WEAKLY SELF-AVOIDING WALK

IN A PARETO-DISTRIBUTED RANDOM POTENTIAL

WOLFGANG KÖNIG
NICOLAS PÉTRÉLIS
RENATO SOARES DOS SANTOS
WILLEM VAN ZUIJLEN

06 June, 2023

ABSTRACT. We investigate a model of continuous-time simple random walk paths in $\mathbb{Z}^d$ undergoing two competing interactions: an attractive one towards the large values of a random potential, and a self-repellent one in the spirit of the well-known weakly self-avoiding random walk. We take the potential to be i.i.d. Pareto-distributed with parameter $\alpha > d$, and we tune the strength of the interactions in such a way that they both contribute on the same scale as $t \to \infty$.

Our main results are (1) the identification of the logarithmic asymptotics of the partition function of the model in terms of a random variational formula, and, (2) the identification of the path behaviour that gives the overwhelming contribution to the partition function for $\alpha > 2d$: the random-walk path follows an optimal trajectory that visits each of a finite number of random lattice sites for a positive random fraction of time. We prove a law of large numbers for this behaviour, i.e., that all other path behaviours give strictly less contribution to the partition function. The joint distribution of the variational problem and of the optimal path can be expressed in terms of a limiting Poisson point process arising by a rescaling of the random potential. The latter convergence is in distribution and is in the spirit of a standard extreme-value setting for a rescaling of an i.i.d. potential in large boxes, like in (KLMS09).

CONTENTS

AMS Subject Classification: Primary 60H25, 60G70. Secondary 82C44, 60F10, 60G55, 60G57.

Keywords: random walk in random potential, random variational problem, parabolic Anderson model, path localisation, intermittent islands, weakly self-avoiding walk, Poisson point process convergence, spatial extreme-value analysis, Pareto potential distribution.

1. Introduction and main results

In the last decades, there was a significant interest in the study of random motions that are attracted to certain regions defined by a surrounding random medium. The most-studied type of models is called a random motion in a random potential, which appears in the study of the parabolic Anderson model (PAM). The methods have been refined and extended significantly in recent years, and a number of situations have been successfully treated in detail. The present paper makes a contribution to this line of research by studying a model that combines attraction with repulsion and shows, as a result, a much more pronounced behaviour.

We explain the background and the main message of this paper in Section 1.1, give some literature remarks in Section 1.7, introduce the model in Section 1.2 and also a crucial rescaling in Section 1.3. The key variational formula and the main results are presented in Section 1.4. In Section 1.5 we provide some heuristic explanations for our results. The remainder of the paper is described in Section 1.6.

1.1 The main purpose

In this paper, we study one of the fundamental models of a random motion in a random medium: the random walk in random potential. This is closely connected to the parabolic Anderson model (PAM), the Cauchy problem for the Laplace operator with random potential. Indeed, in the Feynman–Kac formula for the solution, the random walk is strongly attracted by the local regions in which the potential has particularly high values, in competition with a not too far distance that has to be travelled to that region and to a not too high probabilistic price that the walker pays for staying a long time in that region. The attractive regions are called intermittent islands for the solution of the PAM. Their locations are random and homogeneously distributed. In fact, properly rescaled, they converge to a Poisson point process, as comes out of a spatial extreme-value analysis. The typical behaviour of the walk is that it is attracted to just one of these islands (characterised by an optimal relation between size, height of the potential and distance to the starting site); indeed, it rushes quickly to that region and stays there for the remaining time.

In the present paper, we add another feature to the model: a weak self-repulsion of the path, defined by some negative exponential term in the spirit of the well-known weakly self-avoiding random walk. We tune the strength of this self-repellence in such a way that it effectively competes with the attractive part. In this model, the path would pay significantly if it would spend much time in just one of the islands. We are interested in understanding the typical path behaviour. It is tempting to expect that the path will now visit several of these islands after each other. The purpose of this paper is to find out whether or not this is true and to formulate and prove details about the typical behaviour. We consider an i.i.d. Pareto random potential with parameter $\alpha > d$, for which localization effects are particularly pronounced.

Our main findings are that an interesting and characteristic random variational formula describes the behaviour of the walk, defined on a limiting probability space that carries the limiting Poisson point process. For $\alpha > 2d$, the variational formula is almost surely positive and a subtle analysis of its optimizers shows that the walk visits a certain finite choice of islands and spends in each of them a certain amount of time. The order of the visits is dictated by the length of the shortest trajectory that covers all these islands. Our proof shows that the totality of all other behaviours of the random path has vanishing probability. That is, we prove a detailed localization result for the entire path. The main properties of the variational formula that allow this are the compactness property and the uniqueness of the minimizer, almost surely. Interestingly, the behaviour of this variational formula changes for $\alpha \in (d, 2d)$, we comment on this in Remark 1.10.

1.2 The model

Let $d \in \mathbb{N}$ and $\xi = (\xi(z))_{z\in\mathbb{Z}^d}$ be a random potential with distribution $\mathbf{P}$ that consists of i.i.d. random variables. Let $\mathbb{P}$ be the law of a continuous-time simple random walk $X = (X_s)_{s\geq 0}$ on the lattice $\mathbb{Z}^d$ with generator the discrete Laplacian Δ starting from the origin. We take into account two types of microscopic interactions. The random walk interacts with the random field ξ and undergoes a self-repulsion of strength

β. This leads us to associate with every trajectory X the Hamiltonian

$$H_t^{(\xi,\beta)}(X) = \int_0^t \xi(X_s)\, \mathrm{d}s - \beta \int_0^t \int_0^t \mathbb{1}_{\{X_s = X_u\}}\, \mathrm{d}s\, \mathrm{d}u, \tag{1.1}$$

where $\beta \in [0, \infty)$ is the intensity of the self-repulsion. The first term is the interaction with the random potential ξ, the second is the self-intersection local time (SILT), the amount of time pairs at which the random walk is at the same site. We introduce a polymer measure $\mathbb{P}_t^{(\xi,\beta)}$ that is absolutely continuous with respect to $\mathbb{P}$ (more precisely, to its restriction to paths on $[0, t]$) with Radon-Nikodym derivative given by

$$\frac{\mathrm{d}\mathbb{P}_t^{(\xi,\beta)}}{\mathrm{d}\mathbb{P}}(X) = \frac{\mathrm{e}^{H_t^{(\xi,\beta)}(X)}}{Z_t^{(\xi,\beta)}}, \tag{1.2}$$

where the normalizing constant $Z_t^{(\xi,\beta)} = \mathbb{E}[\mathrm{e}^{H_t^{(\xi,\beta)}}]$ is the partition function of the model. We call this model the weakly self-avoiding random walk in a random potential. We want to study its large-t behaviour.

When $\beta = 0$, the Feynman-Kac formula shows that $Z_t^{(\xi,0)}$ equals the total mass $U(t) = \sum_{x\in\mathbb{Z}^d} u(t, x)$ of the solution u to the parabolic Anderson model (PAM), the heat equation with random potential ξ:

$$\partial_t u(t, x) = \Delta u(t, x) + \xi(x)u(t, x), \qquad x \in \mathbb{Z}^d,$$

with localised initial condition $u(0, 0) = 1$ and $u(0, x) = 0$ for $x \in \mathbb{Z}^d \setminus \{0\}$. On the other side, with $\xi = 0$ and $\beta > 0$, $\mathbb{P}_t^{(0,\beta)}$ is the law of a weakly self-avoiding walk in continuous time. Since the SILT is not an additive functional, there is no obvious connection between this model and any partial differential equation.

It is clear that the Hamiltonian is a function of the walker's local times ℓ_t, given by

$$\ell_t(z) = \ell_t^{(X)}(z) = \int_0^t \mathbb{1}\{X_s = z\}\, \mathrm{d}s, , \qquad z \in \mathbb{Z}^d. \tag{1.3}$$

Indeed,

$$H_t^{(\xi,\beta)}(X) = \sum_{z\in\mathbb{Z}^d} \xi(z)\ell_t(z) - \beta \sum_{z\in\mathbb{Z}^d} \ell_t(z)^2 = \langle \xi, \ell_t\rangle - \beta\|\ell_t\|_2^2, \tag{1.4}$$

where we wrote $\langle\cdot,\cdot\rangle$ for the standard inner product on $\mathbb{Z}^d$ and $\|\cdot\|_2$ for the standard ℓ^2-norm on $\mathbb{Z}^d$.

In earlier work on the PAM, it turned out that the model is the easier to analyse and the resulting picture is more pronounced for heavy-tailed potentials. Here we will assume that the potential variables $\xi(z)$ at all sites $z \in \mathbb{Z}^d$ are independently Pareto-distributed with parameter $\alpha > d$, i.e., have the distribution function

$$F(r) = \mathbf{P}(\xi(z) \le r) = 1 - r^{-\alpha}, \qquad r \ge 1. \tag{1.5}$$

In particular, we have $\xi(z) \ge 1$ for all $z \in \mathbb{Z}^d$, almost surely. This is the most heavy-tailed distribution for which the PAM is well defined; indeed, (GM90, Theorem 2.1) says that $\alpha > d$ is necessary and sufficient for partition function for $\beta = 0$ to be finite. Hence, our model is well-defined for any $\beta \ge 0$.

1.3 Rescaling and point measures

It is the purpose of this paper to study the counterplay between the effects coming from the two terms in the Hamiltonian and the underlying probability distribution of the walk. To make sure that these three effects (i.e., attraction by the potential, self-repulsion and entropy – see the heuristics in Section 1.5) all run on the same scale, we take β depending on t as follows. Fix a parameter $\theta \in (0, \infty)$ and set

$$\beta_t := \theta\, \frac{t^{q-1}}{(\log t)^q}, \qquad \text{where} \quad q = \frac{d}{\alpha - d}. \tag{1.6}$$

Note that q and the large-t behaviour of β_t are increasing in α; for $\alpha \le 2d$, we have that $\beta_t \to 0$ as $t \to \infty$. To reduce the amount of parameters, in the following we write

$$H_t^{(\xi)} := H_t^{(\xi,\beta_t)}, \qquad Z_t^{(\xi)} := Z_t^{(\xi,\beta_t)}, \qquad \mathbb{P}_t^{(\xi)} := \mathbb{P}_t^{(\xi,\beta_t)}. \tag{1.7}$$

We denote by

$$r_t := \Big(\frac{t}{\log t}\Big)^{1+q} \tag{1.8}$$

an important characteristic spatial length scale. More precisely, r_t will turn out to be the typical distance of the relevant islands from the origin at which we will find $(X_s)_{s\in[0,t]}$ with high probability under $\mathbb{P}_t^{(\xi)}$. Furthermore, it will turn out that the largest potential values in boxes of radius $\approx r_t$ are of order $r_t^{d/\alpha}$. It is convenient to express statements like these in terms of point processes, i.e., random variables with values in the set $\mathcal{M}_{\rm p}((0,\infty)\times\mathbb{R}^d)$ of Radon measures on $(0,\infty)\times\mathbb{R}^d$ with values in $\mathbb{N}_0\cup\{\infty\}$, also called point measures since they are of the form $\sum_{n\in\mathbb{N}}\delta_{x_n}$ with $x_n\in(0,\infty)\times\mathbb{R}^d$. The crucial point is that the rescaled point process

$$\Pi_t = \sum_{z\in\mathbb{Z}^d}\delta_{\big(\frac{\xi(z)}{r_t^{d/\alpha}},\frac{z}{r_t}\big)} \tag{1.9}$$

converges as $t\to\infty$, weakly with respect to the vague topology in $\mathcal{M}_{\rm p}((0,\infty)\times\mathbb{R}^d)$, towards a Poisson point process (PPP) Π with intensity measure $\alpha f^{-1-\alpha}\,{\rm d}f\otimes{\rm d}y$:

$$\Pi\sim {\rm PPP}\big((0,\infty)\times\mathbb{R}^d,\alpha f^{-1-\alpha}\,{\rm d}f\otimes{\rm d}y\big). \tag{1.10}$$

This is a basic result from spatial extreme-value analysis; see Lemma 2.4 for the precise statement. We will often write the points in $(0,\infty)\times\mathbb{R}^d$ as (f,y). The process Π may also be seen as a standard Poisson point process in $\mathbb{R}^d$ with Fréchet-distributed i.i.d. marks. We will write the probability with respect to Π also by $\mathbf{P}$.

In order to formulate the path behaviour of the walk in terms of the local times $\ell_t(x)=\ell_t^{(X)}(x)=\int_0^t\mathbb{1}\{X(s)=x\}\,{\rm d}s$, we need to rescale them in time by t and in space by r_t. Those rescaled local times are considered as a density with respect to Π_t of the measure $W_t^{(\xi,X)}$ defined by

$$\frac{{\rm d}W_t^{(\xi,X)}}{{\rm d}\Pi_t}(f,y)=\frac{\ell_t(yr_t)}{t},\qquad (f,y)\in(0,\infty)\times\mathbb{R}^d, \tag{1.11}$$

where we have extended the local times to a function $\ell_t\colon\mathbb{R}^d\to[0,t]$ satisfying $\ell_t=0$ on $\mathbb{R}^d\setminus\mathbb{Z}^d$. We will often omit the superscripts and write simply W_t for $W_t^{(\xi,X)}$. Note that W_t depends on ξ, and assigns to each pair of site and corresponding potential value the proportion of time that the walker has spend on that site by time t. It does not encode the number nor the order of the visits of the random walk to the sites. W_t lies in the set $\mathcal{W}$ of all measures μ on $(0,\infty)\times\mathbb{R}^d$ with total mass $\mu((0,\infty)\times\mathbb{R}^d)\leq 1$. We equip $\mathcal{W}$ with the vague topology as well. By (Kle08, Corollary 13.31) and (Kal83, Theorem 15.7.7), $\mathcal{W}$ is a compact Polish space, which will be convenient for the formulation of our results and proofs. Our main object of study will be W_t. Certainly the rescaled local times are an object of high interest themselves, but their behaviour may be deduced from that of Π_t and W_t.

By using the identities

$$r_t^{d/\alpha}t=r_t\log t\quad\big(\tfrac{d}{\alpha}=1-\tfrac{1}{1+q}\big)\qquad\text{and}\qquad \beta_t t^2=\theta r_t\log t, \tag{1.12}$$

from (1.4) it is easily seen that

$$\begin{aligned} H_t^{(\xi)}(X)&=\sum_{z\in\mathbb{Z}^d}\xi(z)\ell_t(z)-\beta_t\sum_{z\in\mathbb{Z}^d}\ell_t(z)^2=\sum_{z\in\frac{\mathbb{Z}^d}{r_t}}t\xi(zr_t)\frac{\ell_t(zr_t)}{t}-\beta_t t^2\sum_{z\in\frac{\mathbb{Z}^d}{r_t}}\Big(\frac{\ell_t(zr_t)}{t}\Big)^2\\ &=r_t\log t\int_{(0,\infty)\times\mathbb{R}^d}\big[fw(f,y)-\theta w(f,y)^2\big]\,{\rm d}\Pi_t(f,y),\qquad\text{with } w=\frac{{\rm d}W_t}{{\rm d}\Pi_t},\end{aligned} \tag{1.13}$$

i.e., the Hamiltonian is an explicit functional of the rescaled local times and the point process Π_t. This is the starting point of our analysis.

1.4 Main results: convergence towards a variational formula

In this section we formulate and comment on the main results of our paper. In Theorem 1.1 (a) we show that, for any $\alpha > d$, the log of the partition function $Z_t^{(\xi)}$ in (1.2) divided by $r_t \log t$ converges in law by a non-trivial random variable Ξ. In part (b) of Theorem 1.1, we show that, for $\alpha > 2d$, the rescaled local times measure W_t of a typical trajectory sampled from the mixture of $\mathbf{P}$ and $\mathbb{P}_t^{(\xi)}$ converges in distribution as well. Afterwards, in Theorem 1.4, we identify Ξ and the limiting local times process by means of a random variational formula and its unique maximizer. Contrary to the parabolic Anderson model, which corresponds to $\theta = 0$, this maximizer is – with probability larger than 0 – not a Dirac measure. However, which may be quite unexpected, the support of this maximizer is still finite. More precisely, the number of points in the support is a random variable which attains any value of $\mathbb{N}$ with positive probability.

Let us introduce some notation for the statements below. We write $\mathbf{P} \rtimes \mathbb{P}_t^{(\xi)}$ for the mixture of the laws of $\mathbf{P}$ and $\mathbb{P}_t^{(\xi)}$, i.e.,

$$\mathbf{P} \rtimes \mathbb{P}_t^{(\xi)}(A \times B) = \mathbf{E}\big[\mathbb{1}_A(\xi)\, \mathbb{P}_t^{(\xi)}[X \in B]\big]. \tag{1.14}$$

As mentioned above, $\mathcal{W}$ is a compact Polish space. Therefore, by (Kle08, Corollary 13.30) (corollary of Prohorov's theorem) and (Kal83, Theorem 15.7.7), the set of probability measures on $\mathcal{W}$ forms a compact Polish space.

Theorem 1.1 (Convergence towards the variational formula)**.** *Fix $\theta \in (0,\infty)$ and $\alpha \in (d,\infty)$. There exist a random variable Ξ with values in $[0,\infty)$ and a random variable μ^* with values in $\mathcal{W}$, which we may and do assume to live on the $\mathbf{P}$-probability space, such that the following convergences in distribution hold:*

(a) *partition function:*

$$\frac{1}{r_t \log t} \log Z_t^{(\xi)} \overset{t\to\infty}{\Longrightarrow} \Xi. \tag{1.15}$$

(b) *law of the rescaled local times: if $\alpha \in (2d,\infty)$, then, under the mixed measure $\mathbf{P} \rtimes \mathbb{P}_t^{(\xi)}$ as in (1.14), $W_t = W_t^{(\xi,X)}$ converges in distribution to μ^*;*

$$W_t \overset{t\to\infty}{\Longrightarrow} \mu^* \quad \text{in } \mathcal{W}, \tag{1.16}$$

more precisely, for all $g \in C_{\mathrm{b}}(\mathcal{W})$,

$$\mathbf{E}\big[\mathbb{E}_t^{(\xi)}[g(W_t^{(\xi,X)})]\big] \to \mathbf{E}[g(\mu^*)]. \tag{1.17}$$

The proof of Theorem 1.1 is given in Section 5 conditionally on crucial assertions for the lower bound part of the convergence in (a) (which are proved in Section 6) and crucial assertions about the upper bound in (a) and for (b) (which are proved in Section 7).

In Theorem 1.4 below, we describe Ξ more precisely. Namely, $\Xi = \Xi(\Pi)$ (for Π as in (1.10)), where $\Xi(\mathcal{P})$ for a point measure $\mathcal{P}$ on $(0,\infty) \times \mathbb{R}^d$ is defined by

$$\Xi(\mathcal{P}) = \sup_{\mu \in \mathcal{W}} \Psi_{\mathcal{P}}(\mu), \tag{1.18}$$

where $\mathcal{W}$ is the set of subprobability measures on $(0,\infty) \times \mathbb{R}^d$ and Ψ is introduced in the following definition. Moreover, in the case $\alpha > 2d$, Theorem 1.4 states that μ^* is the unique maximizer of (1.18) for $\mathcal{P} = \Pi$.

Definition 1.2. Let $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}((0,\infty) \times \mathbb{R}^d)$. We define

(a) the energy functional $\Phi_{\mathcal{P}} \colon \mathcal{W} \to [-\infty,\infty]$ by

$$\Phi_{\mathcal{P}}(\mu) = \begin{cases} \int_{(0,\infty)\times\mathbb{R}^d} \left[f w(f,y) - \theta w(f,y)^2 \right] \mathrm{d}\mathcal{P}(f,y) & \text{if } \mu \ll \mathcal{P}, w = \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{P}}, \\ -\infty & \text{otherwise,} \end{cases} \tag{1.19}$$

(b) the entropy functional $\mathcal{D}_{\mathcal{P}}\colon \mathcal{W}\to[0,\infty]$ by

$$\mathcal{D}_{\mathcal{P}}(\mu)=\begin{cases}\sup_{Y\subset \mathrm{supp}_{\mathbb{R}^d}\mu,\#Y<\infty} D_0(Y) & \text{if } \mu\ll\mathcal{P},\\ \infty & \text{otherwise,}\end{cases}\tag{1.20}$$

where $D_0(\emptyset)=0$ and $D_0(Y)$ for a finite nonempty set $Y\subset\mathbb{R}^d$ is the smallest possible $|\cdot|$-length of a path from the origin that reaches all points in Y; i.e., the minimum over $\sum_{i=1}^N|\sigma_i-\sigma_{i-1}|$ of bijections $\sigma:\{0,\dots,N\}\to Y\cup\{0\}$ with $\sigma_0=0$. We wrote

$$\mathrm{supp}_{\mathbb{R}^d}\mu=\{y\in\mathbb{R}^d\colon \exists f>0,(f,y)\in\mathrm{supp}\,\mu\},\tag{1.21}$$

for the support of the projection of μ on $\mathbb{R}^d$.

(c) the functional $\Psi_{\mathcal{P}}\colon\mathcal{W}\to[-\infty,\infty)$ by

$$\Psi_{\mathcal{P}}(\mu)=\begin{cases}\Phi_{\mathcal{P}}(\mu)-q\mathcal{D}_{\mathcal{P}}(\mu) & \text{if } \Phi_{\mathcal{P}}(\mu)<\infty,\\ -\infty & \text{otherwise.}\end{cases}\tag{1.22}$$

◇

Remark 1.3 (The appearance of the energy and entropy functionals)**.** We will later (in Section 1.5 on the heuristics) see that $q\,\mathcal{D}_\Pi$ plays the role of the large-deviation rate functional for the rescaled local times on the scale $r_t\log t$; hence we called it an entropy functional. Observe that, see for example (1.13),

$$H_t^{(\xi)}(X)=(r_t\log t)\,\Phi_{\Pi_t}(W_t^{(\xi,X)}),\tag{1.23}$$

and, therewith,

$$Z_t^{(\xi)}=\mathbb{E}\big[\mathrm{e}^{(r_t\log t)\,\Phi_{\Pi_t}(W_t)}\big].\tag{1.24}$$

This says that the random walker gains on the exponential scale $r_t\log t$ the potential reward that is given for $w=\frac{\mathrm{d}W_t}{\mathrm{d}\Pi_t}$ by $\int_{(0,\infty)\times\mathbb{R}^d}fw(f,y)\,\mathrm{d}\mathcal{P}(f,y)$ and it pays the self-repellence price that is given by the expression $\int_{(0,\infty)\times\mathbb{R}^d}\theta w(f,y)^2\,\mathrm{d}\mathcal{P}(f,y)$. See Section 1.5 for a more precise heuristic explanation. ◇

Let us now formulate what we announced before:

Theorem 1.4 (Identification of the limits)**.** *Fix $\theta\in(0,\infty)$ and $\alpha\in(d,\infty)$.*

(a) *Let Ξ be the random variable as in Theorem 1.1. Then, $\mathbf{P}$-almost surely,*

$$\Xi=\Xi(\Pi).$$

(b) *For $\alpha\in(2d,\infty)$, there exists a random variable μ^* with values in $\mathcal{W}$ such that the convergence in (1.16) holds and, $\mathbf{P}$-almost surely,*

$$\Psi_\Pi(\mu^*)=\Xi(\Pi).\tag{1.25}$$

This maximizer is $\mathbf{P}$-almost surely unique in the sense that, $\mathbf{P}$-almost surely, $\Psi_\Pi(\nu)<\Psi_\Pi(\mu^*)$ for any $\nu\in\mathcal{W}\setminus\{\mu^*\}$. Furthermore,

- *$\mathbf{P}$-almost surely, μ^* is a probability measure with finite support and $\mu^*\ll\Pi$. In particular, $\mathbf{P}(\mu^*=0)=\mathbf{P}(\#\,\mathrm{supp}\,\mu^*=0)=0$.*
- *For all $k\in\mathbb{N}$, $\mathbf{P}(\#\,\mathrm{supp}\,\mu^*=k)>0$.*

Consequently,

$$\mathbf{P}(\Xi(\Pi)\in(0,\infty))=1.\tag{1.26}$$

Theorem 1.4 (b) will be a straightforward consequence of some results presented in Section 2, namely Theorem 2.8 (a), Lemma 2.9 and Lemma 2.10. The proof of Theorem 1.4 (a) is part of our proof of Theorem 1.1 in Section 5.

Remark 1.5 (Interpretation of Theorem 1.4 (b))**.** Since μ^* has finite support and is absolutely continuous with respect to Π, there exist (random) $k^* \in \mathbb{N}$ and $(f_1^*, y_1^*), \dots, (f_{k^*}^*, y_{k^*}^*) \in \operatorname{supp}(\Pi)$ and $w_1^*, \dots, w_{k^*}^* \in (0,1]$ satisfying $\sum_{i=1}^{k^*} w_i^* = 1$ such that

$$\mu^* := \sum_{i=1}^{k^*} w_i^* \delta_{(f_i^*, y_i^*)}.$$

Hence, if we interpret the convergence in (1.16) as almost sure convergence, the typical path under $\mathbb{P}_t^{(\xi)}$ spends $\sim w_i^* t$ time units in the site $\sim \lfloor y_i^* r_t \rfloor$ with value $\xi(\lfloor y_i^* r_t \rfloor) \sim f_i^* r_t^{d/\alpha}$ for any $i \in \{1, \dots, k^*\}$, but in all other sites only $o(t)$ time units (or does not even reach them).

The above is an informal interpretation. More formally, we obtain the following consequence from Theorem 1.1 (b). Given $h \in C_c((0,\infty)\times\mathbb{R}^d)$, observe that the function $g : \mathcal{W} \to \mathbb{R}$ defined by $g(\mu) = \int_{(0,\infty)\times\mathbb{R}^d} h \, d\mu$ is continuous and bounded on $\mathcal{W}$, i.e., $g \in C_b(\mathcal{W})$. Because

$$g(W_t^{(\xi,X)}) = \int_{(0,\infty)\times\mathbb{R}^d} h(f,y) \frac{\ell_t(r_t y)}{t} \, d\Pi_t(f,y) = \sum_{z\in\mathbb{Z}^d} h\Big(\frac{\xi(z)}{r_t^{d/\alpha}}, \frac{z}{r_t}\Big) \frac{\ell_t(z)}{t},$$

$$g(\mu^*) = \int_{(0,\infty)\times\mathbb{R}^d} h \, d\mu^* = \sum_{i=1}^{k^*} w_i^* h(f_i^*, y_i^*),$$

and since $W_t \Longrightarrow \mu^*$ in $\mathcal{W}$, see (1.17), we obtain

$$\mathbf{E}\Big[\sum_{z\in\mathbb{Z}^d} \frac{\mathbb{E}_t^{(\xi)}[\ell_t(z)]}{t} h\Big(\frac{\xi(z)}{r_t^{d/\alpha}}, \frac{z}{r_t}\Big)\Big] \longrightarrow \mathbf{E}\Big[\sum_{i=1}^{k^*} w_i^* h(f_i^*, y_i^*)\Big].$$

◇

Remark 1.6. [Generalization of Theorem 1.1 (b)] Actually, we are able to prove a slightly more general but more abstract convergence than in Theorem 1.1 (b). Indeed, let us write $\mathcal{L}_t^{(\xi)}$ for the law of W_t under $\mathbb{P}_t^{(\xi)}$, so that for a Borel set $A \subset \mathcal{W}$ we have

$$\mathcal{L}_t^{(\xi)}(A) = \mathbb{P}_t^{(\xi)}(\{X : W_t^{(\xi,X)} \in A\}) = \mathbb{E}_t^{(\xi)}[\mathbb{1}\{W_t^{(\xi,X)} \in A\}] = \int \mathbb{1}\{W_t^{(\xi,X)} \in A\} \, d\mathbb{P}_t^{(\xi)}(X). \tag{1.27}$$

Then for μ^* as in Theorem 1.1, we can show the following convergence in distribution (see Remark 5.6),

$$\mathcal{L}_t^{(\xi)} \overset{t\to\infty}{\Longrightarrow} \delta_{\mu^*} \text{ with respect to the weak topology in the space of probability measures on } \mathcal{W}. \tag{1.28}$$

More precisely, for all continuous bounded functionals f defined on the set of probability measures on $\mathcal{W}$,

$$\mathbf{E}[f(\mathcal{L}_t^{(\xi)})] \overset{t\to\infty}{\Longrightarrow} \mathbf{E}[f(\delta_{\mu^*})]. \tag{1.29}$$

The convergence in Theorem 1.1 (b) follows from this by taking $f(\nu) = \int g \, d\nu$ in (1.29) for $g \in C_b(\mathcal{W})$ so that $f(\mathcal{L}_t^{(\xi)}) = \mathbb{E}_t^{(\xi)}[g(W_t^{(\xi,X)})]$ and $f(\delta_{\mu^*}) = g(\mu^*)$, implying (1.17). ◇

Remark 1.7 (Large-deviations explanation)**.** Standard ideas from the theory of large deviations applied to the formula in (1.24) already suggest that the statements in Theorems 1.1 and 1.4 are true. Indeed, if $(W_t)_{t\in(0,\infty)}$ would satisfy a large-deviations principle (LDP) on the scale $r_t \log t$ with rate function $q\mathcal{D}_\Pi$, and if the limit $\Pi_t \Longrightarrow \Pi$ could be combined with this LDP, and if the energy functional Φ_Π would have appropriate continuity and boundedness properties, then Varadhan's lemma would imply the validity of our main statement. Roughly, this is also our strategy for proving the theorems, but a lot of technicalities need to be overcome along the way. ◇

Remark 1.8 (The order of visits)**.** The entropy functional $q\mathcal{D}_\Pi$ says nothing about the order in which the trajectory visits all the points of Π, but this is implicitly expressed in the definition of $\mathcal{D}_\Pi$: it is the order that gives the minimal total trajectory length. This is reminiscent of the famous traveling-salesman problem, but we refrain from going into questions about the shortest trajectory here. ◇

Remark 1.9 (A particular case: $d = 1$)**.** In dimension 1, both the expressions of the energy functional Φ_Π in (1.19) and of the entropy functional $\mathcal{D}_\Pi$ in (1.20) turn out to be much easier. To be more specific, we consider $x, z \in [0,\infty)^2$ and $\mu \in \mathcal{W}$ such that $\mu \ll \Pi$ and $\min \operatorname{supp}_{\mathbb{R}} \mu = -x$ and $\max \operatorname{supp}_{\mathbb{R}} \mu = z$. Then, $\mathcal{D}_\Pi(\mu) = (x + z) - \min\{x, z\}$ since it is the shortest distance that one has to travel so that starting from the origin both sites $-x$ and z are visited. Moreover, by screening effect in dimension 1, that is since every site in $[-x, z]$ is visited by a trajectory that reaches both $-x$ and z, any (f, y) belonging to $\operatorname{supp} \Pi$ with $y \in [-x, z]$ may be in the support of such $\mu \in \mathcal{W}$ without increasing the entropy. Note that, $\mathbf{P}$-a.s. $\Pi((0,\infty) \times \{0\}) = 0$ and therefore, it is sufficient to consider (x, z) that are not simultaneously null. We introduce the order statistics $(f^{(i)}_{[-x,z]})_{i\in\mathbb{N}}$ of the field inside $[-x, z]$ such that for a sequence $(y^{(i)}_{[-x,z]})_{i\in\mathbb{N}}$

$$\operatorname{supp} \Pi \cap (0,\infty) \times [-x, z] = \{(f^{(i)}_{[-x,z]}, y^{(i)}_{[-x,z]}) : i \in \mathbb{N}\}.$$

In Section 3.2 a function φ_k for $k \in \mathbb{N}$, see (3.6), is introduced that allows us to describe the supremum over the energy functional over the measures that are supported $(0,\infty) \times [-x, z]$ as follows

$$\sup_{\mu\in\mathcal{W}\colon\ \operatorname{supp}\mu\subset(0,\infty)\times[-x,z]} \Phi_\Pi(\mu) = \varphi_{k_\star}\big(f^{(1)}_{[-x,z]}, \dots, f^{(k_\star)}_{[-x,z]}\big)$$

where $k_\star$ is a function of the order statistics $(f^{(i)}_{[-x,z]})_{i\in\mathbb{N}}$, more precisely,

$$k_\star = \inf\Big\{j \in \mathbb{N}\colon\ j f^{(j+1)}_{[-x,z]} \le \sum_{i=1}^{j} f^{(i)}_{[-x,z]} - 2\theta\Big\}.$$

Thus, formula (1.18) for $\mathcal{P} = \Pi$ can be strongly simplified since the variational formula computed in μ only depends on the leftmost and rightmost points in $\operatorname{supp}\mu$. To be more precise, we obtain

$$\Xi(\Pi) := \sup_{x,y\in[0,\infty)^2\setminus\{(0,0)\}} \varphi_{k_\star}\big(f^{(1)}_{[-x,z]}, \dots, f^{(k^*)}_{[-x,z]}\big) - q(x + z) - q \min\{x, z\}. \tag{1.30}$$

◇

Remark 1.10 (Suggested scenario for $\alpha \in (d, 2d)$)**.** In the course of our proofs for Theorems 1.1 and 1.4 we in particular show that the characteristic variational formula $\Xi(\Pi)$ is finite almost surely for $\alpha > d$ and positive for $\alpha > 2d$. The latter assertion seems crucial for the behaviour of the path in the random potential. It is not easy to give a short argument for that; apparently the PPP possesses sufficiently many sufficiently high potential values with not too large distances between them, such that trajectories exist for which it is worth paying the travels in order to profit from spending time in those large potentials.

This is different for $\alpha \in (d, 2d)$. Indeed, in a forthcoming paper we will show that both $\{\Xi(\Pi) = 0\}$ and $\{\Xi(\Pi) > 0\}$ have positive probability here. This can be roughly explained as follows: With positive probability the PPP, like for $\alpha > 2d$, possesses sufficiently many high potentials with not too large distances. Also, the complement has positive probability, leading to no such preferable locations as it is not worth travelling that far to profit from the large potential. We conjecture that the intermediate order statistics need to be considered to reflect the true behaviour of the random path. A closer description of this scenario will be given in a future work. ◇

1.5 Heuristic explanation

Let us give here an explanation of the main result, the limiting assertion in Theorem 1.1, jointly with the identification of the limit in Theorem 1.4. In Section 6 we will turn the following heuristics into a proof of the lower bound (however, the proof of the upper bound in Section 7 is very different).

We need to understand the large-t behaviour of the partition function $Z^{(\xi)}_t$ defined in (1.7), i.e., the expectation of $\mathrm{e}^{H^{(\xi)}_t}$ with β_t defined in (1.6) and the Hamiltonian as in (1.13). The first step is to understand the scales on which the probability from the simple random walk and the contribution from the potential ξ run, where we first ignore the self-intersection term and concentrate on the potential-interaction term. Hence, this part of the heuristics is the same as for the behaviour of the PAM with Pareto-distributed potential in (KLMS09); let us give an overview now. Note that there is a competition for the random walk between a reward (called 'energy') from staying much time in sites with extremely large values of the potential and the probabilistic cost (called 'entropy') to reach such preferable sites quickly: travelling far, the walker finds a larger potential value, but this is more costly. We need to find an optimal balance.

As in (KLMS09), we obtain a lower bound by inserting the indicator on the event $\mathcal{A}_t^{z,s}$ that the walker wanders on some fixed shortest path to a site z during the time interval $[0, st)$ and stays at z during $[st, t]$. Since the random walk has generator Δ, the probability of this event is

$$\mathbb{P}(\mathcal{A}_t^{z,s}) = \mathrm{Poi}_{2dst}(|z|)(2d)^{-|z|}\mathrm{e}^{-(1-s)2dt}\#\{\text{ shortest paths } 0 \longleftrightarrow z\}.$$

Taking $|z| \gg t$, using Stirling's estimation for the term $|z|!$ that appears in the Poi-term, we see that the dominating terms in the exponent are $|z|\log|z|$ and $|z|\log(st)$, so that, dropping all lower-order terms,

$$\mathbb{P}(\mathcal{A}_t^{z,s}) \approx \exp\left\{|z|\left[\log\frac{t}{|z|} + \log s\right]\right\}.$$

Now let us examine the contribution from the potential ξ. In order to obtain a preferably large lower bound, we pick z as a maximizing point of the potential ξ within a box of radius r. According to the Pareto-tails, we are able to pick z such that $\xi(z) \approx r^{d/\alpha}$, and this site will be close to the boundary of that ball, i.e., $|z| \approx r$. Hence, from the stay at z during $[st, t]$, the potential contributes $\approx \mathrm{e}^{t(1-s)r^{d/\alpha}}$. The potential values that the random walk experiences on the fast rush during $[0, st]$ are negligible. Hence, we have the lower bound

$$Z_t^{(\xi)} \geq \exp\left\{r\left[\log\frac{t}{r} + \log s\right]\right\}\mathrm{e}^{tr^{d/\alpha}}\mathrm{e}^{-str^{d/\alpha}},$$

and we have still the freedom to optimize over small s and large r. The optimal choice of $s \in [0, 1]$ for the second and the last term is $s \approx \frac{1}{t}r^{1-d/\alpha}$, which implies the lower bound

$$Z_t^{(\xi)} \geq \exp\left\{r\log\frac{t}{r} + tr^{d/\alpha} - \log\left(tr^{d/\alpha-1}\right)\right\} = \exp\left\{tr^{d/\alpha} - \frac{d}{\alpha}r\log r\right\}. \tag{1.31}$$

The maximal r satisfies $tr^{d/\alpha-1} = 1 + \log r$, and this is asymptotically satisfied by $r = r_t = (t/\log t)^{1+q}$ as in (1.8) with $q = \frac{d}{\alpha-d}$ as in (1.6). Then both the energy term $tr^{d/\alpha}$ and the entropy term $-\frac{d}{\alpha}r\log r \approx -qr_t\log t$ are on the scale $r_t\log t$. Interestingly, the latter comes exclusively from the probability of the crucial event $\mathcal{A}_t^{z,s}$, after optimizing on $s \approx 1/\log t$, whose choice depends on the potential value. This explains the appearance of the prefactor q in (1.22) and the notion of an 'entropy functional' in Definition 1.2.

So far, this was the first part of the explanation, which applied also to (KLMS09), since we considered only the potential interaction. Now let us become specific to our model, where an additional self-intersection term in the Hamiltonian appears and makes the path paying an extra energy price when staying a long time in a single site. If this time is of order t, then the price is of order $\beta_t t^2 = \theta r_t \log t$ (see (1.12)) i.e., it is on the same scale. Therefore, the strategy has to be improved by not only visiting one site, but several after each other and staying in each of them some time $\asymp t$. Standard assertions from spatial extreme-value theory guarantee that there are not only one, but many sites with potential values $\asymp r_t^{d/\alpha}$, and they are homogeneously distributed over a centered ball with radius $\approx r_t$, so there are many good candidates for sites to be visited. One needs to make a choice of the number of the visited sites and the order in which they are visited during the time interval $[0, t]$. The travel between them costs an additional price of the same order as the first travel from the origin to one of them since the distances of all these travels are on the same scale. The functional $\Phi_{\mathcal{P}}(\mu)$ in (1.19) describes the energetic gain (staying $\approx w(f, y)t$ time units in a site $\approx yr$ with potential value $fr_t^{d/\alpha}$ for all the (f, y) in $\mathcal{P}$ and paying $\theta w(f, y)^2$ for the self-intersections), and the functional $q\mathcal{D}_{\mathcal{P}}$ in (1.20) describes the exponential probabilistic cost payed by the simpe random walk. Hence, the rate functional $\Psi_{\mathcal{P}} = \Phi_{\mathcal{P}} - q\mathcal{D}_{\mathcal{P}}$ in (1.22) gives the entire exponential cost of this path strategy on the scale r_t

for $\mathcal{P} = \Pi_t$, as we explained in Remark 1.3. Then the exponential behaviour of the partition function $Z_t^{(\xi)}$ is given by the maximum of $\mu \mapsto \Psi_\Pi(\mu)$, like in Varadhan's lemma. An additional technical difficulty is the combination of the large-deviation arguments with the point process convergence $\Pi_t \to \Pi$; see Remark 1.3.

1.6 Organization of the paper

The remainder of the paper is organized as follows. In Section 2, we explain our strategy for proving Theorem 1.1 and formulate two types of intermediate results: the first comprises a deterministic version of Theorem 1.4 for point measures that possess certain properties, while the second states that Π_t and Π possess these properties. This version of Theorem 1.4 is proved in Section 3 along with fundamental compactness and continuity properties of the energy functional $\Phi_\mathcal{P}$ and the entropy functional $\mathcal{D}_\mathcal{P}$, as well as the existence and main properties of the maximizer. The proof that Π_t and Π have the good properties is given in Section 4. Hence, Sections 2–4 derive all the properties of the variational formula for $\Xi(\Pi)$ as formulated in Theorem 1.4. In Section 5 we start the proof of the large-t analysis of the model, Theorem 1.1, by formulating two main ingredients for the proof for the lower respectively upper bound (Propositions 5.2 resp. 5.4). The two propositions are proved in Sections 6 and 7, respectively.

1.7 Literature remarks

Let us give some survey on the literature on random motions in random potential and localisation properties. First examples appeared in work by Sznitman on Brownian motion among Poissonian obstacles in the early 1990s, see his monograph (Szn98). Among many other things, he proved almost-sure attraction to one island, but did not identify this island. An analogous localization result (i.e., for the solution of the PAM rather than for the random motion) in the space-discrete setting with an i.i.d. doubly-exponentially distributed potential was (GKM07). Around 2010, it turned out that the strongest attraction to the intermittent islands is present for potentials with heavy tails, since they have a particularly pronounced profile: indeed, the islands are just singletons here. This has been observed for the first time for the most heavy-tailed potential distribution, the Pareto distribution, in (KLMS09) and has been investigated in great detail in (MOS11) and also for the exponential distribution in (LM12); see the survey (Mör11). For double-exponentially distributed potential, localization (and much more) was proved in (BKdS18). Most of these localization results are formulated and proved for the solution of the PAM rather than for the random walk in the Feynmna–Kac formula. See (Ast16, Sect. 6) and (Kön16, Sect. 6.3) for two comprehensive survey texts on such localization results up to 2016.

These two survey texts triggered interest in localization of discrete-time random walks among Bernoulli traps, the (time and space) discrete version of Brownian motion among Poisson obstacles. Deep localization properties were derived in (DFSX20b, DFSX20a, DFSX21) in this setting in dimension $d \geq 2$. Similar results for a correlated random potential in $d = 1$ (with i.i.d. gaps between the obstacles) have been derived recently in (PS).

Earlier work (OR16, OR17, OR18) analysed the strongly related model of a spatial random branching walk in a Pareto-distributed random field of branching rates. For this model, this series of papers derives a description that resembles our model and results quite strongly. It turns out there that the main bulk of the particles is highly concentrated in a number of sites that are defined in terms of a Poisson point process (essentially the same as our Π); more precisely, the branching process subsequently visits points of this point process that are step for step extremal with respect to a compromise between high potential values and short distances. This precise mechanism is different from the one that is detected in the parabolic Anderson model (PAM) in (KLMS09); the main difference to that model being that the branching process is consistent and has no finite time horizon, like the PAM. With respect to our model, an additional difference is the repellent effect from the second part of our Hamiltonian.

The second feature in our model is the Hamiltonian of the famous weakly self-repellent random walk, the negative exponential of the self-intersection local time. It is here only a side-remark that the behaviour of the weakly self-repellent walk is poorly understood in dimensions $d \in \{2, 3, 4\}$, and it was a substantial challenge to investigate it in the other dimensions. See (MS13, Sla11) for surveying texts. Generally, it is expected that the typical behaviour is a more or less uniformly spread-out behaviour in space on a scale t^{γ_d} that is much larger than the scale $t^{1/2}$ of the free walk (at least in $d \leq 4$), but much less than the scale t of a ballistic walk (at least in $d \geq 2$). However, all these effects will not be seen in our model, because of the presence of the random potential. We will necessarily be working on a much rougher scale than those scales that are believed to be responsible for this spread-out behaviour, and the resulting behaviour will be much more spread-out, but for reasons that have to do with the potential and not with the self-repulsion.

1.8 Notation

We write $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \dots, \}$. For the rest of the paper, we fix $d \in \mathbb{N}$, $\theta, \alpha \in (0, \infty)$. We set $Q_R := [-R, R]^d$ for $R \in (0, \infty)$. We write $\Longrightarrow$ for convergence in distribution. We abbreviate 'Poisson point process' by 'PPP'. For $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ we write $|y| = \sum_{i=1}^d |y_i|$ for the ℓ^1 norm.

2. Preparation

In the present section, we prepare for the proof of Theorems 1.1 and 1.4 by analysing the variational formula Ξ in (1.18). On the way, we need to extract several continuity and compactness properties of the energy and entropy functionals $\Phi_{\mathcal{P}}$ and $\mathcal{D}_{\mathcal{P}}$ as functions of $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}((0, \infty) \times \mathbb{R}^d)$. For this, we keep $\mathcal{P}$ deterministic in this section, but restrict to a subclass of such $\mathcal{P}$'s for which we can prove all needed assertions and for which we can prove that the processes Π_t and Π satisfy them. We define in particular a class of good point measures, see Definition 2.7, with the characteristic that if $\mathcal{P}$ is good then $\Psi_{\mathcal{P}}$ has at most one maximizer. In Theorem 2.8 we formulate all the necessary properties for deterministic good point measures, among other things the uniqueness of the maximizer and its continuous dependence on $\mathcal{P}$. Furthermore, in Lemma 2.9 we state that Π_t and Π are almost surely good. The proofs are deferred to later sections.

It will be convenient for us to compactify some sets of $(0, \infty) \times \mathbb{R}^d$ as described next. For $h, s > 0$ we define the cone-shaped set (see also Figure 1) with height h and slope s by

$$\mathcal{H}_h^s := \left\{(f, y) \in (0, \infty) \times \mathbb{R}^d \colon\ f > s|y| + h\right\}. \tag{2.1}$$

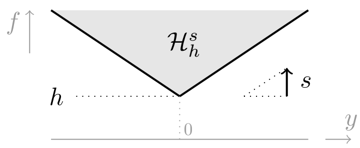

Figure 1. Illustration of $\mathcal{H}_h^s$.

We can embed $(0, \infty) \times \mathbb{R}^d$ continuously and openly into a locally compact Polish space $\mathfrak{E}$ with certain properties, mentioned in the lemma below. For a locally compact metric space E, we write $\mathcal{M}_{\mathrm{p}}(E)$ for the set of point measures on E, i.e., $\mathbb{N}_0 \cup \{\infty\}$-valued Radon measures, or equivalently, due to the fact that the support of each such measure is countable and locally finite, the set of Radon measures that can be written as $\sum_{n\in\mathbb{N}} \delta_{x_n}$ for a sequence $(x_n)_{n\in\mathbb{N}}$ in E. We equip $\mathcal{M}_{\mathrm{p}}(E)$ with the vague topology, i.e., $\mathcal{P}_n \to \mathcal{P}$ in $\mathcal{M}_{\mathrm{p}}(E)$ if and only if $\int \varphi \,\mathrm{d}\mathcal{P}_n \to \int \varphi \,\mathrm{d}\mathcal{P}$ for each continuous compactly supported $\varphi\colon E \to \mathbb{R}$. When $E = \mathfrak{E}$ we will simply write $\mathcal{M}_{\mathrm{p}} = \mathcal{M}_{\mathrm{p}}(\mathfrak{E})$. We denote by $\mathcal{M}_{\mathrm{p}}^\circ$ the set of point measures in $\mathcal{M}_{\mathrm{p}}$ that are supported in $(0, \infty) \times \mathbb{R}^d$ and equip it with the topology from $\mathcal{M}_{\mathrm{p}}$.

Lemma 2.1. *There exists a locally compact Polish space $\mathfrak{E}$, with $(0, \infty) \times \mathbb{R}^d \subset \mathfrak{E}$, such that*

- (i) *for every $h, s > 0$, the open set $\mathcal{H}_h^s$ is relatively compact in $\mathfrak{E}$, and for every compact subset K in $\mathfrak{E}$ there exist $h, s > 0$ such that $K \cap (0, \infty) \times \mathbb{R}^d \subset \mathcal{H}_h^s$,*
- (ii) *the map $\iota : (0, \infty) \times \mathbb{R}^d \to \mathfrak{E}$ given by $\iota((f, y)) = (f, y)$, $(f, y) \in (0, \infty) \times \mathbb{R}^d$ is open and continuous. In other words, $(0, \infty) \times \mathbb{R}^d$ is continuously and openly embedded in $\mathfrak{E}$.*

Moreover,

(a) $\mathcal{M}_{\mathrm{p}}^{\circ}$ *can be viewed as a subspace of* $\mathcal{M}_{\mathrm{p}}((0,\infty)\times\mathbb{R}^d)$*, in the sense that for* $\mathcal{P}\in\mathcal{M}_{\mathrm{p}}^{\circ}$*,* $\mathcal{P}\circ\iota$ *defines a point measure on* $(0,\infty)\times\mathbb{R}^d$*.*

(b) *Let* $\mathcal{P}\in\mathcal{M}_{\mathrm{p}}((0,\infty)\times\mathbb{R}^d)$*. Define* $\overline{\mathcal{P}}$ *on* $\mathfrak{E}$ *by* $\overline{\mathcal{P}}(A)=\mathcal{P}(\iota^{-1}(A))$ *for Borel sets* $A\subset\mathfrak{E}$*. Then* $\overline{\mathcal{P}}$ *is an element of* $\mathcal{M}_{\mathrm{p}}^{\circ}$ *if and only if* $\mathcal{P}(\mathcal{H}_h^s)<\infty$ *for all* $s,h>0$*.*

Proof. The proof is given in Appendix A, below Lemma A.1. □

Remark 2.2. Observe that by the Portmanteau theorem (Kle08, Theorem 13.16), $\mathcal{P}_n\to\mathcal{P}$ in $\mathcal{M}_{\mathrm{p}}(\mathfrak{E})$ implies that $\mathcal{P}_n(A)\to\mathcal{P}(A)$ for all measurable relatively compact $A\subset\mathfrak{E}$. And hence by Lemma 2.1 (b), in particular for all measurable A that are a subset of $\mathcal{H}_h^s$ for some $h,s>0$.

Lemma 2.3. *Let* $t>0$*.* $\overline{\Pi}$ *and* $\overline{\Pi}_t$ *are almost surely in* $\mathcal{M}_{\mathrm{p}}^{\circ}$ *(with* $\overline{\mathcal{P}}$ *as in Lemma 2.1 (b)).*

Proof. Let $h,s>0$. We show that $\mathbf{E}(\Pi(\mathcal{H}_h^s))<\infty$ and $\mathbf{E}(\Pi_t(\mathcal{H}_h^s))<\infty$, so that, e.g., $\mathbf{P}(\Pi(\mathcal{H}_h^s)<\infty)=1$, and therefore $\mathbf{P}(\bigcap_{s,h\in(0,\infty)\cap\mathbb{Q}}\{\Pi(\mathcal{H}_h^s)<\infty\})=1$. Because $\mathcal{H}_h^s\subset\mathcal{H}_j^t$ for $t\le s$ and $j\le h$, this implies $\mathbf{P}(\bigcap_{s,h\in(0,\infty)}\{\Pi(\mathcal{H}_h^s)<\infty\})=1$ and thus, by Lemma 2.1 (b) that $\overline{\Pi}\in\mathcal{M}_{\mathrm{p}}$.

We have

$$\Pi_t(\mathcal{H}_h^s)=\sum_{z\in\mathbb{Z}^d}\delta_{\left(\frac{\xi(z)}{r_t^{d/\alpha}},\frac{z}{r_t}\right)}(\mathcal{H}_h^s)=\sum_{z\in\mathbb{Z}^d}\mathbb{1}\Big\{\frac{\xi(z)}{r_t^{d/\alpha}}>s\Big|\frac{z}{r_t}\Big|+h\Big\}.$$

We calculate

$$\mathbf{P}\Big(\frac{\xi(z)}{r_t^{d/\alpha}}>s\Big|\frac{z}{r_t}\Big|+h\Big)=\Big(r_t^{d/\alpha}\big(s\Big|\frac{z}{r_t}\Big|+h\big)\Big)^{-\alpha}=r_t^{-d}\big(s\Big|\frac{z}{r_t}\Big|+h\big)^{-\alpha}.$$

Therefore, because $\alpha>d$,

$$\mathbf{E}\Big(\Pi_t(\mathcal{H}_h^s)\Big)\le\sum_{z\in\mathbb{Z}^d}r_t^{-d}\big(s\Big|\frac{z}{r_t}\Big|+h\big)^{-\alpha}<\infty.$$

Note that $\Pi(\mathcal{H}_h^s)$ is a Poisson distributed random variable with parameter

$$\int_{\mathbb{R}^d}\int_0^\infty\mathbb{1}_{\mathcal{H}_h^s}(f,y)\frac{\alpha}{f^{\alpha+1}}\,\mathrm{d}f\,\mathrm{d}y=\int_{\mathbb{R}^d}\frac{1}{(s|y|+h)^\alpha}\,\mathrm{d}y,$$

which is finite for $\alpha>d$, so that $\mathbb{E}(\Pi(\mathcal{H}_h^s))<\infty$. □

From here on, we will make abuse of notation and write Π also for $\overline{\Pi}$ and Π_t for $\overline{\Pi}_t$.

In the following lemma we state the convergence of Π_t towards Π, as mentioned between (1.9) and (1.10):

Lemma 2.4 ($\Pi_t\Longrightarrow\Pi$)**.** *Let* $\alpha\in(d,\infty)$*. Let* $t_1,t_2,\dots\in(0,\infty)$ *and* $t_n\to\infty$*. We may view* Π_{t_n} *and* Π *as elements of* $\mathcal{M}_{\mathrm{p}}^{\circ}$ *for all* n*. Then* $\Pi_{t_n}\to\Pi$ *in* $\mathcal{M}_{\mathrm{p}}^{\circ}$ *as* $n\to\infty$*.*

Proof. That we may view Π_{t_n} and Π as elements of $\mathcal{M}_{\mathrm{p}}^{\circ}$ follows by Lemma 2.3.

The convergence follows by (BKdS18, Lemma 7.4) (the fact that we have $(0,\infty)\times\mathbb{R}^d$ instead of $\mathbb{R}\times\mathbb{R}^d$ does not change the validity of the lemma, as the proof builds on (Res87, Proposition 3.21) can be carried out in our situation in the same way). For this we have to check the two conditions, namely (7.17) and (7.18) of that lemma (we take the $\widehat{N}_t$ in that lemma to be equal to zero, furthermore let us mention that in (7.17) there should be "$\frac{t^d}{(2\widehat{N}_t+1)^d}$" instead of "$\frac{t^d}{(2\widehat{N}_t)^d}$"). The first condition, (7.17), follows by

$$\lim_{r\to\infty}r^d\mathbf{P}\Big(\frac{\xi(0)}{r^{d/\alpha}}>s\Big)=\lim_{r\to\infty}r^d(r^{d/\alpha}s)^{-\alpha}=s^{-\alpha}.$$

The second condition, (7.18), follows by the fact that for all $s, h > 0$

$$\sum_{x\in\mathbb{Z}^d:|x|>rn} \mathbf{P}\Big(\frac{\xi(0)}{r^{d/\alpha}} > s\frac{|x|}{r} + h\Big) \le \sum_{x\in\mathbb{Z}^d:|x|>rn} \mathbf{P}\Big(\xi(0) > s|x|r^{\frac{d}{\alpha}-1}\Big)$$
$$\le s^{-\alpha} \sum_{x\in\mathbb{Z}^d:|x|>rn} |x|^{-\alpha} r^{\alpha-d} \le s^{-\alpha}\int_{\frac{rn}{2}}^{\infty} u^{-\alpha} r^{\alpha-d} u^{d-1}\,\mathrm{d}u$$
$$= s^{-\alpha} r^{\alpha-d} \frac{u^{d-\alpha}}{d-\alpha}\Big|_{\frac{rn}{2}}^{\infty} = s^{-\alpha} r^{\alpha-d}\frac{\left(\frac{rn}{2}\right)^{d-\alpha}}{\alpha-d} = s^{-\alpha}\frac{2^{\alpha-d}}{\alpha-d} n^{d-\alpha} \xrightarrow{n\to\infty} 0.$$

So that indeed,

$$\lim_{n\to\infty}\limsup_{r\to\infty} \sum_{x\in\mathbb{Z}^d:|x|>rn} \mathbf{P}\Big(\frac{\xi(0)}{r^{d/\alpha}} > s\frac{|x|}{r} + h\Big) = 0.$$

□

For $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}^{\circ}$ and $R > 0$, define

$$M_R(\mathcal{P}) := \sup\{f\colon (f, y) \in \mathcal{P} \text{ and } y \in Q_R\}. \tag{2.2}$$

Lemma 2.5. *Let $\mathcal{P}, \mathcal{P}_1, \mathcal{P}_2, \ldots$ be in $\mathcal{M}_{\mathrm{p}}^{\circ}$ such that $\mathcal{P}_n \to \mathcal{P}$ in $\mathcal{M}_{\mathrm{p}}$. Then*

$$\sup_{n\in\mathbb{N}} M_R(\mathcal{P}_n) < \infty \quad \textit{for all } R > 0 \qquad \textit{and} \qquad \lim_{R\to\infty}\sup_{n\in\mathbb{N}} \frac{M_R(\mathcal{P}_n)}{R} = 0. \tag{2.3}$$

In particular, $\lim_{R\to\infty} \frac{M_R(\mathcal{P})}{R} = 0$.

Proof. Fix $\varepsilon \in (0, 1)$. First observe that for any $\mathcal{Q} \in \mathcal{M}_{\mathrm{p}}$ and any $h > 0$, $\mathcal{Q}(\mathcal{H}_h^{\varepsilon}) < \infty$ since $\mathcal{H}_h^{\varepsilon}$ is relatively compact (in $\mathfrak{E}$), and thus $V(\mathcal{Q}) := \sup\{f\colon (f, y) \in \operatorname{supp}\mathcal{Q} \cap \mathcal{H}_h^{\varepsilon}\} < \infty$. Fix $h \in (0, \varepsilon)$ such that $\mathcal{P}(\partial\mathcal{H}_h^{\varepsilon}) = 0$. By (Res87, Proposition 3.13), there exists an $n_0 \in \mathbb{N}$ such that $V(\mathcal{P}_n) \le V(\mathcal{P}) + 1$ for all $n \ge n_0$, implying $M = \sup_{n\in\mathbb{N}} V(\mathcal{P}_n) < \infty$. For $R > 0$, note that $(0, \infty) \times Q_R = A \cup B$ where $A \subset (0, \varepsilon(R + 1)] \times \mathbb{R}^d$ and $B \subset \mathcal{H}_h^{\varepsilon}$ (see Figure 2), so that

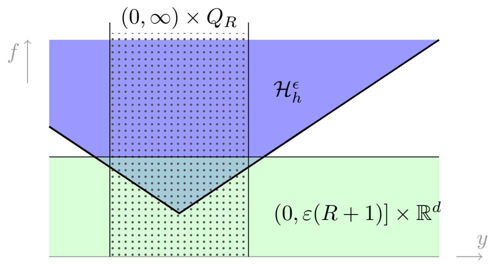

Figure 2. Illustration $(0, \infty) \times Q_R \subset \mathcal{H}_h^{\varepsilon} \cup (0, \varepsilon(R + 1)] \times \mathbb{R}^d$.

$$\sup_{n\in\mathbb{N}} M_R(\mathcal{P}_n) \le \max\{\varepsilon(R + 1), M\} < \infty,$$

implying the first statement in (2.3). For the second statement, divide the above inequality by R, take the $\limsup$ as $R \to \infty$ and then the limit as $\varepsilon \to 0$. □

Recall that $\mathcal{W}$ denotes the set of subprobability measures on $(0,\infty)\times\mathbb{R}^d$ with total mass ≤ 1. For $R>0$ and $\mathcal{P}\in\mathcal{M}_{\rm p}((0,\infty)\times\mathbb{R}^d)$ define the sets

$$\mathfrak{F}(\mathcal{P}) := \{\mu\in\mathcal{W}\colon\ \mu\ll\mathcal{P} \text{ and } \operatorname{supp}(\mu) \text{ is finite}\}, \tag{2.4}$$

$$\mathfrak{F}_1(\mathcal{P}) := \{\mu\in\mathfrak{F}(\mathcal{P}) : \mu \text{ is a probability measure}\}. \tag{2.5}$$

In the following lemma we show that if a point measure $\mathcal{P}$ has sufficiently many points, then one may restrict to take the supremum over elements in $\mathfrak{F}_1(\mathcal{P})$ in the variational formula for $\Xi(\mathcal{P})$ (1.18). Then, we introduce the notion of a good point measure, under which we will prove a conditional version of Theorem 1.4 (b).

Lemma 2.6. *Let $\mathcal{P}\in\mathcal{M}_{\rm p}^{\circ}$. Suppose that*

$$\forall\delta>0\ \exists m\in\mathbb{N}\ \exists \text{ distinct } (f_1,y_1),\dots(f_m,y_m)\in\operatorname{supp}\mathcal{P},\quad \sum_{i=1}^m f_i\ge 2\theta,\ D_0(y_1,\dots,y_m)<\delta. \tag{2.6}$$

Then

$$\Xi(\mathcal{P}) = \sup_{\mu\in\mathcal{W}}\Psi_{\mathcal{P}}(\mu) = \sup_{\mu\in\mathfrak{F}_1(\mathcal{P})}\Psi_{\mathcal{P}}(\mu). \tag{2.7}$$

Definition 2.7 (Good point measure)**.** We say that a point measure $\mathcal{P}\in\mathcal{M}_{\rm p}^{\circ}$ is good if $\Psi_{\mathcal{P}}$ possesses at most one maximizer in $\mathfrak{F}(\mathcal{P})$ in the sense that there exists at most one $\nu\in\mathfrak{F}(\mathcal{P})$ such that $\sup_{\mu\in\mathfrak{F}(\mathcal{P})}\Psi_{\mathcal{P}}(\mu)=\Psi_{\mathcal{P}}(\nu)$, and if it satisfies at least one of the two following conditions:

(i) There exists a $\beta>2$ such that for all $R,C>0$ there exists a $\varepsilon_{R,C}>0$ such that for $\varepsilon\le\varepsilon_{R,C}$ and for all $y\in Q_R$ there exists a $(\widetilde{f},\widetilde{y})\in(C\varepsilon,\infty)\times B(y,\varepsilon^\beta)\cap\operatorname{supp}\mathcal{P}$.

(ii) $\mathcal{P}((0,\infty)\times Q_R)<\infty$ for every $R>0$.

Now we can formulate a deterministic version of Theorem 1.4 (b) (and more) for good processes.

Theorem 2.8 (Analysis of $\Psi_{\mathcal{P}}$ for good $\mathcal{P}$)**.** *If $\mathcal{P},\mathcal{P}_1,\mathcal{P}_2,\dots\in\mathcal{M}_{\rm p}^{\circ}$ are good in the sense of Definition 2.7, then the following statements hold.*

(a) *Maximizer: There exists a unique $\mu^*\in\mathcal{W}$ such that*

$$\Psi_{\mathcal{P}}(\mu^*) = \sup_{\mu\in\mathcal{W}}\Psi_{\mathcal{P}}(\mu). \tag{2.8}$$

This maximizer μ^* has finite support, is a probability measure and satisfies $\mu^*\ll\mathcal{P}$, i.e., $\mu^*\in\mathfrak{F}_1(\mathcal{P})$.

(b) *Multisupport maximizer: Let $k\in\mathbb{N}$, $\varepsilon:=\frac{\theta}{4qk^4}$ and $L>2\theta+(q+1)\varepsilon$. With $B_\varepsilon=\{y\in\mathbb{R}^d : |y|\le\varepsilon\}$ the closed ℓ^1 ball in $\mathbb{R}^d$, define the regions of $(0,\infty)\times\mathbb{R}^d$ (see also Figure 3)*

$$\begin{aligned} G &= \Big[L, L+\frac{2\theta}{k}\Big]\times B_\varepsilon,\quad E^1=(\varepsilon,L)\times B_\varepsilon,\quad E^2=\Big(L+\frac{2\theta}{k},\infty\Big)\times B_\varepsilon,\\ E^3 &= \{(f,y)\in(0,\infty)\times\mathbb{R}^d\colon\ |y|>\varepsilon, f>\varepsilon\vee(|y|-3\theta\}. \end{aligned} \tag{2.9}$$

If $\mathcal{P}$ satisfies

$$\mathcal{P}(G)=k \quad\text{and}\quad \mathcal{P}(E^1)=\mathcal{P}(E^2)=\mathcal{P}(E^3)=0, \tag{2.10}$$

then $\#\operatorname{supp}\mu^*=k$.

(c) *Stability: For any open neighbourhood $O\subset\mathcal{W}$ of μ^*,*

$$\sup_{O^c}\Psi_{\mathcal{P}} < \sup_{\mathcal{W}}\Psi_{\mathcal{P}} = \Psi_{\mathcal{P}}(\mu^*). \tag{2.11}$$

(d) *Continuity of maximizer: If $\mathcal{P}_n\to\mathcal{P}$ in $\mathcal{M}_{\rm p}^{\circ}$, then the maximizers μ_n^* of $\Psi_{\mathcal{P}_n}$ converge towards μ^* as $n\to\infty$ in the vague topology.*

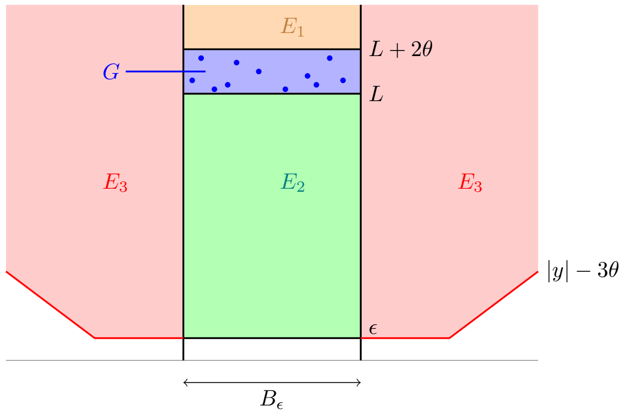

FIGURE 3. Illustration of the regions G, E^1, E^2 and E^3 as in (2.9).

The proofs of Lemma 2.6 and of Theorem 2.8 are postponed to Section 3. (a) and (b) are used only to prove Theorem 1.4, while (c) and (d) are necessary preparations for the proof of Theorem 1.1.

In order to be able to apply Theorem 2.8 to the point processes Π_t defined in (1.9) and its limiting PPP Π defined in (1.10), we use the following lemmas, which proofs are given in Section 4.

Lemma 2.9 (Goodness of Π and Π_t)**.** *Fix $\alpha \in (2d, \infty)$. Then, for any $t \in (0, \infty)$, with probability one, Π and Π_t are good.*

Lemma 2.10. *Let $\alpha \in (d, \infty)$. Let $k \in \mathbb{N}$ and G, E^1, E^2, E^3 be as in (2.9). Then*

$$\mathbf{P}\big(\,\Pi(G) = k, \Pi(E^1 \cup E^2 \cup E^3) = 0\big) > 0.$$

Proof. Since the regions in (2.9) are disjoint and each of them has finite and positive intensity measure, the random variables $\Pi(G)$, $\Pi(E^1)$, $\Pi(E^2)$, $\Pi(E^3)$ are independent and have non-trivial Poisson distributions, so that $\mathbf{P}(\Pi(G) = k, \Pi(E^1) = \Pi(E^2) = \Pi(E^3) = 0)$ has positive probability. □

It is clear that Theorem 1.4 (b) directly follows from Theorem 2.8, combined with Lemma 2.9 and Lemma 2.10.

For the proof of the lower bound in Section 6, we use the following lemma so that we can apply Lemma 2.6 to Π.

Lemma 2.11. *Let $\alpha \in (d, \infty)$. With probability one, Π satisfies (2.6).*

3. Analysis of the variational formula

Here we give the proof of Theorem 2.8; that is, we analyse the maximum of $\Psi_{\mathcal{P}}$ and its maximizer for an arbitrary point measure $\mathcal{P}$ that is good in the sense of Definition 2.7.

Let us first give a short outline of the proof. In Section 3.1, we introduce the crucial tool for handling variational problems, namely the Gamma-convergence, and derive Γ-continuity properties of $\mathcal{P} \mapsto \Phi_{\mathcal{P}}$ and $\mathcal{P} \mapsto \mathcal{D}_{\mathcal{P}}$ and consider the compactness of the objects appearing in the variational formula $\sup_{\mu \in \mathcal{W}} \Psi_{\mathcal{P}}(\mu)$ (the right-hand side of (2.8)): if $\mathcal{P}$ is good, then one can restrict the variational formula to measures in $\mathcal{W}$ that have a compact support with respect to $\mathbb{R}^d$. Then we give the proof of Lemma 2.6. In Section 3.2 we analyse the maximization of $\Phi_{\mathcal{P}}(\mu)$ over μ when the number of points of $\mathcal{P}$ is fixed; this involves only the maximization over the potential values. In Section 3.3 we show that any maximizer of $\Psi_{\mathcal{P}}$ is necessarily of finite support. In Section 3.4 finally we prove Theorem 2.8 (a), putting together the results derived in the preceding sections, namely the Γ-continuity of $\mathcal{P} \mapsto -\Psi_{\mathcal{P}}$, and the fact that we need to optimize $\Psi_{\mathcal{P}}$ only over compact subsets of $\mathcal{W}$. Recall that by our definition of "good", the uniqueness of the maximizer is guaranteed for good $\mathcal{P}$.

3.1 Some topological properties of the variational formula

In this section, we prove that the functional $\mathcal{P} \mapsto -\Psi_{\mathcal{P}}$ introduced in Definition 1.2 is Gamma-continuous in the vague topology, which is the crucial property under which we can find later arguments for the existence of maximizers and continuity properties of the maximizers as a function of $\mathcal{P}$. The main tool of the arguments is a characterization of the vague convergence of point measures in terms of one-by-one convergence of its points.

Let us introduce the crucial sense of convergence for variational formulas.

Definition 3.1 (Gamma convergence)**.** Let X be a metric space. Let $f, f_1, f_2, \dots : X \to [-\infty, \infty]$. We say that the sequence $(f_n)_{n\in\mathbb{N}}$ Gamma converges to f, written $f_n \xrightarrow{\Gamma, n\to\infty} f$, if

(i) for all $x \in X$ and all sequences $(x_n)_{n\in\mathbb{N}}$ in X with $x_n \to x$,

$$f(x) \le \liminf_{n\to\infty} f_n(x_n),$$

(ii) for all $x \in X$ there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $x_n \to x$ and

$$f(x) \ge \limsup_{n\to\infty} f_n(x_n).$$

Remark 3.2. Observe that $f \xrightarrow{\Gamma, n\to\infty} f$ if and only if f is lower semi-continuous. $\diamond$

We use the following statements about Γ-convergence, which are sometimes referred to as the Fundamental theorem(s) of Gamma convergence:

Theorem 3.3. *Let X be a metric space. Let $f, f_1, f_2, \dots : X \to [-\infty, \infty]$. Suppose $f_n \xrightarrow{\Gamma, n\to\infty} f$.*

(a) (Bra02, Proposition 1.18) *For each compact $K \subset X$*

$$\inf_{x\in K} f(x) \le \liminf_{n\to\infty} \inf_{x\in K} f_n(x).$$

(b) (Bra02, Theorem 1.21) *Suppose there exists a compact set $K \subset X$ such that $\inf_{x\in X} f_n(x) = \inf_{x\in K} f_n(x)$ for all $n \in \mathbb{N}$. Suppose that $x_1, x_2, \dots \in X$ are such that $f_n(x_n) = \inf_{x\in X} f_n(x)$ for all $n \in \mathbb{N}$. Then there exists a subsequence of $(x_n)_{n\in\mathbb{N}}$ that converges to an $y \in X$ for which $\inf_{x\in X} f(x) = f(y)$.*

The main result of Section 3.1 is the following proposition. Part (a) will allow us to restrict the search for a maximizer μ^* of $\Psi_{\mathcal{P}}$ to those μ whose $\mathbb{R}^d$-support is within some box in $\mathbb{R}^d$. Recall that $\mathcal{W}$ is the set of subprobability measures on $(0,\infty) \times \mathbb{R}^d$ with total mass ≤ 1, and $Q_R = [-R, R]^d$. Furthermore, we introduce

$$\mathcal{W}_R = \{\mu \in \mathcal{W} \colon \operatorname{supp} \mu \subset (0,\infty) \times Q_R\}, \qquad R > 0. \tag{3.1}$$

Proposition 3.4. *Let $\mathcal{P}, \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \dots$ be in $\mathcal{M}_{\mathrm{p}}^\circ$ such that $\mathcal{P}_n \to \mathcal{P}$ in $\mathcal{M}_{\mathrm{p}}$. Then*

(a) *Compactness*

$$\lim_{R\to\infty} \sup_{\mu\in\mathcal{W}\setminus\mathcal{W}_R} \sup_{n\in\mathbb{N}} \Psi_{\mathcal{P}_n}(\mu) = -\infty.$$

(b) *Gamma convergence of $-\Psi$*

$$-\Psi_{\mathcal{P}_n} \xrightarrow{\Gamma, n\to\infty} -\Psi_{\mathcal{P}}.$$

The proof of this proposition is at the end of this section. We prepare for the proof by citing a well-known result from point-process theory about a characterization of vague convergence by point-wise convergence.

For $\mathcal{P} \in \mathcal{M}_{\rm p}$ and $L > 0$, recalling that $Q_L = [-L, L]^d$, we denote by $\mathcal{P}^{(L)}$ the point measure $\mathbb{1}_{[L^{-1},\infty)\times Q_L}\mathcal{P}$, which means $\frac{\mathrm{d}\mathcal{P}^{(L)}}{\mathrm{d}\mathcal{P}} = \mathbb{1}_{[L^{-1},\infty)\times Q_L}$, i.e.,

$$\mathcal{P}^{(L)}(A) = \mathcal{P}\Big(A \cap [L^{-1}, \infty) \times Q_L\Big) \tag{3.2}$$

for any Borel measurable $A \subset \mathfrak{E}$. For $\mu \in \mathcal{W}$ we also write $\mu^{(L)} = \mathbb{1}_{[L^{-1},\infty)\times Q_L}\mu$.

Observe that as $[L^{-1}, \infty) \times Q_L \subset \mathcal{H}_h^s$ for some $h, s > 0$ (e.g. $s = \frac{1}{4}$ and $h = \frac{L}{2}$), and $\mathcal{H}_h^s$ is relatively compact in $\mathfrak{E}$, $\mathcal{P}([L^{-1}, \infty) \times Q_L) \in \mathbb{N}_0$ for all $\mathcal{P} \in \mathcal{M}_{\rm p}$.

Lemma 3.5. *Let $\mathcal{P}, \mathcal{P}_1, \mathcal{P}_2, \dots \in \mathcal{M}_{\rm p}$ and $L > 0$ be such that $\mathcal{P}_n \to \mathcal{P}$ in $\mathcal{M}_{\rm p}$ and $\mathcal{P}(\partial([L^{-1}, \infty) \times Q_L)) = 0$.*

(a) *Put $k = \mathcal{P}([L^{-1}, \infty) \times Q_L) \in \mathbb{N}_0$. Then there exist $(f_i, y_i), (f_i^n, y_i^n) \in [L^{-1}, \infty) \times Q_L$, for $n \in \mathbb{N}$ and $i \in \{1, \dots, k\}$ such that, for all large enough $n \in \mathbb{N}$ (with empty sums interpreted as zero),*

$$\mathcal{P}_n^{(L)} = \sum_{i=1}^{k} \delta_{(f_i^n, y_i^n)}, \quad \mathcal{P}^{(L)} = \sum_{i=1}^{k} \delta_{(f_i, y_i)}, \quad (f_i^n, y_i^n) \xrightarrow{n\to\infty} (f_i, y_i), \quad i \in \{1, \dots, k\}.$$

(b) *Suppose $\mu, \mu_1, \mu_2, \dots$ are in $\mathcal{W}$ such that $\mu_n \to \mu$ in $\mathcal{W}$ and $\mu_n \ll \mathcal{P}_n$ for all $n \in \mathbb{N}$. Then $\mu \ll \mathcal{P}$ and, with k, $(f_i, y_i), (f_i^n, y_i^n)$ as above, there exist $(w_1^n, \dots, w_k^n), (w_1, \dots, w_k) \in [0, 1]^k$ such that, for all large enough n,*

$$\mu_n^{(L)} = \sum_{i=1}^{k} w_i^n \delta_{(f_i^n, y_i^n)}, \quad \mu^{(L)} = \sum_{i=1}^{k} w_i \delta_{(f_i, y_i)}, \quad w_i^n \xrightarrow{n\to\infty} w_i, \qquad i \in \{1, \dots, k\}.$$

Proof. For the first statement, note that $[L^{-1}, \infty) \times Q_L$ is a relatively compact subset of $\mathfrak{E}$ and apply (Res87, Proposition 3.13). (See also Theorem C.1.) The second statement is a straightforward consequence of the first. That $\mu \ll \mathcal{P}$ follows by the fact that from the convergences one obtains $\mu^{(L)} \ll \mathcal{P}^{(L)}$ for all $L > 0$: Let f_L be a density function which equals zero outside $[L, \infty) \times Q_L$. Then $\mu = \lim_{L\to\infty} \mu^{(L)} = \lim_{L\to\infty} f_L \mathcal{P}^{(L)} = \lim_{L\to\infty} f_L \mathcal{P}$. Hence $f = \lim_{L\to\infty} f_L = \sup_{L\in\mathbb{N}} f_L$ is the density for μ with respect to $\mathcal{P}$. ☐

Here is the main step in the proof of Proposition 3.4.

Lemma 3.6. *Let $\mathcal{P}, \mathcal{P}_1, \mathcal{P}_2, \dots \in \mathcal{M}_{\rm p}^{\circ}$ be such that $\mathcal{P}_n \to \mathcal{P}$ in $\mathcal{M}_{\rm p}$, and let $\mu, \mu_1, \mu_2, \dots \in \mathcal{W}$ be such that $\mu_n \to \mu$ in $\mathcal{W}$. Then*

(a) *$\mathcal{D}_{\mathcal{P}}(\mu) \le \liminf_{n\to\infty} \mathcal{D}_{\mathcal{P}_n}(\mu_n)$.*

(b) *If $\mu_n \ll \mathcal{P}_n$ and there exists a $R > 0$ such that $\mu_n \in \mathcal{W}_R$ for all $n \in \mathbb{N}$, then*

$$\Phi_{\mathcal{P}}(\mu) \ge \limsup_{n\to\infty} \Phi_{\mathcal{P}_n}(\mu_n).$$

Proof. (a) If $\mu \not\ll \mathcal{P}$, then by Lemma 3.5 there exists an $N \in \mathbb{N}$ such that $\mu_n \not\ll \mathcal{P}_n$ for all $n \ge N$, and the conclusion trivially holds. Therefore, we may assume $\mu \ll \mathcal{P}$ and $\mu_n \ll \mathcal{P}_n$ for all $n \in \mathbb{N}$. Moreover, we may assume that $\mu \ne 0$. Let $L > 0$ be such that $\mathcal{P}$ has zero measure on the boundary of $[L^{-1}, \infty) \times Q_L$ and $\mathcal{D}_{\mathcal{P}}(\mu^{(L)}) > 0$, which implies $\mu^{(L)} \ne 0$. Let k, f_i^n, y_i^n, w_i^n, f_i, y_i, w_i be as in Lemma 3.5, and note that $k \ge 1$. Let $i_1, \dots, i_m \in \{1, \dots, k\}$, $m \in \mathbb{N}$, be the distinct indices such that $w_{i_j} > 0$, $j \in \{1, \dots, m\}$, and $w_\ell = 0$ otherwise. We may assume that, for all $i \le k$ and all n large enough, $w_i > 0$ implies $w_i^n > 0$. Then

$$\mathcal{D}_{\mathcal{P}_n}(\mu_n) \ge \mathcal{D}_{\mathcal{P}_n}(\mu_n^{(L)}) \ge D_0(y_{i_1}^n, \dots, y_{i_m}^n) \xrightarrow{n\to\infty} D_0(y_{i_1}, \dots, y_{i_m}) = \mathcal{D}_{\mathcal{P}}(\mu^{(L)}).$$

Therefore, for any $L > 0$, $\liminf_{n\to\infty} \mathcal{D}_{\mathcal{P}_n}(\mu_n) \ge \mathcal{D}_{\mathcal{P}}(\mu^{(L)})$. Since $\mathcal{D}_{\mathcal{P}}(\mu) = \sup_{L>0} \mathcal{D}_{\mathcal{P}}(\mu^{(L)})$, the claim follows.

(b) Let $R > 0$, $\mu_n \in \mathcal{W}_R$, $\mu_n \ll \mathcal{P}_n$ for all $n \in \mathbb{N}$ and $\mu_n \to \mu$ in $\mathcal{W}$. Note that this implies $\mu \in \mathcal{W}_R$ as well and, by Lemma 3.5 (b), $\mu \ll \mathcal{P}$. Let $\varepsilon > 0$. Let us first show that for large $L > 0$

$$\Phi_{\mathcal{P}_n}(\mu_n) < \Phi_{\mathcal{P}_n}(\mu_n^{(L)}) + \varepsilon \;\text{ for all } n \in \mathbb{N} \qquad \text{and} \qquad |\Phi_{\mathcal{P}}(\mu) - \Phi_{\mathcal{P}}(\mu^{(L)})| < \varepsilon. \tag{3.3}$$

Indeed, take $L > R$ such that $L^{-1} < \varepsilon$, $\mu((0, L^{-1}) \times Q_L) < \varepsilon/\theta$ and $\mathcal{P}(\partial([L^{-1}, \infty) \times Q_L)) = 0$. Then

$$\Phi_{\mathcal{P}_n}(\mu_n) - \Phi_{\mathcal{P}_n}(\mu_n^{(L)}) = \int_{(0,L^{-1})\times Q_L} \left[f - \theta \frac{\mathrm{d}\mu_n}{\mathrm{d}\mathcal{P}_n}(f, y) \right] \mathrm{d}\mu_n(f, y) \le L^{-1} < \varepsilon,$$

and the same inequality is valid with $\mathcal{P}_n$, μ_n replaced by $\mathcal{P}$, μ, for which also

$$\Phi_{\mathcal{P}}(\mu) - \Phi_{\mathcal{P}}(\mu^{(L)}) = \int_{(0,L^{-1})\times Q_L} \left[f - \theta \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{P}}(f, y) \right] \mathrm{d}\mu(f, y) \ge -\theta\mu((0, L^{-1}) \times Q_L) > -\varepsilon,$$

where we used that $\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{P}} \le 1$. This concludes (3.3). Now it is enough to show that $\Phi_{\mathcal{P}_n}(\mu_n^{(L)}) \to \Phi_{\mathcal{P}}(\mu^{(L)})$, but this follows by Lemma 3.5: with $k, f_i, f_i^n, y_i, y_i^n, w_i, w_i^n$ as therein and n large enough,

$$\Phi_{\mathcal{P}_n}(\mu_n^{(L)}) = \sum_{i=1}^k \left(w_i^n f_i^n - \theta (w_i^n)^2 \right) \to \sum_{i=1}^k \left(w_i f_i - \theta w_i^2 \right) = \Phi_{\mathcal{P}}(\mu^{(L)}).$$

□

As a by-product of the proof, we obtained:

Lemma 3.7. *Let $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}^{\circ}$ and $\mu \in \mathcal{W}_R$ for some $R \in (0, \infty)$. Then*

$$\Phi_{\mathcal{P}}(\mu^{(L)}) \xrightarrow{L\to\infty} \Phi_{\mathcal{P}}(\mu), \qquad \mathcal{D}_{\mathcal{P}}(\mu^{(L)}) \xrightarrow{L\to\infty} \mathcal{D}_{\mathcal{P}}(\mu), \qquad \Psi_{\mathcal{P}}(\mu^{(L)}) \xrightarrow{L\to\infty} \Psi_{\mathcal{P}}(\mu).$$

Proof. This follows by definition of $\mathcal{D}_{\mathcal{P}}$ and by (3.3). □

Proof of Proposition 3.4. (a) It suffices to show that given any $A > 0$, there exists an $R_0 > 0$ such that,

$$\Psi_{\mathcal{P}_n}(\mu) \le -A \qquad \text{for all } R \ge R_0,\ \mu \in \mathcal{W} \setminus \mathcal{W}_R \text{ and } n \in \mathbb{N}.$$

Recall the definition of M_R from (2) and recall that $q = \frac{d}{\alpha - d} > 0$. By Lemma 2.5, there exists an $R_0 > 0$ such that

$$\max\{A, \sup_{n\in\mathbb{N}} M_R(\mathcal{P}_n)\} \le \tfrac{1}{2} q(R - 1) \qquad \text{for all } R \ge R_0. \tag{3.4}$$

Let $R \ge R_0$, $\mu \in \mathcal{W} \setminus \mathcal{W}_R$ and $n \in \mathbb{N}$. We decompose $(0, \infty) \times \mathbb{R}^d$ into

$$S_R^{(0)} := (0, \infty) \times Q_R, \qquad S_R^{(k)} := (0, \infty) \times \left[Q_{R+k} \setminus Q_{R+k-1} \right] \text{ for } k \in \mathbb{N}.$$

Observe that $\zeta := \mu((0, \infty) \times \mathbb{R}^d)$ is in $(0, 1]$. Write $\zeta = \sum_{k\in\mathbb{N}_0} \zeta_k$ with $\zeta_k := \mu(S_R^{(k)})$ for $k \in \mathbb{N}_0$. Note that $\zeta_k > 0$ implies $\mathcal{D}_{\mathcal{P}_n}(\mu) \ge R + k - 1$, so that $\zeta_k \mathcal{D}_{\mathcal{P}_n}(\mu) \ge \zeta_k (R + k - 1)$ for $k \in \mathbb{N}_0$. Since $\Psi_{\mathcal{P}_n}(\mu) = -\infty$ if $\mu \not\ll \mathcal{P}_n$, we may and do assume $\mu \ll \mathcal{P}_n$. Hence we have the lower bound $\zeta \mathcal{D}_{\mathcal{P}_n}(\mu) = \sum_{k\in\mathbb{N}_0} \zeta_k \mathcal{D}_{\mathcal{P}_n}(\mu) \ge \sum_{k\in\mathbb{N}_0} \zeta_k (R + k - 1)$. Furthermore, we have the upper bound

$$\Phi_{\mathcal{P}_n}(\mu) \le \sum_{k\in\mathbb{N}_0} \int_{S_R^{(k)}} f \,\mathrm{d}\mu(f, y) \le \sum_{k\in\mathbb{N}_0} \zeta_k M_{R+k}(\mathcal{P}_n).$$

Together with (3.4) and $D_{\mathcal{P}_n}(\mu) \ge R \ge R_0$, this gives

$$\begin{aligned}\Psi_{\mathcal{P}_n}(\mu) = \Phi_{\mathcal{P}_n}(\mu) - q\mathcal{D}_{\mathcal{P}_n}(\mu) &\le \sum_{k\in\mathbb{N}_0} \zeta_k \left[\tfrac{1}{2} q(R + k - 1) - q(R + k - 1) \right] - q(1 - \zeta) R \\ &\le \sum_{k\in\mathbb{N}_0} \zeta_k (-A) - (1 - \zeta) A = -A.\end{aligned}$$

(b) Let us first show (i) of Definition 3.1. Pick $\mu, \mu_1, \mu_2, \dots \in \mathcal{W}$ such that $\mu_n \to \mu$. We have to show that $-\Psi_{\mathcal{P}}(\mu) \le \liminf_{n\to\infty} -\Psi_{\mathcal{P}_n}(\mu_n)$, i.e., $\Psi_{\mathcal{P}}(\mu) \ge \limsup_{n\to\infty} \Psi_{\mathcal{P}_n}(\mu_n)$. By (a) we may assume that there exists an $R > 0$ such that $\mu, \mu_n \in \mathcal{W}_R$ for all $n \in \mathbb{N}$, because if such R does not exist, then $\limsup_{n\to\infty} \Psi_{\mathcal{P}_n}(\mu_n) = -\infty$. Passing to subsequences if necessary, we may also assume that $\mu_n \ll \mathcal{P}_n$ for all $n \in \mathbb{N}$ and, by Lemma 3.5 (b), $\mu \ll \mathcal{P}$. In this case, the desired statement is a direct consequence of Lemma 3.6.

To verify (ii) of Definition 3.1, let $\mu \in \mathcal{W}$. Assume first that $\Psi_{\mathcal{P}}(\mu) = -\infty$. Since $\mathcal{P}_n$ is countable for each n, there exists a $(f, y) \in (0, \infty) \times \mathbb{R}^d \setminus \bigcup_{n \in \mathbb{N}} \mathcal{P}_n$. Setting $\mu_n = (1 - \frac{1}{n})\mu + \frac{1}{n}\delta_{(f,y)}$, it is clear that $\mu_n \to \mu$ and $\Psi_{\mathcal{P}_n}(\mu_n) = -\infty$ for every n.

Assume now that $\Psi_{\mathcal{P}}(\mu) > -\infty$, which implies $\mu \ll \mathcal{P}$. As we will soon see, it suffices to show the following: for every $L > 0$, there exists a sequence μ_n with $\mu_n = \mu_n^{(L)}$ such that $\mathcal{D}_{\mathcal{P}_n}(\mu_n) \to \mathcal{D}_{\mathcal{P}}(\mu^{(L)})$ and $\Phi_{\mathcal{P}_n}(\mu_n) \to \Phi_{\mathcal{P}}(\mu^{(L)})$. To prove the latter, fix $L > 0$ and take k, $(f_i, y_i), (f_i^n, y_i^n)$ as in Lemma 3.5. Define $w_i = \mu(f_i, y_i)$ for $i \in \{1, \dots, k\}$ and $\mu_n := \sum_{i=1}^k w_i \delta_{(f_i^n, y_i^n)}$. Then

$$\mathcal{D}_{\mathcal{P}_n}(\mu_n) = D_0(\{y_j^n : w_j > 0\}) \to D_0(\{y_j : w_j > 0\}) = \mathcal{D}_{\mathcal{P}}(\mu^{(L)}),$$

$\mu_n \to \mu^{(L)}$ and $\Phi_{\mathcal{P}_n}(\mu_n) \to \Phi_{\mathcal{P}}(\mu^{(L)})$ as well, as shown in the last line of the proof of Lemma 3.6.

Now, for each $m \in \mathbb{N}$, we can find a sequence $(\nu_{m,n})_{n \in \mathbb{N}}$ in $\mathcal{W}$ with $\nu_{m,n}^{(m)} = \nu_{m,n}$ such that $\mathcal{D}_{\mathcal{P}_n}(\nu_{m,n}) \to \mathcal{D}_{\mathcal{P}}(\mu^{(m)})$ and $\Phi_{\mathcal{P}_n}(\nu_{m,n}) \to \Phi_{\mathcal{P}}(\mu^{(m)})$ as $n \to \infty$. Let $(N_m)_{m \in \mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$ such that

$$\left| \mathcal{D}_{\mathcal{P}_n}(\nu_{m,n}) - \mathcal{D}_{\mathcal{P}}(\mu^{(m)}) \right| \vee \left| \Phi_{\mathcal{P}_n}(\nu_{m,n}) - \Phi_{\mathcal{P}}(\mu^{(m)}) \right| < \frac{1}{m} \qquad \text{for all } n \geq N_m.$$

Define $m_n := \max\{m \in \mathbb{N} : N_m \leq n\}$ and $\mu_n := \nu_{m_n,n}$. Note that $m_n \to \infty$, $\mu_n \to \mu$ and $n \geq N_{m_n}$, so that by Lemma 3.7,

$$\left| \Psi_{\mathcal{P}_n}(\mu_n) - \Psi_{\mathcal{P}}(\mu) \right| \leq \left| \Psi_{\mathcal{P}_n}(\mu_n) - \Psi_{\mathcal{P}}(\mu^{(m_n)}) \right| + \left| \Psi_{\mathcal{P}}(\mu^{(m_n)}) - \Psi_{\mathcal{P}}(\mu) \right| \xrightarrow{n \to \infty} 0.$$

□

With the convergence of Lemma 3.7 and the compactness in Proposition 3.4 (a), we prove Lemma 2.6:

Proof of Lemma 2.6. By Proposition 3.4 (a) it follows that there exists an $R > 0$ such that $\Xi(\mathcal{P})$ equals $\sup_{\nu \in \mathcal{W}_R} \Psi_{\mathcal{P}}(\nu)$. Then, by Lemma 3.7, it follows that $\Xi(\mathcal{P})$ equals $\sup_{\nu \in \mathfrak{F}(\mathcal{P})} \Psi_{\mathcal{P}}(\nu)$. Let $\nu \in \mathfrak{F}(\mathcal{P})$ and $\delta > 0$. We show $\sup_{\mu \in \mathfrak{F}_1(\mathcal{P})} \Psi_{\mathcal{P}}(\mu) \geq \Psi_{\mathcal{P}}(\nu) - 2\delta$. Let $(f_1, y_1), \dots, (f_m, y_m)$ be distinct elements of $\mathcal{P}$ such that $\sum_{i=1}^m f_i \geq 2\theta$ and $D_0(y_1, \dots, y_m) < \delta$. Let $k \in \mathbb{N}_0$ and $(f_{m+1}, y_{m+1}), \dots, (f_{m+k}, y_{m+k})$ be the distinct elements that form the support of ν (so possibly $k = 0$). By Proposition 3.8 there exist w_i for $i \in \{1, \dots, m+k\}$ with $\sum_{i=1}^{m+k} w_i = 0$ such that for $\mu = \sum_{i=1}^{m+k} w_i \delta_{(f_i, y_i)}$ one has

$$\Phi_{\mathcal{P}}(\mu) = \varphi_{k+m}(f_1, \dots, f_{m+k}) \geq \varphi_m(f_1, \dots, f_m) \geq \Phi_{\mathcal{P}}(\nu), \quad \mathcal{D}_{\mathcal{P}}(\mu) \leq D_0(y_1, \dots, y_{m+k}) \leq 2\delta + \mathcal{D}_{\mathcal{P}}(\nu),$$

and thus $\Psi_{\mathcal{P}}(\mu) \geq \Psi_{\mathcal{P}}(\nu) - 2\delta$. □

3.2 Maximization of $\Phi_{\mathcal{P}}$ with fixed number of points

In this section we derive, for a given point measure with finite support, $\mathcal{P} = \sum_{i=1}^k \delta_{(f_i, y_i)}$, explicit information about the maximization of $\Phi_{\mathcal{P}}(\mu)$ over μ. We need slightly adapted notation. Since we optimize here only $\Phi_{\mathcal{P}}(\mu)$ over μ, we can also drop the points $y_1, \dots, y_k$; see Definition 1.2. We obtain explicit information about the maximising vector $w = (w_1, \dots, w_k) = (\mu(f_i, y_i))_{i=1}^k$.

Fix $\theta \in (0, \infty)$ as always. Furthermore, we fix $k \in \mathbb{N}$, assume that $\mathcal{P} = \sum_{i=1}^k \delta_{(f_i, y_i)}$ and therefore may restrict our maximization problem to μ of the form $\mu = \sum_{i=1}^k w_i \delta_{(f_i, y_i)}$. Then a comparison to Definition 1.2 shows that

$$\sup_{\mu \in \mathcal{W}} \Phi_{\mathcal{P}}(\mu) = \varphi_k(f_1, \dots, f_k), \tag{3.5}$$

where $\varphi_k\colon (0,\infty)^k \to [0,\infty)$ is defined as

$$\varphi_k(f_1,\dots,f_k) = \sup_{\substack{w_1,\dots,w_k\geq 0\\ \sum_{i=1}^k w_i\leq 1}} \Big(\sum_{i=1}^k w_i f_i - \theta\sum_{i=1}^k w_i^2\Big). \tag{3.6}$$

We are going to analyze the function φ_k in this section.

Since $\varphi_k(f_1,\dots,f_k)$ does not depend on the order of the f_i, we may assume them to be ordered in a decreasing way. The following is the main result of this section; it identifies the optimal $w_1,\dots,w_k$ and thus the optimal μ, provides some of its properties and shows its uniqueness.

Proposition 3.8 (Analysis of φ_k)**.** *Fix $k\in\mathbb{N}$ and $f_1\geq f_2\geq\dots\geq f_k>0$.*

Case 1: $f_1+\dots+f_k\geq 2\theta$. Let $k_\star = K_\star(f_1,\dots,f_k)$, where

$$K_\star(f_1,\dots,f_k) := \inf\Big\{j\in\{1,\dots,k-1\}\colon\, jf_{j+1}\leq \sum_{i=1}^j f_i - 2\theta\Big\}\wedge k, \tag{3.7}$$

where we interpret $\inf\emptyset=\infty$. Then the unique maximizer in (3.6) is given by

$$w_i := \begin{cases} \frac{1}{2\theta}\big[f_i - \frac{1}{k_\star}(\sum_{j=1}^{k_\star} f_j - 2\theta)\big] & \text{if } i\leq k_\star,\\ 0 & \text{otherwise.}\end{cases} \tag{3.8}$$

Moreover, $w_i>0$ for $i\in\{1,\dots,k_\star\}$, $w_1+\dots+w_{k_*}=1$ and

$$\varphi_k(f_1,\dots,f_k) = \varphi_{k_\star}(f_1,\dots,f_{k_\star}) = \frac{1}{4\theta}\Big(\sum_{i=1}^{k_\star} f_i^2 - \frac{1}{k_\star}\Big(\sum_{i=1}^{k_\star} f_i - 2\theta\Big)^2\Big). \tag{3.9}$$

In particular,

$$\sum_{i=1}^{k_\star-1}\frac{f_i^2}{4\theta} < \varphi_k(f_1,\dots,f_k) \leq \sum_{i=1}^{k_\star}\frac{f_i^2}{4\theta}. \tag{3.10}$$

Case 2: $f_1+\dots+f_k<2\theta$. Then the unique maximizer is given by $w_i=f_i/2\theta$, $1\leq i\leq k$. Moreover, $w_1+\dots+w_k<1$ and

$$\varphi_k(f_1,\dots,f_k) = \sum_{i=1}^k \frac{f_i^2}{4\theta}. \tag{3.11}$$

The proof of Proposition 3.8 builds on the following lemma, and is given below the proof of Lemma 3.9.

Lemma 3.9. *Let $k\in\mathbb{N}$ and $f_1\geq f_2\geq\cdots f_k\geq 0$.*

(a) *The map*

$$(w_1,\dots,w_k)\mapsto \sum_{i=1}^k [w_i f_i - \theta w_i^2] \tag{3.12}$$

is maximized over $[0,\infty)^k$ precisely for $w_i=\frac{f_i}{2\theta}$.

(b) *Suppose $\sum_{i=1}^k \frac{f_i}{2\theta}\leq 1$. Then (3.12) is maximized over $(w_1,\dots,w_k)\in[0,\infty)^k$ under the constraint $\sum_{i=1}^k w_i\leq 1$ precisely by $w_i=\frac{f_i}{2\theta}$.*

(c) *Let $\gamma\in\mathbb{R}$. Then (3.12) (as a function on $\mathbb{R}^k$) is maximized over $(w_1,\dots,w_k)$ in $\mathbb{R}^k$ under the constraint $\sum_{i=1}^k w_i=\gamma$ by*

$$w_j = \frac{f_j}{2\theta} + \frac{1}{k}\Big(\gamma - \sum_{i=1}^k \frac{f_i}{2\theta}\Big), \qquad j\in\{1,\dots,k\}. \tag{3.13}$$

(d) *Let* $\gamma \in [0,\infty)$ *and suppose that* $kf_k + 2\theta\gamma - \sum_{i=1}^k f_i \geq 0$. *Then* (3.12) *is maximized over* $(w_1,\dots,w_k) \in [0,\infty)^k$ *under the constraint* $\sum_{i=1}^k w_i = \gamma$ *by* (3.13)*, and we have*

$$\sum_{i=1}^k w_i f_i - \theta w_i^2 = \frac{1}{4\theta}\sum_{i=1}^k f_i^2 - \frac{\theta}{k}\Big(\gamma - \sum_{i=1}^k \frac{f_i}{2\theta}\Big)^2. \tag{3.14}$$

(e) *Suppose* $\sum_{i=1}^k \frac{f_i}{2\theta} \geq 1$ *and* $kf_k + 2\theta - \sum_{i=1}^k f_i \geq 0$. *Then* (3.12) *is maximized over* $[0,1]^k$ *under the constraint* $\sum_{i=1}^k w_i \leq 1$ *by* (3.13) *with* $\gamma = 1$.

(f) *Suppose* $kf_k + 2\theta - \sum_{i=1}^k f_i \leq 0$ *or equivalently*

$$(k-1)f_k + 2\theta - \sum_{i=1}^{k-1} f_i \leq 0. \tag{3.15}$$

Then, if (3.12) *is maximized by* $(w_1,\dots,w_k) \in [0,1]^k$ *with* $\sum_{i=1}^k w_i \leq 1$*, then* $w_k = 0$.

Proof. (a) follows by the fact that $w_i \mapsto w_i f_i - \theta w_i^2$ is concave for all i, so that the maximum is attained where its derivative equals zero (or at the boundary, i.e., for $w_i = 0$, but this gives an outcome that is clearly less than for $w_i = \frac{f_i}{2\theta}$).

(b) follows immediately from (a).

(c) is proved by using the Lagrange multiplier method: Define $L : \mathbb{R}^{k+1} \to \mathbb{R}$ by

$$L(w_1,\dots,w_k,\lambda) := \sum_{i=1}^k w_i f_i - \theta \sum_{i=1}^k w_i^2 - \lambda\Big(\sum_{i=1}^k w_i - \gamma\Big), \qquad w_1,\dots,w_k,\lambda \in [0,\infty).$$

$(w_1,\dots,w_k,\lambda)$ is the extremal point for L if $\nabla L(w_1,\dots,w_k,\lambda) = 0$, which is the case if

$$f_i - \lambda - 2\theta w_i = 0 \qquad \text{for all } i, \qquad \text{and} \qquad \sum_{i=1}^k w_i = \gamma.$$

Combining gives $\lambda = \frac{1}{k}\sum_{i=1}^k f_i - 2\theta\gamma$ and (3.13). This extremal point for L is the maximizer for L over $\mathbb{R}^k$ as L is concave and because $\lim_{|x|\to\infty} xf_i - \theta x^2 = -\infty$ for all i.

(d) follows from (c) as the condition implies that $kf_j + 2\theta\gamma - \sum_{i=1}^k f_i \geq 0$ (remember $f_j \geq f_k$) for all j and thus $w_j \geq 0$ for w_j as in (3.13), i.e.,

$$w_j = \frac{f_j}{2\theta} + \frac{1}{k}\Big(\gamma - \sum_{i=1}^k \frac{f_i}{2\theta}\Big) \geq 0.$$

As furthermore,

$$\frac{f_j}{\theta} - w_j = \frac{f_j}{2\theta} - \frac{1}{k}\Big(\gamma - \sum_{i=1}^k \frac{f_i}{2\theta}\Big),$$

we have obtain (3.14) by the following equality:

$$\sum_{i=1}^k w_i f_i - \theta w_i^2 = \sum_{i=1}^k w_i(f_i - \theta w_i) = \theta \sum_{i=1}^k \Big(\Big(\frac{f_i}{2\theta}\Big)^2 - \frac{1}{k^2}\big(\gamma - \sum_{i=1}^k \frac{f_i}{2\theta}\big)^2\Big).$$

(e) follows from (d) as one observes that (3.14) is maximal when γ is closest to $\sum_{i=1}^k \frac{f_i}{2\theta}$.

(f) Suppose $\widetilde{w}_1,\dots,\widetilde{w}_k \in [0,1]$ for $k \geq 2$ are such that $\sum_{i=1}^k \widetilde{w}_i \leq 1$ (we may assume $k \geq 2$ as $\theta > 0$ so that (3.15) cannot be satisfied for $k = 1$). Let us define $w_1,\dots,w_k$ by $w_k = 0$ and for $i \in \{1,\dots,k-1\}$

$$w_i := \widetilde{w}_i + \tfrac{1}{k-1}\widetilde{w}_k.$$

Then by writing $\gamma = \sum_{i=1}^k \widetilde{w}_k = \sum_{i=1}^k w_k$, we see that

$$\begin{aligned}\sum_{i=1}^k \widetilde{w}_i f_i - \theta\widetilde{w}_i^2 - \Big(\sum_{i=1}^k w_i f_i - \theta w_i^2\Big) &= \widetilde{w}_k f_k - \theta\widetilde{w}_k^2 + \sum_{i=1}^{k-1}(\widetilde{w}_i - w_i)(f_i - \theta(\widetilde{w}_i + w_i)) \\ &= \widetilde{w}_k f_k - \theta\widetilde{w}_k^2 - \frac{\widetilde{w}_k}{k-1}\Big(\sum_{i=1}^{k-1} f_i - 2\theta\gamma\Big) \\ &= \frac{\widetilde{w}_k}{k-1}\Big((k-1)f_k - \sum_{i=1}^{k-1} f_i + 2\theta - (1-\gamma)2\theta - (k-1)\theta\widetilde{w}_k\Big) \\ &\le -\frac{\widetilde{w}_k}{k-1}\left((1-\gamma)2\theta + (k-1)\theta\widetilde{w}_k\right) \le -\theta\widetilde{w}_k^2.\end{aligned}$$

This proves that the maximizer has to satisfy $w_k = 0$. □

Proof of Proposition 3.8. Case 2 follows directly from Lemma 3.9 (b).

In Case 1, observe first that $jf_{j+1} \le \sum_{i=1}^j f_i - 2\theta$ for all $j > k_\star$, and that $f_1 + \cdots + f_{k_\star} \ge 2\theta$. By definition of $K_\star$ one has $(k_\star - 1)f_{k_\star} > \sum_{j=1}^{k_\star - 1} f_j - 2\theta$ and thus $f_{k_\star} > \frac{1}{k_\star}(\sum_{j=1}^{k_\star} f_j - 2\theta)$ and so $w_1 \ge w_2 \ge \cdots \ge w_{k_\star} > 0$. By Lemma 3.9 (f) it follows that $w_i = 0$ for $i > k_\star$ and so by (e) one completes the proof. □

3.3 Maximizers have finite support

In this section we prove that, if $\mathcal{P}$ is a good point process in $\mathcal{M}_{\mathrm{p}}^\circ$, then every maximizer μ^* of $\Psi_{\mathcal{P}}$ has a finite support. It is this result that needs one of the two conditions (i) or (ii) of Definition 2.7. Indeed, we will use (i) to construct, from a maximization candidate with infinitely many points, a better one with only finitely many points, and we will use (ii) for a simple argument that the maximizer has only finitely many points.

Proposition 3.10 (Maximizers have finite support)**.** *Let $\mathcal{P}$ be a point measure in $\mathcal{M}_{\mathrm{p}}^\circ$ that is good in the sense of Definition 2.7. Then*

- (a) *$\Psi_{\mathcal{P}}$ has at least one maximizer,*
- (b) *there exists a $R > 0$ such that every maximizer of $\Psi_{\mathcal{P}}$ lies in $\mathcal{W}_R$,*
- (c) *every maximizer has finite support,*

and, if $\mathcal{P}$ satisfies (i) of Definition 2.7, then

- (d) *every maximizer ν is a probability measure, i.e., $\nu = \sum_{i=1}^k w_i \delta_{(f_i,y_i)}$ for some $k \in \mathbb{N}$, $w \in [0,1]^k$ $\sum_{i=1}^k w_i = 1$, $f_i \in (0,\infty)$, $y_i \in \mathbb{R}^d$ for $i \in \{1,\dots,k\}$. Moreover, $\sum_{i=1}^k f_i > 2\theta$.*

Proof. (a) By Proposition 3.4, see also Remark 3.2, $\Psi_{\mathcal{P}}$ is upper semicontinuous. $\mathcal{W}$ is sequentially compact by (Kle08, Corollary 13.31). Therefore $\Psi_{\mathcal{P}}$ has at least one maximizer.

(b) By Proposition 3.4 (a), we may pick $R > 0$ so large that $\sup_{\mu\in\mathcal{W}\setminus\mathcal{W}_R} \Psi_{\mathcal{P}}(\mu) < 0 \le \sup_{\mu\in\mathcal{W}_R} \Psi_{\mathcal{P}}(\mu)$. This implies that every maximizer lies in $\mathcal{W}_R$.

(c) Let $\nu \in \mathcal{W}_R$ be a maximizer of $\Psi_{\mathcal{P}}$. Clearly, $\nu \ll \mathcal{P}$ (otherwise $\Psi_{\mathcal{P}}(\nu) = -\infty < \Psi_{\mathcal{P}}(0)$). Under (ii) of Definition 2.7, ν has finite support. Therefore we assume instead that (i) of Definition 2.7 holds. Moreover, without loss of generality we may assume that the supports of $\mathcal{P}$ and ν are infinite. We are going to show that there exists a $\mu \in \mathcal{W}_{R+1}$ such that $\operatorname{supp}\mu$ is a finite set and $\Psi_{\mathcal{P}}(\mu) > \Psi_{\mathcal{P}}(\nu)$, which implies the claim.

Let $(f_i, y_i) \in (0,\infty) \times \mathbb{R}^d$ for $i \in \mathbb{N}$ be distinct and such that $\{(f_i,y_i) : i \in \mathbb{N}\} = \operatorname{supp}\nu$. We may assume that $f_1 \ge f_2 \ge f_3 \ge \dots$ and $f_k \to 0$ as $k \to \infty$ (due to the fact that $[\varepsilon,\infty) \times Q_R$ is relatively compact in $\mathfrak{E}$, because it is a subset of $\mathcal{H}_h^s$ for some $s, h > 0$, and so there are only finitely many i such that $f_i \ge \varepsilon$ for all $\varepsilon > 0$). We separate the proof in two cases, depending on $\sum_{i\in\mathbb{N}} f_i$: Case 1: $\sum_{i\in\mathbb{N}} f_i \in (2\theta, \infty]$, Case 2: $\sum_{i\in\mathbb{N}} f_i \in [0, 2\theta]$.

Case 1 $\sum_{i\in\mathbb{N}} f_i \in (2\theta,\infty]$. The idea is that Proposition 3.8 tells us that $\mathcal{D}_{\mathcal{P}}$ is maximized using a finite number of points such that adding points the Φ part will not enlarge, but the $\mathcal{D}$ part will increase.

Let $\delta > 0$ be such that $\sum_{i\in\mathbb{N}} f_i \geq 2\theta + 2\delta$. Let K_1 be such that $\sum_{i=1}^{K_1} f_i > 2\theta + \delta$. Let $K_2 \geq K_1$ be such that $f_k < \frac{\delta}{K_1}$ for all $k > K_2$. Then for $k \geq K_2$ we have

$$\Phi_{\mathcal{P}}(\nu) \leq \varphi_k(f_1, \dots, f_k) + \delta, \tag{3.16}$$

$$\sum_{i=1}^{k} f_i - k f_{k+1} = \sum_{i=1}^{k}(f_i - f_{k+1}) \geq \sum_{i=1}^{K_1}(f_i - \frac{\delta}{K_1}) \geq 2\theta + \delta - \delta = 2\theta. \tag{3.17}$$

By (3.16) it follows that (as the above can be done for any $\delta > 0$), for φ_k as in (3.6),

$$\Phi_{\mathcal{P}}(\nu) \leq \sup_{k\in\mathbb{N}} \varphi_k(f_1, \dots, f_k).$$

By (3.16) it follows that for all $\delta > 0$ there exists a $K_2 > 0$ such that $\Phi_{\mathcal{P}}(\nu) \leq \varphi_k(f_1, \dots, f_k) + \delta$ for all $k \geq K_2$ and thus $\Phi_{\mathcal{P}}(\nu) \leq \sup_{k\in\mathbb{N}} \varphi_k(f_1, \dots, f_k) + \delta$ for all $\delta > 0$. By (3.17) it follows that there exists a $\ell \in \mathbb{N}$ such that $K_\star(f_1, \dots, f_k) = \ell$ for all $k \geq K_2$, where $K_\star$ is as in (3.7). Therefore, by Proposition 3.8, we have $\varphi_\ell(f_1, \dots, f_k) = \varphi_m(f_1, \dots, f_m)$ for all $m \geq K_2$ and thus, with w_i as in (3.8), for w given by

$$w(f, y) = \begin{cases} w_i & \text{if } i \in \{1, \dots, \ell\} \text{ and } (f, y) = (f_i, y_i), \\ 0 & \text{otherwise}, \end{cases}$$

we have for $\mu = w\mathcal{P}$ that $\Phi_{\mathcal{P}}(\mu) = \varphi_\ell(f_1, \dots, f_\ell) = \sup_{k\in\mathbb{N}} \varphi_k(f_1, \dots, f_k)$ and thus

$$\Phi_{\mathcal{P}}(\mu) \geq \Phi_{\mathcal{P}}(\nu), \qquad \mathcal{D}_{\mathcal{P}}(\mu) < \mathcal{D}_{\mathcal{P}}(\nu) \qquad \text{and therefore} \qquad \Psi_{\mathcal{P}}(\mu) > \Psi_{\mathcal{P}}(\nu).$$

Case 2 $\sum_{i\in\mathbb{N}} f_i \in [0, 2\theta]$. Let us first introduce some objects. For $k \in \mathbb{N}$, let

$$a_k := \frac{\sum_{i=k+1}^{\infty} f_i}{2\theta}.$$

Then $a_k \to 0$ as $k \to \infty$. Let $R_1 := \inf\{s > 0\colon \operatorname{supp}\nu \subset (0,\infty) \times Q_s\}$ the smallest r such that $\nu \in \mathcal{W}_r$. Then $R_1 \leq R$ since $\nu \in \mathcal{W}_R$. Then there exists a function $\varphi\colon \mathbb{N} \to \mathbb{N}$ such that $y_{\varphi(n)}$ converges to $z = (z_1, \dots, z_d) \in \partial Q_{R_1}$ as $n \to \infty$. Assume, without loss of generality, that $z_1 = R_1$. Let $\beta > 2$ and $\varepsilon = \varepsilon_{R_1+1,3\theta}$ as in the condition (i) of Definition 2.7 on $\mathcal{P}$. Pick $k \in \mathbb{N}$ such that $\theta a_k^2 - 8q a_k^\beta > 0$ and $3a_k^\beta \leq 1$ and $a_k < \varepsilon \wedge 1$. Then define

$$a = a_k \qquad \text{and} \qquad \bar{z} = (R_1 + 2a^\beta, z_2, \dots, z_d).$$

Observe that $\bar{z} \in Q_{R_1+1}$. By the assumption (i) on $\mathcal{P}$ from Definition 2.7 there exists a

$$(\widetilde{f}, \widetilde{y}) \in (3\theta a, \infty) \times B(\bar{z}, a^\beta) \cap \operatorname{supp}\mathcal{P}$$

We observe that $(\widetilde{f}, \widetilde{y}) \notin \operatorname{supp}\nu$ since $\widetilde{y} \notin Q_{R_1}$, but $\widetilde{y} \in Q_{R_1+3a^\beta} \subset Q_{R+1}$. Since $\lim_{n\to\infty} y_{\varphi(n)} = z$, there exists a $(\hat{f}, \hat{y}) \in \operatorname{supp}\nu$ such that $\hat{y} \in B(z, a^\beta)$ and therefore $(\widetilde{f}, \widetilde{y}) \in (3\theta a, \infty) \times B(\hat{y}, 4a^\beta)$.

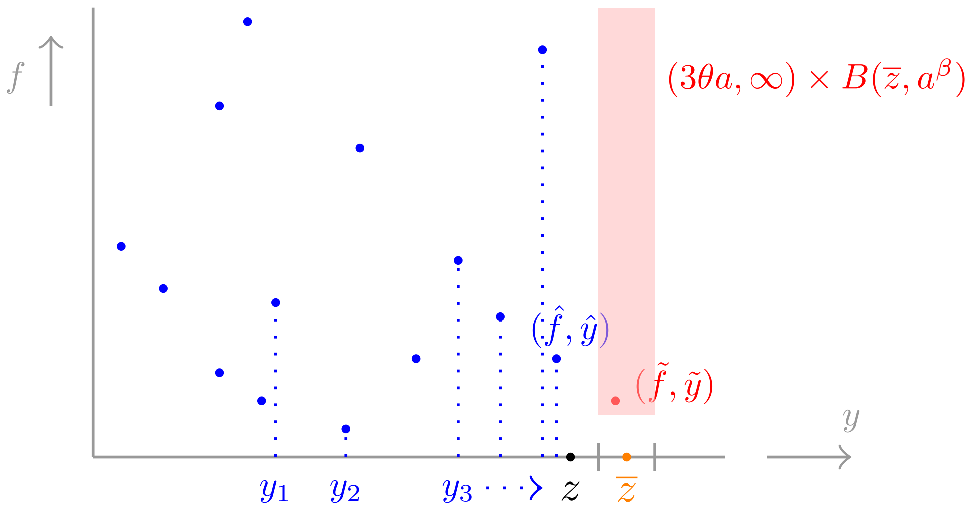

FIGURE 4. Visualisation of z, $\overline{z}$ and $(\tilde{f}, \tilde{y})$ (for $d = 1$).

We recall Proposition 3.8, Case 2, which tells us that the unique maximizer of φ_k as in (3.6), is given by $(\frac{f_1}{2\theta}, \dots, \frac{f_k}{2\theta})$. Define $\mu = w\mathcal{P}$ (i.e., $\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{P}} = w$), where

$$w(f,y) = \begin{cases} \frac{f_i}{2\theta} & \text{if } (f,y) = (f_i, y_i) \text{ for some } i \in \{1, \dots, k\}, \\ a & \text{if } (f,y) = (\widetilde{f}, \widetilde{y}), \\ 0 & \text{otherwise.} \end{cases} \tag{3.18}$$

Because

$$\sum_{i=1}^{k} \frac{f_i}{2\theta} + a = \frac{\sum_{i=1}^{\infty} f_i}{2\theta} \in [0,1],$$

it is clear that $\mu \in \mathcal{W}_{R+1}$ and that μ has only finite support. We are going to show that $\Psi_{\mathcal{P}}(\mu) > \Psi_{\mathcal{P}}(\nu)$.

Observe that

$$\sum_{i=k+1}^{\infty} \frac{f_i^2}{4\theta} \leq \frac{1}{4\theta} \left(\sum_{i=k+1}^{\infty} f_i \right)^2 = \frac{1}{4\theta} (2\theta a)^2 = \theta a^2.$$

Therefore, by using that $\sum_{i=1}^{k} \frac{f_i^2}{4\theta} = \sum_{i=1}^{\infty} \frac{f_i^2}{4\theta} - \sum_{i=k+1}^{\infty} \frac{f_i^2}{4\theta}$, we see that

$$\Phi_{\mathcal{P}}(\mu) \geq \sum_{i=1}^{k} w_i f_i + a\widetilde{f} - \theta\Big(\sum_{i=1}^{k} w_i^2 + a^2 \Big) = \sum_{i=1}^{k} \frac{f_i^2}{4\theta} + a\widetilde{f} - \theta a^2 \geq \sum_{i=1}^{\infty} \frac{f_i^2}{4\theta} + a\widetilde{f}. \tag{3.19}$$

By Proposition 3.8 it follows that

$$\Phi_{\mathcal{P}}(\nu) \leq \sum_{i=1}^{\infty} \frac{f_i^2}{4\theta}. \tag{3.20}$$

Moreover, we recall the definition of $\mathcal{D}_{\mathcal{P}}$ in (1.20) and see that

$$\mathcal{D}_{\mathcal{P}}(\nu) \geq D_0(\{y_1, \dots, y_k, \hat{y}\}) \geq D_0(\{y_1, \dots, y_k, \hat{y}, \widetilde{y}\}) - 8a^\beta, \tag{3.21}$$

where the last inequality holds true since any path starting from 0 and visiting all points in $\{y_1, \dots, y_k, \hat{y}\}$ can be extended to visit $\widetilde{y}$ as well by traveling back and forth along the straight line linking $\hat{y}$ and $\widetilde{y}$ (which are at most $4a^\beta$ apart from each other). Since $\mathrm{supp}_{\mathbb{R}^d}\, \mu = \{y_1, \dots, y_k, \widetilde{y}\}$,

$$D_0(\{y_1, \dots, y_k, \hat{y}, \widetilde{y}\}) \geq D_0(\{y_1, \dots, y_k, \widetilde{y}\}) = \mathcal{D}_{\mathcal{P}}(\mu).$$

Thus, we deduce from (3.21) that

$$\mathcal{D}_{\mathcal{P}}(\mu) \leq \mathcal{D}_{\mathcal{P}}(\nu) + 8a^\beta, \tag{3.22}$$

and therefore, using (3.19), (3.20) and (3.22) in combination with the fact that $\widetilde{f} \geq 3\theta a$ we obtain

$$\Psi_{\mathcal{P}}(\mu) - \Psi_{\mathcal{P}}(\nu) \geq -\sum_{i=k+1}^{\infty} \frac{f_i^2}{4\theta} + a\widetilde{f} - \theta a^2 - 8qa^\beta \geq a\widetilde{f} - 2\theta a^2 - 8qa^\beta > \theta a^2 - 8qa^\beta > 0. \tag{3.23}$$

(d) Suppose that $\mu = \sum_{i=1}^k w_i \delta_{(f_i, y_i)}$. If μ is a maximizer and $\sum_{i=1}^k f_i \geq 2\theta$, then $\sum_{i=1}^k w_i = 1$ because of Proposition 3.8 (as, like in Case 1, $\Phi_{\mathcal{P}}(\mu) = \varphi_k(f_1, \dots, f_k)$), so that μ is a probability measure.

If $\sum_{i=1}^k f_i < 2\theta$ and μ maximizes $\Phi_{\mathcal{P}}$ (observe $\Phi_{\mathcal{P}}$, not $\Psi_{\mathcal{P}}$), then by Proposition 3.8, $\sum_{i=1}^k w_i < 1$, i.e., μ is not a probability measure. Moreover, like in Case 2 above, one can show that μ is not a maximizer of $\Psi_{\mathcal{P}}$: Indeed, one chooses an $a \in (0, \varepsilon \wedge 1)$ with $a \leq 1 - \sum_{i=1}^k w_i$, $\theta a^2 - 8qa^\beta > 0$ and $3a^\beta \leq 1$ and follows the same lines as in Case 2 to find a $(\tilde{f}, \tilde{y})$ such that $\Psi_{\mathcal{P}}(\mu + a\delta_{(\tilde{f}, \tilde{y})}) > \Psi_{\mathcal{P}}(\mu)$. □

3.4 Proof of Theorem 2.8

In this section, we prove Theorem 2.8 subject to Proposition 3.4 (b) (Gamma-convergence of $-\Psi$) and Proposition 3.10 (finiteness of support of maximizers).

Prove of Theorem 2.8. (a) follows directly by Proposition 3.10. (c) follows by the fact that $\Psi_{\mathcal{P}}$ is upper-semicontinuous and $\mathcal{W}$ is sequentially compact (see the beginning of the proof of Proposition 3.10), so that, in particular, O^c is sequentially compact for any open set $O \subset \mathcal{W}$. Therefore there exists a maximizer of $\Psi_{\mathcal{P}}$ on O^c, which by the uniqueness cannot be equal to the maximizer over $\mathcal{W}$, therefore proving the desired inequality. (d) follows from Theorem 3.3 (b) and Proposition 3.4 (b).

Let us now prove (b). The idea is that the points in G are all worth visiting because of their large energy values, but at the same time they are not distinct enough so as to give preference to only a couple of them. Having no points in E^1, E^2 and E^3 contributes to make points outside of G not worth visiting, because either their energy values are too low or their distance too large.

Fix $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}^{\circ}$ with $\mathcal{P}(G) = k$ and $\mathcal{P}(E^i) = 0$, $i = 1, 2, 3$. Denote by $(f_1, y_1), \dots, (f_k, y_k)$ the k points of $\mathcal{P}$ in G, with $f_1 \geq f_2 \geq \dots \geq f_k$. Take w_i as in (3.8) of Proposition 3.8 and let $\nu^* := \sum_{i=1}^k w_i \delta_{(f_i, y_i)}$. Note that, since $f_i \geq L > 2\theta$ for $1 \leq i \leq k$, the relevant formulas from Proposition 3.8 will (mostly) be (3.8) and (3.9). We divide the proof into the following steps:

(Step 1) $w_i > 0$ for all $i \in \{1, \dots, k\}$.
(Step 2) If $\nu \in \mathcal{W}$ and $\mathrm{supp}\, \nu \subset G$ then $\Psi_{\mathcal{P}}(\nu) \leq \Psi_{\mathcal{P}}(\nu^*)$;
(Step 3) If $\nu \in \mathcal{W}$ and $\mathrm{supp}\, \nu \not\subset G$ then $\Psi_{\mathcal{P}}(\nu) \leq \Psi_{\mathcal{P}}(\nu^*)$.

Steps 2–3 together with (a) will then show that $\mu^* = \nu^*$, and this together with Step 1 implies (b).

Step 1 By Proposition 3.8 it suffices to check that $k = K_\star(f_1, \dots, f_k)$, where $K_\star$ is as in (3.7). This follows as for any $j \in \{1, \dots, k-1\}$ we have

$$\sum_{i=1}^{j} f_i - 2\theta \leq j(L + \frac{2\theta}{k}) - 2\theta = jL - \frac{(k-j)2\theta}{k} \leq jf_{j+1} - \frac{2\theta}{k} < jf_{j+1}. \tag{3.24}$$

Step 2 We can assume that $\nu \ll \mathcal{P}$. We will first show the statement for $\nu \in \mathcal{W}$ which are nonzero. Observe that by (3.5), for any $\nu \in \mathcal{W}$ with $\emptyset \neq \operatorname{supp}\nu \subset G$,

$$\Psi_{\mathcal{P}}(\nu) = \Phi_{\mathcal{P}}(\nu) - \mathcal{D}_{\mathcal{P}}(\nu) \leq \varphi_{|J|}((f_j)_{j\in J}) - qD_0((y_j)_{j\in J}),$$

where $J \subset \{1,\dots,k\}$ is such that $\operatorname{supp}\nu = \{(f_j,y_j) : j \in J\}$. Therefore it suffices to show that for all $J \subset \{1,\dots,k\}$, $J \neq \emptyset$, one has

$$\varphi_{|J|}((f_j)_{j\in J} - qD_0((y_j)_{j\in J}) \leq \Psi_{\mathcal{P}}(\nu^*) = \varphi_k(f_1,\dots,f_k) - qD_0(y_1,\dots,y_k).$$

Of course for $k = 1$ the above is clear. Suppose that $k \geq 2$. Let $J \subset \{1,\dots,k\}$ with $m := |J| \leq k-1$ and $\ell \in \{1,\dots,k\} \setminus J$. By an inductive argument on m, it suffices to show that

$$\varphi_m((f_i)_{i\in J}) - qD_0((y_i)_{i\in J}) \leq \varphi_{m+1}((f_i)_{i\in J\cup\{\ell\}}) - qD_0((y_i)_{i\in J\cup\{\ell\}}). \tag{3.25}$$

First of all, a straightforward computation using (3.9) gives setting $S = \sum_{i\in J} f_i - 2\theta$ and using that $\frac{1}{m} - \frac{1}{m+1} = \frac{1}{m(m+1)}$,

$$\varphi_{j+1}((f_i)_{i\in J\cup\{\ell\}}) - \varphi_j((f_i)_{i\in J}) \tag{3.26}$$

$$= \frac{1}{4\theta}\Big(\sum_{i\in J\cup\{\ell\}} f_i^2 - \frac{1}{m+1}\Big(\sum_{i\in J\cup\{\ell\}} f_i - 2\theta\Big)^2\Big) - \frac{1}{4\theta}\Big(\sum_{i\in J} f_i^2 - \frac{1}{m}\Big(\sum_{i\in J} f_i - 2\theta\Big)^2\Big) \tag{3.27}$$

$$= \frac{1}{4\theta}\Big[f_\ell^2 - \frac{1}{m+1}\Big(f_\ell + \sum_{i\in J} f_i - 2\theta\Big)^2\Big) + \frac{1}{m}\Big(\sum_{i\in J} f_i - 2\theta\Big)^2\Big)\Big] \tag{3.28}$$

$$= \frac{1}{4\theta}\Big[f_\ell^2 - \frac{1}{m+1}\Big(f_\ell^2 + 2f_\ell S + S^2\Big) + \frac{1}{m}S^2\Big)\Big] \tag{3.29}$$

$$= \frac{1}{4\theta}\Big[\frac{m}{m+1} f_\ell^2 - \frac{2f_\ell S}{m+1} + \frac{1}{m(m+1)}S^2\Big)\Big] = \frac{1}{4\theta m(m+1)}\Big[m f_\ell - S\Big]^2 \tag{3.30}$$

$$= \frac{1}{4\theta m(m+1)}\Big[m f_\ell - \Big(\sum_{i\in J} f_i - 2\theta\Big)\Big]^2 \geq \frac{1}{4\theta m(m+1)}\big(\frac{2\theta}{k}\big)^2 \geq \frac{\theta}{k^4} = 4q\varepsilon, \tag{3.31}$$

where in the last line we used that

$$\sum_{i\in J} f_i - 2\theta \leq m f_\ell - \frac{2\theta(k-m)}{k}, \tag{3.32}$$

which follows similarly as the estimate in (3.24). From this and the observations

$$D_0((y_i)_{i\in J\cup\{\ell\}}) - D_0((y_i)_{i\in J}) \leq 4\varepsilon, \quad 4q\varepsilon - 4\varepsilon > 0,$$

we deduce (3.25). This basically follows from the observation that

$$|y_1 - y_2| + |y_2 - y_3| \leq |y_1 - y_3| + 2|y_2 - y_3| \leq |y_1 - y_3| + 2|y_2| + 2|y_3|,$$
$$|y_1 - y_2| + |y_2 - y_3| \leq |y_1 - y_2| + |y_2| + |y_3.$$

In order to finish this step, it suffices to observe that (because $L > 2\theta + (q+1)\varepsilon$),

$$\Psi_{\mathcal{P}}(\nu^*) \geq \varphi_1(f_1) - qD_0(y_1) = f_1 - \theta - q|y_1| \geq L - \theta - q\varepsilon > \varepsilon > 0. \tag{3.33}$$

Step 3 By (a) it suffices to show this step for ν with $\operatorname{supp}\nu \subset \mathcal{P}$ finite support. Let $(f,y) \in \operatorname{supp}\nu \setminus G$ be such that f is maximal among such points. Assume first that $f > \varepsilon$. In this case, our assumptions on $\mathcal{P}$ imply that $|y| \geq f + 3\theta$. If $f \geq L + \frac{2\theta}{k}$, then $\Phi_{\mathcal{P}}(\nu) \leq f$, $\mathcal{D}_{\mathcal{P}}(\nu) \geq |y| \geq f + 3\theta \geq f$, and thus

$$\Psi_{\mathcal{P}}(\nu) = \Phi_{\mathcal{P}}(\nu) - q\mathcal{D}_{\mathcal{P}}(\nu) \leq f - qf < 0 < \Psi_{\mathcal{P}}(\nu^*).$$

If instead $\varepsilon < f \le L + 2\theta/k$, then $\Phi_{\mathcal{P}}(\nu) \le L + \frac{2\theta}{k}$ and $|y| \ge f + 3\theta > \varepsilon + \theta + \frac{2\theta}{k}$, so (remember (3.33))

$$\Psi_{\mathcal{P}}(\nu) \le L + \frac{2\theta}{k} - q|y| \le L - (q-1)\frac{2\theta}{k} - q\theta - q\varepsilon < L - \theta - q\varepsilon \le \Psi_{\mathcal{P}}(\nu^*).$$

Assume now that $f \le \varepsilon$. If $\operatorname{supp}\nu \cap G = \emptyset$ then $\Psi_{\mathcal{P}}(\nu) \le \varepsilon \le \Psi_{\mathcal{P}}(\nu^*)$ (because of (3.33) again). Lastly, suppose $\operatorname{supp}\nu \cap G = \{(f_i, y_i)_{i\in J}\} \ne \emptyset$ where $J \subset \{1,\dots,k\}$ with $|J| = m \ge 1$. Let $N \in \mathbb{N}$ be the number of points in $\operatorname{supp}\nu \setminus G$. Denote by $(f_{k+1}, y_{k+1}), \dots, (f_{k+N}, y_{k+N})$ the points in $\operatorname{supp}\nu \setminus G$ with $f = f_{k+1} \ge f_{k+2} \ge \dots \ge f_{k+N}$. Observe that

$$mf \le m\varepsilon < mL - 2\theta \le \sum_{j\in J} f_j - 2\theta,$$

so that for $I = \{k+1, \dots, k+N\}$, $K_\star((f_i)_{i\in J\cup I}) = K_\star((f_i)_{i\in J})$. The latter equals $|J|$ due to (3.32) and thus $\Phi_{\mathcal{P}}(\nu) \le \varphi_{|J\cup I|}((f_i)_{i\in J\cup I}) \le \varphi_m((f_i)_{i\in J}) \le \Phi_{\mathcal{P}}(\nu^*)$ and so $\Psi_{\mathcal{P}}(\nu) \le \Psi_{\mathcal{P}}(\nu^*)$. □

4. Goodness of Π and of Π_t

In order that we can apply Theorem 2.8 to the rescaled process Π_t (for all $t > 0$) defined in (1.9) and to the PPP Π defined in (1.10), we show in this section that they are good in the sense of Definition 2.7 for any $\alpha \in (2d, \infty)$. Moreover, we prove Lemma 2.11 for $\alpha \in (d, \infty)$. That is, we prove Lemma 2.9, see Lemma 4.2 and Lemma 4.4. That both Π and Π_t can be viewed as elements of $\mathcal{M}_{\mathrm{p}}^\circ$, has been shown in Lemma 2.3.

Lemma 4.1. *Let $s, r > 0$ and $x \in \mathbb{R}^d$. Let $V_d \in (0,\infty)$ be the volume of the unit ball in $\mathbb{R}^d$. Then*

$$\mathbf{P}\Big(\Pi\big([s,\infty) \times B(x,r)\big) = 0\Big) = \mathrm{e}^{-V_d s^{-\alpha} r^d}.$$

Proof. The random variable $\Pi([s,\infty) \times B(x,r))$ is Poisson distributed with parameter

$$\int_{[s,\infty)\times B(x,r)} \alpha y^{-(1+\alpha)}\,\mathrm{d}y \otimes \mathrm{d}z = \int_{[s,\infty)} \frac{\alpha}{y^{1+\alpha}}\,\mathrm{d}y\; V_d r^d = V_d s^{-\alpha} r^d. \tag{4.1}$$

□

Lemma 4.2. *Let $t \in (0,\infty)$.*

(a) *Let $\alpha \in (0,\infty)$. With probability one, Π_t satisfies (ii) of Definition 2.7.*
(b) *Let $\alpha \in (2d,\infty)$. With probability one, Π satisfies (i) of Definition 2.7.*

Proof. (a) By construction Π_t satisfies (ii) since $\Pi_t\big((0,\infty) \times Q_R\big)$ is simply the cardinality of $\{x \in \mathbb{Z}^d : |x| \le Rr_t\}$, which is $R^d r_t^d$.

(b) Let $\beta \in (2, \frac{\alpha}{d})$. First observe that Q_R can be covered by balls $B(z, \frac{1}{k})$ with $z \in \frac{1}{k}\mathbb{Z}^d$ and $z \in Q_{R+1}$, i.e.,

$$Q_R \subset \bigcup_{z\in(\frac{1}{k}\mathbb{Z}^d)\cap Q_{R+1}} B(z, \tfrac{1}{k}).$$

Then, observe that therefore, with probability one Π satisfies (i) of Definition 2.7 if (replace ε by $k^{-\frac{1}{\beta}}$)

$$\mathbf{P}\Big(\bigcup_{R,C\in(0,\infty)\cap\mathbb{Q}} \bigcap_{N\in\mathbb{N}} \bigcup_{k\ge N} \bigcup_{z\in(\frac{1}{k}\mathbb{Z}^d)\cap Q_{R+1}} \{\Pi \cap ([Ck^{-\frac{1}{\beta}},\infty) \times B(z,\tfrac{1}{k})) = \emptyset\}\Big) = 0.$$

By the Borel–Cantelli Lemma, the above holds if we can show that for any $R, C \in (0,\infty)$,

$$\sum_{k\in\mathbb{N}} \mathbf{P}\Big(\bigcup_{z\in(\frac{1}{k}\mathbb{Z}^d)\cap Q_{R+1}} \{\operatorname{supp}\Pi \cap ([Ck^{-\frac{1}{\beta}},\infty) \times B(z,\tfrac{1}{k})) = \emptyset\}\Big) < \infty.$$

Let $R, C > 0$. We may assume $R > 2$. By estimating the probability of the union by the sum of the probabilities, and observing that $\#(\frac{1}{k}\mathbb{Z}^d) \cap Q_{R+1} \le (2k(R+1)+1)^d \le (5kR)^d$, (indeed, use that $R+1 \le 2R$ and $4kR+1 \le 5kR$) by Lemma 4.1 (with $s = Ck^{-\frac{1}{\beta}}$, $r = \frac{1}{k}$, so that $s^{-\alpha} r^d = C^{-\alpha} k^{\frac{\alpha}{\beta}-d}$),

$$\sum_{k\in\mathbb{N}} \mathbf{P}\Big(\bigcup_{z\in(\frac{1}{k}\mathbb{Z}^d)\cap Q_{R+1}} \{\operatorname{supp}\Pi \cap ([Ck^{-\frac{1}{\beta}}, \infty) \times B(z, \tfrac{1}{k})) = \emptyset\} \Big) \le \sum_{k\in\mathbb{N}} (5kR)^d \mathrm{e}^{-V_d C^{-\alpha} k^{\frac{\alpha}{\beta}-d}}.$$

Because $\beta < \frac{\alpha}{d}$, we have $\frac{\alpha}{\beta} > d$ and therefore the above sum is finite. □

Lemma 4.3. *Let $\mathcal{P} \in \mathcal{M}_{\mathrm{p}}^{\circ}$. Suppose $\nu \in \mathfrak{F}(\mathcal{P})$, $\nu \neq 0$ and $\Psi_\Pi(\nu) = \sup_{\mu\in\mathfrak{F}(\Pi)} \Psi_{\mathcal{P}}(\mu)$. Let $k \in \mathbb{N}$, $(f_1, y_1), \dots, (f_k, y_k) \in \mathcal{P}$ be such that $\operatorname{supp}\nu = \{(f_1, y_1), \dots, (f_k, y_k)\}$. Then $\Phi_{\mathcal{P}}(\nu) = \widetilde{\varphi}_k(f_1, \dots, f_k)$, where*

$$\widetilde{\varphi}_k(f_1, \dots, f_k) = \frac{1}{4\theta}\Big(\sum_{i=1}^k f_i^2 - \frac{1}{k}\Big(\sum_{i=1}^k f_i - 2\theta \Big)^2 \Big). \tag{4.2}$$

Proof. Observe that (for φ_k as in (3.6))

$$\Phi_{\mathcal{P}}(\nu) = \varphi_k(f_1, \dots, f_k), \qquad \mathcal{D}_{\mathcal{P}}(\nu) = D_0(y_1, \dots, y_k).$$

By the definition of the support, we have $\nu = \sum_{i=1}^k w_i \delta_{(f_i, y_i)}$ for some $w_1, \dots, w_k \in (0, 1]$. By Lemma 3.9 and (f) we may assume that $k \min\{f_1, \dots, f_k\} + 2\theta - \sum_{i=1}^k f_i > 0$ (otherwise $w_i = 0$ for some i). Therefore, by (e) of that lemma, it follows that $\Phi_{\mathcal{P}}(\nu) = \widetilde{\varphi}_k(f_1, \dots, f_k)$ (see also (3.8) and (3.9)). □

Lemma 4.4. *Let $\alpha \in (0, \infty)$ and $t \in (0, \infty)$. Recall the definition of $\mathfrak{F}(\mathcal{P})$ in (2.4).*

(a) *With probability one, Ψ_Π possesses at most one maximizer in $\mathfrak{F}(\Pi)$.*
(b) *With probability one, Ψ_{Π_t} possesses at most one maximizer in $\mathfrak{F}(\Pi_t)$. Moreover, for $L > 0$ and $\Pi_t^{(L)} = \mathbb{1}_{[L^{-1},\infty)\times Q_L} \Pi_t$ (see also (3.2)), the function $\Psi_{\Pi_t^{(L)}}$ possesses at most one maximizer in $\mathfrak{F}(\Pi_t)$.*

Proof. (a) We show that the event that there exist $\mu_1, \mu_2 \in \mathfrak{F}(\Pi)$ with $\mu_1 \neq \mu_2$ and $\Psi_\Pi(\mu_1) = \Psi_\Pi(\mu_2) = \sup_{\mu\in\mathfrak{F}(\Pi)} \Psi_\Pi(\mu)$, has probability zero. For this it suffices to show that $\mathbf{P}(\mathcal{N}_L) = 0$ for any $L > 0$, where, with $S_L := [L^{-1}, \infty) \times Q_L$,

$$\mathcal{N}_L = \Big\{ \exists \mu_1, \mu_2 \in \mathfrak{F}(\Pi)\colon\ \mu_1 \neq \mu_2,\ \operatorname{supp}\mu_i \subset S_L, \Psi_\Pi(\mu_1) = \Psi_\Pi(\mu_2) = \sup_{\mu\in\mathfrak{F}(\Pi)} \Psi_\Pi(\mu) \Big\}.$$

Let $L > 0$. We give an explicit almost sure description of Π on S_L. Let us write Θ for the intensity measure of Π, i.e., $\Theta(\mathrm{d}(f, y)) = \alpha f^{-(1+\alpha)}\,\mathrm{d}f \otimes \mathrm{d}y$. Let N be a Poisson distributed variable with parameter m_L, where $m_L = \Theta(S_L)$. Let $(F_j, Y_j)_{j\in\mathbb{N}}$ be i.i.d. random variables that are independent from N and whose law is given by $\frac{1}{m_L} \mathbb{1}_{S_L} \Theta$. Then,

$$\mathbb{1}_{S_L}\Pi \ \text{ is equal in distribution to } \ \sum_{j=1}^N \delta_{(F_j, Y_j)}.$$

Without loss of generality, we may assume $\mathbb{1}_{S_L}\Pi = \sum_{j=1}^N \delta_{(F_j, Y_j)}$. Then, by Lemma 4.3 $\mathcal{N}_L$ is included in the event (we make abuse of notation and for $m = 0$ we understand $\widetilde{\varphi}_m(F_{j_1}, \dots, F_{j_m})$ and $D_0(Y_{j_1}, \dots, Y_{j_m})$ to be equal to 0)

$$\Big\{ \exists k, m \in \mathbb{N}_0\ \exists (f_1, y_1), \dots, (f_k, y_k), (g_1, z_1), \dots, (g_m, z_m) \in S_L \cap \operatorname{supp}\Pi\colon$$
$$\{(f_1, y_1), \dots, (f_k, y_m)\} \neq \{(g_1, z_1), \dots, (g_m, z_m)\},$$
$$\widetilde{\varphi}_k(f_1, \dots, f_k) = \widetilde{\varphi}_m(g_1, \dots, g_m) + D_0(y_1, \dots, y_k) - D_0(z_1, \dots, z_m) \Big\},$$

therefore in the event

$$\left\{\begin{array}{l}\text{There exist } k,m\in\mathbb{N}_0,\ \text{distinct } i_1,\dots,i_k,i_* \text{ and distinct } j_1,\dots,j_m \text{ in } \{1,\dots,N\}\\ \quad\text{such that } i_*\notin\{j_1,\dots,j_m\} \text{ and}\\ \widetilde{\varphi}_{k+1}(F_{i_1},\dots,F_{i_k},F_{i_*})=\widetilde{\varphi}_m(F_{j_1},\dots,F_{j_m})+D_0(Y_{i_1},\dots,Y_{i_k},Y_{i_*})-D_0(Y_{j_1},\dots,Y_{j_m})\end{array}\right\}. \tag{4.3}$$

The above event is included in the one where we replace $\{1,\dots,N\}$ by $\mathbb{N}$. Therefore, it suffices to let $k,m\in\mathbb{N}_0$, take distinct $i_1,\dots,i_k,i_*$ and distinct $j_1,\dots,j_m$ in $\mathbb{N}$ such that $i_*\notin\{j_1,\dots,j_m\}$ and show that

$$\mathbf{P}\Big(\widetilde{\varphi}_{k+1}(F_{i_1},\dots,F_{i_k},F_{i_*})=\widetilde{\varphi}_m(F_{j_1},\dots,F_{j_m})+D_0(Y_{i_1},\dots,Y_{i_k},Y_{i_*})-D_0(Y_{j_1},\dots,Y_{j_m})\Big)=0.$$

Then, it follows that (4.3) has probability zero by Lemma 4.5.

(b) Follows similar as the above argument: Besides replacing Π by Π_t, replace Θ by the product measure of $\alpha f^{-(1+\alpha)}\mathbb{1}_{[r_t^{-d/\alpha},\infty)}(f)\,\mathrm{d}f$ and $\sum_{z\in r_t^{-1}\mathbb{Z}^d}\delta_z$, and N by $\#(Q_L\cap\mathbb{Z}^d)$. Then again, one can show that the event (4.3) has zero probability by applying Lemma 4.5. From this, the "moreover" part immediately follows too. □

Lemma 4.5. *Suppose that $F_1,F_2,\dots$ are i.i.d. random variables with values in $(0,\infty)$ whose law has a density with respect to the Lebesgue measure. Let $Y_1,Y_2,\dots$ be i.i.d. random variables with values in $\mathbb{R}^d$. Let $k,m\in\mathbb{N}_0$. Suppose that $i_1,\dots,i_k,i_*$ are distinct element of $\mathbb{N}$ and $j_1,\dots,j_m$ are distinct elements in $\mathbb{N}$ such that $i_*\notin\{j_1,\dots,j_m\}$. Then*

$$\mathbf{P}\Big(\widetilde{\varphi}_{k+1}(F_{i_1},\dots,F_{i_k},F_{i_*})=\widetilde{\varphi}_m(F_{j_1},\dots,F_{j_m})+D_0(Y_{i_1},\dots,Y_{i_k},Y_{i_*})-D_0(Y_{j_1},\dots,Y_{j_m})\Big)=0. \tag{4.4}$$

Proof. We explain the following argument in more detail below. If we condition the above event in (4.4) on all variables except F_{i_*}, that is, on $F_{i_1},\dots,F_{i_k},F_{j_1},\dots,F_{j_m}$, $Y_{i_1},\dots,Y_{i_k},Y_{j_1},\dots,Y_{j_m}$ and Y_{i_*}, then by the formula for $\widetilde{\varphi}_k$ (4.2), there exist $C_1,C_2,C_3\in\mathbb{R}$, $C_1\neq 0$ or $C_2\neq 0$ such that the event in the probability of (4.4) becomes

$$C_1F_{i_*}^2+C_2F_{i_*}+C_3=0.$$

The probability of such event is equal to zero as F_{i_*} has a density with respect to the Lebesgue measure.

Indeed, observe that $\widetilde{\varphi}_1(f_{i_*})=\frac{1}{4\theta}(4\theta f_{i_*}-4\theta^2)=f_{i_*}-\theta=A_0f_{i_*}^2+B_0f_{i_*}-C_0$, for $A_0=0$, $B_0=1$ and $C_0=-\theta$, and for $k\in\mathbb{N}$,

$$A_k=\frac{1}{4\theta}\big(1-\frac{4\theta^2}{k+1}\big),\qquad B_k=B_k(f_1,\dots,f_k)=\frac{2}{4\theta(k+1)}\Big(\sum_{i=1}^k f_i-2\theta\Big),$$
$$C_k=C_k(f_1,\dots,f_k)=\frac{1}{4\theta}\Big(\sum_{i=1}^k f_i^2-\frac{1}{k+1}\Big(\sum_{i=1}^k f_i-2\theta\Big)^2\Big),$$

that

$$\widetilde{\varphi}_{k+1}(f_1,\dots,f_k,f_{i_*})=A_kf_{i_*}^2+B_kf_{i_*}+C_k.$$

So that for $\tilde{C}_k=C_k-(\widetilde{\varphi}_m(f_{j_1},\dots,f_{j_m})+D_0(y_{i_1},\dots,y_{i_k},y_{i_*})-D_0(y_{j_1},\dots,y_{j_m})\Big))$, we have

$$\mathbf{P}\left(\begin{matrix}\widetilde{\varphi}_{k+1}(F_{i_1},\dots,F_{i_k},F_{i_*})=\widetilde{\varphi}_m(F_{j_1},\dots,F_{j_m})\\ +D_0(Y_{i_1},\dots,Y_{i_k},Y_{i_*})-D_0(Y_{j_1},\dots,Y_{j_m})\end{matrix}\,\middle|\,\overline{F}=\overline{f},\overline{Y}=\overline{y}\right)=\mathbf{P}(A_kF_{i_*}^2+B_kF_{i_*}+\tilde{C}_k=0), \tag{4.5}$$

where

$$\overline{F}=(F_{i_1},\dots,F_{i_k},F_{j_1},\dots,F_{j_m}),\qquad \overline{f}=(f_{i_1},\dots,f_{i_k},f_{j_1},\dots,f_{j_m}),$$
$$\overline{Y}=(Y_{i_1},\dots,Y_{i_k},Y_{i_*},Y_{j_1},\dots,Y_{j_m}),\qquad \overline{y}=(y_{i_1},\dots,y_{i_k},y_{i_*},y_{j_1},\dots,y_{j_m}).$$

As the law of F_{i_*} has a density with respect to the Lebesgue measure, the right-hand side (and thus the left-hand side) of (4.5) equals zero. □

For the proof of Lemma 2.11, we use the following lemma.

Lemma 4.6. *Let $\lambda \in (0,\infty)$. Let ζ be a PPP on $(0,1)^d$ with intensity λ. If $k \le (\lambda/4)^{1/d}$, then*

$$\mathbf{P}\Big(\exists \text{ distinct } Z_1,\dots,Z_k \in \zeta \colon D_0(Z_1,\dots,Z_k) < d\Big) \ge 1 - \exp\left(-\left(\frac{\lambda}{4}\right)^{\frac{1}{d}}\right).$$

Proof. Let ω be a PPP on $\mathbb{R}^d$ with intensity λ. In an almost sure and inductive sense we define sequences $(R_i)_{i\in\mathbb{N}}$ in $(0,\infty)$ and $(Y_i)_{i\in\mathbb{N}}$ in $[0,\infty)^d$ by setting $R_0 = 0$ and $Y_0 = 0$ and (on the probability one set such that the following infima are finite)

$$R_{i+1} := \inf\left\{r > 0 \colon \ \omega(Y_i + (0,r]^d) > 0\right\} \ ,$$

and by letting $Y_{i+1} \in [0,\infty)^d$ be the unique point in $\operatorname{supp}\omega \cap \big(Y_i + (0,R_{i+1}]^d\big)$ (see also Figure 5).

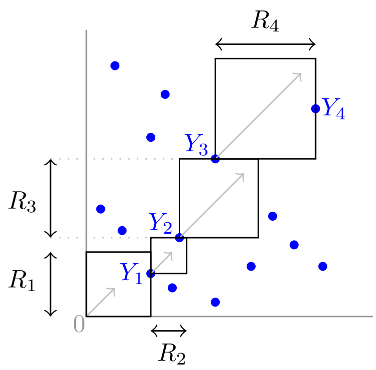

FIGURE 5. Illustration of choosing R_1, R_2, R_3, R_4 and Y_1, Y_2, Y_3, Y_4.

Observe that

$$\mathbf{P}\Big(\exists \text{ distinct } Z_1,\dots,Z_k \in \zeta \colon D_0(Z_1,\dots,Z_k) < d\Big) \ge \mathbf{P}\Big(\{Y_1,\dots,Y_k\} \subset (0,1)^d, \sum_{i=1}^k |Y_i - Y_{i-1}| \le d\Big)$$

Note that $Y_i \in Q_{R_1+\dots+R_i}$, $|Y_i - Y_{i-1}| \le dR_i$ and $(R_1+\dots+R_k)^d \le k^{d-1}(R_1^d+\dots+R_k^d)$. Thus

$$\mathbf{P}\left(\{Y_1,\dots,Y_k\} \not\subset (0,1)^d \text{ or } \sum_{i=1}^k |Y_i - Y_{i-1}| \ge d\right) \le \mathbf{P}\left(\sum_{i=1}^k R_i \ge 1\right) \le \mathbf{P}\left(\sum_{i=1}^k \lambda R_i^d \ge \frac{\lambda}{k^{d-1}}\right).$$

On the other hand, the random variables λR_i^d are i.i.d. Exp(1). Indeed, we have $\mathbf{P}(\lambda R_1^d < t) = \mathbf{P}(\omega((0,(\frac{t}{\lambda})^{\frac{1}{d}})^d) = 0) = \exp(-\lambda\frac{t}{\lambda}) = \exp(-t)$ for all $t \ge 0$. Using $\mathbb{E}[\mathrm{e}^{\frac{1}{2}(\lambda R_1^d - 2)}] = 2\mathrm{e}^{-1} < 1$, $\lambda k^{1-d} - 2k \ge \frac{\lambda}{2}k^{1-d}$, $\frac{\lambda}{4}k^{1-d} \ge (\frac{\lambda}{4})^{\frac{1}{d}}$ and the Markov inequality, we obtain

$$\mathbf{P}\left(\sum_{i=1}^k \lambda R_i^d \ge \lambda k^{1-d}\right) \le \mathbf{P}\left(\sum_{i=1}^k (\lambda R_i^d - 2) \ge \frac{\lambda}{2}k^{1-d}\right) \le \mathrm{e}^{-\frac{\lambda}{4}k^{1-d}}\mathbb{E}[\mathrm{e}^{\frac{1}{2}(\lambda R_1^d-2)}]^k < \mathrm{e}^{-\frac{\lambda}{4}k^{1-d}} \le \mathrm{e}^{-(\frac{\lambda}{4})^{1/d}}.$$

□

Proof of Lemma 2.11. Let $\delta > 0$. For $n, k \in \mathbb{N}$ let

$$\mathcal{Q}_n := \left[\tfrac{1}{n}, \infty\right) \times Q_{\frac{\delta}{d}}, \qquad \lambda_n := (\tfrac{\delta}{d})^d n^\alpha,$$

$$\mathcal{E}_{n,k} := \Big\{ \text{there exist distinct } (f_1, y_1), \dots, (f_k, y_k) \in \operatorname{supp} \Pi \cap \mathcal{Q}_n \text{ such that } D_0(y_1, \dots, y_k) < \delta \Big\}.$$

Then $\Pi(\mathcal{Q}_n)$ is Poisson distributed with parameter λ_n. For $n \in \mathbb{N}$ let $k_n := \lceil \lambda_n^{\frac{1}{d}} \rceil$. Then on the event $\mathcal{E}_{n,k_n}$ we have

$$\sum_{i=1}^{k_n} f_i \geq \frac{k_n}{n} \geq \frac{\delta}{d} n^{\frac{\alpha}{d} - 1},$$

which is larger than 2θ for sufficiently large n.

Hence, for such large n we have $\mathbf{P}(\Pi \text{ satisfies } (2.6)) \geq \mathbf{P}(\mathcal{E}_{n,k_n})$ and so it suffices to show

$$\mathbf{P}(\mathcal{E}_{n,k_n}) \xrightarrow{n\to\infty} 1. \tag{4.6}$$

Note that the projection of $\mathbb{1}_{\mathcal{Q}_n}\Pi$ onto $(0, \delta/d)^d$ is a PPP on $(0, \delta/d)^d$ with intensity $\lambda_n (\delta/d)^{-d}$ (because $\Pi(\mathcal{Q}_n)$ is Poisson distributed with parameter λ_n). Therefore, for a PPP ζ on $(0,1)^d$ with intensity λ_n, we have

$$\mathbf{P}(\mathcal{E}_{n,k_n}) = \mathbb{P}\Big(\exists \text{ distinct } Z_1, \dots, Z_{k_\delta} \in \zeta \colon D_0(Z_1, \dots, Z_{k_\delta}) < d\Big).$$

Therefore, by applying Lemma 4.6, we conclude (4.6). □

5. Proof of Theorem 1.1 and Theorem 1.4

In the present section we will prove Theorem 1.1 (and implicitly also Theorem 1.4 (b)) subject to Proposition 5.2 and Proposition 5.4 below, whose proofs are postponed to Sections 6 and 7, respectively. In some sense, Proposition 5.2 gives us the lower bound of Theorem 1.1 (b) whereas Proposition 5.4 gives us the corresponding upper bound, as well as Theorem 1.1 (b).

Our strategy is the following. We first need to 'compactify' the partition function, i.e., to show that the random walk in the partition function can be restricted to some large box with a diameter on the scale r_t. This is done in Proposition 5.4 (c) in the sense of a convergence in distribution. Furthermore, we derive upper (in Proposition 5.4 (a)) and lower (in Proposition 5.2) bounds for the compactified partition function that lead to the right limit, the variational formula Ξ. Finally we need to upper bound the compactified partition function with W_t outside a neighbourhood of the maximizer against something that has a strictly smaller exponential rate. Here the stability of the variational formula from Theorem 2.8(c) will be crucial.

The upper and lower bounds for the compactified partition function are proved even in the almost-sure sense with respect to ξ, using the Skorohod embedding. That is, we do not work with a fixed trajectory $t \mapsto Z_t^{\xi,\beta}$ for a given realization of ξ, but with a sequence of realizations that are constructed jointly on one probability space. For this, we fix a sequence of times $(t_n)_{n\in\mathbb{N}}$. Since this construction is used several times in the paper, we state it in the following remark.

Remark 5.1 (Skorohod embedding)**.** Let $(t_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence in $(0,\infty)$ such that $t_n \to \infty$. By Skorohod's representation theorem (see, e.g., (Bil99, Theorem 1.6.7)) and Lemma 2.4 we can define on the same probability space a sequence $(\mathbf{\Pi}_n)_{n\in\mathbb{N}}$ of point processes and a Poisson point process $\mathbf{\Pi}$ on $(0,\infty) \times \mathbb{R}^d$ of intensity $\alpha f^{-1-\alpha}\,\mathrm{d}f \otimes \mathrm{d}y$ such that $\mathbf{\Pi}_n$ is the same in distribution as Π_{t_n} (see (1.9)) for every $n \in \mathbb{N}$, and trivially $\mathbf{\Pi}$ is the same in distribution as Π, and

$$\mathbf{\Pi}_n \to \mathbf{\Pi} \quad \text{almost surely in } \mathcal{M}_{\mathrm{p}}^{\circ}.$$

Without loss of generality we may assume that

$$\boldsymbol{\Pi}_n = \sum_{z\in\mathbb{Z}^d} \delta_{\left(\frac{\xi_n(z)}{r_{t_n}^{d/\alpha}}, \frac{z}{r_{t_n}}\right)}, \tag{5.1}$$

for some random variables $\xi_n(z)$ which are the same in distribution as $\xi(z)$, for any $n \in \mathbb{N}$ and $z \in \mathbb{Z}^d$. Fix a metric $\mathfrak{d}$ on $\mathcal{W}$ that is compatible with the vague topology and write $B(\nu, \delta) = \{\mu \in \mathcal{W} : \mathfrak{d}(\nu, \mu) < \delta\}$ for $\nu \in \mathcal{W}$ and $\delta > 0$. We introduce the following notation:

$$\begin{aligned} &\boldsymbol{r}_n = r_{t_n}, \qquad \boldsymbol{\gamma}_n = \boldsymbol{r}_n \log t_n, \\ &\boldsymbol{Z}_n = Z_{t_n}^{\xi_n,\beta_{t_n}}, \qquad \boldsymbol{H}_n(X) = H_{t_n}^{\xi_n,\beta_{t_n}}(X), \qquad \boldsymbol{\mathbb{P}}_n = \mathbb{P}_{t_n}^{(\xi_n)}, \qquad \boldsymbol{W}_n = W_{t_n}^{\xi_n,X}, \end{aligned} \tag{5.2}$$

and for $R, \delta > 0$ and $n \in \mathbb{N}$

$$\boldsymbol{Z}_n^{R,-} = \mathbb{E}\Big[\mathrm{e}^{\boldsymbol{H}_n(X)}\, \mathbb{1}\Big\{\max_{s\in[0,t_n]} |X_s| \le R\boldsymbol{r}_n\Big\}\Big] \qquad \text{and} \qquad \boldsymbol{Z}_n^{R,+} = \boldsymbol{Z}_n - \boldsymbol{Z}_n^{R,-}, \tag{5.3}$$

$$\boldsymbol{Z}_n^{R,-,\delta} = \mathbb{E}\Big[\mathrm{e}^{\boldsymbol{H}_n(X)}\, \mathbb{1}\{\mathfrak{d}(\boldsymbol{W}_n, \mu^*) \ge \delta\}\, \mathbb{1}\Big\{\max_{s\in[0,t_n]} |X_s| \le R\boldsymbol{r}_n\Big\}\Big]. \tag{5.4}$$

Recall (1.18). For a good point measure $\mathcal{P}$ (see Definition 2.7) on $(0, \infty) \times \mathbb{R}^d$ and $\nu^* \in \mathfrak{F}(\mathcal{P})$ the unique maximizer of $\Psi_{\mathcal{P}}$ (the existence is shown in Theorem 2.8) so that $\Xi(\mathcal{P}) = \Psi_{\mathcal{P}}(\nu^*)$, we define

$$\Xi^\delta(\mathcal{P}) = \sup_{\nu\in\mathcal{W}:\mathfrak{d}(\nu,\nu^*)\ge\delta} \Psi_{\mathcal{P}}(\nu). \tag{5.5}$$

For the probability measure on the space where the $\boldsymbol{\Pi}$ and $\boldsymbol{\Pi}_n$'s live, we make abuse of notation and write $\mathbf{P}$ and assume this does not lead to confusion. $\diamond$

We can now formulate the lower bound for the partition function.

Proposition 5.2 (Lower bound)**.** *Fix $\alpha \in (d, \infty)$. Then, with $\mathbf{P}$-probability 1,*

$$\liminf_{n\to\infty} \frac{1}{\boldsymbol{\gamma}_n} \log \boldsymbol{Z}_n \ge \Xi(\boldsymbol{\Pi}). \tag{5.6}$$

The proof of Proposition 5.2 is given in Section 6. Now we formulate the appropriate upper bounds for both assertions of Theorem 1.1. What will be crucial for the proof is the following observation:

Lemma 5.3. *Let $\alpha \in (2d, \infty)$. There exists a random variable μ^* with values in $\mathfrak{F}(\boldsymbol{\Pi})$, such that $\mathbf{P}$-almost surely μ^* is the unique element of $\mathcal{W}$ such that*

$$\Psi_{\boldsymbol{\Pi}}(\mu^*) = \Xi(\boldsymbol{\Pi}). \tag{5.7}$$

Moreover, almost surely

$$\Xi^\delta(\boldsymbol{\Pi}) < \Xi(\boldsymbol{\Pi}). \tag{5.8}$$

Proof. Because $\alpha > 2d$, $\boldsymbol{\Pi}$ is almost surely good by Lemma 2.9. Therefore by Theorem 2.8 (a) such μ^* exists (that it is random in the sense that it is a measurable function on the probability space, is not completely trivial; see Appendix B) and is unique almost surely and by Theorem 2.8 (c), (5.8) holds. □

Proposition 5.4 (Upper bounds)**.** *Fix $\alpha \in (d, \infty)$.*

(a) *Upper bound for compactified $\boldsymbol{Z}_n$:*

$$\mathbf{P}\Big[\limsup_{R\to\infty}\limsup_{n\to\infty} \frac{1}{\boldsymbol{\gamma}_n} \log \boldsymbol{Z}_n^{R,-} \le \Xi(\boldsymbol{\Pi})\Big] = 1. \tag{5.9}$$

(b) *Upper bound for compactified $\boldsymbol{Z}_n$ away from maximizer: Assume that $\alpha > 2d$. Then*

$$\mathbf{P}\Big[\limsup_{R\to\infty}\limsup_{n\to\infty} \frac{1}{\boldsymbol{\gamma}_n} \log \boldsymbol{Z}_n^{R,-,\delta} \le \Xi^\delta(\boldsymbol{\Pi})\Big] = 1, \qquad \delta > 0 \tag{5.10}$$

(c) *Compactification:*

$$\lim_{R\to\infty}\liminf_{n\to\infty}\ \mathbf{P}\Big[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n^{R,+}\le -A\Big]=1,\qquad A>0. \tag{5.11}$$

The proof of Proposition 5.4 is given in Section 7. We extract the following lemma from Proposition 5.2 and Proposition 5.4 (c) which will be used for the proofs of Theorem 1.1 (a) and (b).

Lemma 5.5. *For all $\varepsilon,\eta>0$ there exist an $R>0$ and an $N\in\mathbb{N}$ such that for all $n\ge N$*

$$\mathbf{P}\left[\frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}<\varepsilon\right]\ge 1-\eta.$$

Proof. First, we bound $\boldsymbol{Z}_n^{R,-}$ from below by restricting the expectation to the trajectory that remains at the origin up to time t_n and obtain (because by for example (1.13), for $Y_s=0$ for all $s\in[0,t_n]$, $\boldsymbol{H}_n(Y)\ge -\theta\boldsymbol{\gamma}_n$) Indeed, regarding (1.3), $\ell_{t_n}^{(Y)}=t_n$ and thus

$$\boldsymbol{H}_n(Y)=\sum_{z\in\mathbb{Z}^d}\xi(z)\ell_{t_n}^{(Y)}(z)-\beta_{t_n}\sum_{z\in\mathbb{Z}^d}\ell_{t_n}^{(Y)}(z)^2\ge -\beta_{t_n}t_n^2=\theta\boldsymbol{\gamma}_n,$$

where the latter equality can be found in (1.12).

$$\boldsymbol{Z}_n^{R,-}\ge \mathbb{P}(\ell_{t_n}(0)=t_n)\mathrm{e}^{-\theta\boldsymbol{\gamma}_n}=\mathrm{e}^{-2dt_n-\theta\boldsymbol{\gamma}_n},\qquad n\in\mathbb{N},R>0. \tag{5.12}$$

Because $\boldsymbol{\gamma}_n=\boldsymbol{r}_n\log t_n=t_n^{1+q}(\log t_n)^{-q}$ (see (1.8)), we have $\frac{t_n}{\boldsymbol{\gamma}_n}\to 0$ as $n\to\infty$. Let $\varepsilon,\eta>0$. Then, by also using (5.11) with $A=3\theta$ there exists an $R>0$ and an $N\in\mathbb{N}$ such that for all $n\ge N$,

$$\mathbf{P}\Big[\boldsymbol{Z}_n^{R,+}\le \mathrm{e}^{-3\theta\boldsymbol{\gamma}_n}\Big]\ge 1-\eta,\qquad \boldsymbol{Z}_n^{R,-}\ge \mathrm{e}^{-2\theta\boldsymbol{\gamma}_n},\qquad \mathrm{e}^{-\theta\boldsymbol{\gamma}_n}<\varepsilon$$

so that

$$\mathbf{P}\left[\frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}<\varepsilon\right]\ge \mathbf{P}\left[\frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}<\mathrm{e}^{-\theta\boldsymbol{\gamma}_n}\right]\ge 1-\eta.$$

□

Now we prove Theorem 1.1 and and Theorem 1.4 (a) subject to the above propositions and lemma.

Proof of Theorem 1.1 (a) and Theorem 1.4 (a). Let $\mathcal{D}$ be the continuity set of the distribution function for $\Xi(\boldsymbol{\Pi})$, i.e., the subset of $\mathbb{R}$ containing every continuity point of $x\mapsto \mathbf{P}(\Xi(\boldsymbol{\Pi})\le x)$. We will prove

$$\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n \overset{t\to\infty}{\Longrightarrow} \Xi(\boldsymbol{\Pi}), \tag{5.13}$$

by showing the following two inequalities:

$$\limsup_{n\to\infty}\mathbf{P}\left[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n\le h\right]\ \le\ \mathbf{P}\left(\Xi(\boldsymbol{\Pi})\le h\right),\qquad h\in\mathbb{R}, \tag{5.14}$$

$$\liminf_{n\to\infty}\mathbf{P}\left[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n\le h\right]\ \ge\ \mathbf{P}\left(\Xi(\boldsymbol{\Pi})\le h\right),\qquad h\in\mathcal{D}. \tag{5.15}$$

The proof of (5.15) is more involved. Therefore we focus on (5.15), because (5.14) follows in a similar fashion from Proposition 5.2. Indeed, by Proposition 5.2 it follows that for all $\eta>0$ there exists an $N\in\mathbb{N}$ such that for $\mathcal{B}_N=\{\forall n\ge N:\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n\ge \Xi(\Pi)\}$,

$$\mathbf{P}[\mathcal{B}_N]\ge 1-\eta.$$

For $\eta>0$ and N as such, we have

$$\sup_{n\ge N}\mathbf{P}\Big[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n\le h\Big]\le \sup_{n\ge N}\mathbf{P}\Big[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n\le h,\mathcal{B}_N\Big]+\mathbf{P}[\mathcal{B}_N^c]\le \mathbf{P}[\Xi(\boldsymbol{\Pi})\le h]+\eta.$$

Therefore $\limsup_{n\to\infty} \mathbf{P}\Big[\frac{1}{\gamma_n}\log \boldsymbol{Z}_n \le h, \mathcal{B}_N\Big] \le \mathbf{P}\Big[\Xi(\boldsymbol{\Pi}) \le h\Big] + \eta$ for any $\eta > 0$. Observe that for any $R > 0$, because $\boldsymbol{Z}_n = \boldsymbol{Z}_n^{R,-} + \boldsymbol{Z}_n^{R,+}$ (see (5.3)),

$$\frac{1}{\gamma_n}\log \boldsymbol{Z}_n = \frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} + \frac{1}{\gamma_n}\log\Big(1 + \frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}\Big) \le \frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} + \frac{1}{\gamma_n}\frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}. \tag{5.16}$$

Pick an $\eta > 0$. Let $\mathcal{A}_{n,R} := \{\boldsymbol{Z}_n^{R,+} \le \boldsymbol{Z}_n^{R,-}\}$. By Lemma 5.5 there exist an $R > 0$ and an $N \in \mathbb{N}$ such that $\mathbf{P}(\mathcal{A}_{n,R}) \ge 1 - \eta$ for all $n \ge N$.

Fix $h \in \mathcal{D}$ and pick $\varepsilon > 0$. As $\frac{1}{\gamma_n} \le \varepsilon$ for large n, we have for sufficiently large n that

$$\begin{aligned}\mathbf{P}\left[\frac{1}{\gamma_n}\log \boldsymbol{Z}_n \le h\right] &\ge \mathbf{P}\left[\Big\{\frac{1}{\gamma_n}\log \boldsymbol{Z}_n \le h\Big\} \cap \mathcal{A}_{n,R}\right] \\ &\ge \mathbf{P}\left[\Big\{\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} \le h - \varepsilon\Big\} \cap \mathcal{A}_{n,R}\right] \\ &\ge \mathbf{P}\left[\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} \le h - \varepsilon\right] - \mathbf{P}(\mathcal{A}_{n,R}^{\mathrm{c}}) \\ &\ge \mathbf{P}\left[\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} \le h - \varepsilon\right] - \eta.\end{aligned} \tag{5.17}$$

At this stage, we use Proposition 5.4 (a), i.e., we use (5.9), to infer that (possibly by choosing R larger)

$$\mathbf{P}\Big(\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-} \le \Xi(\boldsymbol{\Pi}) + \varepsilon\Big) = 1, \qquad \text{for large } n. \tag{5.18}$$

Indeed, if we have $\mathbf{P}(\limsup_{n\to\infty} Y_n \le A) = 1$, then $\mathbf{P}(\bigcap_{m\in\mathbb{N}}\bigcup_{N\in\mathbb{N}}\bigcap_{n\ge N}\{Y_n \le A + \frac{1}{m}\}) = 1$, so that for all $\varepsilon > 0$ one has $\mathbf{P}(\bigcup_{N\in\mathbb{N}}\bigcap_{n\ge N}\{Y_n \le A+\varepsilon\}) = 1$, i.e., by monotonicity, $\lim_{N\to\infty}\mathbf{P}(\{Y_n \le A + \varepsilon\}) \ge \lim_{N\to\infty}\mathbf{P}(\bigcap_{n\ge N}\{Y_n \le A+\varepsilon\}) = \mathbf{P}(\bigcup_{N\in\mathbb{N}}\bigcap_{n\ge N}\{Y_n \le A+\varepsilon\}) = 1$. Combining (5.17) and (5.18) gives

$$\liminf_{n\to\infty}\mathbf{P}\left[\frac{1}{\gamma_n}\log \boldsymbol{Z}_n \le h\right] \ge \mathbf{P}\left(\Xi(\boldsymbol{\Pi}) \le h - 2\varepsilon\right) - \eta, \qquad \text{for any } \eta, \varepsilon > 0. \tag{5.19}$$

By letting η and ε converge to zero and by using the continuity at h of $x \mapsto \mathbf{P}\left(\Xi(\boldsymbol{\Pi}) \le x\right)$, this completes the proof of (5.15).

From (5.13) we deduce that $\frac{1}{r_t \log t}\log Z_t^{(\xi)} \implies \Xi(\Pi)$ as we obtained the convergence along diverging sequences of $(t_n)_{n\in\mathbb{N}}$ in $(0,\infty)$. We are left to show that $\Xi(\boldsymbol{\Pi})$ is almost surely in $[0,\infty)$ (which implies also Theorem 1.4 (a)). That $\Xi(\boldsymbol{\Pi}) \ge 0$ follows directly by the fact that it is larger than $\Psi_{\boldsymbol{\Pi}}$ evaluated in the zero measure. That it is finite follows by the fact that $H_t^{(\xi,\beta_t)} \le H_t^{(\xi,0)}$ and thus $Z_t^{(\xi)} = Z_t^{(\xi,\beta_t)} \le Z_t^{(\xi,0)}$. As the limit of $\frac{1}{r_t \log t}\log Z_t^{(\xi,0)}$ is almost surely finite (by e.g. (1.6) in (KLMS09)), so is the limit of $\frac{1}{r_t\log t}\log Z_t^{(\xi)}$, which is $\Xi(\Pi)$. □

Proof of Theorem 1.1 (b). Let $\delta > 0$. First we show that $\mathbb{P}_n[\mathfrak{d}(\boldsymbol{W}_n,\mu^*) > \delta]$ converges to zero in $\mathbf{P}$-probability, i.e., for all $\kappa > 0$,

$$\mathbf{P}\Big[\mathbb{P}_n[\mathfrak{d}(\boldsymbol{W}_n,\mu^*) > \delta] > \kappa\Big] \xrightarrow{t\to\infty} 0. \tag{5.20}$$

Observe that for any $R > 0$

$$\mathbb{P}_n[\mathfrak{d}(\boldsymbol{W}_n,\mu^*) > \delta] \le \frac{\boldsymbol{Z}_n^{R,-,\delta} + \boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n} \le \frac{\boldsymbol{Z}_n^{R,-,\delta}}{\boldsymbol{Z}_n} + \frac{\boldsymbol{Z}_n^{R,+}}{\boldsymbol{Z}_n^{R,-}}.$$

Let $\kappa > 0$. By Lemma 5.5 it is sufficient to show that there exists an $R > 0$ such that $\mathbf{P}[\frac{\boldsymbol{Z}_n^{R,-,\delta}}{\boldsymbol{Z}_n} \le \kappa] \xrightarrow{n\to\infty} 1$. Let $\varepsilon > 0$. By (5.8) there exists an $m \in \mathbb{N}$ such that $\mathrm{e}^{-\frac{1}{3m}} < \kappa$, and

$$\mathbf{P}[\mathcal{B}_{m,\delta}] \ge 1 - \varepsilon, \qquad \text{where } \mathcal{B}_{m,\delta} = \Big\{\Xi^\delta(\boldsymbol{\Pi}) - \Xi(\boldsymbol{\Pi}) < -\frac{1}{m}\Big\}.$$

By Proposition 5.4 (b) and Proposition 5.2 there exists an $R > 0$ and an $N \in \mathbb{N}$ such that for all $n \ge N$

$$\mathbf{P}\Big[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n^{R,-,\delta} \le \Xi^\delta(\boldsymbol{\Pi}) + \frac{1}{3m}\Big] \ge 1 - \varepsilon, \qquad \mathbf{P}\Big[\frac{1}{\boldsymbol{\gamma}_n}\log \boldsymbol{Z}_n^{R} \ge \Xi(\boldsymbol{\Pi}) - \frac{1}{3m}\Big] \ge 1 - \varepsilon,$$

so that

$$1 - 3\varepsilon \le \mathbf{P}\Big[\frac{\boldsymbol{Z}_n^{R,-,\delta}}{\boldsymbol{Z}_n} \le \exp\Big(-\frac{1}{3m}\Big)\Big] \le \mathbf{P}\Big[\frac{\boldsymbol{Z}_n^{R,-,\delta}}{\boldsymbol{Z}_n} \le \kappa\Big].$$

From this we conclude (5.20). From the convergence in probability we deduce the existence of a strictly increasing $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\mathbb{P}_{\varphi(n)}[\mathfrak{d}(\boldsymbol{W}_{\varphi(n)}, \mu^*) > \delta] \to 0$ $\mathbf{P}$-almost surely. This implies $\mathbf{P}$-almost surely that $\boldsymbol{W}_{\varphi(n)} \Longrightarrow \mu^*$ in $\mathcal{W}$, more precisely, $\mathbb{E}_{\varphi(n)}[g(\boldsymbol{W}_{\varphi(n)})] \to g(\mu^*)$ for any $g \in C_{\mathrm{b}}(\mathcal{W})$. Therefore,

$$\mathbf{E}\big[\mathbb{E}_{\varphi(n)}[g(\boldsymbol{W}_{\varphi(n)})]\big] \to \mathbf{E}[g(\mu^*)], \qquad g \in C_{\mathrm{b}}(\mathcal{W}).$$

Therefore, as for each sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to \infty$ there exists a strictly increasing $\varphi : \mathbb{N} \to \mathbb{N}$ such that

$$\mathbf{E}\big[\mathrm{e}^{(\xi)}_{t_{\varphi(n)}}[g(W_{t_{\varphi(n)}})]\big] \to \mathbf{E}[g(\mu^*)], \qquad g \in C_{\mathrm{b}}(\mathcal{W}),$$

it follows that for any sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to \infty$

$$\mathbf{E}\big[\mathrm{e}^{(\xi)}_{t_n}[g(W_{t_n})]\big] \to \mathbf{E}[g(\mu^*)], \qquad g \in C_{\mathrm{b}}(\mathcal{W}),$$

and therefore (1.17). □

Remark 5.6. The proof of the more general convergence in distribution $\mathcal{L}_t^{(\xi)} \overset{t\to\infty}{\Longrightarrow} \delta_{\mu^*}$ as in (1.28) of Remark 1.6 can be deduced from the first part of the proof of Theorem 1.1 (b) as follows.

First we observe that by Portmanteau's theorem, for probability measures $\rho_1, \rho_2, \dots$ on $\mathcal{W}$ and $\mu \in \mathcal{W}$, one has $\rho_n \to \delta_\mu$ weakly if and only if for all closed sets $C \subset \mathcal{W}$ one has $\limsup_{n\to\infty} \rho_n(C) \le \delta_\mu(C)$, which in turns holds if and only if $\lim_{n\to\infty} \rho_n(\mathcal{C}^\delta(\mu)) \to 0$ for all $\delta > 0$, where $\mathcal{C}^\delta(\mu) = \{\nu \in \mathcal{W} : \mathfrak{d}(\nu, \mu) > \delta\}$.

Let $\delta > 0$. Let $\mathcal{C}^\delta = \{\nu \in \mathcal{W} : \mathfrak{d}(\nu, \mu^*) \ge \delta\}$, i.e., $\mathcal{C}^\delta = B(\mu^*, \delta)^c$. We write $\boldsymbol{\mathcal{L}}_n = \mathcal{L}_{t_n}^{(\xi_n)}$. From the fact that

$$\boldsymbol{\mathcal{L}}_n(\mathcal{C}^\delta) = \mathbb{P}_n[\mathfrak{d}(\boldsymbol{W}_n, \mu^*) > \delta],$$

we deduce from (5.20) that $\boldsymbol{\mathcal{L}}_n(\mathcal{C}^\delta)$ converges to zero in $\mathbf{P}$-probability, i.e., for all $\kappa > 0$, $\mathbf{P}[\boldsymbol{\mathcal{L}}_n(\mathcal{C}^\delta) > \kappa] \xrightarrow{n\to\infty} 0$. From this we infer the existence of a strictly increasing $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\boldsymbol{\mathcal{L}}_{\varphi(n)}(\mathcal{C}^\delta) \to 0$ almost surely.

Therefore, by the above observation, it follows that $\boldsymbol{\mathcal{L}}_{\varphi(n)} \to \delta_{\mu^*}$ almost surely, and thus $\mathcal{L}^{(\xi)}_{t_{\varphi(n)}} \Longrightarrow \delta_{\mu^*}$. As for each sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to \infty$ there exists a strictly increasing $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\mathcal{L}^{(\xi)}_{t_{\varphi(n)}} \Longrightarrow \delta_{\mu^*}$, it follows that $\mathcal{L}^{(\xi)}_{t_n} \Longrightarrow \delta_{\mu^*}$ for any sequence $(t_n)_{n\in\mathbb{N}}$ with $t_n \to \infty$, implying (1.28).

6. Lower bound: proof of Proposition 5.2

Our strategy follows the heuristics described in Section 1.5.

Recall the setting introduced at the beginning of Section 5, in particular Remark 5.1 on the Skorohod embedding and the notations in (5.2). Let $\mu \in \mathfrak{F}_1(\Pi)$, i.e., $\mu \in \mathcal{W}$, $\mu \ll \Pi$ and μ be a probability measure (in case $\alpha \in (2d, \infty)$ one may take $\mu = \mu^*$ as in Lemma 5.3). By Lemma 2.6 it is sufficient to show that, with $\mathbf{P}$-probability 1,

$$\liminf_{n\to\infty} \frac{1}{\boldsymbol{\gamma}_n} \log \boldsymbol{Z}_n \ge \Psi_{\boldsymbol{\Pi}}(\mu). \tag{6.1}$$

Our approach to do this is to choose a specific path event $\boldsymbol{\mathcal{A}}_n$ and use the trivial estimate

$$\boldsymbol{Z}_n \geq \mathbb{E}[\mathrm{e}^{\boldsymbol{H}_n(X)} \mathbb{1}_{\boldsymbol{\mathcal{A}}_n}]. \tag{6.2}$$

We describe the event $\boldsymbol{\mathcal{A}}_n$ in Section 6.1, but first give an idea here after introducing the following objects. Since μ is in $\mathfrak{F}_1(\boldsymbol{\Pi})$, there exist ("$\mathbf{P}$-"random) $k \in \mathbb{N}$ and $(f_1, y_1), \dots, (f_k, y_k) \in \mathrm{supp}(\boldsymbol{\Pi})$ and $w_1, \dots, w_k \in (0,1]$ with $\sum_{i=1}^k w_i = 1$ such that

$$\mu = \sum_{i=1}^k w_i \, \delta_{(f_i, y_i)}.$$

We may assume that the order of the $(f_1, y_1), \dots, (f_k, y_k)$ is such that the minimal distance between the $y_1, \dots, y_k$ points is given by $\sum_{i=1}^k |y_i - y_{i-1}|$, where here and in the following we take $y_0 = 0$. Hence,

$$\Xi(\boldsymbol{\Pi}) = \Psi_{\boldsymbol{\Pi}}(\mu) = \sum_{i=1}^k \Big(f_i w_i - \theta(w_i)^2 \Big) - q \sum_{i=1}^k |y_i - y_{i-1}|.$$

Because $\boldsymbol{\Pi}_n \to \boldsymbol{\Pi}$ in $\mathcal{M}_{\mathrm{p}}^\circ$ almost surely, there exists an $N \in \mathbb{N}$ such that for every $n \geq N$ there exist distinct $(f_1^n, y_1^n), \dots, (f_k^n, y_k^n) \in \mathrm{supp}\, \boldsymbol{\Pi}_n$ such that almost surely

$$(f_i^n, y_i^n) \xrightarrow{n\to\infty} (f_i, y_i), \qquad i \in \{1, \dots, k\}. \tag{6.3}$$

Observe that by (5.1),

$$f_i^n = \frac{\xi_n(y_i^n)}{\boldsymbol{r}_n^{d/\alpha}}, \qquad i \in \{1, \dots, k\}, n \geq N. \tag{6.4}$$

We will define the event $\boldsymbol{\mathcal{A}}_n$, such that on this event the path visits the sites $\boldsymbol{r}_n y_1^n, \dots, \boldsymbol{r}_n y_k^n$ in this order, staying $\approx w_i t_n$ time units in $\boldsymbol{r}_n y_i^n$ for any $i \in \{1, \dots, k\}$. We define $\boldsymbol{\mathcal{A}}_n$ and estimate its probability from below in Section 6.1. Then, we bound $\boldsymbol{H}_n(X)$ from below on $\boldsymbol{\mathcal{A}}_n$ in Section 6.2. Finally, in Section 6.3 we combine these bounds and apply them in the framework of the Skorokhod embedding defined above to finish the proof.

6.1 The path event

Let us introduce some useful notation involving paths that will be used to define the set $\boldsymbol{\mathcal{A}}_n$.

Definition 6.1. For $x \in \mathbb{Z}^d$ and $t \in [0, \infty)$ we define the entry time at x after time t, $\tau_x(t)$, and the exit time from x after time t, $\sigma_x(t)$, by

$$\tau_x(t) = \inf\{s > t \colon X_s = x\}, \qquad \sigma_x(t) = \inf\{s > t \colon X_s \neq x\}.$$

Let $t \in (0, \infty)$,

$$\begin{gathered} \delta, s \in (0,1), \quad k \in \mathbb{N}, \quad y_0 := 0, \quad y_1, \dots, y_k \in \mathbb{Z}^d, \quad y = (y_1, \dots, y_k), \\ w_1, \dots, w_k \in [0,1] \text{ with } \sum_{i=1}^k w_i = 1 - s, \quad w = (w_1, \dots, w_k). \end{gathered} \tag{6.5}$$

We define $\mathcal{A}_{t,k}^{\delta,s}(y, w)$ to be the event where the random walk X walks from 0 to y_1 and then to y_2 etcetera. It takes at most $\frac{st}{k}$ time to reach y_1, then it spends at least $(1-\delta)tw_1$ and at most tw_1 time at y_1 before it jumps, then it spends at most $\frac{st}{k}$ time to reach y_2, waits at least $(1-\delta)tw_2$ and at most tw_2 time at y_2 before it jumps, etc. More precisely, first we define inductively the entry τ_y^i and exit times σ_y^i of the y_i, after the time that y_{i-1} and thus all of $0, y_1, \dots, y_{i-1}$ are visited

$$\begin{aligned} \boldsymbol{\tau}_y^0 := 0, \qquad \boldsymbol{\tau}_y^1 := \tau_{y_1}(0), \qquad \boldsymbol{\tau}_y^i &:= \tau_{y_i}(\boldsymbol{\tau}_y^{i-1}), \qquad i \in \{1, \dots, k\}, \\ \boldsymbol{\sigma}_y^i &:= \sigma_{y_i}(\boldsymbol{\tau}_y^i), \qquad i \in \{0, 1, \dots, k\}, \end{aligned}$$

so that (by definition $\boldsymbol{\tau}_y^0 = 0 \le \boldsymbol{\tau}_y^1 \le \boldsymbol{\sigma}_y^1 \le \cdots \boldsymbol{\tau}_y^k \le \boldsymbol{\sigma}_y^k$ and)

$$\mathcal{A}_{t,k}^{\delta,s}(y,w) = \bigcap_{i=1}^k \Big\{ \boldsymbol{\tau}_y^i - \boldsymbol{\sigma}_y^{i-1}\mathbb{1}_{i\ge 2} \le \frac{st}{k},\ \boldsymbol{\sigma}_y^i - \boldsymbol{\tau}_y^i \in [1-\delta,1]tw_i\Big\}.$$

Observe that for $i=1$ we have $\boldsymbol{\tau}_y^i = \boldsymbol{\tau}_y^i - \boldsymbol{\sigma}_y^{i-1}\mathbb{1}_{i\ge 2} \le \frac{st}{k}$, so that the waiting time at 0 plus the "walking time" to y_1 is less or equal to $\frac{st}{k}$. Furthermore, observe that y_k is reached before t, i.e., $\boldsymbol{\sigma}_y^k \le t$, because

$$\begin{aligned}\boldsymbol{\sigma}_y^k &= (\boldsymbol{\sigma}_y^k - \boldsymbol{\tau}_y^k) + (\boldsymbol{\tau}_y^k - \boldsymbol{\sigma}_y^{k-1}) + \cdots + (\boldsymbol{\sigma}_y^1 - \boldsymbol{\tau}_y^1) + (\boldsymbol{\tau}_y^1 - \boldsymbol{\sigma}_y^0) + (\boldsymbol{\sigma}_y^0 - \boldsymbol{\tau}_y^0)\\ &\le tw_k + \frac{st}{k} + tw_{k-1} + \frac{st}{k} + \cdots + tw_1 + \frac{st}{k} = t\Big(s + \sum_{i=1}^k w_i\Big) = t.\end{aligned}$$

Lemma 6.2. *For any $t\in(0,\infty)$, $\delta, s\in(0,1)$, $k\in\mathbb{N}$ and y and w as in (6.5)*

$$\mathbb{P}\Big(\mathcal{A}_{t,k}^{\delta,s}(y,w)\Big) \ge \prod_{i=1}^k \Big[\mathrm{Poi}_{\frac{2dst}{k}}\big(|y_i - y_{i-1}|\big)\mathrm{e}^{-2dtw_i(1-\delta)}[1-\mathrm{e}^{-2dt\delta w_i}]\Big(\frac{1}{2d}\Big)^{|y_i-y_{i-1}|}\Big]. \tag{6.6}$$

Proof. By independence we have

$$\mathbb{P}\Big(\mathcal{A}_{t,k}^{\delta,s}(y,w)\Big) = \prod_{i=1}^k \mathbb{P}\Big(0 \le \boldsymbol{\tau}_y^i - \boldsymbol{\sigma}_y^{i-1}\mathbb{1}_{i\ge 2} \le \frac{st}{k}\Big)\mathbb{P}\Big(\boldsymbol{\sigma}_y^i - \boldsymbol{\tau}_y^i \in [1-\delta,1]tw_i\}\Big).$$

By the strong Markov property and the fact that each jump occurs according to an $\mathrm{Exp}(2d)$ random variable (as we assumed our continuous time random walk to have generator Δ), we have

$$\begin{aligned}\boldsymbol{\sigma}_y^i - \boldsymbol{\tau}_y^i &= \sigma_{y_i}(\boldsymbol{\tau}_y^i) - \boldsymbol{\tau}_y^i = \sigma_{y_i}(\tau_{y_i}(\boldsymbol{\tau}_y^{i-1})) - \tau_{y_i}(\boldsymbol{\tau}_y^{i-1}) \overset{(\mathrm{d})}{=} \sigma_0(0) - \tau_0(0),\\ \boldsymbol{\tau}_y^i - \boldsymbol{\sigma}_y^{i-1}\mathbb{1}_{i\ge 2} &= \tau_{y_i}(\boldsymbol{\tau}_y^{i-1}) - \sigma_{y_{i-1}}(\boldsymbol{\tau}_y^{i-1})\mathbb{1}_{i\ge 2} \overset{(\mathrm{d})}{=} \tau_{y_i - y_{i-1}}(0) - \sigma_0(0)\mathbb{1}_{i\ge 2},\end{aligned}$$

and thus

$$\begin{aligned}\mathbb{P}\Big(\boldsymbol{\sigma}_y^i - \boldsymbol{\tau}_y^i \in [1-\delta,1]tw_i\}\Big) &= \mathbb{P}\Big(\sigma_0(0) - \tau_0(0) \in [1-\delta,1]tw_i\}\Big) = \mathrm{e}^{-2dtw_i(1-\delta)} - \mathrm{e}^{-2dtw_i},\\ \mathbb{P}\Big(0 \le \boldsymbol{\tau}_y^i - \boldsymbol{\sigma}_y^{i-1}\mathbb{1}_{i\ge 2} \le \frac{st}{k}\Big) &= \mathbb{P}\Big(0 \le \tau_{y_i-y_{i-1}}(0) - \sigma_0(0)\mathbb{1}_{i\ge 2} \le \frac{st}{k}\Big) \ge \mathbb{P}\Big(\tau_{y_i-y_{i-1}}(0) \le \frac{st}{k}\Big).\end{aligned}$$

To estimate the latter probability, we use the following estimate for $\rho\in(0,\infty)$ and $z\in\mathbb{Z}^d$, where $N(z)$ denotes the number of direct paths (i.e., of length $|z|$) from 0 to z:

$$\begin{aligned}\mathbb{P}\Big(\tau_z(0) \le \rho\Big) &\ge \mathbb{P}\Big(X \text{ makes } |z| \text{ jumps within } \rho \text{ time, from 0 to } z\Big)\\ &= \mathrm{Poi}_{2d\rho}(|z|)(2d)^{-|z|}N(z) \ge \mathrm{Poi}_{2d\rho}(|z|)(2d)^{-|z|}.\end{aligned}$$

□

6.2 Energetic lower bound

Now we derive a lower bound of $H_t^{(\xi)}$ on $\mathcal{A}_{t,k}^{\delta,s}(y,w)$.

Lemma 6.3 (Lower bound for $H_t^{(\xi)}$)**.** *Let $t\in(0,\infty)$, $\delta,s\in(0,1)$, $k\in\mathbb{N}$ and y and w as in (6.5). Then, on the event $\mathcal{A}_{t,k}^{\delta,s}(y,w)$,*

$$\frac{1}{r_t\log t}H_t^{(\xi)}(X) \ge (1-\delta)\sum_{i=1}^k \frac{\xi(y_i)}{r_t^{d/\alpha}}w_i - \theta\sum_{i=1}^k w_i^2 - (k+5)\theta(\delta+s). \tag{6.7}$$

Proof. We have (see also (1.13))

$$\frac{1}{r_t \log t} H_t^{(\xi)}(X) = \sum_{z\in\mathbb{Z}^d} \Big(\frac{\xi(z)}{r_t^{d/\alpha}} \frac{\ell_t(z)}{t} - \theta\big(\frac{\ell_t(z)}{t}\big)^2 \Big). \tag{6.8}$$

Using that $\xi \geq 0$ and the basic estimate $\sum_{z\in\mathbb{Z}^d} a_z^2 \leq \sum_{z\in A} a_z^2 + (\sum_{z\notin A} a_z)^2$, which can be used to show that the total normalized self-intersection local time (SILT) is not larger than the sum of the normalized SILTs in the $y_1, \dots, y_k$ plus the square of the remaining total local time, we obtain that

$$\frac{1}{r_t \log t} H_t^{(\xi)}(X) \geq \sum_{i=1}^k \Big(\frac{\xi(y_i)}{r_t^{d/\alpha}} \frac{\ell_t(y_i)}{t} - \theta\big(\frac{\ell_t(y_i)}{t}\big)^2 \Big) - \theta\Big(1 - \sum_{i=1}^k \frac{\ell_t(y_i)}{t}\Big)^2. \tag{6.9}$$

Observe that the local time at each y_i is at least the time the random walk waits before jumping away, i.e.,

$$\ell_t(y_i) \geq \boldsymbol{\sigma}_y^i - \boldsymbol{\tau}_y^i \geq (1-\delta) t w_i. \tag{6.10}$$

On $\mathcal{A}_{t,k}^{\delta,s}(y,w)$, in between the times $\boldsymbol{\sigma}_y^{i-1}$ and $\boldsymbol{\tau}_y^i$ the walker is allowed to visit sites y_j for $j \neq i$. Moreover, after $\sigma_k t$ time, each of the y_i may be revisited. We let $m_i \in [0,1]$ be such that $m_i t$ is the amount of time the path visits y_i after $\sigma_k t$, for all $i \in \{1, \dots, k\}$. In particular,

$$(1-\delta) w_i \leq \frac{\ell_t(y_i)}{t} \leq w_i + s + m_i, \qquad i \in \{1, \dots, k\}, \tag{6.11}$$

and consequently,

$$\Big(1 - \sum_{i=1}^k \frac{\ell_t(y_i)}{t}\Big)^2 \leq \Big(1 - (1-\delta)\sum_{i=1}^k w_i\Big)^2 \leq (s + \delta w)^2 \leq (s+\delta)^2. \tag{6.12}$$

Since σ_{y_k} is both bounded from below by $(1-\delta)tw = t(1-\delta)(1-s) \geq t - (s+\delta)t$, we infer $\sum_{i=1}^k m_i t \leq t - \sigma_{y_k} \leq (s+\delta)t$ and therefore deduce from the upper bound in (6.11) that

$$\begin{aligned}\sum_{i=1}^k \frac{\ell_t(y_i)}{t}^2 &\leq \sum_{i=1}^k (w_i+s)^2 + 2\sum_{i=1}^k m_i(w_i+s) + \sum_{i=1}^k m_i^2 \\ &\leq \sum_{i=1}^k w_i^2 + 2sw + ks^2 + 3\sum_{i=1}^k m_i \\ &\leq \sum_{i=1}^k w_i^2 + 2s + ks^2 + 3(s+\delta) \leq \sum_{i=1}^k w_i^2 + (k+5)(s+\delta).\end{aligned} \tag{6.13}$$

Substituting these bounds (6.12) and (**??**) in (6.9) leads to (6.7). □

6.3 Conclusion

We now prove Proposition 5.2, by proving that (6.1) holds $\mathbf{\Pi}$-almost surely.

Proof of Proposition 5.2. Recall the definition of the approximating sequence of vectors in (6.3). We will assume that $n \geq N$ (where N is as mentioned before (6.3)). We set (assuming N is large enough)

$$\delta_n = s_n = \frac{1}{\log t_n}, \quad \boldsymbol{\mathcal{A}}_n := \mathcal{A}_{t_n,k}^{\delta_n,s_n}\Big((\boldsymbol{r}_n y_1^n, \dots, \boldsymbol{r}_n y_k^n), (w_1 - \frac{s_n}{k}, \dots, w_k - \frac{s_n}{k})\Big).$$

By (6.2) we find a lower estimate for $\boldsymbol{Z}_n$ by estimating $\boldsymbol{H}_n(X)$ on $\boldsymbol{\mathcal{A}}_n$ from below and by estimating $\mathbb{P}(\boldsymbol{\mathcal{A}}_n)$ from below. Recalling (5.2) and using Lemma 6.3 we see that on $\boldsymbol{\mathcal{A}}_n$, by using (6.4),

$$\begin{aligned}\frac{1}{\boldsymbol{\gamma}_n}\boldsymbol{H}_n(X) &\geq (1-\delta_n)\sum_{i=1}^{k}\frac{\xi_n(y_i^n)}{\boldsymbol{r}_n^{d/\alpha}}\,w_i^n - \theta\sum_{i=1}^{k}(w_i^n)^2 - (k+5)\theta(\delta_n+s_n)\\ &\geq (1-\delta_n)\sum_{i=1}^{k} f_i^n\,w_i^n - \theta\sum_{i=1}^{k}(w_i^n)^2 - (k+5)\theta(\delta_n+s_n)\\ &\xrightarrow{n\to\infty}\sum_{i=1}^{k}(f_i\,w_i-\theta(w_i)^2).\end{aligned}\tag{6.14}$$

Due to the above limit, for (6.1) we are left to show

$$\liminf_{n\to\infty}\frac{1}{\boldsymbol{\gamma}_n}\log\mathbb{P}(\boldsymbol{\mathcal{A}}_n)\geq -qD_0(y_1,\dots,y_k).\tag{6.15}$$

With $y_0^n:=0$, put

$$\mathfrak{d}_i^n=\boldsymbol{r}_n|y_i^n-y_{i-1}^n|,\qquad i\in\{1,\dots,k\}.$$

From Lemma 6.2 we obtain

$$\mathbb{P}(\boldsymbol{\mathcal{A}}_n)\geq\prod_{i=1}^{k}\Big[\mathrm{Poi}_{\frac{2ds_nt_n}{k}}(\mathfrak{d}_i^n)\mathrm{e}^{-2dt_nw_i^n(1-\delta_n)}[1-\mathrm{e}^{-2dt_n\delta_nw_i^n}]\Big(\frac{1}{2d}\Big)^{\mathfrak{d}_i^n}\Big].\tag{6.16}$$

Since for n large enough $w_i^n=w_i-\frac{s_n}{k}=w_i-\frac{1}{k\log t_n}$ is bounded away from 0 for all $i\in\{1,\dots,k\}$, since $\lim_{n\to\infty}t_n\delta_n=\infty$ and since $\lim_{x\to\infty}\mathrm{e}^{x}(1-\mathrm{e}^{-x})=\infty$, we have that for n large enough - by using that $j^j\geq j!$ for $j\in\mathbb{N}$ and that $\sum_{i=1}^{k}(w_i^n+\frac{s_n}{k})\leq 1$,

$$\begin{aligned}\mathbb{P}(\boldsymbol{\mathcal{A}}_n)&\geq\prod_{i=1}^{k}\Big[\mathrm{Poi}_{\frac{2ds_nt_n}{k}}(\mathfrak{d}_i^n)\mathrm{e}^{-2dt_nw_i^n}\Big(\frac{1}{2d}\Big)^{\mathfrak{d}_i^n}\Big]\\ &=\prod_{i=1}^{k}\Big[\frac{[\frac{2ds_nt_n}{k}]^{\mathfrak{d}_i^n}\mathrm{e}^{-\frac{2ds_nt_n}{k}}}{\mathfrak{d}_i^n!}\mathrm{e}^{-2dt_nw_i^n}\Big(\frac{1}{2d}\Big)^{\mathfrak{d}_i^n}\Big]\\ &\geq\prod_{i=1}^{k}\Big(\Big[\frac{\frac{2ds_nt_n}{k}}{\mathfrak{d}_i^n}\Big]^{\mathfrak{d}_i^n}\mathrm{e}^{-2dt_n(w_i^n+\frac{s_n}{k})}\Big(\frac{1}{2d}\Big)^{\mathfrak{d}_i^n}\Big)\geq\mathrm{e}^{-2dt_n}\prod_{i=1}^{k}\Big[\frac{s_nt_n}{k\mathfrak{d}_i^n}\Big]^{\mathfrak{d}_i^n}.\end{aligned}$$

Clearly, as $\boldsymbol{\gamma}_n=\boldsymbol{r}_n\log t_n=t_n^{1+q}(\log t_n)^q$ (see (1.8)) we have $\boldsymbol{\gamma}_n^{-1}\log(\mathrm{e}^{-2dt_n})\to 0$. Therefore, because $\frac{\mathfrak{d}_i^n}{\boldsymbol{r}_n}=|y_i^n-y_{i-1}^n|\xrightarrow{n\to\infty}|y_i-y_{i-1}|$ for all $i\in\{1,\dots,k\}$ and $n\in\mathbb{N}$ and because $\boldsymbol{\gamma}_n=\boldsymbol{r}_n\log t_n$ and

$\log \frac{\boldsymbol{r}_n}{t_n s_n} = \log r_{t_n} - \log t_n - \log s_n = (1+q)\log t_n - (1+q)\log\log t_n - \log t_n + \log\log t_n = q\log t_n - q\log\log t_n,$

$$
\begin{aligned}
\liminf_{n\to\infty} \frac{1}{\gamma_n} \log \mathbb{P}(\boldsymbol{\mathcal{A}}_n) &\geq \liminf_{n\to\infty} \frac{1}{\gamma_n} \log\Big(\prod_{i=1}^k \Big[\frac{s_n t_n}{k\mathfrak{d}_i^n}\Big]^{\mathfrak{d}_i^n}\Big) \\
&\geq \liminf_{n\to\infty} \frac{1}{\gamma_n}\Big(-\sum_{i=1}^k \Big[\mathfrak{d}_i^n \log \frac{k\mathfrak{d}_i^n}{s_n} + \mathfrak{d}_i^n \log t_n\Big]\Big) \\
&\geq -\limsup_{n\to\infty} \sum_{i=1}^k \frac{\mathfrak{d}_i^n}{\boldsymbol{r}_n}\frac{\boldsymbol{r}_n}{\gamma_n}\Big(\log\frac{\mathfrak{d}_i^n}{\boldsymbol{r}_n} + \log\frac{k\boldsymbol{r}_n}{t_n s_n}\Big) \\
&\geq -\sum_{i=1}^k |y_i - y_{i-1}| \limsup_{n\to\infty} \frac{1}{\log t_n}\Big(\log|y_i-y_{i-1}| + \log k - q\log\log t_n + q\log t_n\Big) \\
&\geq -q\sum_{i=1}^k |y_i - y_{i-1}| = -qD_0(y_1,\dots,y_k),
\end{aligned}
$$

□

7. Upper bounds: proof of Proposition 5.4

Part (c) of Proposition 5.4 is a kind of 'compactification', which we will prove in Section 7.1. Part (a) is proved in Section 7.2 (using the Skorohod embedding of Remark 5.1), and Part (b) in Section 7.3.

7.1 Compactification

In this section, we prove Proposition 5.4 (c). For this we actually do not need to consider a subsequence and the objects considered as in Remark 5.1. That is, we prove the following in this section, from which one directly derives Proposition 5.4 (c):

Proposition 7.1. *Let $\alpha \in (d,\infty)$ and $\theta \in (0,\infty)$. For any $A > 0$,*

$$\lim_{R\to\infty} \liminf_{t\to\infty} \mathbf{P}\Big[\frac{1}{r_t \log t} \log \mathbb{E}\Big[\mathrm{e}^{H_t^{(\xi)}(X)} \mathbb{1}\Big\{\max_{s\in[0,t]} |X_s| > Rr_t\Big\}\Big] \leq -A\Big] = 1. \tag{7.1}$$

Let us first state three auxiliary lemmas. In the first one we estimate the $\mathbb{P}$-probability that the random walk X takes too many jumps before time t (Lemma 7.2) and in the second one we estimate the $\mathbf{P}$-probability of the maximum of a modified version of the field ξ outside a big box centered at the origin (Lemma 7.3). The third one (Lemma 7.4) is a classical representation of the joint distribution of leading values in the ξ-field. The latter is used because it suffices to prove the estimate (7.1) but with $H_t^{(\xi)}$ replaced by the maximum over the ξ values in a box of radius M_t, where

$$M_t := \max_{s\in[0,t]} |X_s|, \qquad t \in [0,\infty). \tag{7.2}$$

Lemma 7.2. *For every $R \geq 1$ and all sufficiently large t,*

$$\mathbb{P}[M_t \geq Rr_t] \leq \exp\Big(-\frac{q}{2} Rr_t \log t\Big). \tag{7.3}$$

Proof. The number of jumps taken by X on the time interval $[0,t]$ is in distribution equal to a Poisson random variable Z with parameter $2dt$. Therefore, $\mathbb{P}[M_t \geq Rr_t] \leq P[Z \geq Rr_t]$. By using Stirling's inequality $n! \geq (\frac{n}{e})^n$, the crude bound $P[Z \geq n] \leq (2dt)^n/n! \leq (2dte/n)^n$ for $n \in \mathbb{N}$ implies (recall that $r_t = (\frac{t}{\log t})^{q+1}$)

$$P[Z \geq Rr_t] \leq \Big(\frac{4det}{Rr_t}\Big)^{Rr_t} \leq \exp\Big(-\frac{q}{2} Rr_t \log t\Big). \tag{7.4}$$

for any large t (we took an additional factor 2 to cover up that Rr_t might not be in $\mathbb{N}$). □

Lemma 7.3. *Let $A, c > 0$ and $\varepsilon \in (0,1)$. Then there exists an $R > 0$ such that for all $r \geq 1$*

$$\mathbf{P}\left(\max_{x \in \mathbb{Z}^d \setminus Q_{Rr}} \left(\frac{\xi(x)}{r^{d/\alpha}} - c\frac{|x|}{r}\right) \leq -A\right) \geq 1 - \varepsilon. \tag{7.5}$$

Proof. Recall that $Q_R = [-R, R]^d$. We write

$$\mathfrak{Q}_R = Q_R \cap \mathbb{Z}^d.$$

Pick a $C > 0$ such that for all $r \geq 1$

$$\#(\mathfrak{Q}_{(n+1)r} \setminus \mathfrak{Q}_{nr}) \leq Cn^{d-1}r^d, \qquad n \in \mathbb{N}.$$

Note that $1 - x \geq \mathrm{e}^{-2x}$ for $x \in [0, \frac{\log 2}{2}]$. It will be clear that $1 - x - \mathrm{e}^{-2x} \geq 0$ for $x = 0$. Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}x} 1 - x - \mathrm{e}^{-2x} = 2\mathrm{e}^{-2x} - 1,$$

which is ≥ 0 if $\mathrm{e}^{-2x} \geq \frac{1}{2}$, i.e., if $-2x \geq -\log 2$, i.e., $x \leq \frac{\log 2}{2}$. Pick $R \in \mathbb{N}$ large enough such that $CR > A$ and $\frac{1}{(cR-A)^\alpha r^d} < \frac{\log 2}{2}$. Then

$$\begin{aligned}
\mathbf{P}\Big(\max_{x \in \mathbb{Z}^d \setminus \mathfrak{Q}_{Rr}} \Big(\frac{\xi(x)}{r^{d/\alpha}} - c\frac{|x|}{r}\Big) \leq -A\Big) &= \prod_{n=R}^{\infty} \mathbf{P}\Big(\max_{x \in \mathfrak{Q}_{(n+1)r} \setminus \mathfrak{Q}_{nr}} \Big(\frac{\xi(x)}{r^{d/\alpha}} - c\frac{|x|}{r}\Big) \leq -A\Big) \\
&\geq \prod_{n=R}^{\infty} \mathbf{P}\Big(\frac{\xi(0)}{r^{d/\alpha}} - cn \leq -A\Big)^{Cn^{d-1}r^d} = \prod_{n=R}^{\infty} \Big(1 - \frac{1}{(cn-A)^\alpha r^d}\Big)^{Cn^{d-1}r^d} \\
&\geq \exp\Big(-2C \sum_{n=R}^{\infty} n^{d-1-\alpha} \Big(\frac{n}{cn-A}\Big)^\alpha\Big).
\end{aligned}$$

The exponential term on the right-hand side does not depend on r and converges to 1 as $R \to \infty$ as $d - 1 - \alpha < -1$ and as $\frac{n}{cn-A}$ is bounded from above. □

We will use the following classical representation of the distribution of the maximum over the ξ-values.

Lemma 7.4. *Let $\alpha \in (0, \infty)$, $n \in \mathbb{N}$ and $Z_1, \ldots, Z_n$ be i.i.d. random variables that are Pareto distributed with parameter α. Then the order statistics $Z_{1:n} \geq \cdots \geq Z_{n:n}$ of $Z_1, \ldots, Z_n$ is given by*

$$\Big(Z_{1:n}, \ldots, Z_{n:n}\Big) \overset{(\mathrm{d})}{=} \left(\Big(\frac{\Gamma_{n+1}}{\Gamma_1}\Big)^{1/\alpha}, \Big(\frac{\Gamma_{n+1}}{\Gamma_2}\Big)^{1/\alpha}, \ldots, \Big(\frac{\Gamma_{n+1}}{\Gamma_n}\Big)^{1/\alpha}\right), \tag{7.6}$$

where $\Gamma_i = E_1 + \cdots E_i$ and $(E_i)_{i \in \mathbb{N}}$ is a sequence of i.i.d. exponentially distributed random variables with parameter one.

Proof. It is a standard exercise, see for example (DVJ03, Exercise 2.1.2), to show that the order statistics of n i.i.d. uniformly distributed random variables in distribution equals $\frac{E_1}{E_1+\cdots+E_{n+1}}, \ldots, \frac{E_1+\cdots+E_n}{E_1+\cdots+E_{n+1}}$. By using that the Pareto distribution function $\mathbf{P}(Z \leq r) = r \mapsto 1 - r^{-\alpha}$ is the composition of the uniform distribution function with $\Phi(s) := (1-s)^{-\frac{1}{\alpha}}$ one finds the order statistics of Pareto distributions and in particular (7.6). Indeed, the function Φ is increasing. Therefore if $U_1, \ldots, U_n$ are i.i.d. uniformly distributed random variables and $U_{1:n} \geq \cdots \geq U_{n:n}$ their order statistics, then $\Phi(U_i) \overset{(\mathrm{d})}{=} Z_i$ and thus the order statistics of the Pareto variables $Z_1, \ldots, Z_n$, say $Z_{1:n} \geq \cdots \geq Z_{n:n}$, satisfies $(Z_{1:n}, \ldots, Z_{n:n}) = (\Phi(U_{1:n}), \ldots, \Phi(U_{n:n}))$, where

$$Z_{i:n} \overset{(\mathrm{d})}{=} \Phi(U_{i:n}) \overset{(\mathrm{d})}{=} \Big(1 - \frac{E_1 + \cdots + E_i}{E_1 + \cdots + E_{n+1}}\Big)^{-\frac{1}{\alpha}} = \Big(\frac{E_{i+1} + \cdots E_{n+1}}{E_1 + \cdots + E_{n+1}}\Big)^{-\frac{1}{\alpha}} \overset{(\mathrm{d})}{=} \Big(\frac{E_1 + \cdots + E_{n+1}}{E_1 + \cdots E_n}\Big)^{\frac{1}{\alpha}}.$$

□

Proof of Proposition 7.1. We write $\xi^*(B) = \max_{x\in B} \xi(x)$ here. We have (recall (1.1))

$$H_t^{(\xi)} \le t\xi^*(\mathfrak{Q}_{M_t}). \tag{7.7}$$

It is then sufficient to prove (7.1) with $t\xi^*(\mathfrak{Q}_{M_t})$ instead of $H_t^{(\xi)}$. Thus, we set

$$C_{t,R} := \mathbb{E}\Big[\mathrm{e}^{t\xi^*(\mathfrak{Q}_{Rr_t})}\mathbb{1}\big\{M_t > Rr_t\big\}\Big] \qquad \text{and} \qquad D_{t,R} := \mathbb{E}\Big[\mathrm{e}^{t\xi^*(\mathfrak{Q}_{M_t}\setminus\mathfrak{Q}_{Rr_t})}\mathbb{1}\big\{M_t > Rr_t\big\}\Big].$$

Because $\xi^*(\mathfrak{Q}_{M_t}) = \xi^*(\mathfrak{Q}_{M_t}\setminus \mathfrak{Q}_{Rr_t}) \vee \xi^*(\mathfrak{Q}_{Rr_t})$, the proof of (7.1) is complete once we show that for every $A > 0$

$$\lim_{R\to\infty} \liminf_{t\to\infty} \mathbf{P}\Big[\frac{1}{r_t\log t}\log C_{t,R} \le -A\Big] = 1, \tag{7.8}$$

and that

$$\lim_{R\to\infty} \liminf_{t\to\infty} \mathbf{P}\Big[\frac{1}{r_t\log t}\log D_{t,R} \le -A\Big] = 1. \tag{7.9}$$

Indeed, by writing

$$E_{t,R} := \mathbb{E}\Big[\mathrm{e}^{t\xi^*(\mathfrak{Q}_{M_t})}\mathbb{1}\big\{M_t > Rr_t\big\}\Big],$$

we have $E_{t,R} \le C_{t,R} + D_{t,R}$ and so

$$\{C_{t,R} \le \exp(-Ar_t\log t)\} \cap \{D_{t,R} \le \exp(-Ar_t\log t)\} \subset \{E_{t,R} \le 2\exp(-Ar_t\log t)\}.$$

Because $\mathfrak{Q}_{M_t}\setminus\mathfrak{Q}_{Rr_t}$ and $\mathfrak{Q}_{Rr_t}$ are disjoint, $\xi^*(\mathfrak{Q}_{M_t}\setminus\mathfrak{Q}_{Rr_t})$ and $\xi^*(\mathfrak{Q}_{Rr_t})$ are independent, and thus so are $C_{t,R}$ and $D_{t,R}$. Therefore

$$\mathbf{P}[E_{t,R} \le 2\exp(-Ar_t\log t)] \ge \mathbf{P}[C_{t,R} \le \exp(-Ar_t\log t)]\mathbf{P}[D_{t,R} \le \exp(-Ar_t\log t)]$$

It remains to observe that

$$\lim_{R\to\infty} \liminf_{t\to\infty} \mathbf{P}[E_{t,R} \le 2\exp(-Ar_t\log t)] = 1, \qquad A > 0,$$

if

$$\lim_{R\to\infty} \liminf_{t\to\infty} \mathbf{P}[E_{t,R} \le \exp(-Br_t\log t)] = 1, \qquad B > 0.$$

Let us begin with proving (7.8). Pick $A > 0$. Recall that $tr_t^{d/\alpha} = r_t\log t$. As $C_{t,R} = \mathrm{e}^{t\xi^*(\mathfrak{Q}_{Rr_t})}\mathbb{P}[M_t > Rr_t]$, from Lemma 7.2 we obtain a $T > 0$ such that for every $R \ge 1$ and $t \ge T$

$$\frac{1}{r_t\log t}\log C_{t,R} \le R\Big(\frac{1}{R^{1-d/\alpha}}\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} - \frac{q}{2}\Big). \tag{7.10}$$

Pick R_0 such that $\frac{qR_0}{4} > A$ and thus, for $t \ge T$

$$\mathbf{P}\Big[R\Big(\frac{1}{R^{1-d/\alpha}}\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} - \frac{q}{2}\Big) \le -\frac{qR_0}{4}\Big] \le \mathbf{P}\Big[\frac{1}{r_t\log t}\log C_{t,R} \le -A\Big]$$

and so for $R \ge R_0$,

$$\begin{aligned}\mathbf{P}\Big[R\Big(\frac{1}{R^{1-d/\alpha}}\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} - \frac{q}{2}\Big) \le -\frac{qR_0}{4}\Big] &\ge \mathbf{P}\Big[R\Big(\frac{1}{R^{1-d/\alpha}}\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} - \frac{q}{2}\Big) \le -\frac{qR}{4}\Big]\\ &= \mathbf{P}\Big[\frac{1}{R^{1-d/\alpha}}\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} \le \frac{q}{4}\Big],\end{aligned}$$

so that for every $t \ge T$ and $R \ge R_0$ one has

$$\mathbf{P}\Big[\frac{1}{r_t\log t}\log C_{t,R} \le -A\Big] \ge \mathbf{P}\Big[\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} \le \frac{q}{4}R^{1-d/\alpha}\Big]. \tag{7.11}$$

We apply Lemma 7.4 to $\xi^*(\mathfrak{Q}_{Rr_t})$ to obtain

$$\mathbf{P}\Big[\frac{\xi^*(\mathfrak{Q}_{Rr_t})}{(Rr_t)^{d/\alpha}} \le \frac{q}{4} R^{1-d/\alpha}\Big] = \mathbf{P}\Big[\frac{\Gamma_{\#(\mathfrak{Q}_{Rr_t})}}{(Rr_t)^d} \le \big(\tfrac{q}{4}\big)^{\alpha} R^{\alpha-d}\Gamma_1\Big].$$

By the weak law of large numbers $(Rr_t)^{-d}\Gamma_{\#(\mathfrak{Q}_{Rr_t})} \Longrightarrow 2^d$ as $t\to\infty$, so that

$$\liminf_{t\to\infty} \mathbf{P}\Big[\frac{\Gamma_{\#(\mathfrak{Q}_{Rr_t})}}{(Rr_t)^d} \le \big(\tfrac{q}{4}\big)^{\alpha} R^{\alpha-d}\Gamma_1\Big] = \mathbf{P}\Big[\Gamma_1 \ge 2^d\big(\tfrac{q}{4}\big)^{-\alpha} R^{d-\alpha}\Big].$$

Then, by letting $R\to\infty$ we conclude (7.8).

It remains to prove (7.9). Since $D_{t,R}$ is decreasing in R, it is sufficient to show that for every $\varepsilon\in(0,1)$, there exists a $R\in\mathbb{N}$ such that

$$\liminf_{t\to\infty}\ \mathbf{P}\Big[\frac{1}{r_t\log t}\log D_{t,R} \le -A\Big] \ge 1-\varepsilon. \tag{7.12}$$

We set

$$x_t^* := \arg\max\{\xi(x)\colon x\in\mathfrak{Q}_{M_t}\setminus\mathfrak{Q}_{Rr_t}\}. \tag{7.13}$$

Set

$$\mathcal{B}_{R,r} := \Big\{\max_{x\in\mathbb{Z}^d\setminus\mathfrak{Q}_{Rr}}\Big(\frac{\xi(x)}{r^{d/\alpha}} - \frac{q}{4}\frac{|x|}{r}\Big) \le -A\Big\}. \tag{7.14}$$

We pick $\varepsilon>0$ and we use Lemma 7.3 with $c=\frac{q}{4}$ to obtain the existence of an $R>1$ such that for every $r\ge1$,

$$\mathbf{P}(\mathcal{B}_{R,r}) \ge 1-\varepsilon. \tag{7.15}$$

As $x_t^*\notin\mathfrak{Q}_{Rr_t}$, on $\mathcal{B}_{R,r_t}$ we derive the following estimates

$$\begin{aligned}
D_{t,R} &\le \mathbb{E}\Big[\mathrm{e}^{t\xi(x_t^*)}\mathbb{1}\{M_t>Rr_t\}\Big]\\
&\le \mathbb{E}\Big[\exp\Big(tr_t^{d/\alpha}\frac{\xi(x_t^*)}{r_t^{d/\alpha}} - \frac{q}{4}(r_t\log t)\frac{|x_t^*|}{r_t}\Big)\ \exp\Big(\frac{q}{4}|x_t^*|\log t\Big)\ \mathbb{1}\{M_t>Rr_t\}\Big]\\
&\le \exp\Big(r_t\log t\max_{x\in\mathbb{Z}^d\setminus\mathfrak{Q}_{Rr_t}}\Big(\frac{\xi(x)}{r_t^{d/\alpha}} - \frac{q}{4}\frac{|x|}{r_t}\Big)\Big)\ \mathbb{E}\Big[\mathrm{e}^{\frac{q}{4}|x_t^*|\log t}\mathbb{1}\{M_t>Rr_t\}\Big]\\
&\le \mathrm{e}^{-Ar_t\log t}\ \ \mathbb{E}\Big[\mathrm{e}^{\frac{q}{4}M_t\log t}\,\mathbb{1}\{M_t>Rr_t\}\Big].
\end{aligned} \tag{7.16}$$

Furthermore, by Lemma 7.2, for large t we have $\mathbb{P}\big[M_t>jr_t\big]\le\exp(-\frac{q}{2}jr_t\log t)$ for all $j\in\mathbb{N}$ and so,

$$\begin{aligned}
\mathbb{E}\Big[\mathrm{e}^{\frac{q}{4}M_t\log t}\,\mathbb{1}\{M_t>Rr_t\}\Big] &\le \sum_{j=R}^{\infty}\mathbb{E}\Big[\mathrm{e}^{\frac{q}{4}M_t\log t}\,\mathbb{1}\{jr_t<M_t\le(j+1)r_t\}\Big]\\
&\le \sum_{j=R}^{\infty}\mathrm{e}^{\frac{q}{4}(j+1)r_t\log t}\,\mathbb{P}\big[M_t>jr_t\big]\\
&\le \mathrm{e}^{\frac{q}{4}r_t\log t}\sum_{j=R}^{\infty}\exp\Big(-\frac{q}{4}jr_t\log t\Big)\\
&= \frac{\exp\big(\frac{q}{4}(1-R)r_t\log t\big)}{1-\exp\big(-\frac{q}{4}r_t\log t\big)}.
\end{aligned}$$

As $R>1$, the latter converges to 0 as $t\to\infty$. Therefore, by combining this with (7.16), we assert that for t large enough, we have

$$\mathcal{B}_{R,r_t} \subset \Big\{\frac{1}{r_t\log t}\log D_{t,R} \le -A\Big\}. \tag{7.17}$$

Indeed, let

$$s_t = \frac{\exp\left(\frac{q}{4}(1-R)r_t\log t\right)}{1-\exp\left(-\frac{q}{4}r_t\log t\right)}.$$

Suppose that $s_t < 1$ for $t \geq T'$ for some $T' \geq T$. Then for $t \geq T'$,

$$\mathcal{B}_{R,r_t} \subset \Big\{ D_{t,R} \leq \exp(-Ar_t\log t)s_t \Big\} \subset \Big\{ D_{t,R} \leq \exp(-Ar_t \log t)s_t \Big\}.$$

It remains to combine (7.15) with (7.17) to derive (7.12). □

7.2 Proof of Proposition 5.4 (a)

We adopt the setting introduced in Remark 5.1; see also (5.2) for abbreviations. Before we start the proof and state a lemma that we will use for it, let us make the following observations. First observe that by (1.23) $\boldsymbol{H}_n(X) = \boldsymbol{\gamma}_n \Phi_{\boldsymbol{\Pi}_n}(\boldsymbol{W}_n)$. Let us write $\boldsymbol{W}_n^\varepsilon$ for the restriction of $\boldsymbol{W}_n$ to $[\varepsilon,\infty)\times\mathbb{R}^d$. Because for $w = \frac{\mathrm{d}\boldsymbol{W}_n}{\mathrm{d}\boldsymbol{\Pi}_n}$ one has $\int_{(0,\varepsilon)\times\mathbb{R}^d} fw(f,y) - \theta w(f,y)^2 \,\mathrm{d}\boldsymbol{\Pi}_n(f,y) \leq \varepsilon$, we have

$$\boldsymbol{H}_n(X) \leq \varepsilon\boldsymbol{\gamma}_n + \Phi_{\boldsymbol{\Pi}_n}(\boldsymbol{W}_n^\varepsilon).$$

Let $\boldsymbol{\Pi}_n^\varepsilon$ also be the restriction of $\boldsymbol{\Pi}_n$ to $[\varepsilon,\infty)\times\mathbb{R}^d$. As on the event $\{\max_{s\in[0,t_n]}|X_s| \leq R\boldsymbol{r}_n\}$ the support of $\boldsymbol{W}_n^\varepsilon$ is a subset of $E_n := \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{\Pi}_n^\varepsilon \cap Q_R$, we have

$$\begin{aligned} \boldsymbol{Z}_n^{R,-} &= \mathbb{E}\Big[\mathrm{e}^{\boldsymbol{H}_n(X)}\mathbb{1}\Big\{\max_{s\in[0,t_n]}|X_s| \leq R\boldsymbol{r}_n\Big\}\Big] \leq \sum_{A\subset E_n} \mathrm{e}^{\varepsilon\boldsymbol{\gamma}_n}\mathbb{E}\Big[\mathrm{e}^{\boldsymbol{\gamma}_n\Phi_{\boldsymbol{\Pi}_n}(\boldsymbol{W}_n^\varepsilon)}\mathbb{1}\Big\{\mathrm{supp}_{\mathbb{R}^d}\boldsymbol{W}_n^\varepsilon = A\Big\}\Big] \\ &\leq \sum_{A\subset E_n} \mathrm{e}^{\varepsilon\boldsymbol{\gamma}_n}\exp\Big(\sup_{\substack{\mu\in\mathcal{W}_R\\ \mathrm{supp}_{\mathbb{R}^d}\mu=A}} \boldsymbol{\gamma}_n\Phi_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)\Big)\mathbb{P}\Big[A = \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{W}_n^\varepsilon\Big]. \end{aligned} \tag{7.18}$$

In the proof we will provide a probabilistic argument that allows us to restrict the A in the summand to those which do not contain elements around zero, i.e., of A that are subsets of $\mathrm{supp}_{\mathbb{R}^d}\boldsymbol{\Pi}_n^\varepsilon \cap Q_R\setminus Q_\delta$. Then we will use that $\{A = \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{W}_n^\varepsilon\} \subset \{A \subset \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{W}_n\}$ and the following lemma (for which we do not need the Skorohod setting, i.e., we do not need to restrict to a sequence of times). To motivate the condition of the lemma, observe that $A\subset E_n$ implies that $A \subset \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{\Pi}_n^\varepsilon \subset \mathrm{supp}_{\mathbb{R}^d}\boldsymbol{\Pi}_n \subset \boldsymbol{r}_n^{-1}\mathbb{Z}^d$.

Lemma 7.5. *Let $R > \delta > 0$. There exists a function $\gamma : (0,\infty)\to\mathbb{R}$ such that $\lim_{t\to\infty}\gamma(t) = 0$, and such that for all $t\in(1,\infty)$ and all $A\subset Q_R\setminus Q_\delta$ with $r_tA\subset\mathbb{Z}^d$,*

$$\mathbb{P}\Big[A \subset \mathrm{supp}_{\mathbb{R}^d} W_t\Big] \leq \exp\Big(-qD_0(A)r_t\log t(1+\gamma(t)).\Big). \tag{7.19}$$

Proof. Let $t\in(0,\infty)$ and A be as mentioned. Without loss of generality we may assume that A is nonempty (because $D_0(\emptyset) = 0$). Write $\widetilde{A} = r_tA$. Observe that $\mathrm{supp}_{\mathbb{R}^d}\Pi_t = r_t^{-1}\mathbb{Z}^d$, so that by definition of W_t, see (1.11),

$$\Big\{A\subset \mathrm{supp}_{\mathbb{R}^d} W_t\Big\} = \Big\{\ell_t(z) > 0 \text{ for all } z\in\widetilde{A}\Big\}.$$

The paths that realise the above event, i.e., that have a strict positive local time at all points of A, they make at least $n = D_0(\widetilde{A})$ jumps. By using Stirling's inequality $n! \geq (\frac{n}{e})^n$ and that $\frac{n!}{(n+m)!} \leq \frac{1}{m!}$, we obtain

$$\begin{aligned} \mathbb{P}[\ell_t(z) > 0 \ \text{ for all } z\in A] &\leq \sum_{m=n}^\infty \mathrm{Poi}_{2dt}(m) = \sum_{m=n}^\infty \mathrm{e}^{-2dt}\frac{(2dt)^m}{m!} \\ &= \frac{(2dt)^n}{n!}\sum_{m=0}^\infty \mathrm{e}^{-2dt}\frac{(2dt)^m}{(m+n)!}n! \leq \Big(\frac{2dte}{n}\Big)^n. \end{aligned} \tag{7.20}$$

Now we use that $n = D_0(\widetilde{A}) = r_t D_0(A)$, that $n > \delta r_t$ (because A is nonempty and a subset of $Q_R \setminus Q_\delta$) and use that $r_t = t^{1+q}(\log t)^{-(1+q)}$ so that $\log t - \log r_t = -q \log t + (1+q) \log\log t$ to obtain

$$\Big(\frac{2dte}{n}\Big)^n \le \Big(\frac{2dte}{\delta r_t}\Big)^{r_t D_0(A)} = \exp\Big(D_0(A) r_t \Big(\log \frac{2de}{\delta} + \log t - \log r_t\Big)\Big)$$
$$= \exp\Big(D_0(A) r_t \Big(\log \frac{2de}{\delta} - q\log t + (1+q)\log\log t\Big)\Big),$$

so that by setting $\gamma(t) = -(\log t)^{-1}(\log \frac{2de}{\delta} + (1+q)\log\log t)$ we obtain the desired inequality. □

Proof of Proposition 5.4 (a). Fix $\varepsilon > 0$ and $\eta > 0$ and choose $\kappa > 0$ so small that

$$\mathbf{P}\big(\mathbf{\Pi}([\varepsilon,\infty)\times Q_\kappa) = 0\big) \ge 1 - \frac{\eta}{2}. \tag{7.21}$$

The $\mathbf{P}$-almost-sure convergence $\mathbf{\Pi}_n \to \mathbf{\Pi}$ in $\mathcal{M}_{\mathrm{p}}^\circ$ (see Lemma 2.4) entails that, see for example Remark 2.2,

$$\mathbf{\Pi}_n([\varepsilon,\infty)\times Q_\kappa) \overset{\mathbf{P}-a.s.}{\underset{n\to\infty}{\longrightarrow}} \mathbf{\Pi}([\varepsilon,\infty)\times Q_\kappa). \tag{7.22}$$

Therefore there exists an $N \in \mathbb{N}$ such that $\mathbf{P}(\mathcal{B}_N) \ge 1 - \eta$, where

$$\mathcal{B}_N := \mathcal{B}_N^{\varepsilon,\kappa} = \{\mathbf{\Pi}_n([\varepsilon,\infty)\times Q_\kappa) = 0 \text{ for all } n \ge N\}.$$

Henceforth, we will work on the event $\mathcal{B}_N$. Let $R > 0$. Observe that on $\mathcal{B}_N$, for any $n \in \mathbb{N}$, one has $\mathrm{supp}_{\mathbb{R}^d} W_n^\varepsilon \subset \mathcal{E}_n$ for

$$\mathcal{E}_n = \mathcal{E}_{n,\varepsilon,R,\kappa} = (\mathrm{supp}_{\mathbb{R}^d} \mathbf{\Pi}_n^\varepsilon) \cap (Q_R \setminus Q_\kappa).$$

Therefore, by adapting the last inequality in (7.18) to restricting to subsets A of $\mathcal{E}_n$, we have on $\mathcal{B}_N$, for all $n \ge N$

$$\boldsymbol{Z}_n^{R,-} \le \sum_{A\subset\mathcal{E}_n} \mathrm{e}^{\varepsilon\gamma_n} \exp\Big(\sup_{\substack{\mu\in\mathcal{W}_R\\ \mathrm{supp}_{\mathbb{R}^d}\mu = A}} \gamma_n \Phi_{\mathbf{\Pi}_n}(\mu)\Big)\mathbb{P}\Big[A \subset \mathrm{supp}_{\mathbb{R}^d} \boldsymbol{W}_n\Big]. \tag{7.23}$$

Therefore, by Lemma 7.5 and because $D_0(A) = \mathcal{D}_{\mathbf{\Pi}_n}(\mu)$ for any $\mu \in \mathcal{W}$ with $\mathrm{supp}_{\mathbb{R}^d}\mu = A$, we have

$$\begin{aligned}\boldsymbol{Z}_n^{R,-} &\le \sum_{A\subset\mathcal{E}_n} \mathrm{e}^{\varepsilon\gamma_n} \exp\Big(\sup_{\substack{\mu\in\mathcal{W}_R\\ \mathrm{supp}_{\mathbb{R}^d}\mu = A}} \Big[\gamma_n \Phi_{\mathbf{\Pi}_n^\varepsilon}(\mu) - \gamma_n(1+\gamma(t_n))q\mathcal{D}_{\mathbf{\Pi}_n^\varepsilon}(\mu)\Big].\Big)\\ &\le \mathrm{e}^{\varepsilon\gamma_n} 2^{\#\mathcal{E}_n} \exp\Big(\sup_{\mu\in\mathcal{W}} \Big[\gamma_n \Phi_{\mathbf{\Pi}_n^\varepsilon}(\mu) - \gamma_n(1+\gamma(t_n))q\mathcal{D}_{\mathbf{\Pi}_n^\varepsilon}(\mu)\Big].\Big)\end{aligned} \tag{7.24}$$

Since $\mathbf{\Pi}_n \to \mathbf{\Pi}$ in $\mathcal{M}_{\mathrm{p}}^\circ$ it follows by Remark 2.2 that for any $\varepsilon > 0$, $\mathbf{\Pi}_n^\varepsilon \to \mathbf{\Pi}^\varepsilon$ in $\mathcal{M}_{\mathrm{p}}^\circ$, where $\mathbf{\Pi}^\varepsilon$ is the restriction of $\mathbf{\Pi}$ to $[\varepsilon,\infty)\times Q_R \setminus Q_\kappa$. Therefore,

$$\#\mathcal{E}_n = \#\mathcal{E}_{n,\varepsilon,R,\kappa} = \mathbf{\Pi}_n^\varepsilon((0,\infty)\times Q_R\setminus Q_\kappa) = \mathbf{\Pi}_n([\varepsilon,\infty)\times Q_R\setminus Q_\kappa) \;\to\; \mathbf{\Pi}([\varepsilon,\infty)\times Q_R\setminus Q_\kappa),$$

and thus

$$\lim_{n\to\infty} \frac{1}{\gamma_n}(\log 2^{\#\mathcal{E}_n}) = 0,$$

and, since $\lim_{n\to\infty}\gamma(t_n) = 0$, we have by Theorem 3.3 (a) and Proposition 3.4 (b),

$$\limsup_{n\to\infty} \sup_{\mu\in\mathcal{W}} \Big[\gamma_n \Phi_{\mathbf{\Pi}_n^\varepsilon}(\mu) - \gamma_n(1+\gamma(t_n))q\mathcal{D}_{\mathbf{\Pi}_n^\varepsilon}(\mu)\Big] \le \sup_{\mu\in\mathcal{W}} [\Phi_{\mathbf{\Pi}^\varepsilon}(\mu) - q\,\mathcal{D}_{\mathbf{\Pi}^\varepsilon}(\mu)] \le \Xi(\mathbf{\Pi}^\varepsilon).$$

Therefore, on $\mathcal{B}_N$, for any $R > 0$,

$$\limsup_{n\to\infty} \frac{1}{\gamma_n} \log \boldsymbol{Z}_n^{R,-} \le \varepsilon + \Xi(\mathbf{\Pi}^\varepsilon).$$

So summarizing the above, for every ε and η in $(0,\infty)$ there exist a $\kappa>0$ and an $N\in\mathbb{N}$ such that $\mathbf{P}[\mathcal{B}_N^{\varepsilon,\kappa}]\geq 1-\eta$ and thus

$$\mathbf{P}\Big[\limsup_{n\to\infty}\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-}\leq \varepsilon+\Xi(\boldsymbol{\Pi}^\varepsilon)\Big]\geq \mathbf{P}[\mathcal{B}_N^{\varepsilon,\kappa}]\geq 1-\eta.$$

As $\boldsymbol{\Pi}^\varepsilon\to\boldsymbol{\Pi}$ in $\mathcal{M}_{\mathrm{p}}^\circ$ almost surely, we have $\limsup_{\varepsilon\downarrow 0}\Xi(\boldsymbol{\Pi}^\varepsilon)\leq\Xi(\boldsymbol{\Pi})$ almost surely by Theorem 3.3 (b) (strictly speaking we should replace the occurrences of "ε" by "ε_k" for a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ that converges to 0 and replace "$\limsup_{\varepsilon\downarrow 0}$" by "$\limsup_{k\to\infty}$"). Therefore, for all $\eta>0$ and $\zeta>0$ there exists an $\varepsilon>0$ such that

$$\mathbf{P}[\varepsilon+\Xi(\boldsymbol{\Pi}^\varepsilon)\leq\Xi(\boldsymbol{\Pi})+\zeta]\geq 1-\eta,$$

and thus

$$\mathbf{P}\Big[\limsup_{n\to\infty}\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-}\leq \Xi(\boldsymbol{\Pi})+\zeta\Big]\geq \mathbf{P}\Big[\limsup_{n\to\infty}\frac{1}{\gamma_n}\log \boldsymbol{Z}_n^{R,-}\leq \varepsilon+\Xi(\boldsymbol{\Pi}^\varepsilon)\Big]\geq 1-\eta.$$

As the above holds for any $\eta>0$ and $\zeta>0$, (by first taking η to zero and then ζ) we obtain (5.9). □

7.3 Proof of Proposition 5.4 (b)

In this section we prove Proposition 5.4 (b) by mentioning where to adapt the proof in of Proposition 5.4 (a) as in the previous section.

Proof of Proposition 5.4 (b). Let us write

$$\mathcal{C}^\delta=\{\nu\in\mathcal{W}:\mathfrak{d}(\nu,\mu^*)\geq\delta\},\qquad \mathfrak{S}^\delta=\bigcup\{\mathrm{supp}_{\mathbb{R}^d}\nu:\nu\in\mathcal{C}^\delta(\mu^*)\}.$$

Thus $\mathfrak{S}^\delta$ is the subset of $\mathbb{R}^d$ where the ν that are at least at distance δ of μ^*, are allowed to be supported. Then, similarly to (7.18) and (7.23), the following estimates hold, with $\mathcal{E}_n^\delta=\mathcal{E}_n\cap\mathfrak{S}^\delta$, on $\mathcal{B}_N$, for $n\geq N$

$$\begin{aligned}\boldsymbol{Z}_n^{R,-,\delta}&=\mathbb{E}\Big[\mathrm{e}^{\boldsymbol{H}_n(X)}\,\mathbb{1}\{\mathfrak{d}(\boldsymbol{W}_n,\mu^*)\geq\delta\}\,\mathbb{1}\Big\{\max_{s\in[0,t_n]}|X_s|\leq R\boldsymbol{r}_n\Big\}\Big]\\&\leq\sum_{A\subset\mathcal{E}_n^\delta}\mathrm{e}^{\varepsilon\boldsymbol{\gamma}_n}\exp\Big(\sup_{\substack{\mu\in\mathcal{W}_R\cap\mathcal{C}^\delta\\ \mathrm{supp}_{\mathbb{R}^d}\mu=A}}\boldsymbol{\gamma}_n\Phi_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)\Big)\mathbb{P}\Big[A\subset\mathrm{supp}_{\mathbb{R}^d}\boldsymbol{W}_n\Big].\end{aligned}$$

Then, similar to (7.24), by using that $\mathcal{E}_n^\delta\subset\mathcal{E}_n$, we obtain (on $\mathcal{B}_N$, for $n\geq N$)

$$\begin{aligned}\boldsymbol{Z}_n^{R,-,\delta}&\leq\sum_{A\subset\mathcal{E}_n^\delta}\mathrm{e}^{\varepsilon\boldsymbol{\gamma}_n}\exp\Big(\sup_{\substack{\mu\in\mathcal{W}_R\\ \mathrm{supp}_{\mathbb{R}^d}\mu=A}}\Big[\boldsymbol{\gamma}_n\Phi_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)-\boldsymbol{\gamma}_n(1+\gamma(t_n))q\mathcal{D}_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)\Big].\Big)\\&\leq\mathrm{e}^{\varepsilon\boldsymbol{\gamma}_n}2^{\#\mathcal{E}_n}\exp\Big(\sup_{\mu\in\mathcal{C}^\delta}\Big[\boldsymbol{\gamma}_n\Phi_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)-\boldsymbol{\gamma}_n(1+\gamma(t_n))q\mathcal{D}_{\boldsymbol{\Pi}_n^\varepsilon}(\mu)\Big].\Big)\end{aligned}$$

The rest of the proof follows in the same fashion as in the proof of Proposition 5.4 (a) in the previous section, by taking Ξ^δ (see (5.5)) instead of Ξ. □

Appendix A. The space $\mathfrak{E}$

Lemma A.1. *Let $\mathfrak{E}$ be the union of $(0,\infty)\times\mathbb{R}^d$ with $(0,\infty]$, $\mathfrak{E}=((0,\infty)\times\mathbb{R}^d)\cup(0,\infty]$. Define $\mathfrak{d}:\mathfrak{E}\times\mathfrak{E}\to[0,\infty)$ by*

$$\begin{aligned}\mathfrak{d}(s,s')&=|s-s'| && s,s'\in(0,\infty],\\ \mathfrak{d}(s,(f,y))&=\frac{1}{f}+\Big|\frac{f}{1\vee|y|}-s\Big| && s\in(0,\infty],(f,y)\in(0,\infty)\times\mathbb{R}^d,\\ \mathfrak{d}((f,y),(f',y'))&=\frac{1}{f\wedge f'}\Big(1-\mathrm{e}^{-|\log f-\log f'|-|y-y'|}\Big)+\Big|\frac{f}{1\vee|y|}-\frac{f'}{1\vee|y'|}\Big| && (f,y),(f',y')\in(0,\infty)\times\mathbb{R}^d.\end{aligned}$$

(a) $\mathfrak{d}$ *is a metric on* $\mathfrak{E}$.
(b) *The function* $\iota : (0,\infty)\times\mathbb{R}^d \to \mathfrak{E}$, $\iota((f,y)) = (f,y)$, $\quad (f,y)\in(0,\infty)\times\mathbb{R}^d$, *is continuous and open.*
(c) $\mathfrak{E}$ *equipped with the topology generated by* $\mathfrak{d}$ *is a locally compact Polish space, such that* (i) *and* (ii) *of Lemma* 2.1 *hold. Moreover, for* $s,h>0$, *the closure of* $\mathcal{H}_h^s$, *is given by*

$$\overline{\mathcal{H}_h^s} = \{(f,y)\in(0,\infty)\times\mathbb{R}^d : f \ge s|y| + h\} \cup [s,\infty]\,, \tag{A.1}$$

and is a compact set.
(d) *For every compact set* K *in* $\mathfrak{E}$ *there exist* $h,s>0$ *such that* $K\cap(0,\infty)\times\mathbb{R}^d \subset \mathcal{H}_h^s$ *for some* $h,s>0$.

Proof. (a) The idea behind this is very similar to the space described in (BKdS18, Section 13) (which considers a larger space than $\mathbb{R}\times\mathbb{R}^d$ instead of $(0,\infty)\times\mathbb{R}^d$). Somehow the homeomorphism $(0,\infty)\times\mathbb{R}^d \to \mathbb{R}\times\mathbb{R}^d$, $(f,y)\mapsto(\log f, y)$ is put inbetween to relate these spaces, but the "second part" of the metric is quite similar to keep the limit as being the "ratio". In order to see that $\mathfrak{d}$ is a metric, we have to show that the triangle inequality is satisfied. If $\mathfrak{a},\mathfrak{b},\mathfrak{c}\in\mathfrak{E}$, then $\mathfrak{d}(\mathfrak{a},\mathfrak{b}) \le \mathfrak{d}(\mathfrak{a},\mathfrak{c}) + \mathfrak{d}(\mathfrak{c},\mathfrak{b})$ follows easily if at least one element of $\mathfrak{a},\mathfrak{b},\mathfrak{c}$ is in $(0,\infty]$. Therefore, we show that $\mathfrak{d}$ satisfies the triangle inequality on $(0,\infty)\times\mathbb{R}^d$. It is rather easy to see that it suffices to prove that $\bar{\mathfrak{d}}$ is a metric on $\mathbb{R}\times\mathbb{R}^d$ (by plugging in $\lambda = \log f$ and $z = y$), where

$$\bar{\mathfrak{d}}((\lambda,z),(\lambda',z')) = \mathrm{e}^{-(\lambda\wedge\lambda')}\left(1 - \mathrm{e}^{-|\lambda-\lambda'|-|z-z'|}\right) \qquad (\lambda,z),(\lambda',z')\in\mathbb{R}\times\mathbb{R}^d.$$

As we will see, this can be boiled down to the fact that $(1-\mathrm{e}^{-a})(1-\mathrm{e}^{-b}) \ge 0$ for $a,b\ge 0$. Let $(\lambda,z),(\lambda',z')$, $(\lambda'',z'')\in\mathbb{R}\times\mathbb{R}^d$. We may assume $\lambda<\lambda'$ and $\lambda' = \lambda + a$, $\lambda'' = \lambda + b$, $a>0$, $b\in\mathbb{R}$. Then with $p = |z-z'|$, $q = |z-z''|$, $r = |z'-z''|$, so that $p \le q + r$ and $\mathrm{e}^{-|z-z'|} = \mathrm{e}^{-p} \ge \mathrm{e}^{-q-r}$,

$$\begin{aligned}
&\mathrm{e}^{\lambda}\left[\bar{\mathfrak{d}}((\lambda,z),(\lambda'',z'')) + \bar{\mathfrak{d}}((\lambda'',z''),(\lambda',z')) - \bar{\mathfrak{d}}((\lambda,z),(\lambda',z'))\right]\\
&= \mathrm{e}^{-(b\wedge 0)}\left(1-\mathrm{e}^{-b-q}\right) + \mathrm{e}^{-(a\wedge b)}\left(1 - \mathrm{e}^{-|b-a|-r}\right) - \left(1-\mathrm{e}^{-a-p}\right)\\
&\ge \left(1-\mathrm{e}^{-b-q}\right) + \mathrm{e}^{-a}\left(1-\mathrm{e}^{-|b-a|-r}\right) - \left(1-\mathrm{e}^{-a-q-r}\right)\\
&= -\mathrm{e}^{-b-q} + \mathrm{e}^{-a} - \mathrm{e}^{-|b-a|-a-r} + \mathrm{e}^{-a-q-r}\\
&= \mathrm{e}^{-a}(1 - \mathrm{e}^{a-b-q} - \mathrm{e}^{-|b-a|-r} + \mathrm{e}^{-q-r})\\
&\begin{cases} \ge \mathrm{e}^{-a}(1-\mathrm{e}^{-q} - \mathrm{e}^{-r} + \mathrm{e}^{-q-r}) = \mathrm{e}^{-a}(1-\mathrm{e}^{-q})(1-\mathrm{e}^{-r}) \ge 0 & b>a,\\ = \mathrm{e}^{-a}(1-\mathrm{e}^{|b-a|-q})(1-\mathrm{e}^{-|b-a|-r}) \ge 0 & b<a. \end{cases}
\end{aligned}$$

(b) It is rather straightforward to check that a sequence $(f_n,y_n)_{n\in\mathbb{N}}$ in $(0,\infty)\times\mathbb{R}^d$ converges to an element (f,y) of $(0,\infty)\times\mathbb{R}^d$ with respect to $\mathfrak{d}$ if and only if it converges with respect to the Euclidean metric on $(0,\infty)\times\mathbb{R}^d$. Therefore ι is continuous and open.

(c) That (A.1) holds follows by the definition of $\mathfrak{d}$. Observe that for a sequence $(a_n)_{n\in\mathbb{N}}$ in $\overline{\mathcal{H}_h^s}$ there either exists a subsequence that is contained in $[s,\infty]$ or a subsequence in $(0,\infty)\times\mathbb{R}^d$ of the form $(f_n,y_n)_{n\in\mathbb{N}}$ for which either

(1) f_n is contained in a set of the form $[h,M]$ for some $M\ge h$ (and thus the y_n are contained in a ball of radius $s(M+h)$), or,
(2) $f_n\to\infty$ (and thus $\liminf_{n\to\infty}\frac{f_n}{|y_n|}\ge s$).

In both cases one can find a subsequence of $(f_n,y_n)_{n\in\mathbb{N}}$ that converges in $\overline{\mathcal{H}_h^s}$. Hence $\overline{\mathcal{H}_h^s}$ is compact.

Every $(f,y)\in(0,\infty)\times\mathbb{R}^d$ has a compact neighbourhood in $(0,\infty)\times\mathbb{R}^d$ and therefore in $\mathfrak{E}$, because ι is continuous. On the other hand, every $t\in(0,\infty]$ has a compact neighbourhood, for example $\overline{\mathcal{H}_h^s}$ for $s<t$ and $h = \frac{1}{t-s}$ (indeed, observe that $\{a\in\mathfrak{E} : \mathfrak{d}(a,t) < t-s\} \subset \mathcal{H}_h^s$). Therefore $\mathfrak{E}$ is locally compact.

Observe that a sequence $(f_n,y_n)_{n\in\mathbb{N}}$ whose elements belong to $(0,\infty)\times\mathbb{R}^d$ is a Cauchy sequence in $\mathfrak{E}$ either if it is a Cauchy sequence in $[h,h^{-1}]\times\mathbb{R}^d$ for some $h\in(0,1)$ or if $f_n\to\infty$ and $\frac{f_n}{|y_n|}\to s$ for some $s\in(0,\infty]$, this s is then the limit in $\mathfrak{E}$. From this we infer that $\mathfrak{E}$ is complete. It is separable as $\mathbb{Q}_{>0} \times \mathbb{Q}^d \cup \mathbb{Q}_{>0}$ is dense, where $\mathbb{Q}_{>0} = (0,\infty) \cap \mathbb{Q}$. Therefore $\mathfrak{E}$ is a Polish space.

(d) This follows by the fact that every compact set in $\mathfrak{E}$ is a subset of $\overline{\mathcal{H}_h^s}$ for some $h, s > 0$. Indeed, first it will be clear that every compact set is a subset of $\{(f,y) \in (0,\infty) \times \mathbb{R}^d : f \geq h\} \cup (0,\infty]$ for some $h > 0$. Secondly, $(0,s]$ is not compact for all $s > 0$ and moreover, if $(f_n, y_n)_{n\in\mathbb{N}}$ is a sequence in $(0,\infty) \times \mathbb{R}^d$ with $f_n, y_n \to \infty$ and $\frac{f_n}{|y_n|} \to 0$, then it does not possess a subsequence that converges in $\mathfrak{E}$. □

Proof of Lemma 2.1. The existence of $\mathfrak{E}$ for which (i) and (ii) hold, follows by Lemma A.1.

(a) Because ι is an open map, $\iota(B)$ is a Borel set in $\mathfrak{E}$ for every Borel set B in $(0,\infty) \times \mathbb{R}^d$. Hence $\mathcal{P} \circ \iota$ defines a measure on $\mathfrak{E}$, clearly with values in $\mathbb{N}_0 \cup \{\infty\}$. It is a Radon measure because ι is continuous and so $\iota(K)$ is compact in $\mathfrak{E}$ for every compact set K in $(0,\infty) \times \mathbb{R}^d$. Therefore it is a Point measure on $(0,\infty) \times \mathbb{R}^d$. because for such K, one has $\mathcal{P}(\iota(K)) < \infty$ because $\mathcal{P}$ is Radon on $\mathfrak{E}$. Observe that not every element of $\mathcal{M}_{\mathrm{p}}((0,\infty) \times \mathbb{R}^d)$ is of the form $\mathcal{P} \circ \iota$ for some $\mathcal{P}$ in $\mathcal{M}_{\mathrm{p}}^{\circ}$. Take $\sum_{n\in\mathbb{N}} \delta_{(|y_n|+1, y_n)}$ for a sequence $(y_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^d$ with $|y_n| \to \infty$ for example (it attains the value ∞ on the set $\mathcal{H}_s^s$ for any $s \in (0,1)$).

(b) Because the embedding is continuous, it follows that $\overline{\mathcal{P}}$ is a measure on $\mathfrak{E}$. Hence it is an element of $\mathcal{M}_{\mathrm{p}}$ if and only if it is a Radon measure. Therefore by (i) it follows that $\overline{\mathcal{P}}$ is an element of $\mathcal{M}_{\mathrm{p}}$ if and only if $\mathcal{P}(\mathcal{H}_h^s) < \infty$ for all $s, h > 0$. Suppose the latter is the case. Then $\operatorname{supp} \overline{\mathcal{P}} \subset (0,\infty) \times \mathbb{R}^d$, as if otherwise, then there exists a sequence $(f_n, y_n)_{n\in\mathbb{N}}$ in this support that converges in $\mathfrak{E}$ to $2s$ for a $s \in (0,\infty]$. This can only be the case if that sequence is contained in $\mathcal{H}_h^s$ for some $h > 0$, in which case $\mathcal{P}(\mathcal{H}_h^s) = \infty$. □

Appendix B. Measurability of the maximizer μ^*

In this section we show the measurability of the μ^* as in Lemma 5.3.

First observe that $\mathbf{\Pi}$, $\mathbf{\Pi}_n$ and $\mathbf{\Pi}_n^{(L)}$ for all $n, L \in \mathbb{N}$ are all good point measures $\mathbf{P}$-almost surely by Lemma 4.4 and Lemma 4.2. Let $(\Omega, \mathcal{F})$ be the underlying (complete) measurable space of $\mathbf{P}$. Let $\Omega_1 \in \mathcal{F}$ be such that on Ω_1, $\mathbf{\Pi}$, $\mathbf{\Pi}_n$ and $\mathbf{\Pi}_n^{(L)}$ for all $n, L \in \mathbb{N}$ are all good point measures and such that $\mathbf{\Pi}_n \to \mathbf{\Pi}$ on Ω_1. Then, for all $\omega \in \Omega_1$ and all $n, L \in \mathbb{N}$, there exist $\mu^*[\omega], \mu_n^*[\omega]$ and $\mu_{n,L}^*[\omega]$ such that

$$\Psi_{\mathbf{\Pi}[\omega]}(\mu^*[\omega]) = \Xi(\mathbf{\Pi}[\omega]), \quad \Psi_{\mathbf{\Pi}_n[\omega]}(\mu^*[\omega]) = \Xi(\mathbf{\Pi}_n[\omega]), \quad \Psi_{\mathbf{\Pi}_n^{(L)}[\omega]}(\mu^*[\omega]) = \Xi(\mathbf{\Pi}_n^{(L)}[\omega]).$$

Let us set $\mu^*[\omega] = \mu_n^*[\omega] = \mu_{n,L}^* = 0$ for all $n, L \in \mathbb{N}$. As, on Ω_1, $\mathbf{\Pi}_n^{(L)} \xrightarrow{L\to\infty} \mathbf{\Pi}_n$ and $\mathbf{\Pi}_n \xrightarrow{L\to\infty} \mathbf{\Pi}$, by Theorem 2.8 (d) we have $\mu_{n,L}^* \xrightarrow{L\to\infty} \mu_n^*$ and $\mu_n^* \xrightarrow{n\to\infty} \mu^*$ on Ω. Therefore it suffices to show that $\mu_{n,L}^*$ is measurable for all $n, L \in \mathbb{N}$ in order to conclude that μ^* is measurable (likewise, μ_n^* for all $n \in \mathbb{N}$).

As $\mathbf{\Pi}$ is almost surely good, there exists an $\Omega_1 \subset \Omega$ with $\mathbf{P}(\Omega_1) = 1$ such that for each $\omega \in \Omega_1$, by Theorem 2.8 (a) there exists a unique $\mu^*[\omega] \in \mathfrak{F}(\mathbf{\Pi}(\omega))$ such that

$$\Psi_{\mathbf{\Pi}[\omega]}(\mu^*[\omega]) = \Xi(\mathbf{\Pi}[\omega]).$$

We will show that there exists an $\Omega^* \subset \Omega$ with $\mathbf{P}(\Omega^*) = 1$ such that $\omega \mapsto \mu^*[\omega]$ is measurable.

This follows because $\mathbf{\Pi}_n^L$ has a finite support: For $\omega \in \Omega_1$, there exist $m[\omega] \in \mathbb{N}_0$ and distinct $(f_1[\omega], y_1[\omega])$, $\dots$, $(f_m[\omega], y_m[\omega])$ in $\operatorname{supp} \mathbf{\Pi}_n^L[\omega]$ such that

$$\Psi_{\mathbf{\Pi}_n^L[\omega]}(\mu_{n:L}^*[\omega]) = \Xi(\mathbf{\Pi}_n^L[\omega]) = \varphi_m(f_1[\omega], \dots, f_m[\omega]) - qD_0(y_1[\omega], \dots, y_m[\omega]).$$

These m, $f_1, \dots, f_m$ and $y_1, \dots, y_m$ are measurable as this is a finite optimization problem. Then $\mu_{n:L}^* = \sum_{i=1}^m w_i \delta_{f_i, y_i}$, where the w_i are measurable functions of the $f_1, \dots, f_m$ due to Proposition 3.8.

(Measurability of a finite optimization problem). Let $U_1, \dots, U_m$ be random variables such that $\mathbf{P}[U_i = U_j] = 0$ for all i, j with $i \neq j$. Suppose Ω_1 is such that $U_i(\omega) \neq U_j(\omega)$ for all i, j with $i \neq j$ and all $\omega \in \Omega_1$. On Ω_1 there exists a function i such that

$$i(\omega) = \arg\max_{j=1}^{m} U_j(\omega), \qquad \omega \in \Omega_1.$$

This function is measurable as for all $k \in \{1, \ldots, m\}$

$$\{\omega \in \Omega_1 : i(\omega) = k\} = \{\omega \in \Omega_1 : U_k(\omega) - \max_{j=1}^{m} U_j(\omega) = 0\},$$

and because U_k and the maximum function $\max_{j=1}^{m} U_j$ are measurable.

References

[Ast16] Arvydas Astrauskas. From extreme values of i.i.d. random fields to extreme eigenvalues of finite-volume Anderson Hamiltonian. Probab. Surv., 13:156–244, 2016.

[Bil99] Patrick Billingsley. Convergence of probability measures. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BKdS18] Marek Biskup, Wolfgang König, and Renato S. dos Santos. Mass concentration and aging in the parabolic Anderson model with doubly-exponential tails. Probab. Theory Related Fields, 171(1-2):251–331, 2018.

[Bra02] Andrea Braides. Γ-convergence for beginners, volume 22 of Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, Oxford, 2002.

[DFSX20a] Jian Ding, Ryoki Fukushima, Rongfeng Sun, and Changji Xu. Biased random walk conditioned on survival among Bernoulli obstacles: subcritical phase. Comm. Math. Phys., 376(3):2161–2195, 2020.

[DFSX20b] Jian Ding, Ryoki Fukushima, Rongfeng Sun, and Changji Xu. Geometry of the random walk range conditioned on survival among Bernoulli obstacles. Probab. Theory Related Fields, 177(1-2):91–145, 2020.

[DFSX21] Jian Ding, Ryoki Fukushima, Rongfeng Sun, and Changji Xu. Distribution of the random walk conditioned on survival among quenched Bernoulli obstacles. Ann. Probab., 49(1):206–243, 2021.

[DVJ03] D. J. Daley and D. Vere-Jones. An introduction to the theory of point processes. Vol. I. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[GKM07] Jürgen Gärtner, Wolfgang König, and Stanislav Molchanov. Geometric characterization of intermittency in the parabolic Anderson model. Ann. Probab., 35(2):439–499, 2007.

[GM90] J. Gärtner and S. A. Molchanov. Parabolic problems for the Anderson model. I. Intermittency and related topics. Comm. Math. Phys., 132(3):613–655, 1990.

[Kal83] Olav Kallenberg. Random measures. Akademie-Verlag, Berlin; Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, third edition, 1983.

[Kle08] Achim Klenke. Probability theory. Universitext. Springer-Verlag London, Ltd., London, 2008. A comprehensive course, Translated from the 2006 German original.

[KLMS09] Wolfgang König, Hubert Lacoin, Peter Mörters, and Nadia Sidorova. A two cities theorem for the parabolic Anderson model. Ann. Probab., 37(1):347–392, 2009.

[Kön16] Wolfgang König. The parabolic Anderson model. Pathways in Mathematics. Birkhäuser/Springer, [Cham], 2016. Random walk in random potential.

[LM12] Hubert Lacoin and Peter Mörters. A scaling limit theorem for the parabolic Anderson model with exponential potential. In Probability in complex physical systems, volume 11 of Springer Proc. Math., pages 247–272. Springer, Heidelberg, 2012.

[Mör11] Peter Mörters. The parabolic Anderson model with heavy-tailed potential. In Surveys in stochastic processes, EMS Ser. Congr. Rep., pages 67–85. Eur. Math. Soc., Zürich, 2011.

[MOS11] Peter Mörters, Marcel Ortgiese, and Nadia Sidorova. Ageing in the parabolic Anderson model. Ann. Inst. Henri Poincaré Probab. Stat., 47(4):969–1000, 2011.

[MS13] Neal Madras and Gordon Slade. The self-avoiding walk. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1993 original.

[OR16] Marcel Ortgiese and Matthew I. Roberts. Intermittency for branching random walk in Pareto environment. Ann. Probab., 44(3):2198–2263, 2016.

[OR17] Marcel Ortgiese and Matthew I. Roberts. One-point localization for branching random walk in Pareto environment. Electron. J. Probab., 22:Paper No. 6, 20, 2017.

[OR18] Marcel Ortgiese and Matthew I. Roberts. Scaling limit and ageing for branching random walk in Pareto environment. Ann. Inst. Henri Poincaré Probab. Stat., 54(3):1291–1313, 2018.

[PS] Julien Poisat and François Simenhaus. Localization of a one-dimensional simple random walk among power-law renewal obstacles. preprint, arXiv:2201.05377 (2022).

[Res87] Sidney I. Resnick. Extreme values, regular variation, and point processes, volume 4 of Applied Probability. A Series of the Applied Probability Trust. Springer-Verlag, New York, 1987.

[Sla11] Gordon Slade. The self-avoiding walk: a brief survey. In Surveys in stochastic processes, EMS Ser. Congr. Rep., pages 181–199. Eur. Math. Soc., Zürich, 2011.

[Szn98] Alain-Sol Sznitman. Brownian motion, obstacles and random media. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.