PROCESS LEVEL MODERATE DEVIATIONS FOR STABILIZING FUNCTIONALS

P. EICHELSBACHER AND T. SCHREIBER

Abstract. Functionals of spatial point process often satisfy a weak spatial dependence condition known as *stabilization*. In this paper we prove process level moderate deviation principles (MDP) for such functionals, which is a level-3 result for empirical point fields as well as a level-2 result for empirical point measures. The level-3 rate function coincides with the so-called specific information. We show that the general result can be applied to prove MDPs for various particular functionals, including random sequential packing, birth-growth models, germ-grain models and nearest neighbor graphs.

1. Introduction and main results

In their seminal works, Föllmer & Orey [4], Olla [10] and Georgii [5] established process-level large deviation principles for (lattice-indexed) Gibbsian random fields, which was further pushed into the point process set-up by Georgii & Zessin [6] and Georgii [7]. The rate function was given there in terms of the relative entropy density. In the present paper we discuss the corresponding process-level moderate deviation problem. The moderate deviation principle we establish applies for well-behaved configuration functionals on Poisson point processes, namely the so-called stabilizing functionals. The rate function turns out to be the specific relative information of null-measures, which is the moderate deviation analogue for the specific relative entropy.

1.1. **Terminology.** Consider a real-valued translation invariant functional $\xi(x, \sigma)$ defined on all pairs (x, σ), where $x \in \mathbb{R}^d$ and σ is a finite point configuration in $\mathbb{R}^d$ containing x. Moreover, for $x \notin \sigma$ write $\xi(x, \sigma) := \xi(x, \sigma \cup \{x\})$. Let $\mathcal{P}$ be a homogeneous Poisson point process on $\mathbb{R}^d$, with a certain intensity $\tau > 0$ to remain fixed throughout the paper, and denote by Π the distribution of $\mathcal{P}$ on the space Σ of locally finite point configurations in $\mathbb{R}^d$. For formal completeness we represent the space Σ as the set of all locally finite and simple (all atoms of mass 1) counting measures σ on $\mathbb{R}^d$, endowed with the σ-field $\mathcal{F}$ generated by the mappings $\Sigma \ni \sigma \mapsto \sigma(A)$ for all bounded Borel $A \subseteq \mathbb{R}^d$.

---

[0] *American Mathematical Society 2000 subject classifications.* Primary 60F05, Secondary 60D05

[0] *Key words and phrases.* Moderate deviations, random Euclidean graphs, random sequential packing

One crucial assumption imposed on ξ throughout this paper is the so-called exponential stabilization, see [1, 11, 12, 13]. We say that ξ is stabilizing (at intensity τ) if for each $x \in \mathbb{R}^d$ there exists an a.s. finite random variable $R(x) := R^\xi(x, \mathcal{P})$ (a radius of stabilization) and $\xi_\infty(x) := \xi_\infty(x, \mathcal{P})$ (the limit of ξ) such that, with probability one, $\xi(x, (\mathcal{P} \cap B_{R(x)}(x)) \cup \sigma) = \xi_\infty(x)$ for all locally finite $\sigma \subseteq \mathbb{R}^d \setminus B_{R(x)}(x)$. More generally, for a locally finite point configuration $\sigma \subseteq \mathbb{R}^d$ we consider the stabilization radius $R(x) := R^\xi(x, \sigma)$ of ξ at x in σ defined so that $\xi(x, (\sigma \cap B_{R(x)}(x)) \cup \sigma')$ takes the same value for all locally finite $\sigma' \subseteq \mathbb{R}^d \setminus B_{R(x)}(x)$. We put $R^\xi(x, \sigma) := +\infty$ if this does not hold for any finite $R(x)$. Our exponential stabilization requirement means that ξ is stabilising (at the intensity τ fixed throughout the paper) and at each point the stabilisation radius exhibits exponentially decaying tail, i.e.

(E) There exists $c > 0$ such that, for r large enough,

$$\mathbb{P}(R(x) > r) \leq \exp(-cr).$$

In [1, 11, 12, 13] a lot of examples of stabilizing functionals are discussed. In Section 2 we will focus on random sequential packing models, birth-growth models, germ-grain models and nearest neighbor graphs.

Under the stabilization condition as stated above, the Poisson point process $\mathcal{P}$ with probability one takes its values in the space $\Sigma^\xi \subseteq \Sigma$, defined to consist of all configurations for which the value of ξ can be uniquely determined at each configuration point. Thus, in order to avoid unnecessary formal subtleties, we simply extend the functional ξ in some artificial way, say by setting $\xi(x, \sigma) := 0$ if $R^\xi(x, \sigma) = +\infty$. Since this happens with probability 0 if σ is given by $\mathcal{P}$, this extension does not affect our results while guaranteeing that $\Sigma^\xi = \Sigma$. For a configuration $\sigma \in \Sigma$ let $\xi[\sigma]$ be its ξ-marked version, where each point $x \in \sigma$ is marked with the corresponding value $\xi(x, \sigma)$. In particular, $\xi[\Sigma]$ is the space of all possible ξ-marked point configurations. We formally represent $\xi[\Sigma]$ as the space of simple point measures of $\mathbb{R}^d \times \mathbb{R}$ and we endow it with the σ-field $\hat{\mathcal{F}}$ generated by the mappings $\xi[\Sigma] \ni \hat{\sigma} \mapsto \hat{\sigma}(A_1 \times A_2)$ for all bounded Borel $A_1 \subseteq \mathbb{R}^d,\ A_2 \subseteq \mathbb{R}$.

For each Borel measurable region $A \subseteq \mathbb{R}^d$ consider the σ-field $\mathcal{F}_A \subseteq \mathcal{F}$ generated by the mappings $\Sigma \ni \sigma \mapsto \sigma(B),\ B \subseteq A,\ B$ bounded and measurable. Define also $\hat{\mathcal{F}}_A$ to be the σ-field generated by the mappings $\xi[\Sigma] \ni \hat{\sigma} \mapsto \hat{\sigma}(B_1 \times B_2)$ with B_1 ranging over bounded Borel subsets of A and with bounded Borel measurable $B_2 \subseteq \mathbb{R}$. We shall write Π_A for the restriction of Π to $\mathcal{F}_A$. We say that a function $\Phi : \Sigma \to \mathbb{R}$ is local if it is measurable with respect to $\mathcal{F}_A$ for some bounded A. Likewise, $\hat{\Phi} : \xi[\Sigma] \to \mathbb{R}$ is local if it is measurable with respect to $\hat{\mathcal{F}}_A$ for some bounded A. Consider the space $B_{loc}(\Sigma)$ consisting of all the bounded local functions on Σ with the topology determined by the convergence: $\Phi_n \to \Phi$ as $n \to \infty$ iff $||\Phi_n - \Phi||_\infty := \sup_{\sigma \in \Sigma} |\Phi_n(\sigma) - \Phi(\sigma)| \to_{n\to\infty} 0$ and there exists bounded $A \subseteq \mathbb{R}^d$ such that $\Phi, \Phi_1, \Phi_2, \ldots$ are all $\mathcal{F}_A$-measurable. The definition of $B_{loc}(\xi[\Sigma])$ is completely analogous.

We say that a set function $\Theta : \mathcal{F} \to \mathbb{R}$ is a signed local measure on Σ iff $\Theta(\bigcup_{i=1}^\infty S_i) = \sum_{i=1}^\infty \Theta(S_i)$ with the RHS series absolutely convergent, whenever S_i are pairwise disjoint and all S_i are $\mathcal{F}_A$-measurable for some bounded $A \subseteq \mathbb{R}^d$. Denote by $\mathcal{M}^{0,\theta}_{loc}(\Sigma)$ the space of all translation invariant signed local measures on Σ with total mass $\int_\Sigma 1 d\Theta = 0$ (and hence referred to as null-measures in the sequel) endowed with the topology $\mathcal{T}$ taken to be the weakest one which makes continuous the mappings $\Theta \mapsto \langle \Phi, \Theta \rangle := \int_\Sigma \Phi d\Theta$ for all $\Phi \in B_{loc}(\Sigma)$. Observe that the mapping $\Phi \mapsto \langle \Phi, \Theta \rangle$ is continuous in $B_{loc}(\Sigma)$ for each $\Theta \in \mathcal{M}^{0,\theta}_{loc}(\Sigma)$. Moreover, it is clear that $B_{loc}(\Sigma)$ is the topological dual of $(\mathcal{M}^{0,\theta}_{loc}(\Sigma), \mathcal{T})$. In a completely analogous way we define the space $\mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$ of translation invariant signed local null-measures on $\xi[\Sigma]$, endowed with the corresponding $B_{loc}(\xi[\Sigma])$-weak topology $\hat{\mathcal{T}}^\xi$. Consider now the natural mark-forgetting mapping $\pi : \xi[\Sigma] \to \Sigma$ and observe that we have $\hat{\Theta} \circ \pi^{-1} \in \mathcal{M}^{0,\theta}_{loc}(\Sigma)$ for $\hat{\Theta} \in \mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$. Define

$$\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma) := \left\{ \Theta \in \mathcal{M}^{0,\theta}_{loc}(\Sigma) \mid \exists \hat{\Theta} \in \mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma]) : \Theta = \hat{\Theta} \circ \pi^{-1} \right\}$$

and endow $\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$ with the topology $\mathcal{T}^\xi := \pi(\hat{\mathcal{T}}^\xi)$. Note that $\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma) \subseteq \mathcal{M}^{0,\theta}_{loc}(\Sigma)$ where the equality may but does not have to occur. Indeed, for $\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$ the existence of $\hat{\Theta} \in \mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$ with $\Theta = \hat{\Theta} \circ \pi^{-1}$ may imply a version of σ-additivity stronger than just local whenever ξ itself is not a local functional. For similar reasons, the topology $\mathcal{T}^\xi$ is stronger or equal to the topology induced by the inclusion $\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma) \subseteq \mathcal{M}^{0,\theta}_{loc}(\Sigma)$. Observe also that $\mathcal{T}^\xi$ can be equivalently characterized as the weakest topology to make continuous the mappings $\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma) \ni \Theta \mapsto [\hat{\Phi}, \Theta]_\xi := \langle \hat{\Phi}, \xi[\Theta] \rangle$ for all $\hat{\Phi} \in B_{loc}(\xi[\Sigma])$. Clearly, $\hat{\Phi} \mapsto [\hat{\Phi}, \Theta]_\xi$ is $B_{loc}(\xi[\Sigma])$-continuous for each Θ in $\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$. Moreover, it is easily seen that $B_{loc}(\xi[\Sigma])$ can be regarded as the topological dual for $(\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma), \mathcal{T}^\xi)$ with respect to the duality $[\cdot, \cdot]_\xi$.

For $\lambda > 0$, let Q_λ be the cube of volume λ centered at 0, i.e. $Q_\lambda = [-\sqrt[d]{\lambda}/2, \sqrt[d]{\lambda}/2]^d$. For a finite point configuration $\sigma \subseteq Q_\lambda$ we define the *empirical point field*

$$\psi^\xi_\lambda(\sigma) := \frac{1}{\lambda} \int_{Q_\lambda} \delta_{\xi[\tau_x \operatorname{Per}_\lambda(\sigma)]} dx, \tag{1}$$

where $\tau_x y := y - x$ stands for the usual shift operator, while $\operatorname{Per}_\lambda(\sigma)$ is the configuration arising by periodically copying σ on disjoint translates of Q_λ, i.e. $\operatorname{Per}_\lambda(\sigma) := \bigcup_{i \in \mathbb{Z}^d} \tau_{\sqrt[d]{\lambda} i} \sigma$. In other words, the empirical process arises as a probability measure on the space $\xi[\Sigma]$ of marked point configurations, by normalized integration over $x \in Q_\lambda$ of unit masses concentrated at ξ-marked $\sqrt[d]{\lambda}$-periodized versions of σ shifted by x. It is clear that $\psi^\xi_\lambda(\sigma)$ is a translation invariant measure.

Throughout this paper we focus our interest on empirical point fields generated by the Poisson point process $\mathcal{P}$

$$\Psi^{\xi}_{\lambda} := \psi^{\xi}_{\lambda}(\mathcal{P} \cap Q_{\lambda}). \tag{2}$$

We consider also the centered versions

$$\bar{\Psi}^{\xi}_{\lambda} := \Psi^{\xi}_{\lambda} - \mathbb{E}\Psi^{\xi}_{\lambda}.$$

Observe that $\bar{\Psi}^{\xi}_{\lambda}$ is a $\mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$-valued random element and that we almost surely have $\pi(\bar{\Psi}^{\xi}_{\lambda}) \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$. It can be shown that the following law of large numbers holds almost surely in $\mathcal{T}^{\xi}$ topology

$$\lim_{\lambda\to\infty} \Psi^{\xi}_{\lambda} = \lim_{\lambda\to\infty} \mathbb{E}\Psi^{\xi}_{\lambda} = \xi[\Pi],$$

this is a consequence of the exponential stabilization and we refer the reader to [15] for details. Note that by $\xi[\Pi]$ we mean here the transport of the law Π on Σ by the map $\xi[\cdot] : \Sigma \to \xi[\Sigma]$. The main goal of this paper is to establish a process level (level-3) and empirical measure level (level-2) moderate deviation principle (MDP) for $\bar{\Psi}^{\xi}_{\lambda}$ under the assumption that the stabilizing ξ satisfies a level-1 moderate deviation principle, as made precise below. The rate function of this MDP turns out to admit representation in terms of the *specific information* functional $I(\cdot|\Pi)$ defined for a local null-measure $\Theta \in \mathcal{M}^{0,\theta}_{loc}(\Sigma)$ by

$$I(\Theta|\Pi) := \frac{1}{2} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_{\lambda}}} \left(\frac{d\Theta_{|Q_{\lambda}}}{d\Pi_{Q_{\lambda}}} \right)^2 d\Pi_{Q_{\lambda}} \tag{3}$$

if $\Theta \ll_{loc} \Pi$ and $I(\Theta) := +\infty$ otherwise. Note that the local absolute continuity requirement stated with $\ll_{loc}$ above means simply that $\Theta_{|Q_{\lambda}} \ll \Pi_{Q_{\lambda}}$ for all λ, with the $|Q_{\lambda}$ operation standing for the restriction of its argument measure to $\mathcal{F}_{Q_{\lambda}}$. The existence of the limit in (3) will be established in Lemma 1, further properties of the specific information functional will be discussed in Section 5.

1.2. **Process level moderate deviation principles.** We say (see [3]) that a family of probability measures $(\mu_{\varepsilon})_{\varepsilon>0}$, on some topological space $\mathcal{Y}$ obeys a large deviation principle (LDP) with speed ε and good rate function $I(\cdot) : \mathcal{Y} \to \mathbb{R}^{+}_{0} \cup \{+\infty\}$ if

- I is lower semi-continuous and has compact level sets $N_L := \{x \in \mathcal{Y} : I(x) \leq L\}$, for every $L \in [0, \infty)$.
- For every open set $G \subseteq \mathcal{Y}$ it holds

$$\liminf_{\varepsilon\to 0} \varepsilon \log \mu_{\varepsilon}(G) \geq - \inf_{x\in G} I(x). \tag{4}$$

- For every closed set $A \subseteq \mathcal{Y}$ it holds

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(A) \leq - \inf_{x \in A} I(x). \tag{5}$$

Similarly we will say that a family of random variables $(Y_\varepsilon)_{\varepsilon > 0}$ with topological state space $\mathcal{Y}$ obeys a large deviation principle with speed ε and good rate function $I(\cdot) : \mathcal{Y} \to \mathbb{R}_0^+ \cup \{+\infty\}$ if the sequence of their distributions does. Formally a moderate deviation principle is nothing but an LDP. However, we will speak about a moderate deviation principle (MDP) for a sequence of random variables, whenever the scaling of the corresponding random variables is between that of an ordinary law of large numbers and that of a central limit theorem.

To proceed with the statement of the moderate deviation principle for $\bar{\Psi}^\xi_\lambda$, we let α_λ be such that $\alpha_\lambda \to \infty$ and $\alpha_\lambda \lambda^{-1/2} \to 0$.

Below, we shall assume that ξ is a bounded exponentially stabilizing functional, as required in **(E)**. From the results and methods of Section 4.3 in [1] it follows that

Proposition 1. *For each $\hat{\Phi} \in B_{loc}(\xi[\Sigma])$ there exists the limit*

$$V[\xi; \hat{\Phi}] := \lim_{\lambda \to \infty} \lambda \operatorname{Var}\left(\langle \hat{\Phi}, \bar{\Psi}^\xi_\lambda \rangle \right) \tag{6}$$

providing the infinite-volume variance density for $\langle \hat{\Phi}, \bar{\Psi}^\xi_\lambda \rangle$. Moreover, for each $R > 0$ we have

$$V_R[\xi] := \sup_{||\hat{\Phi}||_\infty \leq 1,\, D(\hat{\Phi}) \leq R} V[\xi; \hat{\Phi}] < +\infty, \tag{7}$$

where $D(\hat{\Phi})$ stands for the infimum of $r > 0$ such that $\hat{\Phi}$ is $\mathcal{F}_{B(0,r)}$-measurable.

Note that we multiply rather than dividing by λ in (6) because of the normalization for Ψ^ξ_λ being already present in (1) and (2). Further, we impose on ξ the following additional condition

(L) For the log-Laplace functional

$$\Lambda^\xi_{\lambda;\alpha_\lambda}(\hat{\Phi}) := \frac{1}{\alpha_\lambda^2} \log \mathbb{E} \exp \left(\alpha_\lambda \lambda^{1/2} \langle \hat{\Phi}, \bar{\Psi}^\xi_\lambda \rangle \right), \quad \hat{\Phi} \in B_{loc}(\xi[\Sigma]) \tag{8}$$

we have

$$\lim_{\lambda \to \infty} \Lambda^\xi_{\lambda;\alpha_\lambda}(\hat{\Phi}) = \frac{1}{2} V[\xi; \hat{\Phi}]. \tag{9}$$

In fact, this condition is a usual ingredient needed to establish the moderate deviation principle for $\langle \hat{\Phi}, \bar{\Psi}^\xi_\lambda \rangle$ with rate function $\mathbb{R} \ni t \mapsto t^2/(2V[\xi; \hat{\Phi}])$ by an application of the Gärtner-Ellis theorem [Theorem 2.3.6 in [3]], see [2]. In a number of cases the exponential stabilization seems to be enough to guarantee **(L)**, see [2] and Section 2 below, however at present we do not know if the boundedness and exponential stabilization do imply the condition **(L)** in general.

The following process-level moderate deviation theorem is the first main result of our paper.

Theorem 1. *For a bounded functional ξ for which the conditions* **(E)** *and* (**L**) *hold, the family* $(\alpha_\lambda^{-1}\lambda^{1/2}\bar{\Psi}_\lambda^\xi)_\lambda$ *satisfies the moderate deviation principle on the space* $(\mathcal{M}_{loc}^{0,\theta}(\xi[\Sigma]),\hat{\mathcal{T}}^\xi)$ *with speed* α_λ^2 *and with the good rate function* $I^\xi(\cdot|\xi[\Pi])$ *defined for* $\hat{\Theta} \in \mathcal{M}_{loc}^{0,\theta}(\xi[\Sigma])$ *to be* $I(\Theta|\Pi)$ *if* $\hat{\Theta} = \xi[\Theta]$ *and* $+\infty$ *otherwise, with* $\xi[\Theta]$ *standing for the tranport of* Θ *by the map* $\xi[\cdot]$.

At this point, it is very natural to compare our Theorem 1 for stabilizing functionals with the corresponding process level large deviation principles for Gibbs measures, see Föllmer & Orey [4], Olla [10] and Georgii [5], where the rate function was given in terms of the relative entropy density. In fact, the specific information can be roughly interpreted as the (halved) second derivative of the relative entropy density $h(\cdot|\cdot)$ at the equilibrium measure Π in that, vaguely, $h(\Pi+\delta\Theta|\Pi) \approx \delta^2 I(\Theta|\Pi) + o(\delta^2)$. Of course in such formulation this imprecise formula can be given a definite meaning only at the level of finite volume approximations of $h(\cdot|\Pi)$ and $I(\cdot|\Pi)$, yet it provides an intuition that our MDP could be regarded as a *local version* of the process level LDP, *differentiated* at equilibrium. In this context, the role played by the empirical probabilities in the LDP set-up, here is taken over by the *null-measures* arising upon centering laws of the empirical point fields Ψ_λ^ξ. To the best of our knowledge there is no moderate deviations result on process level in the literature.

1.3. **Empirical measure level moderate deviation principles.** Usually as a consequence from the process level (level-3) MDP one obtains its empirical measure level (level-2) counterpart, which is proven via the contraction principle. In our present context we prefer, however, to establish the level-2 MDP directly, thus avoiding certain topological intricacies and getting a simpler formula for the rate function, still in a variational rather than explicit form though.

Let us consider the *empirical point process*

$$Z_\lambda^\xi := \frac{1}{\lambda}\sum_{x\in\mathcal{P}_\tau\cap Q_\lambda} \delta_{\xi(x,\mathrm{Per}_\lambda(\mathcal{P}\cap Q_\lambda))}. \tag{10}$$

and its centered version $\bar{Z}_\lambda^\xi := Z_\lambda^\xi - \mathbb{E}Z_\lambda^\xi$. Moreover let us denote by $\mathcal{M}(\mathbb{R})$ the real vector space of finite variation signed measures on $\mathbb{R}$. Equip $\mathcal{M}(\mathbb{R})$ with the *weak topology* generated by the sets $\{U_{f,x,\delta}, f \in \mathcal{C}_b(\mathbb{R}), x \in \mathbb{R}, \delta > 0\}$, where $\mathcal{C}_b(\mathbb{R})$ is the space of bounded continuous functions on $\mathbb{R}$ and with

$$U_{f,x,\delta} := \{\nu \in \mathcal{M}(\mathbb{R}),\ |\langle f,\nu\rangle - x| < \delta\}.$$

The Borel-σ-field generated by the weak topology is denoted by $\mathcal{B}$. It is well known, that since the collection of linear functionals $\{\nu \mapsto \langle f,\nu\rangle : f \in \mathcal{C}_b(\mathbb{R})\}$ is separating in $\mathcal{M}(\mathbb{R})$, this topology makes $\mathcal{M}(\mathbb{R})$ into a locally convex, Hausdorff topological vector space, whose topological dual is the preceding collection, hereafter identified with $\mathcal{C}_b(\mathbb{R})$.

In analogy with the corresponding results for process level objects, we require that ξ satisfy the exponential stabilization condition **(E)**. Under this conditions, using the results of [15], we get the following almost sure law of large numbers in the $\mathcal{C}_b(\mathbb{R})$-weak topology

$$\lim_{\lambda\to\infty} Z_\lambda^\xi = \lim_{\lambda\to\infty} \mathbb{E} Z_\lambda^\xi = \tau\nu[\xi], \tag{11}$$

where, for Borel $B \subseteq \mathbb{R}$,

$$\nu[\xi](B) := \mathbb{P}(\xi(\mathbf{0}, \mathcal{P}) \in B), \tag{12}$$

that is to say, $\nu[\xi]$ is the law of $\xi(\mathbf{0}, \mathcal{P})$ on $\mathbb{R}$. We recall here that τ is the intensity of the Poisson point process $\mathcal{P}$. Again, using the methods and results of Section 4.3 in [1] we get

Proposition 2. *For each $f \in \mathcal{C}_b(\mathbb{R})$ there exists the limit*

$$V_f[\xi] := \lim_{\lambda\to\infty} \lambda \operatorname{Var}\left(\langle f, \bar{Z}_\lambda^\xi\rangle\right) = \lim_{\lambda\to\infty} \lambda \operatorname{Var}\left(\langle f, Z_\lambda^\xi\rangle\right).$$

Moreover, we have

$$V_f[\xi] = \tau\langle f \otimes f, \mu\rangle, \tag{13}$$

where $\mu \in \mathcal{M}(\mathbb{R} \times \mathbb{R})$ is given by

$$\mu(A_1 \times A_2) := \mathbb{P}(\xi(\mathbf{0}, \mathcal{P}) \in A_1 \cap A_2) +$$
$$\frac{1}{2}\tau \int_{\mathbb{R}^d} [\mathbb{P}(\xi(\mathbf{0}, \mathcal{P} \cup x) \in A_1,\ \xi(x, \mathcal{P} \cup \mathbf{0}) \in A_2) - \mathbb{P}(\xi(\mathbf{0}, \mathcal{P}) \in A_1)\mathbb{P}(\xi(\mathbf{0}, \mathcal{P}) \in A_2)]dx \tag{14}$$

for Borel $A_1, A_2 \subseteq \mathbb{R}^d$ and with $f \otimes f(x, y) := f(x)f(y),\ x, y \in \mathbb{R}$. The convergence of the integral in (14) is guaranteed by the exponential stabilization of ξ.

We note that the multiplication rather than division by λ in the definition of $V_f[\xi]$ above is due to the λ^{-1}-normalization already present in the definition of Z_λ^ξ. The following condition is a natural counterpart of the level-3 condition **(L)**.

(L') The log-Laplace functional

$$L_{\lambda;\alpha_\lambda}^\xi[f] := \frac{1}{\alpha_\lambda^2} \log \mathbb{E} \exp\left(\langle f, \bar{Z}_\lambda^\xi\rangle\right) \tag{15}$$

satisfies

$$\lim_{\lambda\to\infty} L_{\lambda;\alpha_\lambda}^\xi[f] = \frac{1}{2} V_f[\xi]. \tag{16}$$

Under appropriate additional conditions **(L')** would follow as a direct consequence of **(L)**, indeed, taking $\hat{\Phi}_f(\xi[\sigma])$ to be $\sum_{x\in\sigma\cap[0,1]^d} f(\xi(x,\sigma))$ we see that $\langle f, \bar{Z}_\lambda^\xi\rangle$ differs from $\langle \hat{\Phi}_f, \bar{\Psi}_\lambda^\xi\rangle$ just by a boundary-order term, which can be easily dealt with e.g. by considering a periodised version of the process on a torus, thus getting rid of such boundary effects. The point is, though, that thus defined $\hat{\Phi}_f$ is usually not bounded. On the other hand, for all our examples both **(L)** and **(L')** do follow from the same theory developed in [2]. Therefore we have decided to formulate both these conditions separately, without resorting to tedious general considerations which would not add any extra examples to our list of applications.

The following level-2 moderate deviation theorem is our second main result.

Theorem 2. *For ξ satisfying both the exponential stabilization condition* **(E)** *and the condition* **(L')** *, the family $\alpha_\lambda^{-1}\lambda^{1/2}\bar{Z}_\lambda$ satisfies a MDP on $\mathcal{M}(\mathbb{R})$, endowed with the $\mathcal{C}_b(\mathbb{R})$-weak topology, with speed α_λ^2 and a convex, good rate function*

$$J^\xi(\gamma) := \sup_{f \in \mathcal{C}_b(\mathbb{R})} (\langle f, \gamma \rangle - \frac{\tau}{2} \langle f \otimes f, \mu \rangle). \tag{17}$$

2. Examples

Below, we discuss examples of stabilizing functionals for which our general level-3 and level-2 theory applies. Our presentation is borrowed from ([2]) where level-1 moderate deviation principles are established for these functionals. It should be noted that the corresponding central limit theorems, under much milder conditions (no homogeneity required) have been established in [1].

2.1. **Random sequential packing.** The following prototypical random sequential packing model arises in diverse disciplines, including physical, chemical, and biological processes. See [14] for a discussion of the many applications, the many references, and also a discussion of previous mathematical analysis. In one dimension, this model is often referred to as the Rényi car parking model [16].

With $N(\tau\lambda)$ standing for a Poisson random variable with parameter $\tau\lambda$, let $B_{\lambda,1}$, $B_{\lambda,2}$, ..., $B_{\lambda,N(\tau\lambda)}$ be a sequence of d-dimensional balls of volume 1 whose centers are i.i.d. random d-vectors $X_1, ..., X_{N(\lambda)}$ uniformly distributed over $Q_\lambda = [-\sqrt[d]{\lambda}/2, \sqrt[d]{\lambda}/2]^d$. Without loss of generality, assume that the balls are sequenced in the order determined by marks (time coordinates) in $[0, 1]$. Let the first ball $B_{\lambda,1}$ be *packed*, and recursively for $i = 2, 3, \ldots, N(\tau\lambda)$, let the i-th ball $B_{\lambda,i}$ be packed iff $B_{\lambda,i}$ does not overlap any ball in $B_{\lambda,1}, ..., B_{\lambda,i-1}$ which has already been packed. If not packed, the i-th ball is discarded.

For any finite point set $\mathcal{X} \subset \mathbb{R}^d$, assume the points $x \in \mathcal{X}$ have time coordinates which are independent and uniformly distributed over the interval $[0, 1]$. Assume unit volume balls centered at the points of $\mathcal{X}$ arrive sequentially in an order determined by the time coordinates, and assume as before that each ball is packed or discarded according to whether or not it overlaps a previously packed ball. Let $\xi(x;\mathcal{X})$ be either 1 or 0 depending on whether the ball centered at x is packed or discarded. Letting $\mathcal{X} = \mathcal{P}$ we easily see that $\xi(\cdot;\mathcal{P})$ describes the random sequential packing process as constructed above. This process depends not only on the spatial locations of points but also on their $[0,1]$-valued arrival time marks. However, this clearly does fit into our general setting by a simple generalisation to the marked case. A similar remark is due for many of the examples below but will not be repeated there.

From [1, 15] we know that ξ satisfies the exponential stabilization condition (**E**). Moreover, by Section 2 and Subsections 6.1 and 6.2 of [2] we see that ξ satisfies both the **(L)** and **(L')** conditions. In particular, our Theorems 1 and 2 do apply for the random sequential packing functional ξ. To be able to obtain **(L)** and **(L')** in [2], we had to apply stabilization methods, cumulant techniques, and exponential modification of measures.

2.2. **Spatial birth-growth models.** Our results for the prototypical packing measures as described in Subsection 2.1 above, extend to measures arising from more general packing models. Consider for example the following spatial birth-growth model in $\mathbb{R}^d$. Let $\tilde{\mathcal{P}} := \{(X_i, T_i) \in \mathbb{R}^d \times [0,1]\}$ be a spatial-temporal Poisson point process. Seeds appear at uniformly random locations $X_i \in Q_\lambda$ at times T_i i.i.d. and uniform in $[0,1]$. When a seed is born, it has initial radius ρ_i, $0 < L_- < \rho_i \le L_+ < \infty$, and thereafter the radius grows at a constant speed v_i, generating a cell growing radially in all directions. When one expanding cell touches another, they both stop growing in their respective directions. In any event, we assume that the seed radii are deterministically bounded, i.e., they never exceed a fixed cut-off and they stop growing upon reaching it. Moreover, if a seed appears at X_i and if the ball centered at X_i with radius ρ_i overlaps any of the existing cells, then the seed is discarded. Variants of this well-studied process are used to model crystal growth [17].

To proceed, for any finite point set $\mathcal{X} \subset \mathbb{R}^d$, assume the points $x \in \mathcal{X}$ have i.i.d. time marks over $[0,1]$. A mark at $x \in \mathcal{X}$ represents the arrival time of a seed at x. Assume that the seeds are centered at the points of $\mathcal{X}$, that they arrive sequentially in an order determined by the associated marks, and that each seed is accepted or rejected according to the rules above. Let $\xi(x;\mathcal{X})$ be either 1 or 0 according to whether the seed centered at x is accepted or not. Letting $\mathcal{X} = \mathcal{P} \cap Q_\lambda$ we see that $\xi(\cdot;\mathcal{P} \cap Q_\lambda)$ corresponds to the spatial birth-growth model introduced above.

Again, from [1, 15] we know that ξ satisfies the exponential stabilization condition (**E**). Moreover, by Section 3 and Subsections 6.1 and 6.2 of [2] we see that ξ satisfies both the **(L)** and **(L')** conditions. In particular, our Theorems 1 and 2 do apply for the random birth-growth functional ξ.

Remark 1. *The results of the present subsection extend to more general versions of the prototypical packing model. The stabilization analysis of* [14] *combined with* [2] *yields* **(E)** *and* **(L)**,**(L')** *in the finite input setting for the number of packed balls in the following general models: (a) models with balls replaced by particles of random (bounded) size/shape/charge, (b) cooperative sequential adsorption models, and (c) ballistic deposition models (see* [14] *for a complete description of these models). In each case, our general results apply to the functionals* ξ *putting* 1 *in the centers of accepted objects and* 0 *in the centers of rejected objects.*

2.3. **Germ-grain models.** Let $X_i,\ i \geq 1$, be i.i.d. uniformly distributed over Q_λ. Let $T, T_i,\ i \geq 1$, be i.i.d. bounded random variables, independent of the random variables $X_i, 1 \leq i \geq 1$. Consider the random grains $X_i + B_{T_i}(\mathbf{0})$ as well as the random set

$$\Xi_\lambda := \bigcup_{i=1}^{N(\tau\lambda)} (X_i + B_{T_i}(\mathbf{0})),$$

where $B_r(x)$ again denotes the Euclidean ball centered at $x \in \mathbb{R}^d$ of radius $r > 0$. The random set Ξ_λ usually goes under the name of a Boolean model (see e.g. Hall [8], pp. 141, 233 and Molchanov [9] Section 3.2, Example 2.2, p. 35).

For all $u \in \mathbb{R}^d$, let $T(u)$ be i.i.d. random variables with distribution equal to that of T. For all $x \in \mathbb{R}^d$ and all locally finite point sets $\mathcal{X} \subset \mathbb{R}^d$, denote by $V(x, \mathcal{X})$ the Voronoi cell around x with respect to $\mathcal{X}$ and let $\xi(x; \mathcal{X})$ be the Lebesgue measure of the intersection of $\bigcup_{u\in\mathcal{X}} B_{T(u)}(u)$ and $V(x, \mathcal{X})$.

For ξ thus defined, we see that $\sum_{x\in\mathcal{P}\cap Q_\lambda} \xi(x; \mathcal{P})$ is just the Lebesgue measure of Ξ_λ. Likewise, we can easily construct a functional ξ' such that $\sum_{x\in\mathcal{P}\cap Q_\lambda} \xi'(x; \mathcal{P})$ coincides with the surface area measure of Ξ^λ by defining $\xi'(x; \mathcal{X})$ to be the surface area measure of the part of $\partial\Xi_\lambda$ falling into $V(x, \mathcal{X})$.

Using [1] and [2, Section 6.3] we again see that the functionals ξ and ξ' as defined above do satisfy both the **(E)** and **(L)** + **(L')** conditions, whence our general results apply. Note that the arguments used when proving conditions **(L)** + **(L')** for Germ-grain models in [2] differ from those used for the packing models, see Section 6.3 ibidem.

2.4. k**-nearest neighbors random graphs.** Let k be a positive integer. Given a locally finite point set $\mathcal{X} \subset \mathbb{R}^d$, the k-nearest neighbors (undirected) graph on $\mathcal{X}$, denoted $NG(\mathcal{X})$, is the graph with vertex set $\mathcal{X}$ obtained by including $\{x,y\}$ 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. The k-nearest neighbors (directed) graph on $\mathcal{X}$, denoted $NG'(\mathcal{X})$, is the graph with vertex set $\mathcal{X}$ obtained by placing a directed edge between each point and its k nearest neighbors.

For all $t > 0$, let $\xi^t(x;\mathcal{X}) := 1$ if the length of the edge joining x to its nearest neighbor in $\mathcal{X}$ is less than t and zero otherwise. Moreover, for $m \in \mathbb{N}$ we shall consider functionals ξ_m^{NG} and $\xi_m^{NG'}$ taking value 1 if the degree of the vertex x in $NG(\mathcal{X})$ (respectively $NG'(\mathcal{X})$) is m, and value 0 otherwise. Clearly, as usual we shall take $\mathcal{X} := \mathcal{P}$. It follows now from [1] and [2, Section 6.3] that all the functionals ξ^t, ξ_m^{NG} and $\xi_m^{NG'}$ do satisfy **(E)**, **(L)** and **(L$'$)**, whence our general results do apply.

3. Proof of Theorem 1

In view Proposition 1 and condition **(L)** the projective limit technique, see Corollary 4.6.11 in [3], allows us to conclude that $\bar{\Psi}_\lambda^\xi$ satisfies the moderate deviation principle in the *algebraic dual* $[B_{loc}(\xi[\Sigma])]'$ endowed with $B_{loc}(\xi[\Sigma])$-weak topology, with the good rate function

$$[\Lambda^\xi]^*(\hat{\Theta}) := \sup_{\hat{\Phi}\in B_{loc}(\xi[\Sigma])} (\langle \hat{\Phi}, \hat{\Theta}\rangle - \frac{1}{2}V[\xi;\hat{\Phi}]), \quad \hat{\Theta} \in [B_{loc}(\xi[\Sigma])]'. \tag{18}$$

In view of Theorem 4 below, we have

$$[\Lambda^\xi]^*(\hat{\Theta}) = I^\xi(\hat{\Theta}|\xi[\Pi]) \tag{19}$$

for $\hat{\Theta} \in \mathcal{M}_{loc}^{0,\theta}(\xi[\Sigma])$. Further, it is easily seen that $\hat{\mathcal{T}}^\xi$ coincides with the topology on $\mathcal{M}_{loc}^{0,\theta}(\xi[\Sigma])$ induced by the inclusion of this space in $[B_{loc}(\xi[\Sigma])]'$ topologized as above. Thus, in view of Lemma 4.1.5 in [3], Theorem 1 will follow once we show that

$$[\Lambda^\xi]^*(\hat{\Theta}) = +\infty \tag{20}$$

for $\hat{\Theta} \in [B_{loc}(\xi[\Sigma])]' \setminus \mathcal{M}_{loc}^{0,\theta}(\xi[\Sigma])$. To this end, take $\hat{\Theta}$ with $[\Lambda^\xi]^*(\hat{\Theta}) < +\infty$ and use Proposition 1 writing

$$\left\langle \frac{\hat{\Phi}}{||\hat{\Phi}||_\infty}, \hat{\Theta} \right\rangle \le [\Lambda^\xi]^*(\hat{\Theta}) + V\left[\xi; \frac{\hat{\Phi}}{||\hat{\Phi}||_\infty}\right] \le [\Lambda^\xi]^*(\hat{\Theta}) + V_{D(\hat{\Phi})}[\xi]$$

for all $\hat{\Phi} \in B_{loc}(\xi[\Sigma])$. Consequently, we see that $\hat{\Theta}$ is a bounded linear form on $\{\hat{\Phi} \in B_{loc}(\xi[\Sigma]),\ \hat{\Phi}$ is $\mathcal{F}_A-$measurable $\}$ for each bounded Borel $A \subseteq \mathbb{R}^d$. Using the Riesz representation theorem for the restrictions of $\hat{\Theta}$ to subspaces of functions of $B_{loc}(\xi[\Sigma])$ depending only on the marked point configuration within $[-N,N]^d,\ N \to \infty$, we conclude that $\hat{\Theta} \in \mathcal{M}_{loc}(\xi[\Sigma])$. Note that this application of the Riesz representation theorem is justified because for each $N \in \mathbb{N}$ the space of finite point configurations in $[-N,N]^d$ can be embedded in the space of compact subsets of $[-N,N]^d$ endowed with the usual compact Hausdorff metric and with the resulting Borel σ-field coinciding with $\mathcal{F}_{[-N,N]^d}$. To complete the proof it is now enough to exclude the case $\hat{\Theta} \in \mathcal{M}_{loc}(\xi[\Sigma]) \setminus \mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$. However, this is easily done by noting that

$$\mathbb{P}\left(\bar{\Psi}^{\xi}_{\lambda} \in \mathcal{M}_{loc}(\xi[\Sigma]) \setminus \mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])\right) = 0$$

since Ψ^{ξ}_{λ} is translation invariant and has 0 total mass by its definition, and by observing that the space $\mathcal{M}^{0,\theta}_{loc}(\xi[\Sigma])$ is closed in $\mathcal{M}_{loc}(\xi[\Sigma])$ with respect to the $B_{loc}(\xi[\Sigma])$-weak topology. The proof is complete. □

4. Proof of Theorem 2

The proof is organised similar to that of Theorem 1. In view of Proposition 2 and the condition **(L')**, the projective limit technique, see Corollary 4.6.11 in [3], allows us to conclude that $\bar{Z}^{\xi}_{\lambda}$ satisfies the moderate deviation principle in the *algebraic dual* $[\mathcal{C}_b(\mathbb{R})]'$ endowed with $\mathcal{C}_b(\mathbb{R})$-weak topology, with the good rate function

$$[L^{\xi}]^*(\gamma) := \sup_{f \in \mathcal{C}_b(\mathbb{R})} (\langle f, \gamma \rangle - \frac{1}{2} V_f[\xi]), \quad \gamma \in [\mathcal{C}_b(\mathbb{R})]'. \tag{21}$$

In view of Lemma 4.1.5 in [3] and of Proposition 2 guaranteeing that $[L^{\xi}]^*(\gamma) = J^{\xi}(\gamma)$ for $\gamma \in \mathcal{M}(\mathbb{R})$, to complete the proof of Theorem 2 it is now enough to show that

$$[L^{\xi}]^*(\gamma) = +\infty, \quad \gamma \in [\mathcal{C}_b(\mathbb{R})]' \setminus \mathcal{M}(\mathbb{R}). \tag{22}$$

To this end, take γ with $[L^{\xi}]^*(\gamma) < +\infty$ and write for $f \in \mathcal{C}_b(\mathbb{R})$

$$\left\langle \frac{f}{||f||_\infty}, \gamma \right\rangle \leq [L^{\xi}]^*(\gamma) + \frac{1}{2} V_{f/||f||_\infty}[\xi] = [L^{\xi}]^*(\gamma) + \frac{\tau}{2} \left\langle \frac{f}{||f||_\infty} \otimes \frac{f}{||f||_\infty}, \mu \right\rangle.$$

Since the RHS is bounded, this means that γ is a bounded operator on $\mathcal{C}_b(\mathbb{R})$ and hence $\gamma \in \mathcal{M}(\mathbb{R})$ as required. This completes the proof of Theorem 2. □

5. Properties of the specific relative information

In this section we discuss a number of properties of the specific relative information, as introduced in (3). Our main purpose below is to identify the rate function in Theorem 1.

5.1. Existence.

Lemma 1. *For each translation invariant local null measure Θ on Σ there exists the limit*

$$I(\Theta|\Pi) := \frac{1}{2} \lim_{\lambda \to \infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} \left(\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}} \right)^2 d\Pi_{Q_\lambda}.$$

Moreover, we have

$$I(\Theta|\Pi) = \frac{1}{2}\sup_{\lambda>0}\frac{1}{\lambda}\int_{\Sigma_{Q_\lambda}}\left(\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}}\right)^2 d\Pi_{Q_\lambda}. \tag{23}$$

Proof. For a bounded region $A \subseteq \mathbb{R}$ write

$$I_A(\Theta|\Pi) := \frac{1}{2}\int_{\Sigma_A}\left(\frac{d\Theta_{|A}}{d\Pi_A}\right)^2 d\Pi_A. \tag{24}$$

It is clear that, by standard superadditivity argument, the proof will be completed once we show that for bounded and disjoint $A, B \subseteq \mathbb{R}^d$

$$I_{A\cup B}(\Theta|\Pi) \geq I_A(\Theta|\Pi) + I_B(\Theta|\Pi). \tag{25}$$

To establish (25), write ρ_A for the density $d\Theta_{|A}/d\Pi_A$, define ρ_B and $\rho_{A\cup B}$ likewise and let $\rho_{B|A}(\sigma_B|\sigma_A) := \rho_{A\cup B}(\sigma_A \cup \sigma_B) - \rho_A(\sigma_A)$, with σ_A and σ_B standing for generic elements of Σ_A and Σ_B respectively. It is clear that

$$\int_{\Sigma_B}\rho_{A\cup B}(\sigma_A\cup\sigma_B)d\Pi_B(\sigma_B) = \rho_A(\sigma_A) \text{ and hence } \int_{\Sigma_B}\rho_{B|A}(\sigma_B|\sigma_A)d\Pi_B(\sigma_B) = 0 \quad \Pi_A \text{ a.s.} \tag{26}$$

Moreover, since $\int_{\Sigma_A}\rho_A d\Pi_A = \Theta(\Sigma) = 0$, interchanging A and B in (26) we are led to

$$\int_{\Sigma_A}\rho_{A\cup B}(\sigma_A\cup\sigma_B)d\Pi_A(\sigma_A) = \rho_B(\sigma_B) \text{ and } \int_{\Sigma_A}\rho_{B|A}(\sigma_B|\sigma_A)d\Pi_A(\sigma_A) = \rho_B(\sigma_B) \quad \Pi_B \text{ a.s.} \tag{27}$$

With this notation we get, using (26),

$$2I_{A\cup B}(\Theta|\Pi) = \int_{\Sigma_A}\int_{\Sigma_B}\rho^2_{A\cup B}(\sigma_A\cup\sigma_B)d\Pi_B(\sigma_B)d\Pi_A(\sigma_A) =$$
$$\int_{\Sigma_A}\int_{\Sigma_B}(\rho_A(\sigma_A) + \rho_{B|A}(\sigma_B|\sigma_A))^2 d\Pi_B(\sigma_B)d\Pi_A(\sigma_A) =$$
$$\int_{\Sigma_A}\rho_A^2 d\Pi_A + 2\int_{\Sigma_A}\rho_A(\sigma_A)\int_{\Sigma_B}\rho_{B|A}(\sigma_B|\sigma_A)d\Pi_B(\sigma_B)d\Pi_A(\sigma_A) +$$
$$\int_{\Sigma_A}\int_{\Sigma_B}\rho^2_{B|A}(\sigma_B|\sigma_A)d\Pi_B(\sigma_B)d\Pi_A(\sigma_A) =$$
$$\int_{\Sigma_A}\rho_A^2 d\Pi_A + \int_{\Sigma_A}\int_{\Sigma_B}\rho^2_{B|A}(\sigma_B|\sigma_A)d\Pi_B(\sigma_B)d\Pi_A(\sigma_A).$$

Applying Jensen's inequality we come to

$$2I_{A\cup B}(\Theta|\Pi) \geq \int_{\Sigma_A}\rho_A^2 d\Pi_A + \int_{\Sigma_B}\left(\int_{\Sigma_A}\rho_{B|A}(\sigma_B|\sigma_A)d\Pi_A(\sigma_A)\right)^2 d\Pi_B(\sigma_B).$$

Using (27) we obtain finally

$$2I_{A\cup B}(\Theta|\Pi) \geq \int_{\Sigma_A}\rho_A^2 d\Pi_A + \int_{\Sigma_B}\rho_B^2 d\Pi_B = 2I_A(\Theta|\Pi) + 2I_B(\Theta|\Pi)$$

which yields (25) and hence completes the proof of the lemma. □

5.2. **Finite volume variational principle and lower semicontinuity.**

Lemma 2. *For a bounded region $A \subseteq \mathbb{R}^d$ we have for each $\Phi \in B(\Sigma_A)$*

$$\frac{1}{2}\operatorname{Var}(\Phi(\mathcal{P}_A)) = \sup_{\Theta \in \mathcal{M}^0(\Sigma_A)} (\langle \Phi, \Theta \rangle - I_A(\Theta|\Pi)) \tag{28}$$

with $\mathcal{P}_A$ standing for the restriction of $\mathcal{P}$ to A and where $\mathcal{M}^0(\Sigma_A)$ is the collection of all 0-total mass signed measures on Σ_A. Moreover, for each $\Theta \in \mathcal{M}^0(\Sigma_A)$ we have

$$I_A(\Theta|\Pi) = \sup_{\Phi \in B(\Sigma_A)} \left(\langle \Phi, \Theta \rangle - \frac{1}{2}\operatorname{Var}(\Phi(\mathcal{P}_A)) \right). \tag{29}$$

Proof. Fix $\Phi \in B(\Sigma_A)$ and note that for $\Theta \in \mathcal{M}^0(\Sigma_A)$ absolutely continuous w.r.t. Π_A we have

$$\langle \Phi, \Theta \rangle - I_A(\Theta|\Pi) = \int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] d\Theta - \frac{1}{2}\int_{\Sigma_A} \left(\frac{d\Theta}{d\Pi_A}\right)^2 d\Pi_A$$

because $\int_{\Sigma_A} \mathbb{E}\Phi(\mathcal{P}_A) d\Theta = 0$. Consequently,

$$\langle \Phi, \Theta \rangle - I_A(\Theta|\Pi) = \int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] \frac{d\Theta}{d\Pi_A} d\Pi_A - \frac{1}{2}\int_{\Sigma_A} \left(\frac{d\Theta}{d\Pi_A}\right)^2 d\Pi_A \leq$$

$$\frac{1}{2}\int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]^2 d\Pi_A = \frac{1}{2}\operatorname{Var}(\Phi(\mathcal{P}_A)),$$

where the last inequality follows from $f\rho - \frac{1}{2}\rho^2 \leq \frac{1}{2}f^2$ for $f := [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]$ and $\rho := \frac{d\Theta}{d\Pi_A}$. Since Θ was arbitrary with $\Theta \ll \Pi_A$ and $I_A(\Theta|\Pi) = +\infty$ for $\Theta \not\ll \Pi_A$, we conclude that

$$\operatorname{Var}(\Phi(\mathcal{P}_A)) \geq \sup_{\Theta \in \mathcal{M}^0(\Sigma_A)} (\langle \Phi, \Theta \rangle - I_A(\Theta|\Pi)). \tag{30}$$

To proceed, let $\Theta^\Phi \in \mathcal{M}^0(\Sigma_A)$ be given by $d\Theta^\Phi := [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] d\Pi_A$. We have then

$$\frac{1}{2}\operatorname{Var}(\Phi(P_A)) = \frac{1}{2}\int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]^2 d\Pi_A = \int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] d\Theta^\Phi - \frac{1}{2}\int_{\Sigma_A} \left(\frac{d\Theta^\Phi}{d\Pi_A}\right)^2 d\Pi_A =$$

$$\langle \Phi, \Theta^\Phi \rangle - I_A(\Theta^\Phi|\Pi).$$

Combining these equalities with (30) yields now (28).

The proof of (29) is analogous. Fix $\Theta \in \mathcal{M}^0(\Sigma_A)$ and write for $\Phi \in B(\Sigma_A)$

$$\langle \Phi, \Theta \rangle - \operatorname{Var}(\Phi(\mathcal{P}_A)) = \int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] d\Theta - \frac{1}{2}\int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]^2 d\Pi_A =$$

$$\int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)] \frac{d\Theta}{d\Pi_A} d\Pi_A - \frac{1}{2}\int_{\Sigma_A} [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]^2 d\Pi_A \leq$$

$$\frac{1}{2}\int_{\Sigma_A} \left(\frac{d\Theta}{d\Pi_A}\right)^2 d\Pi_A = \frac{1}{2} I_A(\Theta|\Pi),$$

where the last inequality follows from $f\rho - \frac{1}{2}f^2 \leq \frac{1}{2}\rho^2$ for $f := [\Phi - \mathbb{E}\Phi(\mathcal{P}_A)]$ and $\rho := \frac{d\Theta}{d\Pi_A}$. Since Φ was arbitrary, we see that

$$I_A(\Theta|\Pi) \geq \sup_{\Phi \in B(\Sigma_A)} \left(\langle \Phi, \Theta \rangle - \frac{1}{2} \operatorname{Var}(\Phi(\mathcal{P}_A)) \right). \tag{31}$$

To proceed with the proof of the converse inequality observe first that if $\Theta \not\ll \Pi_A$, the expression $\langle \Phi, \Theta \rangle - \frac{1}{2}\operatorname{Var}(\Phi(\mathcal{P}_A))$ can be made arbitrarily large by adjusting Φ on a region in Σ_A of non-zero total variation for Θ to which Π_A assigns zero mass. Now, for $\Theta \ll \Pi_A$ let $\Phi^\Theta := \frac{d\Theta}{d\Pi_A}$. Observe that $\mathbb{E}\Phi^\Theta(\mathcal{P}_A) = \Theta(\Sigma_A) = 0$. Write

$$I_A(\Theta|\Pi) = \frac{1}{2}\int_{\Sigma_A} [\Phi^\Theta]^2 d\Pi_A = \int_{\Sigma_A} \Phi^\Theta d\Theta - \frac{1}{2}\mathbb{E}[\Phi^\Theta]^2 = \langle \Phi^\Theta, \Theta \rangle - \frac{1}{2}\operatorname{Var}(\Phi^\Theta(\mathcal{P}_A)).$$

Putting this together with (31) yields (29). This completes the proof of the lemma. □

Lemma 3. *The mapping*

$$(\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma), \mathcal{T}_\xi) \ni \Theta \mapsto I(\Theta|\Pi)$$

is convex and lower semicontinuous.

Proof. The convexity follows immediately by the definition of $I(\cdot|\Pi)$ in view of the convexity of finite-volume functionals $I_{Q_\lambda}(\cdot|\Pi)$. Further, the variational formula (29) represents the finite volume functionals $I_{Q_\lambda}(\cdot|\Pi)$ as suprema over $\Phi \in B(Q_\lambda)$ of $\mathcal{T}_\xi$-continuous functionals, which yields the $\mathcal{T}_\xi$-lower semicontinuity for $I_{Q_\lambda}(\cdot|\Pi)$. The required $\mathcal{T}_\xi$-lower semicontinuity of $I(\cdot|\Pi)$ follows now by (23). □

5.3. **Infinite volume variational principle.**

Theorem 3. *For each $\hat{\Phi} \in B_{loc}(\xi[\Sigma])$ we have*

$$\frac{1}{2}V[\xi; \hat{\Phi}] = \sup_{\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)} \left(\langle \hat{\Phi}, \xi[\Theta] \rangle - I(\Theta|\Pi) \right).$$

Proof. We claim first that

$$\sup_{\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)} \left(\langle \hat{\Phi}, \xi[\Theta] \rangle - I(\Theta|\Pi) \right) \leq \frac{1}{2}V[\xi; \hat{\Phi}]. \tag{32}$$

For each $\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$ such that $\Theta \ll_{loc} \Pi$ and $I(\Theta|\Pi) < +\infty$ we easily conclude from the exponential stabilization assumption **(E)** and from the translational invariance of Θ that

$$\langle \hat{\Phi}, \xi[\Theta] \rangle - I(\Theta|\Pi) = \lim_{\lambda \to \infty} \int_{\Sigma_{Q_\lambda}} \langle \hat{\Phi}, \psi^\xi_\lambda(\sigma_{Q_\lambda}) \rangle d\Theta_{|Q_\lambda}(\sigma_{Q_\lambda}) - \frac{1}{2} \lim_{\lambda \to \infty} \frac{1}{\lambda} \left(\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}} \right)^2 d\Pi_{Q_\lambda}.$$

Consequently, using that Θ is a null-measure, we come to

$$\langle\hat{\Phi},\xi[\Theta]\rangle - I(\Theta|\Pi) = \lim_{\lambda\to\infty}\frac{1}{\lambda}\int_{\Sigma_{Q_\lambda}}\lambda[\langle\hat{\Phi},\psi^\xi_\lambda(\sigma_{Q_\lambda})\rangle - \mathbb{E}\langle\hat{\Phi},\Psi^\xi_\lambda\rangle]d\Theta_{|Q_\lambda}(\sigma_{Q_\lambda}) - \frac{1}{2}\lim_{\lambda\to\infty}\frac{1}{\lambda}\left(\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}}\right)^2 d\Pi_{Q_\lambda}$$

$$= \lim_{\lambda\to\infty}\frac{1}{\lambda}\int_{\Sigma_{Q_\lambda}}\lambda[\langle\hat{\Phi},\psi^\xi_\lambda(\sigma_{Q_\lambda})\rangle - \mathbb{E}\langle\hat{\Phi},\Psi^\xi_\lambda\rangle]\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}}[\sigma_{Q_\lambda}]d\Pi_{Q_\lambda}(\sigma_{Q_\lambda}) - \frac{1}{2}\lim_{\lambda\to\infty}\frac{1}{\lambda}\left(\frac{d\Theta_{|Q_\lambda}}{d\Pi_{Q_\lambda}}\right)^2 d\Pi_{Q_\lambda} \le$$

$$\frac{1}{2}\lim_{\lambda\to\infty}\frac{1}{\lambda}\int_{Q_\lambda}\left(\lambda[\langle\hat{\Phi},\psi^\xi_\lambda(\sigma_{Q_\lambda})\rangle - \mathbb{E}\langle\hat{\Phi},\Psi^\xi_\lambda\rangle]\right)^2 d\Pi_{Q_\lambda}(\sigma_{Q_\lambda}) = \frac{1}{2}\lim_{\lambda\to\infty}\lambda\,\mathrm{Var}(\langle\hat{\Phi},\bar{\Psi}^\xi_\lambda\rangle),$$

where the last inequality comes from $f\rho - \frac{1}{2}\rho^2 \le \frac{1}{2}f^2$ applied for $f := \lambda[\langle\hat{\Phi},\psi^\xi_\lambda\rangle - \mathbb{E}\langle\hat{\Phi},\Psi^\xi_\lambda\rangle]$ and $\rho := d\Theta_{|Q_\lambda}/d\Pi_{Q_\lambda}$. Now, in view of Proposition 1 the last limit equals $\frac{1}{2}V[\xi;\hat{\Phi}]$. Thus, since Θ was arbitrary with $\Theta \ll_{loc} \Pi$ with $I(\Theta|\Pi) < +\infty$ and since $I(\Theta'|\Pi) = +\infty$ for $\Theta' \not\ll_{loc} \Pi$, we conclude the inequality (32) as required.

To establish the converse inequality, for each $N \in \mathbb{N}$ construct the measure $\Theta^{\Phi\circ\xi}_N$ by partitioning $\mathbb{R}^d$ into translates $Q_N[i],\ i\in\mathbb{Z}^d$ of the cube Q_N and setting

$$\Theta^{\Phi\circ\xi}_N := \frac{1}{N}\int_{Q_N}\tau_x[\bigoplus_{i\in\mathbb{Z}^d}\Theta^{\Phi\circ\xi}_{N:i}]dx \tag{33}$$

with

$$\frac{d\Theta^{\Phi\circ\xi}_{N:i}}{d\Pi_{Q_N[i]}}[\sigma_{Q_N[i]}] := \lambda[\langle\hat{\Phi},\psi^\xi_{Q_N[i]}(\sigma_{Q_N[i]})\rangle - \mathbb{E}\langle\hat{\Phi},\Psi^\xi_N\rangle],\ \ \sigma_{Q_N[i]}\subseteq Q_N[i], \tag{34}$$

where

$$\psi^\xi_{Q_N[i]}(\sigma_{Q_N[i]}) = \frac{1}{N}\int_{Q_N[i]}\delta_{\xi[\tau_x\,\mathrm{Per}_N(\sigma_{Q_N[i]})]}dx,$$

see (1), and where $\bigoplus_{i\in\mathbb{Z}^d}\Theta^{\Phi\circ\xi}_{N:i}$ is given for a cylinder event $S = S_1\times\ldots\times S_k,\ S_j\in\mathcal{F}_{Q_N[j]}$ by

$$[\bigoplus_{i\in\mathbb{Z}^d}\Theta^{\Phi\circ\xi}_{N:i}](S) = \sum_{j=1}^k \Theta^{\phi\circ\xi}_{N:i}(S_j). \tag{35}$$

Note that this definition is consistent since all $\Theta^{\Phi\circ\xi}_{N:i}$ are null-measures (have their total masses 0). Intuitively speaking, the above construction is the counterpart of taking products of probability measures in our null-measure setting. Observe also that, by definition, the measure $\Theta^{\Phi\circ\xi}_{N:i}$ coincides with the translate $\tau_v\Theta^{\Phi\circ\xi}_{N:j}$ where v is the vector joining the center of $Q_N[i]$ to the center of $Q_N[j]$. Again, roughly speaking, this construction can be regarded as a null-measure analogue of taking the product law of i.i.d. random objects.

By exponential stabilization **(E)** is clear that $\Theta^{\Phi\circ\xi}_N \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$. Moreover, by the translation invariance of the Poisson point process Π, writing $Q^*_N[A] := \bigcup_{Q_N[i]\cap A\neq\emptyset} Q_N[i]$ and $Q^\partial_N[A] := Q^*_N[A]\setminus A,\ A\subseteq\mathbb{R}^d$, in view of (33) and (34) above we have

$$[\rho_N]_\lambda(\sigma_{Q_\lambda}) := \frac{d[\Theta^{\Phi\circ\xi}_N]_{Q_\lambda}}{d\Pi_{Q_\lambda}}(\sigma_{Q_\lambda}) =$$

$$\int_{Q_N} \int_{\Sigma_{Q^\partial_N[\tau_x Q_\lambda]}} \sum_{Q_N[i] \subseteq Q^*_N[\tau_x Q_\lambda]} \lambda[\langle \hat{\Phi}, \psi^\xi_{Q_N[i]}([\sigma_{Q_\lambda} \cup \sigma]_{|Q_N[i]}) \rangle - \mathbb{E}\langle \hat{\Phi}, \Psi^\xi_{Q_N} \rangle] d\Pi_{Q^\partial_N[\tau_x Q_\lambda]}(\sigma) dx. \quad (36)$$

Using (36), Proposition 1 and exponential stabilization **(E)**, as a consequence of the method of [1] we get

$$\frac{1}{2} \lim_{N\to\infty} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} [\rho_N]^2_\lambda d\Pi_{Q_\lambda} = \frac{1}{2} V[\xi; \hat{\Phi}]. \quad (37)$$

Combining (36) and (37) with (33) and (34) we can write

$$\frac{1}{2} V[\xi; \hat{\Phi}] = \lim_{N\to\infty} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} [\rho_N]^2_\lambda d\Pi_{Q_\lambda} - \frac{1}{2} \lim_{N\to\infty} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} [\rho_N]^2_\lambda d\Pi_{Q_\lambda} =$$

$$\lim_{N\to\infty} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} [\rho_N]_\lambda d[\Theta^{\Phi\circ\xi}_N]_{|Q_\lambda} - \frac{1}{2} \lim_{N\to\infty} \lim_{\lambda\to\infty} \frac{1}{\lambda} \int_{\Sigma_{Q_\lambda}} \left(\frac{d[\Theta^{\Phi\circ\xi}_N]_{Q_\lambda}}{d\Pi_{Q_\lambda}} \right)^2 d\Pi_{Q_\lambda} =$$

$$\lim_{N\to\infty} \left(\langle \hat{\Phi}, \xi[\Theta^{\Phi\circ\xi}_N] \rangle - I(\Theta^{\Phi\circ\xi}_N | \Pi) \right).$$

This implies that

$$\sup_{\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)} \left(\langle \hat{\Phi}, \xi[\Theta] \rangle - I(\Theta|\Pi) \rangle \right) \geq \frac{1}{2} V[\xi; \hat{\Phi}]. \quad (38)$$

Putting (32) and (38) together completes the proof of the theorem. □

Theorem 4. *For each* $\Theta \in \mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma)$ *we have*

$$I(\Theta|\Pi) = \sup_{\hat{\Phi} \in B_{loc}(\xi[\Sigma^\xi])} \left(\langle \hat{\Phi}, \xi[\Theta] \rangle - \frac{1}{2} V[\xi; \hat{\Phi}] \right).$$

Proof. In view of the convexity and lower semicontinuity of $I(\Theta|\Pi)$ on $(\mathcal{M}^{0,\theta}_{loc,\xi}(\Sigma), \mathcal{T}_\xi)$, as stated in Lemma 3, our assertion follows immediately by the Duality Lemma 4.5.8 in [3] applied for the duality $[\hat{\Phi}, \Theta]_\xi := \langle \hat{\Phi}, \xi[\Theta] \rangle$. □

Acknowledgements. Special thanks are due to Joe E. Yukich, whose ideas and comments have motivated us to prepare this paper. Tomasz Schreiber also wishes to gratefully acknowledge the support from the Polish Minister of Scientific Research and Information Technology grant 1 P03A 018 28 (2005-2007).

References

[1] Yu. Baryshnikov and J. E. Yukich (2005), Gaussian limits for random measures in geometric probability, *Annals Appl. Prob.*, **15**, 1A, 213-253.

[2] Yu. Baryshnikov, P. Eichelsbacher, T. Schreiber and J. E. Yukich (2005), Moderate Deviations for some Point Measures in Geometric Probability, submitted, electronically available on the arXiv, *math.PR/0603022*.

[3] A. Dembo and O. Zeitouni (1998), Large Deviations Techniques and Applications, Springer. Second Ed.

[4] Föllmer, H.; Orey, S. Large Deviations for the Empirical Field of a Gibbs Measure. *Ann. Probab.* **16**, pp. 961-977 (1988).

[5] Georgii, H-O. Large Deviations and Maximum Entropy Principle for Interacting Random Fields on $\mathbb{Z}^d$. *Ann. Probab.* **21**, pp. 1845-1875 (1993).

[6] Georgii, H.-O., Zessin, H. (1993) Large deviations and the maximum entropy principle for marked point random fields. *Probab. Theory Relat. Fields* **96**, 177-204.

[7] Georgii, H.-O. (1994) Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction. *Probab. Theory Relat. Fields* **99**, 171-195.

[8] P. Hall (1988), Introduction to the Theory of Coverage Processes, Wiley, New York.

[9] Molchanov, I.S. (1993) *Limit Theorems for Unions of Random Closed Sets*, Lecture Notes in Mathematics **1561**, Springer.

[10] Olla, S. Large Deviations for Gibbs Random Fields. *Probab. Th. Rel. Fields,* **77**, pp. 343-357 (1988).

[11] M.D. Penrose (2005), Multivariate spatial central limit theorems with applications to percolation and spatial graphs, *Ann. Probab.*, **33**, 1945- 1991.

[12] M.D. Penrose (2005), Convergence of random measures in geometric probability, preprint.

[13] M.D. Penrose and J.E. Yukich (2001), Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11**, 1005-1041.

[14] M.D. Penrose and J.E. Yukich (2002), Limit theory for random sequential packing and deposition, *Ann. Appl. Probab.* **12**, 272-301.

[15] M.D. Penrose and J.E. Yukich (2003), Weak laws of large numbers in geometric probability, *Ann. Appl. Probab.*, **13**, pp. 277-303.

[16] A. Rényi, Théorie des éléments saillants d'une suite d'observations (1962), In *Colloquium on Combinatorial Methods in Probability Theory*, 104-115, Mathematical Institut, Aarhus Universitet, Denmark.

[17] D. Stoyan, W. Kendall, and J. Mecke (1995), Stochastic Geometry and Its Applications, John Wiley and Sons, Second Ed.

Peter Eichelsbacher, Fakultät für Mathematik, Ruhr-Universität Bochum, NA 3/68, 44780 Bochum, Germany: `peter.eichelsbacher@ruhr-uni-bochum.de`

Tomasz Schreiber, Faculty of Mathematics and Computer Science, Nicholas Copernicus University, Toruń, Poland: `tomeks@mat.uni.torun.pl`