_{[\kappa]_1^2}(1)$  and  $\Upsilon_{[\kappa]_1^2}(2)$  have  $(1 + \epsilon/4)$ -moments with the same bound  $M_{1+\epsilon/4}^\xi$ . Here, we use the same argument as in the proof of Item (ii) with  $p = 1 + \epsilon/2$  and  $p' = 1 + \epsilon/4$ . Using the triangle inequality systematically as in (7.26), also using (7.19) and collecting the developed bounds, we obtain:

$$\left| \mathbb{E}_{[\kappa]_1^2} [\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)\xi(\tilde{x}_2, \tilde{\mathcal{P}}_n)] - \mathbb{E} [\Upsilon_{[\kappa]_1^2}(1)\Upsilon_{[\kappa]_1^2}(2)] \right| \leq 4M_{2+\epsilon/2}^\xi a^{-\epsilon/4} + 12a\varphi_2(r),$$

where  $r = d([x]_1^2, \partial W_n)$ . Setting  $a(r) = \varphi_2(r)^{-4/(4+\epsilon)}$ , the right-hand side simplifies to:

$$(4M_{2+\epsilon/2}^\xi + 12)\varphi_2(d([x]_1^2, \partial W_n))^{\epsilon/(4+\epsilon)}.$$

This completes the statement of point (iv).  $\square$

**Remark 7.6** (Distributional limits for BL cluster-localizing score functions). The statements of Lemmas 7.4 and 7.5 remain valid for  $[\xi^{\cup(r)}]_1^p$  (in place of  $[\xi^{(r)}]_1^p$ ) under the sole assumption of BL *union-localization* (4.6), together with the corresponding moment conditions and/or stationarity assumptions. The limiting probability kernels induced by BL cluster-localization may exist even in cases where BL-localization fails; see Example 7.7 below. In situations where the score function  $\xi$  satisfies both BL-localization and BL union-localization, the two types of limits coincide. The proofs of these results carry over verbatim, starting from the Cauchy property derived via the triangle inequality in (7.21).

**Example 7.7** (BL union-localization without cluster-localization). We give an example showing that it is possible for (4.6) to hold but not (4.7) in the definition of BL-cluster localization. This shows that marking functions may be neither BL cluster-localizing nor BL localizing.

Let  $\mu$  be a point process on  $\mathbb{R}^d$  and consider  $\tilde{\mu} = \{(x, u_x)\}_{x \in \mu}$ , a marked point process in  $\mathbb{R}^d \times [0, 1]$ . Define the marking function

$$\xi(\tilde{x}, \tilde{\mu}) = \begin{cases} \operatorname{sgn}\left(\frac{1}{\mu(\mathbb{R}^d)} \sum_{y \in \mu} u_y - 1/2\right) & \text{if } \mu \text{ is finite} \\ \operatorname{sgn}(u_x - 1/2) & \text{if } \mu \text{ is infinite,} \end{cases} \quad (7.27)$$

where  $\operatorname{sgn}(x) := x/|x|$  is the sign function with  $\operatorname{sgn}(0)$  chosen arbitrarily. Let  $\tilde{\mathcal{P}}$  be a homogeneous Poisson process equipped with random variables  $U_x, x \in \mathcal{P}$ , which are i.i.d. uniform on  $[0, 1]$ . The  $U_x$ 's are continuous random variables and hence for all Borel sets  $B \subseteq \mathbb{R}^d$ ,  $\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B), x \in \tilde{\mathcal{P}} \cap B$ , are symmetric Rademacher random variables and are independent iff  $|B| = \infty$ . If  $|B| < \infty$ , then  $\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B) = \xi(\tilde{y}, \tilde{\mathcal{P}} \cap B)$  for all  $x, y \in \tilde{\mathcal{P}} \cap B$ .

For all  $n \in \mathbb{N} \cup \{\infty\}$ , we have for all  $r > 0$  that

$$d_{\text{BL}, [x]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)) = 0,$$

that is (4.6) holds, whereas for all  $r < s([x]_1^q, [x]_{q+1}^p)/2$ , there are constants  $a_p$  such that

$$d_{\text{BL}, [x]_1^p}\left([\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), ([\xi^{\cup(r)}]_1^q(\tilde{x}, \tilde{\mathcal{P}}_n), [\xi^{\cup(r)}]_{q+1}^p(\tilde{x}, \tilde{\mathcal{P}}_n))\right) \geq a_p > 0, \quad (7.28)$$

and hence  $\xi$  fails to satisfy BL cluster-localization. This is because the co-ordinates of  $[\xi^{\cup(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  are all +1 or -1 but  $[\xi^{\cup(r)}]_1^q(\tilde{x}, \tilde{\mathcal{P}}_n)$  and  $[\xi^{\cup(r)}]_{q+1}^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  are independent. It can be verified similarly that  $\xi$  satisfies neither (4.2) nor (4.3), i.e.,  $\xi$  is neither BL-localizing on  $\mathbb{R}^d$  nor on the finite windows of  $\mathbb{R}^d$ .

Further,  $H_n^\xi = \sum_{x \in \mathcal{P}_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) = |\mathcal{P}_n|X$  where  $X$  is an independent symmetric Rademacher random variable. Thus  $\mathbb{E} H_n^\xi = 0$  and  $\operatorname{Var} H_n^\xi = \mathbb{E} |\mathcal{P}_n|^2 = n^2 + n$ . The strong law of large numbers implies that as  $n \rightarrow \infty$

$$\frac{H_n^\xi}{\sqrt{\operatorname{Var} H_n^\xi}} = X \frac{|\mathcal{P}_n|}{\sqrt{n^2 + n}} \xrightarrow{\text{a.s.}} X,$$

and so  $\frac{H_n^\xi}{\sqrt{\operatorname{Var} H_n^\xi}} \Rightarrow X$  as  $n \rightarrow \infty$ . Thus the CLT fails in this case.

## 7.4 Proofs of expectation and variance asymptotics in Section 5.2

*Proof of Proposition 5.3.* (i) We establish the mean asymptotics (5.5). Recall that  $\tilde{x} = (x, U(x))$ . The Campbell-Little-Mecke formula yields

$$\begin{aligned}
& \left| n^{-(d-1)/d} \mathbb{E} \mu_n^\xi(f) - n^{1/d} \rho \mathbb{E}[\Upsilon_{\mathbf{0}}] \int_{W_1} f(x) \, dx \right| \\
&= n^{-(d-1)/d} \left| \mathbb{E} \left[ \sum_{x \in \mathcal{P}_n} f(n^{-1/d}x) \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \right] - n \rho \mathbb{E}[\Upsilon_{\mathbf{0}}] \int_{W_1} f(x) \, dx \right| \\
&\leq n^{-(d-1)/d} \rho \int_{W_n} |f(n^{-1/d}x)| \left| \mathbb{E}_x[\xi(\tilde{x}, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_x] \right| \, dx \\
&\leq (2M_\xi^{1+\epsilon/2} + 6) \|f\|_\infty \rho n^{-(d-1)/d} \int_{W_n} \varphi_1(d(x, \partial W_n))^{\epsilon/(2+\epsilon)} \, dx.
\end{aligned}$$

Here, we used the distributional equality  $\Upsilon_x \stackrel{\text{law}}{=} \Upsilon_{\mathbf{0}}$  (see Lemma 7.4 (iv)) and in the last inequality, we used the estimate (7.23) to bound  $|\mathbb{E}_x[\xi(\tilde{x}, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_x]|$ . The fast decreasing property of  $\varphi_1$  gives

$$n^{-(d-1)/d} \int_{W_n} \varphi_1(d(x, W_n))^{\epsilon/(2+\epsilon)} \, dx = O(1),$$

and thus (5.5) follows. If  $\varphi_1$  is a decreasing function, but not necessarily fast decreasing (as in (2.1)), then

$$n^{-1} \int_{W_n} \varphi_1(d(x, W_n))^{\epsilon/(2+\epsilon)} \, dx = o(1).$$

(ii) Now we establish variance asymptotics (5.6). Using the Campbell-Little-Mecke formula we obtain

$$\begin{aligned}
\text{Var } \mu_n^\xi(f) &= \mathbb{E} \sum_{x \in \mathcal{P}_n} f(n^{-1/d}x)^2 \xi^2(\tilde{x}, \tilde{\mathcal{P}}_n) \\
&\quad + \mathbb{E} \sum_{x, y \in \mathcal{P}_n, x \neq y} f(n^{-1/d}x) f(n^{-1/d}y) \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \xi(\tilde{y}, \tilde{\mathcal{P}}_n) \\
&\quad - \left( \mathbb{E} \sum_{x \in \mathcal{P}_n} f(n^{-1/d}x) \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \right)^2 \\
&= \int_{u \in W_n} f(n^{-1/d}u)^2 \mathbb{E}_x(\xi^2(\tilde{u}, \tilde{\mathcal{P}}_n)) \rho \, du \\
&\quad + \int_{W_n \times W_n} f(n^{-1/d}u) f(n^{-1/d}v) (m_{(2)}(u, v; n) - m_{(1)}(u; n) m_{(1)}(v; n)) \, dudv,
\end{aligned} \tag{7.29}$$

where

$$\begin{aligned}
m_{(1)}(u; n) &:= \mathbb{E}_u \xi(\tilde{u}, \tilde{\mathcal{P}}_n) \rho \\
m_{(2)}(u, v; n) &:= \mathbb{E}_{u, v} [\xi(\tilde{u}, \tilde{\mathcal{P}}_n) \xi(\tilde{v}, \tilde{\mathcal{P}}_n)] \rho^{(2)}(u, v).
\end{aligned}$$

The first term in (7.29), multiplied by  $n^{-1}$ , converges to

$$\rho \mathbb{E}[\Upsilon_0^2] \int_{W_1} f(x)^2 dx < \infty. \quad (7.30)$$

Indeed, we employ similar arguments as in the proof of expectation asymptotics in point (i), now using the estimate (7.23) with  $q = 1$  and  $p = 2$  for  $\left| \mathbb{E}_x[\xi^2(\tilde{x}, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_x^2] \right|$  under the  $(2 + \epsilon)$ -moment condition.

For the second term in (7.29) multiplied by  $n^{-1}$ , we set  $x = n^{-1/d}u$ ,  $z = v - u = v - n^{1/d}x$  and we rewrite it as:

$$\int_{W_1} \int_{W_n - n^{1/d}x} f(x) f(x + n^{-1/d}z) \left[ m_{(2)}(n^{1/d}x, n^{1/d}x + z; \tilde{\mathcal{P}}_n) - m_{(1)}(n^{1/d}x; \tilde{\mathcal{P}}_n) m_{(1)}(n^{1/d}x + z; \tilde{\mathcal{P}}_n) \right] dz dx. \quad (7.31)$$

We claim now that (7.31) converges to the desired value:

$$\int_{W_1} f^2(x) dx \times \int_{\mathbb{R}^d} (\mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1)\Upsilon_{(\mathbf{0}, z)}(2)] \rho^{(2)}(\mathbf{0}, z) - \rho^2 \mathbb{E}[\Upsilon_0]^2) dz < \infty. \quad (7.32)$$

To establish this convergence as  $n \rightarrow \infty$ , we use the Lebesgue dominated convergence theorem. Indeed,  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  exhibits fast BL-mixing correlations of order two as in (3.2) with  $p = q = 1$ ; specifically, it satisfies (5.17). The BL-mixing of correlations is guaranteed by Theorem 4.9 (i) and Corollary 4.10 under the conditions of fast BL-localization of  $\xi$  for  $p = \{1, 2\}$  and summable exponential  $\mathbb{B}$ -mixing correlations of  $\tilde{\mathcal{P}}$  as in Definition 4.8. Then the  $(2 + \varepsilon)$ -moment condition and Lemma 7.3 ensure that the integrand in (7.31) is bounded uniformly for  $x \in W_1$ ,  $z \in W_n - n^{1/d}x$ , and  $n \in \mathbb{N}$ , by a fast-decreasing function  $\tilde{\omega}_2(|z|)$ , which is integrable over  $\mathbb{R}^d$ .

Under the integral, we can exploit the point-wise convergence of the expectations and correlations developed in Lemma 7.5 to obtain

$$\lim_{n \rightarrow \infty} \mathbb{E}_{n^{1/d}x}[\xi(\widetilde{n^{1/d}x}, \tilde{\mathcal{P}}_n)] = \lim_{n \rightarrow \infty} \mathbb{E}_{n^{1/d}x+z}[\xi(\widetilde{n^{1/d}x+z}, \tilde{\mathcal{P}}_n)] = \mathbb{E}[\Upsilon_0], \quad (7.33)$$

$$\lim_{n \rightarrow \infty} \mathbb{E}_{n^{1/d}x, n^{1/d}x+z}[\xi(\widetilde{n^{1/d}x}, \tilde{\mathcal{P}}_n) \xi(\widetilde{n^{1/d}x+z}, \tilde{\mathcal{P}}_n)] = \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1)\Upsilon_{(\mathbf{0}, z)}(2)]. \quad (7.34)$$

Specifically, the stationarity (refer to Lemma 7.4 (iv)), for any  $x, z \in \mathbb{R}^d$ , we observe the following equality in distributions:  $\Upsilon_0 \stackrel{\text{law}}{=} \Upsilon_{n^{1/d}x} \stackrel{\text{law}}{=} \Upsilon_{n^{1/d}x+z}$ , which consequently holds for expectations as well. Additionally, based on estimate (7.23), the differences

$$\left| \mathbb{E}_{n^{1/d}x}[\xi(\widetilde{n^{1/d}x}, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_{n^{1/d}x}] \right| \quad \text{and} \quad \left| \mathbb{E}_{n^{1/d}x+z}[\xi(\widetilde{n^{1/d}x+z}, \tilde{\mathcal{P}}_n)] - \mathbb{E}[\Upsilon_{n^{1/d}x+z}] \right|$$

are bounded by a decreasing function of the distance of  $n^{1/d}x$  and  $n^{1/d}x + z$ , respectively, to the boundary of  $W_n$ . As  $n$  approaches infinity, this distance scales as  $n^{1/d}$  for fixed  $x$  and  $z$  and thus we have shown (7.33).

Similar arguments can be applied to deduce the convergence of correlations in (7.34), utilizing the stationarity for  $\Upsilon_{(n^{1/d}x, n^{1/d}x+z)} \stackrel{\text{law}}{=} \Upsilon_{(\mathbf{0}, z)}$  and the bound (7.25). Again, this involves a decreasing function of the minimum of the distances  $d(n^{1/d}x, \partial W_n)$  and  $d(n^{1/d}x + z, \partial W_n)$ .

Finally, the continuity of  $f$  would suffice for the convergence of  $f(x + z/n^{1/d})$  to  $f(x)$ . In the absence of this assumption, we employ the argument of Lebesgue points of  $f$  which is borrowed from the proof of [115, Theorem 2.1] and was also used in a similar context in [21]. This completes the proof of the variance asymptotics.

The asymptotic of the covariance follows from the polarization identity

$$\text{Cov}(\mu_n^\xi(f), \mu_n^\xi(g)) = \frac{1}{2} (\text{Var}\mu_n^\xi(f + g) - \text{Var}\mu_n^\xi(f) - \text{Var}\mu_n^\xi(g)).$$

□

**Remark 7.8** (Expectation and variance asymptotics under BL cluster-localization). For  $p = 1$ , BL-localization coincides with BL cluster-localization— more precisely, with its first condition (4.6), referred to as BL union-localization. Consequently, Item (i) of Proposition 5.3 (the asymptotic behavior of the mean) remains valid under BL cluster-localization.

The variance asymptotic, stated in Item (ii) of the same proposition, also remains valid under BL cluster-localization, provided the localization holds for  $p \in \{1, 2\}$ . For  $p = 2$ , the Palm limits and the required moment bounds already exist under the first condition (4.6), as noted in Remark 7.6. However, the second condition of BL cluster-localization, namely (4.7), is genuinely required. Indeed, the convergence in (7.34) relies on fast BL-mixing correlations of order two. Such mixing is ensured only when the two conditions defining BL cluster-localization are satisfied jointly, as formalized in Theorem 4.9(ii). Example 7.7 in Section 7.3 provides a counterexample illustrating the necessity of the second condition.

## 7.5 Proof of multivariate central limit theorem in Section 5.2

*Proof of Theorem 5.4.* The asymptotics of the mean and variance of the random variables  $(\mu_n^{\xi_i}(f))_{n \in \mathbb{N}}$ ,  $i = 1, \dots, m$  follow from Proposition 5.3. The limiting covariance is derived from (5.6) and the variance-covariance identity:

$$\text{Cov}(\mu_n^{\xi_i}(f), \mu_n^{\xi_j}(f)) = \frac{1}{2} (\text{Var}(\mu_n^{\xi_i + \xi_j}(f)) - \text{Var}(\mu_n^{\xi_j}(f)) - \text{Var}(\mu_n^{\xi_i}(f))), \quad 1 \leq i \leq j \leq k, \quad (7.35)$$

where the sum  $\xi_i + \xi_j$  forms a new BL– localizing score function (with an order  $p = 2$  sufficing for this) and the assumed  $(2 + \varepsilon)$ -moment condition is satisfied.

The proof of the multivariate central limit theorem for the vector  $(\mu_n^{\xi_1}(f), \dots, \mu_n^{\xi_m}(f))$  uses the Cramér-Wold device and involves relaxing the moment conditions in Theorem 5.2 to some moment  $p > 2$ . For this, we employ an approach suggested by Matthias Schulte, which focuses on reducing the problem of asymptotic normality to establishing three variance asymptotics, specifically (7.36), (7.39) and (7.40) below.

*Cramér-Wold linear combination:* Consider a linear combination  $\xi := \sum_{i=1}^m t_i \xi_i$  of the score functions  $\xi_i$  with given fixed coefficients  $t_i \in \mathbb{R}$ ,  $i = 1, \dots, m$ . We aim to prove the central limit theorem for  $\mu_n^\xi(f) = \sum_{i=1}^m t_i \mu_n^{\xi_i}(f)$  for some given function  $f \in \mathbb{B}(W_1)$ . Observe, by the joint fast BL-localization of  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$  assumed in Theorem 5.4, the linear combination  $\xi$

also satisfies fast BL-localization on finite windows on  $\tilde{\mathcal{P}}$  as in (4.3) with fast decreasing functions  $\varphi_p$  (depending on coefficients  $t_i, i = 1, \dots, m$ ) for all  $p \in \mathbb{N}$ . Consequently, by the assumption of  $\tilde{\mathcal{P}}$  having summable exponential  $\mathbb{B}$ -mixing correlations, Theorem 4.9 (i), and Corollary 4.10,  $\xi$  has fast BL-mixing correlations as in (3.2) on this input process (with  $U$  replaced by  $\xi$ ), with some fast decreasing functions  $\hat{\omega}_k$  (again involving  $t_i, i = 1, \dots, m$ ). Clearly,  $\xi$  inherits the stationarity of  $\xi$  and the  $p > 2$  moment condition for  $\xi_i, i = 1, \dots, m$  (however, the values of the constants  $M_p^\xi$  involve the linear coefficients  $t_i, i = 1, \dots, m$ ).

*BL-truncation of  $\xi$ :* To bypass the requirement of all moment conditions for the central limit theorem in Theorem 5.2, for a given  $M > 0$  we define a function  $g_M : \mathbb{R} \rightarrow \mathbb{R}$  by  $g_M(x) := x\mathbf{1}_{|x| \leq M} + \text{sgn}(x)M\mathbf{1}_{|x| > M}$  and consider two auxiliary score functions being transformations of  $\xi$ :

$$\begin{aligned}\xi_M(x, \tilde{\mathcal{P}}_n) &:= g_M(\xi(x, \tilde{\mathcal{P}}_n)), \\ \xi_{\Delta M}(x, \tilde{\mathcal{P}}_n) &:= \xi(x, \tilde{\mathcal{P}}_n) - \xi_M(x, \tilde{\mathcal{P}}_n).\end{aligned}$$

Observe the equivalent representation

$$\xi_{\Delta M}(x, \tilde{\mathcal{P}}_n) = g_{\Delta M}(\xi(x, \tilde{\mathcal{P}}_n)),$$

with  $g_{\Delta M}(x) := (x - \text{sgn}(x)M)\mathbf{1}_{|x| > M}$ . Since  $g_M$  and  $g_{\Delta M}$  are Lipschitz(1), we have  $f \circ \bigotimes_{i=1}^p g_M, f \circ \bigotimes_{i=1}^p g_{\Delta M} \in \text{BL}(\mathbb{R}^p)$  for any  $f \in \text{BL}(\mathbb{R}^p)$ . Therefore,  $\xi_M$  and  $\xi_{\Delta M}$  satisfy the same fast BL-localization properties (4.3) as does  $\xi$ , with the same fast-decreasing function  $\varphi_p$ , for all  $p \in \mathbb{N}$ , uniformly in  $M > 0$ . Similarly,  $\xi_M$  and  $\xi_{\Delta M}$  have fast BL-mixing correlations as in (3.2) on this input process (with  $U$  replaced by  $\xi_M$  and  $\xi_{\Delta M}$ , respectively), with the same fast decreasing functions  $\hat{\omega}_k$  as for  $\xi$ , uniformly in  $M$ . This is again a consequence of the fact that the functions  $f \circ \bigotimes_{i=1}^p g_M, g \circ \bigotimes_{i=1}^q g_M, f \circ \bigotimes_{i=1}^p g_{\Delta M}$ , and  $g \circ \bigotimes_{i=1}^q g_{\Delta M}$  are Lipschitz(1), when  $f \in \text{BL}(\mathbb{R}^p)$  and  $g \in \text{BL}(\mathbb{R}^q)$ . Clearly,  $\xi_M$  and  $\xi_{\Delta M}$  are stationary just like  $\xi$ . Furthermore,  $\xi_M$  satisfies  $p$ -moment conditions (3.4) for all  $p > 1$ , with constants  $M_p^{\xi_M} \leq \max(1, M^p)$ , while  $\xi_{\Delta M}$  has the same  $p > 2$  moment condition existing for  $\xi$ , with a bound  $M_p^{\xi_{\Delta M}} \leq M_p^\xi$  because  $|g_{\Delta M}(x)| \leq |x|$ .

*Variance asymptotics of  $\xi, \xi_M$ , and  $\xi_{\Delta M}$ :* By the above observations, the three score functions  $\xi, \xi_M$ , and  $\xi_{\Delta M}$  satisfy the assumptions for the variance asymptotics formulated in Proposition 5.3 (ii). We denote these limits by:

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^\xi(f) =: \sigma^2, \quad (7.36)$$

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^{\xi_M}(f) =: \sigma_M^2, \quad (7.37)$$

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^{\xi_{\Delta M}}(f) =: v_M^2. \quad (7.38)$$

Moreover, by Theorem 5.2,  $\mu_n^{\xi_M}(f)$  satisfies the (possibly degenerate) CLT as  $n \rightarrow \infty$ . Thus, to establish (5.11) for  $\mu_n^\xi(f)$ , we assert that it is enough to show the following limits

$$\limsup_{M \rightarrow \infty} v_M^2 \rightarrow 0, \quad (7.39)$$

$$\lim_{M \rightarrow \infty} \sigma_M^2 \rightarrow \sigma^2. \quad (7.40)$$

*Central limit theorem for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$  via three variance asymptotics:* Before establishing (7.39) and (7.40), we demonstrate how these limits and the asymptotic normality of  $(\mu_n^{\xi_M}(f))_{n \in \mathbb{N}}$  implies that of  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ . Denote

$$\bar{\mu}_n^\xi(f) := n^{-1/2}(\mu_n^\xi(f) - \mathbb{E}[\mu_n^\xi(f)]),$$

and similarly  $\bar{\mu}_n^{\xi_M}(f)$  for  $\xi_M$ . The triangle inequality gives, for all  $m \in \mathbb{N}$  and  $\sigma^2 \geq 0$ ,

$$\begin{aligned} d_{\text{BL}}(\bar{\mu}_n^\xi(f), N(0, \sigma^2)) & \\ \leq d_{\text{BL}}(\bar{\mu}_n^\xi(f), \bar{\mu}_n^{\xi_M}(f)) + d_{\text{BL}}(\bar{\mu}_n^{\xi_M}(f), N(0, \sigma_M^2)) + d_{\text{BL}}(N(0, \sigma_M^2), N(0, \sigma^2)). \end{aligned} \quad (7.41)$$

In view of

$$\begin{aligned} d_{\text{BL}}(\bar{\mu}_n^\xi(f), \bar{\mu}_n^{\xi_M}(f)) & \\ \leq \mathbb{E}|\bar{\mu}_n^\xi(f) - \bar{\mu}_n^{\xi_M}(f)| & \leq \sqrt{n^{-1} \text{Var}(\mu_n^\xi(f) - \mu_n^{\xi_M}(f))} = \sqrt{n^{-1} \text{Var} \mu_n^{\xi_{\Delta M}}(f)}, \end{aligned}$$

we can rewrite (7.41) as

$$\begin{aligned} d_{\text{BL}}(\bar{\mu}_n^\xi(f), N(0, \sigma^2)) & \\ \leq \sqrt{n^{-1} \text{Var} \mu_n^{\xi_{\Delta M}}(f)} + d_{\text{BL}}(\bar{\mu}_n^{\xi_M}(f), N(0, \sigma_M^2)) + d_{\text{BL}}(N(0, \sigma_M^2), N(0, \sigma^2)). \end{aligned} \quad (7.42)$$

With  $M$  fixed, we take  $\limsup_{n \rightarrow \infty}$  of both sides of (7.42). On the right-hand side: the first term goes to  $v_M$  by (7.38), while the second term goes to 0 either by  $\sigma_M^2 = 0$  or the asymptotic normality of  $(\bar{\mu}_n^{\xi_M}(f))_{n \in \mathbb{N}}$ . Consequently,

$$\limsup_{n \rightarrow \infty} d_{\text{BL}}(\bar{\mu}_n^\xi(f), N(0, \sigma^2)) \leq v_M + d_{\text{BL}}(N(0, \sigma_M^2), N(0, \sigma^2)).$$

Now letting  $M \rightarrow \infty$  on both sides and using (7.39) and (7.40) establishes (5.11) for  $(\mu_n^\xi(f))_{n \in \mathbb{N}}$ . It is therefore enough to show (7.39) and (7.40), as asserted.

*Proof of limits (7.39) and (7.40):* By (5.6) and (5.8), the asymptotic variance of  $\xi$  can be represented as

$$\sigma^2 = \sigma^2(\Upsilon) \int_{W_1} f(x)^2 dx, \quad (7.43)$$

with

$$\sigma^2(\Upsilon) = \rho \mathbb{E}[(\Upsilon_0)^2] + \int_{\mathbb{R}^d} (\mathbb{E}[\Upsilon_{(0,z)}(1)\Upsilon_{(0,z)}(2)]\rho^{(2)}(\mathbf{0}, z) - \rho^2 \mathbb{E}[\Upsilon_0]^2) dz, \quad (7.44)$$

where  $\Upsilon_0$  and  $\Upsilon_{(0,z)} = (\Upsilon_{(0,z)}(1), \Upsilon_{(0,z)}(2))$  represent the corresponding Palm distributional limits of  $\xi$  defined in (5.3) and (5.4).

Analogously,  $\sigma_M^2$  and  $v_M^2$  can be represented by (7.43)-(7.44) with  $\Upsilon_0$  and  $\Upsilon_{(0,z)}$  replaced, respectively, by  $\Upsilon_0^M$ ,  $\Upsilon_{(0,z)}^M$  and  $\Upsilon_0^{\Delta M}$ ,  $\Upsilon_{(0,z)}^{\Delta M}$ , representing the corresponding Palm distributional limits of  $\xi_M$  and  $\xi_{\Delta M}$ . Observe in (5.3) and (5.4), these limits are for the auxiliary score functions  $\xi_M = g_M(\xi)$

and  $\xi_{\Delta M} = g_{\Delta M}(\xi)$ , which are produced by composing  $\xi$  with Lipschitz(1) functions  $g_M$  and  $g_{\Delta M}$ . Using the continuity of Lipschitz functions with respect to BL-convergence, it follows that with  $M$  fixed, the Palm distributional BL-limits as  $n \rightarrow \infty$  of  $\xi_M$  and  $\xi_{\Delta M}$  can be computed by applying the corresponding Lipschitz(1) functions to the limit of  $\xi$ :

$$\begin{aligned} \xi_M(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n) &\xrightarrow[n \rightarrow \infty]{\text{BL}, \mathbf{0}} \Upsilon_{\mathbf{0}}^M \stackrel{\text{law}}{=} g_M(\Upsilon_{\mathbf{0}}), \\ \left( \xi_M(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n), \xi_M(\tilde{z}, \tilde{\mathcal{P}}_n) \right) &\xrightarrow[n \rightarrow \infty]{\text{BL}, (\mathbf{0}, z)} \Upsilon_{(\mathbf{0}, z)}^M \stackrel{\text{law}}{=} \left( g_M(\Upsilon_{(\mathbf{0}, z)}(1)), g_M(\Upsilon_{(\mathbf{0}, z)}(2)) \right), \end{aligned}$$

and similarly,

$$\begin{aligned} \Upsilon_{\mathbf{0}}^{\Delta M} &\stackrel{\text{law}}{=} g_{\Delta M}(\Upsilon_{\mathbf{0}}), \\ \Upsilon_{(\mathbf{0}, z)}^{\Delta M} &\stackrel{\text{law}}{=} \left( g_{\Delta M}(\Upsilon_{(\mathbf{0}, z)}(1)), g_{\Delta M}(\Upsilon_{(\mathbf{0}, z)}(2)) \right). \end{aligned}$$

Now, by the dominated convergence theorem and the  $p > 2$  moment condition (see Lemma 7.5),

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{\mathbf{0}}^{\Delta M}] &= \mathbb{E}[\lim_{M \rightarrow \infty} g_{\Delta M}(\Upsilon_{\mathbf{0}})] = 0, \\ \lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}^{\Delta M}(i)] &= \mathbb{E}[\lim_{M \rightarrow \infty} g_{\Delta M}(\Upsilon_{(\mathbf{0}, z)}(i))] = 0, \quad i = 1, 2, \end{aligned}$$

and also

$$\lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}^{\Delta M}(1)\Upsilon_{(\mathbf{0}, z)}^{\Delta M}(2)] = 0.$$

Analogously, we have for  $\xi_M$ :

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{\mathbf{0}}^M] &= \mathbb{E}[\lim_{M \rightarrow \infty} g_M(\Upsilon_{\mathbf{0}})] = \mathbb{E}[\Upsilon_{\mathbf{0}}], \\ \lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}^M(i)] &= \mathbb{E}[\lim_{M \rightarrow \infty} g_M(\Upsilon_{(\mathbf{0}, z)}(i))] = \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(i)], \quad i = 1, 2, \\ \lim_{M \rightarrow \infty} \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}^M(1)\Upsilon_{(\mathbf{0}, z)}^M(2)] &= \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}(1)\Upsilon_{(\mathbf{0}, z)}(2)]. \end{aligned}$$

To establish limits (7.39) and (7.40), it remains to justify the passage of  $\lim_{M \rightarrow \infty}$  under the integral  $\int_{\mathbb{R}^d} [\dots] dz$  in (7.44). This is justified by the dominated convergence theorem, evoking, for both  $\xi_M$  and  $\xi_{\Delta M}$ , BL-mixing correlations uniformly in  $M > 0$ ; that is, these properties involve the common functions  $\omega_k$  in (3.2). Indeed, by Lemma 7.3 with  $p = q = 1$ ,  $K = 2$ , we have

$$\left| \mathbb{E}[\Upsilon_{(\mathbf{0}, z)}^M(1)\Upsilon_{(\mathbf{0}, z)}^M(2)]\rho^{(2)}(\mathbf{0}, z) - \rho^2 \mathbb{E}[\Upsilon_{\mathbf{0}}^M]^2 \right| \leq \hat{\omega}_2(|z|)$$

and similarly for  $\Upsilon^{\Delta M}$ , with a fast-decreasing function  $\hat{\omega}_2$  that depends on  $\omega_2$  (and other constants) uniformly in  $M$ .

This completes the proof of limits (7.39) and (7.40), and thus completes the proof of Theorem 5.4.  $\square$

**Remark 7.9** (Bypassing the summable exponential mixing condition in the multivariate CLT). In Theorem 5.4, which establishes a multivariate CLT for statistics of point processes with fast BL-localizing marks under a  $2 + \epsilon$  moment condition, the assumption of summable exponential mixing

correlations of the input process  $\tilde{\mathcal{P}}$  can be relaxed to merely fast mixing of  $\mathcal{P}$ , provided that one can directly verify the fast BL-mixing correlations of the (iterated) marks  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$  associated with  $\mathcal{P}$ . Indeed, the stronger assumption on  $\tilde{\mathcal{P}}$  is only required in the application of Theorem 4.9 to ensure the corresponding mixing property of these new marks. All the crucial steps in the proof of Theorem 5.4—namely, the use of the Cramér–Wold linear combination  $\xi$  and the justification of the CLT for its BL-truncation  $\xi_M$ —rely solely on the fast BL-mixing of  $\xi$  (of all orders), while the derivation of the variance asymptotics depends only on the BL-mixing of  $\xi$  of order 2.

## 7.6 Proof of limit theory for scores on infinite window in Section 5.3

*Proof of Proposition 5.5.* (i) Taking  $\xi_{i,n} = \xi_i := \xi(\tilde{x}_i, \tilde{\mathcal{P}})$ , the central limit theorem for  $(\hat{\mu}_n^\xi(f))_{n \in \mathbb{N}}$  is an easier version of the triangular array central limit theorem given by Theorem 3.3, where the summands of  $\hat{\mu}_n^\xi$ , namely the terms  $\xi_{i,n} := \xi(\tilde{x}_i, \tilde{\mathcal{P}})$ , are equal for  $n \in \mathbb{N}$ . Indeed, the arguments in the proof of Theorem 5.2 apply straightforwardly via Theorem 4.9. The assumptions on  $\tilde{\mathcal{P}}$  and  $\xi$  imply that the family of marked point process  $\tilde{\mathcal{P}}_n := \{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}))\}_{x \in \mathcal{P}_n}$  has fast BL-mixing correlations as in Definition 3.1(iv). Hence Theorem 3.3 yields the central limit theorem for  $(\hat{\mu}_n^\xi(f))_{n \in \mathbb{N}}$ .

(ii) For the limit of the expectation, observe by Campbell-Little-Mecke’s formula and substituting  $x$  for  $n^{-1/d}x$  that

$$\mathbb{E} \left[ \sum_{x \in \mathcal{P}_n} f(n^{-1/d}x) \xi(\tilde{x}, \tilde{\mathcal{P}}) \right] = n\rho \mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})] \int_{W_1} f(x) dx.$$

Hence, the right-hand side of (5.5) is equal to zero when  $\mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]$  is substituted for  $\mathbb{E}[\Upsilon_{\mathbf{0}}]$ .

For the variance, observe in the proof of Proposition 5.3 in Section 7.4 that for all  $n \in \mathbb{N}$

$$m_{(1)}(u; n) = \mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})] \rho \text{ and } m_{(2)}(u, v; n) = \mathbb{E}_{\mathbf{0}, z}[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})\xi(\tilde{z}, \tilde{\mathcal{P}})] \rho^{(2)}(0, z) \text{ with } z = v - u.$$

Consequently, the expression in the bracket  $[\dots]$  in (7.31) is equal to

$$\mathbb{E}_{\mathbf{0}, z}[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})\xi(\tilde{z}, \tilde{\mathcal{P}})\rho^{(2)}(0, z) - \mathbb{E}_0[\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})]^2 \rho^2],$$

which is integrable with respect to  $z$  on  $\mathbb{R}^d$  thanks to the fast mixing correlations of  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  guaranteed by Theorem 4.9 (i). Then, the limit of  $f(x + n^{-1/d}z)$  to  $f(x)$  under the integral relies on the same arguments employed in proof of Propositions 5.3.  $\square$

## Part III

# Applications

In this part, the theoretical framework developed in the previous parts is applied to various spatial random models. The focus is on spin systems (Section 8), interacting diffusions (Section 9), and particle systems on spatial random graphs (Section 10). Each application illustrates the practical relevance of the theory, offering detailed examples and proofs of how these systems can be studied within the framework established in Part II. This part also explores applications to empirical random fields and geostatistical models (Section 11) as well as indicates applications to other spatial random models (Section 12) highlighting the versatility of the methods introduced. Carrying out these various applications often requires formulating analogues of the aforementioned models—originally defined on a fixed graph—in a setting based on random geometry. This then calls for establishing directly BL cluster-localization for spin systems and, indirectly, BL-localization for statistics of other models by leveraging more classical stabilization techniques (such as  $L^2$ -stabilization for interacting diffusions and stopping-set approaches for other classes of models). Moreover, models based on graphs constructed on the input ground process typically require the use of stabilization techniques for the associated graph structures.

For convenience, we recall our standing assumption from Part II.

**Assumption 3.1.** *Recall that our standing assumptions are as follows:*

(i)  $\mathcal{P}$  denotes a simple point process on  $\mathbb{R}^d$  such that the correlation functions  $\rho^{(p)}$ ,  $p \in \mathbb{N}$ , exist and  $\kappa_p := \sup_{[x]_1^p \in (\mathbb{R}^d)^p} \rho^{(p)}([x]_1^p) < \infty$ . We put  $\kappa_0 := \max\{\kappa_1, 1\}$ .

(ii) All marked point processes in this article are assumed to be simple.

Our applications hold for many point processes and spatial random graphs with explicit assumptions made precise in the respective results. However, for simplicity and to get a quick understanding of the applications, the reader may take  $\mathcal{P}$  to be a stationary Poisson process or a determinantal process with fast-decaying kernel, such as the Ginibre process, and one may take the spatial random graph to be the Gilbert graph,  $k$ -nearest neighbor graph or Delaunay graph. Some of these applications—most notably spin systems involving BL cluster-localization—are new even in the context of Poisson processes. We also alert the reader that one may consider a more extended framework for random graphs than the one investigated here. In particular, this could presumably include a random connection model or a Boolean model with random grains in our framework. We remark on this in Section 12.

## 8 Spin systems on spatial random graphs

Spin systems (or Gibbs random fields) defined on spatial random graphs fall within the scope of our main general result, namely Theorem 5.2. Specifically, we establish the asymptotic normality of the total sum of spins. Our Gibbsian models differ from those of Gibbs point processes (e.g., see

Dereudre [39]), where probability measures are defined on locally finite configurations of points,  $\mathcal{N}_{\mathbb{R}^d}$ . In contrast, we consider a spatial random *interaction graph* on a point process. Given its realization, we define a spin system (or Gibbs random field) on the points based on the adjacency relations within the graph. This approach is inspired by lattice spin systems, except that here, the underlying graph is random. Our models serve as spatial analogs to statistical mechanics models on (non-spatial) random graphs, such as those studied by Georgii et al. [64], van der Hofstad [152], Duminil-Copin [48], and Friedli and Velenik [60].

In Section 8.1, we introduce the spin model on spatial random graphs and two key assumptions: *stabilization* of the interaction graph and *averaged weak spatial mixing* of spins on this graph. Stabilization refers to the property that the construction of a random graph, based on an input process of points (interpreted as graph vertices), is governed by the stopping sets of these points. In this context, long-range interactions in the graph structure statistically decay at a fast rate. This property, which also ensures sparsity, is a common characteristic of many proximity graphs. The second assumption, averaged weak spatial mixing, requires that the dependence of spins on the boundary condition in a spatial random graph (accounting for both the Gibbs measure and the randomness of the graph) decays rapidly with respect to the Euclidean distance to the boundary, as the latter tends to infinity. This is an adaptation of the classical weak spatial mixing property for deterministic graphs and typically holds in the high-temperature or low-activity regime of Gibbs models.

In Section 8.2, we combine these two key assumptions to establish a central limit theorem for the total sum of spins. Our central limit theorem accounts for the randomness of both the spin variables and underlying spatial random graph. While the spin variables themselves do not exhibit a stopping set property, the combination of averaged weak spatial mixing and graph stabilization guarantees the BL cluster-localization of the spins. Though establishing sufficient conditions for averaged weak spatial mixing is a distinct challenge, there have been notable contributions of Dobrushin and Shlosman [42], Weitz [154], Sinclair et al. [143], and Regts [130], among others. In Sections 8.3 and 8.4, we employ two complementary approaches providing examples of spin models satisfying the averaged weak spatial mixing condition.

The first is the disagreement percolation method introduced by van den Berg and Maes [151], which is applicable to a wide range of spin models and is detailed in Section 8.3.1. The required disagreement percolation bounds may be obtained using path-counting arguments involving the (averaged) ‘connective constant’ as described in Section 8.3.1 whereas for random graphs exhibiting sharp phase transitions, we may use more direct arguments from continuum percolation theory; see Section 8.3.2. The use of sharp phase transition offers two key advantages. First, it applies to graphs with unbounded degrees and second, since the sharp phase transition assumption is, in principle, weaker than a bound on the (quenched and even averaged) connective constant, it has the potential to yield central limit theorems for a wider range of temperatures or activity parameters. We apply this to the Poisson Gilbert graph in Section 8.3.3.

The second approach to showing averaged weak spatial mixing utilizes a combinatorial technique developed by Weitz [154] and Sinclair et al. [143], specifically designed to show a stronger (and quenched) spatial mixing for Gibbs models and is illustrated with an application to the hard-core model in Section 8.4. For these analyses, the range of the temperature or activity parameter is linked

to the (quenched) connective constant, a concept adapted from Sinclair et al. [143] to measure how the degree distribution of a graph influences spatial mixing.

We refer the reader to Section 1.2 for more on the existing literature of limit theory for spin systems, which is mainly restricted to lattices, save for [65, 66, 28], which treats locally-tree like random graphs with local weak limits.

## 8.1 The spin model

Let  $\mathbb{K}$  be a countably infinite subset of  $\mathbb{R}$ , which we call the spin space. Given a point set of sites (nodes) denoted by  $\mathcal{X} := \{x_1, \dots, x_m\}$  and a graph  $\mathcal{G}(\mathcal{X})$ , we associate a  $\mathbb{K}$ -valued random variable to each site, called the *spin variable*. The spin variables  $[v]_1^m := (v_1, \dots, v_m)$  of points  $[x]_1^m := (x_1, \dots, x_m)$  are chosen randomly, with joint probability proportional to

$$\pi_{[x]_1^m}([v]_1^m) := \exp \left[ \beta \sum_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \Psi(v_i, v_j) + \gamma \sum_{x_i \in \mathcal{X}} \Phi(v_i) \right],$$

where  $\beta \in (0, \infty)$ ,  $\gamma \in \mathbb{R}$ ,  $\Psi : \mathbb{K} \times \mathbb{K} \rightarrow [-\infty, \infty)$  is the symmetric *pairwise potential* and where  $\Phi : \mathbb{K} \rightarrow \mathbb{R}$  is the *external field*. More formally, given  $\mathcal{G}(\mathcal{X})$ , we consider the collection of  $\mathbb{K}$ -valued random variables  $\{V_x\}_{x \in \mathcal{X}} := \{V_x(\mathcal{G}(\mathcal{X}))\}_{x \in \mathcal{X}}$  distributed according to the following probability measure  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}$  on the configuration space  $\mathbb{K}^{\mathcal{X}}$

$$\mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{x_1} = v_1, \dots, V_{x_m} = v_m) = \frac{\pi_{[x]_1^m}([v]_1^m)}{Z(\mathcal{X})} \quad (8.1)$$

where  $Z(\mathcal{X}) := Z(\mathcal{G}(\mathcal{X}))$  is the *normalizing constant*

$$Z(\mathcal{X}) := \sum_{[v]_1^m := (v_1, \dots, v_m) \in \mathbb{K}^m} \pi_{[x]_1^m}([v]_1^m).$$

The probability measure  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}$  at (8.1) is the *Gibbs measure* (or the *specification of the spin model*) on  $\mathcal{G}(\mathcal{X})$ , and the collection of  $\mathbb{K}$ -valued random variables  $\{V_x\}_{x \in \mathcal{X}}$  distributed as  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}$  is the *spin configuration*. The probability measure  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}$  exists whenever  $Z(\mathcal{X}) > 0$ . In the language of statistical physics,  $\beta$  represents the inverse temperature. We refer to Friedli and Velenik [60] and Duminil-Copin [48] for more details on such models. Suitable choices of  $\mathbb{K}$ ,  $\Psi$  and  $\Phi$  yield the hard-core and soft-core models, the Widom-Rowlinson model, the random cluster model, the Ising model and the Potts model (see [48, Section 1.2]). Examples are in Sections 8.3 and 8.4.

Spatial mixing controls the correlations between spin values. Additionally, we need to control the interaction graph, as it impacts the spatial mixing bounds. This is achieved by requiring that the interaction graph has a stabilizing property, specifically by assuming that the radii of stabilization for vertex neighborhoods are not too large. We now formally introduce both the notions of stabilizing interaction graphs and spatial mixing on such graphs, and then, in Section 8.2, we state our main central limit theorem for the random measures  $\mu_n^{\mathcal{G}} := \sum_{x \in \mathcal{P}_n} V_x(\mathcal{G}(\mathcal{P}_n)) \delta_{n^{-1/d_x}}$  induced by the spin systems.

### 8.1.1 Stabilizing interaction graphs

In this section we explain how our random graphs are constructed on the input (ground) process  $\mathcal{P}$  within finite windows.

For any finite point set  $\mu \in \hat{\mathcal{N}}_{W_n}$ , with  $n \in \mathbb{N}$ , let  $\mathcal{G}(\mu)$  be a graph constructed on  $\mu$ , where the edges are determined by a deterministic rule  $\sim$  that takes the entire point set  $\mu$  into account. (Thus  $\mathcal{G}(\mu)$  encodes the interactions between sites in the spin system (8.1), putting an edge between sites if they may potentially interact.) A deterministic construction of the graph  $\mathcal{G}(\mu)$  on  $\mu$  is a modelling restriction. It excludes, for example, graphs built on a Boolean model with random grains or those from a random connection model. In Section 12, we discuss how this framework can be extended to include such graphs.

In order to introduce the notion of stabilization for these graphs (as well as to address the measurability issues of these objects), it is customary to represent the neighbors of a point  $x \in \mu$  within  $\mathcal{G}(\mu)$  as a subset  $N_x(\mu) \subset \mu$ , where  $N_x(\mu) := \{y \in \mu : y \stackrel{\mathcal{G}(\mu)}{\sim} x\}$ . For the sake of translation invariance (if assumed), we center these finite neighborhood-sets as finite counting measures  $\bar{N}_x(\mu) := \sum_{y \in N_x(\mu)} \delta_{y-x} \in \hat{\mathcal{N}}_{\mathbb{R}^d}$  and treat them as auxiliary marks of the ground points  $x \in \mu$ . In essence, when assuming the existence of a graph  $\mathcal{G}(\mu)$  together with its neighborhood structure, we view this as an auxiliary marking  $\{\delta_{(x, \bar{N}_x(\mu))}\}_{x \in \mu}$  of the ground measure  $\mu$ , where the marks satisfy  $\bar{N}_x(\mu) + x \subset \mu$  for all  $x \in \mu$  (neighborhood relation), and  $y \in \bar{N}_x(\mu)$  if and only if  $-y \in \bar{N}_{x+y}(\mu)$  (undirected).

A consequence of assuming a deterministic rule for constructing the graph  $\mathcal{G}(\mu)$  on  $\mu$  is that we may view  $\bar{N}(\cdot) : W_n \times \hat{\mathcal{N}}_{\mathbb{R}^d} \rightarrow \hat{\mathcal{N}}_{\mathbb{R}^d}$  as a marking function in the sense of Definition 4.1 (though here the ground process has no marks). More formally, for us an (undirected) graph corresponds to a *neighborhood marking function*  $\bar{N}$  as above satisfying the two conditions of neighborhood relation and undirectedness and thus measurability of the map  $\mu \mapsto \mathcal{G}(\mu)$  is equivalent to measurability of the marking function  $\bar{N}$ .

Consequently, we define the *interaction range*  $S_n : W_n \times \hat{\mathcal{N}}_{W_n} \rightarrow \mathbb{N}$  of  $x \in \mu \in \hat{\mathcal{N}}_{W_n}$  in the graph  $\mathcal{G}(\mu)$  as the radius of stabilization of this neighborhood marking  $\bar{N}$ , as per Definition 4.5, with

$$S_n(x, \mu) := R_{W_n}^{\bar{N}}(x; \mu). \quad (8.2)$$

By this definition, all graphs  $\mathcal{G}(\mu)$  on  $\mu \in \hat{\mathcal{N}}_{W_n}$  have bounded interaction range  $S_n \leq \lceil \text{diam}(W_n) + 1 \rceil$ . However, we are specifically interested in graphs whose interaction ranges  $S_n(x, \mathcal{P}_n)$  decay rapidly and uniformly in  $x \in W_n, n \in \mathbb{N}$ . This concept, which by the way, implies asymptotic sparsity of the family of graphs  $(\mathcal{G}(\mathcal{P}_n))_{n \in \mathbb{N}}$ , is formalized as follows (in keeping with the general Definition 4.6(ii) of stabilizing marking functions on finite windows):

**Definition 8.1** (Stabilizing graph  $\mathcal{G}$  on the input process  $\mathcal{P}$ ). *We say that a family of graphs  $(\mathcal{G}(\mathcal{P}_n))_{n \in \mathbb{N}}$  (where each graph is generated by the same rule  $\mu \mapsto \mathcal{G}(\mu)$  represented by the marking neighborhood  $\bar{N}$ ) forms a stabilizing interaction graph  $\mathcal{G}$  on finite windows with respect to a point process  $\mathcal{P}$  if for every  $p \in \mathbb{N}$ , there exist a fast decreasing function  $\varphi'_p$  (see (2.1)), such that the interaction range satisfies:*

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[\mathbb{Z}]^p} (S_n(x_1, \mathcal{P}_n) > s) \leq \varphi'_p(s), \quad s > 0. \quad (8.3)$$

Without loss of generality, we assume that  $\varphi'_p(\cdot) \leq \varphi'_{p+1}(\cdot)$  for  $p \in \mathbb{N}$ .

Examples of stabilizing graphs include the Gilbert disc graph (geometric graph), whose range  $S_n$  is bounded by the greatest integer in the sum of 1 plus the disc radius, for all  $n \in \mathbb{N}$ . Proximity graphs from computational geometry, including those defined by  $k$  nearest neighbors, Voronoi tessellations, or spheres of influence, are stabilizing with respect to homogeneous Poisson point processes and stationary  $\alpha$ -determinantal point processes. This is spelled out in Appendix A.

We conclude this section by listing several properties of the interaction ranges of graphs (regardless of their probabilistic stabilization (8.3)). These properties will be used in various places when analyzing stabilizing graphs.

Assuming  $x \in \mu \in \hat{\mathcal{N}}_{W_n}$ , the neighborhood of  $x$  is determined by the points in the ball  $B_{S_n(x,\mu)}(x)$  of radius  $S_n$  around it, i.e.,

$$y \sim x \text{ in } \mathcal{G}(\mu) \text{ iff } y \sim x \text{ in } \mathcal{G}(\mu \cap B_{S_n(x,\mu)}(x)). \quad (8.4)$$

Moreover, the ball  $B_{S_n(x,\mu)}(x)$  is a stopping set, i.e., for all  $s \geq 0$ , we have

$$\{\mu \in \hat{\mathcal{N}}_{W_n} : S_n(x, \mu) \leq s\} = \{\mu \in \hat{\mathcal{N}}_{W_n} : S_n(x, \mu \cap B_s(x)) \leq s\}, \quad (8.5)$$

or, equivalently,

$$S_n(x, \mu) = S_n\left(x, (\mu \cap B_{S_n(x,\mu)}(x)) \cup (\nu \cap B_{S_n(x,\mu)}^c(x))\right) \quad (8.6)$$

for all  $\mu, \nu \in \hat{\mathcal{N}}_{W_n}$ . Finally, note that (8.4) implies

$$x \sim y \text{ in } \mathcal{G}(\mu) \quad \Rightarrow \quad |x - y| \leq \min\{S_n(x, \mu), S_n(y, \mu)\}. \quad (8.7)$$

**Remark 8.2** (Consistency of finite graphs). In general, the stabilizing property (8.3) (defined only for finite windows) does not imply the consistency of the family of graphs  $\mathcal{G}(\mathcal{P}_n)$ ,  $n \in \mathbb{N}$ , in the sense of the existence of some graph  $\mathcal{G}$  on the entire input process  $\mathcal{P}$  (e.g. the Gilbert disc graph or Voronoi tessellation) such that  $\mathcal{G}(\mathcal{P}_n)$  suitably converges to it. Note that our rule  $\mu \mapsto \mathcal{G}(\mu)$  applies only to finite inputs  $\mu$ , and consistently extending this rule to infinite inputs  $\mu \in \mathcal{N}_{\mathbb{R}^d}$  is possible under the additional condition of uniform boundedness of the interaction ranges, i.e.,  $\limsup_{n \rightarrow \infty} S_n(x, \mathcal{P}_n) < \infty$  almost surely for all  $x \in \mathcal{P}$ . See remark (ii) in Section 5.4 for a similar extension of marking functions defined on finite windows and possessing stopping set stabilization on finite windows.

### 8.1.2 Averaged weak spatial mixing of spins

Dating back to the work of Dobrushin and Shlosman [43], spatial mixing properties have been used to characterize the decay of correlations between spins induced by the Gibbs measure (8.1) on a graph  $\mathcal{G}(\mathcal{X})$  or a family of graphs. Among the several formulations of this concept, we adopt a relatively weak version and formulate it for spin models on a random stabilizing graph  $\mathcal{G}$  on  $\mathcal{P}$ .

For subsets  $\mathcal{X}'' \subset \mathcal{X}' \subset \mathcal{X}$ , by  $\mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'})$  we mean the law of the spins on  $\mathcal{X}''$  conditioned on the spin configuration  $v_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}$  on  $\partial\mathcal{X}'$ , assuming that  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}) > 0$ . Here,  $\partial\mathcal{X}' := \partial_{\mathcal{X}}\mathcal{X}'$  is the set of all elements in  $\mathcal{X} \setminus \mathcal{X}'$  that are at unit graph distance from  $\mathcal{X}'$  and

if  $\partial\mathcal{X}' = \emptyset$ , we adopt the convention that  $\mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}) = \mathcal{L}(V_{\mathcal{X}''})$ . When considering conditional laws, we implicitly assume that the boundary condition is non-degenerate as above i.e.,  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}) > 0$ . In general, when referring to boundary conditions we mean specifying spin values on a set  $\mathcal{Z} \subset \mathcal{X}$ .

Recall the *total variation distance*  $d_{\text{TV}}$  between probability measures, as defined in Section 2.1. *Weak spatial mixing* of the Gibbs specification (8.1) on  $\mathcal{G}(\mathcal{X})$  says that two different boundary conditions  $v_{\partial\mathcal{X}'}, z_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}$  on  $\partial\mathcal{X}'$  result in two distributions of spins on  $\mathcal{X}''$  which differ in the total variation distance by a fast decaying function of the graph distance  $d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}')$  between  $\mathcal{X}''$  and  $\partial\mathcal{X}'$ . In what follows, we will primarily take  $\mathcal{X}'$  to be the union of graph balls centered at the points in  $\mathcal{X}''$ . Specifically, for  $x \in \mathcal{X}$  and  $k \in \mathbb{N}$ , the  $\mathcal{G}(\mathcal{X})$ -graph ball of radius  $k$  is

$$\mathfrak{B}_k(x) := \{y \in \mathcal{X} : d_{\mathcal{G}(\mathcal{X})}(x, y) \leq k\},$$

where  $d_{\mathcal{G}(\mathcal{X})}(x, y)$  is the graph distance between  $x$  and  $y$ . For a subset  $\mathcal{X}'' \subset \mathcal{X}$ , define the union of these graph balls by

$$\mathfrak{B}_k(\mathcal{X}'') := \bigcup_{x \in \mathcal{X}''} \mathfrak{B}_k(x).$$

Now we give our formal definition of the spatial mixing property for spin models on random graphs.

**Definition 8.3** (Averaged weak spatial mixing of the Gibbs spin model on a random spatial graph). *The Gibbs spin model (8.1) satisfies averaged weak spatial mixing on the graph  $\mathcal{G} = \mathcal{G}(\cdot)$  with respect to the input process  $\mathcal{P}$  if the following conditions hold:*

(i) *for any finite subset  $\mathcal{X} \subset \mathbb{R}^d$ , there exists a spatial mixing function  $\mathcal{A}_{\mathcal{G}(\mathcal{X})} : 2^{\mathcal{X}} \times 2^{\mathcal{X}} \rightarrow [0, \infty)$  such that*

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}) , \mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'})) \leq \mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}'), \quad (8.8)$$

*for all subsets  $\mathcal{X}'' \subset \mathcal{X}' \subset \mathcal{X}$  and boundary conditions  $v_{\partial\mathcal{X}'}, z_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}$ , and*

(ii) *for all  $p \in \mathbb{N}$  the functions*

$$\mathfrak{a}_p(k) := \sup_{B \in \mathcal{B}_b} \sup_{[x]_1^p \subset B} \mathbb{E}_{[x]_1^p} [\mathcal{A}_{\mathcal{G}(\mathcal{P} \cap B)}([x]_1^p, \partial\mathfrak{B}_k([x]_1^p))], \quad k \in \mathbb{N}, \quad (8.9)$$

*are fast decreasing as  $k \rightarrow \infty$ . Here,  $\mathfrak{B}_k([x]_1^p) = \bigcup_{i=1}^p \mathfrak{B}_k(x_i)$  and the boundary  $\partial$  of  $\mathfrak{B}_k([x]_1^p)$  is again with respect to  $\mathcal{G}(\mathcal{P} \cap B)$ ;  $\mathcal{B}_b$  are the bounded Borel subsets of  $\mathbb{R}^d$ .*

*We assume that the mapping  $((\mathbb{R}^d)^p, \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathcal{N}_{\mathbb{R}^d}}) \ni ([x]_1^p, \tilde{\mu}) = ([x]_1^p, \mathcal{G}) \mapsto \mathcal{A}_{\mathcal{G}}([x]_1^p, \partial^{\mathcal{G}} \mathfrak{B}_k([x]_1^p)) \in \mathbb{R}$  are measurable for all  $k \in \mathbb{N}$ . Without loss of generality, we assume that  $\mathfrak{a}_p$  are increasing in  $p$ .*

Condition (i) is a way of controlling the decay of spin correlations whereas condition (ii) ensures fast decay with respect to graph distance in expectations. More precisely, the function  $\mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}')$  in (8.8) measures the sensitivity of the spin distribution at  $\mathcal{X}''$  to the boundary conditions of the Gibbs measure at  $\partial\mathcal{X}'$  and the property of spatial mixing requires the sensitivity to diminish rapidly with

respect to the graph distance between  $\mathcal{X}''$  and  $\partial\mathcal{X}'$ . Typically, in the setting of deterministic graphs such as lattices, this decay is assumed to be exponential i.e.,

$$A_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}') \leq C|\mathcal{X}''||\partial\mathcal{X}'| \exp(-a \cdot d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}')),$$

for some  $C, a \in (0, \infty)$ ; see [106, (2.5)]. Often this is also taken to be the definition of weak spatial mixing in such settings. While it would have been convenient to assume such an exponential bound a.s. for spatial random graphs, we have opted in (8.9) for a weaker version, one averaging the weak spatial mixing constant over the point process. Hence we call this *averaged weak spatial mixing* and this suffices to establish our general central limit theorem for Gibbs models on spatial random graphs.

This weaker condition brings additional flexibility and enlarges the scope of applications. Indeed, as seen in examples in Section 8.3, weak spatial mixing bounds are derived via connective constant or site percolation probability bounds; see Sections 8.3.1 and 8.3.2. Requiring such graph parameters to be almost surely well-controlled is sometimes very restrictive, possibly requiring almost sure degree bounds on the graph. However, our requirement of averaged weak spatial mixing imposes control on the connective constant (also in its averaged version) or site percolation probability only in expectation and hence it is easier to establish this condition for a wider class of spin models and spatial random graph models as well as for a larger range of temperatures.

In Section 8.4, using an example of hard-core spin model, we will show how a stronger version of spatial mixing for Gibbs models, accounting for the exact location of the differences between the boundary conditions, may allow one to extend the range of model parameters under which the central limit theorem holds for specific Gibbs models.

As highlighted in the introduction, our definition of spatial mixing is inspired by the notion of mixing introduced in Dobrushin and Shlosman [43], a cornerstone in the study of spin systems. For infinite configurations ( $\text{card}(\mathcal{X}) = \infty$ ) and finite spin spaces  $\mathbb{K}$ , weak spatial mixing guarantees the existence of a unique Gibbs measure on the graph configuration  $\mathcal{G}(\mathcal{X})$ ; see [154, Proposition 2.2]. For more on different notions of spatial mixing and their implications, see [106, Section 2].

## 8.2 Main result for spin models

We state a general central limit theorem for spatial spin models satisfying averaged weak spatial mixing. Recall that  $\rho_{[x]_1^p}^{(1)}$  are the correlation functions of order 1 under reduced Palm distributions  $\mathbb{P}_{[x]_1^p}^!$ ; see Section 2.1.

**Theorem 8.4** (CLT for the sum of spins). *Let  $\mathcal{P}$  be a simple point process on  $\mathbb{R}^d$  with summable exponential mixing correlations as in Definition 4.8 and bounded reduced Palm intensity function of the ground process  $\mathcal{P}$ , i.e., for all  $p \in \mathbb{N}$*

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \sup_{y \in \mathbb{R}^d} \rho_{[x]_1^p}^{(1)}(y) \leq \hat{\kappa}_p < \infty. \quad (8.10)$$

*Let  $\mathcal{G} = \mathcal{G}(\cdot)$  be a stabilizing interaction graph on finite windows with respect to  $\mathcal{P}$  as in Definition 8.1. Assume the Gibbs specification (8.1) satisfies averaged weak spatial mixing on  $\mathcal{G}(\mathcal{P})$ , as in*

**Definition 8.3.** Denote by  $\{V_x\}_{x \in \mathcal{P}_n} = \{V_x(\mathcal{G}(\mathcal{P}_n))\}_{x \in \mathcal{P}_n}$  the spin configuration on  $\mathcal{P}_n$  assuming for all  $n \in \mathbb{N}$  that  $Z(\mathcal{P}_n) > 0$ . Assume moreover for all  $p > 1$  the following moment condition

$$\sup_{1 \leq n < \infty} \sup_{q \leq p} \sup_{[x]_1^l \in W_n} \mathbb{E}_{[x]_1^q} [\max(1, |V_{x_1}(\mathcal{G}(\mathcal{P}_n))|^p)] < \infty. \quad (8.11)$$

Put

$$\mu_n^{\mathcal{G}} := \sum_{x \in \mathcal{P}_n} V_x \delta_{n^{-1/d}x}.$$

If  $f \in \mathbb{B}(W_1)$  satisfies  $\text{Var} \mu_n^{\mathcal{G}}(f) = \Omega(n^\nu)$  for some  $\nu > 0$ , then as  $n \rightarrow \infty$

$$(\text{Var} \mu_n^{\mathcal{G}}(f))^{-1/2} (\mu_n^{\mathcal{G}}(f) - \mathbb{E}[\mu_n^{\mathcal{G}}(f)]) \xrightarrow{d} Z.$$

We shall deduce this result from Theorem 5.2 in the next section. Before doing so, we comment on extensions and refinements.

**Remark 8.5** (Mean and variance asymptotics and weaker moment conditions).

(i) If the input process  $\mathcal{P}$  is stationary and the interaction rule  $\mu \mapsto \mathcal{G}(\mu)$  is translation invariant, then under the assumptions of Theorem 8.4, with the moment condition (8.11) satisfied only for  $p = 2 + \epsilon$  for some  $\epsilon > 0$ , then mean asymptotics for  $n^{-1} \mathbb{E}[\mu_n^{\mathcal{G}}(f)]$  and variance asymptotics for  $n^{-1} \text{Var}[\mu_n^{\mathcal{G}}(f)]$  can be expressed as in (5.5) and (5.6), as stated in Proposition 5.3. These results involve Palm distributional limits, given in (5.3) and (5.4), for  $V_x(\mathcal{G}(\mathcal{P}_n))$  and  $(V_{x_1}(\mathcal{G}(\mathcal{P}_n)), V_{x_2}(\mathcal{G}(\mathcal{P}_n)))$ , without requiring the existence of an infinite spin system consistent with  $(\mathcal{G}(\mathcal{P}_n))_{n \in \mathbb{N}}$ . The proof of the central limit theorem in Theorem 8.4 relies on BL cluster-localization for spins as defined in Definition 4.3, see Lemma 8.7. BL cluster-localization is also sufficient to derive the mean and variance asymptotics of Proposition 5.3; see Remark (iv) in Section 5.4 and Remarks 7.6, and 7.8.

(ii) Still assuming stationarity of the model (both the input process and the graph) as above, the central limit theorem for  $(\mu_n^{\mathcal{G}})_{n \in \mathbb{N}}$ , as formulated in Theorem 8.4, holds under only a  $(2 + \epsilon)$ -moment assumption on the spins in (8.11), provided that  $\lim_{n \rightarrow \infty} n^{-1} \text{Var}[\mu_n^{\mathcal{G}}(f)]$  exists and is non-zero. This follows from Theorem 5.4, which moreover extends the result to multivariate spin systems. Many classical spin systems are finite-valued and in these cases, the moment condition is trivially satisfied.

(iii) A general strategy is outlined in [66, Section 1.6] for deriving more explicit expectation and variance asymptotics in terms of susceptibility and pressure but, as noted there, this strategy is more challenging in the averaged quenched setting which corresponds to our framework.

(iv) It is sometimes easier to derive volume-order variance lower bounds in ‘certain spin systems’ than to establish limiting variance asymptotics. For example, using the approach as in [41, Section 3.4.3 of v2], one can derive variance lower bounds for ferromagnetic or positively associated spin systems.

(v) For more general spin systems, one derive variance bounds as follows. Suppose  $\mathcal{I}_n \subset \mathcal{P}_n$  is an independent set of vertices in the graph  $\mathcal{G}(\mathcal{P}_n)$ , i.e.,  $x, z \in \mathcal{I}_n$  implies that  $x \not\sim z$ . Using the conditional

variance formula and the spatial Markov property twice, we derive that

$$\begin{aligned} \text{Var}\left(\sum_{x \in \mathcal{P}_n} V_x\right) &\geq \mathbb{E}\left(\text{Var}\left(\sum_{x \in \mathcal{I}_n} V_x \mid V_y, y \in I_n^c\right)\right) = \mathbb{E}\left(\sum_{x \in \mathcal{I}_n} \text{Var}(V_x \mid V_y, y \in I_n^c)\right) \\ &= \mathbb{E}\left(\sum_{x \in \mathcal{I}_n} \text{Var}(V_x \mid V_y, y \neq x)\right). \end{aligned}$$

Next, we find an independent set  $\mathcal{I}_n$  such that for a positive constant  $c$ , a.s. for all  $x \in \mathcal{I}_n$  we have  $\text{Var}(V_x \mid V_y, y \neq x) \geq c$ . This can be established under conditions on the spin system and possibly requiring uniform bound on degrees of  $\mathcal{I}_n$ ; for example, see [41, Lemma 3.6 of v3]. Then we obtain

$$\text{Var}\left(\sum_{x \in \mathcal{P}_n} V_x\right) \geq c \mathbb{E}(|\mathcal{I}_n|).$$

One may show that  $\mathbb{E}(|\mathcal{I}_n|) \geq c'n$  with  $c' > 0$  for many stabilizing interaction graphs, including, for example, bounded degree graphs such as  $k$ -nearest neighbor graphs or graphs with bounded interaction range such as random geometric graph. Combining the above observations, we obtain volume order variance lower bounds for certain spatial spin systems.

(vi) Averaged weak spatial mixing is here formulated with respect to the graph distance in order to exploit spatial mixing results on deterministic graphs; we may then transfer this to 'spatial mixing with respect to the Euclidean distance' via the stabilization of the interactions graphs. However in certain applications it may be easier to directly establish spatial mixing in the Euclidean distance; see Section 8.3.2.

(vii) Later, in Remark 9.4, we indicate extensions to spatial random graphs models with additional randomness.

(viii) While the measure  $\mu_n^{\mathcal{G}}(\cdot)$  may be interpreted as the *empirical magnetization*—even if this is not immediately evident—we expect that our strategy of the proof of Theorem 8.4 could also be applied to derive limit theorems for *internal energy* and *susceptibility*; see [152, Chapter 5].

### 8.2.1 Proof of CLT for spin models

The specification (8.1) of the laws of the spins  $V_x = V_x(\mathcal{G}(\mathcal{P}_n))$ ,  $x \in \mathcal{P}_n$ , given  $\mathcal{P}_n$  does not involve any particular marking function  $\xi$ . Without such a representation, a direct application of our umbrella Theorem 3.3 to prove a CLT would require verifying fast BL-mixing correlations of these spins on the input process  $\mathcal{P}$  in the sense of Definition 3.1, which appears to be a delicate task.

Instead, we introduce auxiliary i.i.d. pre-marks on  $\tilde{\mathcal{P}}$  and define a *spin marking function* on this extended input process, which is used to construct spins whose joint laws are faithful to the given Gibbs specification. It is clear that, in general, any such marking function cannot satisfy stopping-set stabilization, nor even standard  $L^q$ -stabilization. Consequently, our approach relies crucially on a purely distributional notion of localization—namely bounded Lipschitz localization—of the constructed marking function, and more precisely on BL cluster-localization, developed in Section 4.2.

Establishing this property for our marking function leverages the averaged weak spatial mixing assumption of the spin model on the given input process. As a result, the proof of the CLT for the sum of spins stated in Theorem 8.4 follows as a direct application of Theorem 5.2 for BL cluster-localizing marks.

A more precise working plan for this section is as follows. First, we present Lemma 8.6, which establishes a link between averaged weak spatial mixing (holding almost surely with respect to the input process  $\mathcal{P}$ ) and BL *union-localization* of spins, an analogue of condition (4.6), with localization measured in terms of the graph distance. Next, Lemma 8.7 establishes BL cluster-localization, under the assumption of independence between well-separated clusters in (4.7). We then introduce a spin marking function and show in Lemma 8.8 that it is consistent with the Gibbs specification and, moreover, satisfies the independence property between well-separated clusters assumed above.

This result allows us to conclude the proof of Theorem 8.4 by an application of Theorem 5.2. Finally, auxiliary lemmas—some of which are used in this programme are stated and proved in to Section 8.2.2.

For  $x \in \mathcal{X}$  and  $m \in \mathbb{N}$  recall the notation  $\mathfrak{B}_m(x) := \{y \in \mathcal{X} : d_{\mathcal{G}(x)}(x, y) \leq m\}$  for the  $\mathcal{G}(\mathcal{X})$ -graph ball of radius  $m$ , and their union centered at points in  $\mathcal{X}'' \subset \mathcal{X}$ , namely  $\mathfrak{B}_m(\mathcal{X}'') = \bigcup_{x \in \mathcal{X}''} \mathfrak{B}_m(x)$ . We denote by  $\mathcal{G}_{\mathcal{X}''}^{\cup[m]}(\mathcal{X}) := \mathcal{G}(\mathfrak{B}_m(\mathcal{X}''))$  the *induced subgraph* of  $\mathcal{G}(\mathcal{X})$  on  $\mathfrak{B}_m(\mathcal{X}'')$  and we let  $V_{\mathcal{X}''}^{\cup[m]} := V_{\mathcal{X}''}(\mathcal{G}_{\mathcal{X}''}^{\cup[m]}(\mathcal{X}))$  be the spins with respect to the Gibbs measure on  $\mathcal{G}(\mathfrak{B}_m(\mathcal{X}''))$ .

To emphasize that the  $d_{\text{BL}}$  distance is a property of the laws of the random variables, we shall use here the  $d_{\text{BL}}$  distance between laws of random variables rather than random variables themselves.

**Lemma 8.6** (Averaged weak spatial mixing implies fast BL union-localization of spins). *The spatial mixing function  $\mathcal{A}_{\mathcal{G}(x)}$  at (8.8) controlling the decay of spin correlations induced by the Gibbs measure (8.1) on the graph  $\mathcal{G}(\mathcal{X})$  implies for all  $\mathcal{X}'' \subset \mathcal{X}$  and  $m \in \mathbb{N}$*

$$d_{\text{BL}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup[m]})) \leq d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup[m]})) \leq \mathcal{A}_{\mathcal{G}(x)}(\mathcal{X}'', \partial\mathfrak{B}_{m-1}(\mathcal{X}'')). \quad (8.12)$$

*Proof.* The first inequality follows since  $d_{\text{BL}}$  is bounded by  $d_{\text{TV}}$ . For the second inequality, put  $\mathcal{X}' = \mathfrak{B}_{m-1}(\mathcal{X}'')$ , the  $\mathcal{G}(\mathcal{X})$ -graph ball of radius  $m - 1$  around  $\mathcal{X}''$  and so  $\partial\mathcal{X}' = \partial\mathfrak{B}_{m-1}(\mathcal{X}'')$ . If  $\partial\mathcal{X}' = \emptyset$  (when  $\mathfrak{B}_{m-1}(\mathcal{X}'') = \mathfrak{B}_m(\mathcal{X}'')$ ) then  $d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup[m]})) = 0$ . Otherwise, considering probabilities  $\mathbb{P} = \mathbb{P}_{\mathcal{G}(\mathcal{X})}$  and  $\mathbb{P}' = \mathbb{P}_{\mathcal{G}_{\mathcal{X}''}^{\cup[m]}(\mathcal{X})}$  of the Gibbs model at (8.1), respectively, on  $\mathcal{G}(\mathcal{X})$  and  $\mathcal{G}_{\mathcal{X}''}^{\cup[m]}(\mathcal{X})$ , we have

$$\begin{aligned} & d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup[m]})) \\ &= \sum_{v_{\mathcal{X}''} \in \mathbb{K}^{\mathcal{X}''}} \left| \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \left( \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) - \mathbb{P}'(V_{\mathcal{X}''}^{\cup[m]} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'}^{\cup[m]} = w_{\partial\mathcal{X}'}) \right) \right. \\ & \quad \left. \times \mathbb{P}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}'(V_{\partial\mathcal{X}'}^{\cup[m]} = w_{\partial\mathcal{X}'}) \right| \\ &= \sum_{v_{\mathcal{X}''} \in \mathbb{K}^{\mathcal{X}''}} \left| \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \left( \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) - \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = w_{\partial\mathcal{X}'}) \right) \right. \\ & \quad \left. \times \mathbb{P}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}'(V_{\partial\mathcal{X}'}^{\cup[m]} = w_{\partial\mathcal{X}'}) \right| \\ &\leq \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \sum_{v_{\mathcal{X}''} \in \mathbb{K}^{\mathcal{X}''}} \left| \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) - \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = w_{\partial\mathcal{X}'}) \right| \end{aligned}$$

$$\times \mathbb{P}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}'(V_{\partial\mathcal{X}'}^{\cup[m]} = w_{\partial\mathcal{X}'})$$

$$\leq \mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}'),$$

where we used the consistency of conditional spin distributions (Lemma 8.10) in the second equality and the averaged weak spatial mixing in the last inequality.  $\square$

We introduce some notation for the forthcoming crucial lemma, which establishes BL cluster-localization of our spin marking function on stabilizing random graphs. The vector  $(V_{x_1}(\mathcal{G}(\mathcal{X})), \dots, V_{x_p}(\mathcal{G}(\mathcal{X})))$  is abbreviated by  $V_{[x]_1^p}(\mathcal{G}(\mathcal{X}))$ , where  $x_i \in \mathcal{X}$ ,  $i = 1, \dots, p$ . In the following lemma, we will take  $\mathcal{X}$  to be either the entire input process  $\mathcal{P}_n$  in the window  $W_n$ , or the process restricted to the union of Euclidean balls  $B_r([x]_a^b) = \bigcup_{i=a}^b B_r(x_i)$ ,  $a, b \in \{1, p\}$ ; i.e.,  $\mathcal{X} = \mathcal{P}_n \cap B_r([x]_a^b)$ .

**Lemma 8.7** (Fast BL cluster-localization of spins on stabilizing graphs). *Let  $\mathcal{P}$  be a simple, stationary point process on  $\mathbb{R}^d$  with bounded (reduced) Palm intensity function, as in (8.10). Let  $\mathcal{G} = \mathcal{G}(\cdot)$  be a stabilizing interaction graph on finite windows with respect to  $\mathcal{P}$ , in the sense of Definition 8.1. Assume that the Gibbs specification (8.1) satisfies averaged weak spatial mixing on  $\mathcal{G}$  with respect to  $\mathcal{P}$ , as in Definition 8.3. Then for all  $p \in \mathbb{N}$  there are fast decreasing functions  $\varphi_p$  and constants such that*

$$\sup_{n \in \mathbb{N}} \sup_{x_1, \dots, x_p \in W_n} d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n)), V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p)))) \leq 2\varphi_p(r), \quad r > 0 \quad (8.13)$$

and for all  $l \in \{1, \dots, p-1\}$  and  $x_1, \dots, x_p \in \mathbb{R}^d$  such that  $0 < r < s([x]_1^l, [x]_{l+1}^p)$ , we have

$$d_{\text{BL}, [x]_1^p} \left( V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))), \left( V_{[x]_1^l}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^l))), V_{[x]_{l+1}^p}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_{l+1}^p))) \right) \right) \leq 2\varphi_p(r), \quad (8.14)$$

where  $V_{[x]_1^l}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^l))), V_{[x]_{l+1}^p}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_{l+1}^p)))$  are conditionally independent (given  $\mathcal{P}_n$ ) spin vectors with laws  $\pi_{[x]_1^l}$  and  $\pi_{[x]_{l+1}^p}$  induced respectively by graphs  $\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^l))$  and  $\mathcal{G}(\mathcal{P}_n \cap B_r([x]_{l+1}^p))$ . The above bound also holds under the stronger total variation distance.

*Proof.* Fix  $x_1, \dots, x_p \in W_n$  and consider  $x_i \in \mathcal{P}_n$ ,  $i = 1, \dots, p$ , with the aim of studying  $\mathcal{P}_n$  under the Palm distribution  $\mathbb{P}_{[x]_1^p}$ . Set  $k_r = \lceil r^b \rceil$  for a fixed  $b \in (0, 1)$  and let  $r \geq 1$  be large such that  $k_r < r/3$ .

Denote by  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^p}^{\cup[k_r]}(\mathcal{P}_n)$  the induced subgraph of  $\mathcal{G}(\mathcal{P}_n)$  on  $\mathfrak{B}_{k_r}([x]_1^p)$ —the union of the  $k_r$ -hop balls around  $x_1, \dots, x_p$  in the  $\mathcal{G}(\mathcal{P}_n)$ -graph distance. Denote by  $\mathcal{G}_{[x]_1^p}^{\cup(r)} = \mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))$  the graph constructed on  $\mathcal{P}_n$  restricted to the union of Euclidean balls  $B_r([x]_1^p)$ . We adopt also the shorthand notation  $\mathcal{G} = \mathcal{G}(\mathcal{P}_n)$ .

We will assume the following probability bounds about stabilizing interaction graphs and complete the proofs of (8.13) and (8.14). After that we will derive these two bounds. For all  $p \in \mathbb{N}$ , there exist fast-decreasing functions  $\varphi_p'''$  such that

$$\sup_{n \in \mathbb{N}} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup(r)} \not\cong \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \} \leq \varphi_p'''(r), \quad (8.15)$$

and for all  $l \in \{1, \dots, p-1\}$  and  $x_1, \dots, x_p \in \mathbb{R}^d$  such that  $0 < r < s([x]_1^q, [x]_{q+1}^p)$ , we have

$$\sup_{n \in \mathbb{N}} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \neq \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]} \} \leq \varphi_p'''(r), \quad (8.16)$$

where  $\dot{\cup}$  denotes disjoint union of graphs. We assume without loss of generality that  $\varphi_p'$  are increasing in  $p$ . The bounds show that with high probability, graph balls are contained in large Euclidean balls and graph balls on well-separated vertices are disjoint.

We consider the vector of spins  $V_{[x]_1^p} = V_{[x]_1^p}(\mathcal{G})$ ,  $V_{[x]_1^p}^{\cup(r)} = V_{[x]_1^p}(\mathcal{G}_{[x]_1^p}^{\cup(r)})$  and  $V_{[x]_1^p}^{\cup[k_r]} = V_{[x]_1^p}(\mathcal{G}_{[x]_1^p}^{\cup[k_r]})$  at  $[x]_1^p \in \mathcal{P}_n$  in the respective graphs.

*Proof of (8.13).* Given the inclusion  $\mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]}$ , we can apply the fact that averaged weak spatial mixing implies fast BL-localization of spins from Lemma 8.6 and we have assumed that the inclusion holds with high Palm probability  $\mathbb{P}_{[x]_1^p}$  in (8.15).

Indeed, we apply the triangle inequality

$$d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}, V_{[x]_1^p}^{\cup(r)}) \leq d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}, V_{[x]_1^p}^{\cup[k_r]}) + d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}^{\cup[k_r]}, V_{[x]_1^p}^{\cup(r)}). \quad (8.17)$$

For the first term, we condition on  $\mathcal{P}_n$  and apply averaged weak spatial mixing via (8.12) with  $\mathcal{X} = \mathcal{P}_n$  and obtain:

$$\begin{aligned} d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}, V_{[x]_1^p}^{\cup[k_r]}) &= \sup_{f \in \text{BL}} \left| \mathbb{E}_{[x]_1^p} \left[ \mathbb{E}[f(V_{[x]_1^p}^{\cup[k_r]}) \mid \mathcal{P}_n] - \mathbb{E}[f(V_{[x]_1^p}) \mid \mathcal{P}_n] \right] \right| \\ &\leq \mathbb{E}_{[x]_1^p} \left[ \sup_{f \in \text{BL}} \left| \mathbb{E}[f(V_{[x]_1^p}^{\cup[k_r]}) - f(V_{[x]_1^p}) \mid \mathcal{P}_n] \right| \right] \\ &\leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} \mid \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p} \mid \mathcal{P}_n)) \right] \\ &\leq \mathbb{E}_{[x]_1^p} [\mathcal{A}_{\mathcal{G}(\mathcal{P}_n)}([x]_1^p, \partial \mathfrak{B}_{k_r}([x]_1^p))] \\ &\leq \mathbf{a}_p(k_r). \end{aligned} \quad (8.18)$$

Recalling  $k_r := \lceil r^b \rceil$ , the last expression is fast decreasing in  $r$  by condition (ii) of Definition 8.3.

Now, for the second term in (8.17), on the event  $\mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]}$ , we again apply the averaged weak spatial mixing property via (8.12), with respect to  $\mathcal{G}_{[x]_1^p}^{\cup(r)} = \mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))$ ; on the complementary event, we use the trivial bound  $d_{\text{BL}} \leq 2$ :

$$\begin{aligned} &d_{\text{BL}, [x]_1^p}(V_{[x]_1^p}^{\cup[k_r]}, V_{[x]_1^p}^{\cup(r)}) \\ &\leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} \mid \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup(r)} \mid \mathcal{P}_n)) \right] \\ &\leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} \mid \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup(r)} \mid \mathcal{P}_n)) \mathbf{1}(\mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]}) \right] + 2\mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \} \\ &\leq \mathbb{E}_{[x]_1^p} \left[ \mathcal{A}_{\mathcal{G}_{[x]_1^p}^{\cup(r)}}([x]_1^p, \partial \mathfrak{B}_{k_r}([x]_1^p)) \mathbf{1}(\mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]}) \right] + 2\mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \} \\ &\leq \mathbb{E}_{[x]_1^p} [\mathcal{A}_{\mathcal{G}_{[x]_1^p}^{\cup(r)}}([x]_1^p, \partial \mathfrak{B}_{k_r}([x]_1^p))] + 2\mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \} \end{aligned} \quad (8.19)$$

$$\leq \mathbf{a}_p(k_r) + 2\mathbb{P}_{[x]_1^p} \{ \mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]} \},$$

where in the third inequality we use (8.12) with  $\mathcal{X} = \mathcal{P}_n \cap B_r([x]_1^p)$  knowing that

$$\mathcal{G}(\mathcal{X}) = \mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^p}^{\cup[k_r]}(\mathcal{P}_n) = \mathcal{G}_{[x]_1^p}^{\cup[k_r]}(\mathcal{X}).$$

The first term, again by condition (ii) of Definition 8.3, is fast decreasing in  $r$  and the second term is fast decreasing in  $r$  by (8.15). Thus, we have proven (8.13).

*Proof of (8.14).* We keep the notation of the previous part regarding the points  $[x]_1^p \in W_n$  and the localization radius  $r > 0$ . Moreover, for  $l \in \{1, \dots, p-1\}$  we assume that  $0 < r < s([x]_1^l, [x]_{l+1}^p)$ . Given  $\mathcal{P}_n$  under the Palm distribution  $\mathbb{P}_{[x]_1^p}$ , in addition to  $V_{[x]_1^p}^{\cup(r)} = V_{[x]_1^p}(\mathcal{G}_{[x]_1^p}^{\cup(r)})$ , we consider the random vectors of spins

$$V_{[x]_1^l}^{\cup(r)} = V_{[x]_1^l}(\mathcal{G}_{[x]_1^l}^{\cup(r)}), \quad V_{[x]_{l+1}^p}^{\cup(r)} = V_{[x]_{l+1}^p}(\mathcal{G}_{[x]_{l+1}^p}^{\cup(r)})$$

defined on the respective graphs

$$\mathcal{G}_{[x]_1^l}^{\cup(r)} := \mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^l)), \quad \mathcal{G}_{[x]_{l+1}^p}^{\cup(r)} := \mathcal{G}(\mathcal{P}_n \cap B_r([x]_{l+1}^p)).$$

Recall that we have assumed that  $V_{[x]_1^l}^{\cup(r)}$  and  $V_{[x]_{l+1}^p}^{\cup(r)}$  are conditionally independent given  $\mathcal{P}_n$ .

To evaluate the distance between the law of  $V_{[x]_1^p}^{\cup(r)}$  and the product law of  $(V_{[x]_1^l}^{\cup(r)}, V_{[x]_{l+1}^p}^{\cup(r)})$ , we observe that the approximation  $V_{[x]_1^p}^{\cup[k_r]}$  of  $V_{[x]_1^p}^{\cup(r)}$ , constructed in the preceding part, also admits a product structure. More precisely, on the event  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}$ , which is assumed to hold with high Palm probability  $\mathbb{P}_{[x]_1^p}$  (see (8.16)), the two groups of spins located at  $[x]_1^l$  and  $[x]_{l+1}^p$  do not interact. Consequently,  $V_{[x]_1^p}^{\cup[k_r]}$  decomposes into two independent components. Moreover, these components provide marginal approximations of  $V_{[x]_1^l}^{\cup(r)}$  and  $V_{[x]_{l+1}^p}^{\cup(r)}$ , respectively. Here are the details.

Substituting (8.15) in (8.19), we conclude that  $V_{[x]_1^p}^{\cup[k_r]}$  approximates  $V_{[x]_1^p}^{\cup(r)}$  and, similarly,  $V_{[x]_1^l}^{\cup[k_r]}$  approximates  $V_{[x]_1^l}^{\cup(r)}$  and likewise  $V_{[x]_{l+1}^p}^{\cup[k_r]}$  approximates  $V_{[x]_{l+1}^p}^{\cup(r)}$ , namely we have the bound

$$\max\{d_{\text{BL}}(V_{[x]_1^l}^{\cup(r)}, V_{[x]_1^l}^{\cup[k_r]}), d_{\text{BL}}(V_{[x]_{l+1}^p}^{\cup(r)}, V_{[x]_{l+1}^p}^{\cup[k_r]}), d_{\text{BL}}(V_{[x]_1^p}^{\cup(r)}, V_{[x]_1^p}^{\cup[k_r]})\} \leq \mathbf{a}_p(k_r) + 2\varphi'_p(r) \quad (8.20)$$

Above, we have assumed without loss of generality that  $\mathbf{a}_p$  and  $\varphi'_p$  are increasing in  $p$ .

Further, on the event  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}$  the two groups of spins

$$(V_{[x]_1^l}^{\cup[k_r]}, V_{[x]_{l+1}^p}^{\cup[k_r]}) := V_{[x]_1^p}^{\cup[k_r]}$$

located at respective sites  $[x]_1^l$  and  $[x]_{l+1}^p$  and generated on the graph  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]}$  are independent. Indeed, here, we construct the spin system on the union of the two components. The resulting independence is an intrinsic property of the spin system itself and follows since the two components of the underlying graph are disconnected on the event  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}$ .

Collecting the above arguments, we have

$$\begin{aligned}
& d_{\text{BL}, [x]_1^p} \left( V_{[x]_1^p}^{\cup(r)}, \left( V_{[x]_1^l}^{\cup(r)}, V_{[x]_{l+1}^p}^{\cup(r)} \right) \right) \\
& \leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^l}^{\cup(r)} | \mathcal{P}_n) \times \mathcal{L}(V_{[x]_{l+1}^p}^{\cup(r)} | \mathcal{P}_n) \right) \right] \\
& \leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n) \right) \right. \\
& \quad \left. + d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^l}^{\cup(r)} | \mathcal{P}_n) \times \mathcal{L}(V_{[x]_{l+1}^p}^{\cup(r)} | \mathcal{P}_n) \right) \right] \\
& \leq \mathbb{E}_{[x]_1^p} \left[ \left\{ d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n) \right) \right. \right. \\
& \quad \left. \left. + d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n) \times \mathcal{L}(V_{[x]_{l+1}^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^l}^{\cup(r)} | \mathcal{P}_n) \times \mathcal{L}(V_{[x]_{l+1}^p}^{\cup(r)} | \mathcal{P}_n) \right) \right\} \right. \\
& \quad \left. \mathbf{1}(\mathcal{G}_{[x]_1^p}^{\cup[k_r]} = \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \cup \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}) \right. \\
& \quad \left. + 2\mathbb{P}_{[x]_1^p}(\{\mathcal{G}_{[x]_1^p}^{\cup[k_r]} \neq \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \cup \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}\}) \right] \\
& \leq \mathbb{E}_{[x]_1^p} \left[ \left\{ d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n) \right) + d_{\text{BL}} \left( \mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^l}^{\cup(r)} | \mathcal{P}_n) \right) \right. \right. \\
& \quad \left. \left. + d_{\text{BL}} \left( \mathcal{L}(V_{[x]_{l+1}^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_{l+1}^p}^{\cup(r)} | \mathcal{P}_n) \right) \right\} + 2\mathbb{P}_{[x]_1^p}(\{\mathcal{G}_{[x]_1^p}^{\cup[k_r]} \neq \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \cup \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}\}) \right] \\
& \leq 3\alpha_p(k_r) + 8\varphi_p'''(r),
\end{aligned}$$

where in the last equality we apply (8.20) and (8.16). This justifies the validity of (8.14) and completes the proof of Lemma 8.7 assuming (8.15) and (8.16).

*Proof of (8.15) and (8.16).* We control the event in (8.15) by a larger event and then observe that this larger event also controls the event in (8.16). Then, we derive probability bounds for the larger event to derive both bounds.

Towards this, first observe that

$$\{\mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supseteq \mathcal{G}_{[x]_1^p}^{\cup[k_r]}\} \subset \bigcup_{i=1}^p \{\mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supseteq \mathcal{G}_{x_i}^{+[k_r]}\}, \quad (8.21)$$

where  $\mathcal{G}_{x_i}^{+[k_r]}$  represents the induced subgraph of  $\mathcal{G}(\mathcal{P}_n)$  on  $\mathfrak{B}_{k_r}(x_i)$ , including additionally the edges between points in  $\mathfrak{B}_{k_r}(x_i)$  and  $\mathfrak{B}_{k_r}(x_j)$ ,  $j \neq i$ , that exist in  $\mathcal{G}(\mathcal{P}_n)$  and therefore in  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]}$ . (These edges would otherwise not be included if we simply took the union of edges within each  $\mathfrak{B}_{k_r}(x_i)$ , for  $i = 1, \dots, p$ .)

We control the last event in (8.21) for  $x_1$  and note the same approach works for all  $x_1, \dots, x_p$ . We control the last term for  $x_1$  by exploiting the fact that if a path originating from  $x_1$  of length  $k_r$  in the graph distance is not contained within  $\mathcal{G}_{[x]_1^p}^{\cup(r)}$ , then at least one point in  $\mathcal{P}_n \cap B_r(x_1)$  must have an interaction range which is ‘very large’ in the Euclidean metric. The details go as follows.

Set  $s_r = r^{1-b}/4$  with  $b$  as above. Recall the family of stabilization ranges  $(S_n)_{n \in \mathbb{N}}$  associated to the graph  $\mathcal{G}(\mathcal{P}_n)$  as required by our assumption; see (8.2). Denote by  $A_r = A_{r,n}(x_1)$  the event (configurations of points of  $\mathcal{P}_n$ ) such that for all  $x \in \mathcal{P}_n \cap B_{r/2}(x_1)$ ,  $S_n(x, \mathcal{P}_n) \leq s_r$ .

By the stopping set property (8.6) of  $S_n$ , and  $s_r \leq r/2 \leq r$  (recalling  $r \geq 1$ ), we have for all  $x \in \mathcal{P}_n \cap B_{r/2}(x_1)$

$$S_n(x, \mathcal{P}_n) = S_n(x, \mathcal{P}_n \cap B_{S_n(x, \mathcal{P}_n)}(x_1)) = S_n(x, \mathcal{P}_n \cap B_{s_r}(x_1)) = S_n(x, \mathcal{P}_n \cap B_r([x]_1^p)). \quad (8.22)$$

Thus, for all  $x \in \mathcal{P}_n \cap B_{r/2}(x_1)$ , the graph neighborhood consistency condition (8.4), implies that the graph neighborhoods of  $x$  in  $\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))$  and  $\mathcal{G}(\mathcal{P}_n)$  coincide. Hence

$$\mathcal{G}_{[x]_1^p}^{\cup(r)} = \mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p)) \supset \mathcal{G}(\mathcal{P}_n) \cap B_{r/2}(x_1) \quad \text{on } A_r.$$

Furthermore, on  $A_r$ , any path of (graph) length  $k_r$  originating from  $x_1$  in  $\mathcal{G}(\mathcal{P}_n)$  has Euclidean length of at most  $(k_r + 1)s_r \leq r/2$ . Extending the length  $k_r$  by 1 accounts for possible additional edges going from  $\mathfrak{B}_{k_r}(x_1)$  to  $\mathfrak{B}_{k_r}(x_j)$ ,  $j \neq 1$ , included in  $\mathcal{G}_{x_i}^{+\cup[k_r]}$ . In other words if  $x_1, y_1, \dots, y_{k_r}$  is a path in  $\mathcal{G}(\mathcal{P}_n)$ , then  $|y_i - x_1| \leq r/2$ ,  $i = 1, \dots, k_r$  and so  $y_i \in B_{r/2}(x_1)$ . Hence, by the previous argument, on  $A_r$  all paths in  $\mathcal{G}(\mathcal{P}_n)$  starting from  $x_1$ , of graph length  $k_r$ , are in  $\mathcal{G}_{[x]_1^p}^{\cup(r)}$ , i.e.,  $\mathcal{G}_{[x]_1^p}^{\cup(r)} \supset \mathcal{G}_{x_1}^{+\cup[k_r]}$  on  $A_r$ . So, we have shown that

$$\{\mathcal{G}_{[x]_1^p}^{\cup(r)} \not\supset \mathcal{G}_{[x]_1^p}^{\cup[k_r]}\} \subset \bigcup_{i=1}^p A_{r,n}(x_i)^c. \quad (8.23)$$

Furthermore, on the event  $A_{r,n}(x_i)$  the subgraph  $\mathcal{G}_{x_i}^{\cup[k_r]}$  of  $\mathcal{G}(\mathcal{P}_n)$  induced by  $\mathfrak{B}_{k_r}(x_i)$  is contained in  $B_{r/2}(x_i)$ . Furthermore, if there exist edges in  $\mathcal{G}(\mathcal{P}_n)$  connecting  $\mathfrak{B}_{k_r}(x_i)$  and  $\mathfrak{B}_{k_r}(x_j)$  for  $i \neq j$ , then their endpoints must lie in  $B_{r/2}(x_i) \cap B_{r/2}(x_j)$ . Under the assumption that  $r < s([x]_1^l, [x]_{l+1}^p)$ , this intersection is empty for  $i \in \{1, \dots, q\}$  and  $j \in \{q+1, p\}$ . Hence, on the event  $\bigcap_{i=1}^p A_{r,n}(x_i)$ , the graph  $\mathcal{G}_{[x]_1^p}^{\cup[k_r]}$  is a disjoint union  $\mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}$ . Thus, we also have

$$\{\mathcal{G}_{[x]_1^p}^{\cup[k_r]} \neq \mathcal{G}_{[x]_1^l}^{\cup[k_r]} \dot{\cup} \mathcal{G}_{[x]_{l+1}^p}^{\cup[k_r]}\} \subset \bigcup_{i=1}^p A_{r,n}(x_i)^c. \quad (8.24)$$

Hence, to prove both (8.15) and (8.16), it suffices to bound  $\mathbb{P}_{[x]_1^p}(\bigcup_{i=1}^p A_{r,n}(x_i)^c)$ , which, by the union bound, reduces to bounding  $\mathbb{P}_{[x]_1^p}(A_{r,n}(x_1)^c)$ . Recalling assumption (8.10), which uniformly bounds the Palm intensity functions of  $\mathcal{P}$ , we obtain

$$\begin{aligned} \mathbb{P}_{[x]_1^p}(A_{r,n}(x_1)^c) &\leq \mathbb{P}_{[x]_1^p} \left( \bigcup_{x \in \mathcal{P}_n \cap B_{r/2}(x_1)} \{S_n(x, \mathcal{P}_n) > s_r\} \right) \\ &\leq \int_{B_{r/2}(x_1)} \mathbb{P}_{[x]_1^p, x}(S_n(x, \mathcal{P}_n) > s_r) \rho_{[x]_1^p}^{(1)}(x) dx + \sum_{i=1}^p \mathbb{P}_{[x]_1^p}(S_n(x_i, \mathcal{P}_n) > s_r) \\ &\leq \theta_d \hat{\kappa}_p \left(\frac{r}{2}\right)^d \varphi'_{p+1}(s_r) + p \varphi'_p(s_r) \\ &\leq (\theta_d \hat{\kappa}_p \left(\frac{r}{2}\right)^d + p) \varphi'_{p+1}(s_r) \\ &\leq \varphi_p''(r), \end{aligned} \quad (8.25)$$

for a fast decreasing function  $\varphi_p''$ . Thus (8.25) together with (8.23) and (8.24) yields that (8.15) and (8.16) hold **s** with the fast-decreasing function  $\varphi_p''' = \min\{p\varphi_p'', 1\}$ .  $\square$

Having established in Lemma 8.7 the BL cluster-localization of spins, to complete the proof of Theorem 8.4, we need to represent spins as a marking function of a suitably marked point process. In this regard, we augment the ground process  $\mathcal{P} = \sum_{x \in \mathcal{P}} \delta_x$  to  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, U(x))}$  by auxiliary pre-marks  $U(x)$ , where  $U(x)$  are i.i.d. uniform random variables on  $[0, 1]$ , which allows the spins to be generated in accordance with the Gibbs specification. We define a marking function  $\xi$  that generates the spin system (8.1) on the graph  $\mathcal{G}(\mathcal{P} \cap B)$  at all sites  $x \in \mathcal{P} \cap B$ , by leveraging the randomness of the auxiliary marks  $U(x)$  attached to the points of the process.

To this end, we first introduce a deterministic set-function of the pre-marks, obtained by summing them modulo 1:

$$U(\tilde{\mathcal{P}} \cap B) := \sum_{x \in \mathcal{P} \cap B} U(x) \pmod{1}, \quad (8.26)$$

where  $B \subset \mathbb{R}^d$  is a bounded Borel set.

Next, to encode the law of the spins on  $\mathcal{G}(\mathcal{P} \cap B)$ , recalling that  $\mathbb{K}$  is countable, we construct a partition  $\mathfrak{G}$  of the unit interval  $[0, 1]$ ,

$$\mathfrak{G} = \left\{ \mathfrak{G}_{(v_x: x \in \mathcal{P} \cap B)} \right\}_{(v_x: x \in \mathcal{P} \cap B) \in \mathbb{K}^{\mathcal{P}(B)}},$$

indexed by all possible joint spin configurations on  $\mathcal{P} \cap B$ . Each element of the partition has Lebesgue measure equal to the Gibbs probability of the corresponding configuration:

$$|\mathfrak{G}_{(v_x: x \in \mathcal{P} \cap B)}| = \mathbb{P}_{\mathcal{G}(\mathcal{P} \cap B)}(V_x(\mathcal{G}(\mathcal{P} \cap B)) = v_x : x \in \mathcal{P} \cap B). \quad (8.27)$$

This partition of  $[0, 1]$  is chosen to be a deterministic and measurable function of the graph  $\mathcal{G}(\mathcal{P} \cap B)$ , that is,  $\mathfrak{G} = \mathfrak{G}(\mathcal{G}(\mathcal{P} \cap B))$ . Its construction depends on the interaction potential and external field functions  $\Psi$  and  $\Phi$  appearing in the Gibbs specification (8.1). Moreover, since the graph construction is translation invariant, the partition may be chosen to inherit this invariance.

Using these ingredients, we assign spins to all points  $x \in \mathcal{P} \cap B$  according to the index of the partition element  $\mathfrak{G}_{(v_x: x \in \mathcal{P} \cap B)}$  containing the value  $U(\tilde{\mathcal{P}} \cap B)$ . That is, for a given point  $x \in \mathcal{P} \cap B$ , we define the spin-marking function  $\xi : \mathbb{R}^d \times [0, 1] \times \hat{\mathcal{N}}_{\mathbb{R}^d \times [0, 1]} \rightarrow \mathbb{K}$  by

$$\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B) := v_x \quad \text{whenever} \quad U(\tilde{\mathcal{P}} \cap B) \in \mathfrak{G}_{(v_x: x \in \mathcal{P} \cap B)}. \quad (8.28)$$

The following lemma justifies the relevance of this spin marking function.

**Lemma 8.8** (Marking function for spin systems). *Consider the marked point process  $\tilde{\mathcal{P}} = \sum_{x \in \mathcal{P}} \delta_{(x, U(x))}$  where  $U(x)$  are i.i.d. uniform random variables on  $[0, 1]$ . Consider the spin-marking function  $\xi$  given in (8.28), related to the Gibbs specification (8.1) via (8.27). For any bounded Borel set  $B \subset \mathbb{R}^d$ , jointly the values of this function  $\{\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B)\}_{x \in \mathcal{P} \cap B}$ , given a realization of  $\mathcal{P} \cap B$ , are distributed as the spin configuration  $\{V_x(\mathcal{G}(\mathcal{P} \cap B))\}_{x \in \mathcal{P} \cap B}$  under this Gibbs model. Further, if  $B_1 \cap B_2 = \emptyset$  then the two collections of marks  $\{\xi(x, \mathcal{P} \cap B_1)\}_{x \in \mathcal{P} \cap B_1}$  and  $\{\xi(x, \mathcal{P} \cap B_2)\}_{x \in \mathcal{P} \cap B_2}$  are conditionally independent given  $\mathcal{P}$ . Moreover, if the mapping  $\mu \mapsto \mathcal{G}(\mu)$  is translation invariant, then the marking function can also be constructed to be translation invariant.*

*Proof.* Through the construction, we have

$$\begin{aligned} \mathbb{P}_{\mathcal{P} \cap B}(\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B) = v_x : x \in \mathcal{P} \cap B) &= \mathbb{P}(U(\tilde{\mathcal{P}} \cap B) \in \mathfrak{G}_{(v_x : x \in \mathcal{P} \cap B)}) \\ &= \mathbb{P}_{\mathcal{P} \cap B}(V_x(\mathcal{G}(\mathcal{P} \cap B)) = v_x : x \in \mathcal{P} \cap B) \end{aligned}$$

for all spin configurations  $(v_x : x \in \mathcal{P} \cap B) \in \mathbb{K}^{\mathcal{P}(B)}$  and hence the marking function has the distribution  $\mathbb{P}_{\mathcal{P} \cap B}$  of the spin system (8.1).

Let  $B_1, B_2$  be disjoint sets. The distributions of  $\{\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B_1)\}_{x \in \mathcal{P} \cap B_1}$  and  $\{\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B_2)\}_{x \in \mathcal{P} \cap B_2}$  are functions of  $U(\tilde{\mathcal{P}} \cap B_1)$  and  $U(\tilde{\mathcal{P}} \cap B_2)$  respectively.  $U(\tilde{\mathcal{P}} \cap B_1)$  and  $U(\tilde{\mathcal{P}} \cap B_2)$  are conditionally independent given  $\mathcal{P}$  for disjoint sets  $B_1, B_2$  as they depend on independent collections of uniform random variables  $\{U(x)\}_{x \in \mathcal{P} \cap B_1}$  and  $\{U(x)\}_{x \in \mathcal{P} \cap B_2}$  respectively. Thus, the assertion on conditional independence for marking functions follows. Finally, as explained below (8.27), the asserted translation invariance also holds by construction.  $\square$

The marking function introduced above and considered in Lemma 8.8 does not satisfy BL-localization (4.3). The obstruction stems from overlaps of the balls  $B_r(x_i)$  around points  $x_i \in [x]_1^p$ , which induce arbitrary dependencies among the components of  $[\xi^{(r)}]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$ . These dependencies need not match those present in  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n)$  and typically become more pronounced as  $r$  increases.

Although the construction of spins via a marking function  $\xi$  on the same input process  $\tilde{\mathcal{P}}$  is not unique—for instance, one may modify the function  $U(\tilde{\mathcal{P}} \cap B)$  in (8.26) while preserving uniform marginals on  $[0, 1]$ —it appears difficult to devise a representation that avoids this issue at the level of BL-localization. Nevertheless, Lemma 8.7 shows that the proposed marking function does satisfy BL cluster-localization, positioning us to complete the proof of Theorem 8.4.

*Proof of Theorem 8.4.* We deduce the result from Theorem 5.2. Let  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$ , where  $U(x)$  are i.i.d. uniform  $[0, 1]$ -valued random variables. The assumed summable exponential mixing correlations of  $\mathcal{P}$ , together with Proposition 3.7, imply that  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8. On this extended input process, we consider the spin marking function  $\xi$  of Lemma 8.8, which reproduces the law of the spins:

$$[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \stackrel{d}{=} V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n)).$$

Furthermore by Lemma 8.8, if  $B_r([x]_1^l) \cap B_r([x]_{l+1}^p) = \emptyset$  for some  $l \in \{1, \dots, p-1\}$ , then the random vectors  $[\xi]_1^l(\tilde{x}, \tilde{\mathcal{P}}_n \cap B_r([x]_1^l))$  and  $[\xi]_{l+1}^p(\tilde{x}, \tilde{\mathcal{P}}_n \cap B_r([x]_{l+1}^p))$  are conditionally independent given  $\mathcal{P}_n$  and are distributed as  $V_{[x]_1^l}(\mathcal{G}(\tilde{\mathcal{P}}_n \cap B_r([x]_1^l)))$  and  $V_{[x]_{l+1}^p}(\mathcal{G}(\tilde{\mathcal{P}}_n \cap B_r([x]_{l+1}^p)))$ , respectively.

Thus, Lemma 8.7 guarantees fast BL cluster-localization of this spin marking function on finite windows of  $\tilde{\mathcal{P}}$ , as per Definition 4.3(i) (comprising the two conditions (4.6) and (4.7)), together with the further qualifications (ii) and (iii). Hence, all the assumptions of Theorem 5.2 are satisfied, and the conclusion of Theorem 8.4 follows.  $\square$

## 8.2.2 Auxiliary statements for spin systems

Recall  $\mathcal{X}'' \subset \mathcal{X}' \subset \mathcal{X}$  as well as the notation of graph-ball unions  $\mathfrak{B}_m(\mathcal{X}'')$  centered at points in  $\mathcal{X}''$  and the subgraph  $\mathcal{G}^{[m]}$  induced by them, introduced prior to Lemma 8.6. The factorization property

of Gibbs distributions over cliques implies *the spatial Markov property*, a fundamental tool in the analysis of Gibbs processes. This implies that the law of the spin vector  $V_{\mathcal{X}''} = V_{\mathcal{X}''}(\mathcal{G}(\mathcal{X}))$  at  $\mathcal{X}''$ , when conditioned on spins outside a region is influenced only by the values on the boundary of the region. A consequence is that the spin vector  $V_{\mathcal{X}''} = V_{\mathcal{X}''}(\mathcal{G}(\mathcal{X}))$  on  $\mathcal{X}''$  conditioned on the boundary specification on  $\partial\mathfrak{B}_{m-1}(\mathcal{X}'')$ , coincides with the law of spin  $V_{\mathcal{X}''}^{[m]}$  conditional on the same boundary specification. Formally, this is stated as follows.

**Lemma 8.9** (Spatial Markov property; [99, Proposition 1]). *Let  $\mathcal{Z} \subset (\mathcal{X}' \cup \partial\mathcal{X}')^c$ . Then  $V_{\mathcal{X}''}$  is conditionally independent of  $V_{\mathcal{Z}}$  given  $V_{\partial\mathcal{X}'}$ . Thus, if  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\mathcal{Z} \cup \partial\mathcal{X}'} = v) > 0$  for  $v = v_{\mathcal{Z} \cup \partial\mathcal{X}'}$  then*

$$\mathcal{L}(V_{\mathcal{X}''} \mid V_{\mathcal{Z} \cup \partial\mathcal{X}'} = v) = \mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}).$$

**Lemma 8.10** (Consistency of conditional spin distributions). *When  $\mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\partial\mathfrak{B}_{m-1}(\mathcal{X}'')} = v) > 0$  for  $v = v_{\partial\mathfrak{B}_{m-1}(\mathcal{X}'')}$  we have*

$$\mathcal{L}(V_{\mathcal{X}''} \mid V_{\partial\mathfrak{B}_{m-1}(\mathcal{X}'')} = v) = \mathcal{L}(V_{\mathcal{X}''}^{[m]} \mid V_{\partial\mathfrak{B}_{m-1}(\mathcal{X}'')} = v).$$

The spatial mixing function  $\mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}')$  in (8.8), controlling the decay of spin correlations on the graph  $\mathcal{G}(\mathcal{X})$ , may be refined by introducing a function  $\mathcal{A}'_{\mathcal{G}(\mathcal{X})} : 2^{\mathcal{X}} \times 2^{\mathcal{X}} \rightarrow [0, \infty)$  which explicitly accounts for disagreements in the boundary conditions on  $\mathcal{Z} \subset \mathcal{X}$  with  $\mathcal{Z} \cap \mathcal{X}'' = \emptyset$  and which satisfies an analog of (8.8), namely

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} \mid V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_{\mathcal{X}''} \mid V_{\mathcal{Z}} = z_{\mathcal{Z}})) \leq \mathcal{A}'_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \mathcal{Z}_{\neq}), \quad (8.29)$$

where  $\mathcal{Z}_{\neq} := \mathcal{Z}_{v \neq z} := \{y \in \mathcal{Z} : v_{\mathcal{Z}}(y) \neq z_{\mathcal{Z}}(y)\}$ . Though  $\mathcal{Z}_{\neq}$  depends on the boundary conditions  $v$  and  $z$ , we shall suppress this dependence for notational convenience. Compared to (8.8), the function  $\mathcal{A}'_{\mathcal{G}(\mathcal{X})}$  in (8.29) considers a potentially strict subset of  $\mathcal{Z}$ , which may be farther from  $\mathcal{X}''$  than  $\mathcal{Z}$  itself. The function  $\mathcal{A}'_{\mathcal{G}(\mathcal{X})}$  is typically decreasing and often exhibits exponential decay

$$\mathcal{A}'_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \mathcal{Z}_{\neq}) \leq C |\mathcal{X}''| \exp(-a d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \mathcal{Z}_{\neq})), \quad (8.30)$$

for some constants  $a, C > 0$ . In this case, the spin model is said to exhibit *strong spatial mixing* and we call  $\mathcal{A}'_{\mathcal{G}(\mathcal{X})}$  the *strong spatial mixing bound*. We demonstrate that the strong spatial mixing bound can be expressed as a sum of strong spatial mixing bound over individual sites.

**Lemma 8.11** (Additive representation for the strong spatial mixing bound). *If the function  $\mathcal{A}'_{\mathcal{G}(\mathcal{X})}$  satisfies (8.29) for all singletons  $\mathcal{X}'' = \{x\} \subset \mathcal{X}$  then the following inequality holds for any configuration of sets  $\mathcal{X}''$ ,  $\mathcal{Z} \subset \mathcal{X}$  such that  $\mathcal{Z} \cap \mathcal{X}'' = \emptyset$ ,*

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} \mid V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_{\mathcal{X}''} \mid V_{\mathcal{Z}} = z_{\mathcal{Z}})) \leq \sum_{y \in \mathcal{X}''} \mathcal{A}'(\{y\}, \mathcal{Z}_{\neq}). \quad (8.31)$$

*Proof.* We will show that if (8.31) hold for all subsets  $\mathcal{X}'' \subset \mathcal{X}$  with  $|\mathcal{X}''| \leq k - 1$ , for  $k \geq 2$ , then it also holds for  $|\mathcal{X}''| = k$ . Thus the lemma follows by induction, since by the assumption for a single site, (8.31) holds for  $k = 1$ .

Let  $\mathcal{X}'' \subset \mathcal{X}$ . Let  $w_{\mathcal{X}''} = (w_x : x \in \mathcal{X}'') \in \mathbb{K}^{\partial\mathcal{X}''}$ . We abbreviate the events  $\{V_{\mathcal{Z}} = v_{\mathcal{Z}}\}$  and  $\{V_{\mathcal{Z}} = z_{\mathcal{Z}}\}$  by  $A_v$  and  $A_z$  respectively.

Considering the probability  $\mathbb{P} := \mathbb{P}_{\mathcal{G}(\mathcal{X})}$  of the Gibbs model (8.1), conditioning on  $V_x$  and using the triangle inequality, we obtain

$$\begin{aligned} & |\mathbb{P}(V_{\mathcal{X}''} = w_{\mathcal{X}''} | A_v) - \mathbb{P}(V_{\mathcal{X}''} = w_{\mathcal{X}''} | A_z)| \\ & \leq |\mathbb{P}(V_{\mathcal{X}'' \setminus \{x\}} = w_{\mathcal{X}'' \setminus \{x\}} | V_x = w_x, A_v) - \mathbb{P}(V_{\mathcal{X}'' \setminus \{x\}} = w_{\mathcal{X}'' \setminus \{x\}} | V_x = w_x, A_z)| \cdot \mathbb{P}(V_x = w_x | A_v) \\ & \quad + \mathbb{P}(V_{\mathcal{X}'' \setminus \{x\}} = w_{\mathcal{X}'' \setminus \{x\}} | V_x = w_x, A_z) \cdot |\mathbb{P}(V_x = w_x | A_v) - \mathbb{P}(V_x = w_x | A_z)|. \end{aligned}$$

Summing over  $w_{\mathcal{X}''}$ 's and using that (8.31) holds for  $|\mathcal{X}''| \in \{1, \dots, k-1\}$  we obtain

$$\begin{aligned} & d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} | V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_{\mathcal{X}''} | V_{\mathcal{Z}} = z_{\mathcal{Z}})) \\ & \leq d_{\text{TV}}(\mathcal{L}(V_x | V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_x | V_{\mathcal{Z}} = z_{\mathcal{Z}})) \\ & \quad + \sum_{w_x \in \mathbb{K}} d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}'' \setminus \{x\}} | V_x = w_x, V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_{\mathcal{X}'' \setminus \{x\}} | V_x = w_x, V_{\mathcal{Z}} = z_{\mathcal{Z}})) \\ & \quad \quad \quad \times \mathbb{P}(V_x = w_x | V_{\mathcal{Z}} = v_{\mathcal{Z}}) \\ & \leq \mathcal{A}'(\{x\}, \mathcal{Z}_{\neq}) + \sum_{w_x \in \mathbb{K}} \mathcal{A}'(\mathcal{X}'' \setminus \{x\}, (\mathcal{Z} \cup \{x\})_{\neq}) \mathbb{P}(V_x = w_x | V_{\mathcal{Z}} = v_{\mathcal{Z}}) \\ & \leq \mathcal{A}'(\{x\}, \mathcal{Z}_{\neq}) + \mathcal{A}'(\mathcal{X}'' \setminus \{x\}, \mathcal{Z}_{\neq}) \\ & = \sum_{x \in \mathcal{X}''} \mathcal{A}'(\{x\}, \mathcal{Z}_{\neq}), \end{aligned}$$

where in the penultimate inequality, we have used that  $\mathcal{Z}_{\neq} = (\mathcal{Z} \cup \{x\})_{\neq}$ . This proves (8.31).  $\square$

If strong spatial mixing (8.29) holds with bounds as in (8.30), then it implies averaged weak spatial mixing (8.8) with  $\mathcal{A}(\mathcal{X}'', \partial\mathcal{X}') = \mathcal{A}'(\mathcal{X}'', \partial\mathcal{X}')$ . Furthermore, if one establishes the stronger version only for single sites and uses (8.31), relying on a straightforward extension of the proof of BL union-localization (8.12), we obtain the following :

$$\begin{aligned} & d_{\text{BL}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{[m]})) \\ & \leq \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \mathcal{A}'(\mathcal{X}'', \partial_{\neq}\mathcal{X}') \mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}_{\mathcal{G}(\mathcal{X})}(V_{\partial\mathcal{X}'}^{[m]} = w_{\partial\mathcal{X}'}), \end{aligned} \tag{8.32}$$

where  $\partial_{\neq}\mathcal{X}'$  denotes  $(\partial\mathcal{X}')_{\neq}$  and notation has been slightly modified for convenience. However, the right-hand side of (8.32) depends not only on the graph structure but also on the probabilities associated with the Gibbs specification. In Section 8.4, we will apply this approach to the hard-core model.

### 8.3 Averaged weak spatial mixing models via disagreement percolation

We take up the question of determining which spin models satisfy the averaged weak spatial mixing condition in Definition 8.3. In this section we exploit the *disagreement percolation* approach introduced by van den Berg and Maes [151] and then use it to establish averaged weak spatial mixing for

spin models on random graphs. This method involves analyzing an auxiliary independent site percolation model on the given graph, where open paths represent disagreements in the spin configurations caused by differing boundary conditions. More precisely, this auxiliary model allows one to stochastically dominate any path of disagreement from the boundary to a given site by an independent Bernoulli site percolation on the graph. The probability of this dominating percolation, defined in what follows, is the essential component in establishing averaged weak spatial mixing on random graphs.

To prevent percolation in this auxiliary model (and to ensure the fast decay of the functions  $(\alpha_p)_{p \geq 1}$  in (8.9)), we consider two alternative conditions in this general disagreement percolation approach: In Section 8.3.1, we use bounds on the *averaged connective constant*, cf. Sinclair et al. [143], and a standard self-avoiding walk counting argument from percolation theory. In Section 8.3.2, we refine this approach by incorporating the concept of a *sharp phase transition* (cf. Duminil-Copin and Tassion [50]) into the disagreement percolation framework.

Henceforth we assume that the spin space  $\mathbb{K}$  is finite. The probability of an open site in the auxiliary disagreement percolation model is given by the *maximum influence* of the neighbors on sites, formalized in [151] as

$$q(\mathcal{G}(\mathcal{X})) := \max_{x \in \mathcal{X}} \sup_{v, z \in \mathbb{K}^{|\mathcal{X}|-1}} d_{\text{TV}}(\mathcal{L}(V_x | V_{N_x} = v), \mathcal{L}(V_x | V_{N_x} = z)), \quad (8.33)$$

where the distribution of the spins  $V$  follows the Gibbs specification (8.1) on the graph  $\mathcal{G}(\mathcal{X})$ , and where  $N_x$  denotes the set of neighbors of  $x$ . The maximum influence depends on both the graph structure and also on the specification of the spin model. We state the key result from disagreement percolation [151, Corollary 1] which controls the decay of spin correlations at (8.8) by the decay of disagreement percolation probabilities.

Let  $\mathcal{X}'' \subset \mathcal{X}' \subset \mathcal{X}$  and assume that  $\partial\mathcal{X}' \neq \emptyset$ . Then

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}), \mathcal{L}(V_{\mathcal{X}''} | V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'})) \leq \mathbb{P}_q(\mathcal{X}'' \xleftrightarrow{\mathcal{X}'_{*q}} \partial\mathcal{X}'), \quad (8.34)$$

where,  $\mathcal{X}'_{*q}$  denotes the independent thinning of  $\mathcal{X}'$ , in which vertices are retained (i.e., declared open) with probability  $q = q(\mathcal{G}(\mathcal{X}))$  as defined in (8.33). The event  $\{\mathcal{X}'' \xleftrightarrow{\mathcal{X}'_{*q}} \partial\mathcal{X}'\} = \{\mathcal{X}'' \xleftrightarrow{\mathcal{G}(\mathcal{X}) \cap \mathcal{X}'_{*q}} \partial\mathcal{X}'\}$  in (8.34) thus corresponds to the existence of a path in the graph  $\mathcal{G}(\mathcal{X})$  that connects  $\mathcal{X}''$  to  $\partial\mathcal{X}'$  and that is entirely contained in the vertex set  $\mathcal{X}'_{*q} \cup (\mathcal{X} \setminus \mathcal{X}')$ . Equivalently, this defines a site percolation model on  $\mathcal{G}(\mathcal{X})$  in which vertices in  $\mathcal{X}'$  are independently open with probability  $q$ , while all vertices in  $\mathcal{X} \setminus \mathcal{X}'$  are deterministically open. Moreover, when the underlying graph  $\mathcal{G}(\mathcal{X})$  is clear from the context, we omit it in the above connectivity event. The probability measure  $\mathbb{P}_q$  is taken with respect to this thinning (percolation) randomness. If  $\partial\mathcal{X}' = \emptyset$ , then the left-hand side of (8.34) vanishes and the inequality holds trivially.

We now bound maximum influence when  $\mathcal{X}$  is replaced by  $\mathcal{P}$  and when  $\mathcal{G}(\mathcal{P})$  is a stabilizing graph with not too many self-avoiding paths.

### 8.3.1 Controlling disagreement percolation probabilities via averaged connective constant

Connective constants allow one to bound the (usually exponential) growth in the number of self-avoiding paths of length  $m$  starting from an arbitrary vertex in a graph. We use this notion to bound disagreement percolation probabilities and hence obtain the decay of spin correlations through (8.34). We adapt these concepts to our context of a spatial random graph  $\mathcal{G}$  defined over  $\mathcal{P}$ , defining an averaged connective constant as follows.

**Definition 8.12** (Averaged connective constant for spatial random graph). *For a graph  $\mathcal{G} = \mathcal{G}(\cdot)$  on the input process  $\mathcal{P}$ , the averaged connective constant  $\bar{\Delta} \in [0, \infty]$  is defined as*

$$\bar{\Delta} := \inf \left\{ \delta \geq 0 : \forall p \in \mathbb{N} \exists m_p \geq 1 \text{ such that for all } m \geq m_p \quad \mathfrak{n}_p(m) \leq \delta^m \right\}, \quad (8.35)$$

where

$$\mathfrak{n}_p(m) := \sup_{B \in \mathcal{B}_b} \sup_{[x]_1^p \subset B} \mathbb{E}_{[x]_1^p} [N(x_1, m; \mathcal{G}(\mathcal{P} \cap B))], \quad (8.36)$$

and where  $N(x_1, m; \mathcal{G}(\mathcal{P} \cap B))$  is the number of self-avoiding paths of length  $m$  starting from  $x_1 \in \mathcal{P} \cap B$  in  $\mathcal{G}(\mathcal{P} \cap B)$ .

The *quenched* version of this constant, given in the sequel by  $\Delta$  as in (8.53), is defined by removing the expectation in (8.36) and hence yielding the almost sure inequality  $N(x_1, m; \mathcal{G}(\mathcal{P} \cap B)) \leq \Delta^m$  for large  $m \in \mathbb{N}$ ; this shall be used in applications to hard-core models in Section 8.4. Bounds for this quenched version, which upper-bounds the averaged version, exist for certain random graphs. Notably, this holds for the  $k$ -nearest neighbor graph, which is distinguished by its bounded degree (see Yukich [158, Lemma 8.2]). Lacoïn [96] provides an example illustrating the sharp inequality between averaged and quenched connective constants.

The next result, derived using Theorem 8.4, proves normal convergence for spin systems on spatial random graphs whose maximal influence is strictly smaller than the inverse of the averaged connective constant.

**Proposition 8.13** (CLT for sum of spins for graphs with finite averaged connective constant). *Let  $\mathcal{P}$  be a simple, stationary point process on  $\mathbb{R}^d$  with summable exponential mixing correlations as in Definition 4.8, and assume that the (reduced) Palm intensity function of the ground process  $\mathcal{P}$  is bounded at all orders, as specified in (8.10). Let  $\mathcal{G} = \mathcal{G}(\cdot)$  be a stabilizing interaction graph on finite windows with respect to  $\mathcal{P}$ , with finite averaged connective constant  $\bar{\Delta} < \infty$  defined in Definition 8.12. If, for some  $\epsilon > 0$ , all  $B \in \mathcal{B}_b$ , and almost every realization  $\mathcal{P}$ , the maximum influence  $q(\mathcal{G}(\mathcal{P} \cap B))$  as in (8.33) of the Gibbs specification (8.1) on  $\mathcal{G}(\mathcal{P} \cap B)$  with finite spin space  $\mathbb{K}$  satisfies the inequality*

$$q(\mathcal{G}(\mathcal{P} \cap B)) \leq (\epsilon + \bar{\Delta})^{-1} \quad (8.37)$$

then the spin configuration  $\{V_x(\mathcal{G}(\mathcal{P}_n))\}_{x \in \mathcal{P}_n}$  satisfies the CLT as stated in Theorem 8.4, provided the moment and variance conditions there are also met.

Before giving the proof, we provide several examples where one may verify the assumption (8.37) on the influence parameter  $q(\mathcal{G}(\mathcal{P} \cap B))$ . We choose classical examples from [151, Section 3] and [31, Section 4.3], as upper bounds on the maximum influence  $q(\mathcal{G}(\mathcal{X}))$  have been computed in these models for general graphs. More examples are in [64, Section 3], [63, Section 8]. In the examples below, we take  $\mathcal{X} = \{x_1, \dots, x_m\}$  and  $v_i$ 's are taken to be  $\mathbb{K}$ -valued.

(i) *Hard-core model.* We associate  $\{0, 1\}$ -valued random variables (spins) to each vertex in  $\mathcal{G}(\mathcal{X})$ , ensuring that no neighboring vertices both take the value 1, with joint probability proportional to

$$\pi_{[z]_1^m}([v]_1^m) = \prod_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \mathbf{1}(v_i v_j = 0) \prod_{x_i \in \mathcal{X}} \lambda^{v_i}, \quad v_i \in \{0, 1\}, \quad (8.38)$$

where  $\lambda > 0$  is the *activity parameter*. This is the Gibbs model (8.1), with  $\mathbb{K} = \{0, 1\}$ ,  $\beta = 1$ ,  $\Psi(v_i, v_j) = \log \mathbf{1}(v_i v_j = 0)$ ,  $\lambda = e^\gamma$ , and  $\Phi(v) = v$ . Loosely speaking, small  $\lambda$ -values are expected to result in weak correlations between spins at distant sites, a chief requirement for our general central limit theorems. Indeed, it can be shown that  $q(\mathcal{G}(\mathcal{X})) \leq \frac{\lambda}{1+\lambda}$  for any graph  $\mathcal{G}(\mathcal{X})$ , so Proposition 8.13 already applies for  $\lambda < (\max(0, \bar{\Delta} - 1))^{-1}$  for graphs with finite averaged connective constant  $\bar{\Delta} < \infty$ . In Section 8.4, we will revisit the hard-core model and prove the central limit theorem using a strong version of the spatial mixing approach, though under assumptions on the quenched, rather than averaged, connective constant.

(ii) *Widom-Rowlinson model.* Here  $\mathbb{K} = \{-1, 0, +1\}$  and the spin variables are chosen randomly with joint probability proportional to

$$\pi_{[z]_1^m}([v]_1^m) = \prod_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \mathbf{1}(v_i v_j \neq -1) \prod_{x_i \in \mathcal{X}} \lambda^{|v_i|}, \quad (8.39)$$

where  $\lambda$  is the *activity parameter*. This is the Gibbs model in (8.1) with  $\beta = 1$ ,  $\Psi(v_i, v_j) = \log \mathbf{1}(v_i v_j \neq -1)$ ,  $\lambda = e^\gamma$ ,  $\Phi(v) = |v|$ . Here if  $\mathcal{G}(\mathcal{X})$  has a vertex of degree 2 then  $q(\mathcal{G}(\mathcal{X})) \leq \frac{2\lambda}{1+2\lambda}$ . Thus, Proposition 8.13 holds in this model for  $\lambda < (2 \max(0, \bar{\Delta} - 1))^{-1}$ .

(iii) *Dobrushin's interdependence matrix.* Consider a model that satisfies

$$d_{\text{TV}}(\mathcal{L}(V_x | V_{\mathcal{X} \setminus x} = v), \mathcal{L}(V_x | V_{\mathcal{X} \setminus x} = z)) \leq \sum_{y \in \mathcal{X} \setminus x} A_{x,y} \mathbf{1}(v_y \neq z_y),$$

for some matrix  $A$  on  $\mathcal{X} \times \mathcal{X}$  with non-negative entries and zero entries on the diagonal. Then

$$q(\mathcal{G}(\mathcal{X})) \leq \|A\|_\infty := \max_{x \in \mathcal{X}} \sum_{y \in \mathcal{X}} A_{x,y}.$$

The matrix  $A$  is called ‘Dobrushin’s interdependence matrix’ and there are many examples where an upper bound for  $\|A\|_\infty$  can be derived for bounded degree graphs; for example see [63, Section 8]. We present two specific examples next where this has been computed in [31, Section 4.3] and in particular can be applied to the  $k$ -nearest neighbor graph.

(iv) *Ising model.* Here  $\mathbb{K} = \{-1, +1\}$ ,  $\Psi(u, v) = uv$  and  $\Phi(v) = v$ . The Gibbs measure is given by

$$\pi_{[x]_1^m}([v]_1^m) = \prod_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \exp\left\{\beta \sum_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} v_i v_j + \gamma \sum_{x_i \in \mathcal{X}} v_i\right\},$$

In this case  $\|A\|_\infty \leq 4\beta \deg_{\max}$ , where  $\deg_{\max}$  stands for the maximal degree. Thus in case of interaction graphs with degrees bounded almost surely, say by  $\deg_{\max}$ , then Proposition 8.13 applies for the Ising model with  $\beta < (4\deg_{\max}(\deg_{\max} - 1))^{-1}$ . We have used that  $\bar{\Delta} \leq \deg_{\max} - 1$ .

(v) *Proper  $k$ -colorings.* Here  $\mathbb{K} = \{1, \dots, k\}$ ,  $\Psi(v_i, v_j) = \log \mathbf{1}(v_i \neq v_j)$ ,  $\beta = 1$  and  $\gamma = 0$ . The Gibbs measure is given by

$$\pi_{[x]_1^m}([v]_1^m) = \prod_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \mathbf{1}(v_i \neq v_j).$$

For graphs whose degrees are bounded by  $\deg_{\max}$ , we again have  $\|A\|_\infty \leq \deg_{\max}/(k - \deg_{\max})$ . Thus in case of interaction graphs with degrees bounded almost surely (say by  $\deg_{\max}$ ) then we can verify that Proposition 8.13 applies for the proper  $k$ -colorings with  $k > \deg_{\max}^2$ .

*Proof of Proposition 8.13.* To deduce the result from Theorem 8.4 and (8.34), it suffices to show that the Gibbs model satisfies averaged weak spatial mixing as in Definition 8.3 with some spatial mixing function  $\mathcal{A}$ , since all other assumptions are explicitly assumed. This involves verifying the inequality (8.8) and showing the functions  $(\alpha_p(k))_{p \in \mathbb{N}}$  in (8.9) are fast decreasing in  $k$ , both involving  $\mathcal{A}$ . In what follows we will show that this double condition is met in our model for the spatial mixing function

$$\mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}') := \sum_{x \in \mathcal{X}''} N(x, d_{\mathcal{G}(\mathcal{X})}(x, \partial\mathcal{X}'); \mathcal{G}(\mathcal{X})) a^{d_{\mathcal{G}(\mathcal{X})}(x, \partial\mathcal{X}')}, \quad (8.40)$$

where  $a := 1/(\bar{\Delta} + \epsilon)$ , with  $\epsilon > 0$  fixed, provided  $\partial\mathcal{X}' \neq \emptyset$ , otherwise we put  $\mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \emptyset) = 0$ . We start by studying the functions  $(\alpha_p)_{p \in \mathbb{N}}$  at (8.9). Observe that

$$\mathbb{E}_{[x]_1^p}[\mathcal{A}_{\mathcal{G}(\mathcal{P} \cap B)}([x]_1^p, \partial\mathfrak{B}_k([x]_1^p))] \leq \sum_{i=1}^p \mathbb{E}_{[x]_1^p}[N(x_i, k+1, \mathcal{G}(\mathcal{P} \cap B))] a^{k+1} \leq p \mathfrak{n}_p(k+1) a^{k+1}.$$

Given  $\epsilon > 0$ , by (8.35) we have  $\mathfrak{n}_p(k+1) \leq (\bar{\Delta} + \epsilon/2)^{k+1}$  for  $k \geq m_p$  and thus

$$\mathbb{E}_{[x]_1^p}[\mathcal{A}_{\mathcal{G}(\mathcal{P} \cap B)}([x]_1^p, \partial\mathfrak{B}_k([x]_1^p))] \leq p((\bar{\Delta} + \epsilon/2)a)^{k+1} = p \left( \frac{\bar{\Delta} + \epsilon/2}{\bar{\Delta} + \epsilon} \right)^{k+1},$$

which shows that  $(\alpha_p(k))_{p \in \mathbb{N}}$  are fast decreasing in  $k$ , as required.

We show (8.8) holds with  $\mathcal{A}$  as in (8.40) above by bounding the right-hand side of (8.34) using standard percolation arguments. Consider  $x \in \mathcal{X}''$  and denote by  $\Gamma$  a path from  $x$  to  $\partial\mathcal{X}'$  on  $\mathcal{G}(\mathcal{X})$ , and let  $|\Gamma|$  be the number of sites in  $\Gamma$ . Observe that  $|\Gamma| \geq d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}') =: l$ .

For a given  $q \in [0, 1]$ , in what follows, consider open paths in site percolation on  $\mathcal{G}(\mathcal{X})$  in which vertices in  $\mathcal{X}'$  are independently open with probability  $q$ , while all other vertices are always open. With probability  $\mathbb{P}_q = \mathbb{P}_{\mathcal{G}(\mathcal{X}) \cap \mathcal{X}'_q}$  of this model, Markov's inequality and the above observations yield

$$\mathbb{P}_q(\text{there exists a self-avoiding (s-a) open path from } x \text{ to } \partial\mathcal{X}')$$

$$\begin{aligned}
&\leq \mathbb{P}_q(\text{there exists a s-a open path of length } l \text{ from } x) \\
&\leq \sum_{\Gamma: \text{ s-a path of length } l \text{ in } \mathcal{X}' \text{ from } x} \mathbb{P}_q(\text{all sites in } \Gamma \text{ are open}) \\
&\leq q^l N(x, l; \mathcal{G}(\mathcal{X})), \tag{8.41}
\end{aligned}$$

where we recall that  $N(x, l; \mathcal{G}(\mathcal{X}))$  is the number of self-avoiding walks in  $\mathcal{G}(\mathcal{X})$  of length  $l$  starting at  $x$ . Note that a path of length  $l$  consists of  $l + 1$  vertices. Since only the vertices in  $\mathcal{X}'$  may be removed with probability  $1 - q$ , while those in  $\partial\mathcal{X}'$  are always open, it follows that the probability that the entire path is open is generously bounded by  $q^l$ .

Assume now  $q = q(\mathcal{G}(\mathcal{X}))$ . The union bound and (8.34) yield

$$\begin{aligned}
&d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}|V_{\partial\mathcal{X}'} = v_{\partial\mathcal{X}'}), \mathcal{L}(V_{\mathcal{X}''}|V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'})) \\
&\leq \sum_{x \in \mathcal{X}''} \mathbb{P}_q(\text{there exists a s-a open path from } x \text{ to } \partial\mathcal{X}') \\
&\leq \sum_{x \in \mathcal{X}''} N(x, l; \mathcal{G}(\mathcal{X})) q^l \\
&\leq \sum_{x \in \mathcal{X}''} N(x, d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}'); \mathcal{G}(\mathcal{X})) a^{d_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}')} \\
&= \mathcal{A}_{\mathcal{G}(\mathcal{X})}(\mathcal{X}'', \partial\mathcal{X}'),
\end{aligned}$$

where, the last inequality holds provided the inequality  $q(\mathcal{G}(\mathcal{X})) \leq 1/(\epsilon + \bar{\Delta}) =: a$ . This inequality is assumed almost surely for  $\mathcal{G}(\mathcal{X}) = \mathcal{G}(\mathcal{P} \cap B)$ . This shows that (8.8) holds with  $\mathcal{A}$  as in (8.40) and completes the proof of Proposition 8.13.  $\square$

### 8.3.2 Controlling spin correlations via disagreement percolation and sharp phase transition

In this section, we refine the self-avoiding walk counting argument used in Section 8.3.1 along with disagreement percolation to establish averaged weak spatial mixing for the spin systems. Specifically, we present an alternative argument for the exponential decay of disagreement probabilities, one exploiting the concept of a *sharp phase transition* in an auxiliary independent Bernoulli site percolation model. Essentially, if a sharp phase transition exists in this model, then by definition, for subcritical values  $q$  of the Bernoulli probability, *the 1-arm probabilities* decay exponentially i.e., the probability of paths connecting a point  $x \in \mathcal{X}$  to points outside the ball  $B_r(x)$  of radius  $r$  decays exponentially with  $r$ . (This property, combined with a mean-field lower bound for the percolation function at supercritical  $q$ , formally defines the sharp phase transition in the model.) The sharp phase transition has been proved for many graph models, including the graph of the Boolean model; [51, 52, 159, 109]. In this section, we detail how this concept can be applied to establish the averaged weak spatial mixing property for these graphs.

Since we are dealing with spatial random graphs, sharp phase transition results are often with respect to the Euclidean distance and hence we use a version of averaged weak spatial mixing with respect to the Euclidean distance to prove the same. As will be illustrated by the theorem below,

this variant of averaged spatial mixing can be used in Theorem 8.4 instead of condition (ii) of Definition 8.3. For generality, we begin by defining ‘sharp’ critical probability  $q_c$  in the context of disagreement percolation models on random graphs. Analogously to Definition 8.12, we define it as follows:

**Definition 8.14** (Critical probability in disagreement percolation). *For a graph  $\mathcal{G} = \mathcal{G}(\cdot)$  defined on the input process  $\mathcal{P}$ , the critical probability  $q_c$  for percolation is defined as*

$$q_c := \sup \left\{ q \in [0, 1] : \forall p \in \mathbb{N}, \text{ the function } \tilde{\alpha}_{p,q}(r) \text{ is fast decreasing as } r \rightarrow \infty \right\}, \quad (8.42)$$

where for  $r > 0$

$$\tilde{\alpha}_{p,q}(r) := \sup_{n \geq 1} \sup_{[x]_1^p \subset W_n} \mathbb{P}_{[x]_1^p, q} \left( x_1 \stackrel{(\mathcal{P}_n \cap B_r(x))^{*q}}{\longleftrightarrow} \mathcal{P}_n \cap B_r(x_1)^c \right), \quad (8.43)$$

where the underlying graph is  $\mathcal{G}(\mathcal{P}_n)$  and recall that  $\{x \stackrel{\mathcal{X}'^{*q}}{\longleftrightarrow} \mathcal{Y}\} = \{x \stackrel{\mathcal{G}(\mathcal{X}) \cap \mathcal{X}'^{*q}}{\longleftrightarrow} \mathcal{Y}\}$  means that there exists an open path in  $\mathcal{G}(\mathcal{X})$  from  $x \in \mathcal{X}' \subset \mathcal{X}$  to  $\mathcal{Y} \subset \mathcal{X}$  through vertices in  $\mathcal{X}'^{*q}$ , the thinned subset of  $\mathcal{X}'$  where vertices in  $\mathcal{X}'$  are independently open/retained with probability  $q$ , while all other vertices in  $\mathcal{X}$  are always open, and where  $\mathbb{P}_{[x]_1^p, q}$  incorporates this additional percolation randomness on top of Palm distribution.

The next result, proved similarly to Theorem 8.4, establishes normal convergence for spin systems on spatial random graphs whenever the maximal influence is smaller than the critical probability  $q_c$ .

**Proposition 8.15** (CLT for sum of spins for graphs with sharp phase transition). *Let  $\mathcal{P}$  be a simple, stationary point process on  $\mathbb{R}^d$  with summable exponential mixing correlations as in Definition 4.8, and assume that the (reduced) Palm intensity function of the ground process  $\mathcal{P}$  is bounded at all orders, as specified at (8.10). Let  $\mathcal{G} = \mathcal{G}(\cdot)$  be a stabilizing interaction graph on finite windows with respect to  $\mathcal{P}$  as per Definition 8.1, with the critical probability  $q_c$  as in Definition 8.14. Assume that for some  $\epsilon > 0$  and for all bounded Borel sets  $B \subset \mathbb{R}^d$  and almost every realization  $\mathcal{P}$ , the maximum influence  $q(\mathcal{G}(\mathcal{P} \cap B))$  defined in (8.33) of the Gibbs specification (8.1) with finite spin space  $\mathbb{K}$  satisfies the inequality*

$$q(\mathcal{G}(\mathcal{P} \cap B)) < q_c - \epsilon. \quad (8.44)$$

*Then the system of spins  $\{V_x(\mathcal{G}(\mathcal{P}_n))\}_{x \in \mathcal{P}_n}$  satisfies the CLT as stated in Theorem 8.4, provided the moment and variance conditions there are also met.*

Using bounds as in (8.41) and the quenched version  $\Delta$  of the connective constant (see the comment after Definition 8.12) one obtains  $\Delta^{-1} \leq q'_c$ , where  $q'_c$  is a (quenched) version of the sharp critical percolation in the graph  $\mathcal{G}(\mathcal{P})$  given a realization  $\mathcal{P}$ . The strict inequality has been demonstrated in certain lattice cases where it is known that  $q'_c > \Delta^{-1}$ ; see [156, 49, 143]. Although averaged connective constants or critical probabilities are not easy to compute for spatial random graphs models, we expect the strict inequality  $q_c > \bar{\Delta}^{-1}$  to hold for some models and hence Proposition 8.15 would give a central limit theorem covering a wider range of activity parameters than that given by Proposition 8.13. Moreover, as demonstrated in Section 8.3.3, we can still apply Proposition 8.15 for unbounded degree graphs without needing to compute bounds on the averaged connective constant.

*Proof of Proposition 8.15.* We exploit the critical probability threshold in disagreement percolation, together with the maximum influence bounds of the Gibbs specification, to establish Lemma 8.7 directly, thereby bypassing the verification of averaged weak spatial mixing as defined in Definition 8.3. With this key lemma in place, the proof of Theorem 8.4 then follows by reusing the same core arguments. The main difference between the proofs of (8.13) and (8.14) in Lemma 8.7 lies in the approximation strategy for the spin vector  $V_{[x]_1^p}^{\cup(r)} = V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p)))$ . Instead of passing through the intermediate approximation  $V_{[x]_1^p}^{\cup[kr]}$ , which relies on graph-based neighborhoods of the points  $[x]_1^p$  in  $\mathcal{G}(\mathcal{P}_n)$ , we approximate  $V_{[x]_1^p}^{\cup(r)}$  directly by  $V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_1^p)))$  using a Euclidean truncation of the input configuration. This direct approximation is made possible by the disagreement-percolation bounds described above and avoids the need for graph-neighborhood expansions. The details are given below.

Fix  $n \in \mathbb{N}$  and  $[x]_1^p \subset W_n$ . We set  $\mathcal{X}'' = [x]_1^p$  and under  $\mathbb{P}_{[x]_1^p}$  consider the realization  $\mathcal{P}_n$ . Set  $\mathcal{X}' = \mathcal{P}_n \cap B_{r/2}([x]_1^p)$  where we have suppressed the dependence on  $r, [x]_1^p$  in  $\mathcal{X}'$  for convenience. Define  $A_{r,n}([x]_1^p)$  as the event (configurations of points of  $\mathcal{P}_n$ ) such that for all  $x \in \mathcal{P}_n \cap B_{r/2}([x]_1^p)$ ,  $S_n(x, \mathcal{P}_n) \leq r^b/4$  for a fixed  $b \in (0, 1)$ . Recall that a similar event was defined in the proof of Lemma 8.7 and we have used the same notation for convenience. We know that under the event  $A_{r,n}([x]_1^p)$ , the  $r/2$  neighborhood of  $\mathcal{X}'$  remains the same in  $\mathcal{P}_n$  and  $\mathcal{P}_n \cap B_r([x]_1^p)$  when  $r \geq 1$ . Thus, we have that

$$\partial^{\mathcal{G}(\mathcal{P}_n)} \mathcal{X}' = \partial^{\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))} \mathcal{X}' =: \partial \mathcal{X}' \subset \mathcal{P}_n \cap B_{r/2}([x]_1^p)^c. \quad (8.45)$$

Assume that  $A_{r,n}([x]_1^p)$  holds. Following the proof of Lemma 8.6 verbatim with the above notation, for  $V_{\mathcal{X}''} = V_{\mathcal{X}''}(\mathcal{G}(\mathcal{P}_n))$  and  $V_{\mathcal{X}''}^{\cup(r)} = V_{\mathcal{X}''}(\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p)))$ , we have

$$\begin{aligned} & d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup(r)})) \\ & \leq \sum_{z_{\partial \mathcal{X}'}, w_{\partial \mathcal{X}'} \in \mathbb{K}^{\partial \mathcal{X}'}} \sum_{v_{\mathcal{X}''} \in \mathbb{K}^{\mathcal{X}''}} \left| \mathbb{P}(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = z_{\partial \mathcal{X}'}) - \mathbb{P}'(V_{\mathcal{X}''} = v_{\mathcal{X}''} | V^{\cup(r)}_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'}) \right| \\ & \quad \times \mathbb{P}(V_{\partial \mathcal{X}'} = z_{\partial \mathcal{X}'}) \mathbb{P}'(V^{\cup(r)}_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'}), \end{aligned}$$

with probabilities  $\mathbb{P} = \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}$  and  $\mathbb{P}' = \mathbb{P}_{\mathcal{G}(\mathcal{P}_n \cap B_r([x]_1^p))}$  corresponding to the Gibbs model on the respective graphs. Here we have used that the boundary equivalence as in (8.45) holds under  $A_{r,n}([x]_1^p)$ . Now the spatial Markov property from Lemma 8.9 yields the conditional equality under the two probabilities  $\mathbb{P}$  and  $\mathbb{P}'$

$$\mathcal{L}(V_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'}) = \mathcal{L}(V_{\mathcal{X}''}^{\cup(r)} | V_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'})$$

and thus

$$d_{\text{TV}}(\mathcal{L}(V_{[x]_1^p}), \mathcal{L}(V_{[x]_1^p}^{\cup(r)})) \leq \sup_{z_{\partial \mathcal{X}'}, w_{\partial \mathcal{X}'}} d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = z_{\partial \mathcal{X}'}), \mathcal{L}(V_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'})). \quad (8.46)$$

To further bound the last expression, we consider site percolation on  $\mathcal{G}(\mathcal{P}_n)$  where vertices in  $\mathcal{X}'$  are independently open with probability  $q \in [0, 1]$  and where all other vertices are always open. Denote the probability measure of this model by  $\mathbb{P}_q$  and omit  $\mathcal{G}(\mathcal{P}_n)$  from the path notation  $\longleftrightarrow$ . Applying (8.34) with thinning probability  $q = q(\mathcal{P}_n)$  and then increasing it to  $q' := q_c - \varepsilon$ , we obtain

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = z_{\partial \mathcal{X}'}), \mathcal{L}(V_{\mathcal{X}''} | V_{\partial \mathcal{X}'} = w_{\partial \mathcal{X}'})) \leq \mathbb{P}_{q(\mathcal{G}(\mathcal{P}_n))}(\mathcal{X}'' \xrightarrow{q'} \partial \mathcal{X}')$$

$$\begin{aligned}
&\leq \mathbb{P}_{q'}(\mathcal{X}'' \xleftrightarrow{*q'} \partial\mathcal{X}') \\
&\leq \sum_{x \in \mathcal{X}''} \mathbb{P}_{q'}(x \xleftrightarrow{*q'} \partial\mathcal{X}') \\
&\leq \sum_{x \in \mathcal{X}''} \mathbb{P}_{q'}(x \xleftrightarrow{(\mathcal{P}_n \cap B_{r/2}(x)) * q'} \mathcal{P}_n \cap B_{r/2}(x)^c),
\end{aligned}$$

where the second inequality follows from the assumption (8.44) (almost surely for all  $\mathcal{P}_n$  and  $B = W_n$ ), the third one by the union bound, and the last by (8.45), considering shorter paths and considering  $q'$ -thinning only on points in  $\mathcal{P}_n \cap B_{r/2}(x)$ .

Thus, using (8.46), we have derived that under  $A_{r,n}([x]_1^p)$

$$d_{\text{TV}}(\mathcal{L}(V_{\mathcal{X}''}), \mathcal{L}(V_{\mathcal{X}''}^{\cup(r)})) \leq \sum_{x \in \mathcal{X}''} \mathbb{P}_{q'}(x \xleftrightarrow{(\mathcal{P}_n \cap B_{r/2}(x)) * q'} \mathcal{P}_n \cap B_{r/2}(x)^c). \quad (8.47)$$

Using the above bound, averaging with respect to the input  $\mathcal{P}_n$ , and following the lines of the proof of Lemma 8.7, we find

$$\begin{aligned}
d_{BL, [x]_1^p}(\mathcal{L}(V_{[x]_1^p}), \mathcal{L}(V_{[x]_1^p}^{\cup(r)})) &\leq \mathbb{E}_{[x]_1^p} \left[ d_{BL}(\mathcal{L}(V_{[x]_1^p} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n)) \right] \\
&\leq \mathbb{E}_{[x]_1^p} \left[ d_{\text{TV}}(\mathcal{L}(V_{[x]_1^p} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p}^{\cup(r)} | \mathcal{P}_n)) \mathbf{1}(A_{r,n}([x]_1^p)) \right] + 2\mathbb{P}_{[x]_1^p}(A_{r,n}([x]_1^p)^c) \\
&= \sum_{i=1}^p \mathbb{P}_{[x]_1^p, q'}(x_i \xleftrightarrow{(\mathcal{P}_n \cap B_{r/2}(x_i)) * q'} \mathcal{P}_n \cap B_{r/2}(x_i)^c) + 2 \sum_{i=1}^p \mathbb{P}_{[x]_1^p}(A_{r,n}(x_i)^c) \\
&\leq p \tilde{\alpha}_{p, q'}(r/2) + 2 \sum_{i=1}^p \mathbb{P}_{[x]_1^p}(A_{r,n}(x_i)^c),
\end{aligned}$$

where in the last inequality we used  $\tilde{\alpha}_{p, q'}(r)$  function defined in (8.43) and the fact that  $A_{r,n}(x_i) \subset A_{r,n}([x]_1^p)$  for  $i = 1, \dots, p$ . By Definition 8.14 of  $q_c$ , the fast decay of the first term follows by the assumption of  $q' = q_c - \epsilon$ .

The fast decay of the second term can be shown by observing for all  $i \in \{1, \dots, p\}$ , similarly as in (8.25),

$$\mathbb{P}_{[x]_1^p}(A_{r,n}(x_i)^c) \leq (\theta_d \hat{\kappa}_p (\frac{r}{2})^d + p) \varphi'_{p+1}(r^b/4).$$

Thus, we have established the fast BL union-localization bound (8.13).

We now prove (8.14). Following closely the proof of Lemma 8.7, we fix  $p \geq 2$ , choose  $q \in \{1, \dots, p-1\}$ , and consider points  $[x]_1^p \in W_n$ . Let  $r > 0$  satisfy  $0 < r < s([x]_1^l, [x]_{l+1}^p)$ , and, under the Palm distribution  $\mathbb{P}_{[x]_1^p}$ , consider the spin vectors  $V_{[x]_1^p}^{\cup(r)}$ ,  $V_{[x]_1^l}^{\cup(r)}$ , and  $V_{[x]_{l+1}^p}^{\cup(r)}$  defined as before, with  $V_{[x]_1^l}^{\cup(r)}$  and  $V_{[x]_{l+1}^p}^{\cup(r)}$  assumed to be conditionally independent given  $\mathcal{P}_n$ .

On the event  $A_{r,n}([x]_1^p)$ , the vectors  $V_{[x]_1^l}^{\cup(r)}$  and  $V_{[x]_{l+1}^p}^{\cup(r)}$  are marginally approximated by their counterparts obtained on the graph  $\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_1^l))$  i.e., by  $V_{[x]_1^l}^l(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_1^l)))$  and  $V_{[x]_{l+1}^p}^p(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_{l+1}^p)))$ , respectively. The same approximation holds for  $V_{[x]_1^p}^{\cup(r)}$ . Moreover, the approximation of

$V_{[x]_1^p}^{\cup(r)}$  coincides with the product of the corresponding approximations of  $V_{[x]_1^l}^{\cup(r)}$  and  $V_{[x]_{l+1}^p}^{\cup(r)}$  i.e.,

$$\mathcal{L}(V_{[x]_1^p}(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_1^p)))) = \mathcal{L}(V_{[x]_1^l}(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_1^l)))) \times \mathcal{L}(V_{[x]_{l+1}^p}(\mathcal{G}(\mathcal{P}_n \cap B_{r/2}([x]_{l+1}^p)))) .$$

Using the fact that the probability of the complement of  $A_{r,n}([x]_1^p)$  decays fast in  $r$ , together with the fast decreasing of the error in the above approximations of  $V_{[x]_1^p}^{\cup(r)}$ ,  $V_{[x]_1^l}^{\cup(r)}$ , and  $V_{[x]_{l+1}^p}^{\cup(r)}$ , we conclude that

$$d_{\text{BL}, [x]_1^p}(\mathcal{L}(V_{[x]_1^p}^{\cup(r)}), \mathcal{L}(V_{[x]_1^l}^{\cup(r)}) \times \mathcal{L}(V_{[x]_{l+1}^p}^{\cup(r)}))$$

is fast decreasing in  $r$ . This establishes (8.14).

Finally, applying Theorem 5.2 as in the proof of Theorem 8.4, we conclude the proof.  $\square$

### 8.3.3 Spin systems on the Poisson-Boolean model

We consider an example of a general spin model on the graph generated by a Boolean model, where one can identify the critical probability for disagreement percolation  $q_c$  and also furnish an explicit bound for the maximal influence  $q(\mathcal{G}(\mathcal{P} \cap B))$  of the spin system in (8.44).

Let  $\mathcal{P}$  be a Poisson point process of intensity  $\rho$  on  $\mathbb{R}^d$ ,  $K \subset \mathbb{R}^d$  a compact set with a non-empty interior and without loss of generality, we assume that  $\mathbf{0} \in K$ . The Poisson-Boolean model is  $\mathcal{O} = \mathcal{O}(\mathcal{P}) := \bigcup_{x \in \mathcal{P}} (x \oplus K)$ . For  $B \subset \mathbb{R}^d$ , the Poisson-Boolean graph  $\mathcal{G}(\mathcal{P} \cap B)$  has vertex set  $\mathcal{P} \cap B$ , with vertices  $x, y$  connected by an edge if  $(x \oplus K) \cap (y \oplus K) \neq \emptyset$ .

The critical intensity for percolation in the Boolean model, which determines whether an infinite connected component exists or not in the graph defined above, is denoted by  $\rho_c$ . In other words,  $\rho_c > 0$  denotes the critical intensity below which an infinite connected component does not exist in the Boolean model. The non-triviality of  $\rho_c$  (i.e.,  $0 < \rho_c < \infty$ ) in dimensions  $d \geq 2$  as well as a sharp phase-transition is known; see Hall [68], Meester and Roy [109].

In this section, we exploit the existence of a sub-critical regime where the graph  $\mathcal{G}(\mathcal{P})$  ‘sharply’ fails to percolate when  $\rho \in (0, \rho_c)$ . Even if the actual intensity  $\rho$  is arbitrarily large, we thin the Poisson process with parameter  $q < \rho_c/\rho$ , yielding  $\mathcal{P}_{*q}$  (a Poisson point process of intensity  $q\rho$ ), placing our auxiliary model of disagreement percolation in the subcritical regime. By leveraging the exponential decay of the probability of long paths in this regime, which arises from the sharp phase transition in the Boolean model proved in [159], we identify the value  $\max\{1, \rho_c/\rho\}$  as an upper bound for the critical probability  $q_c$  i.e.,  $q_c \geq \max\{1, \rho_c/\rho\}$ .

Consequently, although the graph has unbounded degree and we cannot establish bounds for the averaged connective constant in (8.36), but we are able to apply Proposition 8.15. This allows us to prove a central limit theorem for the Gibbs specification (8.1), provided the maximum influence remains below  $q_c$ . Importantly, as observed in models like the hard-core model or the Widom-Rowlinson model, this result holds across the entire parameter range  $\rho$  of the Boolean model including the percolating regime ( $\rho > \rho_c$ ), i.e.,  $q_c > 0$  for all  $\rho > 0$ . Furthermore, in the subcritical regime of the Boolean model ( $\rho < \rho_c$ ), we find that  $q_c = 1$ , meaning that any activity level for the Gibbs parameters is admissible, thereby extending the applicability of the central limit theorem without imposing additional constraints.

**Corollary 8.16.** (*CLT for sum of spins on Poisson-Boolean model*) Let  $\mathcal{P}$  be a stationary Poisson process on  $\mathbb{R}^d$  of intensity  $\rho$  and  $K$  be a compact set containing the origin and with a non-empty interior. Let  $\mathcal{G} = \mathcal{G}(\cdot)$  be the Poisson-Boolean graph induced by the Poisson-Boolean model on  $\mathcal{P}$ , as described above. Consider a spin configuration model with the Gibbs distribution given by (8.1), defined on the graph  $\mathcal{G}$  and with finite spin space  $\mathbb{K}$ . Assume that for some  $\epsilon > 0$ , the maximum influence (8.33) satisfies

$$q(\mathcal{G}(\mathcal{P} \cap B)) < \min\{1, \rho_c/\rho\} - \epsilon$$

almost surely for all bounded sets  $B \subset \mathbb{R}^d$  and where  $\rho_c$  is the critical intensity for percolation of the Poisson-Boolean model. Under these conditions, the spin model  $\{V_x(\mathcal{G}(\mathcal{P}_n))\}_{x \in \mathcal{P}_n}$  satisfies the central limit theorem as stated in Theorem 8.4, provided that the moment and variance conditions specified in the theorem are also met.

One can consider a more general Boolean model with random grains, i.e., the fixed compact set  $K$  is replaced by random ‘shapes’  $Z_x, x \in \mathcal{P}$ . In case  $Z_x$ ’s are bounded, our approach can be adapted by extending the framework of stabilizing graphs from Section 8.1.1 to allow for graphs constructed on marked point processes. In particular, the percolation-theoretic results that we shall use are already available in this case; see [159]. The case of unbounded grains requires more careful investigation and in particular suitable moment assumptions on the grains may need to be imposed; see [52]. It may be possible to consider Poisson-Delaunay graphs, leveraging the results of [51] on site percolation in Poisson-Delaunay graphs.

*Proof of Corollary 8.16.* We will apply Proposition 8.15 showing that  $q_c \geq \min\{1, \rho_c/\rho\}$ , where  $q_c$  is defined in Definition 8.14.<sup>3</sup> A key result (see, for example, [159, Theorem 1.1; see also Theorem 3.1]) establishes an exponential bound on the probability that, in the subcritical Boolean model ( $\rho < \rho_c$ ), there exists a *continuum path* in  $\mathcal{O}$  from the origin to the exterior of the Euclidean ball  $B_r(\mathbf{0})$ :

$$\mathbb{P}_{\mathbf{0}} \left( \mathbf{0} \overset{\mathcal{O}}{\longleftrightarrow} B_r(\mathbf{0})^c \right) \leq e^{-cr}, \quad (8.48)$$

for some constant  $c = c(\rho)$ .

This result yields the exponential decay of the functions  $\tilde{\alpha}_{p,q}(r)$  in (8.43) for  $q < q_c$ , seen as follows. Specifically, we recall the notation from Definition 8.14: for  $\mathcal{X}' \subset \mathcal{X}$ ,  $\{x \overset{\mathcal{X}'_q}{\longleftrightarrow} \mathcal{Y}\} = \{x \overset{\mathcal{G}(\mathcal{X}) \cap \mathcal{X}'_q}{\longleftrightarrow} \mathcal{Y}\}$  means that there exists an open path in the graph  $\mathcal{G}(\mathcal{X})$  (here induced by the Boolean model) connecting  $x \in \mathcal{X}'$  to  $\mathcal{Y} \subset \mathcal{X}$  under site percolation, where vertices in  $\mathcal{X}'$  are independently declared open with probability  $q$ , while all other vertices in  $\mathcal{X}$  are always open. The probability measure  $\mathbb{P}_{[\mathcal{z}]_1^p, q}$  incorporates this additional percolation randomness on top of Palm distribution.

<sup>3</sup>In fact, there is equality:  $q_c = \min\{1, \rho_c/\rho\}$ . Indeed, in the subcritical regime ( $\rho < \rho_c$ ), the arguments developed below apply for  $q = 1$ , resulting in  $q_c = 1$ . Conversely, in the supercritical regime ( $\rho > \rho_c$ ), any  $q$ -thinning of the Boolean model with  $q > \rho_c/\rho$  keeps it in the supercritical regime, causing the functions  $\tilde{\alpha}_{p,q}(m)$  in (8.42) to converge to a positive value.

With this notation and assuming also without loss of generality that  $K \subset B_1(\mathbf{0})$ , for  $n \in \mathbb{N}$ ,  $[x]_1^p \in W_n$ , and  $r > 0$ , we have

$$\mathbb{P}_{[x]_1^p, q} \left( x_1 \xleftrightarrow{(\mathcal{P}_n \cap B_r(x_1))^* q} \mathcal{P}_n \cap B_r(x_1)^c \right) \leq \mathbb{P}_{[x]_1^p, q} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_{*q})} B_{\max(0, r-2)}(x_1)^c \right). \quad (8.49)$$

Indeed, applying independent  $q$ -thinning to the original Boolean model of intensity  $\rho$  is same as considering a Boolean model of intensity  $q\rho$ . The inequality follows from the fact that we allow continuum paths to reach the entire complement of  $B_{\max(0, r-2)}(x_1)$ , rather than restricting them to grains of points in  $\mathcal{P}_n \cap B_r(x_1)^c$ ; this relaxation can only increase the probability. Moreover, thinning of  $\mathcal{P}_n$  outside  $B_r(x_1)$  is immaterial after this relaxation. Finally, considering a Boolean model outside  $W_n$  can only further increase the upper bound for a given  $n$ , in case  $B_r(x_1) \cap W_n^c \neq \emptyset$ .

Note that  $\mathcal{P}_{*q}$  is distributed as a Poisson process  $\mathcal{P}_\circ$  of intensity  $q\rho$ . Thus, we have that

$$\mathbb{P}_{[x]_1^p, q} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_{*q})} B_r(x_1)^c \right) = \mathbb{P}_{[x]_1^p} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_\circ)} B_r(x_1)^c \right).$$

We assert that, uniformly over  $[x]_1^p \subset \mathbb{R}^d$

$$\mathbb{P}_{[x]_1^p} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_\circ)} B_r(x_1)^c \right) \leq e^{-cr}, \quad (8.50)$$

for some constant  $c > 0$  and all sufficiently large  $r > 0$ , provided the Boolean model is thinned with probability  $q$  such that  $q < \rho_c/\rho$ , where  $\rho_c$  is the critical intensity for percolation of this model.

To derive the bound (8.50) from (8.48), we need to control the impact of several fixed atoms of the input process induced by higher-order Palm probabilities, which may result in looser bounds compared to the analysis in [159], where Palm conditioning is applied at a single point placed at the origin.

Fix  $r > 4p$ . By the Slivnyak-Mecke theorem, (see e.g. [8, Theorem 3.2.4]) we have that

$$\mathbb{P}_{[x]_1^p} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_{*})} B_r(x_1)^c \right) = \mathbb{P} \left( x_1 \xleftrightarrow{\mathcal{O}(\mathcal{P}_\circ + \sum_{i=1}^p \delta_{x_i})} B_r(x_1)^c \right)$$

where  $\mathcal{O}(\mathcal{P}_\circ + \sum_{i=1}^p \delta_{x_i})$  denotes the existence of a path in the Boolean model augmented with the fixed points  $x_i$  and grains  $K$  at these sites i.e.,  $\mathcal{O}(\mathcal{P}_\circ + \sum_{i=1}^p \delta_{x_i}) := \bigcup_{x \in \mathcal{P}_\circ \cup \{x_1, \dots, x_p\}} (x \oplus K)$ .

Under the assumption  $q < \min\{1, \rho_c/\rho\}$ , the model  $\mathcal{P}_\circ$  is in its subcritical regime. To leverage equation (8.48), we must address the ‘undesired’ extra points arising from  $\sum_{i=1}^p \delta_{x_i}$ . To do this, we subdivide  $B_r(x_1)$  into annuli  $A_1, \dots, A_{10p}$  of width  $\frac{r}{10p}$ , where:

$$A_1 := B_{r/(10p)}(x_1), \quad A_i := B_{ir/(10p)}(x_1) \setminus B_{(i-1)r/(10p)}(x_1), \quad i = 2, \dots, 10p.$$

By the pigeonhole principle, the choice of  $r$ , and the set inclusion  $K \subset B_1(\mathbf{0})$ , for sufficiently large  $r$  (precisely when  $2 < \frac{5r}{10p}$ ), there exist five consecutive annuli, say  $A_j, \dots, A_{j+4}$  for some  $j \in \{1, \dots, 10p - 4\}$ , such that

$$\bigcup_{i=1}^p B_1(x_i) \cap (A_j \cup \dots \cup A_{j+4}) = \emptyset.$$

As a result, these five annuli remain unaffected by the extra points and their grains in  $\sum_{i=1}^p \delta_{x_i}$ . If we further increase  $r$  such that  $2 < \frac{r}{10p}$ , then any path from  $x_1$  to  $B_r(x_1)^c$  must pass through a Poisson point  $x \in \mathcal{P}_\circ \cap A_{j+2}$  since the grain  $K$  is within a unit ball. Consequently, we have:

$$\begin{aligned} & \mathbb{P} \left( x_1 \xrightarrow[\leftarrow]{\mathcal{O}(\mathcal{P}_\circ + \sum_{i=1}^p \delta_{x_i})} B_r(x_1)^c \right) \\ & \leq \sum_{j=3}^{10p-2} \mathbb{P} \left( \text{there exists } x \in \mathcal{P}_* \cap A_j \text{ such that } x \xrightarrow[\leftarrow]{\mathcal{O}(\mathcal{P}_\circ)} B_{r/(10p)}(x)^c \right). \end{aligned}$$

For a fixed  $j = 3, \dots, 10p - 2$ , we now bound the probability using the Campbell-Little-Mecke formula (2.3) and the bound (8.48):

$$\begin{aligned} & \mathbb{P} \left( \text{there exists } x \in \mathcal{P}_* \cap A_{j+2} \text{ such that } x \xrightarrow[\leftarrow]{\mathcal{O}(\mathcal{P}_\circ)} B_{r/(10p)}(x)^c \right) \\ & \leq \mathbb{E} \left[ \sum_{x \in \mathcal{P}_* \cap A_{j+2}} \mathbf{1}(x \xrightarrow[\leftarrow]{\mathcal{O}(\mathcal{P}_\circ)} B_{r/(10p)}(x)^c) \right] \\ & = q\rho \int_{A_j} \mathbb{P}_x \left( x \xrightarrow[\leftarrow]{\mathcal{O}(\mathcal{P}_\circ)} B_{r/(10p)}(x)^c \right) dx \\ & \leq q\rho \text{Vol}(A_j) e^{-cr/(10p)}, \end{aligned}$$

where in the last inequality, we used stationarity and equation (8.48) with  $\rho$  replaced by  $q\rho < \rho_c$ . Here,  $\text{Vol}(A_j)$  denotes the volume of the annulus  $A_j$ .

Summing over  $j = 3, \dots, 10p - 2$  and combining with the above derivations, we obtain the necessary exponential bound for the left hand side in (8.50):

$$q\rho \text{Vol}(B_r) e^{-cr/(10p)} = q\rho\theta_d r^d e^{-cr/(10p)} = Cq\rho e^{-cr/(10p)},$$

for some constant  $C$  depending on  $p$ . Accounting for the maximal grain diameter by shifting the argument  $r$  by 2 in the exponential function above yields the exponential bound for (8.49).

This establishes the exponential decay of the functions  $\tilde{\alpha}_{p,q}(r)$  in (8.42) and thus the corollary as argued before (8.50).  $\square$

## 8.4 Strong spatial mixing for the hard-core model via combinatorial techniques

Recall that the hard-core model is specified by the probability distribution

$$\pi_{[z]_1^m}([v]_1^m) = \prod_{(x_i, x_j) \in \mathcal{G}(\mathcal{X})} \mathbf{1}(v_i v_j = 0) \prod_{x_i \in \mathcal{X}} \lambda^{v_i}, \quad v_i \in \{0, 1\}, \quad (8.51)$$

where  $\lambda > 0$  is the *activity parameter*,  $\mathcal{G}(\mathcal{X})$  is the interaction graph on  $\mathcal{X} = \{x_1, \dots, x_m\}$ . In this section, we revisit this model using the combinatorial techniques of Weitz [154] and Sinclair et al. [143]. Their analysis establishes a stronger form of spatial mixing, extending the range of low-activity parameters for the Gibbs hard-core model beyond the maximum influence conditions typically

employed to demonstrate weak spatial mixing via disagreement percolation. This is achieved by making use of a stronger spatial mixing property established for this Gibbs model in the aforementioned works. Specifically, this approach demonstrates the strong mixing bound (8.29), at the level of individual sites  $\mathcal{X}'' = \{x\}$ . Using the additive representation provided by Lemma 8.11 we revisit the proof of Lemma 8.7, a key argument in the main result of Theorem 8.4, and which is based on averaged weak spatial mixing. Consequently, we consider the range of the activity parameter  $\lambda$  above the critical value

$$\lambda_c(\Delta) := \frac{\Delta^\Delta}{(\Delta - 1)^{\Delta+1}} > \frac{1}{\Delta - 1}, \quad \text{for } \Delta \geq 1 \text{ and } \lambda_c(0) = \infty, \quad (8.52)$$

where the (*quenched*) *connective constant*  $\Delta \in [0, \infty]$  is defined as the smallest constant such that for all  $m \geq m_0 \in \mathbb{N}$  and for all  $x \in \mathcal{X}$ , it holds that

$$N(x, m; \mathcal{G}(\mathcal{X})) \leq \Delta^m, \quad (8.53)$$

with  $N(x, m; \mathcal{G}(\mathcal{X}))$  denoting the number of self-avoiding paths of length  $m$  starting from  $x \in \mathcal{X}$  in the graph  $\mathcal{G}(X)$ ; see [143, Section 2.6]. However working with this quenched connective constant comes at the cost of requiring almost sure (quenched) control over the graph's connectivity. Observe that the quenched connective constant cannot take fractional values in the range  $\Delta \in (0, 1)$ . Additionally, a value of  $\Delta = 0$  indicates that the graph consists of uniformly bounded components and one can take large activity parameters as well, which is consistent with  $\lambda_c(0) = \infty$ . The combinatorial techniques can also be applied to Monomer-Dimer model and Ising model with zero external field and hence our results also can be extended to these models; see the discussion at the beginning of Subsection 8.4.1. The critical value  $1/\bar{\Delta}$  established in Proposition 8.13 may not be comparable to  $\lambda_c(\Delta)$  as  $\bar{\Delta} \leq \Delta$ .

For deterministic graphs, the critical value  $\lambda_c(\Delta)$  was first identified for hard-core models on infinite  $d$ -ary trees with  $\Delta = d$ , as it characterizes the regime where correlations decay exponentially and consequently enabled efficient algorithms for sampling and computation of the partition function. The work [143] further generalized it to all finite or infinite graphs with a connective constant bounded by  $\Delta$ . These works, however, do not address limit theorems for statistics of random graphs. Recall that  $N(x, m; \mathcal{G}(\mathcal{X}))$  is the number of self-avoiding paths of length  $m$  starting from  $x \in \mathcal{X}$  in the graph  $\mathcal{G}(\mathcal{X})$ .

**Proposition 8.17** (CLT for the sum of spins in the hard-core model). *Let  $\mathcal{P}$  be a stationary process on  $\mathbb{R}^d$  with summable exponential mixing correlations as in Definition 4.8 and having bounded Palm intensity functions as at (8.10). Let  $\mathcal{G} = \mathcal{G}(\sim, \cdot)$  be a stabilizing interaction graph on  $\mathcal{P}$  and satisfying the following a.s. connective constant bound: there exists  $\Delta \in [1, \infty)$  and  $m_0 \in \mathbb{N}$  such that for all  $B \in \mathcal{B}_b$ ,  $x \in \mathcal{P} \cap B$ , and  $m \geq m_0$*

$$N(x, m; \mathcal{G}(\mathcal{P} \cap B)) \leq \Delta^m. \quad (8.54)$$

*Consider the hard-core Gibbs model (8.51) with activity parameter  $\lambda$  such that*

$$\lambda < \lambda_c(\Delta),$$

where the critical value  $\lambda_c(\Delta)$  is at (8.52). Under these conditions, the spin configuration  $\{V_x(\mathcal{G}(\mathcal{P}_n))\}_{x \in \mathcal{P}_n}$  satisfies the CLT as stated in Theorem 8.4, provided that the moment and variance conditions there are also satisfied.

*Proof.* We cannot directly use Theorem 8.4 but rather follow the approach therein by establishing Lemma 8.7 by suitably adapting the proof by borrowing the bounds from [143].

The case of  $\Delta = 0$  is straightforward, as the central limit theorem follows directly from the presence of uniformly bounded components in the graph. For  $\Delta \geq 1$ , the proof hinges on the single-site strong spatial mixing property of the hard-core model on graphs, as demonstrated in [143]. This property relies on the ‘self-avoiding walk representation’ introduced by [155], which simplifies the problem by reducing it to trees. The (quenched) spatial mixing property is established and analyzed through the study of occupation ratios. Specifically, in this model, the distribution of the spin at a given site can be described by the occupation probability:

$$p(x, \mathcal{X}; v_{\mathcal{Z}}) := \mathbb{P}(V_x = 1 \mid V_{\mathcal{Z}} = v_{\mathcal{Z}}),$$

which represents the probability that the spin at node  $x \in \mathcal{X} \setminus \mathcal{Z}$  in the graph  $\mathcal{G}(\mathcal{X})$  equals 1, conditioned on the spins at the subset  $\mathcal{Z} \not\ni x$  being fixed to the values  $v_{\mathcal{Z}}$ . We analyze the decay of the dependence of the spin  $V_x$  on the boundary condition  $V_{\mathcal{Z}} = v_{\mathcal{Z}}$  through the *occupation ratio* at node  $x$ , namely

$$R(x, \mathcal{X}; v_{\mathcal{Z}}) := \frac{p(x, \mathcal{X}; v_{\mathcal{Z}})}{1 - p(x, \mathcal{X}; v_{\mathcal{Z}})}.$$

The total variation distance between the conditional distributions of  $V_x$ , given two different boundary conditions, is bounded as follows:

$$\begin{aligned} d_{\text{TV}}(\mathcal{L}(V_x \mid V_{\mathcal{Z}} = v_{\mathcal{Z}}), \mathcal{L}(V_x \mid V_{\mathcal{Z}} = z_{\mathcal{Z}})) &= |p(x, \mathcal{X}; v_{\mathcal{Z}}) - p(x, \mathcal{X}; z_{\mathcal{Z}})| \\ &\leq |R(x, \mathcal{X}; v_{\mathcal{Z}}) - R(x, \mathcal{X}; z_{\mathcal{Z}})|. \end{aligned} \quad (8.55)$$

Thus, a ‘spatial mixing property of the occupation ratios’ implies spatial mixing for the model.

In the hard-core model (8.51) on a general graph  $\mathcal{G}(\mathcal{X})$ , the strong spatial mixing of occupation ratios’ as established in [143], may be stated as follows: For any vertex  $x \in \mathcal{X}$  and any subset  $\mathcal{Z} \subset \mathcal{X}$  not containing  $x$ , it is the case that there exists  $\alpha < 1/\Delta$  such that

$$|R(x, \mathcal{X}; v_{\mathcal{Z}}) - R(x, \mathcal{X}; z_{\mathcal{Z}})| \leq CN(x, l - 1; \mathcal{G}(\mathcal{X}))^{\frac{1}{q}} \alpha^{\frac{l-1}{q}}, \quad (8.56)$$

where  $l = d_{\mathcal{G}(\mathcal{X})}(x, \mathcal{Z}_{\neq})$  is the distance between  $x$  and points in  $\mathcal{Z}$  where the boundary conditions  $v_{\mathcal{Z}}$  and  $z_{\mathcal{Z}}$  differ, and where it is required that  $l$  is non-zero. The constants  $C, q$  depend on  $\lambda$  and are specified explicitly in the sequel. The decay constant  $\alpha$  is defined to be  $\alpha := 1/\Delta_c(\lambda)$ , where  $\Delta_c(\lambda)$  is determined as the unique solution to the equation:

$$\lambda = \frac{t^t}{(t-1)^{t+1}}. \quad (8.57)$$

Observe that the function  $t \mapsto t^t/(t-1)^{t+1}$  in (8.57) is decreasing. Recall that  $\Delta \in [1, \infty)$  is a bound on the connective constant, as in (8.54) and  $\lambda_c(\Delta)$  is the critical value as in (8.52). Thus,

our assumption  $\lambda < \lambda_c(\Delta)$  implies that  $\Delta_c(\lambda)$ , the solution in  $t$  of the equation (8.57), satisfies  $\Delta_c(\lambda) > \Delta \geq 1$ . Consequently,  $\alpha := 1/\Delta_c(\lambda) < 1/\Delta$ .

This formulation (8.56) provides a single-site strong spatial mixing bound based on occupation ratios and as demonstrated in (8.55), it can be used in further considerations. The key ideas and steps behind the derivation of (8.56) are in Section 8.4.1.

We recognize that  $\alpha$  serves as a refinement of the requirement that the maximum influence  $q = q(\mathcal{G}(\mathcal{X}))$ <sup>4</sup> in Proposition 8.13 is smaller than the (albeit averaged) connective constant  $\bar{\Delta}$ . Although multi-site weak spatial mixing can be derived (see (8.32)) from the strong spatial mixing single-site bound (8.56), the resulting estimate would be too crude to establish BL-stabilization (8.13), crucial to proving Theorem 8.4. Instead, we shall establish (8.13) by combining the multi-site bound (8.32) with the additive representation for strong spatial mixing provided in Lemma 8.11, using a stronger, quenched (almost sure) bound on the connective constant  $\Delta$ .

To prove BL-stabilization (8.13) under the new assumptions, we only need to refine the inequalities (8.18) and (8.19) in the proof of Lemma 8.7. Using this strong spatial mixing bound (8.32) and letting  $\mathcal{X}' = \mathfrak{B}_{k_r}([x]_1^p)$ , we have:

$$\begin{aligned} & \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p} | \mathcal{P}_n)) \right] \\ & \leq \mathbb{E}_{[x]_1^p} \left[ \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \mathcal{A}'([x]_1^p, \partial_{z \neq w} \mathcal{X}') \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'}^{[k_r]} = w_{\partial\mathcal{X}'}) \right]. \end{aligned}$$

Next, applying the additive representation (8.31) for strong spatial mixing along with (8.56) (linked to (8.55)) and taking the function

$$\mathcal{A}'_{\mathcal{G}(\mathcal{P}_n)}(\{x\}, \mathcal{Z}_{\neq}) = N(x, d_{\mathcal{G}(\mathcal{P}_n)}(x, \mathcal{Z}_{\neq}) - 1; \mathcal{G}(\mathcal{P}_n))^{1/q} \alpha^{(d_{\mathcal{G}(\mathcal{P}_n)}(x, \mathcal{Z}_{\neq}) - 1)/q},$$

this gives:

$$\begin{aligned} & \mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p} | \mathcal{P}_n)) \right] \tag{8.58} \\ & \leq C \mathbb{E}_{[x]_1^p} \left[ \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \sum_{i=1}^p N(x_i, d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial_{z \neq w} \mathcal{X}') - 1; \mathcal{G}(\mathcal{P}_n))^{1/q} \alpha^{(d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial_{z \neq w} \mathcal{X}') - 1)/q} \right. \\ & \quad \left. \times \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'}^{[k_r]} = w_{\partial\mathcal{X}'} \right]. \end{aligned}$$

Since  $\mathcal{X}' = \mathfrak{B}_{k_r}([x]_1^p)$  is the subgraph induced in  $\mathcal{P}_n$  by the graph balls centered at  $x_i$ ,  $i = 1, \dots, p$ , of radius  $k_r$ , we have  $d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial_{z \neq w} \mathcal{X}') > k_r$  (interpreted as  $= \infty$  if  $\partial_{z \neq w} \mathcal{X}' = \emptyset$ ), for all  $z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}$  and  $i = 1, \dots, p$ . By assumption (8.54), for  $k_r > m_0$ ,

$$N(x_i, d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial_{z \neq w} \mathcal{X}') - 1; \mathcal{G}(\mathcal{P}_n)) \leq \Delta^{d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial_{z \neq w} \mathcal{X}') - 1},$$

Substituting this into (8.58), we obtain:

$$\mathbb{E}_{[x]_1^p} \left[ d_{\text{BL}}(\mathcal{L}(V_{[x]_1^p}^{\cup[k_r]} | \mathcal{P}_n), \mathcal{L}(V_{[x]_1^p} | \mathcal{P}_n)) \right]$$

---

<sup>4</sup>The maximum influence  $q = q(\mathcal{G}(\mathcal{X}))$ , defined in (8.33) and not used any more in this proof, should not be confused with the constant  $q$  in (8.56). Both depend on the activity  $\lambda$  of the hard-core model, but the former also depends on the underlying graph.

$$\begin{aligned}
&\leq C \mathbb{E}_{[x]_1^p} \left[ \sum_{z_{\partial\mathcal{X}'}, w_{\partial\mathcal{X}'} \in \mathbb{K}^{\partial\mathcal{X}'}} \sum_{i=1}^p (\Delta\alpha)^{(d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial\mathcal{X}')-1)/q} \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'} = z_{\partial\mathcal{X}'}) \mathbb{P}_{\mathcal{G}(\mathcal{P}_n)}(V_{\partial\mathcal{X}'}^{[k_r]} = w_{\partial\mathcal{X}'}) \right] \\
&\leq C p(\Delta\alpha)^{(k_r)/q},
\end{aligned}$$

where, in the last inequality, we used  $d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial\mathcal{X}') \geq d_{\mathcal{G}(\mathcal{P}_n)}(x_i, \partial\mathcal{X}') > k_r$  and the fact the  $\Delta\alpha < 1$  (as a consequence of our assumption  $\lambda < \Delta^\Delta/(\Delta - 1)^{\Delta+1}$  and the representation of  $\alpha = 1/\Delta_c(\lambda) < 1/\Delta$ ). Choosing  $k_r = \lceil r^b \rceil$ , with  $b \in (0, 1)$ , the term  $(\Delta\alpha)^{k_r/q} \leq (\Delta\alpha)^{(r^b)/q}$  decreases rapidly as  $r$  increases. Thus, we achieve the same conclusion as in (8.18).

The same reasoning applies when modifying (8.19), thereby confirming the result of Lemma 8.7 whenever the strong spatial mixing bound in (8.32) decays rapidly. The latter is guaranteed by our assumption on  $\lambda$  and (8.56). This allows us to use Lemma 8.7 in the proof of Theorem 8.4 giving the asymptotic normality for the sum of spins in the hard-core model.  $\square$

#### 8.4.1 Derivation of (8.56)—Strong spatial mixing for the hard-core model; [143], [155]

In the following, we provide additional details on how the strong spatial mixing (8.56) is established in [143] for the hard-core model. Since the specific form of the bound we use may not be transparent in the proof of Theorem 1 in [143], we describe the key steps of the derivation here for self-containment. Broadly the proof of (8.56) involves two steps, firstly reducing the right-hand side in (8.56) to that of a tree via the ‘Weitz reduction’ and then a very involved analysis of the same for a tree using message passing recursions. In particular, for the second step, we refer many times to specific claims in [143] for proofs. We shall explain these steps and then describe how the proof follows from these two steps.

A similar approach was used in [143] to establish strong spatial mixing for the monomer-dimer model, where instead of independent vertices in the hard-core model, random matchings on the graph are considered, also controlled by the activity parameter; see also [126]. One can also apply this approach to the Ising model and random cluster model; see [142, Theorem 28]. This suggests that the central limit theorem should hold for these models as well under an analogous range of the activity parameters.

**FIRST STEP:** The ‘Weitz reduction’ or ‘self-avoiding walk (SAW) representation’ as stated in [143, Theorem 4] is based on the original result of [155]. The SAW tree (where SAW stands for Self-Avoiding Walk)  $\mathcal{T}_{SAW} = \mathcal{T}_{SAW}(x, \mathcal{G}(\mathcal{X}))$  represents all possible self-avoiding paths in  $\mathcal{G}(\mathcal{X})$  that start from  $x$ . Its root corresponds to the trivial self-avoiding walk that begins and ends at  $x$ . The tree is built recursively: for each node  $w$  in  $\mathcal{T}_{SAW}$ , its children represent all possible extensions of the self-avoiding path associated with  $w$ , each extended by exactly one additional step. The theorem asserts that, for any vertex  $x$  in a (finite or locally finite, infinite) graph  $\mathcal{G}(\mathcal{X})$ , the tree  $\mathcal{T}_{SAW}$  rooted at  $x$  may be equipped with a boundary condition  $w_{\mathcal{W}}$  on a subset  $\mathcal{W} \subset \mathcal{T}_{SAW}$  (see [143, Appendix A] and [155, Section 3]) such that, for the hard-core Gibbs model on  $\mathcal{G}(\mathcal{X})$  specified by (8.51) with activity parameter  $\lambda > 0$ , and any boundary condition  $v_{\mathcal{Z}}$  on  $\mathcal{Z} \subset \mathcal{G}(\mathcal{X})$ , we have:

$$R(x, \mathcal{X}; v_{\mathcal{Z}}) = R(x, \mathcal{T}_{SAW}; w_{\mathcal{W}} \cup v_{\mathcal{Z}}), \quad (8.59)$$

where  $v_{\mathcal{Z}}$  on the right-hand side denotes the translation of the boundary condition from  $\mathcal{G}(\mathcal{X})$  to  $\mathcal{T}_{SAW}$ . The combined boundary condition  $w_{\mathcal{W}} \cup v_{\mathcal{Z}}$  is defined by first applying  $w_{\mathcal{W}}$  and then  $v_{\mathcal{Z}}$ , with  $v_{\mathcal{Z}}$  overriding  $w_{\mathcal{W}}$  wherever they both apply and differ. Thanks to the above reduction, to prove (8.56) it now suffices to bound

$$|R(x, \mathcal{T}_{SAW}; w_{\mathcal{W}} \cup v_{\mathcal{Z}}) - R(x, \mathcal{T}_{SAW}; w_{\mathcal{W}} \cup z_{\mathcal{Z}})|. \quad (8.60)$$

Thus the difference on the boundary due to the conditions  $w_{\mathcal{W}} \cup v_{\mathcal{Z}}$ ,  $w_{\mathcal{W}} \cup z_{\mathcal{Z}}$  arises from difference between  $v_{\mathcal{Z}}$  and  $z_{\mathcal{Z}}$ .

**SECOND STEP:** We describe how to derive bounds for (8.60) for any tree. Consider a tree  $\mathcal{T}$  rooted at  $x$  and denote the graph metric on  $\mathcal{T}$  by  $d_{\mathcal{T}}$ . For a finite tree  $\mathcal{T}$ , a *cutset*  $\mathcal{C}$  is a set of vertices in  $\mathcal{T}$  such that (i) any path from the root  $x$  to a leaf  $y$  where the graph distance  $d_{\mathcal{T}}(x, y) \geq \max_{y' \in \mathcal{C}} d_{\mathcal{T}}(x, y')$  must pass through  $\mathcal{C}$ , and (ii) the vertices in  $\mathcal{C}$  form an ‘antichain’, meaning that for any two vertices in  $\mathcal{C}$ , neither is an ancestor of the other in  $\mathcal{T}$ . For rooted, locally finite, infinite trees, condition (i) is replaced by the condition that any infinite path starting from the root  $x$  must pass through  $\mathcal{C}$ . A key example of a cutset we will use later is the set of all vertices at a fixed distance  $l$  from  $x$  in  $\mathcal{T}$ . For a cutset  $\mathcal{C}$ , let  $\mathcal{T}_{\leq \mathcal{C}}$  be the subtree of  $\mathcal{T}$  obtained by removing the descendants of vertices in  $\mathcal{C}$ , and let  $\mathcal{T}_{< \mathcal{C}}$  denote the subtree rooted at  $x$  obtained by removing the vertices in  $\mathcal{C}$ . Also, let  $\mathcal{C}^+$  denote the set of the children of the vertices in  $\mathcal{C}$  (here  $\mathcal{C}$  may not necessarily form a cutset). For  $y \in \mathcal{T}$ , we denote by  $\mathcal{T}_y$  the sub-tree of  $\mathcal{T}$  rooted at  $y$ .

We assert that the following adaptation of [143, Lemma 3] applies: For the hard-core Gibbs model specified by (8.51) with activity parameter  $\lambda > 0$ , there exist constants  $C = C(\lambda) < \infty$ ,  $q = q(\lambda) > 1$ , and  $\alpha = \alpha(\lambda) > 0$  such that this specification, when applied to any finite or infinite locally finite tree  $\mathcal{T}$  rooted at  $x$ , satisfies the following decay of the occupation ratio at the root:

$$|R(x, \mathcal{T}; v_{\mathcal{Z}}) - R(x, \mathcal{T}; z_{\mathcal{Z}})|^q \leq C^q \sum_{y \in \mathcal{C}} \alpha^{d_{\mathcal{T}}(x, y)}, \quad (8.61)$$

where  $\mathcal{C}$  is a cutset in  $\mathcal{T}$  located at a distance of at least 1 from the root  $x$ . Here,  $\mathcal{Z}$  is a subset of  $\mathcal{T}_{\leq \mathcal{C}^+}$  not containing  $x$ , and  $v_{\mathcal{Z}}$  and  $z_{\mathcal{Z}}$  are two boundary configurations on  $\mathcal{Z}$  that differ only on  $\mathcal{C}^+$  (agreeing on the spins fixed in  $\mathcal{Z} \cap \mathcal{T}_{\leq \mathcal{C}}$ ).

Note that the above claim constitutes a slight, two-fold modification of [143, Lemma 3]: we have dropped the requirement that  $\mathcal{C}$  has no leaves, and we formulate the conditional probability of the spin (and hence the occupation ratio) at  $x$  using our boundary condition terminology—represented by  $v_{\mathcal{Z}}, z_{\mathcal{Z}}$ —rather than the ‘initial conditions’ used in the original work, which refer to fixed values of the occupation ratios  $R$  at the vertices in  $\mathcal{Z}$ . Note that the occupation ratio takes real values in  $[0, \lambda]$  at all vertices, except it is infinite at points where the boundary condition sets the spin to 1. To avoid this issue in the proof,  $\mathcal{C}$  was previously assumed to have no leaves.

The rest of this step is to explain choices of  $C, q, \alpha$  along with details showing how (8.61) can be derived from [143, Lemma 3] by substituting specific choices from [143, Section 4] pertaining to the hard-core model. Our notation coincides with that in [143]. As in [143], we choose  $\phi(x) = \sinh^{-1}(\sqrt{x}) = \operatorname{arcsinh}(\sqrt{x})$  with derivative  $\phi'(x) = 1/(2\sqrt{x(1+x)})$  and so we have that  $M :=$

$\sup_{x \in [0, \lambda]} \phi(x) = \sinh^{-1}(\sqrt{\lambda})$  and  $L := \inf_{x \in [0, \lambda]} \phi'(x) = 1/(2\sqrt{\lambda(1+\lambda)})$ . Since the occupation ratio at any internal node of the tree is upper bounded by  $\lambda$ , we have via the mean-value theorem that

$$|R(x, \mathcal{T}; v_{\mathcal{Z}}) - R(x, \mathcal{T}; z_{\mathcal{Z}})| \leq L^{-1} |\phi(R(x, \mathcal{T}; v_{\mathcal{Z}})) - \phi(R(x, \mathcal{T}; z_{\mathcal{Z}}))|.$$

Putting the constant  $C$  in (8.61) to be

$$C := \frac{M}{L} = 2 \sinh^{-1}(\sqrt{\lambda}) \sqrt{\lambda(1+\lambda)}$$

it suffices to show

$$|\phi(R(x, \mathcal{T}; v_{\mathcal{Z}})) - \phi(R(x, \mathcal{T}; z_{\mathcal{Z}}))|^q \leq M^q \sum_{y \in \mathfrak{C}} \alpha^{d_{\mathcal{T}}(x, y)}. \quad (8.62)$$

This claim is the analogue of [143, (20)] but with boundary conditions on  $\mathcal{Z} \subset \mathcal{T}_{\leq \mathfrak{C}^+}$  (instead of  $R$ -related initial condition on  $\mathcal{T}_{\leq \mathfrak{C}^+}$ , as discussed above) and is similarly proved inductively on  $d_{\mathcal{T}}(x, \mathfrak{C}) = \inf_{y \in \mathfrak{C}} d_{\mathcal{T}}(x, y)$ . Before sketching the proof, we need to mention the choices of  $q, \alpha$  for which the above claim holds.

Let  $\lambda_c(t) : (1, \infty) \rightarrow (0, \infty)$  be the strictly decreasing function  $t \mapsto \frac{t^t}{(t-1)^{t+1}}$  and let  $\Delta_c(\lambda)$  be the unique solution in  $t$  to the equation  $\lambda_c(t) = \lambda$  for  $\lambda \in (0, \infty)$  i.e.,  $\lambda_c(\Delta_c(\lambda)) = \lambda$ . As explained below (8.56),  $\Delta_c(\lambda) > \Delta \geq 1$  if  $\lambda < \lambda_c(\Delta)$  where  $\Delta$  is the upper bound on the connective constant. Assuming this, the decay constant  $\alpha = \alpha(\lambda)$  (defined in [143, (9)]) and the exponent  $q = q(\lambda)$  are defined as follows

$$\alpha(\lambda) := \frac{1}{\Delta_c(\lambda)} \text{ and } q := 1 - \frac{\Delta_c(\lambda) - 1}{2} \log\left(1 + \frac{\Delta_c(\lambda) - 1}{2}\right),$$

where the well-definedness follows because  $\Delta_c(\lambda) > 1$ ; see Lemma 6 of [143]. Additionally, as in [143], take  $1/a = 1 - 1/q$ , though this exponent will not appear explicitly in our derivations.

Now we sketch why (8.62) holds with the above choices of  $\alpha, q$ . Consider the base case  $d_{\mathcal{T}}(x, \mathfrak{C}) = 1$ . Let  $y_1, \dots, y_k$  be the immediate descendants of  $x$  in the tree  $\mathcal{T}$ . Suppose  $R(y_i, \mathcal{T}_{y_i}; v_{\mathcal{Z}}) = \infty$  for some  $y_i \in \mathcal{Z} \cap \mathfrak{C}$ . Then by the assumption that boundary conditions are differing only on  $\mathfrak{C}^+$ ,  $R(y_i, \mathcal{T}_{y_i}; z_{\mathcal{Z}}) = \infty$  as well and in this case  $R(x, \mathcal{T}; v_{\mathcal{Z}}) = R(x, \mathcal{T}; z_{\mathcal{Z}}) = 0$ . Thus (8.62) holds trivially here. Otherwise, we have  $R(y, \mathcal{T}_y; v_{\mathcal{Z}}) \in [0, \lambda]$  for all  $y \in \mathfrak{C}$  and [143, Lemma 2] along with definition of  $M$  gives that

$$|\phi(R(x, \mathcal{T}; v_{\mathcal{Z}})) - \phi(R(x, \mathcal{T}; z_{\mathcal{Z}}))|^q \leq \alpha \sum_{y_i \in \mathfrak{C}} |\phi(R(y_i, \mathcal{T}_{y_i}; v_{\mathcal{Z}})) - \phi(R(y_i, \mathcal{T}_{y_i}; z_{\mathcal{Z}}))|^q \leq M^q \sum_{y \in \mathfrak{C}} \alpha,$$

thereby immediately proving (8.62) in this case as well. Thus the base case holds in (8.62).

Now, for some  $k \in \mathbb{N}$ , assume (8.62) for  $d_{\mathcal{T}}(x, \mathfrak{C}) \leq k$ . As above, consider immediate descendants  $y_i$  of  $x$  in the tree  $\mathcal{T}$ . Either accounting for possibility that  $R(y_i, \mathcal{T}_{y_i}; v_{\mathcal{Z}}) = R(y_i, \mathcal{T}_{y_i}; z_{\mathcal{Z}}) = \infty$  for  $y_i \in \mathcal{Z} \cap \mathcal{T}_{\leq C}$  or using [143, Lemma 2] as in the first inequality above and induction hypothesis, one

can obtain (8.62). As argued before (8.62), this suffices to prove (8.61).

PROOF OF (8.56) FROM THE FIRST TWO STEPS: Choose  $\mathfrak{C}$  to be set of vertices at distance  $l-1$  from  $x$  in  $\mathcal{T}_{SAW}$ ; by the construction of  $\mathcal{T}_{SAW}$  these are all self-avoiding paths of length  $l-1$  from  $x$  in  $\mathcal{G}(\mathcal{X})$  and therefore  $|\mathfrak{C}| = N(x, l-1; \mathcal{G}(\mathcal{X}))$ . So  $\mathfrak{C}^+$  corresponds to the set of all self-avoiding paths from  $x$  to  $\partial\mathfrak{B}_{l-1}(x)$  where  $\mathfrak{B}_{l-1}(x) = \{y \in \mathcal{G}(\mathcal{X}) : d_{\mathcal{G}(\mathcal{X})}(x, y) \leq l-1\}$ . Since the boundary conditions  $v, z$  differ on  $\mathcal{Z}$  in  $\mathcal{G}(\mathcal{X})$  in (8.56) only when the graph distance to  $x$  is greater than or equal to  $l$ , we have that  $v_{\mathcal{Z}}, z_{\mathcal{Z}}$  may differ on paths being elements in  $\mathfrak{C}^+$  but not in  $\mathfrak{C}$ . Moreover, due to the cutset nature of  $\mathfrak{C}$  and the spatial Markov property, the spin values assigned in  $\mathcal{Z}$  on paths beyond  $\mathfrak{C}^+$  do not affect the distribution of the spin at  $x$ . Hence (8.61) holds in this case and using that  $d_{\mathcal{T}_{SAW}}(x, y) = l-1$  for  $y \in \mathfrak{C}$ , we obtain

$$\begin{aligned} |R(x, \mathcal{T}_{SAW}; w_{\mathcal{W}} \cup v_{\mathcal{Z}}) - R(x, \mathcal{T}_{SAW}; w_{\mathcal{W}} \cup z_{\mathcal{Z}})|^q &\leq C^q \sum_{y \in \mathfrak{C}} \alpha^{d_{\mathcal{T}_{SAW}}(x, y)} \\ &= C^q |\mathfrak{C}| \alpha^{l-1} = C^q N(x, l-1; \mathcal{G}(\mathcal{X})) \alpha^{l-1}. \end{aligned}$$

Thus (8.56) now follows from (8.59) and (8.60). The choices of  $C, q$  and  $\alpha$  have been explained in the Second Step above.

## 9 Interacting diffusions on spatial random graphs

In this section we define a model of interacting diffusions on spatial random graphs and establish limit theorems for statistics of these diffusions. In contrast to mean field models the diffusions considered here interact only with their finitely many neighbors. Similar to the spin variables encountered in the previous section, diffusions do not have a stopping set property, but they nonetheless satisfy fast BL-localization, as well as the even stronger  $L^2$ -stabilization, when defined on stabilizing interaction graphs. The diffusion model on graphs considered here was studied in [95], which provides conditions ensuring the existence of the model in infinite space. Building on this prior work—and in contrast to the finite-window approach commonly used for spin systems—we study in this section the limiting behavior of sums of scores of the interacting diffusions over the infinite window, using the results from Section 5.3. We refer to [95, Section 1] and [127] regarding the motivation and applications of this model. An interacting diffusion model on deterministic point sets has been considered in [30] and one may consider such models on random graphs using the methods here.

### 9.1 The diffusion model

Using the same graph notation as for spin systems in Section 8.1.1, we now consider a finite or countable set of sites (nodes), denoted by  $\mathcal{X}$ , and a locally finite graph  $\mathcal{G}(\mathcal{X})$  defined on  $\mathcal{X}$ . For each  $x \in \mathcal{X}$ , we denote by

$$N_x = N_x(\mathcal{X}) := \{x' \in \mathcal{X} : x' \sim x\}$$

the (finite) set of neighbors of  $x$  in  $\mathcal{G}(\mathcal{X})$ . Fix a time horizon  $t_0 \in [0, \infty)$  and a diffusion dimension  $d' \in \mathbb{N}$ . We consider a system of interacting  $\mathbb{R}^{d'}$ -valued diffusions

$$M(x, t) := M^{\mathcal{G}(\mathcal{X})}(x, t), \quad x \in \mathcal{X}, \quad t \in [0, t_0],$$

evolving on vertices of the graph  $\mathcal{G}(\mathcal{X})$ , interacting only via its edges, and defined by the stochastic differential equations:

$$dM(x, t) = b(t, M[x, t], M[N_x, t]) dt + \sigma(t, M[x, t], M[N_x, t]) dZ_x(t), \quad (9.1)$$

where  $b$  and  $\sigma$  are the *drift* and *diffusion* coefficients, respectively, understood as  $\mathbb{R}^{d'}$ -valued and  $\mathbb{R}^{d'} \times \mathbb{R}^{d'}$ -valued functions, and  $\{Z_x(\cdot)\}_{x \in \mathcal{X}}$  are i.i.d. standard Brownian motions in  $\mathbb{R}^{d'}$ ; in this notation  $M[x, t]$  denotes the path of  $M(x, \cdot)$  in the interval  $[0, t]$  and  $M[N_x, t] := \sum_{x' \in N_x} \delta_{M[x', t]}$ ;  $\sigma$  here is matrix-valued but treated equivalently as a vector in  $\mathbb{R}^{(d')^2}$ . Indeed, following [95], we allow path-dependence in the coefficients  $b$  and  $\sigma$  in (9.1) "both because such interactions arise in applications and because this does not complicate the arguments". The SDE (9.1) says that particles interact directly only with their (finitely many) neighbors in the graph. As is the case with the interacting particle system models considered in the next section, this is the sparse regime model, in contrast to the mean field regime model, where particles interact with all other particles.

When the *initial conditions*  $\{M(x, 0)\}_{x \in \mathcal{X}}$  are *uniformly bounded* and the drift and diffusion coefficients  $b, \sigma$  are *Lipschitz functions*, as detailed in the next section, the system (9.1) admits a *unique solution with the family of processes*  $\{M(x, t)\}_{x \in \mathcal{X}}$ , taking values in the space of continuous paths with  $t \in [0, t_0]$ , as established in [95]. In what follows we will call this collection of processes *the family of interacting diffusions* on graph  $\mathcal{G}(\mathcal{X})$ .

Large ensembles of interacting diffusions are used to model complex dynamical behavior arising in statistical physics, biology, and neural networks; see [127] and also Lucon and Stannat [102]. Mean field models have received considerable attention and their dynamics are relatively well understood when the graph  $\mathcal{G}(\mathcal{X})$  is the complete graph. The ' $P$ -nearest-neighbor model',  $P \in (0, 1)$ , has also been studied [102], where here the neighbors of a site consist of all particles within a window centered at the site and whose volume equals  $Pn$ . In this model and others, one often takes the diffusion  $\sigma$  to be a fixed constant. Here we consider diffusion models which almost surely satisfy  $P = o(1)$ , a regime which is not well studied.

**Example 9.1** (Markovian drift). We give an example of a simple Markovian model of drift and diffusion. This model has been investigated in mean-field settings; see for example [129, Example 2.1] or [102, Section 1.2.2]. In this model, we put  $\sigma \equiv 1$  (i.e., the constant function) and take  $b$  as follows:

$$b(t, M[x, t], M[N_x, t]) = b(M(x, t), M(N_x, t)) = F_1(M(x, t)) + \frac{1}{|N_x|} \sum_{y \sim x} F_2(M(x, t), M(y, t)),$$

where  $F_1, F_2$  are measurable functions on  $\mathbb{R}^{d'}$  and  $\mathbb{R}^{d'} \times \mathbb{R}^{d'}$  respectively.

### 9.1.1 Lipschitz assumption on coefficients

For classical diffusion processes defined by stochastic differential equations, Lipschitz conditions on drift and diffusion coefficients are essential to ensure the existence and uniqueness of strong solutions via Grönwall's inequality. A similar principle applies to interacting diffusions on graphs, where such conditions not only guarantee well-posedness but also imply (spatial)  $L^2$ -stabilization: given a fixed time horizon, the state at a given node becomes increasingly independent of distant nodes as their separation grows, capturing how influence disperses across the graph.

Before introducing this crucial Lipschitz condition, we need to establish some notation: Let  $\mathbb{M} := \mathcal{C}([0, t_0], \mathbb{R}^d)$  denote the path space of continuous functions, equipped with the topology of uniform convergence on compact sets. For truncated paths up to time  $t$ , where  $0 \leq t \leq t_0$ , we use the notation  $\mathbb{M}(t) := \mathcal{C}([0, t], \mathbb{R}^d)$ . The paths of the processes  $M(x, \cdot)$ ,  $x \in \mathcal{X}$ , defined in equation (9.1) and subsequently referred to as  $M[x]$ , are elements of  $\mathbb{M}$  (i.e.,  $M[x] \in \mathbb{M}$ ). Their truncations are denoted by  $M[x, t] \in \mathbb{M}(t)$  in (9.1).

More generally, for a given function  $m[x] \in M(t)$  representing a possible entire trajectory  $m(x, s)$  over  $s \in [0, t_0]$ , at  $x \in \mathcal{X}$ , its truncation to  $s \in [0, t]$  is noted by  $m[x, t]$  and its norm is defined as  $\|m[x, t]\| := \sup_{s \in [0, t]} |m(x, s)|$ , where  $|\cdot|$  denotes the Euclidean norm in  $\mathbb{R}^d$ .

With the above notation, following [95, Lipschitz Assumption A'], we impose the following Lipschitz conditions on the coefficient functions  $b$  and  $\sigma$ . These conditions hold trivially for the Markovian drift model in Example 9.1 if  $F_1, F_2$  therein satisfy the Lipschitz assumptions.

**Definition 9.2** (Lipschitz diffusion coefficients). *The jointly measurable functions  $(b, \sigma) : [0, \infty) \times \mathbb{M} \times \hat{\mathcal{N}}_{\mathbb{M}} \rightarrow \mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{R}^d)$  are said to be Lipschitz diffusion coefficients if they satisfy the non-anticipative property, meaning that for all  $t \in [0, t_0]$ ,  $x \in \mathcal{X}$ , any finite subset  $\mathcal{X}' \subset \mathcal{X}$ , and for  $m_1[x], m_2[x] \in \mathbb{M}$ , with  $m_1[\mathcal{X}'] := \sum_{x' \in \mathcal{X}'} \delta_{m_1[x']}$   $\in \hat{\mathcal{N}}_{\mathbb{M}}$ , and similarly for  $m_2[\mathcal{X}'] \in \hat{\mathcal{N}}_{\mathbb{M}}$ , we have:*

$$\begin{aligned} b(t, m_1[x], m_1[\mathcal{X}']) &= b(t, m_2[x], m_2[\mathcal{X}']), \\ \sigma(t, m_1[x], m_1[\mathcal{X}']) &= \sigma(t, m_2[x], m_2[\mathcal{X}']) \end{aligned}$$

whenever

$$\begin{aligned} m_1[x, t] &= m_2[x, t], \\ m_1[\mathcal{X}', t] &= m_2[\mathcal{X}', t]. \end{aligned}$$

Furthermore, there exists a Lipschitz diffusion constant  $K_t < \infty$  such that:

$$\begin{aligned} &\sup_{s \in [0, t]} |b(s, m_1[x], m_1[\mathcal{X}']) - b(s, m_2[x], m_2[\mathcal{X}'])| + |\sigma(s, m_1[x], m_1[\mathcal{X}']) - \sigma(s, m_2[x], m_2[\mathcal{X}'])| \\ &\leq K_t \left( \|m_1[x, t] - m_2[x, t]\| + \frac{1}{|\mathcal{X}'|} \sum_{x' \in \mathcal{X}'} \|m_1[x', t] - m_2[x', t]\| \right), \end{aligned}$$

with the convention that  $\frac{1}{|\mathcal{X}'|} \sum_{x' \in \mathcal{X}'} (\dots) = 0$  if  $\mathcal{X}' = \emptyset$ . We assume, moreover, that  $\sup_{t \in [0, t_0]} K_t < \infty$  and under this condition, we can assume that  $K_t$  are increasing in  $t$ . Finally, using the zero-function notation  $0[x] \in \mathbb{M}$  and  $0[\mathcal{X}'] := \sum_{x' \in \mathcal{X}'} \delta_{0[x']}$ , the coefficients  $(b, \sigma)$  satisfy:

$$\int_0^{t_0} (|b(s, 0, 0[\mathcal{X}'])|^2 + |\sigma(s, 0, 0[\mathcal{X}'])|^2) ds < \infty.$$

The Lipschitz assumptions imposed on the diffusion coefficients lead to a crucial property, which can be thought of as a *pre-Grönwall's inequality for interacting diffusions*. This property is formulated in Lemma 9.5 in Section 9.3. Firstly, this property is a cornerstone of the standard Picard iteration argument (applying Grönwall's inequality) to establish the existence and uniqueness of the diffusion described in (9.1) on any arbitrary, locally finite graph  $\mathcal{G}(\mathcal{X})$ , provided that the initial conditions are uniformly bounded. This is demonstrated in [95, Theorem 3.1, see also Appendix C], which we shall recall in Lemma 9.6. Furthermore, this property will be instrumental in demonstrating *L<sup>2</sup>-stabilization of the diffusions on graphs*, see Lemma 9.7 in Section 9.3.

### 9.1.2 Stabilizing random graphs

In this section, we shift our focus to the graph  $\mathcal{G}(\mathcal{P})$  whose vertices are the points  $x$  of a simple point process  $\mathcal{P}$  in  $\mathbb{R}^d$ , each hosting diffusion paths in dimension  $d'$ , which may differ from the ambient space dimension  $d$ .

Specifically, as in Section 8.1.1, we represent  $\mathcal{G}(\mathcal{P})$  as an auxiliary measurable marking  $\{\delta_{(x, \bar{N}_x)}\}_{x \in \mathcal{P}}$ , where the marks  $\bar{N}_x$  represent neighbors of  $x$  in  $\mathcal{G}(\mathcal{P})$  whose locations are expressed with respect to  $x$ ; i.e., they satisfy  $\bar{N}_x + x \subset \mathcal{P}$  for all  $x \in \mathcal{P}$ , and  $y \in \bar{N}_x$  if and only if  $-y \in \bar{N}_{x+y}$ . Assuming  $\bar{N}_x \in \hat{\mathcal{N}}_{\mathbb{R}^d}$  ensures local finiteness of the graph  $\mathcal{G}(\mathcal{P})$ .

We make also the simplifying assumption that the marks  $\bar{N}_x$  (and hence the graph  $\mathcal{G}(\mathcal{P})$ ) are (measurable) *deterministic functions of solely the input process*  $\mathcal{P}$ ,  $\bar{N}_x = \bar{N}_x(\mathcal{P})$  considering  $\bar{N}(\cdot) : \mathbb{R}^d \times \mathcal{N}_{\mathbb{R}^d} \rightarrow \hat{\mathcal{N}}_{\mathbb{R}^d}$  as a marking function as per Definition 4.1, admitting this time all input configurations of points  $\mu \in \mathcal{N}_{\mathbb{R}^d}$ , including those with infinite input.

Again as in Section 8.1.1, but for the infinite window  $W_\infty = \mathbb{R}^d$ , this marking function  $\bar{N}$  admits an interaction range of stabilization  $S : \mathbb{R}^d \times \mathcal{N}_{\mathbb{R}^d} \rightarrow \mathbb{N} \cup \{\infty\}$  of the graph  $\mathcal{G}(\mu)$  as per Definition 4.5, with  $S(x, \mu) := R^{\bar{N}}(x; \mu)$ . Clearly,  $S$  satisfies (8.4), (8.5), and (8.6).

In this graph construction scenario, a key stabilization assumption for the graph  $\mathcal{G}(\mathcal{P})$  is formulated analogously to Definition 8.1, but applied directly in the infinite setting via (4.11). Specifically, for every  $p \in \mathbb{N}$ , we assume that there exists a fast decreasing function  $\varphi'_p$ , as defined in (2.1), such that

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{P}_{[\cdot]_1^p} (S(x_1, \mathcal{P}) > s) \leq \varphi'_p(s), \quad s > 0. \quad (9.2)$$

We remark that the above assumption of stabilization of the graph  $\mathcal{G}(\mathcal{P})$  allows it to be approximated by local graphs. This makes a subtle difference from the setting of spin systems in Section 8.1.1, where the local graphs  $\mathcal{G}(\mathcal{P} \cap W_n)$ ,  $n \in \mathbb{N}$ , are, in fact, the only available representations of the interaction, and their stabilization only on finite windows ( $W_n$ ,  $n \in \mathbb{N}$ ) ensures consistency of these representations across increasing windows.

In Section 12, we will discuss how this framework can be extended to include more general randomized graph models.

## 9.2 Main results for the diffusion model

Our goal is to study certain scores associated with the individual paths of the diffusion on the infinite graph  $\mathcal{G}(\mathcal{P})$ . Specifically, the sums of these scores over points  $x \in \mathcal{P} \cap W_n$  will be the subject of the Gaussian fluctuations presented in this section.

To fully define the diffusion on this graph, we introduce the following marked point process as an input (upon which we will construct diffusion trajectories  $M[x]$ , for  $x \in \mathcal{P}$ ): For ambient and diffusion dimensions  $d, d' \in \mathbb{N}$ , consider a marked point process  $\tilde{\mathcal{P}} := \{\delta_{(x, M(x), Z_x)}\}_{x \in \mathcal{P}}$ , where  $x \in \mathbb{R}^d$  are points (sites) of the diffusions,  $M(x) =: M(x, 0) \in \mathbb{R}^{d'}$  are initial condition of the diffusions at these sites,  $Z_x \in \mathbb{M} = \mathcal{C}([0, t_0], \mathbb{R}^{d'})$  refer to standard Brownian motions on the interval  $[0, t_0]$  in  $\mathbb{R}^{d'}$ , assumed to be independent given the object  $\{\delta_{(x, M(x))}\}_{x \in \mathcal{P}}$ . In other words, this latter entire object is measurable at time  $t = 0$  with respect to a filtration that accommodates the Brownian motions.

Recall that the graph  $\mathcal{G}(\mathcal{P})$  is a function of the ground process  $\mathcal{P}$  (represented by the neighborhood marks  $\tilde{N}_x$  introduced above in Section 9.1.2, which are not part of the constitutive marks of input process  $\tilde{\mathcal{P}}$  directly related to the diffusion).

Assuming the Lipschitz coefficients  $(b, \sigma)$  are as defined in Definition 9.2, and given that the initial conditions  $M(x)$  are almost surely deterministically, uniformly bounded, then based on [95, Theorem 3.1] (see also Lemma 9.6), almost surely for  $\tilde{\mathcal{P}}$ , a path-wise unique strong solution  $M = M^{\mathcal{G}(\mathcal{P})}$  to the spatial system (9.1) on  $\mathcal{G}(\mathcal{P})$  exists. By its very construction, the individual paths  $M[x] = M^{\mathcal{G}(\mathcal{P})}[x] \in \mathbb{M}$  become new, measurable marks for points  $x \in \mathcal{P}$  (deterministically constructed given the input process  $\tilde{\mathcal{P}}$  and the graph  $\mathcal{G}(\mathcal{P})$ ).

Finally, we consider real-valued scores of the trajectories:

$$\xi(\tilde{x}, \tilde{\mathcal{P}}) := \xi^{(h, \mathcal{G}(\mathcal{P}))}(\tilde{x}, \tilde{\mathcal{P}}) := h(M^{\mathcal{G}(\mathcal{P})}[x]), \quad (9.3)$$

where  $h : \mathbb{M} \rightarrow \mathbb{R}$  is a Lipschitz function with respect to the sup-norm on  $\mathbb{M} = \mathcal{C}([0, t_0], \mathbb{R}^{d'})$ .

Given  $p \in [1, \infty)$ , we say that  $\xi = \xi^{(h, \mathcal{G}(\mathcal{P}))}$  satisfies the  $p$ -moment condition if

$$\sup_{1 \leq q \leq p} \sup_{x_1, \dots, x_q \in \mathbb{R}^d} \mathbb{E}_{[\tilde{x}]_1^q} \left[ \max \left( 1, |\xi(\tilde{x}, \tilde{\mathcal{P}})|^p \right) \right] \leq M_p^\xi < \infty \quad (9.4)$$

for some value  $M_p^\xi$ , which is assumed to be no smaller than  $M_{p'}$  for all  $p' \in [1, p]$ .

Uniform boundedness of the initial conditions  $M(x)$  and the Lipschitz form of the diffusion coefficient functions ensure that the diffusion trajectories  $M[x]$  are uniformly  $L^2$ -bounded; see Lemma 9.6. Consequently, given a Lipschitz function  $h$ , this implies  $M_p^\xi < \infty$  for  $p \in [1, 2]$  and hence guarantees that  $\xi$  satisfies the requisite  $(1 + \varepsilon)$ -moment condition needed to prove expectation asymptotics given in the next limit theorem.

Now, we present a main result of this application—the central limit theorem for interacting diffusions on spatial random graphs.

**Theorem 9.3** (Limit theory for statistics of interacting diffusions). *For  $d, d' \in \mathbb{N}$ ,  $t_0 \in (0, \infty)$ , let  $\tilde{\mathcal{P}} := \{\delta_{(x, M(x), Z_x)}\}_{x \in \mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{M}$  such that  $Z_x \in \mathbb{M} = \mathcal{C}([0, t_0], \mathbb{R}^{d'})$  are i.i.d. standard Brownian motions on the interval  $[0, t_0]$  in  $\mathbb{R}^{d'}$ , independent given the object*

$\{\delta_{(x, M(x))}\}_{x \in \mathcal{P}}$ . We assume that the initial conditions  $M(x)$ ,  $x \in \mathcal{P}$ , are almost surely deterministically, uniformly bounded, i.e.,  $\sup_{x \in \mathcal{P}} |M(x)| \leq L < \infty$ . Moreover, we assume that the marked point process  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8 and has bounded reduced Palm intensity function as in (8.10). Consider a graph  $\mathcal{G}(\mathcal{P})$  constructed deterministically on  $\mathcal{P}$  fast stabilizing with the interaction range satisfying (9.2). Given Lipschitz diffusion coefficients  $(b, \sigma)$  as in Definition 9.2 consider the strong solution to the interacting diffusions system (9.1) on  $\mathcal{G}(\mathcal{P})$ , realized by individual paths  $M^{\mathcal{G}(\mathcal{P})}[x] \in \mathbb{M}$  for points  $x \in \mathcal{P}$ . Finally, define the measures

$$\hat{\mu}_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}) \delta_{n^{-1/d}x}, \quad (9.5)$$

where  $\xi = \xi^{(h, \mathcal{G}(\mathcal{P}))}$  are real-valued scores of the individual paths  $M^{\mathcal{G}(\mathcal{P})}[x] \in \mathbb{M}$  given by (9.3) with a Lipschitz function  $h : \mathbb{M} \rightarrow \mathbb{R}$ .

(i) If the score function  $\xi$  in (9.3) satisfies the  $p$ -moment condition (9.4) for all  $p \in (1, \infty)$ , then the random measures  $(\hat{\mu}_n^\xi)_{n \in \mathbb{N}}$  at (9.5) satisfy the central limit theorem, i.e., for  $f \in \mathbb{B}(W_1)$  such that  $\text{Var} \hat{\mu}_n^\xi(f) = \Omega(n^\nu)$  for some  $\nu > 0$ , we have

$$(\text{Var} \hat{\mu}_n^\xi(f))^{-1/2} (\hat{\mu}_n^\xi(f) - \mathbb{E} \hat{\mu}_n^\xi(f)) \xrightarrow{d} Z.$$

(ii) If the input process  $\{\delta_{(x, M(x))}\}_{x \in \mathcal{P}}$  is stationary and the graph construction  $\mathcal{G}(\mathcal{P})$  is translation invariant, then the mean asymptotics can be expressed for all  $f \in \mathbb{B}(W_1)$  as:

$$\left| n^{-1} \mathbb{E} \hat{\mu}_n^\xi(f) - \rho \mathbb{E}_0 \xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \int_{W_1} f(x) dx \right| = O(n^{-1/d}), \quad (9.6)$$

whereas under an extra  $p = (2 + \epsilon)$ -moment condition, for some  $\epsilon > 0$ , the variance asymptotics can be expressed as

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \hat{\mu}_n^\xi(f) = \sigma^2(\xi) \int_{W_1} f(x)^2 dx \in [0, \infty), \quad (9.7)$$

where

$$\begin{aligned} \sigma^2(\xi) &:= \rho \mathbb{E}_0 \xi^2(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \\ &+ \int_{\mathbb{R}^d} \left( \mathbb{E}_{\tilde{\mathbf{0}}, x}(\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \xi(\tilde{x}, \tilde{\mathcal{P}})) \rho^{(2)}(\mathbf{0}, x) - \rho^2 \mathbb{E}_0(\xi(\tilde{\mathbf{0}}, \tilde{\mathcal{P}})) \mathbb{E}_x(\xi(\tilde{x}, \tilde{\mathcal{P}})) \right) dx \in [0, \infty). \end{aligned} \quad (9.8)$$

*Proof.* We deduce all statements from Proposition 5.5. Given the mixing and moment conditions on  $\tilde{\mathcal{P}}$  and  $\xi$ , it remains to localize the scores  $\xi(\tilde{x}, \tilde{\mathcal{P}}) = \xi^{(h, \mathcal{G}(\mathcal{P}))}(\tilde{x}, \tilde{\mathcal{P}})$  calculated for the diffusion trajectories in the infinite model. That is, we need to define a marking function  $\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B)$ , measurable with respect to the input  $\tilde{\mathcal{P}}$  truncated to a bounded set  $B \subset \mathbb{R}^d$ , such that these localized scores approximate (in the BL-sense) the original scores  $\xi(\tilde{x}, \tilde{\mathcal{P}})$  rapidly as  $B \uparrow \mathbb{R}^d$ . This corresponds exactly to the fast BL-localization of  $\xi$  as required in Definition 4.2(i) and (iii), and invoked in Proposition 5.5.

To achieve this, it is natural to define  $\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B)$  as the value of the score function  $h$  applied to the trajectory at  $x$  resulting from the interacting diffusions run only on the truncated input  $\tilde{\mathcal{P}} \cap B$ —that is, using the points  $y \in \mathcal{P} \cap B$  along with their original initial conditions  $M(y)$  and Brownian motions  $Z_y$ . What remains is to determine how the graph structure should be incorporated in this local diffusion construction.

It would be straightforward to consider these diffusions on the truncated graph  $\mathcal{G}(\mathcal{P}) \cap B$  and to apply Lemma 9.7 on  $L^2$ -stabilization of the interacting diffusions (which also implies BL-stabilization), since  $\mathcal{G}(\mathcal{P}) \cap B \subset \mathcal{G}(\mathcal{P})$ . However, we cannot use the truncation  $\mathcal{G}(\mathcal{P}) \cap B$  as the system of edges to propagate these local diffusions at  $x \in \mathcal{P} \cap B$ , because in our setting,  $\mathcal{G}(\mathcal{P}) \cap B$  is not measurable with respect to  $\tilde{\mathcal{P}} \cap B$ . Indeed, the auxiliary marks  $\tilde{N}_x$  (which determine  $\mathcal{G}(\mathcal{P}) \cap B$ ) are not included in  $\tilde{\mathcal{P}}$ .

Instead, we will rely on a local construction  $\mathcal{G}(\mathcal{P} \cap B)$  of the graph, consider the diffusion defined on it (rather than on  $\mathcal{G}(\mathcal{P}) \cap B$ ), and apply the function  $h$  to the trajectory at  $x$  resulting from the interacting diffusions governed by this alternative graph structure.

Specifically, we consider the following score function

$$\xi(\tilde{x}, \tilde{\mathcal{P}} \cap B) := \xi^{(h, \mathcal{G}(\mathcal{P} \cap B))}(\tilde{x}, \tilde{\mathcal{P}} \cap B) := h(M^{\mathcal{G}(\mathcal{P} \cap B)}[x]), \quad x \in \mathcal{P} \cap B. \quad (9.9)$$

Note that the local graph construction  $\mathcal{G}(\mathcal{P} \cap B)$  only approximates  $\mathcal{G}(\mathcal{P}) \cap B$ , and does so thanks to the assumed stabilization property of the graph construction. For a large enough set  $B$ , the multi-hop neighborhood of  $x$  in  $\mathcal{G}(\mathcal{P} \cap B)$  will coincide with that in  $\mathcal{G}(\mathcal{P}) \cap B$ . This will allow us to invoke Lemma 9.7 to establish  $L^2$ -stabilization of the interacting diffusions. This strategy is developed in detail in what follows.

More precisely, we aim to prove the following BL-localizing property

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{E}_{[x_1^p]} \left[ \left| f([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})) - f([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}} \cap B_r([x]_1^p))) \right| \right] \leq 2\varphi_p(r), \quad r \geq 1. \quad (9.10)$$

satisfied for all test functions  $f \in \text{BL}(\mathbb{R}^p)$ , for all  $p \in \mathbb{N}$ , and some fast decreasing functions  $\varphi_p$ . Here  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}})$  stands for the  $p$ -vector

$$(\xi^{(h, \mathcal{G}(\mathcal{P}))}(\tilde{x}_1, \tilde{\mathcal{P}}), \dots, \xi^{(h, \mathcal{G}(\mathcal{P}))}(\tilde{x}_p, \tilde{\mathcal{P}})),$$

and,  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}} \cap B_r([x]_1^p))$  similarly for a  $p$ -vector with entries  $\xi^{(h, \mathcal{G}(\mathcal{P} \cap B_r([x]_1^p)))}(\tilde{x}_i, \tilde{\mathcal{P}} \cap B_r([x]_1^p))$ , where  $B_r([x]_1^p) = \cup_{i=1}^p B_r(x_i)$  is the union of balls of radius  $r$  centred at points  $x_i$ ,  $i = 1, \dots, p$ .

By Remark 4.7, it will be enough to show the stabilization in  $L^2$ ; i.e., that

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{E}_{[x_1^p]} \left[ \left| \xi^{(h, \mathcal{G}(\mathcal{P}))}(\tilde{x}_1, \tilde{\mathcal{P}}) - \xi^{(h, \mathcal{G}(\mathcal{P} \cap B_r([x]_1^p)))}(\tilde{x}_1, \tilde{\mathcal{P}} \cap B_r([x]_1^p)) \right|^2 \right] \leq 2\varphi_p(r), \quad r \geq 1, \quad (9.11)$$

for all  $p \in \mathbb{N}$ , and some fast decreasing functions  $\varphi_p$ . By the Lipschitz(1) property of  $h$ , it is enough to show

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{E}_{[x_1^p]} \left\| M^{\mathcal{G}(\mathcal{P})}[x_1] - M^{\mathcal{G}(\mathcal{P} \cap B_r([x]_1^p))}[x_1] \right\|^2 \leq 2\varphi_p(r), \quad r \geq 1.$$

We may bound the required difference in the left-hand side by using the spatial  $L^2$ -stabilization of the diffusion on graphs (see Lemma 9.7) and the stabilization of the graph.

Indeed, set  $k_r := r^b$  for a fixed  $b \in (0, 1)$  and let  $\mathcal{G}_{[x_1^p]}^{\cup(r)} := \mathcal{G}(\mathcal{P} \cap B_r([x_1^p]))$  and  $\mathcal{G}_{[x_1^p]}^{[k_r]}$  be a subgraph of  $\mathcal{G}(\mathcal{P})$  induced by its graph ball  $\mathfrak{B}_{k_r}(x_1)$  of radius  $k_r$  centered at  $x_1$  (both graphs are considered under Palm conditioning  $x_1, \dots, x_p \in \mathcal{P}$ ). We have for all  $x_1, \dots, x_p \in \mathbb{R}^d$

$$\begin{aligned} & \mathbb{E}_{[x_1^p]} \|M^{\mathcal{G}(\mathcal{P})}[x_1] - M^{\mathcal{G}(\mathcal{P} \cap B_r([x_1^p]))}[x_1]\|^2 \\ & \leq \mathbb{E}_{[x_1^p]} [\mathbb{E}[\|M^{\mathcal{G}(\mathcal{P})}[x_1] - M^{\mathcal{G}(\mathcal{P} \cap B_r([x_1^p]))}[x_1]\|^2 \mid \mathcal{P}] \mathbf{1}(\mathcal{G}_{[x_1^p]}^{\cup(r)} \supset \mathcal{G}_{[x_1^p]}^{[k_r]}) \\ & \quad + 2C' \mathbb{P}_{[x_1^p]} \{\mathcal{G}_{[x_1^p]}^{\cup(r)} \not\supset \mathcal{G}_{[x_1^p]}^{[k_r]}\} \\ & \leq \frac{C^{k_r}}{k_r!} + 2C' \mathbb{P}_{[x_1^p]} \{\mathcal{G}_{[x_1^p]}^{\cup(r)} \not\supset \mathcal{G}_{[x_1^p]}^{[k_r]}\}, \end{aligned} \tag{9.12}$$

where the constant  $C'$  uniformly bounds in  $L^2$  all diffusion trajectories,  $\sup_{y \in \mathcal{X}} \mathbb{E}[\|M[y]\|^2] < C'$ , given the bound of the initial conditions and Lipschitz coefficients but regardless of the graph (see Lemma 9.6) and the first term in (9.12) is due to  $L^2$  stabilization of the diffusion with respect to the graph distance, specifically from (9.15) in the Lemma 9.7.

In order to show that the probability bound in (9.12) is a fast decreasing function of  $r$ , and thus conclude the proof, we follow the arguments used for Lemma 8.7 (regarding interaction graphs for spin systems). Specifically, we bound the probability in this expression by exploiting the fact that if a path of length  $k_r$  from  $x_1$  in  $\mathcal{G}(\mathcal{P})$  is not contained within  $\mathcal{G}_{[x_1^p]}^{\cup(r)}$ , then at least one point  $y$  in  $\mathcal{P} \cap B_r([x_1^p])$  must have a ‘very large’ interaction range  $S(y, \mathcal{P})$ . This is hardly likely since the graph is stabilizing. The details of this argument are nearly identical to those in the proof of Lemma 8.7; it suffices to repeat the arguments leading to the derivation of (8.25), replacing  $\mathcal{P}_n$  with  $\mathcal{P}$  and using the stabilization property (9.2).  $\square$

We conclude this section with a few remarks.

**Remark 9.4.** (i) *Summable exponential  $\mathbb{B}$ -mixing* of points and initial marks is sufficient. In Theorem 9.3, we assumed that the Brownian motions  $Z_x$  are i.i.d. given  $\{\delta_{(x, M(x))}\}_{x \in \mathcal{P}}$ . Consequently, in the summable exponential  $\mathbb{B}$ -mixing assumption for  $\tilde{\mathcal{P}}$ , one can omit these Brownian marks and verify the condition only for  $\{\delta_{(x, M(x))}\}_{x \in \mathcal{P}}$ .

(ii) *CLT for diffusions on finite windows.* One can establish the central limit theorem for diffusions considered only on finite windows, similarly to the spin systems discussed in Section 8, whether or not a limiting infinite graph exists.

(iii) *Convergence of interacting diffusions.* Our work on interacting diffusions was inspired by [95], which focused on diffusion limits over sequences of converging graphs. The typical setting there involves finite, uniformly rooted graphs converging in the local weak sense to infinite unimodular graphs, viewed as equivalence classes under graph isomorphism (e.g., sparse Erdős–Rényi graphs converging to a Bienaymé–Galton–Watson tree). A key assumption in their analysis is the uniform boundedness of the initial conditions and the Lipschitz continuity of the diffusion coefficients.

Our work—of independent interest—extends this framework to interacting diffusions on finite spatial random graphs, which converge to their limiting infinite graphs in a stabilization sense (see Definition 8.1). While our main convergence notion relies on graph stabilization, weaker forms of convergence (e.g., in  $L^2$ ) should also be considered as natural in this context. It may also be interesting to introduce a dependence of the diffusion coefficients on the Euclidean positions of the vertices of the graphs.

### 9.3 Auxiliary statements for diffusions on graphs

In this section, we collect results concerning interacting diffusions, including its existence, uniqueness, and the boundedness of trajectories. These results were established in [95], but specifically in the context of graphs considered as elements of the space  $\mathcal{G}^*$  of equivalence classes under the isomorphism relation on locally finite, rooted graphs. This setting is too restrictive for many random graphs built on point processes. However, the key arguments developed in that work remain valid in our more general setting. For completeness and self-containment, we briefly recall these arguments here.

The main new result of this section is Lemma 9.7, which establishes spatial  $L^2$ -stabilization of the diffusion on graphs.

**Lemma 9.5** ((Pre-)Grönwall’s inequality for interacting diffusions). *Consider Lipschitz diffusion coefficients  $(b, \sigma)$  as defined in Definition 9.2. Let  $\{X[x]\}_{x \in \mathcal{X}}$  and  $\{X'[x]\}_{x \in \mathcal{X}}$  be two families of continuous processes in  $\mathbb{R}^d$  adapted to the filtration generated by the family of i.i.d. standard Brownian motions  $\{Z_x(\cdot)\}_{x \in \mathcal{X}}$  in  $\mathbb{R}^d$  on  $[0, t_0]$  and  $\mathbb{E}[\|X[x]\|^2] < \infty$ ,  $\mathbb{E}[\|X'[x]\|^2] < \infty$  for all  $x \in \mathcal{X}$ . Consider the family of stochastic processes  $\{Y[x]\}_{x \in \mathcal{X}}$  on  $\mathbb{M}$ , defined on  $[0, t_0]$  by*

$$Y(x, t) := Y(x, 0) + \int_0^t b(t, X[x], X[N_x]) dt + \sigma(t, X[x], X[N_x]) dZ_x(t), \quad (9.13)$$

and similarly the family of stochastic processes  $Y'(x, t)$  on  $\mathbb{M}$  defined as above but with  $X'$  instead of  $X$ . These processes are indexed by the vertices  $x \in \mathcal{X}$ , utilize the graph neighborhood  $N_x = \{y \in \mathcal{X} : y \sim x\}$  on a given graph  $\mathcal{G}(\mathcal{X})$  and share the same initial conditions  $Y(x, 0) = Y'(x, 0)$ . Then, for any  $x \in \mathcal{X}$  and  $t \in [0, t_0]$ , we have:

$$\mathbb{E}[\|Y[x, t] - Y'[x, t]\|^2] \leq 32 \max(t, 1) K_t^2 \int_0^t \max_{y \in \mathfrak{B}_1(x)} \mathbb{E}[\|X[y, s] - X'[y, s]\|^2] ds, \quad (9.14)$$

where  $\mathfrak{B}_1(x)$  is graph ball on  $\mathcal{G}(\mathcal{X})$  of radius 1 centred at  $x$ .

*Proof.* For  $x \in \mathcal{X}$ , using the triangle inequality, we have for all  $t \in [0, t_0]$

$$\begin{aligned} & \|Y[x, t] - Y'[x, t]\|^2 \\ & \leq 2 \left[ \left( \int_0^t |b(s, X[x], X[N_x]) - b(s, X'[x], X'[N_x])| ds \right)^2 \right. \\ & \quad \left. + \sup_{s \leq t} \left( \int_0^s [\sigma(u, X[x], X[N_x]) - \sigma(u, X'[x], X'[N_x])] dZ(x, u) \right)^2 \right]. \end{aligned}$$

By the progressive measurability assumption on  $\sigma$  in Definition 9.2, the finiteness of the second moments of  $\|X[x]\|$  and  $\|X'[x]\|$ , and the Lipschitz assumption on  $\sigma$  (making  $\mathbb{E}[\int_0^{t_0} |\sigma(u, X[x], X[N_v])|^2 du < \infty$  and similarly for  $\sigma(u, X'[x], X'[N_v])$ ) the second integral in the right-hand side is a martingale; see [85, Section 3.2(A,B)]. Thus the square of the integral is a submartingale. Using the Cauchy-Schwarz inequality for the first integral and Doob's maximal inequality for submartingales (see [85, Section 1.3, Theorem 3.8(iv)] for example) followed by the Itô isometry for the second integral, we obtain

$$\begin{aligned} \Delta_Y[x, t] &:= \mathbb{E}[\|Y[x, t] - Y'[x, t]\|^2] = \mathbb{E}[\sup_{s \leq t} |Y(x, s) - Y'(x, s)|^2] \\ &\leq 2 \left[ t \int_0^t \mathbb{E} \left[ \left| b(s, X[x], X[N_x]) - b(s, X'[x], X'[N_x]) \right|^2 \right] ds \right. \\ &\quad \left. + 4 \int_0^t \mathbb{E} \left[ \left| \sigma(s, X[x], X[N_x]) - \sigma(s, X'[x], X'[N_x]) \right|^2 \right] ds \right]. \end{aligned}$$

Using the Lipschitz property of  $(b, \sigma)$  in Definition 9.2 and then applying Cauchy-Schwarz to the second term, we obtain

$$\Delta_Y[x, t] \leq 16 \max(t, 1) K_t^2 \int_0^t \left( \mathbb{E}[\|X[x, s] - X'[x, s]\|^2] + \frac{1}{|N_x|} \sum_{y \in N_x} \mathbb{E}[\|X[y, s] - X'[w, s]\|^2] \right) ds.$$

This yields

$$\begin{aligned} \Delta_Y[x, t] &\leq 16 \max(t, 1) K_t^2 \int_0^t \left( \Delta_X[x, s] + \frac{1}{|N_x|} \sum_{y \in N_x} \Delta_X[y, s] \right) ds \\ &\leq 32 \max(t, 1) K_t^2 \int_0^t \max_{y \in \mathfrak{B}_1(x)} \Delta_X[y, s] ds, \end{aligned}$$

where  $\Delta_X[x, t] := \mathbb{E}[\|X[x, t] - X'[x, t]\|^2]$ . This completes the proof.  $\square$

**Lemma 9.6** (Unique, path-wise solution of the interacting diffusions and uniform  $L^2$ -norm bound; [95, Theorem 3.1, Lemma 4.1]). *Consider finite or locally finite, infinite graph  $\mathcal{G}(\mathcal{X})$  on  $\mathcal{X}$  equipped with initial conditions  $\{M(x, 0)\}_{x \in \mathcal{X}}$  in  $\mathbb{R}^{d'}$ ,  $d' \in \mathbb{N}$ , uniformly bounded in  $\mathcal{X}$ :  $\sup_{y \in \mathcal{X}} |M(y, 0)| \leq L$ , and i.i.d. standard Brownian motions  $\{Z_x(\cdot)\}_{x \in \mathcal{X}}$  in  $\mathbb{R}^{d'}$  on  $[0, t_0]$ , for  $t_0 \in (0, \infty)$ . Consider the drift and diffusion coefficients  $b, \sigma$  being Lipschitz functions as per Definition 9.2. Then the system (9.1) admits a unique solution with the family of processes  $\{M(x, t)\}_{x \in \mathcal{X}}$ , taking values in the space of continuous paths with  $t \in [0, t_0]$ . Moreover, the trajectories  $M[x]$  are uniformly bounded in  $L^2$ :*

$$\sup_{y \in \mathcal{X}} \mathbb{E}[\|M[y]\|^2] < C',$$

where the constant  $C'$  depends only on  $t_0$ , the Lipschitz constant  $K_{t_0}$ , as well on the bound  $L$  of the initial conditions, and not on the graph structure.

*Proof.* Using the same arguments as in the proof of Lemma 9.5 we show first that the iterations  $m \in \mathbb{N}$ :

$$X_m(x, t) = M(x, 0) + \int_0^t b(t, X_{m-1}[x], X_{m-1}[N_x]) dt + \sigma(t, X_{m-1}[x], X_{m-1}[N_x]) dZ_x(t),$$

with,  $X_0(x, t) \equiv 0$  on  $[0, t_0]$ , are uniformly bounded in  $L^2$ :  $\sup_{m \in \mathbb{N}, x \in \mathcal{X}} \mathbb{E}[\|X_m[x]\|^2] < \infty$ . Then, using Lemma 9.5, by the Cauchy property, we show that  $X_m$  converge in  $L^2$ -norm of the supremum to  $M$  i.e.,  $\sup_{x \in \mathcal{X}} \mathbb{E}[\|X_m[x] - M[x]\|^2] \rightarrow 0$  as  $m \rightarrow \infty$ . Uniqueness uses the same Lemma.  $\square$

**Lemma 9.7** ( $L^2$ -stabilization of the diffusion on graphs). *Consider the interacting diffusions  $\{M(y, t)\}$  in the interval  $t \in [0, t_0]$  given in (9.1) on a graph  $\mathcal{G}(\mathcal{X})$ , with Lipschitz diffusion coefficients  $(b, \sigma)$  as in Definition 9.2, and bounded initial conditions  $\sup_{y \in \mathcal{X}} |M(y, 0)| \leq L$  for a (deterministic) constant  $L < \infty$ . For  $x \in \mathcal{X}$ , consider a graph  $\mathcal{G}_x^{[m]}(\mathcal{X}) := \mathcal{G}(\mathfrak{B}_m(x))$  induced by  $\mathcal{G}(\mathcal{X})$  on the ball  $\mathfrak{B}_m(x)$  of radius  $m$  centred at  $x$  in  $\mathcal{G}(\mathcal{X})$ , and consider the interacting diffusions  $\{M^{\mathcal{G}_x^{[m]}(\mathcal{X})}(y, t)\}$  for  $t \in [0, t_0]$  on  $y \in \mathfrak{B}_m(x)$  with the same coefficients  $(b, \sigma)$ , generated by the same Brownian motions and sharing the same initial conditions  $M^{\mathcal{G}_x^{[m]}(\mathcal{X})}(y, 0) = M(y, 0)$  for  $y \in \mathfrak{B}_m(x)$ . Then there exists a constant  $C = C(L, t_0, K_{t_0}) < \infty$  (not depending on the graph  $\mathcal{G}(\mathcal{X})$ ) such that for all  $t \in [0, t_0]$*

$$\mathbb{E}[\|M[x, t] - M^{\mathcal{G}_x^{[m]}(\mathcal{X})}[x, t]\|^2] \leq \frac{C^m}{m!}. \quad (9.15)$$

*Proof.* We use Lemma 9.5 considering stochastic processes  $X(y, t) := M(y, t)$  and  $X'(y, t) := M^{\mathcal{G}_x^{[m]}(\mathcal{X})}(y, t)$ ,  $t \in [0, t_0]$ , considering  $y$  on the graph  $\mathcal{G}_x^{[m]}(\mathcal{X})$ . Set

$$\Delta[x, t] := \mathbb{E}[\|M[x, t] - M^{\mathcal{G}_x^{[m]}(\mathcal{X})}[x, t]\|^2].$$

Since  $M$ ,  $M^{\mathcal{G}_x^{[m]}(\mathcal{X})}$  satisfy the integral representation (9.13) with  $Y(y, t) = M(y, t)$  and  $Y'(y, t) = M^{\mathcal{G}_x^{[m]}(\mathcal{X})}(y, t)$  respectively, using (9.14), we derive

$$\Delta[x, t] \leq 32 \max(t, 1) K_t^2 \int_0^t \max_{y \in \mathfrak{B}_1(x)} \Delta[y, s] ds.$$

Iterating the inequality  $m$  times we obtain

$$\begin{aligned} \Delta[x, t] &\leq (32 \max(t, 1) K_t^2)^m \int_0^t \int_0^{s_1} \dots \int_0^{s_{m-1}} \max_{y \in \mathfrak{B}_m(x)} \Delta[y, s_m] ds_m \dots dt_2 ds_1 \\ &\leq \frac{(32 \max(t, 1) K_t^2)^m}{m!} \max_{y \in \mathfrak{B}_m(x)} \Delta[y, s_m]. \end{aligned}$$

By Lemma 9.6 we have  $\sup_{y \in \mathcal{X}} \mathbb{E}[\|M[y]\|^2] < C'$ , where the constant  $C'$  depends only on  $t_0$ , the Lipschitz constant  $K_{t_0}$ , as well on the bound  $L$  of the initial conditions, and *not* on the graph structure. This implies  $\max_{y \in \mathfrak{B}_m(x)} \Delta[y, s_m] \leq 2C'$  completing the proof of Lemma 9.7.  $\square$

## 10 Interacting particle systems on spatial random graphs

This section uses the main theoretical results of Section 5 to establish the limit theory for summary statistics of discrete and continuous time interacting particle systems in the continuum. Interacting particle systems are usually studied on fixed geometries, including lattices and trees. Here we study statistics of such systems on random geometries given by a point process, or, more precisely on a graph

on a random point process. There are thus two sources of randomness, namely the random set of sites and the interacting particle system evolution. Statistics of interacting particle systems are expressed as sums of real-valued marking functions with stopping sets satisfying stabilization criteria, putting us in the set-up of Section 5. This contrasts with the statistics encountered in Sections 8 and 9, where the underlying point processes were equipped with marks which did not have stopping sets.

We work in a particle system framework, which is more general than that considered in most classical models. As mentioned in Section 1, the framework allows for these conditions: (i) initial states may be dependent, (ii) interaction neighborhoods may be unbounded, (iii) states of particles may take values in Polish spaces and they may be a function of the entire time-evolved history of states in the interaction neighborhood and (iv) particle locations may be correlated.

Roughly speaking, the dynamics of these general interacting particle system go as follows. The particles of a point process having dependent initial states are equipped with independent Poisson clocks, whose ‘rings’ at a given site trigger an update at that site (through a specified update function) as well as at neighboring sites. The update may depend on the prior history at the site as well as as the histories at neighboring sites, formalized below at (10.2). Interactions depend on neighbors which are not ‘too far away’, meaning that the neighbors should be within some random distance having a decaying tail; this is formalized via the notion of stabilizing interaction graphs given in the previous two sections. The particle locations are required to belong to the realization of a point process having exponentially decaying correlations and satisfying an additional growth condition on their Palm correlation functions, a generalization of condition (8.10) also needed when considering spin systems or interacting diffusions.

This approach enlarges the scope of previous studies, e.g. in [117], which have assumed bounded interaction ranges, independent initial states, and independent particle locations, namely those given by a Poisson point process. Our set-up does not require discretization or lattice-based particle systems and allows for models defined by a graphical structure based on geometry, arguably more realistic than the Erdős-Rényi graphical structure. However, we emphasize that we shall work in the sparse regime, where the particles interact with only neighboring particles and not all the particles as in mean-field models.

In Section 10.1, we introduce a general continuous-time interacting particle system model and state and prove limit theorems for statistics of such models in Section 10.2. Markovian particle systems and various examples are discussed in Section 10.3. Discrete-time interacting particle systems, which are of independent interest, are taken up in Section 10.4. The discrete-time model features globally synchronous updates whereas the continuous-time model will have only locally-synchronous updates. Our list of explicit models is illustrative and not exhaustive.

## 10.1 General model assumptions

We introduce the key ingredients in our framework for continuous-time spatial interacting particle systems.

### 10.1.1 Sites and states of particles

We consider a simple marked point process  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$ , where the ground process  $\mathcal{P}$  defines the *sites (locations)* of the particles on  $\mathbb{R}^d$  and their marks  $U(x) = (M(x), \tilde{\tau}_x) \in \mathbb{K}$  take values in a Polish space having the following form

$$\mathbb{K} := \mathbb{M} \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{L}}, \quad (10.1)$$

where

- $M(x) \in \mathbb{M}$  is an *initial state* at site  $x \in \mathcal{P}$  at time 0;  $\mathbb{M}$  is a Polish space and represents for example, the occupation of the site by a particle of some color, infection status, etc;  $\{M(x)\}_{x \in \mathcal{P}}$  are possibly dependent given the ground process  $\mathcal{P}$ .
- $\tilde{\tau}_x \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{L}}$  is an independently *marked Poisson clock* at site  $x \in \mathcal{P}$ ;  $\tilde{\tau}_x$  are i.i.d. given  $\{(x, M(x))\}_{x \in \mathcal{P}}$ . More precisely,  $\tilde{\tau}_x := \{(T_i = T_i(x), L_i = L_i(x))\}$ , where  $\tau_x := \{T_i(x)\}$  is a unit rate Poisson point process on  $[0, t_0]$  for some finite time horizon  $t_0 < \infty$ —it denotes the clock whose rings trigger updates of the states of site  $x$  and its neighbors in  $\mathbb{M}$  during the  $[0, t_0]$  window (cf. admissible update rules below)—and  $L_i(x)$  are i.i.d. (given  $\tau_x$ ) random elements with values in a Polish space  $\mathbb{L}$ . These *time-marks*  $L_i(x)$  allow for additional independent randomness (given  $\tau_x$ ) in the admissible update rule (for example, to define Markovian updates). In many examples  $\mathbb{L} = [0, 1]$  but for some particle systems such as ballistic deposition, it is useful to allow  $\mathbb{L}$  to be a space of shapes; see Section 10.3. To lighten the exposition we often suppress the dependence on  $x$  when writing  $T_i$  and  $L_i$ .

In Section 5 we considered simple marked point processes on  $\mathbb{R}^d \times \mathbb{K}$ ,  $\mathbb{K}$  a Polish space. Taking  $\mathbb{K}$  of the form (10.1), with both Polish  $\mathbb{M}$  and  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{L}}$  (the latter being the space of finite subsets of  $[0, t_0] \times \mathbb{L}$ , equipped with the weak topology and corresponding Borel  $\sigma$ -algebra), provides an input framework for studying the limit theory for continuous-time interacting particle systems.

We emphasize that we neither assume that  $\mathcal{P}$  is a Poisson point process nor do we assume independence of the collection of initial states  $\{M(x)\}_{x \in \mathcal{P}}$ .

### 10.1.2 Update rules and the evolution of states

The ringing of clocks  $\tau_x$  are the times that trigger updates of the state in  $\mathbb{M}$  (for example signalling arrivals or departures of particles at the considered locations or updating the particles' color, infection status, etc.) at site  $x$ , and possibly at its neighbors  $y \sim x$ . Here the relation  $y \sim x$  is in terms of graph neighbors, where the graph belongs to the set of stabilizing interaction graphs recalled in the next Section 10.1.3. Also, we shall write  $y \simeq x$  if  $y \sim x$  or  $y = x$ . The times at which a site  $x \in \mathcal{P}$  updates (or re-updates) its state in  $\mathbb{M}$  are denoted by  $t_j \in [0, t_0]$ . The corresponding states at these times are denoted  $M(x, t_j)$ , with  $M(x, 0) = M(x)$  representing the initial state. The *history (of the evolution)* of these states visited by site  $x$  is formally represented as a time-marked point process  $M(x, \cdot) := \{(t_j, M(x, t_j))\} \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$ , with analogous notation used for marked clocks  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{L}}$ . However, it is often more intuitive to view this history as a càdlàg function  $M(x, t) := M(x, \max\{t_j : t_j \leq t\})$  for

$t \in [0, t_0]$ , i.e., a function that is right-continuous with left limits, defined over continuous time. The process of the evolution on the states  $M(x, \cdot) = M_n(x, \cdot)$  will be constructed on *finite windows*  $W_n$ ,  $n \in \mathbb{N}$ ; i.e., for locations  $x \in \mathcal{P}_n = \mathcal{P} \cap W_n$ , with their neighborhood  $\sim$  possibly specific for  $W_n$ ; see Section 10.1.3 below. This will be constructed recursively, with an *updating function*  $\Phi$  that will be applied recurrently at all times  $t \in \tau_x$  of all  $x \in \mathcal{P}_n$ .

We define the inductive rules of the updates of states. The *update rules are admissible* if the following assumptions are met:

- (a) For  $t = 0$  we set  $M_n(x, \cdot) := M(x, \cdot) = \{(0, M(x))\}$  for all  $x \in \mathcal{P}_n$ ; that is, the history of the evolution of the states in  $\mathbb{M}$  of all sites  $x \in \mathcal{P}_n$  consists only of the input initial states  $M(x, 0) = M(x)$ .
- (b) For any marked clock event  $(T, L) \in \tilde{\tau}_x$  of some site  $x \in \mathcal{P}_n$ , the prior histories  $\{(t, M(y, t))\}_{t < T}$  at sites  $y \simeq x$ , i.e. at  $x$  and its neighbors in  $\mathcal{P}_n$ , are expanded to  $\{(t, M(y, t))\}_{t < T} \cup \{(T, M(y, T))\}$ , which include events  $(T, M(y, T)) \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$  given by an updating function  $\Phi$  as specified below in (10.2).
- (c) The updating function  $\Phi : \mathbb{R}^d \times \mathbb{L} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}} \rightarrow \hat{\mathcal{N}}_{\mathbb{R}^d \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}}$  is measurable. Here the argument in  $\mathbb{R}^d$  identifies the site  $x \in \mathcal{P}$ ,  $\mathbb{L}$  contains time-marks corresponding to the clock ringing at  $x$  and  $\hat{\mathcal{N}}_{\mathbb{R}^d \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}}$  denotes a marked point process on  $\mathbb{R}^d$ , representing the relative locations  $y - x$  of the neighbors  $y$  of  $x$  (including  $x$ ), each marked by its own process of state history, which lies in  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$ . The function  $\Phi$  is applied to these arguments and returns a new point process in the same space, representing the *new states*:

$$\Phi \left( x, L, \left\{ \left( y - x, \{(t, M(y, t))\}_{t < T} \right) \right\}_{\{y \in \mathcal{P}_n : y \simeq x\}} \right) =: \left\{ \left( y - x, (T, M(y, T)) \right) \right\}_{\{y \in \mathcal{P}_n : y \simeq x\}} \quad (10.2)$$

to be incorporated into the updated state histories of the neighboring sites. Sometimes, we also explicitly assume that  $\Phi$  is translation-invariant, meaning  $\Phi(x + z, \cdot, \{(y, \cdot)\}) \equiv \Phi(x, \cdot, \{(y, \cdot)\})$  for all  $x, z \in \mathbb{R}^d$ .

Observe, for a marked clock event  $(T, L) \in \tilde{\tau}_x$  at the site  $x \in \mathcal{P}_n$ , besides the location of the site  $x$  and the time mark  $L$ , the function  $\Phi$  takes as its third argument the previous histories  $M(y, t)$ ,  $t < T$ , at all neighboring locations  $y \simeq x$  and its value consists in producing the new events  $(T, M(y, T))$ , which represent entering (possibly re-entering) these states at time  $T$  to the neighboring sites. In other words, at time  $T$ , the histories are updated only for  $y \simeq x$  to  $\{(t, M(y, t))\}_{t < T} \cup \{(T, M(y, T))\}$ , where  $M(y, T)$  are defined in the right-hand side of (10.2).

The *complete history* of a given site  $x \in \mathcal{P}_n$ —resulting from the execution of the function  $\Phi$  at all clock rings  $\tilde{\tau}_y$  for  $y \in \mathcal{P}_n$ —is denoted by  $M_n[x] \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$ . These histories are treated as an additional marking of the point  $x \in \mathcal{P}_n$ . Note, this marking is measurable as a measurable mapping

$$g : (\mathbb{R}^d \times \mathbb{K}) \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \ni (\tilde{x}, \tilde{\mu}) \mapsto M[x] = (M(x, t))_{t \in [0, t_0]} \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}, \quad (10.3)$$

constructed recursively using the measurable function  $\Phi$  described above. This mapping describes the history of marks  $M_n[x]$ ,  $x \in \mathcal{P}_n$ , when applied simultaneously on all  $x \in \mathcal{P}_n$ . For more discussion see Remark 10.2 (iii) and (viii) below.

### 10.1.3 Stabilizing interaction graphs

The admissible update rules with the update function  $\Phi$  depend on the interaction relation specified via graphs on  $\mathcal{P}_n$ . As in Section 8.1.1, we restrict attention to stabilizing graphs—that is, graphs whose interaction ranges  $S_n$  on finite windows, as defined in (8.2) and satisfying properties (8.4)–(8.7), fulfill the fundamental assumption stated in Definition 8.1, recalled here for convenience:

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[\mathbb{Z}]_1^p} (S_n(x_1, \mathcal{P}_n) > s) \leq \varphi'_p(s), \quad s > 0, \quad (10.4)$$

for all  $p \in \mathbb{N}$ , where the functions  $(\varphi'_p)_{p \in \mathbb{N}}$  are fast-decreasing as at (2.1).

Recall these conditions imply that the expected degree of a vertex in  $G(\mathcal{P}_n)$  is finite, i.e., they imply sparsity. As noted in Section 8.1.1 and Appendix A, the class of stabilizing interaction graphs includes certain proximity graphs from computational geometry.

## 10.2 Asymptotic normality for statistics of particle systems

Our goal is to study certain scores associated with the history of the marks  $M_n[x]$  for all  $x \in \mathcal{P}_n$  (describing the evolution of the updates of the states in  $\mathbb{M}$  of  $x$  during the time interval  $[0, t_0]$ ,  $t_0 < \infty$ ). We restrict to the real-valued score function

$$\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \xi^{(h, \Phi, \mathcal{G})}(\tilde{x}, \tilde{\mathcal{P}}_n) := h(M_n[x]), \quad (10.5)$$

where  $h : \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}} \rightarrow \mathbb{R}$  is measurable. We aim to establish Gaussian fluctuations for the summary statistics

$$H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \quad (10.6)$$

and the (possibly signed) random measures

$$\mu_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d_x}}. \quad (10.7)$$

For example, for a suitable choice of  $h$ , the statistic  $H_n^\xi$  could either count the total number of sites in a given state  $M_0 \subset \mathbb{M}$  at time  $t_0$ , the total number of sites spending at least time  $t_{\min}$  in the state  $M_0$ , or the total time sites are in a given state  $M_0$ . Section 10.3 provides detailed examples.

Finally, given  $p \in [1, \infty)$ , we say that the score function  $\xi$  in (10.5) satisfies the *p-moment condition on finite windows* with respect to  $\tilde{\mathcal{P}}$  if

$$\sup_{1 \leq n < \infty} \sup_{1 \leq q \leq p} \sup_{x_1, \dots, x_q \in W_n} \mathbb{E}_{[\mathbb{Z}]_1^q} [\max(1, |\xi(\tilde{x}, \tilde{\mathcal{P}}_n)|^p)] \leq M_p^\xi < \infty, \quad (10.8)$$

where  $M_p^\xi$  is assumed to be non-decreasing in  $p$ .

All of our applications in Section 10.3 showing asymptotic normality will depend on the following main result or its variants. Recall the notion of Palm distributions and correlations from Sections 2.1 and 2.2.

**Theorem 10.1** (CLT for continuous time interacting particle systems). *Let the input process to the interacting particle system be*

$$\tilde{\mathcal{P}} = \left\{ (x, M(x), \tilde{\tau}_x) \in \mathbb{R}^d \times \mathbb{M} \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{L}} \right\}, \quad (10.9)$$

where  $\tilde{\tau}_x$  are i.i.d. and independently marked Poisson clocks, which are also independent of sites and initial states, i.e., independent of  $\sum_x \delta_{(x, M(x))}$ . Assume that:

(i)  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations, as in Definition 4.8, and has uniformly bounded correlation functions under reduced Palm distributions, i.e., for all  $p \in \mathbb{N}$ , there exists a constant  $\hat{\kappa}_p$  such that for all  $m \in \mathbb{N}$ ,

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \sup_{y_1, \dots, y_m \in \mathbb{R}^d} \rho_{[x]_1^p}^{(m)}([y]_1^m) \leq (\hat{\kappa}_p)^m, \quad (10.10)$$

where  $\rho_{[x]_1^p}^{(m)}$  denotes the  $m$ -th order correlation function of  $\mathcal{P}$  under its reduced Palm distribution  $\mathbb{P}_{[x]_1^p}^!$ .

(ii) An admissible update rule given by a measurable function  $\Phi$  as in (10.2) acts on a stabilizing interaction graph  $\mathcal{G}$  (on finite windows) with respect to  $\mathcal{P}$ , such that the interaction range satisfies the decay assumption (10.4).

(iii) A score function  $\xi = \xi^{(h, \Phi, \mathcal{G})}(\tilde{x}, \tilde{\mathcal{P}}_n) := h(M_n[x])$  defined in (10.5), satisfy the  $p$ -moment condition (10.8) for all  $p \in [1, \infty)$ .

Then the sequence of random measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  defined in (10.7) satisfies the central limit theorem. That is, for every test function  $f \in \mathbb{B}(W_1)$  such that  $\text{Var}[\mu_n^\xi(f)] = \Omega(n^\nu)$  for some  $\nu > 0$ , we have

$$(\text{Var}[\mu_n^\xi(f)])^{-1/2} (\mu_n^\xi(f) - \mathbb{E}[\mu_n^\xi(f)]) \xrightarrow{d} Z,$$

where  $Z$  is a standard Gaussian random variable.

In the sequel, when we say that Theorem 10.1 holds, we implicitly mean that it is valid subject to a variance lower bound  $\text{Var}[\mu_n^\xi(f)] = \Omega(n^\nu)$  for some  $\nu > 0$ , which, as always, we regard as a separate problem.

The Palm correlation bound condition (10.10) implies that all factorial (and ordinary) moment measures under Palm distributions are dominated by those of a Poisson point process with intensity  $\hat{\kappa}_p$ . This motivates the definition of *sub-Poisson processes* in the sense of moment measures, as introduced in [18], with determinantal processes being a prominent example. See Appendix A for further examples.

**Remark 10.2** (Further comments and comparison with previous work).

(i) (*Summable exponential  $\mathbb{B}$ -mixing* of points and initial marks is sufficient.) In Theorem 10.1, we assumed that the Poisson marked clocks  $\tilde{\tau}_x \in \hat{\mathcal{N}}_{[0,t_0] \times \mathbb{L}}$  are i.i.d. given sites and initial conditions. Consequently, in the summable exponential  $\mathbb{B}$ -mixing assumption for  $\tilde{\mathcal{P}}$ , one can omit these clocks and verify the condition only for  $\{\tilde{\delta}_{(x, M(x))}\}_{x \in \mathcal{P}}$ .

(ii) (*LLN and variance asymptotics.*) Under the assumptions of Theorem 10.1, but with the  $p$ -moment condition (10.8) weakened to hold for some  $p \in (2, \infty)$ , and without any assumption on the variance rate of  $\mu_n^\xi(f)$ , suppose further that the input process  $\tilde{\mathcal{P}}$  is stationary and that both the graph  $\mathcal{G} = \mathcal{G}(\mathcal{P})$  and the update function  $\Phi$  in (10.2) are translation invariant. Then, by Proposition 5.3, the mean and variance asymptotics of  $n^{-1}\mu_n^\xi(f)$  are given by (5.5) and (5.6), involving Palm-distributional limits in (5.3) and (5.4). If, in addition, the interaction ranges  $S_n(x, \mathcal{P}_n)$  are uniformly bounded in  $n$  almost surely for all  $x \in \mathcal{P}$ , i.e. satisfy  $\sup_n S_n(x, \mathcal{P}_n) < \infty$  a.s., then the score function  $\xi = \xi^{(h, \Phi, \mathcal{G})}$  has a limit

$$\lim_{n \rightarrow \infty} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \xi_\infty(\tilde{x}, \tilde{\mathcal{P}}) \quad \text{a.s. for all } x \in \mathcal{P}.$$

Interaction ranges for the spatial random graph examples in Appendix A can be shown to satisfy the uniform bound required above.

Moreover, the mean and variance asymptotics admit representations involving this limit:

$$\begin{aligned} \left| n^{-1} \mathbb{E} \mu_n^\xi(f) - \rho \mathbb{E}_0 \xi_\infty(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \int_{W_1} f(x) dx \right| &= O(n^{-1/d}), \\ \lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^\xi(f) &= \sigma^2(\xi_\infty) \int_{W_1} f(x)^2 dx \in [0, \infty), \end{aligned} \quad (10.11)$$

where

$$\xi_\infty(\tilde{x}, \tilde{\mu}) = \xi_\infty^{(h, \Phi, \mathcal{G})}(\tilde{x}, \tilde{\mu}) := \lim_{n \rightarrow \infty} \xi(\tilde{x}, \tilde{\mu} \cap W_n), \quad (10.12)$$

$$\sigma^2(\xi_\infty) := \rho \mathbb{E}_0 \xi_\infty^2(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \quad (10.13)$$

$$+ \int_{\mathbb{R}^d} \left[ \mathbb{E}_{\mathbf{0}, x}(\xi_\infty(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \xi_\infty(\tilde{x}, \tilde{\mathcal{P}})) \rho^{(2)}(\mathbf{0}, x) - \rho^2 \mathbb{E}_0 \xi_\infty(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}) \mathbb{E}_x \xi_\infty(\tilde{x}, \tilde{\mathcal{P}}) \right] dx \in [0, \infty).$$

This representation, as described in Remark (ii) (Section 5.4), can be justified via Lemma B.1. This requires deriving a uniform in  $n \in \mathbb{N}$  bound on the stabilization radius  $R_{W_n}^\xi$  of  $\xi$  (see (10.14) in the proof of Theorem 10.1). The bound on  $R_{W_n}^\xi$  is derived using a graphical construction for interacting particle systems and controlling the diameter of the corresponding backward space-time cluster, which tracks all the nodes influencing updates at a given node. Though not straightforward, one can derive uniform bounds for the diameter by exploiting the uniform boundedness of the interaction range  $S_n$  for  $n \in \mathbb{N}$ .

(iii) (*Local synchronous, global asynchronous update rules.*) The function  $\Phi$  in (10.2) is applied successively with each ringing of some clock  $\tau_x, x \in \mathcal{P}_n$ . At time  $T \in \tau_x$ , the application of  $\Phi$  creates a *local, synchronous update* of the history at  $x$  and its neighbors  $y \sim x$ : by the stopping set property (8.5) of  $S_n$ , this update is local because the neighborhood of  $x$  has a radius bounded by  $S_n(x, \mathcal{P}_n)$ , also, it is

synchronous because the new states at  $x$  and its neighbors are evaluated jointly based on all histories at  $x$  and its neighbors (with, if necessary, priorities assigned according to the location of  $x$ , its time-mark  $L$  triggering this update, and the positions of the neighbors). At the same time  $T$ , all locations  $y'$  which are not neighbors of  $x$  are not updated and their histories do not contribute to the evaluation of the local update around  $x$ . This is hence *asynchronous updating occurring globally*. It is in contrast to a discrete, globally synchronous updates considered in Section 10.4.

(iv) (Non-Markovian updates.) Observe, locations  $x \in \mathcal{P}_n$  are typically updated multiple times in the period  $[0, t_0]$ . Hence, the process of the history  $M[x] = M(x, \cdot)$  may be viewed as *a pure jump type process*, but it need not be Markovian as in [117, Section 2.1]. This is because the update rule depends upon the entire history of the process at a site and its neighboring sites. (This was also the case with the diffusion model in Section 9.1). Allowing the update to depend only upon the current configuration makes the process Markovian and hence encompasses well-studied examples in the literature. Such examples are described in detail in Section 10.3.

(v) (Comparison with Penrose [117].) We compare the theorems of this section with the closely related results of [117]. Theorem 10.1 requires neither Poisson input, independence of initial states, nor bounded interaction range. In this way we extend upon the thermodynamic (LLN) and Gaussian limit theory of [117] and address some questions raised in this article (see especially Remark 6 in section 3 of [117], also Section 5.2 of [117]). Penrose's proof of the thermodynamic and Gaussian limits rely upon an add-one cost stabilization condition for the functionals  $H_n^\xi$  (which is distinct from the stabilization and localization of scores described in Section 4.3), discretization, and a martingale central limit theorem, an approach relying upon the existence of the infinite particle system. By contrast, our approach, which uses the method of cumulants, only requires stabilization of the scores  $\xi$  on particle systems confined to the window  $W_n$ , bypassing the question of existence. We show stabilization of the scores  $\xi$  via a graphical construction. Showing existence of the limit of the finite systems to a limiting dynamic system on the infinite input  $\tilde{\mathcal{P}}$  lies outside the scope of this paper though it may be possible to use the approaches from [117, 108] in combination with our derivations. Though graphical constructions have been previously used to study interacting particle systems (see [117, 119] for example), the analysis of our graphical construction is more involved due to the spatial dependencies of the underlying point processes and the unbounded interaction range.

(vi) (Related literature.) With the exception of Qi [125], the limit theory for interacting particle systems has assumed finite range interactions. The paper [125] proves a functional central limit theorem (FCLT) for spatial birth-death processes with possibly unbounded interaction range, provided the processes are obtained as a solution to stochastic equations. The study of asymptotic Gaussianity was initiated by Holley and Stroock [76], who used classical tightness arguments to establish the FCLT for lattice based systems. Later, Doukhan et al. [46] worked on general transitive graphs, sometimes requiring bounded degree, but still assuming finite range interactions. More recently, Onaran et al. [111] establish a FCLT for spatial birth-death processes on sites given by Poisson input. They deduce convergence of finite-dimensional distributions from Lachièze-Rey et al. [93] and establish tightness using a classical criteria from Ethier and Kurtz [54]. In case of birth-death processes with diffusive

dynamics, finite-dimensional convergence of local statistics has been shown in Onaran et al. [112].

(vii) (Extensions.) Setting  $\xi_i = \xi_i^{(h, \Phi, \mathcal{G})} = h(M(x, t_i))$ ,  $1 \leq i \leq k$ , for  $h : \mathbb{K} \rightarrow \mathbb{R}$  bounded and measurable and given  $t_1, \dots, t_k \in [0, \infty)$ , we may obtain a multivariate central limit theorem for  $(\mu_n^{\xi_i})_{i=1, \dots, k}$  provided the covariances of the entries converge at some scale  $n^{-\mu}$  for some positive  $\mu$ ; see the discussion below Theorem 5.4. As in Item (ii) above, when  $\nu = 1$ , one can deduce convergence of the covariances provided the interaction ranges  $S_n$  of the graph  $\mathcal{G}$  are uniformly bounded. One expects that central limit theorems could be extended to functional central limit theorems yielding convergence of scaled versions of the process  $(\sum_{x \in \mathcal{P}_n} h((M_n(x, t))_{t \in [0, t_0]}))$ , thus extending Theorem 3.3 of [117], which assumes independent initial marks and bounded interaction range, and also extending the functional central limit theorem of [125] given for certain birth death processes having polynomial decay of interactions. Furthermore, one may expect to establish a functional central limit theorem for the function-indexed process  $(\mu_n(f))_{f \in \mathcal{F}}$ , with  $\mathcal{F}$  being a suitable function class and  $t = t_0$  fixed, but to our knowledge such a general functional limit theorem is not available even in the simpler setting consisting of Poisson input, independent marks, and unbounded interaction range. We expect that such a general functional central limit theorem would extend those given in more specialized settings considered by Biscio et al. [16], [15].

(viii) (Measurability of the score function  $\xi^{(h, \Phi, \mathcal{G})}$ .) We now justify the measurability of  $g$  defined at (10.3). We need to show that the map  $g$  from  $\mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}}$  to  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$  given by  $(\tilde{x}, \tilde{\mu}) \mapsto (M(x, t))_{t \in [0, t_0]}$  is measurable. Note  $M[x] = (M(x, t))_{t \in [0, t_0]}$  is the final mark of  $x$  in  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$ , measurable as a consequence of the iteration of the measurable marking function  $\Phi$  given in (10.2), each time being executed based on the measurable input consisting of: the location  $x \in \mathcal{P}_n$ , its clock event  $(T, L) \in \tilde{\tau}_x$ , and the set of (prior) histories  $\{(y, M(y, \cdot))\}_{y \in \mathcal{P}_n: y \simeq x} \in \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathcal{N}_{[0, t_0] \times \mathbb{M}}}$ , measurably by the induction starting from the measurable initial states  $M(y, 0)$ . Here, the measurability of the neighborhood  $\sim$  (in the sense of the neighborhood-sets described in Section 8.1.1) is required in the specification of the arguments of  $\Phi$ .

### 10.2.1 Proofs

We next prove Theorem 10.1. The proof relies upon deviation inequalities for number counts of the point process  $\mathcal{P}$  as well as for degrees and stabilization radii of spatial random graphs. These estimates will be proved at the end of section in Lemmas 10.3, 10.4 and 10.5.

*Proof of Theorem 10.1.* We show that  $(\mu_n^\xi)_{n \in \mathbb{N}}$  satisfy the assumptions of Theorem 5.2, with  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) := h(M[x]) = h((M_n(x, t))_{t \in [0, t_0]})$  as at (10.5). The  $p$ -moment condition (10.8) of  $\xi = \xi^{(h, \Phi, \mathcal{G})}$  for all  $p \in (1, \infty)$  with respect to  $\mathcal{P}$  is a verbatim repetition of the moment condition (5.2) required in this theorem. The variance lower bound holds by assumption and thus to verify the assumptions in Theorem 5.2, we need only show that  $\xi$  is fast BL-localizing on finite windows as in Definition 4.2 (ii) and (iii). In this regard, recall from Remark 4.7 that this is implied by fast stabilization on finite windows, as specified in Definition 4.6 (ii), (iii), which we will aim to establish.

Our proof is based on the *study of backward clusters of ‘information’ propagation*. For a site  $x$  this cluster consists of the set of sites in  $\mathcal{P}_n$  from which its history  $M_n(x, t)$  can be influenced at any

time  $t \in [0, t_0]$ . This is introduced in *Step 1*. Next in *Step 2*, we *control the diameter of the backward clusters*, i.e., we upper-bound the (appropriate Palm) probability that the state at  $x$  is influenced by sites outside the ball  $B_m(x)$  of radius  $m \geq 1$ . We show that this probability is fast-decreasing to zero as a function of  $m \rightarrow \infty$ , in the sense of (2.1). This is essentially because there are not enough ‘opportunities’, i.e. not enough ordered events with respect to clock rings on paths from  $B_m^c(x)$  to  $x$ , especially taking into account probabilistic constraints on edge lengths and on degrees of sites in our stabilizing interaction graph  $\mathcal{G}$  on  $\mathcal{P}$  satisfying decay bounds (10.4). Finally in *Step 3*, we add a *buffer stabilizing region* around  $B_m(x)$  to account for stabilization of graph edges making  $\xi$  fast stabilizing at  $x$ . The proof details are as follows.

*Step 1. Definition of backward clusters.* Given the graph  $G(\mathcal{P}_n)$ ,  $n \in \mathbb{N}$ , consider the ‘super-clocks’ at locations  $x \in \mathcal{P}_n$ ,

$$\hat{\tau}_x := \hat{\tau}_{n,x} = \tau_x \cup \bigcup_{y \in \mathcal{P}_n: y \sim x} \tau_y = \bigcup_{y \in \mathcal{P}_n: y \simeq x} \tau_y \subset [0, t_0],$$

i.e., the union of the clock at  $x$  and all neighboring clocks, with the neighbor relation respecting (8.4). Since the (original) Poisson clocks are i.i.d., we may assume that a.s. the clocks do not ‘ring’ together, i.e.,  $\tau_x \cap \tau_y = \emptyset$  for all  $x, y \in \mathcal{P}_n$ ,  $x \neq y$ . Observe that  $M(x, \cdot)$  is updated only at ringing times of the ‘super-clock’  $\hat{\tau}_x$ , i.e., the history  $M(x, \cdot)$  has events  $(t, M(x, t))$  only for  $t \in \hat{\tau}_x$  created according to the update function (10.2), apart from its own history  $\{M(x, s)\}_{s < t}$  and influenced by the prior histories  $\{M(y', s)\}_{s < t}$  of the sites  $y'$  whose ‘super-clocks’ ring at the same time  $t \in \hat{\tau}_y$ .

In order to track sites (the points of  $x \in \mathcal{P}_n$ ) and times of their ‘super-clocks’  $\hat{\tau}_x$  which mutually influence their updates we define an *oriented graph*  $\hat{\mathcal{G}}(\mathcal{P}_n)$  as follows: The vertices of the oriented graph  $\hat{\mathcal{G}}(\mathcal{P}_n)$  are  $V = V_n := \{(x, t) : x \in \mathcal{P}_n, t \in \hat{\tau}_x \cup \{0\}\}$  and the oriented edges are  $(x, t) \rightarrow (x', t')$  if  $t < t'$  and  $(x, t') \in V$  (i.e.,  $t' \in \hat{\tau}_x$ ). Since the clocks  $\tau_x, x \in \mathcal{P}_n$  have no common ‘rings’, the condition that  $(x', t'), (x, t) \in V$  is equivalent to either (i)  $t' \in \tau_x \cup \tau_{x'}$  and  $x \sim x'$  or (ii)  $t' \in \tau_{x''}$  for some  $x''$  where  $x'' \sim x', x'' \sim x$ . In the first case, the clock at  $x$  or  $x'$  rings at time  $t'$  and because of update rule (10.2), the update at  $x$  or  $x'$  depends on previous updates at neighboring points including the other point. In the second case, the clock at a mutual neighbor  $x''$  rings at time  $t'$  and the update at  $x$  and  $x'$  is influenced by previous updates at neighboring points including the point  $x''$  due to the update rule (10.2). Thus the oriented graph captures space-time dependency between updates. We shall now make this more precise.

Define  $C'(x, t) = C'_n(x, t)$ ,  $(x, t) \in V = V_n$ , to be the *backward (in time) cluster* of  $(x, t)$ , i.e.,  $(x', t') \in C'(x, t)$  if there exists a path in  $\hat{\mathcal{G}}(\mathcal{P}_n)$  from  $(x', t')$  to  $(x, t)$ . The events at times  $s$  prior to  $t'$  of the histories of sites  $x' \in \mathcal{P}_n$

$$\{((s, M(x', s)))_{s < t'} : (x', t') \in C'(x, t)\}$$

are precisely those which contribute to the evaluation of the value of the updated  $(t, M(x, t)) \in M(x, \cdot)$  by the function (10.2), given the initial states  $M(x', 0)$ .

By cylinder of radius  $r$  with axis through  $x \in \mathbb{R}^d$  we mean the set  $B_r(x) \times [0, t_0]$ . Let  $\hat{t}_x$  be the last time of the ‘super-clock’  $\hat{\tau}_x$ . The radius of stabilization  $R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  of  $\xi = \xi^{(h, \Phi, \mathcal{G})}$  as in Definition 4.5 may be upper bounded by the radius of the smallest cylinder containing the backwards

cluster  $C'(x, \hat{t}_x) = C'_n(x, \hat{t}_x)$ , centered at  $x \in \mathcal{P}_n$ , plus the largest graph stabilization radius  $S_n(x', \mathcal{P}_n)$  as in (8.2) among points  $x' \in \mathcal{P}_n$  in the basis of this cylinder, i.e.,

$$R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n) \leq \max_{(x', t') \in C'_n(x, \hat{t}_x)} |x' - x| + \max_{(x', t') \in C'_n(x, \hat{t}_x)} S_n(x', \mathcal{P}_n). \quad (10.14)$$

Indeed, the histories  $(s, M(x', s))_{s \leq t'}$  for  $(x', t') \in C'_n(x, \hat{t}_x)$  are sufficient to construct the complete history  $(s, M(x, s))_{s \leq t_0}$  of  $x \in \mathcal{P}_n$ , and the second term in (10.14) corresponds to a buffer region of  $C'_n(x, \hat{t}_x)$ , such that the addition of points outside this buffer region will not affect the backwards cluster  $C'_n(x, \hat{t}_x)$ , and hence will not affect the score at  $x$ .

The two terms in the right-hand-side of (10.14) may be controlled using the Palm correlation functions of the ground process  $\mathcal{P}$  as in (10.10), the fast decay of the graph stabilization radius  $S_n$  as in (10.4), and the ‘asymptotic independence’ of ‘super-clocks’. This goes as follows: We shall bound the first term on the right-hand side of (10.14) in Step 2, which forms the bulk of the remaining proof. In Step 3, using Lemma 10.5 at the end of this section, we bound the second term on the right-hand side of (10.14). The requisite deviation bounds and degree bounds are derived at the end of the section in Lemmas 10.3 and 10.4.

*Step 2. Controlling the diameter of the backward clusters.* In this step we control the growth of

$$D(\tilde{x}, \tilde{\mathcal{P}}_n) := \max_{(x', t') \in C'_n(x, \hat{t}_x)} |x' - x|. \quad (10.15)$$

Fix  $\alpha \in (0, 1)$  and  $\beta \in (0, (1 - \alpha)/(d + 1))$  (where  $d$  is the dimension of  $\mathbb{R}^d$ ) and for  $m \in \mathbb{N}$  set  $r_m := m^\beta$ ,  $k_m := m^\alpha$ . For  $x \in \mathcal{P}_n$ , consider the event  $A_m(x) := A_{n,m}(x)$  such that there is a site  $y \in \mathcal{P}_n \cap B_m(x)$  such that either (i) its degree in the graph  $\mathcal{G}(\mathcal{P}_n)$ , denoted  $\text{Deg}(y, \mathcal{P}_n)$ , is at least  $k_m$ , i.e.,  $\text{Deg}(y, \mathcal{P}_n) \geq k_m$ , or (ii) one of the edges incident to it has length at least  $r_m$ . Conditioning on  $A_m(x_1)$  and its complement  $A_m^c(x_1)$  we get

$$\mathbb{P}_{x_1, \dots, x_p}(D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m) \leq \mathbb{P}_{[x_1]^p}(A_m(x_1)) + \mathbb{P}_{[x_1]^p}(\{D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m\} \cap A_m^c(x_1)). \quad (10.16)$$

For the event  $A_m(x_1)$  to occur under  $\mathbb{P}_{x_1, \dots, x_p}$ , in view of the rule for the existence of edges (8.5), there must be a site  $x \in \mathcal{P}_n \cap B_m(x_1)$  such that  $S_n(x, \mathcal{P}_n) \geq r_m$  or  $\text{Deg}(x, \mathcal{P}_n) \geq k_m$ . Applying the union bound, the Campbell-Little-Mecke formula (2.5), (10.10), we have

$$\begin{aligned} \mathbb{P}_{[x_1]^p}(A_m(x_1)) &\leq \int_{B_m(x)} \mathbb{P}_{x, [x_1]^p}(S_n(x, \mathcal{P}) \geq r_m \text{ or } \text{Deg}(x, \mathcal{P}_n) \geq k_m) \rho_{[x_1]^p}^{(1)}(x) dx \\ &\quad + \sum_{i=1}^p \mathbb{P}_{[x_1]^p} \left( (S_n(x_i, \mathcal{P}_n) > r_m) + p \mathbb{P}_{[x_1]^p}(\text{Deg}(x_i, \mathcal{P}_n) \geq k_m) \right) \\ &\leq \theta_d \hat{\kappa}_p m^d (\varphi'_{p+1}(r_m) + \varphi''_{p+1}(k_m)) + p(\varphi'_p(r_m) + \varphi''_p(k_m)), \end{aligned} \quad (10.17)$$

where  $\varphi'_p$  is at (10.4) and where  $\varphi''_p$ , defined at (10.21), controls the degree bound given in Lemma 10.4 at the end of this section.

Recalling that  $r_m = m^\beta$  and  $k_m = m^\alpha$  we find that for (any)  $\alpha, \beta > 0$  the expression in (10.17) is a fast decreasing function of  $m \in \mathbb{N}$  and so is the first term on the right-hand side of (10.16).

We now bound the second term in (10.16), as a fast decreasing of  $m$ . We define the sequence  $l_m := \lfloor m^{1-\beta}/(2(p+1)) \rfloor$ ,  $m \in \mathbb{N}$ . Observe, on  $A_m^c(x_1)$ , if  $D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m$  then there exist *distinct* sites  $y_1, \dots, y_{l_m} \in \mathcal{P}_n \cap B_m(x_1)$ , *different from*  $x_1, \dots, x_p$ , such that  $|y_i - y_{i-1}| < 2(p+1)r_m$ ,  $i = 1, \dots, l_m$ , where we set  $y_0 := x_1$ , whose super-clocks (when they update their states) ring at some strictly positive times in  $\hat{\tau}_{y_i}$ , with the exception of the earliest time being possibly the initial time zero. Specifically, looking backwards in time from  $t_0$ , we enumerate these nodes  $y_i$  and their corresponding updating times in reverse chronological order  $t_0 \geq \hat{T}_1 > \hat{T}_2 > \dots > \hat{T}_{l_m}$  with  $\hat{T}_i \in \hat{\tau}_{y_i}$  for  $i = 1, \dots, l_m$  and  $\hat{T}_{l_m} \in \hat{\tau}_{y_i} \cup \{0\}$ .

Note, the degrees at all sites  $y_i$  are bounded by  $\text{Deg}(y_i, \mathcal{P}_n) < k_m$ . Indeed, when all edges of  $\mathcal{G}(\mathcal{P}_n)$  in  $B_m(x_1)$  have length smaller than  $r_m$  (complement of the condition (ii) of the event  $A_m(x_1)$ ), then the corresponding oriented edges in  $\hat{\mathcal{G}}(\mathcal{P}_m)$  (with chronologically oriented times of clocks), when projected on  $\mathbb{R}^d$ , are shorter than  $2r_m$ , and hence, one needs at least  $l'_m := \lfloor m/(2r_m) \rfloor$  distinct sites  $\{y'_i : i = 1, \dots, l'_m\}$  in  $\mathcal{P}_n$  to cover the distance from  $x_0$  to  $B_m^c(x_1)$  on  $\hat{\mathcal{G}}(\mathcal{P}_n)$ . Among these sites  $y'_i$  there might be fixed atoms  $x_1, \dots, x_p$  under  $\mathbb{P}_{[x]_1^p}$ ; removing them one obtains the considered path  $\{y_i : i = 1, \dots, l_m\}$ . Clearly  $\text{Deg}(y_i, \mathcal{P}_n) < k_m$  results from the complement of the condition (i) of  $A_m(x_1)$ .

Considering the path  $\{y_i : i = 1, \dots, l_m\}$ , using the Markov inequality and the Campbell-Little-Mecke formula one bounds the considered probability by

$$\begin{aligned} & \mathbb{P}_{[x]_1^p}(\{D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m\} \cap A_m^c(x_1)) \\ & \leq \mathbb{E}_{[x]_1^p} \left[ \sum_{\substack{(y_1, \dots, y_{l_m}) \in (\mathcal{P}_n \setminus \{x_1, \dots, x_p\})^{(l_m)} \\ \hat{T}_i \in \hat{\tau}_{y_i} : i=1, \dots, l_m}} \prod_{i=1}^{l_m} \mathbf{1}(|y_i - y_{i-1}| < 2(p+1)r_m) \mathbf{1}(\hat{T}_i > \hat{T}_{i+1}) \mathbf{1}(\text{Deg}(y_i, \mathcal{P}_n) < k_m) \right] \\ & \leq \int_{[y]_1^{l_m} \in (\mathbb{R}^d)^{l_m}} \rho_{[x]_1^p}^{(l_m)}([y]_1^{l_m}) \prod_{i=1}^{l_m} \mathbf{1}(|y_i - y_{i-1}| < 2(p+1)r_m) \\ & \quad \times \mathbb{E}_{[x]_1^p \sqcup [y]_1^{l_m}} \left[ \sum_{\hat{T}_i \in \hat{\tau}_{y_i} : i=1, \dots, l_m} \prod_{i=1}^{l_m} \mathbf{1}(\hat{T}_i > \hat{T}_{i+1}) \mathbf{1}(\text{Deg}(y_i, \mathcal{P}_n) < k_m) \right] d[y]_1^{l_m} \quad (10.18) \end{aligned}$$

where, for the convenience, we set  $y_0 := x_1$  and  $\hat{T}_{l_m+1} := -1$ . Note the super-clocks  $\hat{\tau}_{y_i}$  are, in general, not independent since the sets of neighbors of sites  $y_i$  may not be disjoint. However, under  $\mathbb{P}_{[x]_1^p \sqcup [y]_1^{l_m}}$  with  $\text{Deg}(y_i, \mathcal{P}_n) < k_m$ , these aggregated clocks are still homogeneous Poisson point processes on  $[0, t_0]$  of intensities bounded by  $k_m$ , with the last (index  $l_m$ ) appended by the initial time event  $t = 0$ . First assuming  $\hat{T}_{l_m} > 0$ , using the strong Markov property for the aggregated clocks one gets

$$\mathbb{E}_{[x]_1^p \sqcup [y]_1^{l_m}} \left[ \sum_{\hat{T}_i \in \hat{\tau}_{y_i} : i=1, \dots, l_m} \prod_{i=1}^{l_m} \mathbf{1}(\hat{T}_i > \hat{T}_{i+1}) \mathbf{1}(D(y_i, \mathcal{P}_n) < k_m) \right]$$

$$\begin{aligned}
&\leq k_m^{l_m} \int_0^{t_0} ds_{l_m} \int_{s_{l_m}}^{t_0} ds_{l_m-1} \cdots \int_{s_3}^{t_0} ds_2 \int_{s_2}^{t_0} ds_1 \\
&= \frac{k_m^{l_m} t_0^{l_m}}{l_m!}.
\end{aligned}$$

Considering separately  $\hat{T}_{l_m} = 0$  one obtains similarly that the expected number of the considered paths is equal to  $k_m^{l_m-1} t_0^{l_m-1} / (l_m - 1)!$ . Combining and substituting into (10.18) and using (10.10) one obtains

$$\mathbb{P}_{[\mathbb{Z}]_1^p}(\{D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m\} \cap A_m^c(x_1)) \leq \hat{\kappa}_p^{l_m} (\theta_d (2(p+1)r_m)^d)^{l_m} t_0^{l_m} \left(1 + \frac{l_m}{k_m t_0}\right) \frac{k_m^{l_m}}{l_m!}. \quad (10.19)$$

Recall  $r_m = m^\beta$ ,  $l_m = \lfloor m^{1-\beta} / (2(p+1)) \rfloor$ , and  $k_m = m^\alpha$ , with  $\alpha \in (0, 1)$  and  $\beta \in (0, (1-\alpha)/(d+1))$ . Using a crude Stirling bound  $l_m! \geq C e^{\frac{1}{2} l_m \log l_m}$ , with  $C > 0$  some small constant, and applying it to (10.19), one obtains the following bounds, with some further positive, finite constants  $C_1, C_2$  (not dependent on  $n, m$  and  $\alpha, \beta$ )

$$\begin{aligned}
\mathbb{P}_{[\mathbb{Z}]_1^p}(\{D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq m\} \cap A_m^c(x_1)) &\leq C_1 \left(1 + \frac{l_m}{k_m}\right) \exp\left(C_2 l_m + \frac{1}{2} l_m \log \frac{r_m^d k_m}{l_m}\right) \\
&\leq C_1 (1 + m^{1-\beta-\alpha}) \exp\left(m^{1-\beta} (C_2 + \frac{1}{2} \log m^{\beta(d+1)+\alpha-1})\right).
\end{aligned} \quad (10.20)$$

Observe  $\beta(d+1) + \alpha - 1 < 0$ , whence the last bound is a fast-decreasing function of  $m$ . This shows that the second term in (10.16) is fast-decreasing in  $m$  and together with the same statement for the bound at (10.17), this yields that the random variable  $D(\tilde{x}_0, \tilde{\mathcal{P}}_n)$  has a fast decreasing tail.

*Step 3. Controlling the width of the buffer region.* Here we wish to control the term

$$\mathcal{W}(\tilde{x}, \tilde{\mathcal{P}}_n) := \max_{(x', t') \in C'_n(x, \hat{t}_x)} S_n(x', \mathcal{P}_n)$$

in (10.14). Notice that by (10.15),

$$\mathcal{W}(\tilde{x}, \tilde{\mathcal{P}}_n) \leq \sup_{|x' - x| \leq D(\tilde{x}, \tilde{\mathcal{P}}_n)} S_n(x', \mathcal{P}_n).$$

Since  $D(\tilde{x}, \tilde{\mathcal{P}}_n)$  and  $S_n(x, \mathcal{P}_n)$  are fast decreasing, it follows by Lemma 10.5 with  $Q_1$  set to  $D$  and  $Q_2$  set to  $S_n$ , that  $\mathcal{W}(\tilde{x}, \tilde{\mathcal{P}}_n)$  is fast decreasing.

Since  $R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n) \leq D(\tilde{x}, \tilde{\mathcal{P}}_n) + \mathcal{W}(\tilde{x}, \tilde{\mathcal{P}}_n)$ , where both  $D(\tilde{x}, \tilde{\mathcal{P}}_n)$  and  $\mathcal{W}(\tilde{x}, \tilde{\mathcal{P}}_n)$  are fast decreasing, it follows that  $R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  is fast decreasing on finite windows as in Definition 4.6 (ii), (iii).  $\square$

We prove three lemmas used in the proof of Theorem 10.1.

**Lemma 10.3** (Deviation inequality for number counts). *Let  $\mathcal{P}$  be a simple point process satisfying the Palm correlation bound (10.10). Then for a bounded Borel set  $B \subset \mathbb{R}^d$ ,*

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{P}_{[\mathbb{Z}]_1^p}(\mathcal{P}(B) \geq \hat{\kappa}_p |B| + p + t) \leq \exp\left(-t \left(\log\left(1 + \frac{t}{\hat{\kappa}_p |B|}\right) - 1\right)\right), \quad p \in \mathbb{N}, \quad t > 0.$$

*Proof.* By the relation between Palm and reduced Palm, we have

$$\mathbb{P}_{[x]_1^p}(\mathcal{P}(B) \geq \hat{\kappa}_p|B| + p + t) \leq \mathbb{P}_{[x]_1^p}(\mathcal{P}(B) \geq \hat{\kappa}_p|B| + t),$$

and so we shall bound the latter probability. To do so, we bound  $\mathcal{P}(B)$  by a Poisson random variable  $N(B)$  with mean  $\hat{\kappa}_p|B|$  whose (factorial, hence non-factorial) exponential moments are larger; cf. [18, 19]. Indeed, using definition of Palm correlation functions and (10.10), we have that

$$\mathbb{E}_{[x]_1^p}^! \mathcal{P}(B)^{(k)} = \int_B \rho_{[x]_1^p}([y]_1^k) dy_1 \dots dy_k \leq (\hat{\kappa}_p|B|)^k = \mathbb{E}N(B)^{(k)}$$

and hence for  $u \geq 0$  we have  $\mathbb{E}_{[x]_1^p}^![e^{u\mathcal{P}(B)}] \leq \mathbb{E}[e^{uN(B)}]$ . The rest of the proof now follows from the classical Chernoff bounds for a Poisson random variable:

$$\begin{aligned} \mathbb{P}_{[x]_1^p}^!(\mathcal{P}(B) \geq \hat{\kappa}_p|B| + t) &\leq \exp\left(-\sup_{u \geq 0} \left\{ u(\hat{\kappa}_p|B| + t) - \log \mathbb{E}_{[x]_1^p}^![e^{u\mathcal{P}(B)}] \right\}\right) \\ &\leq \exp\left(-\sup_{u \geq 0} \left\{ u(\hat{\kappa}_p|B| + t) - \log \mathbb{E}[e^{uN(B)}] \right\}\right) \\ &= \exp\left(-(\hat{\kappa}_p|B| + t) \log\left(\frac{\hat{\kappa}_p|B| + t}{\hat{\kappa}_p|B|}\right) + t\right) \\ &= \exp\left(-t \log\left(\frac{\hat{\kappa}_p|B| + t}{\hat{\kappa}_p|B|}\right) - 1 - \hat{\kappa}_p|B| \log\left(\frac{\hat{\kappa}_p|B| + t}{\hat{\kappa}_p|B|}\right)\right) \\ &\leq \exp\left(-t \log\left(\frac{\hat{\kappa}_p|B| + t}{\hat{\kappa}_p|B|}\right) - 1\right), \end{aligned}$$

where we used in the first equality  $\sup_{u \geq 0} (ua - \log \mathbb{E}[e^{uN(B)}]) = a \log(a/\mathbb{E}[N(B)]) - a + \mathbb{E}[N(B)]$  (see e.g. [47, Example 1.10]).  $\square$

**Lemma 10.4** (Degree bounds in stabilizing interaction graphs). *Let  $\mathcal{P}$  be a simple point process on  $\mathbb{R}^d$  satisfying Palm correlation bound (10.10) and let  $\mathcal{G}$  be a stabilizing interaction graph defined on the finite windows of  $\mathbb{R}^d$ , with  $\mathcal{G}(\mathcal{P}_n)$  satisfying (8.4)-(8.3). Denote by  $\text{Deg}(x, \mathcal{P}_n)$  the degree of the site  $x$  in  $\mathcal{G}(\mathcal{P}_n)$ . Then for all  $p \in \mathbb{N}$  there are fast decreasing functions  $\varphi_p''$  such that*

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[x]_1^p}(\text{Deg}(x_1, \mathcal{P}_n) > k) \leq \varphi_p''(k), \quad k \in \mathbb{N}. \quad (10.21)$$

*Proof.* Fix  $p \in \mathbb{N}$  and assume without loss of generality that  $k \geq 2(p+1)$ . Given such a  $k \geq 1$ , choose such that  $\hat{\kappa}_p \theta_d r^d + p = k/2$  and set  $t = k/2$ . By stabilization of the graph, we have that if  $S_n(x_1, \mathcal{P}_n) \leq r$  then  $\text{Deg}(x_1, \mathcal{P}_n) \leq \mathcal{P}(B_r(x_1))$ . By the choice of  $t, r$ , we have

$$\mathbb{P}_{[x]_1^p}(\text{Deg}(x_1, \mathcal{P}_n) > k) \leq \mathbb{P}_{[x]_1^p}(\mathcal{P}(B_r(x_1)) \geq \hat{\kappa}_p \theta_d r^d + p + t) + \mathbb{P}_{[x]_1^p}(S_n(x_1, \mathcal{P}_n) \geq r).$$

Using again the choice of  $t, r$ , the proof follows from the fast stabilization of  $S_n$  as at (10.4) and the deviation inequality Lemma 10.3 applied to  $B = B_r(x_1)$ .  $\square$

**Lemma 10.5** (Tail bounds for the maximum of fast decreasing marks). *Let  $Q_1, Q_2$  be two marking functions on finite marked point processes. Let  $\tilde{\mathcal{P}}$  be a marked point process such that the ground process  $\mathcal{P}$  satisfies the Palm correlation bound (10.10) for  $m = 1$ . Suppose for  $p \in \mathbb{N}$ ,  $i = 1, 2$  that there are fast decreasing functions  $\varphi_{i,p}$  as in (2.1) such that*

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[x]_1^p} (Q_i(\tilde{x}_1, \tilde{\mathcal{P}}_n) > r) \leq \varphi_{i,p}(r), \quad r > 0. \quad (10.22)$$

Then for all  $p \in \mathbb{N}$  there are fast decreasing functions  $\tilde{\varphi}_p$  such that

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{P}_{[x]_1^p} \left( \max_{x \in \mathcal{P}_n: |x - x_1| \leq Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)} Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r \right) \leq \tilde{\varphi}_p(r), \quad r > 0.$$

If the assumption (10.22) holds with the supremum taken up to  $n = \infty$ , i.e.,  $\sup_{1 \leq n \leq \infty}$ , then the conclusion remains valid with  $\sup_{1 \leq n \leq \infty}$ .

*Proof.* Assume that  $Q_i(x, \tilde{\mathcal{P}}_n), i \in \{1, 2\}$  satisfy (10.22). Recall that  $\theta_d$  is the volume of the  $d$ -dimensional unit ball. Thus using the union bound and the Campbell-Little-Mecke formula, we obtain

$$\begin{aligned} & \mathbb{P}_{[x]_1^p} \left( \max_{x \in \mathcal{P}_n: |x - x_1| \leq Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)} Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r \right) \quad (10.23) \\ & \leq \mathbb{E}_{[x]_1^p} \left[ \sum_{x \in \mathcal{P} \cap B_{Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)}(x_1)} \mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \right] \\ & \leq \sum_{x \in [x]_1^p} \mathbb{E}_{[x]_1^p} [\mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r)] \\ & \quad + \mathbb{E}_{[x]_1^p} \left[ \sum_{x \in (\mathcal{P} \setminus [x]_1^p) \cap B_{Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)}(x_1)} \mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \mathbf{1}(Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n) \in [0, 1]) \right] \\ & \quad + \mathbb{E}_{[x]_1^p} \left[ \sum_{j=1}^{\infty} \sum_{x \in (\mathcal{P} \setminus [x]_1^p) \cap B_{Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)}(x_1)} \mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \mathbf{1}(Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n) \in [2^{j-1}, 2^j]) \right] \\ & \leq \sum_{x \in [x]_1^p} \mathbb{E}_{[x]_1^p} [\mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r)] \\ & \quad + \mathbb{E}_{[x]_1^p} \left[ \sum_{x \in \mathcal{P} \cap B_{Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n)}(x_1)} \mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \right] \\ & \quad + \sum_{j=1}^{\infty} \mathbb{E}_{[x]_1^p} \left[ \sum_{x \in \mathcal{P}_n \cap B_{2^j}(x_1)} \mathbf{1}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \mathbf{1}(Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq 2^{j-1}) \right] \\ & \leq p\varphi_{2,p}(r) + \int_{B_1(x_1)} \mathbb{P}_{x, [x]_1^p} (Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r) \rho_{[x]_1^p}^{(1)}(x) dx \\ & \quad + \sum_{j=1}^{\infty} \int_{B_{2^j}(x_1)} \mathbb{P}_{x, [x]_1^p} (Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r, Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq 2^{j-1}) \rho_{[x]_1^p}^{(1)}(x) dx \quad (10.24) \end{aligned}$$

with  $\mathbb{P}_{x, [x]_1^p} = \mathbb{P}_{x, x_1, \dots, x_p}$ , and where the first term in the last bound is a bound for the sum of the probabilities

$$\mathbb{P}_{[x]_1^p}(Q_2(\tilde{x}_j, \tilde{\mathcal{P}}_n) > r) \leq \varphi_{2,p}(r)$$

of the tails of the  $Q_2$  marks of the atoms  $x_j$ ,  $j = 1, \dots, p$ , some of which might be in  $B_{2^j}(x_1)$  but not counted in the Campbell-Little-Mecke integral formula involving the intensity  $\rho_{[x]_1^p}^{(1)}(x)$  of the reduced Palm  $\mathbb{P}_{[x]_1^p}^\dagger$ . This in turn is a fast decreasing function, as  $\varphi_{2,p}$  is fast decreasing as well.

The sum of the integral formulas  $\sum_{j=1}^{\infty} \int_{B_{2^j}(x_1)} \mathbb{P}_{x, [x]_1^p}(\dots) \rho_{[x]_1^p}^{(1)}(x) dx$ , using the Cauchy-Schwartz inequality and the condition (10.10), can be further bounded by

$$\begin{aligned} & \sum_{j=0}^{\infty} \int_{B_{2^j}(x_1)} \mathbb{P}_{x, [x]_1^p}(Q_2(\tilde{x}, \tilde{\mathcal{P}}_n) > r)^{1/2} \mathbb{P}_{x, [x]_1^p}(Q_1(\tilde{x}_1, \tilde{\mathcal{P}}_n) \geq 2^{j-1})^{1/2} \rho_{[x]_1^p}^{(1)}(x) dx \\ & \leq \hat{\kappa}_p(\varphi_{2,p+1}(r))^{1/2} \sum_{j=0}^{\infty} \theta_d 2^{jd} (\varphi_{1,p+1}(2^{j-1}))^{1/2}. \end{aligned}$$

Since  $\varphi_{1,p+1}$  is fast decreasing, the last sum is finite and so the right-hand side in the last line is also fast decreasing as  $\varphi_{2,p+1}$  is fast decreasing as well.

The proof of the version of the result with  $\sup_{1 \leq n \leq \infty}$  replacing  $\sup_{1 \leq n < \infty}$  follows the same lines.  $\square$

### 10.3 Markovian interacting particle systems

We give examples of well-known continuum interacting particle systems, as well as some of their variants, all falling within the framework presented in Section 10.1. We describe update rules in detail and refer the reader to Appendix A for examples of admissible point processes and stabilizing interaction graphs. We will highlight some examples of simple admissible score functions that are of interest. Concerning admissible update rules, we shall focus only on Markovian updates, though the framework allows for analysis of non-Markovian updates.

We again remind the reader that we are studying interacting particle systems in the finite time horizon regime. However, in some models where update rules satisfy certain monotonicity conditions (e.g., the RSA model in Section 10.3.1) it may be possible to adapt our proof techniques to cover the infinite time horizon regime. As noted before, there is considerable literature on interacting systems focussed on long-term behaviour and phase-transitions. Such studies are often model-specific and involve challenging questions outside the scope of this paper.

**Examples of admissible score functions.** We list some relevant choices of the function  $h$  used to define the scores  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = h(M(x, t)_{t \in [0, t_0]})$  and the functional  $H_n^\xi = \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  at (10.6). These choices of  $h$  yield the following summary statistics, which figure prominently in the upcoming applications.

S1. Let  $T \subset [0, t_0]$  be a finite set. Put  $h(M(x, t)_{t \in [0, t_0]}) := \mathbf{1}(M(x, t) \in M_0, \forall t \in T)$  for some measurable subset  $M_0 \subset \mathbb{M}$ , giving that  $H_n^\xi$  is the total number of sites in  $\mathcal{P} \cap W_n$  which are

in state  $M_0$  at (discrete) times in  $T$ . This can be extended to Borel subsets  $T \subset [0, t_0]$  by taking  $h(M(x, t)_{t \in [0, t_0]}) := \exp\{\int_T \log \mathbf{1}(M(x, t) \in M_0) dt\}$ .

S2 Put  $h(M(x, t)_{t \in [0, t_0]}) := \int_0^{t_0} \mathbf{1}(M(x, t) \in M_0) dt$  for some measurable subset  $M_0 \subset \mathbb{M}$ , giving that  $H_n^\xi$  is the total time the constituent sites in  $\mathcal{P} \cap W_n$  are in state  $M_0$ .

S3. Given  $t_{\min} \in [0, t_0]$  and a measurable subset  $M_0 \subset \mathbb{M}$ , put  $h(M(x, t)_{t \in [0, t_0]}) := \mathbf{1}(\int_0^{t_0} \mathbf{1}(M(x, t) \in M_0) dt \geq t_{\min})$  giving that  $H_n^\xi$  is the total number of sites in  $\mathcal{P} \cap W_n$  which are in state  $M_0$  for at least time  $t_{\min}$ .

S4. When  $\mathbb{M} = \mathbb{R}$ , put  $h(M(x, t)_{t \in [0, t_0]}) := M(x, t_0)$ , giving that  $H_n^\xi$  is the total sum of all states at time  $t_0$ .

S5. Put  $h(M(x, t)_{t \in [0, t_0]}) := h'(M(x, t_1), \dots, M(x, t_m))$ , where  $t_1, \dots, t_m \in [0, t_0]$  are fixed and  $h' : \mathbb{M}^m \rightarrow \mathbb{R}$  is a measurable function. This yields the study of finite-dimensional distributions of the process  $(M(x, t))_{t \in [0, t_0]}$ .

Note that  $\xi$  in S1–S3 trivially satisfy the moment condition (10.8) and, in case of S4 and S5, we refer to Sections 2.1 and 2.2 of [21] for conditions on the point process  $\mathcal{P}$  insuring that  $\xi$  satisfy the requisite moment conditions.

We illustrate how measurability may be verified in the above examples. Consider the example of score function S1. Remember, the history of state modifications at site  $x$  is formalized as a counting measure  $M(x, t)_{t \in [0, t_0]} = \sum_j \delta_{(t_j, M(x, t_j))} \in \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$  where  $t_j$ 's are the update times of  $M(x, \cdot)$  and are ordered in increasing order. Observe that this increasing ordering is a measurable function of  $M$  from  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$  to  $\hat{\mathcal{N}}_{[0, t_0]}$ . With this representation, for any  $t \in [0, t_0]$  and  $M_0 \subset \mathbb{M}$ , we may write

$$\mathbf{1}(M(x, t) \in M_0) = \sum_j \mathbf{1}(t_j \leq t < t_{j+1}) \mathbf{1}(M(x, t_j) \in M_0),$$

and note that the sums and indicators are also compositions of measurable real-valued functions from  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$ . Thus  $\mathbf{1}(M(x, t) \in M_0)$  is a measurable real-valued function from  $\hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}$  and hence the finite product as in score function S1 for finite  $T$  is also a measurable function. Similarly the other score functions can be checked to be measurable.

**Markovian update rules.** The evolution is said to be *Markovian* if the update rule  $\Phi$  at (10.2) depends only on the configuration at the time of update, i.e., it does not depend on the entire history but only the current state at the site and its neighbors. To be more precise, for any marked clock event  $(T, L) \in \tilde{\tau}_x$  of some site  $x \in \mathcal{P}_n$ , we set  $M(x, T-) \in \mathbb{M}$  to be the value of the mark of the last event in  $M(x, \cdot)$  strictly before time  $T$ , and hence  $M(x, T-)$  is a measurable function of  $(M(x, t))_{t < T}$ . If the update rule  $\Phi$  in (10.2) has the form

$$\Phi \left( x, L, \left\{ \left( y - x, \{(t, M(y, t))\}_{t < T} \right) \right\}_{\{y \in \mathcal{P}_n : y \simeq x\}} \right) = \tilde{\Phi} \left( x, L, \left\{ (y - x, M(y, T-)) \right\}_{\{y \in \mathcal{P}_n : y \simeq x\}} \right) \quad (10.25)$$

$$= \left\{ \left( y - x, (T, M(x, T)) \right) \right\} \in \hat{\mathcal{N}}_{\mathbb{R}^d \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}}$$

for some function  $\tilde{\Phi} : \mathbb{R}^d \times \mathbb{L} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}} \rightarrow \hat{\mathcal{N}}_{\mathbb{R}^d \times \hat{\mathcal{N}}_{[0, t_0] \times \mathbb{M}}}$  which is jointly measurable with respect to the product  $\sigma$ -algebra, then we say that both the rule  $\tilde{\Phi}$  and the interacting particle system are *Markovian*. The specification of the update rule corresponds to the random mapping representation of a Markov process. We do not specify the generator as it is not necessary for our analysis.

In the coming subsections, we shall explicitly specify update rules satisfying condition (10.25). These extend some well-known Markovian models of interacting particle systems by allowing them to have dependent initial states, unbounded admissible interaction neighborhoods, and continuum particle locations given by admissible input  $\mathcal{P}$ , and hence the locations may be spatially correlated. We will also show that Theorem 10.1 immediately establishes Gaussian fluctuations for statistics of these models described by admissible score functions of the type S1–S5, including univariate and multivariate central limit theorems given by Theorem 5.2 and Theorem 5.4. To show that Theorem 10.1 is indeed applicable, *we only need to show that the update rules  $\tilde{\Phi}$  for these models are admissible, the only missing ingredient*.

Our first three models do not involve synchronous updating and thus do not utilise the full generality of our update rule (10.25). On the other hand, the last three models in Section 10.3.4 and Section 10.3.5 involve synchronous updating in an essential way.

### 10.3.1 Cooperative and random sequential adsorption (CSA and RSA)

Cooperative sequential adsorption (CSA) is a model of adsorption of particles in the continuum [141]. We may establish Gaussian fluctuations for this model for unbounded interaction ranges, addressing an open problem in [117, Section 5.2]. Informally, particles arrive at random times at sites given by an admissible point process  $\mathcal{P}$  and the probability that an arriving particle centered at  $x$  is accepted depends on the configuration of previously arrived particles in the neighborhood of  $x$  in a stabilizing interaction graph and the acceptance rule may have a stochastic component as well. The choice of unbounded interaction range is in keeping with many classical models; see e.g. [55]. We now present more formally a CSA model introduced in [114, Section 7] and which includes RSA as a special case. RSA models on the random geometric graph are of interest in statistical physics (see [132]) and in wireless networks. These models on fixed bounded degree graphs were studied in [119], but we do not require bounded degree. In case of Poisson input  $\mathcal{P}$ , a space-time variant of this model has been investigated in [67], [122], [120], [136].

Let  $p : \hat{\mathcal{N}}_{\mathbb{R}^d} \rightarrow [0, 1]$  be a measurable function. Informally, particles arrive at random times at sites on the substrate  $\mathbb{R}^d$ . The sites are given by an admissible point process  $\mathcal{P}$  and if the configurations of occupied neighbors of a site  $x$  is  $\mathcal{X}$  then the particle is placed at the site (i.e., it is adsorbed) with probability  $p(\mathcal{X} - x)$ . We consider states  $M(x, t) \in \{0, 1\}$ , with 1 denoting the presence of a particle at site  $x$  and time  $t$  and 0 denoting its absence.

Denote by  $M(\cdot, 0) = M(\cdot)$  the initial states, not necessarily independent. When a marked clock rings at  $(T, L) \in \tilde{\tau}_x$  at some site  $x \in \mathcal{P}_n$ , where  $L$  is a uniform  $[0, 1]$  random variable, we define the update rule as follows: If  $M(x, T-) = 0$  and  $\{y : M(y, T-) = 1, y \sim x\} =: N_{1,x}$  then we set

$M(x, T) = 1$  with probability  $p(N_{1,x} - x)$ , i.e., when  $L \leq p(N_{1,x} - x)$ . In all other cases we have  $M(x, T) := M(x, T-)$ , i.e., the state does not change at  $x$ . The neighbors of  $x$  are not updated. More formally, the Markovian update rule  $\tilde{\Phi}$  in (10.25) is considered with

$$M(x, T) := \mathbf{1}(M(x, T-) = 1) + \mathbf{1}(M(x, T-) = 0)\mathbf{1}(L \leq p(N_{1,x} - x)). \quad (10.26)$$

The measurability of  $\tilde{\Phi}$  follows since the indicator restricts a point process to a measurable subset of  $\mathbb{R}^d \times \{1\}$ . Thus  $\tilde{\Phi}$  is admissible.

As a corollary of Theorem 10.1, for a given  $t_0 \in (0, \infty)$ , we obtain the limit theory for  $(\mu_n^\xi)_{n \in \mathbb{N}}$  with  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = M(x, t_0)$ . In this case  $H_n^\xi$  is simply the total number of accepted particles in the window  $W_n$  under CSA dynamics up to time  $t_0$ . Under some additional assumptions, we deduce the limit theory for  $(\mu_n^\xi)_{n \in \mathbb{N}}$  when  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = M_n(x, \infty) := \lim_{t \rightarrow \infty} M_n(x, t)$ .

The following result is an immediate consequence of Theorem 10.1.

**Corollary 10.6** (Limit theory for the number of accepted particles in CSA). *Consider the CSA model with input point process  $\tilde{\mathcal{P}}$  given in (10.9), where states  $\mathbb{M} = \{0, 1\}$  and uniformly marks of clocks in  $\mathbb{L} = [0, 1]$ , as in Theorem 10.1 having summable exponential  $\mathbb{B}$ -mixing correlations, and Palm moment bounds (10.10) for all  $p, m \in \mathbb{N}$ . We assume Markovian update rules (10.25) with (10.26) on a stabilizing interaction graph  $\mathcal{G}$  on finite windows with respect to  $\mathcal{P}$  as in (10.4). Put  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = M_n(x, t_0)$  for  $t_0 \in (0, \infty)$ . Then the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (10.7) satisfy the conclusions of Theorem 10.1. When  $f \equiv 1$  this yields the asymptotic normality for the total number of accepted particles in  $W_n$  up to time  $t_0$ ,*

Choosing  $p(\mathcal{X}) = \mathbf{1}(\mathcal{X} = \emptyset)$  we obtain *the random sequential adsorption (RSA) model on a graph*. In other words, a particle is adsorbed at an empty site if there are no occupied neighboring sites. Note that the first clock ring of the Poisson process definitively fixes the state (0 or 1) of the site  $x$ , provided the clock rings before the finite time  $t_0 < \infty$ . Otherwise,  $x$  retains its initial state  $M(x) = M(x, 0)$ .

**Remark 10.7** (Further comments and comparison with previous work).

(i) (Comparison with literature for the case  $t_0 < \infty$ .) In the case of Poisson point with bounded range of interactions, Corollary 10.6 can be deduced from [117]. However the case of unbounded interactions, even in the case of Poisson input, was left open in [117]. Corollary 10.6 adds to the central limit theory of previously studied RSA models having finite time horizon ([122], [120] and references therein), which is confined to Poisson input, finite range interactions and independent initial states.

(ii) (Multivariate CLT.) More generally, given  $t_0 \in (0, \infty)$ , subsets  $T_i \subset [0, t_0]$ , scores  $\xi_i(\tilde{x}, \tilde{\mathcal{P}}_n) = \int_{T_i} \mathbf{1}(M(x, t) = 1) dt$  and assuming convergence of covariances, we find that Theorem 5.4 (along with Remarks 10.2(ii) and (vii)) establish a multivariate central limit theorem for the  $k$ -vector whose  $i$ th entry is the number of accepted particles in the time period  $T_i$ .

(iii) (Graph based RSA models with random particle sizes,  $t_0 \in (0, \infty)$ .) Let marked clocks have rings  $(T, L) \in \tilde{\tau}_x$  with  $L$  being a random particle shape, taking value in the space  $\mathcal{K}$  of nonempty, compact

subsets of  $\mathbb{R}^d$ , with diameter bounded by a fixed constant  $r$ . The update Markovian rule (10.25) for  $M(x, t) \in \mathcal{K}$  is as follows: If  $M(x, T-) = \emptyset$  and for all  $y \in \mathcal{P}_n$ ,  $y \sim x$ ,  $M(y, T-) \cap (x \oplus L) = \emptyset$ , then we set  $M(x, T) := x \oplus L$ . Otherwise, the state of  $x$  does not change and we put  $M(x, T) = M(x, T-)$ . Consider the score function  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \mathbf{1}(M(x, t_0) \neq \emptyset)$  for  $t_0 \in (0, \infty)$ . Assuming admissible  $\tilde{\mathcal{P}}$  and letting the stabilizing interaction graph be the Gilbert (geometric) graph with radius  $2r$ , we may check that all assumptions of Theorem 10.1 are satisfied, thus giving the limit theory in particular for the total number  $H_n^\xi$  of accepted particles in this model up to time  $t_0 < \infty$  in  $W_n$ . Assuming that there exists a ‘minimal’ shape  $\mathbf{L}$  such that  $\mathbb{P}(\mathbf{L} \subset L) = 1$  and  $\mathbb{P}(L = \mathbf{L}) = \epsilon > 0$ , then  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = M(x, \infty) := \lim_{t \rightarrow \infty} M(x, t)$  exists almost surely, and Theorem 10.1 gives in particular the results when  $t_0 = \infty$ . Indeed, we can use the approach as in Corollary 10.6(ii) with  $\hat{T}_x^n$  being the last modification of the history  $M(x, t)$ , admitting stochastic bound by exponential random variable of parameter  $\epsilon > 0$ , as in the proof of Corollary 10.6.

(iv) (Further extensions.) Variants of the CSA model including dimer RSA, the annihilation process, and bootstrap percolation may be analysed using the framework of this subsection; see also [119, Section 2].

### 10.3.2 Epidemic spread and voter models

In the SIR epidemic model, particles are of three types: susceptible (S), infected (I) and recovered (R). Particle locations and initial states are given by the realization of  $\tilde{\mathcal{P}}$ . Whether a susceptible particle becomes infected is a function of the number of infected neighbors. However, once infected, a particle may recover on its own, and similarly, a recovered particle may again become susceptible. We shall consider more general dynamics where the updates are based on the relative locations and states of the neighboring sites. We provide here conditions under which the number of infected individuals in the window  $W_n$  at time  $t_0$  has Gaussian fluctuations. We adopt a Markovian set-up, but we could also consider a non-Markovian example, whereby updates are based on the entire time-evolved history of neighbors.

Formally, set  $\mathbb{M} = \{S, I, R\}$ . Let  $p : (\mathbb{M} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}}) \times \mathbb{M} \rightarrow [0, 1]$  be a jointly measurable probability kernel governing the transition from S, I, R to S, I, R and depending on the configuration of neighbors, i.e., for  $\tilde{\mu} \in \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}}$ ,  $u \in \mathbb{M}$ ,  $p(u, \tilde{\mu}, \cdot)$  is a probability distribution function on  $\mathbb{M}$ . As before,  $L_i, i \geq 0$ , are i.i.d. uniform random variables with values in  $\mathbb{L} = [0, 1]$ . Let the initial states be given by  $M(x, 0)$ ,  $x \in \mathcal{P}$ , which are not assumed to be independent.

Set

$$p_S := p(M(x, t-), \{(y - x, M(y, t-))\}_{y \sim x}, S),$$

$$p_I := p(M(y, t-), \{(y - x, M(y, t-))\}_{y \sim x}, I)$$

and put  $p_R = 1 - p_S - p_I$ . Define the function  $\tilde{\Phi}$  in (10.25) as follows:

$$M(x, t) := \mathbf{S1}(L \leq p_S) + \mathbf{I1}(p_S < L \leq p_S + p_I) + \mathbf{R1}(L > p_S + p_I), \quad (10.27)$$

where the sum is to taken as a formal sum. Only one of the indicators is non-zero and hence  $\tilde{\Phi}$  is well-defined.

Put  $\xi^{(1)}(\tilde{x}, \tilde{\mathcal{P}}_n) = \mathbf{1}(M(x, t_0) = \text{I})$  so that  $H_n^{\xi^{(1)}}$  is the total number of infected individuals at time  $t_0$ . Letting  $\xi^{(2)}(\tilde{x}, \tilde{\mathcal{P}}_n) = \int_0^{t_0} \mathbf{1}(M(x, t) = \text{I}) dt$  gives that  $H_n^{\xi^{(2)}}$  is the total time all individuals are infected. To deduce the limit theory for these statistics we apply directly Theorem 10.1.

**Corollary 10.8** (Gaussian fluctuations for the number of infected individuals and for the total infection time). *Consider the SIR epidemic model with input point process  $\tilde{\mathcal{P}}$  given in (10.9), with states  $\mathbb{M} = \{\text{S}, \text{I}, \text{R}\}$  and clocks having uniformly distributed marks in  $\mathbb{L} = [0, 1]$ . As in Theorem 10.1, we assume  $\tilde{\mathcal{P}}$  has summable exponential  $\mathbb{B}$ -mixing correlations and Palm moment bounds (10.10) for all  $p, m \in \mathbb{N}$ . We assume the Markovian update rules (10.25) along with (10.27) on a stabilizing interaction graph  $\mathcal{G}$  on finite windows with respect to  $\mathcal{P}$  as in (10.4). Then the measures  $(\mu_n^{\xi^{(i)}})_{n \in \mathbb{N}}, i = 1, 2$ , at (10.7) induced by the score functions  $\xi^{(i)}, i = 1, 2$  above, respectively satisfy the conclusion of Theorem 10.1. When  $f \equiv 1$  this yields the asymptotic normality for the total number of infected individuals at time  $t_0$  as well as the total time the individuals are infected during the period  $[0, t_0]$ .*

This corollary extends upon Section 4.5 of [117], which is constrained to models having finite range and independent initial states. The models described here arise in the study of epidemics and also in the propagation of viruses in networks. Suitably choosing the probability kernel yields classical models including the Richardson model, the contact process, the voter model, and the chase-escape model (see [79, Chapter 8] and [47]). The more general kinetically constrained models which include bootstrap percolation and its stochastic counterpart given by the Fredrickson-Andersen model (see [70, 69]), can also be handled by the above framework.

### 10.3.3 Majority dynamics in the continuum

In this model, we again denote the sites by  $\mathcal{P}$ , and now the state at each site  $x \in \mathcal{P}$  is  $\{+1, -1\}$ -valued and at each clock ring, the state is updated to that of the majority of its neighbors. In case of a tie, the state is unchanged. More formally, we take  $\mathbb{M} = \{-1, +1\}$  (and in general our setting, the initial states  $M(x), x \in \mathcal{P}$ , and they are not necessarily independent). When the clock at site  $x$  rings at time  $t$ , it undergoes a Markovian update as follows : If  $\sum_{y \sim x} M(y, t-) \neq 0$ , then  $M(x, t) = \text{sgn}(\sum_{y \sim x} M(y, t-))$  where  $\text{sgn}$  is the sign function. Otherwise we put  $M(x, t) = M(x, t-)$ . Formally, the update rule  $\tilde{\Phi}$  in (10.25) is defined as follows :

$$M(x, t) := \text{sgn}\left(\sum_{y \sim x} M(y, t-)\right) \mathbf{1}\left(\sum_{y \sim x} M(y, t-) \neq 0\right) + M(x, t-) \mathbf{1}\left(\sum_{y \sim x} M(y, t-) = 0\right). \quad (10.28)$$

The measurability of  $\tilde{\Phi}$  here can also be verified by writing it as composition of projections, sums, and the sign function. Thus  $\tilde{\Phi}$  is admissible.

Put  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \mathbf{1}(M(x, t_0) \in 1)$  so that  $H_n^\xi$  is the total number of sites in state 1 at time  $t_0$ . Theorem 10.1 immediately gives the following corollary.

**Corollary 10.9** (Gaussian fluctuations for the size of the majority/minority). *Consider the majority/minority model with input point process  $\tilde{\mathcal{P}}$  given in (10.9), with states  $\mathbb{M} = \{-1, +1\}$  and clocks*

with no marks. As in Theorem 10.1, we assume  $\tilde{\mathcal{P}}$  having summable exponential  $\mathbb{B}$ -mixing correlations, and Palm moment bounds (10.10) for all  $p, m \in \mathbb{N}$ . We assume Markovian update rules (10.25) with (10.28) on a stabilizing interaction graph  $\mathcal{G}$  on finite windows with respect to  $\mathcal{P}$  as in (10.4). Then the measure  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (10.7) induced by the score functions  $\xi$  above satisfies the conclusions of Theorem 10.1. When  $f \equiv 1$  this yields the asymptotic normality for the total number of sites in state 1 at time  $t_0$ .

One could also consider this model of majority dynamics in discrete time with synchronous updates; see Section 10.4. These results add to the central limit theorem for majority dynamics in a dense Erdős-Rényi random graph (cf. [12, Theorem 1]). More general opinion dynamics such as spatial graph versions of *edge-averaging dynamics* or  $\ell^p$ -minimization dynamics can also be considered; see for example [2, 3].

### 10.3.4 Ballistic deposition

We sketch the model and refer the reader to [122] and [118] a discussion of similar models. Particles rain down at sites given by the realization of a point process  $\mathcal{P}$  and they stick to the first point of contact. We assume that the particles are random shapes with random heights. This is accounted by the clock marks  $L = (L^{(1)}, L^{(2)}) \in [0, \infty) \times \mathcal{K}$ , where  $L^{(1)}$  denotes the height of the particle (arriving at some clock event time  $T_i$  of some site  $x$ ), and  $L^{(2)}$  is a shape variable—a random element of the space of compact sets  $\mathcal{K}$  of  $\mathbb{R}^d$ —assumed to be nonempty, contain the origin  $\mathbf{0}$ . As a technical assumption, we suppose that the height  $L^{(1)}$  is possibly unbounded but has all moments finite, while the diameter of  $L^{(2)}$  (the particle shapes in  $\mathcal{K}$ ) is bounded by a fixed finite constant  $r > 0$ , and  $L^{(1)}$ ,  $L^{(2)}$  may be mutually dependent.

We now formally define a specific model. As the state space, we consider  $\mathbb{M} := \mathbb{R} \times \mathcal{K}$ , where each state  $M = (M_1, M_2) \in \mathbb{M}$  represents a particle  $M_2$  positioned at level  $M_1$ , with  $M_1$  indicating the height at which  $M_2$  is aligned at the top. The assumption that the particle diameter is bounded by a given  $r > 0$  allows one (without loss of generality regarding the principle of ‘point of contact’) to consider the Gilbert graph on  $\mathcal{P}_n$ ,  $\mathcal{G}(\mathcal{P}_n)$ , where  $x \sim y$  if and only if  $|x - y| < 2r$ . With this specification, when a clock rings at time  $t$  at a site  $x$  with a mark  $L = (L^{(1)}, L^{(2)})$ , we update  $M(x, t)$  and  $M(y, t)$  for  $y \sim x$  by the function  $\tilde{\Phi}$  in (10.25) with

$$M_1(x, t) := L^{(1)} + \max_{y: y \sim x, (y \oplus M_2(y, t-)) \cap (x \oplus L^{(2)}) \neq \emptyset} M_1(y, t-), \quad M_2(x, t) := L^{(2)}, \quad (10.29)$$

$$M_1(y, t) := M_1(y, t-), \quad M_2(y, t) := \left( (y \oplus M_2(y, t-)) \setminus (x \oplus L^{(2)}) \right) \oplus (-y). \quad (10.30)$$

Note, the neighbors  $y \sim x$  keep their top level  $M_1(y, t-)$  unchanged while their particles  $M_2(y, t-)$  are eroded by the shadow of the particle  $L^{(2)}$  deposited at  $x$ . Note, the *height function*

$$H(y, t) := \max_{x \in \mathcal{P}_n: x \oplus M_2(x, t) \ni y} M_1(x, t), \quad y \in \mathbb{R}^d,$$

represents the growth level of the ( $d'$ -dimensional hyper-)surface in this ballistic deposition model.

Putting  $\xi^{(1)}(\tilde{x}, \mathcal{P}_n) = M_1(x, t_0)|M_2(x, t_0)|$  gives that  $H_n^{\xi^{(1)}} = \sum_{x \in \mathcal{P}_n} \xi^{(1)}(\tilde{x}, \mathcal{P}_n)$  is the total volume under the height function  $H$  at time  $t_0$ . Alternatively, putting  $\xi^{(2)}(\tilde{x}, \mathcal{P}_n) = |M_2(x, t_0)| \mathbf{1}(M_1(x, t_0) \in [h_{\min}, \infty))$ , we obtain that  $H_n^{\xi^{(2)}}$  is the total surface area visible at height at least  $h_{\min}$  at time  $t_0$ .

**Corollary 10.10** (Gaussian fluctuations for the volume under the height function and the surface area at a given height). *Consider the ballistic deposition model with input point process  $\tilde{\mathcal{P}}$  given in (10.9), with initial states  $M(x) = (0, \emptyset)$  and the clocks as described above, with the particles' diameter satisfying  $\text{diam}(L^{(2)}) < r$  for some  $r < \infty$ , and with the height having moments  $\mathbb{E}(L^{(1)})^p < \infty$  for all  $p > 1$ . We assume  $\mathcal{P}$  has summable exponential mixing correlations and Palm moment bounds (10.10) for all  $p, m \in \mathbb{N}$ . We assume Markovian update rules (10.25) along with (10.29)–(10.30) (assuming  $x \sim y$  iff  $|x - y| \leq 2r$ ). Then the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (10.7) induced by the score functions  $\xi = \xi^{(i)}$ ,  $i = 1, 2$  above satisfy the conclusions of Theorem 10.1. When  $f \equiv 1$ , this yields the asymptotic normality for the total volume  $H_n^{\xi^{(1)}}$  under the height function and the total surface  $H_n^{\xi^{(2)}}$  visible above some given height  $h_{\min}$ .*

*Proof.* This result follows from Theorem 10.1, provided the considered score functions  $\xi^{(i)}$  satisfy moment conditions (10.8) for all  $p > 1$ . Indeed, the other assumptions are met: the Gilbert graph stabilizes— $x \sim y$  iff  $|x - y| \leq 2r$ —for any input, and the initial conditions  $M(x)$  together with the marked clocks are i.i.d., thus ensuring summable exponential  $\mathbb{B}$ -mixing correlations for  $\tilde{\mathcal{P}}$ .

We now prove the moment conditions for  $M_1(x, t_0) =: M_1(x, \tilde{\mathcal{P}}_n)$ , which suffices for those of  $\xi^{(i)}(\tilde{x}, \tilde{\mathcal{P}}_n)$ , since the surface of the particles  $|M_2(x, t)|$  is bounded for all  $t \in [0, t_0]$ . This the latter assumption is equivalent to showing that the probability  $\mathbb{P}_{[\tilde{x}_1]}(M_1(x_1, \tilde{\mathcal{P}}_n) > m)$  is fast decreasing to 0 as a function of  $m \rightarrow \infty$ .

For proving this, observe that

$$M_1(x_1, \tilde{\mathcal{P}}_n) \leq \sum_{y: |y-x_1| \leq D(\tilde{x}_1, \tilde{\mathcal{P}}_n)} \sum_{T_i(y) \in \hat{\tau}_y} L_i^{(1)}(y), \quad (10.31)$$

that is, the height function of  $x_1$  at time  $t_0$  is bounded by the sum of all particle heights  $L_i^{(1)}(y)$  corresponding to arrivals at the clock times  $T_i(y)$  during the interval  $(0, t_0]$ , over all sites  $y$  in the backward cluster of  $x_1$ , which are within distance  $D(\tilde{x}_1, \tilde{\mathcal{P}}_n)$  as defined in (10.15).

Also note that  $L_i^{(1)}(y), \hat{\tau}_y$ 's are i.i.d. for  $y \in \mathcal{P}_n$  given  $\mathcal{P}$ . In particular for all  $y \in \mathcal{P}_n$ ,  $\sum_{T_i(y) \in \hat{\tau}_y} L_i^{(1)}(y)$  has distribution of  $\sum_{j=1}^{N_0} L_j^{(1)}$ , where  $L_j^{(1)}$  are i.i.d. random variables distributed as  $L^{(1)}$  and  $N_0$  is an independent Poisson random variable with mean  $t_0$ . Thus, from the assumed finiteness of moments of  $L^{(1)}$  and tail property of  $N_0$ , we have the fast decay of the tail of the above sum i.e.,

$$\mathbb{P}\left(\sum_{j=1}^{N_0} L_j^{(1)} \geq m\right) \text{ decays fast in } m \text{ as } m \rightarrow \infty.$$

Now using the above fast decay, Lemma 10.5, and the fast decay of  $D(\tilde{x}_1, \tilde{\mathcal{P}}_n)$  (see Step 2 of the proof of the Theorem 10.1; in particular (10.16), (10.17) and (10.20)), we deduce that for any  $\alpha > 0$ , there

exists a fast decreasing function  $\tilde{\varphi}_p$  such that

$$\mathbb{P}_{[x_1]^p} \left( \max_{y \in \mathcal{P}_n: |y-x_1| \leq D(\tilde{x}_1, \tilde{\mathcal{P}}_n)} \sum_{T_i(y) \in \hat{\tau}_y} L_i^{(1)}(y) > m^\alpha \right) \leq \tilde{\varphi}_p(m). \quad (10.32)$$

Denote by  $N_1 = |\{|y - x_1| \leq D(\tilde{x}_1, \tilde{\mathcal{P}}_n)\}|$  the number of sites under the maximum. We fix some  $\alpha \in (0, 1)$  and  $\beta = (1 - \alpha)/(d + 1)$ . If the above fast decay event does not happen and  $M_1(x_1, \tilde{\mathcal{P}}_n) > m$ , then we must have  $N_1 > m^{1-\alpha}$ . Combining these observations we have

$$\begin{aligned} & \mathbb{P}_{[x_1]^p}(M_1(x_1, \tilde{\mathcal{P}}_n) > m) \\ & \leq \mathbb{P}_{[x_1]^p} \left( \sum_{T_i(y) \in \hat{\tau}_y} L_i^{(1)}(y) > m^\alpha \text{ for some } y \in \mathcal{P}_n \text{ with } |y - x_1| \leq D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \right) \\ & \quad + \mathbb{P}_{[x_1]^p}(N_1 > m^{1-\alpha}) \\ & \leq \tilde{\varphi}_p(m) + \mathbb{P}_{[x_1]^p}(N_1 > m^{1-\alpha}, D(\tilde{x}_1, \tilde{\mathcal{P}}_n) > m^\beta) + \mathbb{P}_{[x_1]^p}(N_1 > m^{1-\alpha}, D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \leq m^\beta) \\ & \leq \tilde{\varphi}_p(m) + \mathbb{P}_{[x_1]^p}(D(\tilde{x}_1, \tilde{\mathcal{P}}_n) > m^\beta) + \mathbb{P}_{[x_1]^p}(\mathcal{P}(B_{m^\beta}(x_1)) \geq m^{1-\alpha}), \end{aligned}$$

where in the penultimate inequality we have divided into cases whether  $D(\tilde{x}_1, \tilde{\mathcal{P}}_n) > m^\beta$  or not and in the last inequality we have used the simple fact that if  $D(\tilde{x}_1, \tilde{\mathcal{P}}_n) \leq m^\beta$  and  $N_1 > m^{1-\alpha}$ , then  $B_{m^\beta}(x_1)$  must contain at least  $m^{1-\alpha}$  points from  $\mathcal{P}$ .

It remains to justify the fast decay of second and third terms in the right-hand side of the last inequality. The fast decay of the second term again follows from Step 2 of the proof of the Theorem 10.1; see (10.16), (10.17) and (10.20). The fast decay of the third term follows from the deviation inequality in Lemma 10.3.  $\square$

This result extends the central limit theorem which could be deduced from the methods of [117, Section 4.4] which requires the stronger assumptions that the interaction ranges are bounded, the initial states are i.i.d., and the input  $\mathcal{P}$  is Poisson. One could likewise modify the score  $\xi$  to obtain Gaussian fluctuations for other statistics, including the total number of contacts, further extending results of [117, Section 4.4] to non-Poisson input and unbounded interaction ranges.

The above model is but a simple prototype which admits extensions and variants. A further generalization, amenable to our analysis, involves including queueing dynamics whereby the particles start slowly diminishing in height and vanish [6].

### 10.3.5 Further models

We describe two additional models which may be treated by Theorem 10.1 and Remark 10.2(ii). For examples of further update rules, and thus additional examples of spatial interacting particle systems satisfying our framework, we refer to [117, Section 4], [79, Chapter 8], [47, Chapters 8 and 9], [147, Chapter 1], [103], [56] and [69]. Section 12 describes additional models falling within our framework but which require separate investigation.

(i) **Exclusion processes.** In this model, within the framework of the update rules at (10.25), some sites are initially occupied by particles which attempt to hop at independent times according to an

exclusion rule. In the simplest case, when the clock at an occupied site  $x$  rings, the particle at  $x$  chooses a neighboring site at random and jumps there if it is unoccupied, otherwise it remains at  $x$ . Hence, each site contains at most one particle, and the total number of particles is conserved. Neighboring sites need not lie within a bounded interaction range, so the model goes beyond of [117, example (c) in Section 4.5]. Moreover, the sites are not assumed to form a homogeneous Poisson process, but only a realization of an admissible point process under the assumptions of Theorem 10.1, and the initial states need not be independent. This prototype can be extended to allow multiple particles per site and more general jump rules, covering, for instance, some of the birth–death–migration models in [37] and [117, Section 4.1].

To handle this simple exclusion process, we proceed as follows. We set  $\mathbb{M} = \{0, 1\}$  and let the  $L$ 's be i.i.d. uniform random variables on  $[0, 1]$ . Consider a jointly measurable probability kernel  $p : \hat{\mathcal{N}}_{\mathbb{R}^d} \setminus \{\emptyset\} \times \mathbb{R}^d \rightarrow [0, 1]$  which selects a site in  $\mu$  based on its location and assume that  $p(\mu, \cdot)$  is supported on  $\mu$  for all  $\mu \in \hat{\mathcal{N}}_{\mathbb{R}^d} \setminus \{\emptyset\}$ , i.e.,  $p(\mu, \cdot)$  gives a probability distribution on  $\mu$ . For example, we may choose  $p(\mu, y)$  to be proportional to  $|y|$  for  $y \in \mu$ .

Given  $\mu$ , we define  $X'(\mu)$  to be the random element in  $\mu$  with distribution  $p(\mu, \cdot)$ . For convenience, set  $X' := X'(\{y - x : y \simeq x\}) = X'(\{y - x : y \sim x\} \cup \{\mathbf{0}\})$ . Define  $\tilde{\Phi}$  in (10.25) according to the rule for  $x$  and  $y \sim x$ :

$$\begin{aligned} M(x, t) &:= M(x, t-)M(x + X', t-), \\ M(y, t) &:= \left( M(x, t-)(1 - M(x + X', t-)) + (1 - M(x, t-))M(x + X', t-) \right) \mathbf{1}(x + X' = y) \\ &\quad + M(y, t-) \mathbf{1}(x + X' \neq y), \end{aligned}$$

where  $L$  is used to choose the random element  $X'$  (see a similar random mapping representation as in Section 10.3.2).

Choosing a score function related to the occupations  $M(x, t)$  at  $t \in [0, t_0]$  of the site  $x$  (see examples given at the beginning of Section 10.3) we obtain the asymptotic normality for the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (10.7) leveraging Theorem 10.1.

(ii) **Divisible sandpile dynamics.** In the classical Abelian sandpile model (see, e.g., [81]), each site contains an integer number of grains. When the number of grains at a site exceeds a fixed threshold, the site becomes unstable and topples, sending one grain to each of its finitely many neighbors. Such topplings may trigger further instabilities, leading to a cascade of updates known as an avalanche, which continues until the system reaches a stable configuration.

Within the framework of (10.25), we assume that at time  $t = 0$  each site carries an initial amount of mass, and at independent exponential times the site topples part of this mass to its neighbors. Importantly, the topplings are not followed by the complete avalanche mechanism; that is, we do not enforce further updates until stability is restored, but rather there is only a single, local redistribution at the ringing site. However, unlike the classical sandpile, the set of neighbors updated in a toppling need not lie within a fixed deterministic distance.

We put  $\mathbb{M} = \mathbb{R}$  and we interpret  $M(x, t) \in \mathbb{R}$  as the amount of (signed) mass at site  $x$ , with  $x \in \mathcal{P}$ . The initial masses are given by  $M(x, 0)$ , not necessarily independent. When the clock at  $x$  rings at time  $t \in (0, t_0]$ , we update  $\tilde{\Phi}$  in (10.25) according to the rule for  $x$  and  $y \sim x$ :

$$\begin{aligned} M(x, t) &:= M(x, t-)\mathbf{1}(M(x, t-) < 1) + \mathbf{1}(M(x, t-) \geq 1), \\ M(y, t) &:= M(y, t-) + \frac{M(x, t-) - 1}{|N_x|}\mathbf{1}(M(x, t-) \geq 1), \end{aligned}$$

where  $N_x := \{x : y \sim x\}$ . In other words,  $x$  retains unit mass for itself and distributes the remaining mass equally among its neighbors. If  $M(x, t-) < 1$ , then there is no update at  $x$  or its neighbors.

For score functions  $\xi$  related to the mass evolution  $M(x, t)$  at  $t \in [0, t_0]$  of the site  $x$  (see examples given at the beginning of Section 10.3) we obtain the asymptotic normality for the measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  at (10.7) leveraging Theorem 10.1. In particular, if we set

$$\xi(\tilde{x}, \tilde{\mathcal{P}}_n) = \sum_{t \in \tau_x} [M(x, t-) - 1] \cdot \mathbf{1}[M(x, t-) \geq 1],$$

then  $H_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  yields the ‘odometer’ up to time  $t_0$ , i.e., the total amount of mass distributed to neighbors until time  $t_0$ .

The above model, where excess mass is proportionally dispatched to neighbors, is not Abelian since the order of topplings affects the mass distribution. Our framework, however, also accommodates more classical Abelian dynamics and variants such as the chip-firing game, the dollar game as well as more general Abelian networks; see [25, 58, 81, 87, 35].

## 10.4 Discrete-time interacting particle systems

We now consider discrete-time particle systems. In many models such as probabilistic cellular automata or kinetically constrained models (see for example [103, 69]), it is natural to consider globally synchronous dynamics and these are modelled better by discrete-time particle systems. We shall now modify our framework for the same and then state and prove our limit theorems for discrete-time particle systems.

Let  $\mathbb{M}$  and  $\mathbb{L}$  be Polish spaces as before, hosting respectively the states of the sites and the marks of the clocks (see Section 10.1.1). In contrast to the previous setting with individual (independent Poisson) clocks  $\tau_x$ ,  $x \in \mathcal{P}$ , we now replace them by a universal clock ringing at times  $1, 2, \dots, J$  for some finite  $J < \infty$ , simultaneously for all  $x \in \mathcal{P}$ . That is, the input process  $\tilde{\mathcal{P}} = \{(x, U(x))\}_{x \in \mathcal{P}} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  is defined by the ground process  $\mathcal{P}$  (specifying the locations of the sites), equipped with marks  $U(x) = (M(x), \mathbf{L}_x) \in \mathbb{K}$  taking values in the Polish space  $\mathbb{K} := \mathbb{M} \times \mathbb{L}^J$ . Here, as before,  $M(x) = M(x, 0) \in \mathbb{M}$  denotes the initial state of site  $x$ , while

$$\mathbf{L}_x = (L_1(x), \dots, L_J(x)) \in \mathbb{L}^J$$

is an individual vector of marks (related to the universal clock  $i \in [J] = \{1, \dots, J\}$ ), where  $L_i(x)$ ,  $i \in [J]$ ,  $x \in \mathcal{P}$  are i.i.d. (conditionally on  $\{(x, M(x))\}_{x \in \mathcal{P}}$ )  $\mathbb{L}$ -valued random elements. For simplicity

of notation, we often suppress the dependence on  $x$  when writing  $L_i$ . We emphasize again that we do not assume  $\mathcal{P}$  to be a Poisson point process, nor do we assume independence of the initial states  $\{M(x)\}_{x \in \mathcal{P}}$ .

We assume the framework of stabilizing interaction graphs from Section 10.1.3. For the update rules and state evolution, we refer to Section 10.1.2, modifying it as follows: At each universal clock tick  $j \in [J]$ , the state at all sites  $x \in \mathcal{P}_n$  is updated by adding event  $M(x, j)$  to the previous history

$$\mathbf{M}_{j-1}(x) = \mathbf{M}_{j-1}(x; n) := (M(x, 0), M(x, 1), \dots, M(x, j-1)) \in \mathbb{M}^j.$$

The update is performed via the function  $\Phi_j$ , which depends on the clock mark  $L_j(x)$  and the histories of  $x$  and its neighbors  $y \sim x$ , possibly accounting for their relative positions. More formally, for each  $j \in [J]$ , we assume the existence of a measurable updating function  $\Phi_j : \mathbb{R}^d \times \mathbb{L} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}^j} \rightarrow \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}^{j+1}}$  and put

$$M(x, j) := \Phi_j \left( x, L, \left\{ (y - x, \mathbf{M}_{j-1}(y)) \right\}_{\{y \in \mathcal{P}_n : y \sim x\}} \right). \quad (10.33)$$

The complete history of a given site  $x \in \mathcal{P}_n$ —resulting from the execution of the functions  $\Phi_j$  at all clock rings  $j \in [J]$ —is denoted by  $M_n[x] = (M(x, 0), \dots, M(x, J)) \in \mathbb{M}^{J+1}$ . These histories are treated as an additional marking of the point  $x \in \mathcal{P}_n$ . Note, this marking is measurable as a measurable mapping

$$g : (\mathbb{R}^d \times \mathbb{K}) \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \ni (\tilde{x}, \tilde{\mu}) \mapsto M[x] \in \mathbb{M}^{J+1}, \quad (10.34)$$

constructed iteratively using the measurable functions  $\Phi_j$  described above. This mapping describes the history of marks  $M_n[x]$ ,  $x \in \mathcal{P}_n$ , when applied simultaneously on all  $x \in \mathcal{P}_n$ .

Note that, unlike the continuous-time model of Section 10.1, the states at sites  $x \in \mathcal{P}$  are updated at globally synchronous times  $t = 1, \dots, J$ . However, each site  $x$  updates only its own state, using the previous histories of its neighbors.

Put  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) := h(M_n[x])$ , where  $h : \mathbb{M}^{J+1} \rightarrow \mathbb{R}$  is a measurable function. We now state our central limit theorem for linear statistics of  $(\mu_n^\xi)_{n \in \mathbb{N}}$  in (10.7), similar to Theorem 10.7, and then we shall show how certain Markovian interacting particle systems (including all of those in Section 10.3) fall within this framework.

**Theorem 10.11** (CLT for discrete-time interacting particle systems). *Let the input process for the discrete-time interacting particle system be  $\tilde{\mathcal{P}} = \{(x, M(x), \mathbf{L}_x) \in \mathbb{R}^d \times \mathbb{M} \times \mathbb{L}^J\}$ , where  $J \in \mathbb{N}$  and  $\mathbf{L}_x$  are vectors of i.i.d. elements of  $\mathbb{L}$ , independent of  $\sum_x \delta_{(x, M(x))}$ . Put*

$$\mu_n^\xi := \sum_{x \in \mathcal{P} \cap W_n} \xi(\tilde{x}, \tilde{\mathcal{P}}_n) \delta_{n^{-1/d}x}. \quad (10.35)$$

*If  $\tilde{\mathcal{P}}$ ,  $\{\Phi_j\}_{j=1}^J$ , and  $\xi$  satisfy the assumptions (i)-(iii) in Theorem 10.1, where the  $\Phi_j$  are at (10.33), then the random measures  $(\mu_n^\xi)_{n \in \mathbb{N}}$  satisfy the central limit theorem. That is, for every test function  $f \in \mathbb{B}(W_1)$  such that  $\text{Var}[\mu_n^\xi(f)] = \Omega(n^\nu)$  for some  $\nu > 0$ , we have*

$$(\text{Var}[\mu_n^\xi(f)])^{-1/2} (\mu_n^\xi(f) - \mathbb{E}[\mu_n^\xi(f)]) \xrightarrow{d} Z,$$

*where  $Z$  is a standard Gaussian random variable.*

**Remark 10.12** (Further comments and comparison with previous work).

(i) By imposing additional assumptions of translation-invariance on score-functions, stationarity on  $\tilde{\mathcal{P}}$ , and a uniform bound for the interaction ranges across all windows for every  $x \in \mathcal{P}$  one may obtain expectation and variance asymptotics for  $(\mu_n^\xi)_{n \in \mathbb{N}}$  as in Remark 10.2(ii).

(ii) When  $\mathcal{P}$  is Poisson and when the states  $M(x), x \in \mathcal{P}$ , are independent, then this result follows from the central limit theorems of [115] and [93]. We require neither Poisson input nor independent initial states.

(iii) (Markovian interacting particle systems) Let  $M(x) \in \mathbb{M}$ . Let the updates at discrete times  $j \in [J]$  be as follows

$$M(x, j) := \Phi \left( x, L, \left\{ (y - x, M(y, j - 1)) \right\}_{\{y \in \mathcal{P}_n : y \simeq x\}} \right).$$

where  $\Phi : \mathbb{R}^d \times \mathbb{L} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{M}} \rightarrow \mathbb{M}$  is a measurable function. Informally, at each step, a vertex is updated according to its own state, the states of its neighbors and also the location of its neighbors.

(iv) Update rules defined in Sections 10.3.2–10.3.3 can be trivially adapted in the above Markovian framework but with the caveat that the evolution of the particle systems can be considerably different. For example, in the RSA model in discrete time, it is possible that neighboring sites can become occupied together.

(v) Examples in Sections 10.3.4 and 10.3.5 need to be modified suitably to fit into the above framework as the update rules therein are more general than those in Sections 10.3.2–10.3.3. For example in the divisible sandpile dynamics, there could be an exchange of mass as every site will re-distribute its excess mass in parallel. More specifically, we may define the updates as follows. For  $j = 1, \dots, J$

$$\begin{aligned} M(x, j) &= M(x, j - 1) \mathbf{1}(M(x, j - 1) < 1) + \mathbf{1}(M(x, t_{j-1}) \geq 1) \\ &\quad + \sum_{y: y \sim x} \frac{M(y, j - 1) - 1}{|N_x|} \mathbf{1}(M(y, t_{j-1}) \geq 1). \end{aligned}$$

*Proof of Theorem 10.11.* We show that  $(\mu_n^\xi)_{n \in \mathbb{N}}$  satisfy the assumptions of Theorem 5.2. We only need to establish the fast stabilization of  $\xi$  on finite windows. Due to the updates occurring in parallel at discrete times, the proof is simpler than in the continuous-time case. Precisely, for every  $x \in \mathcal{P}_n$  we consider a backward cluster  $C'_n(x, j)$  at discrete time instances  $j \in [J]$ , consisting of all space-time points  $(y, j')$ ,  $y \in \mathcal{P}_n$ ,  $j' \in \{0, \dots, j\}$ , that (possibly indirectly) contribute to the update of  $x$  up to time  $j$ . By the nature of discrete time and the update rule (10.33), this backward cluster grows backward from  $j'$  to  $j' - 1$ ,  $j' \in [j]$ , by adding neighbors of points  $y$  in the cluster at time  $j'$ .

Consequently, the projection of the space-time cluster  $C'_n(x, J)$  onto  $\mathcal{P}_n$  gives, besides  $x$ , the union of all  $j$ -nearest neighbors of  $x$  in  $\mathcal{P}_n$ , for  $j \in [J]$ . Call this set  $\mathcal{N}_J(x, \mathcal{P}_n)$ . Enlarging this set by the graph stabilization radius of its elements,  $S_n(y, \mathcal{P}_n)$ ,  $y \in \mathcal{N}_J(x, \mathcal{P}_n)$ , as in (8.2), we obtain, for  $x \in \mathcal{P}_n$ , the bound

$$R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}}_n) \leq \max_{y \in \mathcal{N}_J(x, \mathcal{P}_n)} (|y - x| + S_n(y, \mathcal{P}_n)). \quad (10.36)$$

Using (8.7) and the stabilization of the graph on finite windows, iteratively, together with Lemma 10.5, this yields the required fast stabilization of  $\xi$  on finite windows for general  $J$  and completes the proof of Theorem 10.11.  $\square$

## 11 Empirical random fields and geostatistical models

In the three subsequent subsections, we consider models involving two random objects defined on  $\mathbb{R}^d$ : a random field  $M$  with values possibly in a general Polish space, and a point process  $\mathcal{P}$ . A key feature of these objects is that they possess fast mixing correlations. We consider the following two scenarios:

- The point process  $\mathcal{P}$  is assumed independent of the random field  $M$ . Viewed as a *sampling process* in an increasing window  $W_n$ , it allows us to estimate statistics of the random field  $M$ , such as its distribution function or covariogram. We study the properties of the corresponding estimators, including their asymptotic mean, variance, and Gaussian fluctuations. These results are presented in Sections 11.1 and 11.2.
- In Section 11.3, still assuming independence between  $M$  and  $\mathcal{P}$ , we consider models based on  $\mathcal{P}$  where the random field  $M$  is seen as providing “external” pre-marks in the construction of the model. These are referred to as *geostatistical models* (see [134, 1]). We illustrate this with a Boolean model of this type and present results related to the total length of the edges induced by the model.

In both scenarios mentioned above, the theoretical framework developed in Part II allows us to achieve these goals, providing either new results or reaffirming existing results under assumptions that are easier to verify in practice.

Specifically, let  $M := \{M(x)\}_{x \in \mathbb{R}^d}$  be a random field taking values in a Polish space  $\mathbb{M}$ . As mentioned above, the assumptions and results regarding this random field  $M$  concern its mixing correlations, which we formulate below, naturally extending the definitions given in Definitions 3.5 and 3.6.

**Definition 11.1** (Mixing correlations of the random field).

(i) *The random field  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  with values in  $\mathbb{M}$  satisfies BL-mixing correlations if there exists a family of decreasing functions  $\omega_k : [0, \infty) \rightarrow [0, 2]$ ,  $k \in \mathbb{N}$ , (called correlation decay functions), such that  $\omega_k \leq \omega_{k+1}$  for all  $k \in \mathbb{N}$  and  $\lim_{s \rightarrow \infty} \omega_k(s) = 0$ , and for all  $p, q \in \mathbb{N}$ ,  $x_1, \dots, x_{p+q} \in \mathbb{R}^d$ , and all  $f \in \text{BL}(\mathbb{M}^p)$ , and  $g \in \text{BL}(\mathbb{M}^q)$  we have*

$$|\mathbb{E}[f([M(x_i)]_1^p) g([M(x_i)]_{p+1}^{p+q})] - \mathbb{E}[f([M(x_i)]_1^p)] \mathbb{E}[g([M(x_i)]_{p+1}^{p+q})]| \leq \omega_{p+q}(s), \quad (11.1)$$

where  $s := d([x]_1^p, [x]_{p+1}^{p+q})$  is at (3.3).

(ii) *The random field  $M$  satisfies  $\mathbb{B}$ -mixing correlations if (11.1) holds when BL is replaced by  $\mathbb{B}$ .*

(iii) *Under Item (i) (resp. Item (ii)), the random field  $M$  has fast BL-mixing correlations (resp. fast  $\mathbb{B}$ -mixing correlations) if  $(\omega_k)_{k \in \mathbb{N}}$  are fast decreasing.*

In the final part of this section, Section 11.4, we present explicit constructions of random fields  $M$  satisfying the mixing correlation properties introduced above, which are required for the results developed in Sections 11.1–11.3. This is done by constructing the random field as a marking function of an independent marked point process  $\tilde{\Xi}$  which serves as a *covariate* for the random field  $M$ . This is analogous to Theorem 4.9.

## 11.1 Limit theory for empirical random fields

For a real valued measurable function  $\zeta : \mathbb{R}^d \times \mathbb{M} \rightarrow \mathbb{R}$ , we consider the following random field  $\zeta(M) := \{\zeta(x, M)\}_{x \in \mathbb{R}^d}$  sampled at  $\mathcal{P}$ :

$$\hat{\sigma}_n^{\zeta(M)} := \sum_{x \in \mathcal{P}_n} \zeta(x, M(x)) \delta_{n^{-1/d}x}, \quad n \rightarrow \infty, \quad (11.2)$$

where  $\mathcal{P}_n := \mathcal{P} \cap W_n$ . We call  $\hat{\sigma}_n^{\zeta(M)}$  *empirical random field*.

In what follows, we establish Gaussian fluctuations for  $(\hat{\sigma}_n^{\zeta(M)})_{n \in \mathbb{N}}$  under fast mixing correlations of  $\mathcal{P}$  and  $M$ , and suitable moment assumptions on  $\zeta(M)$ . We say that  $\zeta$  satisfies the  $p$ -moment condition with respect to  $M$  if

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left[ \max \left( 1, |\zeta(x, M(x))|^p \right) \right] \leq M_p^\zeta < \infty. \quad (11.3)$$

**Proposition 11.2** (Limit theory for empirical distribution function of random field). *Let  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  be a random field in  $\mathbb{R}^d$  with values in a Polish space  $\mathbb{M}$  and  $\zeta : \mathbb{R}^d \times \mathbb{M} \rightarrow \mathbb{R}$  be a measurable function. Let  $\mathcal{P}$  be a simple point process on  $\mathbb{R}^d$ , independent of the random field  $M$ . Consider empirical random measures  $(\hat{\sigma}_n^{\zeta(M)})_{n \in \mathbb{N}}$  as in (11.2).*

(i) *Assume that  $\mathcal{P}$  has fast mixing correlations as in Definition 3.4(ii). Assume that either (a) the random field  $M$  satisfies the fast  $\mathbb{B}$ -mixing condition of Definition 11.1(ii)(iii) or (b)  $M$  satisfies the fast BL-mixing condition of Definition 11.1 and  $\zeta(x, \cdot)$  is Lipschitz. If  $\zeta$  satisfies the  $p$ -moment condition (11.3) with respect to  $M$  for all  $p > 1$  then the central limit theorem holds for  $(\hat{\sigma}_n^{\zeta(M)}(f))_{n \in \mathbb{N}}$ , i.e., as  $n \rightarrow \infty$*

$$(\text{Var } \hat{\sigma}_n^{\zeta(M)}(f))^{-1/2} \left( \hat{\sigma}_n^{\zeta(M)}(f) - \mathbb{E} \hat{\sigma}_n^{\zeta(M)}(f) \right) \xrightarrow{d} Z,$$

for  $f \in \mathbb{B}(W_1)$ , provided  $\text{Var } \hat{\sigma}_n^{\zeta(M)}(f) = \Omega(n^\nu)$  for some  $\nu > 0$ .

(ii) *Suppose the random field  $M$  and the point process  $\mathcal{P}$  are stationary (necessarily jointly so, due to their independence) with  $\mathcal{P}$  having intensity  $\rho$ , and suppose that  $\zeta$  is translation invariant, i.e.,  $\zeta(x, \cdot) = \zeta(\cdot)$ . Also assume  $\zeta$  satisfies the  $(1 + \epsilon)$ -moment condition (11.3) with respect to  $M$  for some  $\epsilon > 0$ . Then, the mean empirical field satisfies*

$$n^{-1} \mathbb{E} \hat{\sigma}_n^{\zeta(M)}(f) = \rho \mathbb{E} \zeta(\mathbf{0}, M(\mathbf{0})) \int_{W_1} f(x) dx, \quad n \in \mathbb{N}, \quad (11.4)$$

which holds without assuming mixing conditions on  $\mathcal{P}$  and  $M$  nor Lipschitz conditions on  $\zeta$ . Moreover, under either of the mixing assumptions (a) or (b) of Item (i) we obtain

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var} \hat{\sigma}_n^{\zeta(M)}(f) = \sigma^2(\zeta) \int_{W_1} f(x)^2 dx,$$

where

$$\begin{aligned} \sigma^2(\zeta) &:= \rho \mathbb{E} \zeta^2(\mathbf{0}, M(\mathbf{0})) \\ &+ \int_{\mathbb{R}^d} \left( \mathbb{E}[\zeta(\mathbf{0}, M(\mathbf{0})) \zeta(x, M(x))] \rho^{(2)}(\mathbf{0}, x) - \rho^2 \mathbb{E}[\zeta(\mathbf{0}, M(\mathbf{0}))] \mathbb{E}[\zeta(x, M(x))] \right) dx \in [0, \infty) \end{aligned} \quad (11.5)$$

provided that  $\zeta$  satisfies the  $(2 + \epsilon)$ -moment condition (11.3) with respect to  $M$  for some  $\epsilon > 0$ .

(iii) Let  $\zeta := (\zeta_1, \dots, \zeta_m)$  be a vector of translation-invariant, real-valued measurable functions on  $\mathbb{R}^d \times \mathbb{M}$ , with each  $\zeta_i$  satisfying the  $(2 + \epsilon)$ -moment condition (11.3) with respect to  $M$  for some  $\epsilon > 0$ . Under either of the mixing assumptions (a) or (b) of Item (i), and under the stationarity assumption of Item (ii) the multivariate central limit theorem holds as in Theorem 5.4.

**Remark 11.3.** Central limit theorems for empirical measures of random fields were established by Pawlas [113], who assumes strong mixing conditions (in terms of  $\sigma$ -fields) on both the point process  $\mathcal{P}$  and the random field  $M$ . He also points out that, in practice, these conditions are often difficult to verify. In contrast, assuming fast mixing correlations of  $\mathcal{P}$  together with fast  $\mathbb{B}$ - or BL-mixing correlations of  $M$  constitutes a weaker and often more tractable alternative in applications; see Section 11.4 for examples of such  $M$ .

*Proof of Proposition 11.2.* We first make some common observations on the proofs of the three statements. Consider the marked point process  $\tilde{\mathcal{P}} := \{(x, \zeta(x, M(x)))\}_{x \in \mathcal{P}}$ . Due to the independence between  $\mathcal{P}$  and  $M$ , the corresponding Palm distributions of  $\zeta(x, M(x))$  (with respect to multiple locations of  $x$  in  $\mathcal{P}$ ) coincide with the (unconditional) joint finite-dimensional distributions of  $\zeta$ , i.e., for all bounded functions  $f$  and  $x_1, \dots, x_p \in \mathbb{R}^d$ , we have

$$\mathbb{E}_{[x]_1^p} \left[ f([\xi(x_i, \zeta(x_i, M(x_i)))]_1^p) \right] = \mathbb{E} \left[ f([\xi(x_i, \zeta(x_i, M(x_i)))]_1^p) \right], \quad (11.6)$$

where in the left-hand side the Palm expectation is with respect to  $\mathcal{P}$ .

We prove now statement (i) where we directly apply the umbrella result, Theorem 3.3. Indeed, it suffices to show that  $\tilde{\mathcal{P}}$  has fast BL-mixing correlations as in Definition 3.1(i) and (iii) (since here  $\mathcal{P}_n = \mathcal{P} \cap W_n$ ). For this, we rely on the result on joint mixing correlations of points and marks given in Proposition 3.7. Due to the independence of the random field  $M$  and the point process  $\mathcal{P}$ , the assumed fast  $\mathbb{B}$ -mixing correlations of  $M$  translate directly into the fast BL-mixing of the marks  $\zeta(x, M(x))$ ,  $x \in \mathcal{P}$ . Indeed, fast  $\mathbb{B}$ -mixing correlations of  $\{M(x)\}_{x \in \mathbb{R}^d}$  implies fast  $\mathbb{B}$ - (hence BL-) mixing correlations of the field  $\{\zeta(x, M(x))\}_{x \in \mathbb{R}^d}$ . If the function  $\zeta(x, \cdot)$  is Lipschitz, then fast BL-mixing of  $M$  is sufficient for this last observation. Thus, in view of (11.6) and under either of the assumptions (a) or (b) of (i), we have fast BL-mixing correlations of the marked point process  $\tilde{\mathcal{P}} = \{\zeta(x, M(x))\}_{x \in \mathcal{P}}$ . This justifies statement (i) via Theorem 3.3.

We now establish statements (ii) and (iii). The first of these, concerning expectations in (11.4), follows directly from the Campbell formula under the stationarity and independence assumptions on  $\mathcal{P}$  and  $M$ .

All other statements (i.e., variance asymptotics (11.5) and the multivariate CLT in Item (iii)) share a common feature: the marked point process  $\tilde{\mathcal{P}} = \{(x, \zeta(x, M(x)))\}_{x \in \mathcal{P}}$  exhibits fast BL-mixing correlations. This property follows directly from the mixing properties of  $\mathcal{P}$  and  $M$ , as explained in the proof of Item (i). The asymptotics of the variance and the multivariate central limit theorem under the  $(2 + \epsilon)$ -moment condition then rely on the fast BL-mixing of  $\tilde{\mathcal{P}}$  in the framework of Theorem 5.5. Note that we do not require summable fast BL-mixing correlations of  $\mathcal{P}$  as is to be expected from Theorem 5.5 or Proposition 5.3. In those results, such an assumption on  $\tilde{\mathcal{P}}$  (which corresponds to  $\mathcal{P}$  here) is to establish fast BL-mixing of  $\tilde{\mathcal{P}}$  (which corresponds to  $\tilde{\mathcal{P}}$  here) but here our assumptions on  $M, \mathcal{P}$  already guarantee fast BL-mixing correlations of  $\tilde{\mathcal{P}}$ .  $\square$

## 11.2 Limit theory for covariograms

In this section, we are interested in the *covariogram*, a tool for capturing the spatial dependence structure of a real-valued, stationary random field. Let  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  denote such a  $\mathbb{R}$ -valued random field. The *theoretical covariogram* of  $M$  is the deterministic, real-valued function defined by

$$\gamma(h) = \gamma_M(h) := \frac{1}{2} \mathbb{E}[(M(\mathbf{0}) - M(h))^2], \quad h \in \mathbb{R}^d,$$

see, e.g., [134, Section 2].

To estimate the value of  $\gamma(h)$ , we consider sampling the random field  $M$  at locations given by a stationary, simple point process  $\mathcal{P}$  on  $\mathbb{R}^d$ , independent of  $M$ . A straightforward approach would be to average the quantities  $\frac{1}{2}|M(x) - M(y)|^2$  over all pairs of points  $x, y$  in some observation window  $W_n$  such that  $y - x = \pm h$ , where we admit  $\pm h$  since the function  $\gamma$  is even. However, unless the point process  $\mathcal{P}$  possesses a lattice structure aligned with the specific vector  $h$ , there will typically be no such pairs satisfying this exact relation.

To overcome this issue, we introduce a *tolerance parameter*  $\delta > 0$ , such that  $\delta < |h|$ , and instead consider pairs  $x, y \in W_n$  such that  $y - x \in B_\delta(\mathbf{0} \pm h)$ , with the notation  $B_\delta(a \pm h) := B_\delta(a + h) \cup B_\delta(a - h)$  for  $a \in \mathbb{R}^d$ . Under this approach, we define the following score function for the points of  $\mathcal{P}$ , involving also  $M$ :

$$\xi(x, \mathcal{P}_n, M) := \frac{1}{4\theta_d \delta^d} \sum_{y \in \mathcal{P}_n: y-x \in B_\delta(\mathbf{0} \pm h)} (M(x) - M(y))^2, \quad (11.7)$$

where  $\theta_d$  denotes the volume of the unit ball in  $\mathbb{R}^d$ ; the factor 4 (instead of 2) compensates for the fact that the term  $(M(x) - M(y))^2$  is counted twice in the sum, and  $\delta < |h|$  prevents taking  $x = y$ . It is natural to postulate that this score satisfies the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[ \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n, M) \right] = \frac{\rho}{4\theta_d \delta^d} \mathbb{E}^0 \left[ \int_{B_\delta(\mathbf{0} \pm h)} (M(\mathbf{0}) - M(y))^2 \mathcal{P}(dy) \right]$$

$$\begin{aligned}
&= \frac{\rho}{4\theta_d \delta^d} \int_{B_\delta(\mathbf{0} \pm h)} \mathbb{E}[(M(\mathbf{0}) - M(y))^2] \rho^{(2)}(\mathbf{0}, y) \, dy \\
&= \frac{\rho}{\theta_d \delta^d} \int_{B_\delta(h)} \gamma(y) \rho^{(2)}(\mathbf{0}, y) \, dy =: \rho \bar{\gamma}(h),
\end{aligned} \tag{11.8}$$

where  $\rho = \rho^{(1)}$  and  $\rho^{(2)}$  denote, respectively, the first- and second-order correlation functions of  $\mathcal{P}$  and, in the last equality, we used the translation invariance and symmetry of  $\rho^{(2)}$ , and the evenness of  $\gamma$ . For small  $\delta > 0$ , this limit is expected to be close to  $\rho^{(2)}(\mathbf{0}, h) \gamma(h)$ , thus providing a practical means to estimate the covariogram using the sampling process  $\mathcal{P}$ .

The following result provides a more complete limit theory for the *empirical covariogram*  $\gamma(h)$  of the random field  $M$

$$\sigma_n^\gamma = \sigma_n^{\gamma(h)} := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n, M) \delta_{n^{-1/d}x}, \tag{11.9}$$

for  $h \in \mathbb{R}^d$ , based on the independent sampling point process  $\mathcal{P}$ , where we suppress the dependence on the parameter  $\delta$ ,  $0 < \delta < |h|$ , related to the tolerance.

**Proposition 11.4** (Limit theory for empirical covariogram). *Let  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  be a real-valued random field in  $\mathbb{R}^d$  having fast  $\mathbb{B}$ -mixing correlations as in Definition 11.1(ii)(iii). Let  $\mathcal{P}$  be a simple point process on  $\mathbb{R}^d$  having fast mixing correlations as in Definition 3.4(ii), bounded reduced Palm intensity function (8.10), and independent of the random field  $M$ . Consider empirical covariogram  $\sigma_n^\gamma$  of  $M$  based on  $\mathcal{P}$ , as in (11.9), for some value  $h$  and tolerance  $\delta > 0$ .*

(i) *If  $M$  has  $p$ -moment conditions for all  $p > 1$  (i.e., (11.3) holds with  $\zeta(x, M(x)) = M(x)$ ) then the central limit theorem holds for  $(\sigma_n^\gamma(f))_{n \in \mathbb{N}}$ , i.e., as  $n \rightarrow \infty$*

$$(\text{Var } \sigma_n^\gamma(f))^{-1/2} \left( \sigma_n^\gamma(f) - \mathbb{E} \sigma_n^\gamma(f) \right) \xrightarrow{d} Z$$

for  $f \in \mathbb{B}(W_1)$ , provided the random  $\text{Var } \sigma_n^\gamma(f) = \Omega(n^\nu)$  for some  $\nu > 0$ .

(ii) *Further, if the random field  $M$  and the point process  $\mathcal{P}$  are stationary (necessarily jointly so, due to their independence), with  $\rho$  the intensity of  $\mathcal{P}$ , then the expectation and variance asymptotics and multivariate central limit theorem hold mutatis mutandis for  $(\sigma_n^\gamma(f))_{n \in \mathbb{N}}$  as in Proposition 5.3 and Theorem 5.4 under  $p = (2 + \epsilon)$ - or  $p = (4 + \epsilon)$ -moment condition for  $M$  i.e., (11.3) with  $\zeta(x, M(x)) = M(x)$ , respectively. Precisely, the limiting expectation takes the form  $\rho \bar{\gamma}(h)$  of (11.8) and the covariance of the empirical covariogram at two values  $\sigma_n^{\gamma(h)}(f)$  and  $\sigma_n^{\gamma(h')}(f)$ , with  $h, h' \in \mathbb{R}^d$ , (and for the simplicity, we assume the same tolerance  $\delta > 0$  such that  $\delta < \min\{|h|, |h'|, |h - h'|, |h + h'|\}$ ) takes the form*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Cov}(\sigma_n^{\gamma(h)}(f), \sigma_n^{\gamma(h')}(f)) = \sigma^2(h, h') \int_{W_1} f(x)^2 \, dx \in (-\infty, \infty),$$

where

$$\sigma^2(h, h') := \tag{11.10}$$

$$\begin{aligned}
& \frac{\rho}{16\theta_d^2\delta^{2d}} \int_{B_\delta(\mathbf{0}\pm h)} \int_{B_\delta(\mathbf{0}\pm h')} \mathbb{E}[(M(\mathbf{0}) - M(y))^2(M(\mathbf{0}) - M(y'))^2] \rho^{(3)}(\mathbf{0}, y, y') \, dy \, dy' \\
& + \int_{\mathbb{R}^d} \left( \frac{1}{16\theta_d^2\delta^{2d}} \int_{B_\delta(\mathbf{0}\pm h) \cap B_\delta(z\pm h')} \mathbb{E}[(M(\mathbf{0}) - M(y))^2(M(z) - M(y))^2] \rho^{(3)}(\mathbf{0}, z, y) \, dy \right. \\
& \quad \left. + \frac{1}{16\theta_d^2\delta^{2d}} \int_{B_\delta(\mathbf{0}\pm h)} \int_{B_\delta(z\pm h')} \mathbb{E}[(M(\mathbf{0}) - M(y))^2(M(z) - M(y'))^2] \rho^{(4)}(\mathbf{0}, z, y, y') \, dy \, dy' \right. \\
& \quad \left. - \rho^2 \bar{\gamma}(h) \bar{\gamma}(h') \right) dz \in (-\infty, \infty).
\end{aligned}$$

*Proof.* First, we indicate how to verify the moment conditions. Using the independence of  $\mathcal{P}$  and  $M$  we can apply the Minkowski inequality and then the Campbell-Little-Mecke formula (2.5) to derive the following bound:

$$\begin{aligned}
& \mathbb{E}_{[\mathcal{P}]_1^q} [\max(1, |\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)|^p)] \\
& \leq \left( 1 + \frac{1}{4\theta_d\delta^d} \left( \sum_{i=1}^q \mathbb{E}[(M(x_1) - M(x_i))^{2p}]^{1/p} + \int_{B_\delta(x_1\pm h)} \mathbb{E}[(M(x_1) - M(y))^{2p}]^{1/p} \rho_{[\mathcal{P}]_1^1}^{(1)}(y) \, dy \right) \right)^p.
\end{aligned}$$

Now the finiteness of the  $p$ th moment of the score function (11.7) follows from the existence of  $2p$ -moments of the random field  $M$  and the moment assumption (8.10) on the reduced Palm intensity of  $\mathcal{P}$ .

We now prove points (i) and (ii). Observe that the score function  $\xi$  in (11.7) has the specific form of a local  $U$ -statistic, whose general expression is given in (6.20). The limit theory (i.e., Theorem 5.2, Proposition 5.3 and Theorem 5.4) for such scores can be established under merely fast  $\mathbb{B}$ -mixing correlations of the marked input process  $\tilde{\mathcal{P}}$ , as stated in Corollary 6.8. More precisely, under these fast mixing correlations—which are weaker than the summable exponential mixing correlation assumption—the central limit theorem in Item (i) follows from Theorem 5.2.

Furthermore, in the stationary setting, the limits of the mean and variance follow from Proposition 5.3, under the  $(1 + \epsilon)$ - and  $(2 + \epsilon)$ -moment conditions for the score  $\xi$ , respectively. Since this score involves the square of the values of the random field  $M$ , these requirements correspond to the  $p = (2 + \epsilon)$  and  $p = (4 + \epsilon)$  moment conditions (11.3), with  $\zeta(x, M(x)) = M(x)$ , respectively. Similarly, Theorem 5.4, under these reduced mixing assumptions, establishes the multivariate central limit theorem under the  $p = 2 + \epsilon$  moment condition for  $\xi$ , which again translates to the  $p = (4 + \epsilon)$  moment condition (11.3) with  $\zeta(x, M(x)) = M(x)$ .

Finally, the explicit expressions of the limiting quantities follow from a suitable application of the Campbell-Little-Mecke formula (2.5) involving  $\mathcal{P}$  and  $M$ .  $\square$

### 11.3 Geostatistical Boolean models

In the two previous sections, our main object of interest was a random field  $M = \{M(x)\}_{x \in \mathbb{R}^d}$ . We considered a point process  $\mathcal{P}$  in  $\mathbb{R}^d$ , independent of this field, used to estimate distributional properties of  $M$ . In this context, the process  $\mathcal{P}$  was regarded merely as a ‘sampling point process’. In the present

section, our focus shifts to more complex models that are based on the process  $\mathcal{P}$  as a ground process and involve pre-marks  $U(x)$  attached to its points  $x \in \mathcal{P}$ . This terminology was introduced in Section 4 and has been used throughout the applicative part of our work. However, here, these pre-marks take the form  $U(x) = M(x)$ , where the values are driven by an “external” random field  $M$ , independent of  $\mathcal{P}$ . In this setting, the marks are referred to as *geostatistical marks*, and the corresponding models are called *geostatistical models*, as suggested in the seminal paper [134]. Despite this paradigmatic change in the roles of  $M$  and  $\mathcal{P}$ , one can easily recognize in the covariogram considered in Section 11.2 a particular instance of a geostatistical model. In this section, we revisit another model of this type, namely the *geostatistical Boolean model* studied in Ahlberg and Tykesson [1], thus illustrating the relevance of the limit theory developed in Section 5 in this context as well.

As in previous subsections, consider a real, non-negative valued random field  $M$  on  $\mathbb{R}^d$ , and let  $\mathcal{P}$  be a point process on  $\mathbb{R}^d$ , independent of  $M$ . Following [1], we define the *geostatistical random geometric graph*  $\mathcal{G}(\mathcal{P}, M)$ , which places an edge between two points  $x, y \in \mathcal{P}$  whenever  $|x - y| \leq M(x) + M(y)$ , i.e., the balls at both the points intersect each other. Equivalently, this is the graph induced by the *spherical geostatistical Boolean model (GeoBM)* on  $\mathcal{P}$ , where to each point  $x \in \mathcal{P}$  we attach a ball of radius  $M(x)$ , given by the value of the (independent) random field  $M$ . See the upcoming Example 11.6 for a Poisson-Voronoi driven geostatistical Boolean model.

We establish below the limit theory for the total edge length of  $\mathcal{G}(\mathcal{P}_n, M)$ , thus taking up a statistic suggested by Ahlberg and Tykesson [1], but whose limit theory was not analyzed. Specifically, consider the following score function for  $x \in \mathcal{P}_n$ :

$$\xi(x, \mathcal{P}_n, M) := \frac{1}{2} \sum_{y \in \mathcal{P}_n} |x - y| \mathbf{1}(|x - y| \leq M(x) + M(y)), \quad (11.11)$$

which corresponds to one half of the sum of the lengths of the edges in  $\mathcal{G}(\mathcal{P}_n, M)$  that are incident to  $x \in \mathcal{P}_n$ . Define

$$\sigma_n := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n, M) \delta_{n^{-1/d}x}, \quad (11.12)$$

and observe that the total mass of  $\sigma_n$  is equal to the total edge length of the graph  $\mathcal{G}(\mathcal{P}_n, M)$ .

**Proposition 11.5** (Limit theory for total edge lengths in GeoBM). *Let  $M = \{M(x)\}_{x \in \mathbb{R}^d}$  be a non-negative, real-valued random field on  $\mathbb{R}^d$  having fast  $\mathbb{B}$ -mixing correlations as in Definition 11.1(ii)(iii). Assume that  $M$  is bounded some constant  $R$ , that is,  $\sup_{x \in \mathbb{R}^d} M(x) \leq R < \infty$ . Let  $\mathcal{P}$  be a simple point process on  $\mathbb{R}^d$  having fast mixing correlations as in Definition 3.4(ii), bounded reduced Palm intensity function (8.10), and independent of the random field  $M$ . Consider the total edge length  $\sigma_n$  of the geostatistical graph  $\mathcal{G}(\mathcal{P}_n, M)$  defined in (11.12).*

(i) *A central limit theorem holds for the family  $(\sigma_n(f))_{n \in \mathbb{N}}$ ; that is, as  $n \rightarrow \infty$ ,*

$$(\text{Var } \sigma_n(f))^{-1/2} (\sigma_n(f) - \mathbb{E} \sigma_n(f)) \xrightarrow{d} Z,$$

*for all  $f \in \mathbb{B}(W_1)$ , provided that the variance satisfies  $\text{Var } \sigma_n(f) = \Omega(n^\nu)$  for some  $\nu > 0$ .*

(ii) Further, if the random field  $M$  and the point process  $\mathcal{P}$  are stationary (necessarily jointly so, due to their independence), with  $\rho$  denoting the intensity of  $\mathcal{P}$ , then the expectation and variance asymptotics hold for  $(\sigma_n(f))_{n \in \mathbb{N}}$ , as stated in Proposition 5.3.

*Proof.* The boundedness of the random field  $M$  implies that the score function  $\xi$  in (11.11) has the specific form of a local  $U$ -statistic, whose general expression is given in (6.20). The limit theory for such scores can be established under merely fast  $\mathbb{B}$ -mixing correlations of the marked input process  $\tilde{\mathcal{P}} = \{(x, M(x))\}_{x \in \mathcal{P}}$ , as stated in Corollary 6.8, in the same way as for the covariogram considered in Proposition 11.4 (the bound for the reduced Palm intensity of  $\mathcal{P}$  is only required, no for higher moments due to the boundedness of  $M$ ).  $\square$

**Example 11.6** (Poisson–Voronoi–driven GeoBM). Consider the following Boolean model whose grains are defined according to a geostatistical marking. Given a Poisson point process  $\mathcal{X}$ , let  $\text{Vor}(\mathcal{X})$  be the Poisson-Voronoi tessellation of  $\mathbb{R}^d$  induced by  $\mathcal{X}$ , with cells  $C(y, \mathcal{X}), y \in \mathcal{X}$ . For  $x \in \mathbb{R}^d$ ,  $C(x, \mathcal{X})$  denotes the cell containing  $x$ . In other words, if  $y \in \mathcal{X}$  is the nearest point of  $\mathcal{X}$  to  $x \in \mathbb{R}^d$  (which is a.s. defined), then we have that  $x \in C(y, \mathcal{X})$ . We assign i.i.d. real-valued random variables  $M(y)$  to each  $C(y, \mathcal{X}), y \in \mathcal{X}$ . We assume the random variables are bounded by  $R \in (0, \infty)$ .

Next, given the ground point process  $\mathcal{P}$  on  $\mathbb{R}^d$ , assumed independent of  $\mathcal{X}$ , we construct a Boolean model by putting a ball around each point  $x \in \mathcal{P}$  and whose radius equals the random variable  $M(x)$  assigned to  $C(x, \mathcal{X})$  (i.e.,  $C(y, \mathcal{X})$  such that  $x \in C(y, \mathcal{X})$ ). The grains  $B(x, M(x)), x \in \mathcal{P}$ , are spatially dependent, as their radii are a function of the underlying Poisson-Voronoi tessellation  $\text{Vor}(\mathcal{X})$ . As in the second example of the introduction to [1], this is a geostatistical Boolean model on  $\mathcal{P}$ , with cell radii determined by the external Poisson point process  $\mathcal{X}$ . This model generates a geostatistical random geometric graph  $\mathcal{G}(\mathcal{P}_n, M)$ . Proposition 11.5 gives conditions on  $\mathcal{P}$  such that the total edge length of  $\mathcal{G}(\mathcal{P}_n, M)$  satisfies the CLT. Under the same conditions on  $\mathcal{P}$ , one could similarly show that the edge count of  $\mathcal{G}(\mathcal{P}_n, M)$  is also asymptotically normal. Questions regarding such statistics of  $\mathcal{G}(\mathcal{P}_n, M)$  are indicated at the conclusion of the introduction to [1].

**Remark 11.7** (Towards more general geostatistical models).

(i) *Multivariate statistics of GeoBM.* Under the assumptions of Proposition 11.5, the result can be straightforwardly extended to a larger class of score functions  $\xi$  related to GeoBM, as given in Corollary 6.8, allowing, for example, the same statements for the total number of edges in  $\mathcal{G}(\mathcal{P}_n, M)$ . Moreover, in the stationary setting, one obtains a multivariate central limit theorem for several such scores by leveraging Theorem 5.4.

(ii) *Unbounded random field in GeoBM.* Extending the result of Proposition 11.5 to the case where  $\sup_{x \in \mathbb{R}^d} M(x)$  has fast-decreasing tail probabilities is possible, but requires summable exponential  $\mathbb{B}$ -mixing of  $\tilde{\mathcal{P}} = \{(x, M(x))\}$  as in Definition 4.8.

(iii) *Beyond GeoBM.* The assumption of summable exponential  $\mathbb{B}$ -mixing of  $\tilde{\mathcal{P}} = \{(x, M(x))\}$  also allows one to obtain the same limit theorems for BL-localizing score functions in fairly general geostatistical models.

(iv) *Geostatistical models with a dependent random field.* The paper [134], in the first lines of page 80, discusses the desirability of relaxing the fundamental assumption of independence in geostatistical marking, namely between the sampling locations (ground process  $\mathcal{P}$ ) and the values of the spatial random field  $M$ . Our framework, in its full generality, allows one to address such generalized models, capturing the dependence between  $M$  and  $\mathcal{P}$  through the Palm distributions of the pre-marks  $M(x)$  for  $x \in \mathcal{P}$ . (Recall that in the classical geostatistical models these Palm distributions coincide with the original finite-dimensional distributions of  $M$ .) In this setting, again, summable exponential  $\mathbb{B}$ -mixing of  $\tilde{\mathcal{P}} = \{(x, M(x))\}$  and BL-localizing score functions constitute the right tools. It then becomes primarily a modeling question what kind of dependence between  $M$  and  $\mathcal{P}$  still allows one to obtain the required mixing properties of  $\tilde{\mathcal{P}}$ .

## 11.4 Mixing correlations of random fields with a covariate point process

In the previous section, the mixing of the random field as in Definition 11.1 was taken as a constitutive assumption, allowing one to obtain various limiting results in geostatistical models. In this section, we provide a tool to establish the mixing of the random field when it is itself constructed from a *covariate point process*.

Consider a marked point process  $\tilde{\Xi} = \{(x, U(x))\}_{x \in \Xi}$  on  $\mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  with marks in a Polish space  $\mathbb{K}$ . Let

$$F : \mathbb{R}^d \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}} \longrightarrow \mathbb{M}$$

be a measurable function, where  $\mathbb{M}$  is a Polish space. Define the random field on  $\mathbb{R}^d$  with values in  $\mathbb{M}$  by

$$M(x) = F(x, \tilde{\Xi}), \quad x \in \mathbb{R}^d.$$

The process  $\tilde{\Xi}$  plays the role of a *covariate point process* for this field.

If the function  $F$  is BL-localizing or stabilizing on the marked point process  $\tilde{\Xi}$ , then the random field  $M$  exhibits, respectively, BL-mixing or  $\mathbb{B}$ -mixing correlations, in the sense of Definition 11.1.

**Lemma 11.8** (Fast mixing correlations of random fields with covariate point process). *Let  $\tilde{\Xi}$  be a simple, marked point process on  $\mathbb{R}^d \times \mathbb{K}$  having summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8, and with decay function  $\hat{\omega}_k(s) = \hat{C}_k \hat{\phi}(s)$ . Consider random field  $M(x) := F(x, \tilde{\Xi})$ ,  $x \in \mathbb{R}^d$ , where  $F$  takes values in Polish space  $\mathbb{M}$  and is BL-localizing as in Definition 4.2 (ii) and (iii) (resp. fast stabilizing as in Definition 4.6(i) and (iii)) but under the (unconditional) probability  $\mathbb{P}$  of  $\tilde{\Xi}$  in both cases. Then  $M$  exhibits fast BL-mixing correlations as in Definition 11.1(i) and (iii) (respectively, fast  $\mathbb{B}$ -mixing correlations as in Definition 11.1(ii) and (iii)).*

*Proof.* By considering the restricted score  $F^{(r)}$  (i.e.,  $F^{(r)}(x, \tilde{\Xi}) = F(x, \tilde{\Xi} \cap B_r(x))$ ,  $x \in \mathbb{R}^d$ ), it is evident that the radius of stabilization  $R^{F^{(r)}}$  satisfies  $R^{F^{(r)}} \leq r$ . We use the FME expansion (Theorem 6.1) and follow the proof of Proposition 6.4 to obtain

$$|\mathbb{E}[f([F^{(r)}]_1^p)g([F^{(r)}]_{p+1}^{p+q})] - \mathbb{E}[f([F^{(r)}]_1^p)] \mathbb{E}[g([F^{(r)}]_{p+1}^{p+q})]| \leq \hat{\phi}\left(\frac{s}{2}\right) \sum_{l=1}^{\infty} \frac{(4(p+q)\theta_d r^d)^l}{l!} \hat{C}_l.$$

The difference between this and (6.13) is that the ground point process  $\tilde{\Xi}$  is not conditioned to have points at  $x_1, \dots, x_{p+q}$ . Consequently the constants  $\hat{C}_l$  have no dependence on  $(p+q)$  and only depend on  $l$ . We may extend the arguments to unrestricted score functions  $F$  satisfying fast BL-localization, by following the proof of part (i) of Theorem 4.9 (resp. for fast stabilization of  $F$ , part (iii) of the same theorem); see Section 6.2 and in particular, the bound (6.30) (resp. (6.29)).  $\square$

**Corollary 11.9** (Shot-noise random field with fast mixing correlations). *Consider a simple marked point process  $\tilde{\Xi} = \{\tilde{x} = (x, U(x))\}_{x \in \Xi}$  on  $\mathbb{R}^d$  with i.i.d. marks  $U(x)$  taking values in a Polish space  $\mathbb{K}$ . Let the shot-noise field  $M$  on  $\mathbb{R}^d$  with values in  $\mathbb{R}$  be defined by*

$$M(x) := \sum_{y \in \Xi} \zeta(x, \tilde{y}), \quad x \in \mathbb{R}^d,$$

for some measurable function  $\zeta : \mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{K}) \rightarrow \mathbb{R}$  such that  $\zeta(\tilde{x}, \tilde{y}) \leq \phi(|x - y|)$  for a fast-decreasing function  $\phi$  as in (2.1). If the ground point process  $\Xi$  has fast summable exponential mixing as in Definition 4.8(iii), then the random field  $M$  exhibits fast BL-mixing correlations as in Definition 11.1(i) and (iii).

*Proof.* The result can be derived from Lemma 11.8 by setting  $F(x, \tilde{\Xi}) := M(x)$ ,  $x \in \mathbb{R}^d$ . Indeed, the i.i.d. marking ensures that  $\tilde{\Xi}$  is fast summable exponential  $\mathbb{B}$ -mixing via Proposition 6.4. Further, we assert that the score function  $F$  is fast  $L^1$ -stabilizing (hence BL-localizing, see Remark 4.7(ii)) on  $\tilde{\Xi}$ , leveraging the boundedness of the first intensity function of the ground process. To see this, let  $p \in \mathbb{N}$ ,  $f$  a bounded Lipschitz function on  $\mathbb{R}^p$  and  $x_1, \dots, x_p \in \mathbb{R}^d$ . Using the Lipschitzness of  $f$ , the definition of  $F$ , and the Campbell-Little-Mecke formula, we obtain

$$\begin{aligned} |\mathbb{E}f([F(x_i, \tilde{\Xi})]_1^p) - \mathbb{E}f([F(x_i, \tilde{\Xi} \cap B_r(x_i))]_1^p)| &\leq \sum_{i=1}^p \mathbb{E} |F(x_i, \tilde{\Xi}) - F(x_i, \tilde{\Xi} \cap B_r(x_i))| \\ &= \sum_{i=1}^p \mathbb{E} \left| \sum_{y \in \Xi \cap B_r(x_i)^c} \zeta(x_i, \tilde{y}) \right| \\ &\leq \kappa_0 \sum_{i=1}^p \int_{y \in B_r(x_i)^c} \phi(|y - x_i|) dy \\ &\leq p\kappa_0 \int_{y \in B_r(\mathbf{0})^c} \phi(|y|) dy \\ &\leq p\kappa_0 \theta_{d-1} \int_r^\infty s^{d-1} \phi(s) ds, \end{aligned}$$

which is fast-decreasing in  $r$  as  $\phi$  is fast-decreasing. Thus  $F$  is fast  $L^1$ -stabilizing and hence is also fast BL-localizing.  $\square$

## 12 Further applications and directions

In this section we indicate some random geometric structures that deserve separate investigation and where we expect that the general limit theory of Part II will be applicable. More precisely the gen-

eral theorems of Section 5 may be used to establish the limit theory for statistics of other geometric structures, including those having an underlying graph which is randomized or which has a dynamic component. We list three such structures and include five directions for future research.

(i) *Statistics of weighted Voronoi tessellations.* Recall, given a weight function  $\rho : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ , each  $x \in \mathcal{P}$  generates weighted cells

$$C(x, \mathcal{P}) := \{y \in \mathbb{R}^d : \rho(y, x) \leq \rho(y, z) \text{ for all } z \in \mathcal{P}\},$$

the collection of which generates the weighted Voronoi tessellation. It is not difficult to establish conditions on the weights and the cell centers  $\mathcal{P}$  under which geometric characteristics (including total number of edges, total length of edges) may be expressed as a sum of fast BL-localizing score functions satisfying the conditions of Proposition 5.3 and Theorem 5.4. The work of Flimmel et al. [59] shows that if the weight  $\rho$  is bounded, then this gives rise to score functions satisfying stopping set stabilization; if an unbounded  $\rho$  satisfies a suitable integrability condition, then it is likely that this would give rise to score functions satisfying fast BL-localization.

(ii) *Euclidean minimal spanning trees.* One could apply the law of large numbers given by Proposition 5.3 to deduce the first order limit theory for statistics (total edge length, number of vertices of a given degree) of the minimal spanning tree on the ergodic point sets  $\mathcal{P}$  in Subsection 5.2 having fast mixing correlations. This extends Penrose and Yukich [123, Theorem 2.3] and is proved by making straightforward modifications to the proof of that theorem, which is confined to Poisson input. In particular we expect to require that (i) there is a constant  $\rho$  such that the complement of the union of closed balls of radius  $\rho$  centered at points of  $\mathcal{P}$  a.s. has no unbounded component and (ii) the infinite component in continuum percolation on  $\mathcal{P}$  is a.s. unique.

(iii) *Spatial birth-growth models.* In the classical spatial birth-growth model, seeds (grains) arrive independently and at random on a substrate. Upon contact with the substrate the seeds grow radially in all directions with independent growth rates. One is typically interested in geometric statistics of this model, including the total volume covered at time  $t_0$ . Our general results provide the limit theory for statistics of these models under conditions which are less restrictive than the classical ones. This includes possibly dependent arrival locations on the substrate, including arrival locations specified by the realization of a point process having fast mixing correlations. It also allows for the initial grain sizes to be a function of previously arrived grains falling within some fixed neighborhood and it also allows the growth of the grains to be dependent on growth rates of grains within some fixed distance. This gives rise to a Boolean model having dependencies in grain locations, initial grain sizes, as well as grain growth rates. Theorem 5.2 may be used to establish the central limit theory for statistics of such models, including, for example, statistics involving sums of scores  $\xi(x, \mathcal{P}) := h((\text{Vol}(x, \mathcal{P}, t))_{t \in [0, t_0]})$ ,  $x \in \mathcal{P}$ , where  $\text{Vol}(x, \mathcal{P}, t)$  is the volume contribution at time  $t$  to the birth-growth model associated to the seed at  $x$  and  $h$  is an appropriate measurable function.

We list additional models where we expect our main general results would establish limit theorems for relevant statistics of these structures. These models merit an independent investigation.

(i) *Randomized graphs  $\mathcal{G}$  on  $\mathcal{P}$ .* The graphs underlying the spin models, diffusion models and interacting particle systems, as developed in Sections 8.1.1, 9.1.2 and 10.1 respectively, can be naturally extended to accommodate extra randomization—for instance, Laguerre tessellations with random weights or, more generally, weighted Delaunay/Voronoi graphs. In these models, each site is assigned a random weight that modifies the effective distance in the classical Euclidean construction; see Chiu et al. [33, Section 9.2]. In these cases, the graph  $\mathcal{G}$  is  $\mathcal{G}(\hat{\mathcal{P}})$  where  $\hat{\mathcal{P}} = \{(x, W(x))\}_{x \in \mathcal{P}}$  is a marked point process including auxiliary marks  $W(x)$ , for example, the weights of the aforementioned tessellations. Using our framework of neighbor-marking functions  $\bar{N}$  of  $\hat{\mathcal{P}}$ , one can extend the notion of stabilizing interaction graphs to include these randomized graphs  $\mathcal{G}(\hat{\mathcal{P}})$ . Graphs arising from the random connection model and from Boolean models with random grains may not exhibit this type of stabilization if the weights or grains are unbounded. In such cases, a promising alternative approach goes as follows.

We sketch this approach in the case of interacting diffusions and it can be adapted suitably to the case of interacting particle systems as well. We extend the input process to  $\tilde{\mathcal{P}} = \{(x, M(x), Z_x, \bar{N}_x)\}_{x \in \mathcal{P}}$ , with  $M(x)$  representing the initial conditions of the model (states of particles or diffusions) and  $Z_x$  representing Brownian motions in the case of diffusions, and  $\bar{N}_x$  representing the neighborhood of  $x$  in the spatial random graph  $\mathcal{G} = \mathcal{G}(\hat{\mathcal{P}})$  where  $\hat{\mathcal{P}}$  is marked point process. Assuming summable exponential  $\mathbb{B}$ -mixing for this extended input process  $\tilde{\mathcal{P}}$ , it would be enough to assume an ‘ $L^1$ -type stabilization’ condition for the neighborhood distance instead of the graph stabilization condition at (9.2); namely it would be enough to assume a condition of the form

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{P}_{[x]_1^p} \left( \max_{y \in \bar{N}_{x_1}} |y| > s \right) \leq \varphi'_p(s), \quad s > 0, \quad (12.1)$$

instead of our usual stabilization condition on the interaction range

$$\sup_{x_1, \dots, x_p \in \mathbb{R}^d} \mathbb{P}_{[x]_1^p} (S(x_1, \mathcal{P}) > s) \leq \varphi'_p(s), \quad s > 0,$$

for a rapidly decaying function  $\varphi'_p$ . However, we caution that checking the summable exponential  $\mathbb{B}$ -mixing of  $\tilde{\mathcal{P}}$  and the above ‘ $L^1$ -type stabilization’ condition may be non-trivial beyond the Poisson-Boolean or random connection model.

(ii) *Dynamic random geometric graphs (RGG).* It would be worthwhile to use our general results to establish the limit theory for statistics of dynamic graphs which evolve according to rules depending on the local geometry. Until recently the study of statistics of the random geometric graph has been confined to the static case. Under appropriate conditions, our general set-up should provide the limit theory for statistics of the RGG on dynamic point sets, including those given by interacting particle system models. Alternatively, one could potentially use Theorem 5.2 to establish the limit theory for statistics of the RGG when the input  $\mathcal{P}$  undergoes dynamic re-positioning. The points in  $\mathcal{P}$  are allowed to undergo perturbations which may depend on nearby point configurations. This gives rise to non-static RGG which play a role in modelling. They are used in the study of mobile networks, a topic considered by Díaz et al. [40], as well as the case of mobile random geometric graphs as considered

by Peres et al. [124]. More recently, they also feature in modelling SIS epidemics, particularly when one is interested in understanding the contact process on dynamic point sets. See the paper of Baccelli and Ramesan [7]. It is also of interest to consider dynamic repositioning via a dependent spatial birth-death process as in Qi [125], Onaran et al. [111, 112], which prove functional central limit theorems in the time domain.

(iii) *Dynamic Voronoi tessellations.* The study of statistics of the Voronoi tessellation has been mainly confined to the static case, namely for random input which does not evolve with time. In many situations, including the study of algorithms in nearest neighbor search queries Kaplan et al. [84], sites are dynamically inserted and deleted, giving rise to Voronoi diagrams which evolve in time. The sites may be viewed as the realization of an interacting particle system. Subject to appropriate conditions on the dynamics, one may potentially use Theorem 5.2 to establish the limit theory for statistics of the Voronoi tessellation on dynamically changing point sets.

(iv) *Shot-noise and Markov random fields.* Apart from the empirical random field in Section 11.4, two other random field models constructed via marked point processes include shot-noise random fields and Markov random fields (MRF). Normal approximation of Poissonian shot-noise random fields has been investigated in Lachièze-Rey [90], Lachièze-Rey and Muirhead [92], Last et al. [98] and we expect our results in Section 5 to be useful in studying general point process counterparts. To date we are unaware of a systematic investigation of the limit theory for stabilizing statistics of MRF in the continuum. When the underlying graph is a stabilizing interaction graph as in Section 8.2, we expect that the approach described here could lead to the limit theory for general statistics of MRF, including those arising in the detection of Gauss MRF as in [29], as well as those going beyond functionals occurring in specialized settings, as in [4].

(v) *Models of population genetics.* The birth-death-migration example described in Section 10.3 can be considered as a simple evolutionary model in population genetics. Using our general framework, especially Theorem 10.1, we may potentially establish the limit theory for statistics of more advanced models on general point processes, such as the Kimura’s stepping stone model (see Cox and Durrett [36]) and the Fleming-Viot model (see Dawson and Greven [37]).

## Acknowledgements

The authors are grateful to Jesper Möller for his comments on a preliminary version of this paper. B. Błaszczyszyn and D. Yogeshwaran were partially supported by the IFCAM project “Geometric statistics of stationary point processes”. B. Błaszczyszyn research was partially supported by the European Research Council (ERC-NEMO-788851). D. Yogeshwaran’s research was partially supported by SERB-MATRICES Grant and CPDA from the Indian Statistical Institute. Part of the work was done when DY was hosted by CNRS and Laboratoire MAP5 of Université Paris Cité. He is thankful to Rajat Subhra Hazra for pointing out some references on sandpile models, to Manjunath Krishnapur for discussions on variance asymptotics, to Mathew Joseph and Kavita Ramanan for explaining various calculations regarding interacting diffusions and to Sunder Sethuraman for comments on CLT for

particle systems, to Piyush Srivastava for numerous elaborate discussions on spin systems, particularly regarding Section 8.4.1. The research of J. Yukich was partially supported by a Simons collaboration grant. He is grateful for support from the CNRS and the Institut Henri Poincaré, where part of this research was conducted. He thanks Matthias Schulte for explaining the truncation approach in the proof of Theorem 5.4. He also thanks Anima Anandkumar for discussions and who pointed out long ago relationships between weak spatial mixing and decay of correlations.

## Appendices

### A Mixing point processes and interaction graphs in our framework

We provide examples of point processes satisfying specific assumptions underlying our main theoretical results in Part II and also the study of some applications in Part III. These include the assumptions of summable exponential mixing correlations as in Definition 4.8, which is stronger than the merely fast mixing correlations of Definition 3.1, as well as bounds on the Palm correlation functions (8.10) and (10.10), which are stronger than the ‘usual’ bounds on Palm intensity functions required in Item (i) of Assumption 3.1 required throughout the general theory. Stabilizing interaction graphs, introduced in Section 8.1.1, are employed to demonstrate the localization or stabilization of marking functions in the models considered in Sections 8, 9, and 10.

#### A.1 Mixing point processes

As indicated after Definition 4.8, marked point processes  $\tilde{\mathcal{P}}$  with summable exponential mixing correlations can be constructed by combining a ground point process  $\mathcal{P}$  satisfying such an assumption together with appropriately mixing marks, such as i.i.d. marks. Hence we now provide examples of point processes  $\mathcal{P}$  with summable exponential mixing correlations. *The stationary Poisson point process is a trivial example.*

**Example A.1** (Point processes with summable exponential mixing correlations and bounded Palm intensity functions). Point processes  $\mathcal{P}$  with summable exponential mixing correlations (eqref:sum-(4.15)) include stationary  $\alpha$ -determinantal point processes ( $-1/\alpha \in \mathbb{N}$ ) with exponentially decreasing kernels, and certain Cox point processes; see [21, Section 2.2.2]. Moreover, mixing of correlations for *subcritical* Gibbs point processes is established in [13, Section 3.3.5]. Under suitable assumptions—in particular, boundedness of the Papangelou intensity and exponential decay of cluster-dependent interactions—using Theorem 3.34 therein, it can be shown that subcritical Gibbs point processes satisfy (4.14)–(4.15).

For these Gibbs point processes, the bound on Palm correlation functions in (10.10) still remains to be verified. In what follows, we establish this property for stationary  $\alpha$ -determinantal point pro-

cesses. For stationary determinantal point process (i.e.,  $\alpha = -1$ ) as follows: Using the determinantal structure, Fischer's inequality for determinants [77, Theorem 7.8.5] twice and stationarity, we have that

$$\rho^{(p+q)}([x]_1^p, [y]_1^q) \leq \rho^{(p)}([x]_1^p) \rho^{(q)}([y]_1^q) \leq \kappa_1^{p+q},$$

and now plugging this into (2.7) gives that  $\rho_{[x]_1^p}^{(q)}([y]_1^q) \leq \kappa_1^q$ . Now using [22, (1.11)], we can show that (10.10) holds for  $\alpha$ -determinantal,  $-1/\alpha \in \mathbb{N}$ , point processes as the latter is an independent superposition of  $-1/\alpha$  many determinantal point processes.

## A.2 Stabilizing interaction graphs

Recall the definition of stabilizing interaction graphs given in Definition 8.1, where the interaction range  $S_n(x, \mu)$  is defined as the radius of stabilization of the neighborhood marking function  $\bar{N}$ ; see (8.2) and Definition 4.5.

Graphs with finite-range interactions (i.e.,  $x \sim y$  implies  $|x - y| \leq r$  for some  $r > 0$ ) trivially admit the bound  $S_n(x, \mu) \leq r + 1$  (recall that the radius of stabilization is defined to take integer values) and hence satisfy (8.3) with  $\varphi'_p(s) = 0$  for  $s > r + 1$ .

We next mention some non-trivial examples. Though these graphs are standard in the stochastic geometry literature, we shall briefly describe their constructions and indicate the relevant known results that help us to verify the condition (8.3). In the three graph examples below, we use a classical argument involving conical regions. Though similar bounds are available in the literature, we give details in planar case (i.e.,  $d = 2$ ) here tailored to our stabilization definition and for self-containment. The examples explain how the requisite stabilization bounds can be derived for stationary  $\alpha$ -determinantal point processes with  $-1/\alpha \in \mathbb{N}$  and with exponentially decreasing kernels. The key to this is availability of Palm void probability bounds or concentration bounds for number counts (for example, see (A.3)) of this point process. Though we expect such bounds to be true for more point processes, we do not pursue this here. For example, Palm void probability bounds of first two orders is derived for permenantal and Gibbs point processes in [86, Section 6.3].

Recall  $W_n = [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$  and  $W_\infty = \mathbb{R}^d$ . Below we consider  $W_n, n \in \mathbb{N} \cup \{\infty\}$ .

**Example A.2** (Undirected  $k$ -nearest neighbors graph). Let  $\mathcal{G}(\mu)$  be the undirected  $k$ -nearest neighbors graph on  $\mu \in \mathcal{N}_{\mathbb{R}^d}$ , obtained by including  $x \sim y$  as an edge whenever  $y$  is one of the  $k$  nearest neighbors of  $x$  and/or  $x$  is one of the  $k$  nearest neighbors of  $y$ . We first construct an upper bound for the interaction range  $S_n(x, \mu)$  and then show that it satisfies (8.3) for stationary  $\alpha$ -determinantal point processes,  $-1/\alpha \in \mathbb{N}$ . Notice that it is not enough to let  $S_n(x, \mu)$  be the distance to the  $k$ th nearest neighbor of  $x$  in the point set  $\mu$ , as (8.7) would fail in general.

Instead, in dimension  $d = 2$  we proceed as follows. For each  $t > 0$  and  $x \in \mathbb{R}^2$ , let  $\mathcal{T}^{(j)}(x, t)$ ,  $1 \leq j \leq 6$ , denote six disjoint sectors with apex at  $x$ , radius  $t$ , and opening angle  $\pi/3$  (their common rotation will be fixed later). Let

$$\mathcal{T}^{(j)}(x) := \bigcup_{t>0} \mathcal{T}^{(j)}(x, t), \quad 1 \leq j \leq 6,$$

denote the corresponding infinite cones with apex at  $x$  and opening angle  $\pi/3$ . Furthermore, for  $n \in \mathbb{N}$ , define  $t_n^{(j)}(x)$  as the smallest radius  $t$  for which the truncated sector  $\mathcal{T}^{(j)}(x, t)$  already covers the entire intersection of the infinite cone with the window  $W_n$ , that is,

$$t_n^{(j)}(x) := \inf \left\{ t \geq 0 : (\mathcal{T}^{(j)}(x) \setminus \mathcal{T}^{(j)}(x, t)) \cap W_n = \emptyset \right\}.$$

We say that the sector  $\mathcal{T}^{(j)}(x, t)$  is *saturated on  $W_n$*  if  $t \geq t_n^{(j)}(x)$ .

For  $n \in \mathbb{N}$ ,  $x \in W_n$  and all  $1 \leq j \leq 6$ , we put

$$T_n^{(j)}(x) := \begin{cases} \inf \{ t \in [0, \infty) : |\mu \cap \mathcal{T}^{(j)}(x, t) \cap W_n| \geq k + 1 \} & \text{if } |\mu \cap \mathcal{T}^{(j)}(x) \cap W_n| \geq k + 1 \\ t_n^{(j)}(x) & \text{otherwise.} \end{cases} \quad (\text{A.1})$$

We claim that the interaction range of the  $k$ -nearest neighbors graph  $\mathcal{G}(\mu)$  satisfies for  $x \in W_n$

$$S_n(x, \mu) \leq 2 \max_{j=1, \dots, 6} T_n^{(j)}(x, \mu) + 1. \quad (\text{A.2})$$

Indeed, let  $T := \max_{j=1, \dots, 6} T_n^{(j)}(x, \mu)$ , and observe that if  $T = \max_{j=1, \dots, 6} t_n^{(j)}(x)$ , then  $2T + 1 \geq \lceil \text{diam}(W_n) \rceil$  and hence trivially  $S_n \leq 2T + 1$ .

In the opposite case, when  $T < \max_{j=1, \dots, 6} t_n^{(j)}(x)$ , there exist sectors  $\mathcal{T}^{(j)}(x, T)$  which are not saturated on  $W_n$ . We restrict attention to modifications of the configuration  $\mu$  in the regions of  $W_n$  not covered by these sectors, and examine how such modifications may affect the edges incident to  $x$  in the graph  $\mathcal{G}(\mu)$ , assuming  $\mu \subset W_n$ .

Observe first that all *primary-type neighbors* of  $x$  in  $\mathcal{G}(\mu)$ —those which are among  $k$ -nearest neighbors of  $x$ —are, clearly, in the ball  $B_x(T) \cap W_n$ . Moreover, this ball contains also all points of  $\mu$  for which  $x$  is among their  $k$ -nearest neighbors; we call such points *secondary-type neighbors* of  $x$  in  $\mathcal{G}(\mu)$ . Indeed, consider a point

$$y \in W_n \cap (\mathcal{T}^{(j)}(x) \setminus B_x(T))$$

for some sector  $j$  for which the intersection above is not empty. Consider the open ball  $B_y^\circ(|x - y|)$  centered at  $y$  with radius  $|x - y|$ . This ball completely covers the sector  $\mathcal{T}^{(j)}(x, |x - y|)$  except its center  $x$ , and consequently

$$B_y^\circ(|x - y|) \supseteq \mathcal{T}^{(j)}(x, |x - y|) \setminus \{x\} \supseteq \mathcal{T}^{(j)}(x, T) \setminus \{x\} \supseteq \mathcal{T}^{(j)}(x, T_n^{(j)}(x, \mu)) \setminus \{x\}.$$

By the definition of  $T_n^{(j)}(x, \mu)$ , the smallest set in the above chain of inclusions contains (exactly)  $k$  points of  $\mu$ . Consequently,  $B_y^\circ(|x - y|) \setminus \{y\}$  contains at least  $k$  points of  $\mu$  closer to  $y$  than  $x$ , so  $x$  cannot be among the  $k$  nearest neighbors of  $y$ . Hence,  $y$  cannot be a secondary-type neighbor of  $x$  in  $\mathcal{G}(\mu)$ .

It follows that adding points to  $\mu$  in  $W_n$  outside  $B_x(T)$  cannot create new edges incident to  $x$  in  $\mathcal{G}(\mu)$ . However, adding points outside  $B_x(T)$  may destroy existing edges incident to  $x$  by ‘attracting’ points that were previously secondary-type neighbors of  $x$ , and removing points there may create new edges incident to  $x$  by the inverse mechanism. Nevertheless, neither of these effects is possible for

modifications of  $\mu$  outside  $B_x(2T)$ . Indeed, any point that could become, or cease to be, a secondary-type neighbor of  $x$  in  $\mu$  must lie in  $B_x(T)$ , and is therefore closer to  $x$  than any point  $y \in B_x^c(2T)$ . Consequently, changes to the configuration outside  $B_x(2T)$  cannot influence the adjacency relations of  $x$  in  $\mathcal{G}(\mu)$ . Finally, adding 1 in (A.2) ensures that  $2T + 1$  bounds the interaction range  $S_n(x, \mu)$  of the graph  $\mathcal{G}(\mu)$ , where  $S_n(x, \mu)$  is the integer radius of stabilization as in (8.2) and Definition 4.5.

We now establish the probabilistic bound (8.3) under the assumption that  $\mathcal{P}$  is a stationary  $\alpha$ -determinantal point process with an exponentially decreasing kernel. The first steps of the argument are general, whereas the later, model-specific steps can be adapted to other classes of input point processes along similar lines.

From the previous considerations, for  $t \geq 0$  we have the bound,

$$\mathbb{P}_{[x]_1^p}(S_n(x_1, \mathcal{P}_n) > t) \leq \mathbb{P}_{[x]_1^p}\left(\max_{j=1, \dots, 6} T_n^{(j)}(x_1, \mathcal{P}_n) > (t-1)/2\right) \leq \sum_{j=1}^6 \mathbb{P}_{[x]_1^p}(T_n^{(j)}(x_1, \mathcal{P}_n) > (t-1)/2).$$

Moreover,

$$\mathbb{P}_{[x]_1^p}(T_n^{(j)}(x_1, \mathcal{P}_n) > t) = \begin{cases} \mathbb{P}_{[x]_1^p}(\mathcal{P}(\mathcal{T}^{(j)}(x_1, t) \cap W_n) \leq k) & \text{for } t \leq t_n^{(j)}(x_1) \\ 0 & \text{for } t > t_n^{(j)}(x_1). \end{cases} \quad (\text{A.3})$$

Now we fix the orientation of the six-sector partition so that the sector boundaries are inclined at an angle of  $\pi/12$  relative to the coordinate axes. With this choice, there exists a constant  $c_0 = \pi/24$  such that, uniformly in  $n \in \mathbb{N}$ ,  $x \in W_n$ , and  $j = 1, \dots, 6$ ,

$$|\mathcal{T}^{(j)}(x, t) \cap W_n| \geq c_0 t^2, \quad \text{for all } t < t_n^{(j)}(x).$$

Then the probability in (A.3) admits a fast decreasing upper bound as a function of  $t$ , which follows from standard concentration inequalities (such as Chernoff bounds). This is because the number  $k$  is fixed, whereas the volume of the sets  $\mathcal{T}^{(j)}(x, t) \cap W_n$  grows with  $t$ .

A further technical point is to ensure that these bounds hold uniformly with respect to the Palm conditioning at  $[x]_1^p$ . Such uniformity is established, for instance, for stationary  $\alpha$ -determinantal point processes with exponentially decaying kernels; see [21, Theorem 2.5] and [22, Corollary 1.10]. This completes the proof of stabilization of the  $k$ -nearest-neighbor graph in the plane with respect to the aforementioned input processes; i.e., it satisfies conditions of Definition 8.1.

A similar geometric construction applies in dimensions  $d \geq 3$ , with planar sectors replaced by cones. Hence,  $\mathcal{G}(\cdot)$  is a stabilizing interaction graph on finite windows in higher dimensions as well, provided that the above concentration bounds for the Palm distributions of the input process  $\mathcal{P}$  hold.

**Example A.3** (Delaunay graph). The Voronoi cell associated with  $x \in \mu \in \mathcal{N}_{\mathbb{R}^d}$  is the set  $C(x, \mu) := \{y \in \mathbb{R}^d : |x - y| \leq |z - y|, \text{ for all } z \in \mu, z \neq x\}$ . The point  $x$  is called the generator of the cell  $C(x, \mu)$ . The Voronoi tessellation of  $\mathbb{R}^d$  induced by  $\mu$  is the collection of cells  $\{C(x, \mu)\}_{x \in \mu}$ . Given  $\mu \in \mathcal{N}_{\mathbb{R}^d}$  with  $\mu$  in general position, let  $\mathcal{G}(\mu)$  be the Delaunay graph on  $\mu \in \mathcal{N}_{\mathbb{R}^d}$ , i.e.  $x \sim y$  iff  $C(x, \mu) \cap C(y, \mu) \neq \emptyset$ .

It is well known (indeed, this is the original definition of Delaunay [38], based on the empty circumsphere property) that  $x \sim y$  in  $\mathcal{G}(\mu)$  if and only if there exist points  $z_1, \dots, z_{d-1} \in \mu$ , all distinct from  $x$  and  $y$  and pairwise distinct, such that the open ball  $B^\circ(x, y, z_1, \dots, z_{d-1})$  circumscribed about the points  $(x, y, z_1, \dots, z_{d-1})$  contains no further points of  $\mu$ , that is,

$$\mu(B^\circ(x, y, z_1, \dots, z_{d-1})) = 0.$$

Using this characterization, one can bound the interaction range  $S_n(x, \mu)$ . We illustrate this argument in dimension  $d = 2$  as before. Taking  $T_n^{(j)}(x, \mu)$  as given in (A.1) with  $k = 1$ , we can show that

$$S_n(x, \mu) \leq 2T := 2 \max_{j=1, \dots, 6} T_n^{(j)}(x, \mu). \quad (\text{A.4})$$

Indeed, let  $y \in \mathbb{R}^2$  satisfy

$$|x - y| > 2T,$$

with  $T := \max_{j=1, \dots, 6} T_n^{(j)}(x, \mu)$ . Then, for any ball having both  $x$  and  $y$  on its boundary, the corresponding radius  $r$  must satisfy  $r > T$ , and hence the center  $w$  must satisfy  $|w - x| > T$ . Using the same geometric argument as in Example A.2, the open ball  $B_w^\circ(r)$  contains the sector

$$\mathcal{T}^{(j)}(x, T_n^{(j)}(x)) \setminus \{x\} \subset B_w^\circ(r),$$

for the index  $j$  such that  $w \in \mathcal{T}^{(j)}(x)$ . By the definition of  $T_n^{(j)}(x)$ , we have

$$\mu\left(\mathcal{T}^{(j)}(x, T_n^{(j)}(x)) \setminus \{x\}\right) = 1,$$

which implies that the open ball  $B_w^\circ(r)$  is nonempty with respect to the points of  $\mu$ . This rules out  $y$  as a neighbor of  $x$  in the Delaunay graph on  $\mu$  and thereby verifying (A.4). This argument can be extended to  $d \geq 3$ .

Following verbatim the discussion in Example A.2, it follows that  $S_n(x, \mathcal{P})$  satisfies (8.3) if  $\mathcal{P}$  is a stationary  $\alpha$ -determinantal point process,  $-1/\alpha \in \mathbb{N}$ , with an exponentially decaying kernel. Consequently,  $\mathcal{G}(\mathcal{P})$  is a stabilizing graph on finite windows with respect to  $\mathcal{P}$  for such processes.

**Example A.4** (Sphere of influence graph). Given a point set  $\mu \in \mathcal{N}_{\mathbb{R}^d}$ , the sphere of influence graph  $\mathcal{G}(\mu)$  is constructed as follows. For each  $x \in \mu$ , let  $B_{\text{NN}}(x)$  be the closed ball centered at  $x$  with radius equal to the distance between  $x$  and its nearest neighbor in  $\mu$ . The sphere of influence graph puts an edge between  $x$  and  $y$  iff  $B_{\text{NN}}(x) \cap B_{\text{NN}}(y) \neq \emptyset$ . To demonstrate the stabilization of this graph on some interesting input processes, we need an upper bound for  $S_n$ , for which it again suffices to modify Example A.2

More precisely, in dimension  $d = 2$ , for  $x \in \mathbb{R}^2$ , we consider the same six disjoint sectors  $\mathcal{T}^{(j)}(x, t)$ ,  $j = 1, \dots, 6$  with apex at  $x$ , radius  $t$ , and opening angle  $\pi/3$  (their common rotation will be fixed later). As before, let their saturation radii on the windows  $W_n$  be  $t_n^{(j)}(x)$ . For  $n \in \mathbb{N}$ ,

$x \in W_n$ , and  $1 \leq j \leq 6$ , we define

$$T_n^{(j)}(x) := \begin{cases} \min \left\{ k \in \mathbb{N} : 1 \leq k \leq t_n^{(j)}(x), |\mu \cap \mathcal{T}^{(j)}(x, k/3) \cap W_n| \geq 1, \right. \\ \quad \left. \text{and } |\mu \cap (\mathcal{T}^{(j)}(x, k) \setminus \mathcal{T}^{(j)}(x, 2k/3)) \cap W_n| \geq 1 \right\}, & \text{if this set is nonempty,} \\ t_n^{(j)}(x), & \text{otherwise.} \end{cases} \quad (\text{A.5})$$

In other words,  $T_n^{(j)}(x) = k$  is the smallest integer such that, within the sector  $\mathcal{T}^{(j)}(x, k)$ , there is at least one point of  $\mu$  in both the inner part  $\mathcal{T}^{(j)}(x, k/3)$  and the outer annular part  $\mathcal{T}^{(j)}(x, k) \setminus \mathcal{T}^{(j)}(x, 2k/3)$  in  $W_n$ . If no such  $k$  exists, we set  $T_n^{(j)}(x)$  equal to the saturation radius  $t_n^{(j)}(x)$  of this sector on  $W_n$ .

We claim that the interaction range of the sphere of influence graph  $\mathcal{G}(\mu)$  satisfies for  $x \in W_n$

$$S_n(x, \mu) \leq \max_{j=1, \dots, 6} T_n^{(j)}(x, \mu) + 1. \quad (\text{A.6})$$

Indeed, the presence of a point in the inner part of each sector  $\mathcal{T}^{(j)}(x, k)$ ,  $j = 1, \dots, 6$ , when  $\max_{j=1, \dots, 6} T_n^{(j)}(x) = k$ , yields that  $B_{NN}(x) \subset B_{k/3}(x)$ . On the other hand, the presence of points in the outer parts of the sectors  $\mathcal{T}^{(j)}(x, T_n^{(j)}(x))$  ensures that, for any point  $y$  with  $|y - x| > k$ , we have that  $B_{NN}(y) \subset B_{k/3}^c(x)$  i.e., the corresponding ball  $B_{NN}(y)$  does not intersect  $B_{NN}(x)$ .

To establish the probabilistic bound (8.3), we observe again for  $t \geq 0$

$$\mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}(S_n(x_1, \mathcal{P}_n) > t) \leq \sum_{j=1}^6 \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}(T_n^{(j)}(x_1, \mathcal{P}_n) > t - 1)$$

and for  $t \geq 1$

$$\begin{aligned} & \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}(T_n^{(j)}(x_1, \mathcal{P}_n) > t) \\ & \leq \begin{cases} \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}\left(\bigcap_{k \in \mathbb{N} \cap [1, t]} \left\{ \mathcal{P}(\mathcal{T}^{(j)}(x_1, k/3) \cap W_n) = 0 \right. \right. \\ \quad \left. \left. \text{or } \mathcal{P}((\mathcal{T}^{(j)}(x_1, k) \setminus \mathcal{T}^{(j)}(x_1, 2k/3)) \cap W_n) = 0 \right\}\right) & \text{for } t \leq t_n^{(j)}(x_1) \\ 0 & \text{for } t > t_n^{(j)}(x_1). \end{cases} \quad (\text{A.7}) \end{aligned}$$

Finally

$$\begin{aligned} & \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}\left(\bigcap_{k \in \mathbb{N} \cap [1, t]} \left\{ \mathcal{P}(\mathcal{T}^{(j)}(x_1, k/3) \cap W_n) = 0 \text{ or } \mathcal{P}((\mathcal{T}^{(j)}(x_1, k) \setminus \mathcal{T}^{(j)}(x_1, 2k/3)) \cap W_n) = 0 \right\}\right) \\ & \leq \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}(\mathcal{P}(\mathcal{T}^{(j)}(x_1, \lfloor t \rfloor / 3) \cap W_n) = 0) + \mathbb{P}_{[\mathfrak{z}]_1^{\mathcal{P}}}(\mathcal{P}((\mathcal{T}^{(j)}(x_1, \lfloor t \rfloor) \setminus \mathcal{T}^{(j)}(x_1, 2\lfloor t \rfloor / 3)) \cap W_n) = 0). \end{aligned}$$

Following the discussion in Example A.2, the probability in (A.7) admits a fast decreasing upper bound as a function of  $t$  for input processes  $\mathcal{P}$  whose Palm distributions satisfy standard concentration inequalities. This includes, in particular, stationary  $\alpha$ -determinantal point processes,  $-1/\alpha \in \mathbb{N}$ , with exponentially decaying kernels.

Consequently, the sphere of influence graph  $\mathcal{G}(\cdot)$  on the plane is a stabilizing graph on finite windows with respect to such processes  $\mathcal{P}$ , for instance for stationary  $\alpha$ -determinantal point processes with exponentially decaying kernels. A similar geometric construction applies in dimensions  $d \geq 3$ .

## B Some consequences of stabilization via stopping sets

In this appendix, we develop some consequences of stopping-set stabilization, as defined in Section 4.3. Though our results mainly use the weaker notion of BL-localization, some of our results may be strengthened under the stronger assumption of stopping-set stabilization. Stopping-set stabilization is satisfied by many statistics of spatial random models (for example, see Sections 10, 11.3). As it is central to many applications we therefore include formal statements and proofs of these properties.

For  $\tilde{\mu} \in \mathcal{N}_{\mathbb{R}^d \times K}$  and  $x \in \tilde{\mu}$ , we put

$$R_{\infty}^{\xi}(\tilde{x}, \tilde{\mu}) := \limsup_{n \rightarrow \infty} R_{W_n}^{\xi}(\tilde{x}, \tilde{\mu}_n)$$

and when  $R_{\infty}^{\xi}(\tilde{x}, \tilde{\mu}) < \infty$ , we define the *infinite input score*

$$\xi_{\infty}(\tilde{x}, \tilde{\mu}) := \xi(\tilde{x}, \tilde{\mu} \cap B_{R_{\infty}^{\xi}(\tilde{x}, \tilde{\mu})}(x)). \quad (\text{B.1})$$

The following result justifies defining input scores via (B.1) and provides additional properties of  $\xi_{\infty}$  and  $R_{\infty}^{\xi}$ .

**Lemma B.1** (Properties of the infinite input score  $\xi_{\infty}(\tilde{x}, \tilde{\mathcal{P}})$ ). *Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a stabilizing marking function satisfying almost surely  $R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}) < \infty$  for all  $\tilde{x} \in \tilde{\mathcal{P}}$  and consider  $\xi_{\infty}$  as defined in (B.1). The following statements hold.*

(i) *There exist (finite) random variables  $n_0 := n_0(\tilde{x}, \tilde{\mathcal{P}})$  such that for all  $k \geq n_0$  we have  $R_{W_k}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_k) = R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}})$  and  $\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = \xi_{\infty}(\tilde{x}, \tilde{\mathcal{P}})$ . In particular, a.s. for all  $\tilde{x} \in \tilde{\mathcal{P}}$ , we have that  $\xi_{\infty}(\tilde{x}, \tilde{\mathcal{P}}) = \lim_{n \rightarrow \infty} \xi(\tilde{x}, \tilde{\mathcal{P}}_n)$  i.e., the equivalence (5.16) holds.*

(ii)  *$R_{\infty}^{\xi}$  is a.s. an upper bound for the radius of stabilization for  $\xi_{\infty}$  as in Definition 4.5.*

(iii) *If  $\xi$  satisfies the  $p$ -moment condition on finite windows (3.4) then (3.4) also holds for  $\xi_{\infty}$  (i.e., replacing  $\xi_{i,n}$  by  $\xi_{\infty}(\tilde{x}_i, \tilde{\mathcal{P}})$ ) with the same moment bound  $M_p^{\xi}$ . For this implication to hold, it is enough that  $\xi_{\infty}$  is defined as a limit as in (5.16), without guaranteeing that  $R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}) < \infty$ .*

(iv) *If  $\xi$  is translation invariant then is so  $R_{\infty}^{\xi}$  and hence  $\xi_{\infty}$ .*

*Suppose now that  $\xi$  is defined on  $\mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$  and satisfies stopping-set stabilization on  $\tilde{\mathcal{P}}$  (4.11) with  $p = 1$ . Then*

(v)  *$R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}) < \infty$  for all  $\tilde{x} \in \tilde{\mathcal{P}}$ , a.s.. Consequently all properties (i)–(iv) hold and moreover  $R_{\infty}^{\xi} = R^{\xi}$  and  $\xi_{\infty} = \xi$  for all  $\tilde{x} \in \tilde{\mathcal{P}}$ , a.s.*

*Proof.* (i) By assumption, almost surely for all  $\tilde{x} \in \tilde{\mathcal{P}}$ ,  $\limsup_{n \rightarrow \infty} R_{W_n}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_n) = R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}) < \infty$  and hence there exists a finite random variable  $n_0 := n_0(\tilde{x}, \tilde{\mathcal{P}})$  such that  $R_{W_{n_0}}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_{\infty}^{\xi}(\tilde{x}, \tilde{\mathcal{P}})$  (remember, the stabilizing radii are defined as integers),  $B_{R_{W_{n_0}}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_{n_0})}(x) \subset W_{n_0}$  and for all  $k \geq n_0$  we have  $R_{W_k}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_k) \leq R_{W_{n_0}}^{\xi}(\tilde{x}, \tilde{\mathcal{P}}_{n_0})$ . Consequently, for  $k \geq n_0$ , by the stopping property of

$R_{W_{n_0}}^\xi, R_{W_k}^\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = R_{W_{n_0}}^\xi(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}})$  and consequently  $\xi_\infty(\tilde{x}, \tilde{\mathcal{P}}) = \xi(\tilde{x}, \tilde{\mathcal{P}}_k) = \xi(\tilde{x}, \tilde{\mathcal{P}} \cap B_{R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}})}(x))$  guaranteeing also the validity of (5.16).

(ii) We begin by noting the following general consequence of the stopping property (4.10), which holds for all  $k \in \mathbb{N}$ ,  $\tilde{\mu} \in \hat{\mathcal{N}}_{W_k \times \mathbb{K}}$ , and  $\tilde{x} \in \tilde{\mu}$ :

$$R_k(\tilde{x}, \tilde{\mu}) = R_k\left(\tilde{x}, (\tilde{\mu} \cap B_{R_k(\tilde{x}, \tilde{\mu})}(x)) \cup (B_{R_k(\tilde{x}, \tilde{\mu})}^c(x) \cap \tilde{\nu})\right) \quad (\text{B.2})$$

for all  $\tilde{\nu} \in \hat{\mathcal{N}}_{W_k \times \mathbb{K}}$ , where for  $k \in \mathbb{N} \cup \infty$  we make the notational simplification  $R_k := R_{W_k}^\xi$ . Hence, the stabilization radius  $R_k$  stabilizes itself, which is a direct consequence of defining  $R$  as the *smallest*  $r \in \mathbb{N}$  satisfying (4.8).

To prove that  $R_\infty := R_\infty^\xi$  is an upper bound for the radius of stabilization of  $\xi_\infty$ , we need to justify the equality

$$\xi_\infty\left(\tilde{x}, (\tilde{\mathcal{P}} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}(x))\right) = \xi_\infty\left(\tilde{x}, (\tilde{\mathcal{P}} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}(x)) \cup (\tilde{\nu} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}^c(x))\right) \quad (\text{B.3})$$

for almost all input  $\tilde{\mathcal{P}}$  and all (possibly infinite)  $\tilde{\nu} \in \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ . In this regard, denote the point process arguments on both sides of (B.3) respectively by  $\mathcal{Q}_1 := \tilde{\mathcal{P}} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}(x)$  and  $\mathcal{Q}_2 := (\tilde{\mathcal{P}} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}(x)) \cup (\tilde{\nu} \cap B_{R_\infty(\tilde{x}, \tilde{\mathcal{P}})}^c(x))$ . With the definition of  $\xi_\infty$  at (B.1) and this notation, equality (B.3) is equivalent to

$$\xi(\tilde{x}, \mathcal{Q}_1 \cap B_{R_\infty(\tilde{x}, \mathcal{Q}_1)}) = \xi(\tilde{x}, \mathcal{Q}_2 \cap B_{R_\infty(\tilde{x}, \mathcal{Q}_2)}), \quad (\text{B.4})$$

which is itself implied by

$$\mathcal{Q}_1 \cap B_{R_\infty(\tilde{x}, \mathcal{Q}_1)} = \mathcal{Q}_2 \cap B_{R_\infty(\tilde{x}, \mathcal{Q}_2)} = \tilde{\mathcal{P}} \cap B_{R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}})}. \quad (\text{B.5})$$

To prove this last equality we study first the function  $R_\infty(\tilde{x}, \cdot) := \limsup_{n \rightarrow \infty} R_n(\tilde{x}, \cdot)$  applied to  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ . As in point (i) above, we take  $n_0 := n_0(\tilde{x}, \tilde{\mathcal{P}})$  such that  $R_{n_0}(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_\infty(\tilde{x}, \tilde{\mathcal{P}})$ ,  $B_{R_{n_0}(\tilde{x}, \tilde{\mathcal{P}}_{n_0})}(x) \subset W_{n_0}$ , and for all  $k \geq n_0$  we have  $R_k(\tilde{x}, \tilde{\mathcal{P}}_k) \leq R_{n_0}(\tilde{x}, \tilde{\mathcal{P}}_{n_0})$ . Then, for  $k > n_0$ , by  $\mathcal{Q}_1 \in \hat{\mathcal{N}}_{W_k \times \mathbb{K}}$ , we have

$$R_k(\tilde{x}, \mathcal{Q}_1) = R_k(\tilde{x}, \tilde{\mathcal{P}}_k) \leq R_{n_0}(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_\infty(\tilde{x}, \tilde{\mathcal{P}}),$$

where we used (B.2) for the first equality. Since there are infinitely many values  $k \in \mathbb{N}$  such that  $R_k(\tilde{x}, \tilde{\mathcal{P}}_k) = R_{n_0}(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_\infty(\tilde{x}, \tilde{\mathcal{P}})$  we have  $R_\infty(\tilde{x}, \mathcal{Q}_1) = R_\infty(\tilde{x}, \tilde{\mathcal{P}})$ .

Similarly,  $R_\infty(\tilde{x}, \mathcal{Q}_2) = R_\infty(\tilde{x}, \tilde{\mathcal{P}})$  given that  $\tilde{\nu}$  is finite. Indeed, in this case, it suffices to increase the value of  $k$  used for  $\mathcal{Q}_1$ , if necessary, choosing it such that  $\tilde{\nu} \in \hat{\mathcal{N}}_{W_k \times \mathbb{K}}$ . With the definition of  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ , this justifies (B.5) for finite  $\tilde{\nu}$ .

Suppose now  $\tilde{\nu}$  is infinite. Then recall  $R_\infty(\tilde{x}, \mathcal{Q}_2) = \limsup_{n \rightarrow \infty} R_n(\tilde{x}, \mathcal{Q}_2 \cap W_n)$ . For all large enough  $n \in \mathbb{N}$ , by the previous argument, we have  $R_\infty(\tilde{x}, \mathcal{Q}_2 \cap W_n) = R_\infty(\tilde{x}, \tilde{\mathcal{P}})$ , which establishes (B.5) for  $\tilde{\nu}$  infinite and thus concludes the proof that  $R_\infty$  is an upper bound for the radius of stabilization of  $\xi_\infty$ .

(iii) For  $1 \leq q \leq p$ , consider a bounded, non-negative function  $h : \mathbb{R}^q \rightarrow [0, \infty)$  with bounded support. Using the Campbell-Little-Mecke formula for  $\xi_\infty$  and  $\xi$ , we have

$$\begin{aligned}
& \int_{\mathbb{R}^q} h([x]_1^l) \mathbb{E}_{[x]_1^l} [\max(1, |\xi_\infty(\tilde{x}_1, \tilde{\mathcal{P}})|^q)] \rho^{(q)}([x]_1^l) d[x]_1^l \\
&= \mathbb{E} \left[ \sum_{\text{distinct } [x]_1^l \subset \mathcal{P}} h([x]_1^l) \max(1, |\xi_\infty(\tilde{X}_1, \tilde{\mathcal{P}})|^q) \right] \\
&= \mathbb{E} \left[ \sum_{\text{distinct } [x]_1^l \subset \mathcal{P}} h([x]_1^l) \lim_{n \rightarrow \infty} \max(1, |\xi(\tilde{X}_1, \tilde{\mathcal{P}}_n)|^q) \right] \\
&\leq \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \sum_{\text{distinct } [x]_1^l \subset \mathcal{P}} h([x]_1^l) \max(1, |\xi(\tilde{X}_1, \tilde{\mathcal{P}}_n)|^q) \right] \\
&= \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^q} h([x]_1^l) \mathbb{E}_{[x]_1^l} [\max(1, |\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n)|^q)] \rho^{(q)}([x]_1^l) d[x]_1^l \\
&\leq M_p^\xi \int_{\mathbb{R}^q} h([x]_1^l) \rho^{(q)}([x]_1^l) d[x]_1^l
\end{aligned}$$

where the first inequality holds by Fatou's lemma and the last one by the  $p$ -moment assumption (3.4) for  $\xi$  on finite windows. This shows that  $\xi_\infty$  satisfies the  $p$ -moment condition, provided it is almost surely well defined as the limit (or even  $\liminf$ ) in (5.16).

(iv) Assume  $\xi$  is translation invariant and consider  $a \in \mathbb{R}^d$ . Note, in general, for some  $n$  we might have  $R_{W_n}^\xi(\tilde{x}, \tilde{\mathcal{P}} \cap W_n) \neq R_{W_n}^\xi(\tilde{x} + a, (\tilde{\mathcal{P}} + a) \cap W_n)$ . However, when both radii  $R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}})$  and  $R_\infty^\xi(\tilde{x} + a, \tilde{\mathcal{P}} + a)$  are finite (which holds a.s.) then for all  $k$  large enough  $R_{W_k}^\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = R_{W_{n_0}}^\xi(\tilde{x}, \tilde{\mathcal{P}}_{n_0}) = R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}})$  and  $R_{W_k}^\xi(\tilde{x} + a, (\tilde{\mathcal{P}} + a) \cap W_k) = R_{W_{n_0^a}}^\xi(\tilde{x} + a, (\tilde{\mathcal{P}} + a) \cap W_{n_0^a}) = R_\infty^\xi(\tilde{x} + a, \tilde{\mathcal{P}} + a)$ , with  $n_0 = n_0(\tilde{x}, \tilde{\mathcal{P}})$  and  $n_0^a := n_0(\tilde{x} + a, \tilde{\mathcal{P}} + a)$  as in point (i). Moreover,  $B_{R_{W_{n_0^a}}^\xi(\tilde{x} + a, \tilde{\mathcal{P}} \cap W_{n_0^a})}(x) \subset W_k$  and  $B_{R_{W_{n_0}}^\xi(\tilde{x}, \tilde{\mathcal{P}}_{n_0})}(x + a) \subset W_k$ . Then, by the stopping property of  $R_{W_k}^\xi(\tilde{x}, \tilde{\mathcal{P}}_k)$  and  $R_{W_k}^\xi(\tilde{x} + a, \tilde{\mathcal{P}} + a)$  and the translation invariance of  $\xi$ , we have

$$R_{W_k}^\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = R_{W_k}^\xi(\tilde{x} + a, (\tilde{\mathcal{P}} + a) \cap W_k)$$

and, consequently  $\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = \xi(\tilde{x} + a, (\tilde{\mathcal{P}} + a) \cap W_k) = \xi_\infty(\tilde{x}, \tilde{\mathcal{P}}) = \xi_\infty(\tilde{x} + a, \tilde{\mathcal{P}} + a)$ .

(v) By (4.11) with  $p = 1$  and the Campbell-Little-Mecke formula, we have  $R^\xi(\tilde{x}, \tilde{\mathcal{P}}) < \infty$  a.s. for all  $\tilde{x} \in \tilde{\mathcal{P}}$  under  $\mathbb{P}$ . Consequently for some finite (random variable)  $n_0 := n_0(\tilde{x}, \tilde{\mathcal{P}})$ , we have  $B_{R^\xi(\tilde{x}, \tilde{\mathcal{P}})}(x) \subset W_{n_0}$  and hence, for  $k \geq n_0$ , a.s. we have  $R_{W_k}^\xi(\tilde{x}, \tilde{\mathcal{P}}_k) \leq R^\xi(\tilde{x}, \tilde{\mathcal{P}}) < \infty$ , and consequently  $\xi(\tilde{x}, \tilde{\mathcal{P}}_k) = \xi(\tilde{x}, \tilde{\mathcal{P}})$ . Thus Items (i)-(iv) follow as well as the identities  $R_\infty^\xi(\tilde{x}, \tilde{\mathcal{P}}) = R^\xi(\tilde{x}, \tilde{\mathcal{P}})$ ,  $\xi_\infty(\tilde{x}, \tilde{\mathcal{P}}) = \xi(\tilde{x}, \tilde{\mathcal{P}})$  for all  $\tilde{x} \in \tilde{\mathcal{P}}$ , a.s..  $\square$

Finally, we summarize the observations regarding the relationship between stopping set stabilization and BL-localization. The main message is that the former implies the latter.

**Lemma B.2** (Stopping set stabilization implies BL-localization). *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ . Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a marking function.*

(i) If  $\xi$  is stabilizing (resp. fast stabilizing) on finite windows as in Definition 4.6(ii) (resp. (iii)) then  $\xi$  is BL-localizing (resp. fast BL-localizing) on finite windows as in Definition 4.2(ii) (resp. (iii)). The same implications hold if  $\xi$  is defined on the whole space  $\mathbb{R}^d \times \mathbb{K} \times \mathcal{N}_{\mathbb{R}^d \times \mathbb{K}}$ , involving the stopping set Definition 4.6(i) and the BL stabilizing Definition 4.2 (i).

(ii) If  $\xi$  is stabilizing (resp. fast stabilizing) on finite windows and if moreover

$$\tilde{\xi}_\infty(\tilde{x}_i, \tilde{\mathcal{P}}) := \lim_{n \rightarrow \infty} \xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$$

exists,  $\mathbb{P}_{x_1, \dots, x_p}$ -a.s. for all  $p \in \mathbb{N}$  and all  $x_1, \dots, x_p \in \mathbb{R}^d$  then  $\tilde{\xi}_\infty$  is BL-localizing (resp. fast BL-localizing) as in (4.2).

The main thrust of Item (ii) above is that BL-localization of the limiting score can be deduced without imposing finiteness of the radius of stabilization  $R_\infty^\xi$ . This offers the advantage that we can derive limit theorems for fast (stopping-set) stabilizing score functions on finite windows without requiring uniform bounds on the tail of the stabilization radii.

*Proof.* (i) The arguments are given in the second Item of Remark 4.7.

(ii) Fix  $p \in \mathbb{N}$ ,  $x_1, \dots, x_p \in \mathbb{R}^d$  and  $f \in \text{BL}(\mathbb{R}^p)$ . Without loss of generality, let  $r \geq 1$  and choose  $n > (2(r + \max_{i=1, \dots, p} |x_i|))^d$ . Thus we have that  $\mathcal{P}_n \cap B_r([x]_1^p) = \mathcal{P} \cap B_r([x]_1^p)$ . Using the fast BL-localization of  $\xi$  and the equivalence  $\tilde{\xi}_\infty = \xi$  on finite point-sets, we obtain

$$\begin{aligned} & \left| \mathbb{E}_{[x]_1^p} [f(\tilde{\xi}_\infty^p_1(\tilde{x}, \tilde{\mathcal{P}}))] - \mathbb{E}_{[x]_1^p} [f(\tilde{\xi}_\infty^p_1(\tilde{x}, \tilde{\mathcal{P}} \cap B_r([x]_1^p)))] \right| \\ & \leq \left| \mathbb{E}_{[x]_1^p} [f(\tilde{\xi}_\infty^p_1(\tilde{x}, \tilde{\mathcal{P}}))] - \mathbb{E}_{[x]_1^p} [f(\xi^p_1(\tilde{x}, \tilde{\mathcal{P}}_n))] \right| \\ & \quad + \left| \mathbb{E}_{[x]_1^p} [f(\tilde{\xi}_\infty^p_1(\tilde{x}, \tilde{\mathcal{P}}_n))] - \mathbb{E}_{[x]_1^p} [f(\xi^p_1(\tilde{x}, \tilde{\mathcal{P}}_n \cap B_r([x]_1^p)))] \right| \\ & \leq \left| \mathbb{E}_{[x]_1^p} [f(\tilde{\xi}_\infty^p_1(\tilde{x}, \tilde{\mathcal{P}}))] - \mathbb{E}_{[x]_1^p} [f(\xi^p_1(\tilde{x}, \tilde{\mathcal{P}}_n))] \right| + 2\varphi_p(r). \end{aligned}$$

Letting  $n \rightarrow \infty$  in the first term and using the assumed a.s. convergence of  $\xi(\tilde{x}_i, \tilde{\mathcal{P}}_n)$  to  $\tilde{\xi}_\infty(\tilde{x}_i, \tilde{\mathcal{P}})$  for  $i = 1, \dots, p$ , we obtain fast BL-localization of  $\tilde{\xi}_\infty$ .  $\square$

**Corollary B.3.** Consider a stationary marked point process  $\tilde{\mathcal{P}}$  on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ , having summable exponential  $\mathbb{B}$ -mixing correlations as in Definition 4.8. Let  $\xi_i : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  be translation invariant score functions such that they are fast-stabilizing on finite windows of  $\tilde{\mathcal{P}}$  as in (4.12) for  $p \in \{1, 2\}$ . Assume that for some  $\varepsilon > 0$ , the  $(2 + \varepsilon)$ -moment condition (5.2) holds uniformly for  $r > 0$  large enough, for all  $\xi_i^{(r)}$ ,  $i = 1, \dots, m$ . Then the conclusions of Theorem 5.4 hold.

*Proof.* In view of the proof of Theorem 5.4, the key task here is to reduce the required Palm order to  $p \in \{1, 2\}$  in case of stabilization.

For the last statement of Theorem 5.4, as previously, we consider the linear combination  $\mu_n^\xi(f) = \sum_{i=1}^m t_i \mu_n^{\xi_i}(f)$  of the score functions  $\xi_i$  with given fixed coefficients  $t_i \in \mathbb{R}$ ,  $i = 1, \dots, m$  and for some given function  $f \in \mathbb{B}(W_1)$ . Now, instead of truncating the value of  $\mu_n^\xi(f)$  we restrict individually

the score functions  $\xi_i^{(r)}(\tilde{x}, \tilde{\mu}) := \xi_i(\tilde{x}, \tilde{\mu} \cap B_r(x))$ ,  $i = 1, \dots, m$ , for  $r > 0$ , and consider their differences  $\Delta\xi_i^{(r)} := \xi_i - \xi_i^{(r)}$ . Furthermore, we consider the respective linear combinations  $\xi^{(r)} := \sum_{i=1}^m t_i \xi_i^{(r)}$ ,  $\Delta\xi^{(r)} := \sum_{i=1}^m t_i \Delta\xi_i^{(r)}$ , and their integrals  $\mu_n^{\xi^{(r)}}(f) = \sum_{i=1}^m t_i \mu_n^{\xi_i^{(r)}}(f)$  and  $\mu_n^{\Delta\xi^{(r)}}(f) = \sum_{i=1}^m t_i \mu_n^{\Delta\xi_i^{(r)}}(f)$ . Besides the limit (7.36) for the variance of  $\mu_n^{\xi}(f)$ , which exists as before (under fast stabilization—hence fast BL-localization—under Palm distributions of order  $p = \{1, 2\}$ ), we consider two other limits:

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^{\xi^{(r)}}(f) &=: \sigma_r^2, \\ \lim_{n \rightarrow \infty} n^{-1} \text{Var} \mu_n^{\Delta\xi^{(r)}}(f) &=: v_r^2. \end{aligned}$$

These limits exist too, because the fast stabilization of  $\xi_i$  of order  $p = \{1, 2\}$  implies stabilization of the same order of  $\xi_i^{(r)}$  and  $\Delta\xi_i^{(r)}$ . (Note, that this statement for  $\Delta\xi_i^{(r)}$  does not necessarily hold under BL-localization of  $\xi_i$ .) Moreover, since  $\xi_i^{(r)}$  is stabilizing at fixed radius  $r$  (and hence fast BL-localizing) of all Palm orders, under  $2 + \epsilon$  moment conditions for it,  $\mu_n^{\xi^{(r)}}(f)$  satisfies the CLT as  $n \rightarrow \infty$ , provided  $\sigma_r^2 > 0$ . This follows from the statement of Theorem 5.4 already proved. Now, using the same arguments developed in the proof of central limit theorem for  $(\mu_n^{\xi}(f))_{n \in \mathbb{N}}$  in Theorem 5.4 via three variance asymptotics, it remains to establish limits (7.39) and (7.40) for  $\sigma_r^2$  and  $v_r^2$  as  $r \rightarrow \infty$ .

In this regard, observe that the vector  $\xi^{(r)} := (\xi_i, \xi_i^{(r)}, \Delta\xi_i^{(r)})_{i \in \{1, \dots, m\}}$  is fast stabilizing, hence (jointly) fast BL-localizing of Palm order  $p = \{1, 2\}$  (this restriction of  $p$  is due to the assumption on  $\xi$ ). Denote by  $\Upsilon_0^{(r)}$  and  $\Upsilon_{(0,z)}^{(r)}$  the BL-limits of Palm distributions of order 1 and 2, of  $\xi^{(r)}$ :

$$\begin{aligned} \xi^{(r)}(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n) &\xrightarrow[n \rightarrow \infty]{\text{BL}, \mathbf{0}} \Upsilon_0^{(r)}, \\ \left( \xi^{(r)}(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n), \xi^{(r)}(\tilde{z}, \tilde{\mathcal{P}}_n) \right) &\xrightarrow[n \rightarrow \infty]{\text{BL}, (\mathbf{0}, z)} \Upsilon_{(0,z)}^{(r)} \stackrel{\text{law}}{=} \left( \Upsilon_{(0,z)}^{(r)}(1), \Upsilon_{(0,z)}^{(r)}(2) \right). \end{aligned}$$

By  $(2 + \epsilon)$ -moment conditions the above limits hold also for the expectations and all cross-products. Moreover, by the triangle inequalities, the moments of limits involving  $\Delta\xi_i^{(r)}$  are zero. Indeed, observe for  $i \in \{1, \dots, m\}$ , BL-distance between  $\xi_i^{(r)}(\tilde{x}_n, \tilde{\mathcal{P}}_n)$  and  $\xi_i(\tilde{x}_n, \tilde{\mathcal{P}}_n)$  goes to 0 for  $r \rightarrow \infty$  uniformly in  $n \in \mathbb{N}$  (this is BL-localization, here considered for two Palm orders only).

In order to lift the above observations to the linear combinations of  $\xi$  and  $\Delta\xi^{(r)}$ , consider the function  $h$  mapping  $(x_i^1, x_i^2, x_i^3)_{i \in \{1, \dots, m\}} \in \mathbb{R}^{3 \times m}$  to  $\sum_{i=1}^m t_i (x_i^1 - x_i^2) \in \mathbb{R}$ . This function is Lipschitz. Consequently, by the  $(2 + \epsilon)$ -moment condition which is satisfied by all co-ordinates of  $\xi^{(r)}$ , the BL-limits of random variables  $\Delta\xi^{(r)}$  satisfy

$$\begin{aligned} \Delta\xi^{(r)}(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n) &\xrightarrow[n \rightarrow \infty]{\text{BL}, \mathbf{0}} h^{(r)}(\Upsilon_0^{(r)}), \\ \left( \Delta\xi^{(r)}(\tilde{\mathbf{0}}, \tilde{\mathcal{P}}_n), \Delta\xi^{(r)}(\tilde{z}, \tilde{\mathcal{P}}_n) \right) &\xrightarrow[n \rightarrow \infty]{\text{BL}, (\mathbf{0}, z)} \left( h^{(r)}(\Upsilon_{(0,z)}^{(r)}(1)), h^{(r)}(\Upsilon_{(0,z)}^{(r)}(2)) \right). \end{aligned}$$

Again, by  $(2 + \epsilon)$ -moment conditions the above limits hold also for the expectations and all cross-products and the cross-moments of limits involving  $\Delta\xi^{(r)}$  are zero, by the previous observation for the individual score functions  $\xi_i$ .

To establish versions of the limits (7.39) and (7.40) for  $\sigma_r^2$  and  $v_r^2$ , it remains to justify the passage of  $\lim_{r \rightarrow \infty}$  under the integral  $\int_{\mathbb{R}^d} [\dots] dz$  in (7.44) where respective variables  $\Upsilon$  represent now the corresponding BL-limits of  $\xi^{(r)}$  and  $\Delta\xi^{(r)}$ . This is justified again by the dominated convergence theorem, evoking, for both  $\xi^{(r)}$  and  $\Delta\xi^{(r)}$ , BL-mixing correlations of order 2, uniformly in  $r > 0$ ; that is, these properties involve the common functions  $\omega_k$  in (3.2).  $\square$

## C Some consequences of $L^1$ -stabilization

We now consider a notion of stabilization, referred to as  $L^1$ -stabilization, which is stronger than BL-localization but weaker than stopping-set stabilization. It enables us to show that the point processes  $\tilde{\mathcal{P}}_n^\xi := \sum_{x \in \mathcal{P}_n} \delta_{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))}$  converge as  $n \rightarrow \infty$  to the ground process equipped with new marks  $\tilde{\mathcal{P}}^\Upsilon := \sum_{x \in \mathcal{P}} \delta_{(x, \Upsilon(x))}$  over all of  $\mathbb{R}^d$ . In this approach, the existence of  $\tilde{\mathcal{P}}^\Upsilon \cap K$  arises as a consequence of  $(\tilde{\mathcal{P}}_n^\xi \cap K)_{n \in \mathbb{N}}$  forming a Cauchy sequence in the bounded Lipschitz distance for finite point processes introduced in Schuhmacher and Xia [137] (see also Barbour and Brown [9]). This Cauchy property is proved using  $L^1$ -stabilization. Once this is done, we demonstrate the consistency of the limits across different subsets  $K \subset \mathbb{R}^d$  by leveraging the common ground processes, thus guaranteeing the new marking  $\tilde{\mathcal{P}}^\Upsilon$  over the whole space. Finally, we show that the limits  $\Upsilon_{[x]_1^p}$  of the Palm distributions developed in Lemma 7.4 indeed correspond to the Palm mark distributions  $\mathcal{M}_{[x]_1^p}^\Upsilon$  of the limiting point process  $\tilde{\mathcal{P}}^\Upsilon$ . The approach involving the Cauchy property is similar to the convergence of the Palm distributions of  $\xi$  stated in Lemma 7.4 under BL-localization. However this bounded BL convergence for point processes holds under the stronger  $L^1$ -stabilization of the score function  $\xi$  on finite windows of  $\tilde{\mathcal{P}}$  and is clearly implied by the stopping set stabilization (cf. Remark 4.7).

A general definition of this new stabilization is a modification of the BL-localization formulation in Definition 4.2 (ii). Specifically, for  $q \geq 1$ , we say that  $\xi$  is a  $L^q$ -stabilizing marking function on finite windows of  $\tilde{\mathcal{P}}$ , if for all  $p \in \mathbb{N}$  there are decreasing functions  $\varphi_p = \varphi_{p,q} : (0, \infty) \rightarrow [0, \infty)$  such that

$$\sup_{1 \leq n < \infty} \sup_{x_1, \dots, x_p \in W_n} \mathbb{E}_{[x]_1^p} [|\xi(\tilde{x}_1, \tilde{\mathcal{P}}_n) - \xi(\tilde{x}_1, \tilde{\mathcal{P}}_n \cap B_r([x]_1^p))|^{q}]^{1/q} \leq \varphi_{p,q}(r), \quad r > 0. \quad (\text{C.1})$$

Observe,  $L^q$ -stabilization is formulated only for marginal Palm distributions, as with stopping set stabilization (cf. (4.12)). Also, recall that (C.1) implies BL-localization as per (4.3); see (4.13).

**Lemma C.1.** (convergence of point processes with  $L^1$ -stabilizing marks) *Let  $\tilde{\mathcal{P}}$  be a marked point process on  $\mathbb{R}^d \times \mathbb{K}$  with non-null, finite intensity  $\rho$ . Let  $\xi : \mathbb{R}^d \times \mathbb{K} \times \hat{\mathcal{N}}_{\mathbb{R}^d \times \mathbb{K}} \rightarrow \mathbb{R}$  be a  $\mathbb{R}$ -valued marking function which is  $L^1$ -stabilizing on finite windows of  $\tilde{\mathcal{P}}$  as in (C.1) with  $p = 1$  and satisfies the first moment condition on finite windows as in (3.4) with  $p = 1$ .*

(i) *Then the ground point process marked by the score function  $\xi$  on finite windows converges weakly in the vague topology to a stationary marking  $\tilde{\mathcal{P}}^\Upsilon$  of this ground point process on the entire space, i.e.,*

$$\tilde{\mathcal{P}}_n^\xi = \sum_{x \in \mathcal{P}_n} \delta_{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))} \xrightarrow[n \rightarrow \infty]{d} \tilde{\mathcal{P}}^\Upsilon = \sum_{x \in \mathcal{P}} \delta_{(x, \Upsilon(x))}. \quad (\text{C.2})$$

(ii) If moreover  $\xi$  satisfies (C.1) for some  $p \in \mathbb{N}$  then the Palm distributions  $\mathcal{M}_{[x]_1^p}^\Upsilon$  of  $\tilde{\mathcal{P}}^\Upsilon$  of order  $p$  satisfy

$$\mathcal{M}_{[x]_1^p}^\Upsilon(\cdot) = \mathbb{P}\{\Upsilon_{[x]_1^p} \in \cdot\},$$

where  $\Upsilon_{[x]_1^p}$  are limits of the Palm distributions of the score function  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[n \rightarrow \infty]{\text{BL}, [x]_1^p} \Upsilon_{[x]_1^p}$  stated in (7.18).

A key element of the proof of the convergence (C.2) is demonstrating the Cauchy property in the bounded Lipschitz distance  $d_{2,BL}$  (defined in (C.6) below) for each sequence  $\tilde{\mathcal{P}}_{n,K}^\xi := \tilde{\mathcal{P}}_n^\xi \cap K$  on compact windows  $K \subset \mathbb{R}^d$  as  $n \rightarrow \infty$ . This metric  $d_{2,BL}$  completely metrizes the weak convergence of finite point processes. Thus using the Cauchy property, we will construct the limit  $\tilde{\mathcal{P}}^\Upsilon$  on compact windows  $K$  and, by Kolmogorov consistency, extend the limit to the entire space  $\mathbb{R}^d$ . We now define the required metrics and discuss some of their properties before proving the above lemma.

*Compactification of points and marks:* To use the framework of the BL metric of  $d_{2,BL}$  for finite point processes in a compact space, we consider the points of the ground process within a compact window  $K \subset \mathbb{R}^d$ . Moreover, to handle the real marks  $\xi(\tilde{x}, \tilde{\mathcal{P}}_n) \in \mathbb{R}$  of these points, we use the one-point (Alexandroff) compactification  $\bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$  of  $\mathbb{R}$ . On the compact space  $\bar{\mathbb{R}}$ , we consider a complete metric  $\bar{d}$  induced and bounded by the Euclidean distance on  $\mathbb{R}$ : for  $x, y \in \mathbb{R}$ ,

$$\bar{d}(x, y) := \min(|x - y|, h(x), h(y)) \quad (\text{C.3})$$

with  $h(z) = 1/(1 + |z|)$  for  $z \in \mathbb{R}$ , and  $\bar{d}(\infty, x) := h(x)$ ; cf [104]. With a compact subset  $K \subset \mathbb{R}^d$  representing the locations of points and the compactification  $\bar{\mathbb{R}}$  of the real space of their marks, we next consider a metric  $d_0$  on  $\mathfrak{X} := K \times \bar{\mathbb{R}}$  defined as:

$$d_0(\tilde{x}, \tilde{x}') := D_K(|x - x'| + \bar{d}(u, u')), \quad \tilde{x} = (x, u), \quad \tilde{x}' = (x', u') \in \mathfrak{X}, \quad (\text{C.4})$$

where  $D_K$  is a constant ensuring  $d_0 \leq 1$ . The compact metric space  $(\mathfrak{X}, d_0)$  of ‘marked points’ is both separable and complete.

*Metric on the counting measures:* Following [137], we define a metric  $d_1$  on the space of finite counting measures  $\hat{\mathcal{N}}_{\mathfrak{X}}$  on  $\mathfrak{X}$ . Specifically, we define the symmetric map  $\hat{\mathcal{N}}_{\mathfrak{X}}^2 \rightarrow [0, \infty)$  for  $\tilde{\mu} = \sum_{\tilde{x} \in \tilde{\mu}} \delta_{\tilde{x}}$  and  $\tilde{\mu}' = \sum_{\tilde{x}' \in \tilde{\mu}'} \delta_{\tilde{x}'}$  with  $\tilde{\mu}'(\mathfrak{X}) \geq \max(\tilde{\mu}(\mathfrak{X}), 1)$  by

$$d_1(\tilde{\mu}, \tilde{\mu}') := \frac{1}{n} \left( \min_{\pi \in \Pi_n} \sum_{i=1}^m d_0(\tilde{x}_i, \tilde{x}'_{\pi(i)}) + (n - m) \right) \quad (\text{C.5})$$

where  $n := \tilde{\mu}'(\mathfrak{X})$ ,  $m := \tilde{\mu}(\mathfrak{X})$ , and  $\Pi_n$  denotes the set of permutations of  $\{1, 2, \dots, n\}$ , and we put  $d_1(\tilde{\mu}, \tilde{\mu}') = 0$  if both counting measures are null.

In what follows, we will consider marked points sharing the same locations on  $K$ , hence we will always have  $\tilde{\mu}(\mathfrak{X}) = \tilde{\mu}'(\mathfrak{X})$ , as in (C.7) below. In this case, the metric  $d_1$  simplifies to the one considered in Barbour and Brown [9]. The metric  $d_1$  metrizes the vague topology (the same as the weak topology) of the finite counting measures  $\hat{\mathcal{N}}_{\mathfrak{X}}$ , making  $(\hat{\mathcal{N}}_{\mathfrak{X}}, d_1)$  locally compact, complete and separable [137, Proposition 2.B].

*Bounded Lipschitz distance on marked point processes:* Finally, consider the set  $\text{BL}(\hat{\mathcal{N}}_{\mathfrak{X}})$  of bounded Lip(1) functions on  $(\hat{\mathcal{N}}_{\mathfrak{X}}, d_1)$ , the space of finite counting measures on  $\mathfrak{X}$ . Here, we define a metric  $d_{2,BL}$  for the probability distributions on  $\hat{\mathcal{N}}_{\mathfrak{X}}$ , capturing the distance between the distributions of finite point processes  $\tilde{\mathcal{P}}, \tilde{\mathcal{P}}'$  on  $\mathfrak{X}$ :

$$d_{2,BL}(\tilde{\mathcal{P}}, \tilde{\mathcal{P}}') := \sup_{f \in \text{BL}(\hat{\mathcal{N}}_{\mathfrak{X}})} \left| \mathbb{E}[f(\tilde{\mathcal{P}})] - \mathbb{E}[f(\tilde{\mathcal{P}}')] \right|. \quad (\text{C.6})$$

This metric is denoted by  $d_2$  in [137] but we have chosen the above notation to align with our general terminology for BL metrics. We remark that  $d_{2,BL}$  is a distance between the laws of the marked point processes rather than the point processes themselves. Since  $(\hat{\mathcal{N}}_{\mathfrak{X}}, d_1)$  is complete and separable, the metric  $d_{2,BL}$  completely metrizes the weak topology of point process distributions on  $\mathfrak{X}$ ; [23, Theorems 8.3.2 and 8.10.43], see also [137, Proposition 2.C]. This makes it suitable for constructing limits of the marked point processes  $(\tilde{\mathcal{P}}_{n,K}^\xi)_{n \in \mathbb{N}}$  via the Cauchy property.

*Proof of Lemma C.1. Cauchy property and  $d_{2,BL}$  convergence of  $\tilde{\mathcal{P}}_{n,K}^\xi$ :* For a given compact  $K \subset \mathbb{R}^d$  we prove that the sequence of point processes  $(\tilde{\mathcal{P}}_{n,K}^\xi)_{n \in \mathbb{N}}$  on  $\mathfrak{X} = K \times \bar{\mathbb{R}}$ , is a Cauchy sequence in the  $d_{2,BL}$  metric, thus justifying the convergence

$$\tilde{\mathcal{P}}_{n,K}^\xi = \sum_{x \in \mathcal{P}_n \cap K} \delta_{(x, \xi(\tilde{x}, \tilde{\mathcal{P}}_n))} \xrightarrow[n \rightarrow \infty]{d_{2,BL}} \tilde{\mathcal{P}}_K^\Upsilon \quad (\text{C.7})$$

to some random counting measure  $\tilde{\mathcal{P}}_K^\Upsilon \in \hat{\mathcal{N}}_{\mathfrak{X}}$ . Consider  $1 \leq n \leq n'$  and a function  $f \in \text{BL}(\hat{\mathcal{N}}_{\mathfrak{X}})$ . Coupling  $\tilde{\mathcal{P}}_{n,K}^\xi$  and  $\tilde{\mathcal{P}}_{n',K}^\xi$  on the common process  $\tilde{\mathcal{P}}$ , using in (C.5) the identity permutation  $\pi \in \Pi_{\mathcal{P}(K)}$  with respect to any numbering of the ground process, and (C.3), (C.4) we have

$$\begin{aligned} & \left| \mathbb{E}[f(\tilde{\mathcal{P}}_{n,K}^\xi)] - \mathbb{E}[f(\tilde{\mathcal{P}}_{n',K}^\xi)] \right| \quad (\text{C.8}) \\ & \leq \mathbb{E} \left[ \frac{\mathbf{1}(\mathcal{P}(K) > 0)}{\mathcal{P}(K)} \sum_{x \in \mathcal{P} \cap K} D_K \bar{d}(\xi(\tilde{x}, \tilde{\mathcal{P}}_n), \xi(\tilde{x}, \tilde{\mathcal{P}}_{n'})) \right] \\ & \leq D_K \mathbb{E} \left[ \sum_{x \in \mathcal{P} \cap K} |\xi(\tilde{x}, \tilde{\mathcal{P}}_n) - \xi(\tilde{x}, \tilde{\mathcal{P}}_{n'})| \right] \\ & = \rho D_K \int_K \mathbb{E}_x [|\xi(\tilde{x}, \tilde{\mathcal{P}}_n) - \xi(\tilde{x}, \tilde{\mathcal{P}}_{n'})|] dx. \end{aligned}$$

We now use  $L^1$ -stabilization, with (C.1) holding for  $p = 1$  and with a decreasing function  $\varphi_1 = \varphi_{1,1} \in [0, 1]$ . For  $r = d(x, \partial W_n)$ , we obtain the bound:

$$\begin{aligned} & \mathbb{E}_x [|\xi(\tilde{x}, \tilde{\mathcal{P}}_n) - \xi(\tilde{x}, \tilde{\mathcal{P}}_{n'})|] \quad (\text{C.9}) \\ & \leq \mathbb{E}_x [|\xi^{(r)}(\tilde{x}, \tilde{\mathcal{P}}_n) - \xi(\tilde{x}, \tilde{\mathcal{P}}_n)|] + \mathbb{E}_x [|\xi^{(r)}(\tilde{x}, \tilde{\mathcal{P}}_{n'}) - \xi(\tilde{x}, \tilde{\mathcal{P}}_{n'})|] \\ & \leq 2\varphi_1(d(x, \partial W_n)). \end{aligned}$$

Substituting into (C.8) and applying the Monotone Convergence Theorem, we have:

$$\lim_{n \rightarrow \infty} \left| \mathbb{E}[f(\tilde{\mathcal{P}}_{n,K}^\xi)] - \mathbb{E}[f(\tilde{\mathcal{P}}_{n',K}^\xi)] \right| \leq 2D_K \int_K \lim_{n \rightarrow \infty} \varphi_1(d(x, \partial W_n)) dx = 0. \quad (\text{C.10})$$

This confirms statement (C.7).

*Real marked point process  $\tilde{\mathcal{P}}_K^\Upsilon$ :* The random counting measure  $\tilde{\mathcal{P}}_K^\Upsilon \in \hat{\mathcal{N}}_{\mathfrak{X}}$  is indeed a simple ground process in  $K$  marked by values in  $\bar{\mathbb{R}}$ . This follows from the fact that the projection of  $\tilde{\mathcal{P}}_{n,K}^\xi$  onto  $K$  is equal in distribution to  $\mathcal{P} \cap K$  whenever  $K \subset W_n$ . However, in principle, the marked point process limit  $\tilde{\mathcal{P}}_K^\Upsilon$  could have marks at  $\infty \in \bar{\mathbb{R}}$ . This is ruled out by the moment condition, which implies the tightness of the point processes  $\tilde{\mathcal{P}}_{n,K}^\xi$  uniformly in  $n \in \mathbb{N}$ . Indeed, by Campbell's formula and the  $p = 1$  moment condition (3.4), we have for any  $t > 0$ ,

$$\mathbb{P}(\exists x \in \mathcal{P} \cap K : \xi(x, \tilde{\mathcal{P}}_n) > t) \leq \rho \int_K \mathbb{P}_x(\xi(x, \tilde{\mathcal{P}}_n) > t) dx \leq \rho \text{Vol}(K) M_1^\xi/t.$$

This along with (C.7), justifies that  $\tilde{\mathcal{P}}_K^\Upsilon$  is a marked point processes on  $K \times \mathbb{R}$ .

*Consistency of  $\tilde{\mathcal{P}}_K^\Upsilon$  as  $K \uparrow \mathbb{R}^d$ :* Note that the distributions of  $\tilde{\mathcal{P}}_K^\Upsilon$  are consistent across  $K$ . Indeed, consider  $K \subset K'$ , take  $\epsilon > 0$ , and due to (C.7), choose  $n$  large enough such that  $d_{2,BL}(\tilde{\mathcal{P}}_{n,K}^\xi, \tilde{\mathcal{P}}_K^\Upsilon) < \epsilon$  and similarly for  $K'$ . By the triangle inequality we obtain

$$d_{2,BL}(\tilde{\mathcal{P}}_K^\Upsilon, \tilde{\mathcal{P}}_{K'}^\Upsilon \cap K) \leq 2\epsilon + d_{2,BL}(\tilde{\mathcal{P}}_{n,K}^\xi, \tilde{\mathcal{P}}_{n,K'}^\xi \cap K) = 2\epsilon,$$

where the last equality is by consistency of distributions of  $\tilde{\mathcal{P}}_{n,K}^\xi$  across  $K$ . This establishes the existence of the marked point process  $\tilde{\mathcal{P}}^\Upsilon$  on  $\mathbb{R}^d \times \mathbb{R}$  and concludes the proof of the statement in Item (i).

We next prove statement (ii). For this, observe first that  $L^1$ -stabilization as in (C.1) implies BL-localization as in (4.3) and hence, for every  $p \in \mathbb{N}$ , by Lemma 7.4 (iii), there exist limits of the Palm distributions of the score function  $[\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n) \xrightarrow[n \rightarrow \infty]{\text{BL}, [\xi]_1^p} \Upsilon_{[x]_1^p}$  stated in (7.18). Consider now two functions  $f$  and  $h$  defined, respectively, on  $\mathbb{R}^{pd}$  and  $\mathbb{R}^p$ , bounded by 1, with bounded support, such that  $h$  is a bounded Lipschitz(1) on  $\mathbb{R}^p$  and the function

$$\hat{\mathcal{N}}_{\mathfrak{X}} \ni \tilde{\mu} = \sum_{\tilde{x}=(x,u) \in \tilde{\mu}} \delta_{(x,u)} \mapsto \sum_{\tilde{x}_{i_1}, \dots, \tilde{x}_{i_p} \neq} f([x]_1^p) h([u]_1^p) \in \mathbb{R}$$

is in  $\text{BL}(\hat{\mathcal{N}}_{\mathfrak{X}})$  where  $\mathfrak{X} = K \times \bar{\mathbb{R}}$  and  $K \subset \mathbb{R}^d$  is such that  $K^p \subset \mathbb{R}^{dp}$  includes the support of  $f$ . For a given  $\epsilon > 0$ , using (C.7), take  $n$  large enough such that  $d_{2,BL}(\tilde{\mathcal{P}}_K^\Upsilon, \tilde{\mathcal{P}}_{n,K}^\xi) \leq \epsilon$ . We have

$$\begin{aligned} & \left| \int_{K^p} f([x]_1^p) \int_{\mathbb{R}^d} h([u]_1^p) \mathcal{M}_{[x]_1^p}^\Upsilon(d[u]_1^p) \rho^{(p)}(d[x]_1^p) - \int_{K^p} f([x]_1^p) \mathbb{E}[h(\Upsilon_{[x]_1^p})] \rho^{(p)}(d[x]_1^p) \right| \\ & \leq d_{2,BL}(\tilde{\mathcal{P}}_K^\Upsilon, \tilde{\mathcal{P}}_{n,K}^\xi) + \left| \int_{K^p} f([x]_1^p) \mathbb{E}_{[x]_1^p}[h([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n))] \rho^{(p)}(d[x]_1^p) - \int_{K^p} f([x]_1^p) \mathbb{E}[h(\Upsilon_{[x]_1^p})] \rho^{(p)}(d[x]_1^p) \right| \\ & \leq \epsilon + \kappa_p \int_{K^p} \left| \mathbb{E}_{[x]_1^p}[h([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n))] - \mathbb{E}[h(\Upsilon_{[x]_1^p})] \right| d[x]_1^p \\ & \leq \epsilon + \kappa_p \int_{K^p} d_{\text{BL}, [\xi]_1^p}([\xi]_1^p(\tilde{x}, \tilde{\mathcal{P}}_n), \Upsilon_{[x]_1^p}) d[x]_1^p \end{aligned}$$

$$\leq \epsilon + 6\kappa_p \int_{K^p} \varphi_1(d([x]_1^p, \partial W_n)) d[x]_1^p,$$

where we used (7.19) in the last inequality. The last term converges to 0 when  $n \rightarrow \infty$ . This justifies the statement in Item (ii) and completes the proof of Lemma C.1.  $\square$

## References

- [1] D. Ahlberg and J. Tykesson. Gilbert’s disc model with geostatistical marking. *Adv. in Appl. Probab.*, 50(4):1075–1094, 2018. 152, 158, 159
- [2] D. Aldous and D. Lanoue. A lecture on the averaging process. *Probability Surveys*, 9:90 – 102, 2012. 145
- [3] G. Amir, F. Nazarov, and Y. Peres. Convergence rate of  $\ell^p$ -energy minimization on graphs: sharp polynomial bounds and a phase transition at  $p = 3$ . *arXiv:2508.19411*, 2025. 145
- [4] A. Anandkumar, J. E. Yukich, L. Tong, and A. Swami. Energy scaling laws for distributed inference in random fusion networks. *IEEE J. Select. Areas Comm.*, 27(7):1203–1217, 2009. 164
- [5] M. Austern and P. Orbanz. Limit theorems for distributions invariant under groups of transformations. *Ann. Stat.*, 50(4):1960–1991, 2022. 43
- [6] F. Baccelli and S. Foss. Poisson hail on a hot ground. *J. Appl. Probab.*, 48(A):343–366, 2011. 147
- [7] F. Baccelli and N. Ramesan. A computational framework for evaluating the role of mobility on the propagation of epidemics on point processes. *J. Math. Biol.*, 84:1–40, 2022. 164
- [8] F. Baccelli, B. Błaszczyszyn, and M. Karray. *Random Measures, Point Processes, and Stochastic Geometry*. <https://hal.inria.fr/hal-02460214>, 2020. 15, 19, 20, 106
- [9] A. Barbour and T. Brown. Stein’s method and point process approximation. *Stoch. Proc. Applns.*, 43(1):9–31, 1992. 176, 177
- [10] Y. Baryshnikov and J.E. Yukich. Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.*, 15(1A):213–253, 2005. 12
- [11] V. Beneš, C. Hofer-Temmel, G. Last, and J. Večeřa. Decorrelation of a class of Gibbs particle processes and asymptotic properties of U-statistics. *J. Appl. Probab.*, 57(3):928–955, 2020. 13
- [12] R. Berkowitz and P. Devlin. Central limit theorem for majority dynamics: Bribing three voters suffices. *Stoch. Proc. Applns.*, 146:187–206, 2022. 145

- [13] S. Betsch. *A point process theoretic study of Gibbs measures in abstract state spaces*. PhD thesis, Dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), 2022. URL <https://publikationen.bibliothek.kit.edu/1000150234>. 165
- [14] C. A. N. Biscio and F. Lavancier. Brillinger mixing of determinantal point processes and statistical applications. *Elec. J. Stat.*, 10(1):582–607, 2016. 59
- [15] C. A. N. Biscio and A. M. Svane. A functional central limit theorem for the empirical Ripley’s  $k$ -function. *Elec. J. Stat.*, 16(1):3060–3098, 2022. 132
- [16] C. A. N. Biscio, N. Chenavier, C. Hirsch, and A. M. Svane. Testing goodness of fit for point processes via topological data analysis. *Elec. J. Stat.*, 14(1):1024–1074, 2020. 132
- [17] B. Błaszczyszyn. Factorial-moment expansion for stochastic systems. *Stoch. Proc. Appl.*, 56:321–335, 1995. 10, 45, 46
- [18] B. Błaszczyszyn and D. Yogeshwaran. On comparison of clustering properties of point processes. *Adv. in Appl. Probab.*, 46(1):1–20, 2014. 129, 137
- [19] B. Błaszczyszyn and D. Yogeshwaran. Clustering comparison of point processes, with applications to random geometric models. In *Stochastic Geometry, Spatial Statistics and Random Fields*, pages 31–71. Springer, 2015. 137
- [20] B. Błaszczyszyn, E. Merzbach, and V. Schmidt. A note on expansions for functionals of spatial marked processes. *Statist. Prob. Lett.*, 36:299–306, 1997. 10, 45, 46
- [21] B. Błaszczyszyn, D. Yogeshwaran, and J. E. Yukich. Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Probab.*, 47(2):835–895, 2019. 9, 12, 13, 22, 24, 31, 34, 37, 42, 43, 44, 47, 51, 54, 55, 57, 58, 72, 140, 165, 168
- [22] B. Błaszczyszyn, D. Yogeshwaran, and J. E. Yukich. Supplement to ‘Limit theory for geometric statistics of point processes having fast decay of correlations’. DOI:10.1214/18-AOP1273SUPP, 2019. 166, 168
- [23] V. I. Bogachev. *Measure Theory*, volume 2. Springer, 2007. 16, 66, 178
- [24] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Ann. Probab.*, pages 1047–1050, 1982. 13
- [25] B. Bond and L. Levine. Abelian networks I. Foundations and examples. *SIAM J. Discrete Math.*, 30(2):856–874, 2016. 149
- [26] T. Brox and H. Rost. Equilibrium fluctuations of stochastic particle systems: the role of conserved quantities. *Ann. Probab.*, 12:742–759, 1984. 14

- [27] A. Bulinski and E. Spodarev. Central limit theorems for weakly dependent random fields. In *Stochastic Geometry, Spatial Statistics and Random Fields*, pages 337–383. Springer, 2013. 9, 13
- [28] V. H. Can. Annealed limit theorems for the Ising model on random regular graphs. *Ann. Appl. Probab.*, 29(3):1398–1445, 2019. 13, 79
- [29] E. Castro, S. Bubeck, G. Lugosi, and N. Verzelen. Detecting Markov random fields hidden in white noise. *Bernoulli*, 24((4B)):3628–3656, 2018. 164
- [30] G. Chargaziya and A. Daletskii. Stochastic dynamics of particle systems on unbounded degree graphs. *J. Math. Phys.*, 66(2), 2025. 114
- [31] S. Chatterjee. *Concentration inequalities with exchangeable pairs*. Stanford University, Ph.D. Thesis, 2005. 13, 98
- [32] L. H. Y. Chen, A. Röllin, and A. Xia. Palm theory, random measures and stein couplings. *Ann. Appl. Probab.*, 31(6):2881–2923, 2021. 12, 43
- [33] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications*. John Wiley & Sons, 2013. 163
- [34] T. Cong and A. Xia. Convergence rate for geometric statistics of point processes with fast decay dependence. *Elec. J. Probab.*, 28:1–35, 2023. 12, 43
- [35] S. Corry and D. Perkinson. *Divisors and Sandpiles: An Introduction to Chip-firing*. American Mathematical Society, 2018. 149
- [36] J. T. Cox and R. Durrett. The stepping stone model: New formulas expose old myths. *Ann. Appl. Probab.*, 12(4):1348–1377, 2002. 164
- [37] D. A. Dawson and A. Greven. *Spatial Fleming-Viot Models with Selection and Mutation*. Springer, 2014. 148, 164
- [38] Boris Delaunay. Sur la sphère vide. à la mémoire de georges voronoï. *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et naturelles*, 6:793–800, 1934. 169
- [39] D. Dereudre. Introduction to the theory of Gibbs point processes. In *Stochastic Geometry*, pages 181–229. Springer, 2019. 78
- [40] J. Díaz, D. Mitsche, and X. Pérez-Giménez. Large connectivity for dynamic random geometric graphs. *IEEE Tran. Mob. Comput.*, 8(6):821–835, 2009. 163
- [41] T.-C. Dinh, S. Ghosh, H. Tran, and M. Tran. Quantitative Marcinkiewicz’s theorem and central limit theorems: applications to spin systems and point processes. *arXiv:2107.08469*, 2021. 13, 43, 84, 85

- [42] R. L. Dobrushin and S. B. Shlosman. Completely analytical interactions: constructive description. *J. Stat. Phys.*, 46:983–1014, 1987. [14](#), [78](#)
- [43] R.L. Dobrushin and S.B. Shlosman. Constructive criterion for the uniqueness of gibbs field. In *Statistical Physics and Dynamical Systems: Rigorous Results*, pages 347–370. Springer, 1985. [13](#), [81](#), [83](#)
- [44] R.L. Dobrushin and B. Tirozzi. The central limit theorem and the problem of equivalence of ensembles. *Comm. Math. Phys.*, 54:173–192, 1977. [13](#)
- [45] H. Döring, S. Jansen, and K. Schubert. The method of cumulants for the normal approximation. *Prob. Surv.*, 19:185–270, 2022. [43](#)
- [46] P. Doukhan, G. Lang, S. Louhichi, and B. Ycart. A functional central limit theorem for interacting particle systems on transitive graphs. *Markov Proc. Relat. Fields*, 14:79–114, 2008. [14](#), [131](#)
- [47] M. Draief and L. Massoulié. *Epidemics and Rumours in Complex Networks*. Cambridge University Press, 2010. [137](#), [144](#), [147](#)
- [48] H. Duminil-Copin. Geometric representations of lattice spin models. *Spartacus Graduate, Cours Peccot, Collège de France*, 2015. [13](#), [78](#), [79](#)
- [49] H. Duminil-Copin and S. Smirnov. The connective constant of the honeycomb lattice equals  $\sqrt{2 + \sqrt{2}}$ . *Ann. Math.*, 175(3):1653–1665, 2012. [101](#)
- [50] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation on  $\mathbb{Z}^d$ . *L’Enseign. Math.*, 62(1):199–206, 2017. [96](#)
- [51] H. Duminil-Copin, A. Raoufi, and V. Tassion. Exponential decay of connection probabilities for subcritical Voronoi percolation in  $R^d$ . *Prob. Th. Rel. Fields*, 173:479–490, 2019. [100](#), [105](#)
- [52] H. Duminil-Copin, A. Raoufi, and V. Tassion. Subcritical phase of  $d$ -dimensional Poisson–Boolean percolation and its vacant set. *Ann. Hen. Leb.*, 3:677–700, 2020. [100](#), [105](#)
- [53] R. S. Ellis. *Entropy, large deviations, and statistical mechanics*. Springer, 2007. [13](#)
- [54] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009. [131](#)
- [55] J. W. Evans. Random and cooperative sequential adsorption. *Rev. Mode. Phys.*, 65(4):1281, 1993. [141](#)
- [56] M. R. Evans and T. Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A: Math. Gen.*, 38(19):R195, 2005. [147](#)

- [57] M. Fenzl. Asymptotic results for stabilizing functionals of point processes having fast decay of correlations. *arXiv:1909.13274*, 2019. 9, 12, 22, 43
- [58] A. Fey, R. Meester, and F. Redig. Stabilizability and percolation in the infinite volume sandpile model. *Ann. Probab.*, 37(2):654–675, 2009. 149
- [59] D. Flimmel, Z. Pawlas, and J. E. Yukich. Limit theory for unbiased and consistent estimators of statistics of random tessellations. *J. Appl. Probab.*, 57:679–702, 2020. 162
- [60] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press, 2017. 13, 78, 79
- [61] A. Gabrielli, M. Joyce, and S. Torquato. Tilings of space and superhomogeneous point processes. *Phy. Rev. E—Stat., Nonlinear, and Soft Matter Physics*, 77(3):031125, 2008. 44
- [62] D. Gamarnik. Correlation decay and the absence of zeros property of partition functions. *Rand. Struct. Alg.*, 62(1):155–180, 2023. 14
- [63] H-O. Georgii. *Gibbs measures and phase transitions*, volume 9. Walter de Gruyter, 2011. 98
- [64] H-O. Georgii, O. Häggström, and C. Maes. The random geometry of equilibrium phases. In *Phase transitions and critical phenomena*, volume 18, pages 1–142. Elsevier, 2001. 78, 98
- [65] C. Giardinà, C. Giberti, R. van der Hofstad, and M. L. Prioriello. Quenched central limit theorems for the ising model on random graphs. *J. Stat. Phys.*, 160(6):1623–1657, 2015. 13, 79
- [66] C. Giardinà, C. Giberti, R. van der Hofstad, and M. L. Prioriello. Annealed central limit theorems for the ising model on random graphs. *ALEA*, 13:121–161, 2016. 13, 79, 84
- [67] A. Gloria and M. D. Penrose. Random parking, Euclidean functionals, and rubber elasticity. *Comm. Math. Phys.*, 321(1):1–31, 2013. 141
- [68] P. Hall. On continuum percolation. *Ann. Probab.*, pages 1250–1266, 1985. 104
- [69] I. Hartarsky and C. Toninelli. Kinetically constrained models. *arXiv:2412.13634*, 2024. 144, 147, 149
- [70] I. Hartarsky, F. Martinelli, and C. Toninelli. Fredrickson–Andersen model in two dimensions. *ESAIM: Proc. Surv.*, 74:108–118, 2023. 144
- [71] L. Heinrich. Asymptotic methods in statistics of random point processes. In *Stochastic Geometry, Spatial Statistics and Random Fields*, pages 115–150. Springer, 2013. 9, 12, 23
- [72] L. Heinrich and Z. Pawlas. Weak and strong convergence of empirical distribution functions from germ-grain processes. *Statistics*, 42(1):49–65, 2008. 9, 12, 23

- [73] L. Heinrich, S. Lück, and V. Schmidt. Asymptotic goodness-of-fit tests for the Palm mark distribution of stationary point processes with correlated marks. *Bernoulli*, 20(4):1673–1697, 2014. [9](#), [12](#), [23](#)
- [74] C. Hirsch, M. Otto, and A. M. Svane. Normal approximation for gibbs processes via disagreement couplings. *Elec. J. Prob.*, 30:1–56, 2025. [12](#), [43](#)
- [75] R. Holley and D. Stroock. Generalized Ornstein-Uhlenbeck processes as limits of interacting systems. *Stochastic Integrals*, page 152, 1980. [14](#)
- [76] R. Holley and D. W. Stroock. Central limit phenomena of various interacting systems. *Ann. Math.*, 110(2):333–393, 1979. [13](#), [14](#), [131](#)
- [77] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012. [166](#)
- [78] G. Ivanoff. Central limit theorems for point processes. *Stoch. Proc. Appl.*, 12(2):171–186, 1982. [9](#), [12](#), [23](#), [59](#)
- [79] B. Jahnel and W. König. *Probabilistic Methods in Telecommunications*. Birkhäuser Basel, 2020. [144](#), [147](#)
- [80] J. Jalowy and H. Stange. Box-covariances of hyperuniform point processes. *arXiv:2506.13661*, 2025. [44](#)
- [81] A. A. Járai. Sandpile models. *Prob. Surv.*, 15:243–306, 2018. [148](#), [149](#)
- [82] N. Jenish and I. R. Prucha. Central limit theorems and uniform laws of large numbers for arrays of random fields. *J. Econ.*, 150(1):86–98, 2009. [9](#), [12](#), [23](#)
- [83] O. Kallenberg. *Random Measures, Theory and Applications*. Springer, 2017. [15](#), [17](#), [18](#)
- [84] H. Kaplan, W. Mulzer, L. Roditty, P. Seiferth, and M. Sharir. Dynamic planar Voronoi diagrams for general distance functions and their algorithmic applications. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2495–2504, 2017. [164](#)
- [85] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113. Springer, 2014. [123](#)
- [86] M. A. Klatt, G. Last, L. Lotz, and D. Yogeshwaran. Invariant transports of stationary random measures: asymptotic variance, hyperuniformity, and examples. *arXiv:2506.05907*, 2025. [166](#)
- [87] C. J. Klivans. *The Mathematics of Chip-Firing*. CRC Press, 2018. [149](#)
- [88] M. Krishnapur and D. Yogeshwaran. Stationary random measures : Covariance asymptotics, variance bounds and central limit theorems. *arXiv:2411.08848*, 2024. [44](#)

- [89] H. Künsch. Decay of correlations under dobrushin’s uniqueness condition and its applications. *Comm. Math. Phys.*, 84:207–222, 1982. [13](#)
- [90] R. Lachièze-Rey. Normal convergence of nonlocalised geometric functionals and shot-noise excursions. *Ann. Appl. Prob.*, 29(5):2613–2653, 2019. [12](#), [33](#), [164](#)
- [91] R. Lachièze-Rey. Rigidity of random stationary measures and applications to point processes. *arXiv:2409.18519*, 2024. [44](#)
- [92] R. Lachièze-Rey and S. Muirhead. Asymptotics for the critical level and a strong invariance principle for high intensity shot noise fields. *Stoch. Proc. Applns.*, 147:175–209, 2022. [164](#)
- [93] R. Lachièze-Rey, M. Schulte, and J. E. Yukich. Normal approximation for stabilizing functionals. *Ann. Appl. Probab.*, 29(2):931–993, 2019. [10](#), [12](#), [31](#), [43](#), [131](#), [151](#)
- [94] R. Lachièze-Rey, G. Peccati, and X. Yang. Quantitative two-scale stabilization on the Poisson space. *Ann. Appl. Probab.*, 32(4):3085–3145, 2022. [12](#)
- [95] D. Lacker, K. Ramanan, and R. Wu. Local weak convergence for sparse networks of interacting processes. *Ann. Appl. Prob.*, 33(2):643 – 688, 2023. [11](#), [114](#), [115](#), [116](#), [117](#), [118](#), [121](#), [122](#), [123](#)
- [96] H. Lacoin. Non-coincidence of quenched and annealed connective constants on the supercritical planar percolation cluster. *Prob. Th. Rel. Fields*, 159:777–808, 2014. [97](#)
- [97] G. Last and M. Penrose. *Lectures on the Poisson Process*, volume 7. Cambridge University Press, 2017. [15](#), [45](#)
- [98] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler’s formula, second order Poincaré inequalities and stabilization. *Prob. Th. Rel. Fields*, 165(3):667–723, 2016. [10](#), [12](#), [164](#)
- [99] S. L. Lauritzen, A. P. Dawid, B. N. Larsen, and H-G. Leimer. Independence properties of directed markov fields. *Networks*, 20(5):491–505, 1990. [94](#)
- [100] T. M. Liggett. *Interacting Particle Systems*, volume 276. Springer Science and Business Media, 2012. [14](#)
- [101] J. Liu, A. Sinclair, and P. Srivastava. Correlation decay and partition function zeros: Algorithms and phase transitions. *SIAM J. Comp.*, (0):FOCS19–200, 2022. [14](#)
- [102] E. Lucon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *Annals Appl. Prob.*, 24(5):1946–1993, 2014. [115](#)
- [103] J. Mairesse and I. Marcovici. Around probabilistic cellular automata. *Theoretical Computer Science*, 559:42–72, 2014. [147](#), [149](#)

- [104] M. Mandelkern. Metrization of the one-point compactification. *Proc. Am. Math. Soc.*, 107(4): 1111–1115, 1989. [177](#)
- [105] Ph. A. Martin and T. Yalcin. The charge fluctuations in classical Coulomb systems. *J. Stat. Phys.*, 22(4):435–463, 1980. [9](#)
- [106] F. Martinelli and E. Olivieri. Approach to equilibrium of glauber dynamics in the one phase region: I. the attractive case. *Comm. Math. Phys.*, 161(3):447–486, 1994. [13](#), [83](#)
- [107] G. Mastrilli, B. Błaszczyszyn, and F. Lavancier. Estimating the hyperuniformity exponent of point processes. *arXiv:2407.16797*, 2024. [44](#)
- [108] K. G. Mazumder. Existence and convergence of interacting particle systems on graphs. *arXiv:2408.07976*, 2024. [14](#), [131](#)
- [109] R. Meester and R. Roy. *Continuum percolation*, volume 119. Cambridge University Press, 1996. [100](#), [104](#)
- [110] F. Nazarov and M. Sodin. Correlation functions for random complex zeroes: Strong clustering and local universality. *Comm. Math. Phys.*, 310(1):75–98, 2012. [9](#), [13](#), [37](#), [44](#), [57](#)
- [111] E. Onaran, O. Bobrowski, and R. J. Adler. Functional central limit theorems for local statistics of spatial birth–death processes in the thermodynamic regime. *Ann. Appl. Probab.*, 33(5):3958–3986, 2023. [14](#), [131](#), [164](#)
- [112] E. Onaran, O. Bobrowski, and R. J. Adler. Functional limit theorems for local functionals of dynamic point processes. *arXiv:2310.17775*, 2023. [14](#), [132](#), [164](#)
- [113] Z. Pawlas. Empirical distributions in marked point processes. *Stoch. Proc. Appl.*, 119(12): 4194–4209, 2009. [14](#), [23](#), [154](#)
- [114] M. D. Penrose. Random parking, sequential adsorption, and the jamming limit. *Comm. Math. Phys.*, 218(1):153–176, 2001. [141](#)
- [115] M. D. Penrose. Gaussian limits for random geometric measures. *Elec. J. Probab.*, 12(35):989–1035, 2007. [12](#), [72](#), [151](#)
- [116] M. D. Penrose. Laws of large numbers in stochastic geometry with statistical applications. *Bernoulli*, 13(4):1124–1150, 2007. [31](#)
- [117] M. D. Penrose. Existence and spatial limit theorems for lattice and continuum particle systems. *Prob. Surv.*, 5:1–36, 2008. [14](#), [125](#), [131](#), [132](#), [141](#), [142](#), [144](#), [147](#), [148](#)
- [118] M. D. Penrose. Growth and roughness of the interface for ballistic deposition. *J. Stat. Phys.*, 131(2):247–268, 2008. [145](#)

- [119] M. D. Penrose and A. Sudbury. Exact and approximate results for deposition and annihilation processes on graphs. *Ann. Appl. Probab.*, 15(1B):853–889, 2005. 131, 141, 143
- [120] M. D. Penrose and J. E. Yukich. Mathematics of random growing interfaces. *J. Phys. A: Math Gen.*, 34(2):6239–6247, 2001. 14, 141, 142
- [121] M. D. Penrose and J. E. Yukich. Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.*, 11(4):1005–1041, 2001. 12
- [122] M. D. Penrose and J. E. Yukich. Limit theory for random sequential packing and deposition. *Ann. Appl. Probab.*, 12(1):272–301, 2002. 12, 141, 142, 145
- [123] M. D. Penrose and J. E. Yukich. Weak laws of large numbers in geometric probability. *Ann. Appl. Probab.*, 13:277–303, 2003. 31, 162
- [124] Y. Peres, A. Sinclair, P. Sousi, and A. Stauffer. Mobile geometric graphs: Detection, coverage and percolation. *Prob. Th. Rel. Fields*, 156(1-2):273–305, 2013. 164
- [125] X. Qi. A functional central limit theorem for spatial birth and death processes. *Adv. in Appl. Probab.*, 40(3):759–797, 2008. 14, 131, 132, 164
- [126] A. Quitmann. Decay of correlations in the monomer-dimer model. *J. Math. Phys.*, 65(10), 2024. 111
- [127] K. Ramanan. Interacting stochastic processes on sparse random graphs. In *Proc. Int. Cong. Math. 2022*, volume 6, pages 4394–4425. 2022. 114, 115
- [128] T. R. Reddy, S. Vadlamani, and D. Yogeshwaran. Central limit theorem for exponentially quasi-local statistics of spin models on Cayley graphs. *J. Stat. Phys.*, 173(3-4):941–984, 2018. 13
- [129] F. Redig, S. Roelly, and W. Ruszel. Short-time Gibbsianness for infinite-dimensional diffusions with space-time interaction. *J. Stat. Phys.*, 138(6):1124–1144, 2010. 115
- [130] G. Regts. Absence of zeros implies strong spatial mixing. *Prob. Th. Rel. Fields*, 186:621–641, 2023. 14, 78
- [131] H. Rost. A central limit theorem for a system of interacting particles. In *Stochastic Space—Time Models and Limit Theorems*, pages 243–248. Springer, 1985. 14
- [132] J. Sanders, M. Jonckheere, and S. Kokkelmans. Sub-Poissonian statistics of jamming limits in ultracold Rydberg gases. *Phy. Rev. Lett.*, 115(4):043002, 2015. 141
- [133] L. Saulis and V. A. Statulevicius. *Limit theorems for large deviations*, volume 73. Springer Science and Business Media, 1991. 43
- [134] M. Schlather, P. J. Riberio, and P. Diggle. Detecting dependence between marks and locations of marked point processes. *J. Roy. Stat. Soc. B*, 66(1):79–93, 2004. 152, 155, 158, 160

- [135] T. Schreiber and J. E. Yukich. Limit theorems for geometric functionals of Gibbs point processes. *Ann. I.H.P. Prob. Stat.*, 49(4):1158–1182, 2013. [12](#)
- [136] T. Schreiber, M. Penrose, and J. E. Yukich. Gaussian limits for multidimensional random sequential packing at saturation. *Comm. Math. Phys.*, 272:167–183, 2007. [141](#)
- [137] D. Schuhmacher and A. Xia. A new metric between distributions of point processes. *Adv. in Appl. Probab.*, 40(3):651–672, 2008. [176](#), [177](#), [178](#)
- [138] M. Schulte and V. Trapp. Lower bounds for variances of Poisson functionals. *Elec. J. Probab.*, 29:1–43, 2024. [44](#)
- [139] M. Schulte and J. E. Yukich. Multivariate second order Poincaré inequalities for Poisson functionals. *Elec. J. Probab.*, 24:1–42, 2019. [12](#)
- [140] M. Schulte and J. E. Yukich. Rates of multivariate normal approximation for statistics in geometric probability. *Ann. Appl. Probab.*, 33(1):507–548, 2023. [12](#), [31](#)
- [141] V. Shcherbakov. Probabilistic models motivated by cooperative sequential adsorption. *arXiv:2312.13731*, 2023. [141](#)
- [142] S. Shuai and S. Ke. Zero-freeness is all you need: A Weitz-type fptas for the entire Lee-Yang zero-free region. *arXiv:2509.06623*, 2025. [111](#)
- [143] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin. Spatial mixing and the connective constant: Optimal bounds. *Prob. Th. Rel. Fields*, 168:153–197, 2017. [11](#), [13](#), [78](#), [79](#), [96](#), [101](#), [107](#), [108](#), [109](#), [111](#), [112](#), [113](#)
- [144] M. Sodin, A. Wennman, and O. Yakir. The random Weierstrass zeta function II. Fluctuations of the electric flux through rectifiable curves. *J. Stat. Phys.*, 190(10):164, 2023. [44](#)
- [145] A. Soshnikov. Determinantal random point fields. *Uspekhi Mat. Nauk*, 55(5(335)):107–160, 2000. [13](#)
- [146] A. Soshnikov. Gaussian limit for determinantal random point fields. *Ann. Probab.*, 30(1):171–187, 2002. [60](#)
- [147] J. M. Swart. A course in interacting particle systems. *arXiv:1703.10007*, 2017. [147](#)
- [148] T. Trauthwein. Quantitative CLTs on the Poisson space via skorohod estimates and  $p$ -Poincaré inequalities. *Ann. Appl. Probab.*, 2025. [10](#), [12](#)
- [149] T. Trauthwein and J. E. Yukich. Normal approximation on the Poisson space. *preprint*, 2026. [10](#), [43](#)
- [150] J. van den Berg. A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet. *Comm. Math. Phys.*, 152:161–166, 1993. [13](#)

- [151] J. van den Berg and C. Maes. Disagreement percolation in the study of Markov fields. *Ann. Probab.*, 22(2):749–763, 1994. [11](#), [13](#), [78](#), [95](#), [96](#), [98](#)
- [152] R. van der Hofstad. Stochastic processes on random graphs. *Lecture notes for the 47th Summer School in Probability Saint-Flour*, 2017. [https://www.win.tue.nl/~rhofstad/SaintFlour\\_SPoRG.pdf](https://www.win.tue.nl/~rhofstad/SaintFlour_SPoRG.pdf). [78](#), [85](#)
- [153] C. Villani. *Optimal Transport: Old and New*, volume 338. Springer, 2009. [16](#)
- [154] D. Weitz. Combinatorial criteria for uniqueness of Gibbs measures. *Rand. Struct. Alg.*, 27(4):445–475, 2005. [78](#), [83](#), [107](#)
- [155] D. Weitz. Counting independent sets up to the tree threshold. *Proc. Annual ACM Symp. Th. Computing*, 2006:140–149, 2006. [109](#), [111](#)
- [156] J. C. Wierman. New bounds for the site percolation threshold of the hexagonal lattice. *J. Phys. A: Math. Theor.*, 55(22):224017, 2022. [101](#)
- [157] A. Xia and J. E. Yukich. Normal approximation for statistics in geometric probability. *Adv. in Appl. Probab.*, page 934–972, 2015. [12](#)
- [158] J. E. Yukich. *Probability Theory of Classical Euclidean Optimization Problems*. Springer, 1998. [97](#)
- [159] S. Ziesche. Sharpness of the phase transition and lower bounds for the critical intensity in continuum percolation on  $\mathbb{R}^d$ . *Ann. l’I.H.P. Prob. Stat.*, 54(2):866–878, 2018. [11](#), [100](#), [104](#), [105](#), [106](#)